Optimization of cluster analysis based on drug resistance profiles of MRSA isolates

HIROYA TANI¹,₂, TAKAHIKO KISHI², MINEHIRO GOTOH², YUKA YAMAGISHI³ and HIROSHIGE MIKAMO¹,³

¹Department of Clinical Infectious Diseases, Aichi Medical University
Graduate School of Medicine
²Department of Clinical Laboratory, Aichi Medical University Hospital
³Department of Clinical Infectious Diseases, Aichi Medical University Hospital

(Received for publication August 25, 2015)

We examined 402 methicillin-resistant Staphylococcus aureus (MRSA) strains isolated from clinical specimens in our hospital between November 19, 2010 and December 27, 2011 to evaluate the similarity between cluster analysis of drug susceptibility tests and pulsed-field gel electrophoresis (PFGE). The results showed that the 402 strains tested were classified into 27 PFGE patterns (151 subtypes of patterns). Cluster analyses of drug susceptibility tests with the cut-off distance yielding a similar classification capability showed favorable results—when the MIC method was used, and minimum inhibitory concentration (MIC) values were used directly in the method, the level of agreement with PFGE was 74.2% when 15 drugs were tested. The Unweighted Pair Group Method with Arithmetic mean (UPGMA) method was effective when the cut-off distance was 16. Using the SIR method in which susceptible (S), intermediate (I), and resistant (R) were coded as 0, 2, and 3, respectively, according to the Clinical and Laboratory Standards Institute (CLSI) criteria, the level of agreement with PFGE was 75.9% when the number of drugs tested was 17, the method used for clustering was the UPGMA, and the cut-off distance was 3.6. In addition, to assess the reproducibility of the results, 10 strains were randomly sampled from the overall test and subjected to cluster analysis. This was repeated 100 times under the same conditions. The results indicated good reproducibility of the results, with the level of agreement with PFGE showing a mean of 82.0%, standard deviation of 12.1%, and mode of 90.0% for the MIC method and a mean of 80.0%, standard deviation of 13.4%, and mode of 90.0% for the SIR method. In summary, cluster analysis for drug susceptibility tests is useful for the epidemiological analysis of MRSA.
Introduction

Methicillin-resistant *Staphylococcus aureus* (MRSA), first reported in the United Kingdom in 1961\(^1\), includes approximately 50% of the nosocomial isolates of *S. aureus*\(^2\) and remains the most important microorganism in healthcare-associated infections. Epidemiological analyses of MRSA have become increasingly important because MRSA presents problems such as infections or outbreaks in long-term hospitalized patients, dialysis patients, patients with indwelling medical devices, or compromised hosts in hospital settings\(^3\sim^5\). Additionally, community-acquired MRSA has also attracted attention as the causative agent of community-acquired infections in recent years\(^6\), \(^7\).

Pulsed-field gel electrophoresis (PFGE) is the gold standard used for the epidemiological analyses of MRSA\(^8\), \(^9\), but this method has many disadvantages including procedural complexity, the long period of 5 days or more required to obtain results, and the need for special equipment\(^10\sim^12\). Thus, new techniques such as multi-locus sequence typing\(^13\), phage open reading frame typing\(^14\), and repetitive sequence-based PCR\(^15\) have been developed\(^16\sim^20\). However, although these methods are more rapid and convenient than PFGE, they are difficult to perform in routine laboratories because of the need to familiarize the operating personnel with the dedicated equipment or procedures. In contrast, although drug susceptibility testing is routinely performed in hospital laboratories and produces a wide range of results, there are limited reports of the use of drug susceptibility testing for epidemiological analysis.

In this study, we investigated the similarity between PFGE and cluster analysis of drug susceptibility tests (“drug cluster analysis”) to evaluate the usefulness of drug cluster analysis in epidemiological analyses of MRSA.

**Materials and Methods**

1. **Bacterial strains used in this study**

A total of 402 MRSA was strains isolated from clinical specimens in our hospital between November 19, 2010 and December 27, 2011. When two or more strains were detected in a single patient, only the first detected strain was included if they were the same strains. All detected strains were included if they were different.

2. **PFGE**

Heart infusion broth (Nippon Becton Dickinson Company, Ltd., Tokyo, Japan) was inoculated and incubated with test strains at 37°C overnight and the obtained bacterial suspension was used. A DNA plug was prepared using the CHEF Bacterial Genomic DNA Plug Kit (Bio-Rad, Hercules, CA, USA) and the restriction enzyme *Sma*I (Takara Bio, Shiga, Japan). Electrophoresis
was performed in a CHEF-DR III system (Bio-Rad) according to the program 5 for the CHEF-Mapper: voltage of 6.0 V/cm, angle of 120°, total run time of 20.0 h, initial switch time of 5.3 s, and final switch time of 34.9 s. Lambda Ladder (Bio-Rad) was used as a control.

To analyze the PFGE patterns, a dendrogram was generated using FP Quest Plus software (Bio-Rad) and patterns with a similarity of 80% or higher were considered to represent the same strain\(^{21-23}\). Similarity among strains was determined by the Dice method and clustering was performed using the unweighted pair group method with arithmetic mean (UPGMA) method.

### 3. Drug susceptibility testing

Drug susceptibility testing was performed using RAISUS (Nissui Pharmaceutical, Tokyo, Japan), a routinely used and fully automated system for identification and susceptibility testing, using the broth microdilution method. Isolates were tested against a total of 17 drugs, including oxacillin (MPIPC), ampicillin (ABPC), ABPC/sulbactam (SBT), cefazolin (CEZ), cefoxitin (CFX), imipenem (IPM), gentamicin (GM), arbekacin (ABK), clarithromycin (CAM), clindamycin (CLDM), levofloxacin (LVFX), minocycline (MINO), linezolid (LZD), vancomycin (VCM), teicoplanin (TEIC), fosfomycin (FOM), and sulfamethoxazole/trimethoprim (ST), and interpreted according to the Clinical and Laboratory Standards Institute (CLSI) criteria\(^{24,25}\). *Staphylococcus aureus* isolates with a minimum inhibitory concentration (MIC) of 4 μg/mL or more for MPIPC were defined as MRSA.

### 4. Cluster analysis

Drug cluster analysis was performed using the MIC method, in which MIC values were directly used as data, and the SIR method, in which based on the method of SATO, et al.\(^{26}\), isolates that were interpreted as susceptible (S), intermediate (I), and resistant (R) according to CLSI criteria\(^{24,25}\) were coded as 0, 2, and 3, respectively. A dendrogram was generated by selecting squared Euclidean distances to determine similarity, and the UPGMA method\(^{27,28}\), which is commonly used in PFGE, and Ward’s method\(^{29}\), which typically yields clear clusters for clustering. Dendrograms were generated using SAS version 9.3 (SAS Institute Inc., Cary, NC, USA). In the MIC method, the results for ABPC/SBT and ST were excluded from the data because they were difficult to directly quantify, and results for the other drugs with an inequality sign but no equals sign (\(>\) or \(<\)) were used; a value higher or lower by one dilution than the figure and results with an inequality sign with an equal sign (\(\geq\) or \(\leq\)) were directly used as the figure. In the SIR method, the MIC of VCM was converted to 1 if the value was not more than 1 μg/mL and to 2 if the value was 2 μg/mL based on VCM MIC creeping\(^{30-33}\). Using these methods, cluster analyses were performed to investigate cut-off distances that yielded the number of clusters similar to those observed using PFGE, the number of clusters, and the level of agreement with PFGE in the following cases: i) when all drugs (15 drugs for the MIC method and 17 drugs for the SIR
method) were included, ii) drugs that were highly correlated to other antimicrobials (a Pearson correlation coefficient of $\geq 0.7$) were excluded (14 drugs, excluding IPM, for the MIC method and 15 drugs, excluding ABPC/SBT and CLDM, for the SIR method), iii) 5 representative drugs from antimicrobial classes (penicillin: ABPC, cephem: CEZ, carbapenem: IPM, aminoglycoside: GM, and quinolone: LVFX) were selected, and iv) 9 drugs, consisting the 5 representative drugs plus 4 anti-MRSA drugs (ABK, LZD, VCM, and TEIC), were selected.

The level of agreement was calculated based on the assumption that 2 clusters form a pair if they contained a large number of shared strains. The number of strains that were shared with pairs of drug clusters based on PFGE clusters was counted.

5. Reproducibility of cluster analyses results
Ten strains were randomly sampled from the overall population of 402 test strains using the SAS9.3 and drug cluster analysis was performed under the conditions with the highest level of agreement with PFGE method using the MIC and SIR methods. This was repeated 100 times to assess reproducibility by calculating the mean, standard deviation, and mode of the agreement level with PFGE.

Results

1. PFGE method
According to PFGE results, the isolates tested were classified into the following 27 patterns (151 subtypes) from type A to type AA. In addition, 309 strains belonged to type C, accounting for 76.9% of the total strains (Figure 1, Table 1).

2. Drug susceptibility testing
As shown in Table 2, drug susceptibility testing revealed that the isolates generally showed resistance to many antimicrobials. Among anti-MRSA drugs, 5 strains were considered as I only for ABK, but the susceptibility to VCM, TEIC, or LZD was S in all strains.

3. Cluster analysis
The cut-off distance yielding a number of clusters close to that obtained using PFGE, the number of clusters, and the level of agreement with PFGE were shown in Table 3. The results revealed a higher level of agreement with PFGE using the UPGMA method compared to using the Ward’s method. In addition, the level of agreement with PFGE was generally higher for the UPGMA method when increasing number of drugs were tested. Moreover, the level of agreement between C type of PFGE method and a cluster analysis under the analysis conditions for the highest level of agreement with the PFGE method (Table 4) was 86.7% in MIC method and 88.7% in SIR method.
Fig. 1. Dendrogram of PFGE method
Table 1. Result of PFGE method

<table>
<thead>
<tr>
<th>PFGE type</th>
<th>The number of subtypes</th>
<th>The number of strains</th>
<th>Detection rate (%)</th>
<th>PFGE type</th>
<th>The number of subtypes</th>
<th>The number of strains</th>
<th>Detection rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1</td>
<td>1</td>
<td>0.2</td>
<td>O</td>
<td>3</td>
<td>3</td>
<td>0.7</td>
</tr>
<tr>
<td>B</td>
<td>1</td>
<td>1</td>
<td>0.2</td>
<td>P</td>
<td>6</td>
<td>11</td>
<td>2.7</td>
</tr>
<tr>
<td>C</td>
<td>79</td>
<td>309</td>
<td>76.9</td>
<td>Q</td>
<td>5</td>
<td>6</td>
<td>1.5</td>
</tr>
<tr>
<td>D</td>
<td>5</td>
<td>5</td>
<td>1.2</td>
<td>R</td>
<td>2</td>
<td>2</td>
<td>0.5</td>
</tr>
<tr>
<td>E</td>
<td>1</td>
<td>1</td>
<td>0.2</td>
<td>S</td>
<td>5</td>
<td>8</td>
<td>2.0</td>
</tr>
<tr>
<td>F</td>
<td>5</td>
<td>5</td>
<td>1.2</td>
<td>T</td>
<td>3</td>
<td>6</td>
<td>1.5</td>
</tr>
<tr>
<td>G</td>
<td>2</td>
<td>2</td>
<td>0.5</td>
<td>U</td>
<td>5</td>
<td>6</td>
<td>1.5</td>
</tr>
<tr>
<td>H</td>
<td>1</td>
<td>1</td>
<td>0.2</td>
<td>V</td>
<td>3</td>
<td>3</td>
<td>0.7</td>
</tr>
<tr>
<td>I</td>
<td>5</td>
<td>9</td>
<td>2.2</td>
<td>W</td>
<td>1</td>
<td>1</td>
<td>0.2</td>
</tr>
<tr>
<td>J</td>
<td>1</td>
<td>1</td>
<td>0.2</td>
<td>X</td>
<td>2</td>
<td>2</td>
<td>0.5</td>
</tr>
<tr>
<td>K</td>
<td>2</td>
<td>2</td>
<td>0.5</td>
<td>Y</td>
<td>1</td>
<td>1</td>
<td>0.2</td>
</tr>
<tr>
<td>L</td>
<td>2</td>
<td>2</td>
<td>0.5</td>
<td>Z</td>
<td>5</td>
<td>6</td>
<td>1.5</td>
</tr>
<tr>
<td>M</td>
<td>1</td>
<td>1</td>
<td>0.2</td>
<td>AA</td>
<td>1</td>
<td>1</td>
<td>0.2</td>
</tr>
<tr>
<td>N</td>
<td>3</td>
<td>6</td>
<td>1.5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 2. Result of drug susceptibility testing

<table>
<thead>
<tr>
<th>Antimicrobial name</th>
<th>S (strain)</th>
<th>I (strain)</th>
<th>R (strain)</th>
<th>S rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>oxacillin (MPPC)</td>
<td>0</td>
<td>0</td>
<td>402</td>
<td>0.0%</td>
</tr>
<tr>
<td>ampicillin (ABPC)</td>
<td>0</td>
<td>0</td>
<td>402</td>
<td>0.0%</td>
</tr>
<tr>
<td>ampicillin/sulbactam (ABPC/SBT)</td>
<td>61</td>
<td>86</td>
<td>255</td>
<td>15.2%</td>
</tr>
<tr>
<td>cefazolin (CEZ)</td>
<td>35</td>
<td>9</td>
<td>358</td>
<td>8.7%</td>
</tr>
<tr>
<td>cefotixin (CFX)</td>
<td>16</td>
<td>0</td>
<td>386</td>
<td>4.0%</td>
</tr>
<tr>
<td>imipenem (IPM)</td>
<td>74</td>
<td>9</td>
<td>319</td>
<td>18.4%</td>
</tr>
<tr>
<td>gentamicin (GM)</td>
<td>109</td>
<td>10</td>
<td>283</td>
<td>27.1%</td>
</tr>
<tr>
<td>arbekacin (ABK)</td>
<td>397</td>
<td>5</td>
<td>0</td>
<td>98.8%</td>
</tr>
<tr>
<td>clarithromycin (CAM)</td>
<td>36</td>
<td>0</td>
<td>366</td>
<td>9.0%</td>
</tr>
<tr>
<td>clindamycin (CLDM)</td>
<td>64</td>
<td>2</td>
<td>336</td>
<td>15.9%</td>
</tr>
<tr>
<td>levofoxacin (LVFX)</td>
<td>42</td>
<td>11</td>
<td>349</td>
<td>10.4%</td>
</tr>
<tr>
<td>minocycline (MINO)</td>
<td>124</td>
<td>116</td>
<td>162</td>
<td>30.8%</td>
</tr>
<tr>
<td>linezolid (LZD)</td>
<td>402</td>
<td>0</td>
<td>0</td>
<td>100.0%</td>
</tr>
<tr>
<td>vancomycin (VCM)</td>
<td>402</td>
<td>0</td>
<td>0</td>
<td>100.0%</td>
</tr>
<tr>
<td>teicoplanin (TEIC)</td>
<td>402</td>
<td>0</td>
<td>0</td>
<td>100.0%</td>
</tr>
<tr>
<td>fosfomycin (FOM)</td>
<td>62</td>
<td>14</td>
<td>326</td>
<td>15.4%</td>
</tr>
<tr>
<td>sulfamethoxazole/trimethoprim (ST)</td>
<td>402</td>
<td>0</td>
<td>0</td>
<td>100.0%</td>
</tr>
</tbody>
</table>

S: susceptible  I: intermediate  R: resistant  S rate: Susceptible rate in the drug susceptibility testing result

*1 µg/mL or less: 393 strains, 2 µg/mL: 9 strains
4. Reproducibility of results of cluster analyses

The reproducibility of the analysis conditions for the highest level of agreement with the PFGE method (Table 4) was favorable using both the MIC method, with a mean of 82.0%, standard deviation of 12.1%, and mode of 90.0%, and the SIR method, with a mean of 80.0%, standard deviation of 13.4%, and mode of 90.0% (Table 5). There was no significant difference in reproducibility between these two methods ($P=0.3920$: Wilcoxon rank sum test).

<table>
<thead>
<tr>
<th>Cases</th>
<th>Cluster analysis method</th>
<th>Cut-off distance (Squared distance)</th>
<th>The number of clusters</th>
<th>The level of agreement with PFGE (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>i</td>
<td>UPGMA</td>
<td>16</td>
<td>27</td>
<td>74.2</td>
</tr>
<tr>
<td>ii</td>
<td>UPGMA</td>
<td>14.9</td>
<td>28</td>
<td>62.0</td>
</tr>
<tr>
<td>iii</td>
<td>UPGMA</td>
<td>6.5</td>
<td>29</td>
<td>57.6</td>
</tr>
<tr>
<td>iv</td>
<td>UPGMA</td>
<td>6.5</td>
<td>29</td>
<td>57.8</td>
</tr>
<tr>
<td></td>
<td>Ward's</td>
<td>361</td>
<td>27</td>
<td>39.3</td>
</tr>
<tr>
<td></td>
<td>Ward's</td>
<td>361</td>
<td>28</td>
<td>34.9</td>
</tr>
<tr>
<td></td>
<td>Ward's</td>
<td>52</td>
<td>29</td>
<td>52.9</td>
</tr>
<tr>
<td></td>
<td>Ward's</td>
<td>56</td>
<td>28</td>
<td>52.1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cases</th>
<th>Cluster analysis method</th>
<th>Cut-off distance (Squared distance)</th>
<th>The number of clusters</th>
<th>The level of agreement with PFGE (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>i</td>
<td>UPGMA</td>
<td>3.6</td>
<td>28</td>
<td>75.9</td>
</tr>
<tr>
<td>ii</td>
<td>UPGMA</td>
<td>3.2</td>
<td>26</td>
<td>64.2</td>
</tr>
<tr>
<td>iii</td>
<td>UPGMA</td>
<td>0.5</td>
<td>27</td>
<td>61.0</td>
</tr>
<tr>
<td>iv</td>
<td>UPGMA</td>
<td>0.5</td>
<td>31</td>
<td>60.3</td>
</tr>
<tr>
<td></td>
<td>Ward's</td>
<td>18.5</td>
<td>27</td>
<td>37.6</td>
</tr>
<tr>
<td></td>
<td>Ward's</td>
<td>15</td>
<td>27</td>
<td>39.9</td>
</tr>
<tr>
<td></td>
<td>Ward's</td>
<td>0.4</td>
<td>27</td>
<td>61.0</td>
</tr>
<tr>
<td></td>
<td>Ward's</td>
<td>0.8</td>
<td>29</td>
<td>59.1</td>
</tr>
</tbody>
</table>

Cases

i) When all drugs were included.
ii) When drugs that were highly correlated to other antimicrobials (a Pearson correlation coefficient of $\geq 0.7$) were excluded.
iii) When 5 representative drugs from antimicrobial classes were selected.
iv) When 9 drugs, including the 5 representative drugs plus 4 anti-MRSA drugs, were selected.

Table 3. Result of cluster analysis

<table>
<thead>
<tr>
<th>MIC method</th>
<th>SIR method</th>
</tr>
</thead>
<tbody>
<tr>
<td>i (15 drugs)</td>
<td>i (17 drugs)</td>
</tr>
<tr>
<td>UPGMA</td>
<td>UPGMA</td>
</tr>
<tr>
<td>Ward's</td>
<td>Ward's</td>
</tr>
<tr>
<td>16</td>
<td>3.6</td>
</tr>
<tr>
<td>27</td>
<td>28</td>
</tr>
<tr>
<td>74.2</td>
<td>75.9</td>
</tr>
</tbody>
</table>

Table 4. The analysis conditions for the highest level of agreement with the PFGE method

<table>
<thead>
<tr>
<th>The number of antimicrobials</th>
<th>Cluster analysis method</th>
<th>Cut-off distance (Squared distance)</th>
<th>The level of agreement with PFGE (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MIC method</td>
<td>15</td>
<td>UPGMA</td>
<td>16</td>
</tr>
<tr>
<td>SIR method</td>
<td>17</td>
<td>UPGMA</td>
<td>3.6</td>
</tr>
</tbody>
</table>
Discussion

The PFGE results were classified into 27 types, ranging from type A to type AA, and each type was further divided into 1~79 subtypes (151 subtypes in total). Individual types ranged from those comprising a single strain to those comprising numerous strains, such as type C (which included 309 strains, accounting for 76.9% of the total) and were divided into major types and minor types showing diversity. Our results were consistent with those of previous studies\(^\text{34, 35}\), and the predominant type C strains were detected nearly continuously during the study period covered by the study, suggesting that these strains may cause healthcare-related infections in our hospital.

Although drug susceptibility testing of the strains showed a trend towards resistance to multiple drugs, except anti-MRSA drugs (ABK, LZD, VCM, and TEIC) and ST, the susceptibility to antimicrobials varied among drugs.

Based on these results, drug cluster analyses were conducted using the UPGMA and Ward’s methods and by changing the types of drugs tested. The results revealed a higher level of agreement with PFGE using the UPGMA method compared to using the Ward’s method. In addition, the level of agreement with PFGE was generally higher for the UPGMA method when increasing number of drugs were tested, with the highest level of agreement of 75.9% observed using the SIR method when all drugs were included. This was likely due to the following reasons: for the former result of the higher level of agreement with PFGE in the UPGMA method than the Ward’s method, the UPGMA method was considered appropriate for drug cluster analyses since cluster analysis is a multivariate analysis that does not include clear classification criteria\(^\text{36}\) and can be performed using the method most convenient for the analyst\(^\text{29}\); and for the latter result of the tendency for a higher level of agreement with PFGE with increasing numbers of the drugs tested, the variability in the results of the drug susceptibility tests was increased with increasing numbers of drugs. In addition, this trend was observed using both the MIC method and the SIR method, suggesting that application of the MIC method without requiring data conversion in facilities using a commonly employed drug susceptibility panel or application of the SIR method in facilities using a breakpoint panel only for qualitative interpretation may enable facilities to perform drug susceptibility tests for strain typing. Moreover, the reproducibility of conditions showing the best

<table>
<thead>
<tr>
<th></th>
<th>Mean(%)</th>
<th>SD(%)</th>
<th>Mode(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MIC method</td>
<td>82.0</td>
<td>12.1</td>
<td>90.0</td>
</tr>
<tr>
<td>SIR method</td>
<td>80.0</td>
<td>13.4</td>
<td>90.0</td>
</tr>
</tbody>
</table>
agreement with PFGE was highly favorable for both the MIC method and the SIR method, indicating that the results were not specific to the population studied. However, because the cut-off distance associated with clustering varied according the type of drugs tested, a facility performing drug cluster analyses should determine an appropriate cut-off distance based on the drugs tested in that facility. Although the present study showed that if a standard drug susceptibility panel is used, the same cut-off distance can be used by another facility that uses the same panel, the procedure for determining cut-off distances cannot be omitted in facilities using an ordered panel; therefore, a simple method for determining cut-off distances remains to be identified.

Although few studies have examined these issues, there have been a variety of reports regarding the relationship between drug cluster analysis and PFGE, including reports showing that drug cluster analysis is useful for strain typing\textsuperscript{26, 37} and that this analysis has a limited relationship with PFGE and a low capability of typing strains\textsuperscript{9, 38}, indicating that the results are inconclusive. This may be because most reports to date evaluated the usefulness of drug cluster analysis based only on the results of the studies and did not verify whether the analysis parameters included were suitable for other situations. The present study demonstrated that application of the UPGMA method for clustering with increasing numbers of drugs tested yielded a high level of similarity to the results of PFGE, although drug cluster analysis requires determination of an appropriate cut-off distance based on the drugs used by facilities.

Drug susceptibility tests are essential for determining the type and dose of antimicrobials that should be used to treat infections. In recent years, drug susceptibility testing has also been applied as antibiograms in empirical therapy\textsuperscript{39, 40} and has become increasingly important. Our results suggest that drug susceptibility testing is useful for epidemiological analyses and that the results indicate the importance of further drug susceptibility tests.

**Disclosure:** The authors have declared no conflicts of interest.

**References**

24) Clinical and Laboratory Standards Institute (CLSI): Performance standards for antimicrobial susceptibility testing. 20th Informational supplement M100-S20. Wayne, PA, USA: CLSI; 2010