JB Journals.ASM.org

Genome Sequence-Based Discriminator for Vancomycin-Intermediate *Staphylococcus aureus*

Lavanya Rishishwar,[a,b] Robert A. Petit III,[c] Colleen S. Kraft,[c] I. King Jordan[a,b]

School of Biology, Georgia Institute of Technology, Atlanta, Georgia, USA[a]; PanAmerican Bioinformatics Institute, Santa Marta, Colombia[b]; Division of Infectious Diseases, Emory University, Atlanta, Georgia, USA[c]

Vancomycin is the mainstay of treatment for patients with *Staphylococcus aureus* infections, and reduced susceptibility to vancomycin is becoming increasingly common. Accordingly, the development of rapid and accurate assays for the diagnosis of vancomycin-intermediate *S. aureus* (VISA) will be critical. We developed and applied a genome-based machine-learning approach for discrimination between VISA and vancomycin-susceptible *S. aureus* (VSSA) using 25 whole-genome sequences. The resulting machine-learning model, based on 14 gene parameters, including 3 molecular typing markers and 11 genes implicated in reduced vancomycin susceptibility, is able to unambiguously distinguish between the VISA and VSSA isolates analyzed here despite the fact that they do not form evolutionarily distinct groups. As such, the model is able to discriminate based on specific genomic markers of antibiotic susceptibility rather than overall sequence relatedness. Subsequent evaluation of the model using leave-one-out validation yielded a classification accuracy of 84%. The machine-learning approach described here provides a generalized framework for the application of genome sequence analysis to the classification of bacteria that differ with respect to clinically relevant phenotypes and should be particularly useful in defining the genomic features that underlie antibiotic resistance.

Staphylococcus aureus is the leading cause of a broad spectrum of infections, including skin and soft-tissue infections, bacteremia, and endocarditis, in patients in the United States (1). Mortality from severe *S. aureus* infections can reach 20 to 30% (2), and treatment requires a prolonged course of antibiotics (3). Antibiotic resistance presents a major challenge to the effective treatment of *S. aureus* infections. The first methicillin-resistant isolates of *S. aureus* were identified in the early 1960s, and the mechanism of resistance to methicillin, which involves the acquisition of the *mecA* gene (methicillin-resistant *S. aureus* [MRSA]) (Fig. 1), is well understood (4). *S. aureus* isolates identified as methicillin resistant are typically treated with vancomycin, and strains with various degrees of vancomycin resistance are also becoming increasingly prevalent (5, 6).

S. aureus isolates that have a high MIC to vancomycin of >16 μg/μl are considered fully resistant (VRSA) (Fig. 1), and this phenotype results from the acquisition of the *vanA* gene found on the Tn*1546* transposon (7, 8). Isolates of *S. aureus* with reduced susceptibility, but not full resistance, to vancomycin form the category of vancomycin-intermediate *S. aureus* (VISA) (Fig. 1) (9). Unlike the single-gene determinants of full methicillin and vancomycin resistance, the genetic basis of the vancomycin-intermediate phenotype appears to be more complex, likely involving multiple genes, and is less well understood (10). Studies of the phenotypic characteristics of VISA as well as the identification of single-nucleotide polymorphisms (SNPs) that distinguish VISA strains (10) have led to the proposal of a number of candidate genes that may underlie vancomycin-intermediate susceptibility in *S. aureus*.

The genes that have been implicated as vancomycin intermediate can be functionally classified into three groups: (i) genes related to cellular processing and signaling, (ii) genes related to information storage and processing, and (iii) genes related to metabolism (Fig. 1) (11). A number of these genes have functions that are related to cell wall integrity, since vancomycin acts to disrupt cell wall biosynthesis by binding to peptidoglycan precursors (9). Indeed, VISA isolates appear to have thicker cell walls (12) along with an increase in the D-alanyl-D-alanine cell wall constituent residues, which are thought to be the target for the vancomycin glycopeptide, thereby reducing its efficacy (13). SNPs within several classes of cell wall-related genes have been associated with reduced vancomycin susceptibility, including a number of transcriptional regulators, *agrA*, *agrC*, and *rpoB*; members of two-component regulatory systems, *vraSR*, *graSR*, and *walKR*; penicillin binding proteins *pbp4* and *pbpB*; and the phosphatase *stp1* (10, 13–21).

It should be noted that the VISA phenotype was initially thought to be exclusively associated with a subset of the regulators described above, including the *agr* genes, but VISA has now been demonstrated in strains lacking any changes in these and other regulators (9). As such, it is likely to be the case that vancomycin-intermediate susceptibility can also be encoded by genes related to other aspects of antimicrobial resistance. Consistent with this idea, genes that encode the antibiotic degrading enzyme beta-lactamase recently have been implicated in the VISA phenotype (Fig. 1) (22, 23). In the case of vancomycin, beta-lactamase and related genes are thought to mitigate susceptibility by stimulating pathways involved in cell wall maintenance and repair.

The ability to distinguish *S. aureus* clinical isolates with different antibiotic-resistant phenotypes early on is critically important

Received 5 December 2013 **Accepted** 12 December 2013

Published ahead of print 20 December 2013

Address correspondence to I. King Jordan, king.jordan@biology.gatech.edu.

Supplemental material for this article may be found at http://dx.doi.org/10.1128/JB.01410-13.

doi:10.1128/JB.01410-13

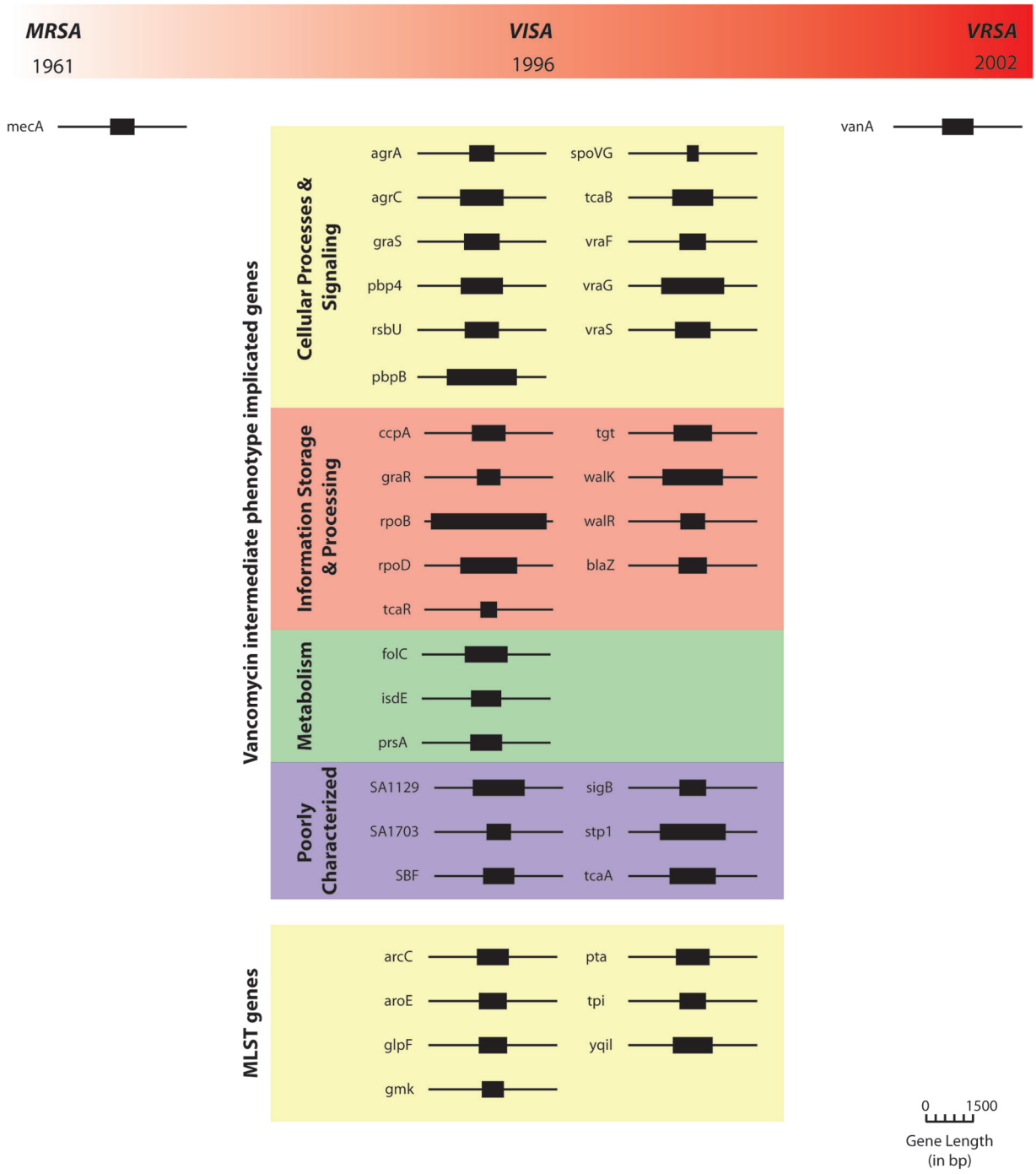

FIG 1 Genetic determinants of reduced antibiotic susceptibility in *S. aureus*. Single genes known to determine resistance to methicillin (MRSA) and vancomycin (VRSA) are shown along with groups of genes associated with reduced susceptibility to vancomycin (VISA).

for patient treatment decisions. Identification of MRSA and VRSA isolates is generally straightforward given the respective single-gene determinants of these phenotypes. The goal of this study was to take advantage of the availability of the VISA-associated genes described above, together with a number of other genomic characteristics of *S. aureus* isolates, in order to build a genome-based discriminator between VISA and VSSA isolates. To do this, we chose a machine-learning computational approach (24) that entails the use of multiple genomic features, i.e., genes and genome-wide characteristics, for discrimination of VISA and VSSA isolates. The machine-learning computational framework provides for maximum discriminative power along with information as to which of the genomic features employed by the algorithm are ultimately most informative for distinguishing between VISA and VSSA isolates. We show that this approach allows for discrimination between VISA and VSSA isolates even though isolates from

TABLE 1 Publicly available genome sequences of VISA and VSSA strains[a]

Isolate type and ID	Name	GenBank accession no.	PMID
VISA			
GB1	*S. aureus* Mu50	NC_002758	11418146
GB2	*S. aureus* JH9	NC_009487	17517606
GB3	*S. aureus* Mu3	NC_009782	17954695
GB4	*S. aureus* JKD6008	NC_017341	20802046
VSSA			
GB5	*S. aureus* N315	NC_002745	11418146
GB6	*S. aureus* COL	NC_002951	15774886
GB7	*S. aureus* MRSA252	NC_002952	15213324
GB8	*S. aureus* MW2	NC_003923	12044378
GB9	*S. aureus* USA300_FPR3757	NC_007793	16517273
GB10	*S. aureus* Newman	NC_009641	17951380
GB11	*S. aureus* USA300_TCH1516	NC_010079	17986343
GB12	*S. aureus* MSSA476	NC_002953	15213324
GB13	*S. aureus* NCTC 8325	NC_007795	22417616
GB14	*S. aureus* JH1	NC_009632	17517606

[a] Strain names, NCBI GenBank (GB) RefSeq accession numbers, and PubMed identifiers (PMID) are shown for *S. aureus* strains designated VISA or VSSA.

these phenotypically defined classes do not form evolutionarily distinct groups.

(These data were reported in part at the 113th General Meeting of the American Society for Microbiology [25].)

MATERIALS AND METHODS

Publicly available genome sequences. Complete *S. aureus* genome sequences that were publicly available as of May 2012 were downloaded from the NCBI RefSeq database (26) (http://www.ncbi.nlm.nih.gov/genomes/geblast.cgi?taxid=1280), and individual strains were characterized as VISA or VSSA based on published evidence (Table 1). The evolutionary relationships among the complete genome sequences were characterized using 16S rRNA (see Fig. S1 in the supplemental material) and multilocus sequence typing (MLST) analysis (see Fig. 3A). Individual 16S rRNA and MLST locus sequences were taken from the NCBI RefSeq genome annotations, and sequences were aligned using ClustalW (27). The 16S rRNA phylogenetic tree was reconstructed based on the number of differences between sequences using the neighbor-joining (NJ) method (28) implemented in the program MEGA (29). A concatenated sequence alignment of all 7 MLST loci was used to reconstruct the MLST phylogenetic tree using the NJ method in MEGA with Poisson-corrected distances and 500 bootstrap replicates. MLST sequence types (STs) were assigned to individual genome sequences using the *S. aureus* MLST database (http://saureus.mlst.net) (30). Whole-genome sequence trees were reconstructed based on the average nucleotide identities (ANI) computed between pairs of genomes using the program MUMmer (31).

Patient isolates and genome sequencing. *S. aureus* isolates from 11 patients presenting with respiratory symptoms, skin and soft-tissue infections, septic arthritis, or bloodstream infection (Table 2) were collected and cultured in the Emory Healthcare microbiology laboratory (under Institutional Review Board approval 50685). The *S. aureus* isolates were subjected to automated susceptibility testing using the MicroScan Walkaway96 Plus system (Siemens Healthcare Diagnostics Inc., Tarrytown, NY), and vancomycin susceptibility was confirmed by Etest (bioMérieux, Inc., Durham, NC). Isolate DNA extraction was performed using the cetyl trimethyl ammonium bromide (CTAB) protocol. Isolate genome sequencing was performed using either the Ion Torrent PGM 314 chip (isolates CI1 to CI6) or the Illumina MiSeq (isolates CI7 to CI11). This resulted in high-coverage draft genome sequences (~36× for Ion Torrent and ~99× for MiSeq) that were assembled but not finished to yield a single contig.

Machine-learning approach used to distinguish VISA and VSSA isolates. A schematic overview of the machine-learning approach used to discriminate between VISA and VSSA isolates is shown in Fig. 2, and the genome-wide and gene-based parameters are shown in Table 3. Briefly, complete genome sequences of 25 *S. aureus* isolates (step 1) were parameterized in order to quantify the variability seen for a number of genome-wide attributes and specific gene sequences between groups of VISA and VSSA strains (step 2). Based on these parameters and their grouping between VISA and VSSA, a model to distinguish these two groups was built. As part of the model construction, attribute selection was performed in order to converge on the minimal set of maximally informative attributes (step 3). Once the VISA and VSSA classification model was built in this way (step 4), it was used to make predictions about which group an as-yet unseen *S. aureus* genome sequence would belong to (step 5). Cross-fold validation with $K = 25$ (i.e., leave-one-out validation) was used to test the classification accuracy of the model with respect to each individual isolate genome sequence. Details of the machine-learning approach used here can be found in the supplemental material.

SRA accession numbers. Sequencing data have been deposited in the Sequence Read Archive (SRA) database under accession numbers SRX282043, SRX282049, SRX282051, SRX282052, SRX282055, SRX282056, SRX374117, SRX381343, SRX381344, SRX381345, and SRX382695.

TABLE 2 *S. aureus* isolate sources, phenotypes, and genome sequences for clinical isolates characterized from presenting patients

ID	Patient source	Site of isolation	Vancomycin finding		No. of genome reads	Genome coverage	MLST	SRA accession no.
			MIC (μg/ml)	Phenotype				
CI1	Respiratory infection	Cystic fibrosis sputum	2.0	VSSA	157,897	8.86×	ST8[a]	SRX282043
CI2	Soft-tissue infection	Buttocks	2.0	VSSA	229,038	13.36×	ST8[a]	SRX282049
CI3	Respiratory infection	Minibronchoalveolar lavage specimen	2.0	VSSA	405,583	21.94×	ST105[a]	SRX282051
CI4	Soft tissue infection	Abdominal	2.0	VSSA	282,749	15.18×	ST5[a]	SRX282052
CI5	Respiratory infection	Endotracheal aspirate	2.0	VSSA	346,920	18.70×	ST105[a]	SRX282055
CI6	Soft tissue infection	Abdomen	2.0	VSSA	2,511,097	137.53×	ST5[a]	SRX282056
CI7	Synovial fluid	Knee	1.5	VSSA	4,198,318	220.37×	ST105[b]	SRX374117
CI8	Respiratory infection	Sputum	3.0	VISA	1,100,692	85.95×	ST8[b]	SRX381343
CI9	Synovial fluid	Knee	3.0	VISA	901,940	70.20×	ST5[b]	SRX381344
CI10	Bacteremia	Catheter tip	3.0	VISA	616,116	99.26×	ST5[b]	SRX381345
CI11	Respiratory infection	Cystic fibrosis sputum	3.0	VISA	455,327	65.98×	ST225[b]	SRX382695

[a] Reported by Ion Torrent.
[b] Determined from sequence similarity.

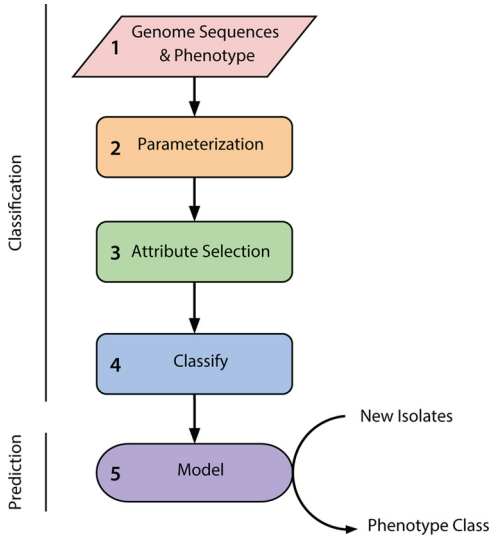

FIG 2 General scheme of the machine-learning framework used to discriminate VISA and VSSA isolates. A more detailed scheme of the machine-learning algorithm is shown in Fig. S1 in the supplemental material.

RESULTS

Evolutionary relationships of VISA and VSSA genomes. A set of 25 *S. aureus* whole-genome sequences (8 VISA and 17 VSSA) were evaluated using several different approaches in order to characterize their evolutionary relationships. Fourteen complete *S. aureus* genome sequences were taken from the NCBI RefSeq genome database (26), and 11 draft *S. aureus* genome sequences, corresponding to clinical isolates from the Emory Healthcare microbiology laboratory, were characterized as described in Materials and Methods.

The *S. aureus* isolates analyzed here are very closely related and cannot be readily separated into distinct VISA and VSSA groups using traditional methods of molecular typing. The close relationships among these strains is supported by high levels of pairwise 16S rRNA sequence similarity; on average, the 16S rRNA sequences of these strains are >99.9% identical. There are three groups of strains that share 100% identical 16S rRNA sequences, with only three nucleotide differences between the most distantly related pairs of sequences. Furthermore, groups of strains with identical 16S rRNA sequences include representatives from both the VISA and VSSA sets (see Fig. S2 in the supplemental material). The inability of 16S rRNA sequences to distinguish VISA from VSSA strains could be due to the high levels of sequence similarity resulting in poor resolution or indicate that the VISA and VSSA groups do not represent two distinct evolutionary lineages.

MLST analysis provided greater resolution for typing of the *S. aureus* strains analyzed here than 16S rRNA analysis (Fig. 3A). There are 10 MLST sequence types (STs) represented among the *S. aureus* isolates analyzed here, and the MLST loci show an average pairwise sequence identity of 93.6%. However, as with 16S rRNA, the MLST analysis does not distinguish VISA from VSSA strains. For example, there are two groups of strains that share identical STs and include both VISA and VSSA strains (Fig. 3A, ST-5 and ST-105). Overall, VISA and VSSA strains are mixed across the MLST tree.

Finally, a whole-genome phylogeny of the *S. aureus* isolates was reconstructed using pairwise average nucleotide identities (ANI) between genomes. As can be seen for the 16S rRNA and MLST analyses, VISA and VSSA isolates show high genome-wide sequence identity and are mixed on the ANI tree (Fig. 3B). Considered together, these three comparative sequence analyses indicate that the VISA and VSSA strains analyzed here do not represent distinct evolutionary lineages. Thus, any attempt to discriminate between these phenotypic groups based on genome sequence comparison would have to entail an analysis of the sequence-level determinants of their vancomycin susceptibility differences as opposed to molecular typing markers alone.

Genome-based discrimination of VISA and VSSA isolates. Despite the fact that VISA and VSSA isolates are not evolutionarily distinct, we reasoned that there may be genomic features that can collectively distinguish between the two groups. These could be gene-specific features, genome-wide characteristics, or some combination thereof. In particular, genes relevant to the vancomycin susceptibility phenotype (Fig. 1) may be expected to be able to distinguish VISA and VSSA isolates. Accordingly, we attempted to build a machine-learning protocol to distinguish VISA from VSSA isolates using a combination of 29 gene-specific metrics, 7 genome-scale parameters, and 9 molecular typing parameters (Table 3). The gene metrics were chosen based on their potential involvement in the vancomycin susceptibility phenotype, and the genome-wide and molecular typing attributes were chosen in an effort to provide increased resolution for isolate discrimination. The machine-learning approach we employed (see Materials and Methods; also see the supplemental material) provides a number of advantages, including the ability to identify which metrics are

TABLE 3 List of metrics used in the preliminary analysis

Assembly based		Gene based[a]					
		Molecular typing marker		Vancomycin resistance-implicated phenotype			
De novo	Reference	16S rRNA	MLST based	Cellular processes and signaling	Information storage and processing	Metabolism	Poorly characterized
ANI, aligned bases	No. of contigs, assembly size, N_{50}, % alignment, assembly score	16S aligned bases, 16S ANI	*arcC*, aroE, glpF*, gmk*, pta, tpi, yqil*	*agrA, agrC, graS, pbp4*, pbpB, rsbU*, spoVG, tcaB, vraF, vraG, vraS**	*ccpA, graR*, rpoB, rpoD*, tcaR, tgt*, walK, walR*, blaZ*	*folC*, isdE*, prsA*	*SA1129*, SA1703, sbf, sigB*, stp1, tcaA*

[a] An asterisk marks metrics chosen by attribute selection and used in the final model generation. Assembly and gene-based metrics used in the study are categorized by conceptual/functional class. Detailed functional assignments are shown in Table S2 in the supplemental material.

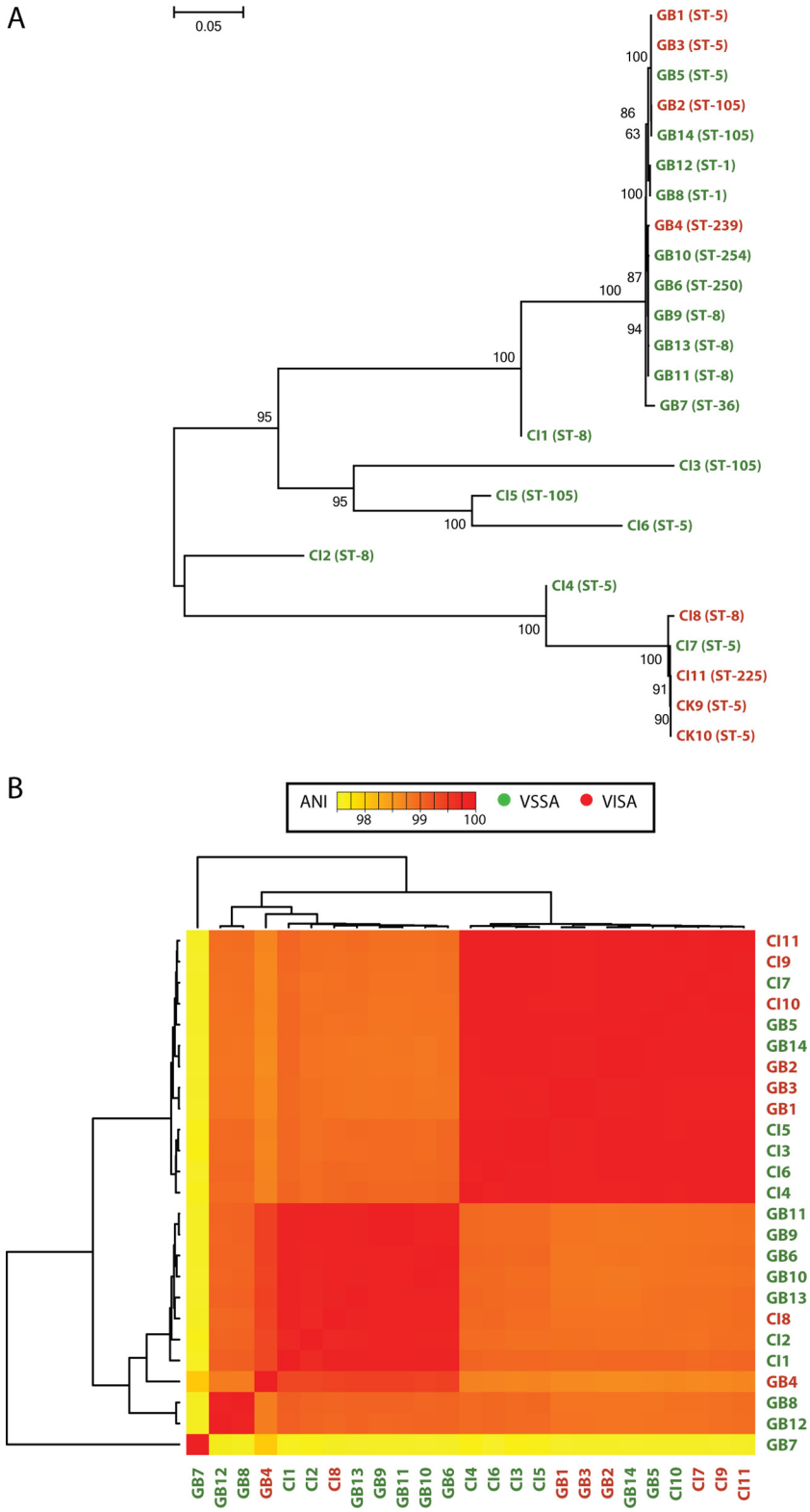

FIG 3 Evolutionary relationships among VISA (red) and VSSA (green) isolates. (A) Phylogeny based on a concatenated alignment of 7 MLST gene coding sequences. The MLST sequence type (ST) is also shown for each isolate, as are bootstrap values in support of internal nodes. (B) Heat map showing average pairwise nucleotide identities along with a dendrogram showing inferred evolutionary relationships for VISA and VSSA isolate genomes.

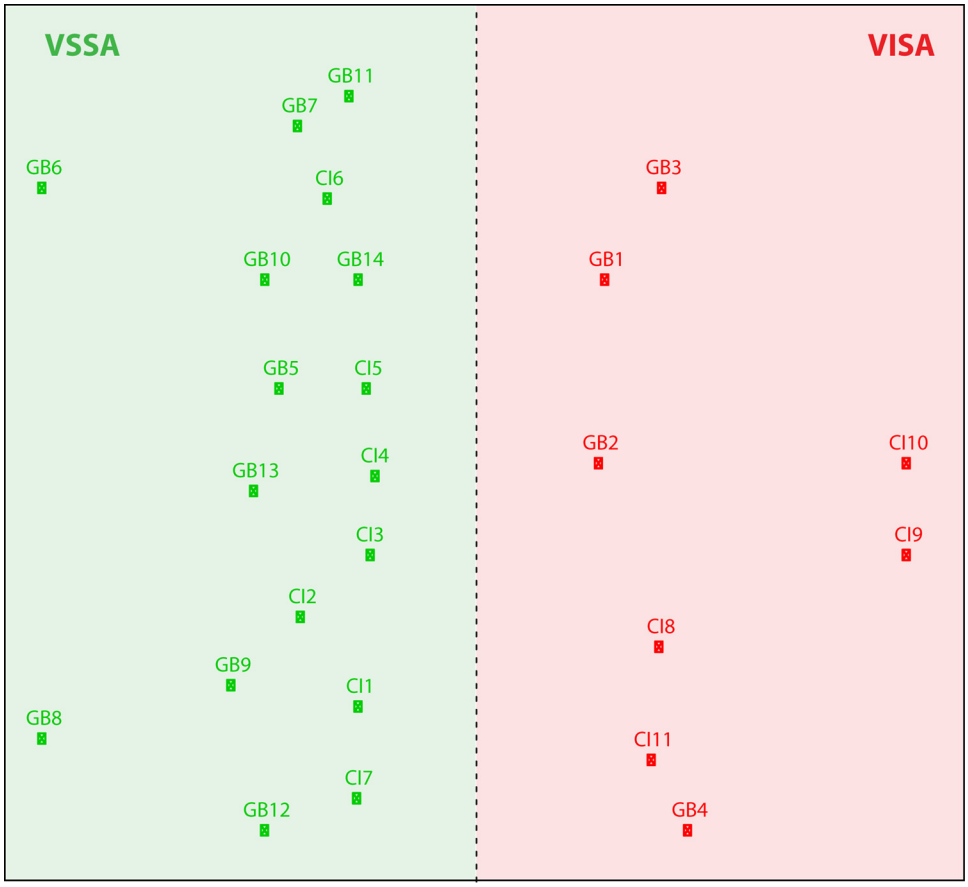

FIG 4 Machine-learning genome-based discrimination of VISA and VSSA isolates. VISA (red) and VSSA (green) isolates are shown on a two-dimensional projection of the 14-dimensional logistic regression equation used for model building. Isolate coordinates represent log-transformed output values of the regression model.

most useful for the discrimination between VISA and VSSA. In light of this utility, we initially used numerous gene-based and genome-wide parameters to build the classifier and then allowed the algorithm to inform us as to which parameters are relevant. This provides the additional insight into which genes are most relevant to the reduced vancomycin susceptibility phenotype.

Attribute selection resulted in a reduction from 45 initial parameters to 14 final parameters, including 3 molecular typing markers and 11 genes implicated in the reduced vancomycin susceptibility phenotype. Using the final set of 14 parameters, *S. aureus* classifier models were generated using 6 different machine-learning algorithms, and the different algorithms were evaluated based on accuracy, precision, and recall with 10-fold cross validation (see Table S1 and supplemental methods in the supplemental material). Based on this comparison, logistic regression was chosen for the final classifier regeneration owing to its superior performance and its simplicity. The final logistic regression machine-learning model unambiguously discriminated between all VISA and VSSA isolates analyzed here (Fig. 4). Subsequently, the accuracy of the final machine-learning model was evaluated individually for each isolate using leave-one-out cross validation, yielding an accuracy of 84% (i.e., 21 out of 25 accurate classifications).

DISCUSSION

In this study, we provide a proof of principle that a machine-learning approach can distinguish between VISA and VSSA isolates even when they do not form evolutionarily distinct groups. This approach provides several advantages for clinical applications, including speed, accuracy, and ease of use. Furthermore, the process of attribute selection provides information on the identities of genes that are most likely to be involved in the reduced vancomycin susceptibility phenotype.

Attribute selection and genes implicated in the VISA phenotype. Attribute selection uncovered a number of genes involved in synthesis and/or modification of the cell wall. For example, members of two-component regulatory systems selected here, *graR*, *vraS*, and *walR*, have been implicated in the VISA phenotype via their effects on cell wall composition and capsule expression (4, 32–34). It has also been shown that mutations to other regulatory genes, including *rpoD* selected here, can lead to the VISA phenotype in the Mu3 lineage (35). The selection of the cell wall synthesis enzyme *pbp4*, here implicated in the VISA phenotype, has been confirmed in other studies (35, 36) and is related to its role in peptidoglycan cross-linking.

The attribute selection algorithm also identified a number of genes that may exert their effects on reduced vancomycin susceptibility via alternative mechanisms. For example, the sigma factor *sigB* is a master regulator that controls the expression of numerous genes, and *rsbU* is required for its activation (37). *rsbU* has previously been implicated in glycopeptide resistance (38), and modification of these synergistically acting genes presumably could af-

fect the expression of any number of downstream targets, leading to changes in vancomycin susceptibility. Translation has also been implicated in the VISA phenotype here via the selection of the *tgt* gene (10), which encodes queuine tRNA-ribosyltransferase, as has metabolism via the selection of the *folC* (20) and *isdE* genes (9).

It should be noted that reduced vancomycin susceptibility may be encoded by any number or any combination of these implicated genes. In other words, any given VISA strain may harbor mutations in one or more different members of this set of implicated genes. This is consistent with the fact that the reduced vancomycin phenotype is found dispersed among numerous strains from evolutionarily distinct groups.

Need for a molecular assay for VISA. Antibiotic therapy with vancomycin is the mainstay of treatment for patients with *S. aureus* bloodstream infections. Vancomycin is typically given before knowledge as to the possibility of reduced susceptibility to the drug, which has become increasingly prevalent over the last 10 years (5, 6). The ability to detect *S. aureus* isolates with reduced vancomycin susceptibility is critical given the poor clinical outcomes for patients infected with VISA (39). Currently, patients infected with VISA may be treated with vancomycin for up to 72 h before an elevated MIC is determined by routine antibiotic susceptibility testing. The availability of a rapid and accurate VISA diagnostic tool (i.e., a tool to detect reduced vancomycin susceptibility or clinical failure with vancomycin use) would help to avoid such situations and ensure an improved standard of care for patients infected with *S. aureus*.

Building a molecular assay for VISA. Molecular assays are increasingly used for antibiotic susceptibility and have been shown to work well for MRSA in particular (40). A PCR-based assay for MRSA is feasible, since the phenotype is encoded by a single gene that is either present in resistant isolates or absent from susceptible isolates, except in rare cases. However, given the multigenic determinants of the VISA phenotype, development of a molecular-based assay for VISA would seem to benefit from a broadly targeted genome-wide approach. Fortunately, genome-scale or multigene sequence-based approaches to molecular typing and diagnosis are becoming increasingly affordable and tractable owing to advances in DNA sequencing and bioinformatics technologies. Here, we have taken such a genome-scale approach to the discrimination between VISA and VSSA isolates. This approach will serve as a guide for the future development of more targeted multigene or SNP-based assays that will be able to accurately detect VISA isolates directly from clinical samples.

Genomics and machine learning for molecular assay development. The machine-learning analytical framework provides a generalizable, flexible, and powerful set of solutions for classification based on large data sets characterized by numerous features. As such, it is highly amenable to classification of organisms based on genome sequences and features. machine-learning approaches provide the added value of selecting which of the attributes initially chosen for classification actually contribute to the model's discriminatory power. This allows for both a rapid and streamlined classifier based on the minimal possible set of parameters and also yields biological insight, in the case of the gene-based parameters employed here, into the genetic determinants of the phenotypes that are being distinguished.

Machine-learning approaches also have clinical relevance in the sense that they can provide for rapid and automated turnkey solutions to health care providers in need of clinical diagnoses. The vast majority of the computational effort in machine learning occurs far in advance of the actual clinical analysis in order to select the relevant parameters and build the model that will be used for classification. Once such a model is in place, the machine-learning tool can be easily applied to the clinical setting and its execution will be computationally efficient and rapid. The turnkey performance of the genome-based machine-learning approach employed here is exemplified by the fact that the clinical isolates were analyzed as draft genome sequences rather than complete assembled genome sequences, starting with sequence reads and proceeding through the machine-learning classification process in a fully automated fashion. In light of these features, machine-learning-based approaches hold great promise for the development of genome-scale molecular assays.

Current limitations and future developments of the approach. There remain limitations to the approach we have employed here along with prospects for future development and improvement. The machine-learning models were built using a number of candidate genes that were selected based on current knowledge as to the possible genetic determinants of vancomycin susceptibility. Given that this phenotype is not entirely understood, it is possible that there are additional relevant genes that were not included in our assay. With the machine-learning framework developed here in place, we could readily evaluate additional genes whose identities may become available as knowledge on the genetic determinants of vancomycin resistance grows. Another caveat is that there is a formal possibility that as additional VISA and VSSA genome sequences become available the discriminatory power of the algorithm could change, as could the set of relevant parameters employed to build the model. Once additional genome sequences become available, they could be iteratively added to the machine-learning process until it converges on a stable set of discriminating parameters.

An alternative to the candidate gene approach employed here is a more agnostic genome-wide approach that identifies target genes for VISA versus VSSA discrimination based on a genome-wide association (GWAS)-type study that would point to a set of SNPs that collectively distinguish between phenotypic groups. Owing to the statistical properties (i.e., the power) of currently employed GWAS approaches, many more complete genome sequences would need to be analyzed in order for such an approach to be feasible.

ACKNOWLEDGMENTS

We acknowledge Monica Farley for helpful comments and discussion along with editorial input and the Emory Integrated Genomics Core for providing genome sequencing services for the clinical isolates.

We have no conflicts of interest to report.

This work was supported by NIH/NCRR KL2 TR000455 and UL1TR000454 to C.S.K. L.R. and K.J. were supported by the Georgia Tech School of Biology and Bioinformatics Graduate Program.

REFERENCES

1. **Lowy FD.** 1998. Staphylococcus aureus infections. N. Engl. J. Med. **339:** 520–532. http://dx.doi.org/10.1056/NEJM199808203390806.
2. **Laupland KB, Ross T, Gregson DB.** 2008. Staphylococcus aureus bloodstream infections: risk factors, outcomes, and the influence of methicillin resistance in Calgary, Canada, 2000–2006. J. Infect. Dis. **198:**336–343. http://dx.doi.org/10.1086/589717.
3. **Liu C, Bayer A, Cosgrove SE, Daum RS, Fridkin SK, Gorwitz RJ, Kaplan SL, Karchmer AW, Levine DP, Murray BE, Rybak JM, Talan DA, Chambers HF.** 2011. Clinical practice guidelines by the Infectious Dis-

eases Society of America for the treatment of methicillin-resistant Staphylococcus aureus infections in adults and children: executive summary. Clin. Infect. Dis. **52:**285–292. http://dx.doi.org/10.1093/cid/cir034.

4. **Chen CJ, Lin MH, Shu JC, Lu JJ.** 2013. Reduced susceptibility to vancomycin in isogenic Staphylococcus aureus strains of sequence type 59: tracking evolution and identifying mutations by whole-genome sequencing. J. Antimicrob. Chemother. [Epub ahead of print.] http://dx.doi.org/10.1093/jac/dkt395.
5. **Tenover FC, Biddle JW, Lancaster MV.** 2001. Increasing resistance to vancomycin and other glycopeptides in Staphylococcus aureus. Emerg. Infect. Dis. **7:**327–332. http://dx.doi.org/10.3201/eid0702.010237.
6. **Tenover FC, Moellering RC, Jr.** 2007. The rationale for revising the Clinical and Laboratory Standards Institute vancomycin minimal inhibitory concentration interpretive criteria for Staphylococcus aureus. Clin. Infect. Dis. **44:**1208–1215. http://dx.doi.org/10.1086/513203.
7. **de Niederhausern S, Bondi M, Messi P, Iseppi R, Sabia C, Manicardi G, Anacarso I.** 2011. Vancomycin-resistance transferability from VanA enterococci to Staphylococcus aureus. Curr. Microbiol. **62:**1363–1367. http://dx.doi.org/10.1007/s00284-011-9868-6.
8. **Kos VN, Desjardins CA, Griggs A, Cerqueira G, Van Tonder A, Holden MT, Godfrey P, Palmer KL, Bodi K, Mongodin EF, Wortman J, Feldgarden M, Lawley T, Gill SR, Haas BJ, Birren B, Gilmore MS.** 2012. Comparative genomics of vancomycin-resistant Staphylococcus aureus strains and their positions within the clade most commonly associated with Methicillin-resistant S. aureus hospital-acquired infection in the United States. mBio **3:**e00112–12. http://dx.doi.org/10.1128/mBio.00112-12.
9. **Howden BP, Davies JK, Johnson PD, Stinear TP, Grayson ML.** 2010. Reduced vancomycin susceptibility in Staphylococcus aureus, including vancomycin-intermediate and heterogeneous vancomycin-intermediate strains: resistance mechanisms, laboratory detection, and clinical implications. Clin. Microbiol. Rev. **23:**99–139. http://dx.doi.org/10.1128/CMR.00042-09.
10. **Hafer C, Lin Y, Kornblum J, Lowy FD, Uhlemann AC.** 2012. Contribution of selected gene mutations to resistance in clinical isolates of vancomycin-intermediate Staphylococcus aureus. Antimicrob. Agents Chemother. **56:**5845–5851. http://dx.doi.org/10.1128/AAC.01139-12.
11. **Tatusov RL, Galperin MY, Natale DA, Koonin EV.** 2000. The COG database: a tool for genome-scale analysis of protein functions and evolution. Nucleic Acids Res. **28:**33–36. http://dx.doi.org/10.1093/nar/28.1.33.
12. **Cui L, Ma X, Sato K, Okuma K, Tenover FC, Mamizuka EM, Gemmell CG, Kim MN, Ploy MC, El-Solh N, Ferraz V, Hiramatsu K.** 2003. Cell wall thickening is a common feature of vancomycin resistance in Staphylococcus aureus. J. Clin. Microbiol. **41:**5–14. http://dx.doi.org/10.1128/JCM.41.1.5-14.2003.
13. **Pootoolal J, Neu J, Wright GD.** 2002. Glycopeptide antibiotic resistance. Annu. Rev. Pharmacol. Toxicol. **42:**381–408. http://dx.doi.org/10.1146/annurev.pharmtox.42.091601.142813.
14. **Cameron DR, Ward DV, Kostoulias X, Howden BP, Moellering RC, Jr, Eliopoulos GM, Peleg AY.** 2012. Serine/threonine phosphatase Stp1 contributes to reduced susceptibility to vancomycin and virulence in Staphylococcus aureus. J. Infect. Dis. **205:**1677–1687. http://dx.doi.org/10.1093/infdis/jis252.
15. **Cui L, Isii T, Fukuda M, Ochiai T, Neoh HM, Camargo IL, Watanabe Y, Shoji M, Hishinuma T, Hiramatsu K.** 2010. An RpoB mutation confers dual heteroresistance to daptomycin and vancomycin in Staphylococcus aureus. Antimicrob. Agents Chemother. **54:**5222–5233. http://dx.doi.org/10.1128/AAC.00437-10.
16. **Ho CM, Hsueh PR, Liu CY, Lee SY, Chiueh TS, Shyr JM, Tsao SM, Chuang YC, Yan JJ, Wang LS, Wang JH, Ho MW, Tien N, Lu JJ.** 2010. Prevalence and accessory gene regulator (agr) analysis of vancomycin-intermediate Staphylococcus aureus among methicillin-resistant isolates in Taiwan–SMART program, 2003. Eur. J. Clin. Microbiol. Infect. Dis. **29:**383–389. http://dx.doi.org/10.1007/s10096-009-0868-4.
17. **Howden BP, McEvoy CR, Allen DL, Chua K, Gao W, Harrison PF, Bell J, Coombs G, Bennett-Wood V, Porter JL, Robins-Browne R, Davies JK, Seemann T, Stinear TP.** 2011. Evolution of multidrug resistance during Staphylococcus aureus infection involves mutation of the essential two component regulator WalKR. PLoS Pathog. **7:**e1002359. http://dx.doi.org/10.1371/journal.ppat.1002359.
18. **Kato Y, Suzuki T, Ida T, Maebashi K.** 2010. Genetic changes associated with glycopeptide resistance in Staphylococcus aureus: predominance of amino acid substitutions in YvqF/VraSR. J. Antimicrob. Chemother. **65:**37–45. http://dx.doi.org/10.1093/jac/dkp394.
19. **McEvoy CR, Tsuji B, Gao W, Seemann T, Porter JL, Doig K, Ngo D, Howden BP, Stinear TP.** 2013. Decreased vancomycin susceptibility in Staphylococcus aureus caused by IS256 tempering of WalKR expression. Antimicrob. Agents Chemother. **57:**3240-3249. http://dx.doi.org/10.1128/AAC.00279-13.
20. **Passalacqua KD, Satola SW, Crispell EK, Read TD.** 2012. A mutation in the PP2C phosphatase gene in a Staphylococcus aureus USA300 clinical isolate with reduced susceptibility to vancomycin and daptomycin. Antimicrob. Agents Chemother. **56:**5212–5223. http://dx.doi.org/10.1128/AAC.05770-11.
21. **Yoo JI, Kim JW, Kang GS, Kim HS, Yoo JS, Lee YS.** 2013. Prevalence of amino acid changes in the yvqF, vraSR, graSR, and tcaRAB genes from vancomycin intermediate resistant Staphylococcus aureus. J. Microbiol. **51:**160–165. http://dx.doi.org/10.1007/s12275-013-3088-7.
22. **Takata T, Miyazaki M, Futo M, Hara S, Shiotsuka S, Kamimura H, Yoshimura H, Matsunaga A, Nishida T, Ishikura H, Ishikawa T, Tamura K, Tsuji BT.** 2013. Presence of both heterogeneous vancomycin-intermediate resistance and beta-lactam antibiotic-induced vancomycin resistance phenotypes is associated with the outcome in methicillin-resistant Staphylococcus aureus bloodstream infection. Scand. J. Infect. Dis. **45:**203–212. http://dx.doi.org/10.3109/00365548.2012.723221.
23. **Hirao Y, Ikeda-Dantsuji Y, Matsui H, Yoshida M, Hori S, Sunakawa K, Nakae T, Hanaki H.** 2012. Low level ss-lactamase production in methicillin-resistant Staphylococcus aureus strains with ss-lactam antibiotics-induced vancomycin resistance. BMC Microbiol. **12:**69. http://dx.doi.org/10.1186/1471-2180-12-69.
24. **Frank E, Hall M, Trigg L, Holmes G, Witten IH.** 2004. Data mining in bioinformatics using Weka. Bioinformatics **20:**2479–2481. http://dx.doi.org/10.1093/bioinformatics/bth261.
25. **Rishishwar L, Petit RA, III, Kraft CS, Jordan IK.** 2013. Abstr. 113th Gen. Meet. Am. Soc. Microbiol., poster 2594.
26. **Pruitt KD, Tatusova T, Brown GR, Maglott DR.** 2012. NCBI Reference Sequences (RefSeq): current status, new features and genome annotation policy. Nucleic Acids Res. **40:**D130–D135. http://dx.doi.org/10.1093/nar/gkr1079.
27. **Thompson JD, Higgins DG, Gibson TJ.** 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22:**4673–4680. http://dx.doi.org/10.1093/nar/22.22.4673.
28. **Saitou N, Nei M.** 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. **4:**406–425.
29. **Kumar S, Nei M, Dudley J, Tamura K.** 2008. MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences. Brief. Bioinform. **9:**299–306. http://dx.doi.org/10.1093/bib/bbn017.
30. **Enright MC, Day NP, Davies CE, Peacock SJ, Spratt BG.** 2000. Multilocus sequence typing for characterization of methicillin-resistant and methicillin-susceptible clones of Staphylococcus aureus. J. Clin. Microbiol. **38:**1008–1015.
31. **Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M, Antonescu C, Salzberg SL.** 2004. Versatile and open software for comparing large genomes. Genome Biol. **5:**R12. http://dx.doi.org/10.1186/gb-2004-5-2-r12.
32. **Cui L, Neoh HM, Shoji M, Hiramatsu K.** 2009. Contribution of vraSR and graSR point mutations to vancomycin resistance in vancomycin-intermediate Staphylococcus aureus. Antimicrob. Agents Chemother. **53:**1231–1234. http://dx.doi.org/10.1128/AAC.01173-08.
33. **Howden BP, Stinear TP, Allen DL, Johnson PD, Ward PB, Davies JK.** 2008. Genomic analysis reveals a point mutation in the two-component sensor gene graS that leads to intermediate vancomycin resistance in clinical Staphylococcus aureus. Antimicrob. Agents Chemother. **52:**3755–3762. http://dx.doi.org/10.1128/AAC.01613-07.
34. **Neoh HM, Cui L, Yuzawa H, Takeuchi F, Matsuo M, Hiramatsu K.** 2008. Mutated response regulator graR is responsible for phenotypic conversion of Staphylococcus aureus from heterogeneous vancomycin-intermediate resistance to vancomycin-intermediate resistance. Antimicrob. Agents Chemother. **52:**45–53. http://dx.doi.org/10.1128/AAC.00534-07.
35. **Matsuo M, Cui L, Kim J, Hiramatsu K.** 2013. Comprehensive identification of mutations responsible for heterogeneous vancomycin-intermediate Staphylococcus aureus (hVISA)-to-VISA conversion in laboratory-generated VISA strains derived from hVISA clinical strain Mu3. Antimicrob. Agents Chemother. **57:**5843–5853. http://dx.doi.org/10.1128/AAC.00425-13.

36. **Atilano ML, Pereira PM, Yates J, Reed P, Veiga H, Pinho MG, Filipe SR.** 2010. Teichoic acids are temporal and spatial regulators of peptidoglycan cross-linking in Staphylococcus aureus. Proc. Natl. Acad. Sci. U. S. A. **107:**18991–18996. http://dx.doi.org/10.1073/pnas.1004304107.
37. **Pane-Farre J, Jonas B, Hardwick SW, Gronau K, Lewis RJ, Hecker M, Engelmann S.** 2009. Role of RsbU in controlling SigB activity in Staphylococcus aureus following alkaline stress. J. Bacteriol. **191:**2561–2573. http://dx.doi.org/10.1128/JB.01514-08.
38. **Bischoff M, Berger-Bachi B.** 2001. Teicoplanin stress-selected mutations increasing sigma(B) activity in Staphylococcus aureus. Antimicrob. Agents Chemother. **45:**1714–1720. http://dx.doi.org/10.1128/AAC.45.6.1714-1720.2001.
39. **van Hal SJ, Jensen SO, Vaska VL, Espedido BA, Paterson DL, Gosbell IB.** 2012. Predictors of mortality in Staphylococcus aureus bacteremia. Clin. Microbiol. Rev. **25:**362–386. http://dx.doi.org/10.1128/CMR.05022-11.
40. **Parta M, Goebel M, Thomas J, Matloobi M, Stager C, Musher DM.** 2010. Impact of an assay that enables rapid determination of Staphylococcus species and their drug susceptibility on the treatment of patients with positive blood culture results. Infect. Control Hosp. Epidemiol. **31:**1043–1048. http://dx.doi.org/10.1086/656248.