Genomic Update of Phenotypic Prediction Rule for Methicillin-Resistant *Staphylococcus aureus* (MRSA) USA300 Discloses Jail Transmission Networks with Increased Resistance

Sarah E. Sansom,[a] Emily Benedict,[b] Stephanie N. Thiede,[b] Bala Hota,[a] Alla Aroutcheva,[a] Darjai Payne,[a] Chad Zawitz,[c] Evan S. Snitkin,[b] Stefan J. Green,[a] Robert A. Weinstein,[a] Kyle J. Popovich[a]

[a]Division of Infectious Diseases, Rush University Medical Center/Cook County Health, Chicago, Illinois, USA
[b]Department of Microbiology and Immunology, University of Michigan, Ann Arbor, Michigan, USA
[c]Division of Infectious Diseases, Cook County Health/Cermak Health Services, Chicago, Illinois, USA

ABSTRACT Methicillin-resistant *Staphylococcus aureus* (MRSA) is a leading cause of health care-associated (HA) and community-associated (CA) infections. USA300 strains are historically CA-MRSA, while USA100 strains are HA-MRSA. Here, we update an antibiotic prediction rule to distinguish these two genotypes based on antibiotic resistance phenotype using whole-genome sequencing (WGS), a more discriminatory methodology than pulsed-field gel electrophoresis (PFGE). MRSA clinical isolates collected from 2007 to 2017 underwent WGS; associated epidemiologic data were ascertained. In developing the rule, we examined MRSA isolates that included a population with a history of incarceration. Performance characteristics of antibiotic susceptibility for predicting USA300 compared to USA100, as defined by WGS, were examined. Phylogenetic analysis was performed to examine resistant USA300 clades. We identified 275 isolates (221 USA300, 54 USA100). Combination susceptibility to clindamycin or levofloxacin performed the best overall (sensitivity 80.7%, specificity 75.9%) to identify USA300. The average number of antibiotic classes with resistance was higher for USA100 (3 versus 2, $P < 0.001$). Resistance to ≤ 2 classes was predictive for USA300 (area under the curve (AUC) 0.84, 95% confidence interval 0.78 to 0.90). Phylogenetic analysis identified a cluster of USA300 strains characterized by increased resistance among incarcerated individuals. Using a combination of clindamycin or levofloxacin susceptibility, or resistance to ≤ 2 antibiotic classes, was predictive of USA300 as defined by WGS. Increased resistance was observed among individuals with incarceration exposure, suggesting circulation of a more resistant USA300 clade among at-risk community networks. Our phenotypic prediction rule could be used as an epidemiologic tool to describe community and nosocomial shifts in USA300 MRSA and quickly identify emergence of lineages with increased resistance.

IMPORTANCE Methicillin-resistant *Staphylococcus aureus* (MRSA) is an important cause of health care-associated (HA) and community-associated (CA) infections, but the epidemiology of these strains (USA100 and USA300, respectively) now overlaps in health care settings. Although sequencing technology has become more available, many health care facilities still lack the capabilities to perform these analyses. In this study, we update a simple prediction rule based on antibiotic resistance phenotype with integration of whole-genome sequencing (WGS) to predict strain type based on antibiotic resistance profiles that can be used in settings without access to molecular strain typing methods. This prediction rule has many potential epidemiologic applications, such as analysis of retrospective data sets, regional monitoring, and ongoing surveillance of CA-MRSA infection trends. We demonstrate application of this rule to identify an emerging USA300 strain with increased antibiotic resistance among incarcerated individuals that deviates from the rule.

Methicillin-resistant *Staphylococcus aureus* (MRSA) has been an increasingly important cause of both health care-associated (HA) and community-associated (CA) infections over the past several decades (1–4). CA-MRSA is increasingly recognized as an important cause of nosocomial infections (5–7). By pulsed-field gel electrophoresis (PFGE), USA300 has been identified as the most common strain of CA-MRSA and can lead to severe infections in previously healthy individuals (8–10). USA100 is a strain of MRSA traditionally associated with health care exposures and remains an important cause of bloodstream infection in hospitals (11). While HA-MRSA infection rates have declined (12), the incidence of invasive CA-MRSA infections remains largely unchanged (13–15). As USA300 has emerged as an important cause of hospital-onset MRSA infections (5, 16–18), the ability to differentiate between USA300 and USA100 can have useful epidemiologic and treatment implications.

Historically, USA300 strains have increased susceptibility to non-β-lactam antibiotics. By contrast, USA100 strains often display resistance to multiple antibiotic classes (19). Clindamycin and levofloxacin resistance have been suggested as potentially useful phenotypic markers to distinguish between the USA100 and USA300 strains (20). Preserved susceptibility to trimethoprim-sulfamethoxazole and tetracyclines has been described at greater than 90% prevalence in CA-MRSA skin and soft tissue infections (21). However, multidrug-resistant (MDR) USA300 MRSA is emerging as a community-associated infection in some high-risk populations, including men who have sex with men (22).

While PFGE has traditionally been used to distinguish strains of MRSA, it has limited ability to differentiate between endemic MRSA strains. In addition, studies using WGS have shown that PFGE can misclassify strains (23, 24). WGS resolution is superior to PFGE and is now considered the gold standard for MRSA strain typing (25, 26). Our previous work (27) used performance characteristics of the antimicrobial susceptibility phenotype, a "phenotypic rule", to predict genotype as identified by PFGE. Our aim is to update this phenotypic prediction rule for distinguishing between common MRSA genotypes using WGS and current antibiotic resistance patterns. We examined performance of antimicrobial susceptibility and number of resistant antibiotics as a tool to distinguish between USA300 and USA100 in clinical MRSA isolates. This rule is intended to be used for discrimination of USA300 in both community and hospital isolates to study the evolving epidemiology of MRSA. We also demonstrate how our approach disclosed possible networks of MRSA transmission based on deviations from the established rule in a population with a history of incarceration.

RESULTS

MRSA isolates (n = 275) were included in the analysis, with 221 isolates identified as USA300 and 54 isolates identified as USA100. Performance characteristics for predicting genotype, for selected individual, and for combinations of antibiotics are listed in Table 1. Susceptibility to clindamycin (sensitivity [Sn] 63.3%, specificity [Sp] 88.9%, likelihood ratio [LR] 5.7 [95% confidence interval (CI) 2.7 to 12.2]) performed the best for an individual antibiotic to predict USA300, followed by levofloxacin (Sn 62.4%, Sp 81.5%, LR 3.4 [95% CI 1.9 to 6.0]). Tetracycline, gentamicin, and rifampin had high sensitivity but poor specificity for predicting USA300 (Sn 99.1%, 100%, and 99.1%, respectively; Sp 18.5%, 7.4%, and 5.6%, respectively). Combination of susceptibility to clindamycin or levofloxacin performed the best overall to predict USA300 (Sn 80.7%, Sp 75.9%, LR 3.4 [95% CI 2.1 to 5.4]). Combination of susceptibility to both clindamycin and levofloxacin had excellent specificity but poor sensitivity (Sn 45.1%, Sp 94.4%, LR 8.2 [95% CI 2.7 to 24.8]).

The positive predictive value (PPV) of each rule for identifying USA300 strains increased with higher prevalence. When the prevalence of USA300 strains was 25%, 50%, and 75%, the PPVs for identifying USA300 were 65.5%, 85.1%, and 94.5% for the

TABLE 1 Antibiotic susceptibility phenotype: performance characteristics to predict USA300

Susceptible to antibiotic[a]	Sensitivity (%)	Specificity (%)	Positive LR (95% CI)	Prevalence, 25%		Prevalence, 50%		Prevalence, 75%	
				PPV (%)	NPV (%)	PPV (%)	NPV (%)	PPV (%)	NPV (%)
Levofloxacin and clindamycin	45.4	94.4	8.2 (2.7–24.8)	73.2	83.9	89.1	63.4	96.1	36.6
Clindamycin and tetracycline	62.8	90.7	6.8 (2.9–15.7)	69.4	88.0	87.2	71.0	95.3	44.9
Clindamycin[b]	**63.3**	**88.9**	**5.7 (2.7–12.2)**	**65.5**	**87.9**	**85.1**	**70.8**	**94.5**	**44.7**
Levofloxacin	**62.4**	**81.5**	**3.4 (1.9–6.0)**	**53.0**	**86.7**	**77.1**	**68.5**	**91.0**	**42.0**
Levofloxacin or clindamycin	**80.7**	**75.9**	**3.4 (2.1–5.4)**	**52.8**	**92.2**	**77.0**	**79.8**	**91.0**	**56.8**
Tetracycline	99.1	18.5	1.2 (1.1–1.4)	28.9	98.4	54.9	95.3	78.5	87.2
Gentamicin	100	7.4	1.1 (1.0–1.2)	26.5	100	51.9	100	76.4	100
Rifampin	99.1	5.6	1.1 (1.0–1.1)	25.9	94.9	51.2	86.0	75.9	67.2

[a]Antibiotics excluded due to resistance >90% included erythromycin; antibiotics excluded due to susceptibility >99% included trimethoprim-sulfamethoxazole, linezolid, Synercid, and vancomycin.
[b]Bolded text indicates the most discriminatory antibiotic(s) to predict USA300 MRSA.

clindamycin-susceptible rule and 53.0%, 77.1%, and 91.0% for the levofloxacin-suscep-tible rule, respectively. The PPVs for identifying USA300 strains for the combined clin-damycin or levofloxacin-susceptible rule were 52.8%, 77.0%, and 91.0% when the prev-alence of USA300 was 25%, 50%, and 75%, respectively.

The Matthews correlation coefficient (MCC) for USA300 identification by the com-bined clindamycin or levofloxacin-susceptible rule, which inherently considers the dominance of USA300 in our data, is fair for all isolates (0.393) but strong for blood-stream-only isolates (0.754). Upon closer evaluation, the improved performance of the prediction rule in bloodstream-only isolates appears to be related to a group of

Concordant
Discordant

No data
Clindamycin resistant
Susceptible
Levofloxacin resistant
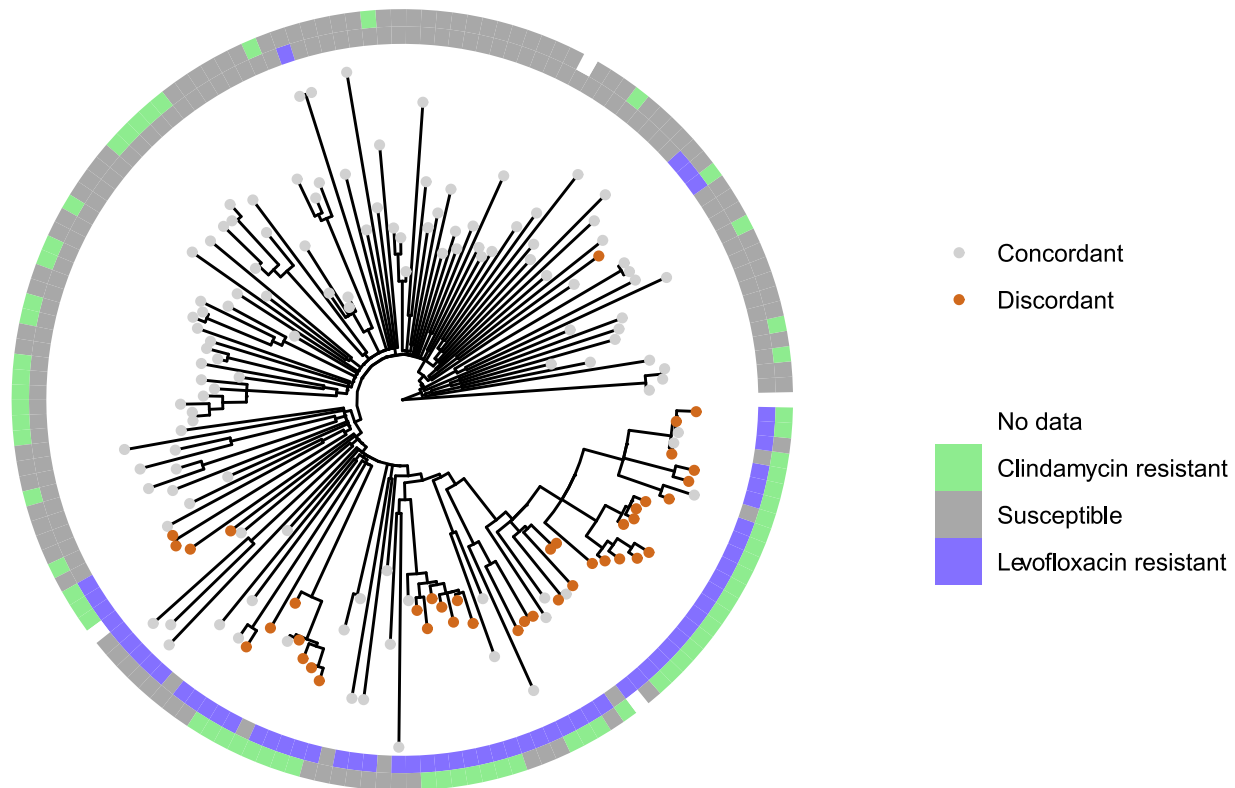
FIG 1 Phylogenetic tree of USA300 incarceration isolates annotated with prediction concordance or discordance. A recombination-masked whole-genome alignment was used to make a maximum likelihood phylogeny of USA300 clinical isolates from incarcerated individuals. The tree is midpoint rooted and annotated with resistance or susceptibility to clindamycin and levofloxacin. Tree tips are labeled with prediction discordance (resistance to both clindamycin and levofloxacin) or prediction concordance (susceptibility to clindamycin, levofloxacin, or both). Clustering of isolates discordant with the prediction rule is observed, indicating circulation of dual clindamycin-levofloxacin-resistant USA300.

Sansom et al.
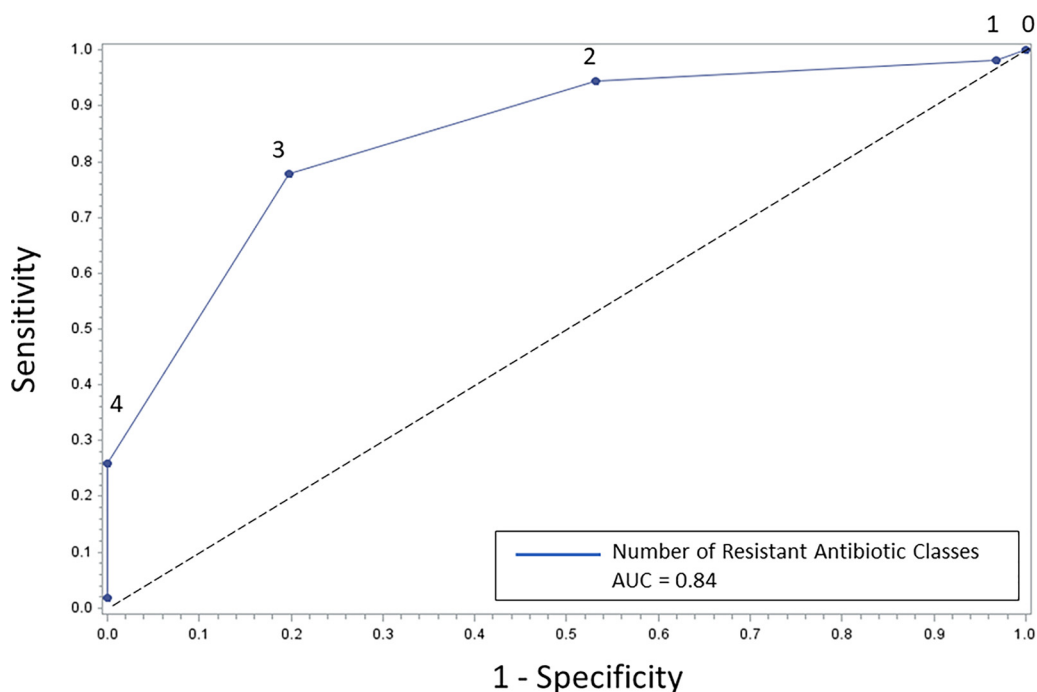
FIG 2 Number of resistant antibiotic classes as a predictor of USA300 MRSA. The numbers above the curve refer to the number of antibiotic classes to which the isolates expressed resistance. A cutoff point of resistance to ≤2 classes for USA300 and to ≥3 classes for USA100 isolates is seen (AUC 0.84, 95% CI 0.78 to 0.90). The cutoff point was calculated using the Youden index (48).

USA300 isolates collected from the jail that were resistant to both clindamycin and levofloxacin. Phylogenetic analysis of USA300 jail isolates revealed a cluster characterized by resistance to both antibiotics (Fig. 1).

In a second analysis, we assessed resistance to antibiotics from six different classes. The average number of resistant antibiotic classes was higher for USA100 than for USA300 (3 versus 2, $P < 0.001$). The majority of USA300 isolates were resistant to ≤2 antibiotic classes (80.3%), with no USA300 isolates resistant to >3 antibiotics. The majority of USA100 isolates were resistant to ≥3 classes (77.8%). A cutoff of resistance to less than 3 antibiotic classes had the best diagnostic characteristics on a generated receiver-operator curve (ROC) (sensitivity 80.3%, specificity 77.8%, LR 3.6 [95% CI 2.2 to 6.0]; area under the curve [AUC] 0.84 [95% CI 0.78 to 0.90]). (Fig. 2) This suggests an optimal cutoff point of resistance to ≤2 classes to predict USA300 and ≥3 classes to predict USA100.

DISCUSSION

Molecular subtyping methods, such as PFGE, do not sufficiently differentiate endemic MRSA strain types; therefore, use of these methods has previously limited our ability to understand transmission patterns of MRSA (24, 26). WGS can detect genetic differences between MRSA strains that are unidentifiable by other molecular epidemiologic methods and more accurately categorize MRSA strain types (26, 28). In this study, we evaluated antibiotic resistance phenotype as a predictor of genotype for WGS-defined USA300 strains from clinical MRSA isolates. We found that clindamycin and levofloxacin susceptibility were the best individual antibiotics for prediction of USA300, with the best overall performance seen with combined susceptibility to clindamycin or levofloxacin. The number of resistant antibiotics was useful, with a cutoff of ≤2 resistant antibiotic classes being predictive of USA300. Importantly, our analysis also identified a novel cluster of USA300 MRSA strains with increased resistance among individuals with incarceration history.

Volume 9   Issue 1   e00376-21   MicrobiolSpectrum.asm.org   **4**

In comparison to our previously developed MRSA phenotypic prediction rule (27), clindamycin and fluoroquinolone susceptibility was found to be less sensitive and specific for prediction of CA-MRSA. In the current study, we used genomic sequencing to categorize MRSA strains as USA300 or USA100 and did not rely on PFGE as in our prior study (27), thus improving the accuracy of our derived rules. Additionally, the prior analysis utilized combinations of HA-MRSA (i.e., USA100, USA500, and USA800) and CA-MRSA (USA300 and USA400) isolates instead of limiting to the most common strain types (USA100 and USA300). Another potential explanation for these observed changes in the rule performance could be increasing prevalence of CA-MRSA antibiotic resistance. Increasing rates of clindamycin resistance among USA300 strains have been described (29), and regional variation exists for USA300 susceptibility patterns (22, 30). Although there are variations across the United States in rates of resistance of USA300 MRSA isolates to clindamycin, there are several studies that show nonsignificant trends that clindamycin may be better for reduction of recurrence and may be favored in some subgroups, such as children (31–33). Levofloxacin is not an agent that is typically used for the treatment of MRSA, but it remains clinically relevant for epidemiologic analyses. Our results support the importance of routinely updating antibiograms for MRSA isolates in both community and acute care settings. Clinicians should continue to submit isolates for culture and antimicrobial susceptibility to ensure continued monitoring for emergence of resistant strains and to inform epidemiologic analyses.

Previous attempts to classify MRSA as CA and HA strains included important clinical differences, such as antimicrobial resistance (24). However, exceptions to these rules have become more common (6, 22, 34). Additionally, as USA300 becomes a more ubiquitous cause of HA-MRSA infections, it has become less clear where acquisition of strains may be occurring (6, 19). Beyond the increased granularity of strain typing enabled by WGS, the inclusion of fully sequenced isolates in our study allowed phylogenetic analysis that revealed a closely related subset of USA300 jail isolates with increased antimicrobial resistance. This finding highlights the added strength of this phenotypic rule in that deviations from the rule may be useful to identify emergence of new resistant strains. This strain is unique from MDR USA300 clusters described in other cities, including San Francisco, California (22), and Austin, Texas (34). The prevalence of MRSA colonization at jail intake at our facility has been observed at nearly 1 in 5 inmates, with colonization associated with high-risk exposures in the community (i.e., injection drug use, HIV infection, unstable housing, and methamphetamine use) (35). High rates of recidivism and prior demonstration of MRSA community transmission networks among detainees (35) suggests that this more resistant USA300 MRSA cluster could have far-reaching implications for the treatment of MRSA infection in both the incarcerated population and associated at-risk community networks (i.e., where these individuals live).

A limitation of this study is that we included only the two most common MRSA clinical isolates (USA300, USA100) for analysis; other circulating MRSA strains (e.g., USA500 and USA800) were not included as they are less commonly isolated. So, it is unclear how the phenotypic prediction rule would perform when applied to non-USA300/non-USA100 MRSA strains. Additionally, this analysis was focused regionally and included fewer USA100 isolates than USA300 isolates. It may be useful for a community or health care setting to be familiar with circulating MRSA clones; our methodology could be adapted to a setting with differing prevalence of various MRSA strain types. Another limitation is that samples were collected from clinical isolates over time and likely reflect secular changes. Finally, the phenotypes described here reflect local antimicrobial resistance patterns (i.e., our local antibiogram) and may not be generalizable to all clinical settings. Differences in regional antimicrobial resistance patterns should be considered when applying these suggested phenotypic rules. A predictive phenotypic rule can easily be updated to account for changing antimicrobial resistance patterns. For example, to adapt this rule to a new region, researchers would submit a representative sample of clinical MRSA isolates for strain typing followed by analysis of

antimicrobial resistance patterns. Similarly, changes in the antibiogram for MRSA, such as increasing antimicrobial resistance, would prompt the application of the same methodology to update the predictive phenotypic rule.

A phenotypic rule based on commonly available clinical information to predict genotype may be useful for epidemiologic applications, especially in the analysis of retrospective data or in a setting where strain typing is not available. It is reassuring that the prior PFGE-derived phenotypic rule has changed little with updates in the WGS strain typing methods among bloodstream isolates. However, a cluster of more resistant USA300 MRSA isolates among individuals with incarceration exposures suggests that phenotypic rules should be interpreted carefully, as local epidemiology is key for accuracy. As multidrug resistance becomes an increasing challenge and emerges in the community, this rule can be adapted to accommodate these shifts.

In this study, we updated a previously derived phenotypic prediction rule for USA300 by using WGS. Clindamycin and levofloxacin susceptibility were the best individual predictors for USA300, with the best overall performance seen with a two-antibiotic prediction rule with susceptibility to clindamycin or levofloxacin. However, such phenotypic rules should be applied and interpreted in the context of local MRSA epidemiology. Emergence of more resistant USA300 MRSA among individuals with exposure to incarceration has important treatment and infection control implications for at-risk community networks.

MATERIALS AND METHODS

Study design. Cook County Health (CCH) is the major safety net health care network in Cook County and has a comprehensive, integrated system of health care throughout Chicago, Illinois, and suburban Cook County. Included in this system are hospitals (including Stroger Hospital of Cook County), ambulatory and community health network clinics, a correctional health care facility, and an outpatient infectious disease clinic. We utilized a deduplicated convenience sample of existing MRSA clinical isolates from 2007 to 2017 cultured from individuals 18 years of age and older who received care at CCH. Clinical isolates were primarily from skin and soft tissue infections and bloodstream infections, but body fluid, urine, and respiratory tract isolates were also included. Nares surveillance cultures were excluded. Isolates were selected based on previously identified strain type using WGS. Antibiotic susceptibility was determined by the clinical microbiology laboratory with a MicroScan WalkAway 96 plus (Brea, CA). The study was approved with waiver of consent by the CCH institutional review board.

Genomic analysis. We utilized MRSA isolates related to USA300 or USA100 that had already undergone genomic sequencing as well as sequenced additional historical PFGE-predicted isolates to enrich the sample with USA100 strains. A total of 170 previously sequenced isolates from the Cook County Jail (8 USA100, 162 USA300), predominantly from skin and soft tissue infections (BioProject PRJNA638400) (35), as well as 71 previously sequenced isolates from MRSA bacteremia at Cook County Hospital (17 USA100, 54 USA300) were included (BioProject PRJNA345238) (6). Due to the low number of available previously sequenced USA100 isolates, an additional 34 contemporaneous PFGE-predicted USA100 bacteremia isolates at Cook County Hospital were submitted for sequencing (29 USA100, 5 USA300). Genomic DNA extracted from MRSA isolates submitted under BioProject PRJNA744410 was prepared for sequencing using a Nextera XT library preparation kit (Illumina, San Diego, CA) according to the manufacturer's instructions. Sequencing was performed on an Illumina NextSeq500 instrument using a high-output kit with paired-end 2 × 150 bp reads. Library preparation and sequencing were performed at the University of Illinois Chicago Sequencing Core (UICSQC). Details on sequenced strains are available in Table S1 in the supplemental material. The quality of sequencing reads was assessed with FastQC (36), and Trimmomatic (37) was used for trimming adapter sequences and low-quality bases.

Phylogenetic analysis and *in silico* PFGE classification. Results of our revised prediction rule suggested existence of a more resistant strain population that we studied with additional phylogenetic analyses. Single nucleotide variants (SNVs) were identified by mapping cleaned sequencing reads against a finished USA300 MRSA reference genome and applying a series of filters to retain high quality variants. The variant calling pipeline can be found on GitHub at https://github.com/Snitkin-Lab-Umich/variant_calling_pipeline. After generating the whole-genome alignment for the isolates from the Cook County jail (PRJNA638400), we then masked sites identified as recombinant by Gubbins (38) and used this masked whole-genome alignment to build a maximum likelihood phylogeny with IQ-TREE version 1.6.12 (39) using the ModelFinder option to choose the best fit substitution model and ultrafast bootstrap with 1,000 replicates (-bb 1000) (40). The tree was midpoint rooted using the midpoint.root function in phytools (41). Non-USA300 tips were dropped using the drop.tip function in ape (42). Resistance data were overlaid on the tree with ggtree (43).

We classified isolates as USA300 using sequence probes provided by Bowers et al. (44) that included the probes innerCC8, CC8f, and CC8e or CC8e and CC8f. Isolates underwent multilocus sequence typing with ARIBA (45). Isolates that were not part of CC8 were further investigated for relatedness to USA100. Two USA100 genomes and a USA800 genome (USA100-NRS642, USA100-NRS720, USA800-NRS387 from BioProject PRJNA374337) and non-CC8 genomes from our study were assembled with an internal pipeline available at https://github.com/Snitkin-Lab-Umich/assemblage and aligned with cognac (46), and a

phylogenetic tree was created with fasttree (47). Isolates with ST5 or that clustered with the USA100 reference genomes were considered USA100 (Fig. S1).

Statistical analysis. Analysis was performed using SAS software (version 9.4, SAS Institute, Cary, NC), Python (version 3.7, Python Software Foundation, available at http://www.python.org), and R (R version 4.0.3, R Foundation for Statistical Computing, Vienna, Austria.). The sensitivity (Sn), specificity (Sp), positive and negative predictive value (PPV and NPV), and likelihood ratios (LRs) were calculated for each unique antibiotic phenotype and antibiotic combinations. For these analyses, "true positive" was defined as USA300 that was susceptible to the antibiotic(s), "false negative" was defined as USA300 that was resistant to the antibiotic(s), "false positive" was defined as USA100 that was susceptible to the antibiotic(s), and "true negative" was defined as USA100 that was resistant to the antibiotic(s). The Matthews correlation coefficient (MCC) was calculated to assess performance of the highest-performing predictor to determine its predictive performance while accounting for unbalanced data sets. Antibiotics with very high levels of resistance (>90%; erythromycin) or susceptibility (>99%; trimethoprim-sulfamethoxazole, linezolid, Synercid, and vancomycin) were excluded from this analysis because they were unlikely to be discriminatory. The antibiotic phenotype for each unique isolate was then used to calculate the number of resistant antibiotic classes (using tetracycline, levofloxacin, rifampin, gentamicin, erythromycin, and clindamycin), from which a receiver-operator curve (ROC) was constructed.

Data availability. Sequencing data sets are available and can be found at BioProject PRJNA638400 (35), PRJNA345238 (6), and BioProject PRJNA744410.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.
SUPPLEMENTAL FILE 1, XLSX file, 0.02 MB.
SUPPLEMENTAL FILE 2, PDF file, 0.03 MB.
SUPPLEMENTAL FILE 3, PDF file, 0.1 MB.

ACKNOWLEDGMENTS

REFERENCES

1. Grundmann H, Aires-de-Sousa M, Boyce J, Tiemersma E. 2006. Emergence and resurgence of meticillin-resistant *Staphylococcus aureus* as a public-health threat. Lancet 368:874–885. https://doi.org/10.1016/S0140-6736(06)68853-3.

2. Uhlemann AC, Otto M, Lowy FD, DeLeo FR. 2014. Evolution of community- and healthcare-associated methicillin-resistant *Staphylococcus aureus*. Infect Genet Evol 21:563–574. https://doi.org/10.1016/j.meegid.2013.04.030.

3. Roberts RB, Chung M, de Lencastre H, Hargrave J, Tomasz A, Nicolau DP, John JF, Jr, Korzeniowski O, Tri-State MRSA Collaborative Study Group. 2000. Distribution of methicillin-resistant *Staphylococcus aureus* clones among health care facilities in Connecticut, New Jersey, and Pennsylvania. Microb Drug Resist 6:245–251. https://doi.org/10.1089/mdr.2000.6.245.

4. Ray GT, Suaya JA, Baxter R. 2012. Trends and characteristics of culture-confirmed *Staphylococcus aureus* infections in a large U.S. integrated health care organization. J Clin Microbiol 50:1950–1957. https://doi.org/10.1128/JCM.00134-12.

5. Rhee Y, Aroutcheva A, Hota B, Weinstein RA, Popovich KJ. 2015. Evolving epidemiology of *Staphylococcus aureus* bacteremia. Infect Control Hosp Epidemiol 36:1417–1422. https://doi.org/10.1017/ice.2015.213.

6. Popovich KJ, Snitkin ES, Hota B, Green SJ, Pirani A, Aroutcheva A, Weinstein RA. 2017. Genomic and epidemiological evidence for community origins of hospital-onset methicillin-resistant *Staphylococcus aureus* bloodstream infections. J Infect Dis 215:1640–1647. https://doi.org/10.1093/infdis/jiw647.

7. Seybold U, Kourbatova EV, Johnson JG, Halvosa SJ, Wang YF, King MD, Ray SM, Blumberg HM. 2006. Emergence of community-associated methicillin-resistant *Staphylococcus aureus* USA300 genotype as a major cause of health care-associated blood stream infections. Clin Infect Dis 42:647–656. https://doi.org/10.1086/499815.

8. Fridkin SK, Hageman JC, Morrison M, Sanza LT, Como-Sabetti K, Jernigan JA, Harriman K, Harrison LH, Lynfield R, Farley MM, Active Bacterial Core Surveillance Program of the Emerging Infections Program Network. 2005. Methicillin-resistant *Staphylococcus aureus* disease in three communities. N Engl J Med 352:1436–1444. https://doi.org/10.1056/NEJMoa043252.

9. Herold BC, Immergluck LC, Maranan MC, Lauderdale DS, Gaskin RE, Boyle-Vavra S, Leitch CD, Daum RS. 1998. Community-acquired methicillin-resistant *Staphylococcus aureus* in children with no identified predisposing risk. JAMA 279:593–598. https://doi.org/10.1001/jama.279.8.593.

10. David MZ, Daum RS. 2010. Community-associated methicillin-resistant *Staphylococcus aureus*: epidemiology and clinical consequences of an emerging epidemic. Clin Microbiol Rev 23:616–687. https://doi.org/10.1128/CMR.00081-09.

11. Pardos de la Gandara M, Curry M, Berger J, Burstein D, Della-Latta P, Kopetz V, Quale J, Spitzer E, Tan R, Urban C, Wang G, Whittier S, de Lencastre H, Tomasz A. 2016. MRSA causing infections in hospitals in greater metropolitan New York: major shift in the dominant clonal type between 1996 and 2014. PLoS One 11:e0156924. https://doi.org/10.1371/journal.pone.0156924.

12. Kallen AJ, Mu Y, Bulens S, Reingold A, Petit S, Gershman K, Ray SM, Harrison LH, Lynfield R, Dumyati G, Townes JM, Schaffner W, Patel PR, Fridkin SK, Active Bacterial Core surveillance (ABCs) MRSA Investigators of the Emerging Infections Program. 2010. Health care-associated invasive MRSA infections, 2005-2008. JAMA 304:641–648. https://doi.org/10.1001/jama.2010.1115.

13. Iwamoto M, Mu Y, Lynfield R, Bulens SN, Nadle J, Aragon D, Petit S, Ray SM, Harrison LH, Dumyati G, Townes JM, Schaffner W, Gorwitz RJ, Lessa FC. 2013. Trends in invasive methicillin-resistant *Staphylococcus aureus* infections. Pediatrics 132:e817–e824. https://doi.org/10.1542/peds.2013-1112.

14. Dantes R, Mu Y, Belflower R, Aragon D, Dumyati G, Harrison LH, Lessa FC, Lynfield R, Nadle J, Petit S, Ray SM, Schaffner W, Townes J, Fridkin S, Emerging Infections Program-Active Bacterial Core Surveillance MRSA Surveillance Investigators. 2013. National burden of invasive methicillin-resistant *Staphylococcus aureus* infections, United States, 2011. JAMA Intern Med 173:1970–1978. https://doi.org/10.1001/jamainternmed.2013.10423.

15. Nguyen DB, Lessa FC, Belflower R, Mu Y, Wise M, Nadle J, Bamberg WM, Petit S, Ray SM, Harrison LH, Lynfield R, Dumyati G, Thompson J,

Schaffner W, Patel PR, Active Bacterial Core Surveillance (ABCs) MRSA Investigators of the Emerging Infections Program. 2013. Invasive methicillin-resistant *Staphylococcus aureus* infections among patients on chronic dialysis in the United States, 2005-2011. Clin Infect Dis 57:1393–1400. https://doi.org/10.1093/cid/cit546.

16. Popovich KJ, Weinstein RA, Hota B. 2008. Are community-associated methicillin-resistant *Staphylococcus aureus* (MRSA) strains replacing traditional nosocomial MRSA strains? Clin Infect Dis 46:787–794. https://doi.org/10.1086/528716.

17. Tenover FC, Tickler IA, Goering RV, Kreiswirth BN, Mediavilla JR, Persing DH, Consortium M, MRSA Consortium. 2012. Characterization of nasal and blood culture isolates of methicillin-resistant *Staphylococcus aureus* from patients in United States Hospitals. Antimicrob Agents Chemother 56:1324–1330. https://doi.org/10.1128/AAC.05804-11.

18. King MD, Humphrey BJ, Wang YF, Kourbatova EV, Ray SM, Blumberg HM. 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. https://doi.org/10.7326/0003-4819-144-5-200603070-00005.

19. Popovich KJ, Weinstein RA. 2009. Commentary: the graying of methicillin-resistant *Staphylococcus aureus*. Infect Control Hosp Epidemiol 30:9–12. https://doi.org/10.1086/592709.

20. Diekema DJ, Richter SS, Heilmann KP, Dohrn CL, Riahi F, Tendolkar S, McDanel JS, Doern GV. 2014. Continued emergence of USA300 methicillin-resistant *Staphylococcus aureus* in the United States: results from a nationwide surveillance study. Infect Control Hosp Epidemiol 35:285–292. https://doi.org/10.1086/675283.

21. Talan DA, Krishnadasan A, Gorwitz RJ, Fosheim GE, Limbago B, Albrecht V, Moran GJ, EMERGEncy ID Net Study Group. 2011. Comparison of *Staphylococcus aureus* from skin and soft-tissue infections in US emergency department patients, 2004 and 2008. Clin Infect Dis 53:144–149. https://doi.org/10.1093/cid/cir308.

22. Diep BA, Chambers HF, Graber CJ, Szumowski JD, Miller LG, Han LL, Chen JH, Lin F, Lin J, Phan TH, Carleton HA, McDougal LK, Tenover FC, Cohen DE, Mayer KH, Sensabaugh GF, Perdreau-Remington F. 2008. Emergence of multidrug-resistant, community-associated, methicillin-resistant *Staphylococcus aureus* clone USA300 in men who have sex with men. Ann Intern Med 148:249–257. https://doi.org/10.7326/0003-4819-148-4-200802190-00204.

23. Popovich KJ, Snitkin E, Green SJ, Aroutcheva A, Hayden MK, Hota B, Weinstein RA. 2016. Genomic epidemiology of USA300 methicillin-resistant *Staphylococcus aureus* in an urban community. Clin Infect Dis 62:37–44. https://doi.org/10.1093/cid/civ794.

24. David MZ, Taylor A, Lynfield R, Boxrud DJ, Short G, Zychowski D, Boyle-Vavra S, Daum RS. 2013. Comparing pulsed-field gel electrophoresis with multilocus sequence typing, spa typing, staphylococcal cassette chromosome *mec* (SCC*mec*) typing, and PCR for panton-valentine leukocidin, *arcA*, and *opp3* in methicillin-resistant *Staphylococcus aureus* isolates at a U.S. medical center. J Clin Microbiol 51:814–819. https://doi.org/10.1128/JCM.02429-12.

25. Salipante SJ, SenGupta DJ, Cummings LA, Land TA, Hoogestraat DR, Cookson BT. 2015. Application of whole-genome sequencing for bacterial strain typing in molecular epidemiology. J Clin Microbiol 53:1072–1079. https://doi.org/10.1128/JCM.03385-14.

26. Popovich KJ, Snitkin ES. 2017. Whole genome sequencing-implications for infection prevention and outbreak investigations. Curr Infect Dis Rep 19:15. https://doi.org/10.1007/s11908-017-0570-0.

27. Popovich K, Hota B, Rice T, Aroutcheva A, Weinstein RA. 2007. Phenotypic prediction rule for community-associated methicillin-resistant *Staphylococcus aureus*. J Clin Microbiol 45:2293–2295. https://doi.org/10.1128/JCM.00044-07.

28. Durand G, Javerliat F, Bes M, Veyrieras JB, Guigon G, Mugnier N, Schicklin S, Kaneko G, Santiago-Allexant E, Bouchiat C, Martins-Simoes P, Laurent F, Van Belkum A, Vandenesch F, Tristan A. 2018. Routine whole-genome sequencing for outbreak investigations of *Staphylococcus aureus* in a national reference center. Front Microbiol 9:511. https://doi.org/10.3389/fmicb.2018.00511.

29. Sader HS, Mendes RE, Streit JM, Flamm RK. 2017. Antimicrobial susceptibility trends among *Staphylococcus aureus* isolates from U.S. hospitals: results from 7 years of the ceftaroline (AWARE) surveillance program, 2010 to 2016. Antimicrob Agents Chemother 61:e01043-17. https://doi.org/10.1128/AAC.01043-17.

30. Han LL, McDougal LK, Gorwitz RJ, Mayer KH, Patel JB, Sennott JM, Fontana JL. 2007. High frequencies of clindamycin and tetracycline

resistance in methicillin-resistant *Staphylococcus aureus* pulsed-field type USA300 isolates collected at a Boston ambulatory health center. J Clin Microbiol 45:1350–1352. https://doi.org/10.1128/JCM.02274-06.

31. Daum RS, Miller LG, Immergluck L, Fritz S, Creech CB, Young D, Kumar N, Downing M, Pettibone S, Hoagland R, Eells SJ, Boyle MG, Parker TC, Chambers HF, DMID 07-0051 Team. 2017. A placebo-controlled trial of antibiotics for smaller skin abscesses. N Engl J Med 376:2545–2555. https://doi.org/10.1056/NEJMoa1607033.

32. Miller LG, Daum RS, Creech CB, Young D, Downing MD, Eells SJ, Pettibone S, Hoagland RJ, Chambers HF, Team D, DMID 07–0051 Team. 2015. Clindamycin versus trimethoprim-sulfamethoxazole for uncomplicated skin infections. N Engl J Med 372:1093–1103. https://doi.org/10.1056/NEJMoa1403789.

33. Cluzet VC, Gerber JS, Nachamkin I, Metlay JP, Zaoutis TE, Davis MF, Julian KG, Linkin DR, Coffin SE, Margolis DJ, Hollander JE, Bilker WB, Han X, Mistry RD, Gavin LJ, Tolomeo P, Wise JA, Wheeler MK, Hu B, Fishman NO, Royer D, Lautenbach E. 2015. Risk factors for recurrent colonization with methicillin-resistant *Staphylococcus aureus* in community-dwelling adults and children. Infect Control Hosp Epidemiol 36:786–793. https://doi.org/10.1017/ice.2015.76.

34. Lee GC, Dallas SD, Wang Y, Olsen RJ, Lawson KA, Wilson J, Frei CR. 2017. Emerging multidrug resistance in community-associated *Staphylococcus aureus* involved in skin and soft tissue infections and nasal colonization. J Antimicrob Chemother 72:2461–2468. https://doi.org/10.1093/jac/dkx200.

35. Popovich KJ, Snitkin ES, Zawitz C, Aroutcheva A, Payne D, Thiede SN, Schoeny M, Green SJ, Hayden MK, Hota B, Weinstein RA. 2020. Frequent methicillin-resistant *Staphylococcus aureus* introductions into an inner-city jail: indications of community transmission networks. Clin Infect Dis 71:323–331. https://doi.org/10.1093/cid/ciz818.

36. Babraham Institute. Babraham Bioinformatics - FastQC a quality control tool for high throughput sequence data. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/. Accessed June 16 2020.

37. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120. https://doi.org/10.1093/bioinformatics/btu170.

38. Croucher NJ, Page AJ, Connor TR, Delaney AJ, Keane JA, Bentley SD, Parkhill J, Harris SR. 2015. Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. Nucleic Acids Res 43:e15. https://doi.org/10.1093/nar/gku1196.

39. Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol 32:268–274. https://doi.org/10.1093/molbev/msu300.

40. Minh BQ, Nguyen MA, von Haeseler A. 2013. Ultrafast approximation for phylogenetic bootstrap. Mol Biol Evol 30:1188–1195. https://doi.org/10.1093/molbev/mst024.

41. Revell LJ. 2012. phytools: an R package for phylogenetic comparative biology (and other things). Methods Ecol and Evol 3:217–223. https://doi.org/10.1111/j.2041-210X.2011.00169.x.

42. Paradis E, Schliep K. 2019. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics 35:526–528. https://doi.org/10.1093/bioinformatics/bty633.

43. Yu G, Smith DK, Zhu H, Guan Y, Lam TT. 2017. ggtree: an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. Methods Ecol Evol 8:28–36. https://doi.org/10.1111/2041-210X.12628.

44. Bowers JR, Driebe EM, Albrecht V, McDougal LK, Granade M, Roe CC, Lemmer D, Rasheed JK, Engelthaler DM, Keim P, Limbago BM. 2018. Improved subtyping of *Staphylococcus aureus* clonal complex 8 strains based on whole-genome phylogenetic analysis. mSphere 3:e00464-17. https://doi.org/10.1128/mSphere.00464-17.

45. Hunt M, Mather AE, Sanchez-Buso L, Page AJ, Parkhill J, Keane JA, Harris SR. 2017. ARIBA: rapid antimicrobial resistance genotyping directly from sequencing reads. Microb Genom 3:e000131. https://doi.org/10.1099/mgen.0.000131.

46. Crawford RD, Snitkin ES. 2021. cognac: rapid generation of concatenated gene alignments for phylogenetic inference from large, bacterial whole genome sequencing datasets. BMC Bioinformatics 22:70. https://doi.org/10.1186/s12859-021-03981-4.

47. Price MN, Dehal PS, Arkin AP. 2010. FastTree 2–approximately maximum-likelihood trees for large alignments. PLoS One 5:e9490. https://doi.org/10.1371/journal.pone.0009490.

48. Youden WJ. 1950. Index for rating diagnostic tests. Cancer 3:32–35. https://doi.org/10.1002/1097-0142(1950)3:1<32::AID-CNCR2820030106>3.0.CO;2-3.