



# Accurate and Rapid Differentiation of *Acinetobacter baumannii* Strains by Raman Spectroscopy: a Comparative Study

Meron Ghebremedhin,<sup>a,b</sup> Rae Heitkamp,<sup>c,d</sup> Shubha Yesupriya,<sup>a</sup> Bradford Clay,<sup>e</sup> Nicole J. Crane<sup>a,b,f</sup>

Regenerative Medicine Department, Naval Medical Research Center, Silver Spring, Maryland, USA<sup>a</sup>; Henry M. Jackson Foundation for the Advancement of Military Medicine, Bethesda, Maryland, USA<sup>b</sup>; Wound Infections Department, Walter Reed Army Institute of Research, Silver Spring, Maryland, USA<sup>c</sup>; Clinical Research Management, Hinckley, Ohio, USA<sup>d</sup>; bioMérieux, Inc., Instrumentation R&D, Hazelwood, Missouri, USA<sup>e</sup>; Department of Surgery, Walter Reed and Uniformed Services University of Health Sciences, Bethesda, Maryland, USA<sup>f</sup>

**ABSTRACT** In recent years, matrix-assisted laser desorption ionization–time of flight mass spectrometry (MALDI-TOF MS) has become the standard for routine bacterial species identification due to its rapidity and low costs for consumables compared to those of traditional DNA-based methods. However, it has been observed that strains of some bacterial species, such as *Acinetobacter baumannii* strains, cannot be reliably identified using mass spectrometry (MS). Raman spectroscopy is a rapid technique, as fast as MALDI-TOF, and has been shown to accurately identify bacterial strains and species. In this study, we compared hierarchical clustering results for MS, genomic, and antimicrobial susceptibility test data to hierarchical clustering results from Raman spectroscopic data for 31 *A. baumannii* clinical isolates labeled according to their pulsed-field gel electrophoresis data for strain differentiation. In addition to performing hierarchical cluster analysis (HCA), multiple chemometric methods of analysis, including principal-component analysis (PCA) and partial least-squares discriminant analysis (PLSDA), were performed on the MS and Raman spectral data, along with a variety of spectral preprocessing techniques for best discriminative results. Finally, simple HCA algorithms were performed on all of the data sets to explore the relationships between, and natural groupings of, the strains and to compare results for the four data sets. To obtain numerical comparison values of the clustering results, the external cluster evaluation criteria of the Rand index of the HCA dendrograms were calculated. With a Rand index value of 0.88, Raman spectroscopy outperformed the other techniques, including MS (with a Rand index value of 0.58).

**KEYWORDS** AST, *Acinetobacter baumannii*, DNA sequencing, hierarchical cluster analysis, Raman spectroscopy, mass spectrometry

The emergence of multidrug-resistant (MDR) nosocomial infections caused by bacteria such as *Acinetobacter baumannii* has become a global problem for both civilian and military populations. MDR infections can be problematic due to the small number of effective drugs for treatment (1). A rapid and accurate strain typing method becomes important in order to detect epidemiological outbreaks of nosocomial infections and provide appropriate and timely intervention in infection control (2). Traditional DNA-based methods of typing, such as multilocus sequence typing (MLST), microsatellite genotyping, amplification fragment length polymorphism (AFLP), DNA microsatellite typing, and pulsed-field gel electrophoresis (PFGE), can be used to obtain

**Received** 18 August 2016 **Returned for modification** 9 September 2016 **Accepted** 28 April 2017

**Accepted manuscript posted online** 7 June 2017

**Citation** Ghebremedhin M, Heitkamp R, Yesupriya S, Clay B, Crane NJ. 2017. Accurate and rapid differentiation of *Acinetobacter baumannii* strains by Raman spectroscopy: a comparative study. *J Clin Microbiol* 55:2480–2490. <https://doi.org/10.1128/JCM.01744-16>.

**Editor** Erik Munson, Marquette University

**Copyright** © 2017 Ghebremedhin et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Meron Ghebremedhin, [meron.ghebremedhin.ctr@mail.mil](mailto:meron.ghebremedhin.ctr@mail.mil).

accurate and reliable results (3, 4). However, routine application of such methods has become impractical due to their high cost and time-consuming processes (4).

Matrix-assisted laser desorption ionization–time of flight mass spectrometry (MALDI-TOF MS) has become an ideal solution for overcoming these hurdles by providing rapid and accurate results and, as a result, has moved to the forefront of routine bacterial identification (5–7). However, some studies have shown the unsuitability of MALDI-TOF MS for strain-level discrimination in closely related bacterial species, including *A. baumannii* (4, 8–11). This inconsistency of MS in differentiating between closely related strains is hypothesized to arise primarily because the closely related species or strains express many similar, if not identical, proteins (12). This can result in multiple identical MS peaks, making discrimination between species or strains difficult.

In recent years, advances have been made in the application of Raman spectroscopy for bacterial identification. Raman spectroscopy is a rapid and noninvasive vibrational spectroscopic technique that yields molecular fingerprint information from biological samples such as bacteria. Numerous studies have successfully identified bacteria at the strain level by using Raman spectroscopy and surface-enhanced Raman spectroscopy (SERS) spectral databases of the microorganisms (13–25). In our previous work (25), Raman spectra of 42 bacterial samples that encompassed a wide range of bacterial species (10 different genera) were analyzed. Models were developed and successfully validated using a test set at the Gram and genus taxonomic levels. In this work, we developed hierarchical cluster analysis (HCA) models to differentiate between *A. baumannii* strains, with the goal of using Raman spectroscopy for strain-level identification in the future.

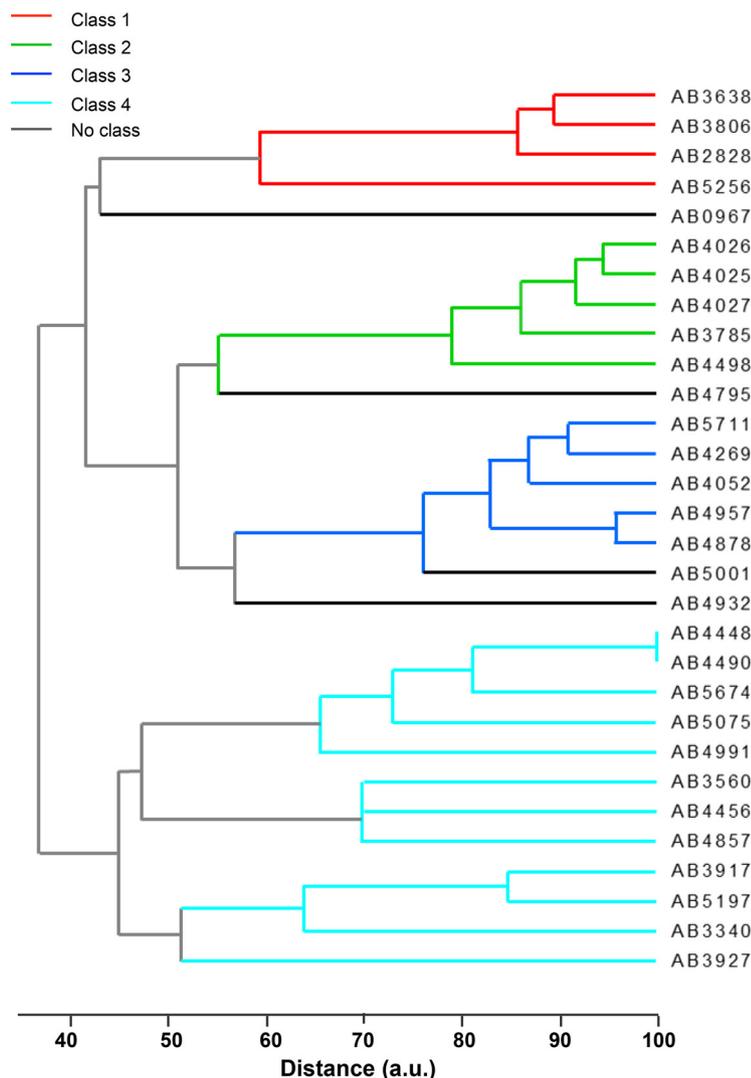
Although Raman spectroscopy has the capability of providing detailed molecular information by generating vast amounts of data, this can also be a drawback because the actual spectral interpretation is complex and requires robust data analysis and a comprehensive database (26). Moreover, fluorescence background can sometimes overwhelm Raman spectra, diminishing or hiding the Raman spectral peaks. Here, we evaluate the capability of Raman spectroscopy to reliably differentiate the *A. baumannii* strains responsible for clinical infections using hierarchical cluster analysis. Additionally, the HCA results for Raman spectral data are compared to the clustering results for three other data types, including data from mass spectrometry, DNA sequencing, and antimicrobial susceptibility testing (AST). For each resulting dendrogram, the strain type (based on PFGE type) is indicated on branches.

## RESULTS

Initially, a dendrogram was plotted using the PFGE banding pattern outcomes (Fig. 1). Clades on the resulting dendrogram were separated into PFGE classes 1, 2, 3, and 4. Some strains could not be assigned to a PFGE class. The PFGE classes were then utilized as class labels for the rest of the data sets. As demonstrated in Fig. 2, four dendrograms were plotted using DNA sequencing, AST, MS, and Raman spectral data sets of 31 *A. baumannii* strains.

In order to quantitatively compare the performance of the Raman spectral hierarchical clustering to the clustering of the other data types, external criteria of Rand index values were calculated using the *a priori* knowledge of Apal digestion-PFGE classes (Fig. 1). The Rand index is defined as the number of pairs of objects that are either in the same groups in both partitions or in different groups in both partitions, divided by the total number of pairs of objects (27, 28). The value of the Rand index lies between 0 and 1, where 1 indicates that two partitions agree perfectly, and 0 indicates that the two partitions disagree completely (29). This method is useful in comparing cluster results with different numbers of clusters between the groups being compared, which is the case in the data sets reported here.

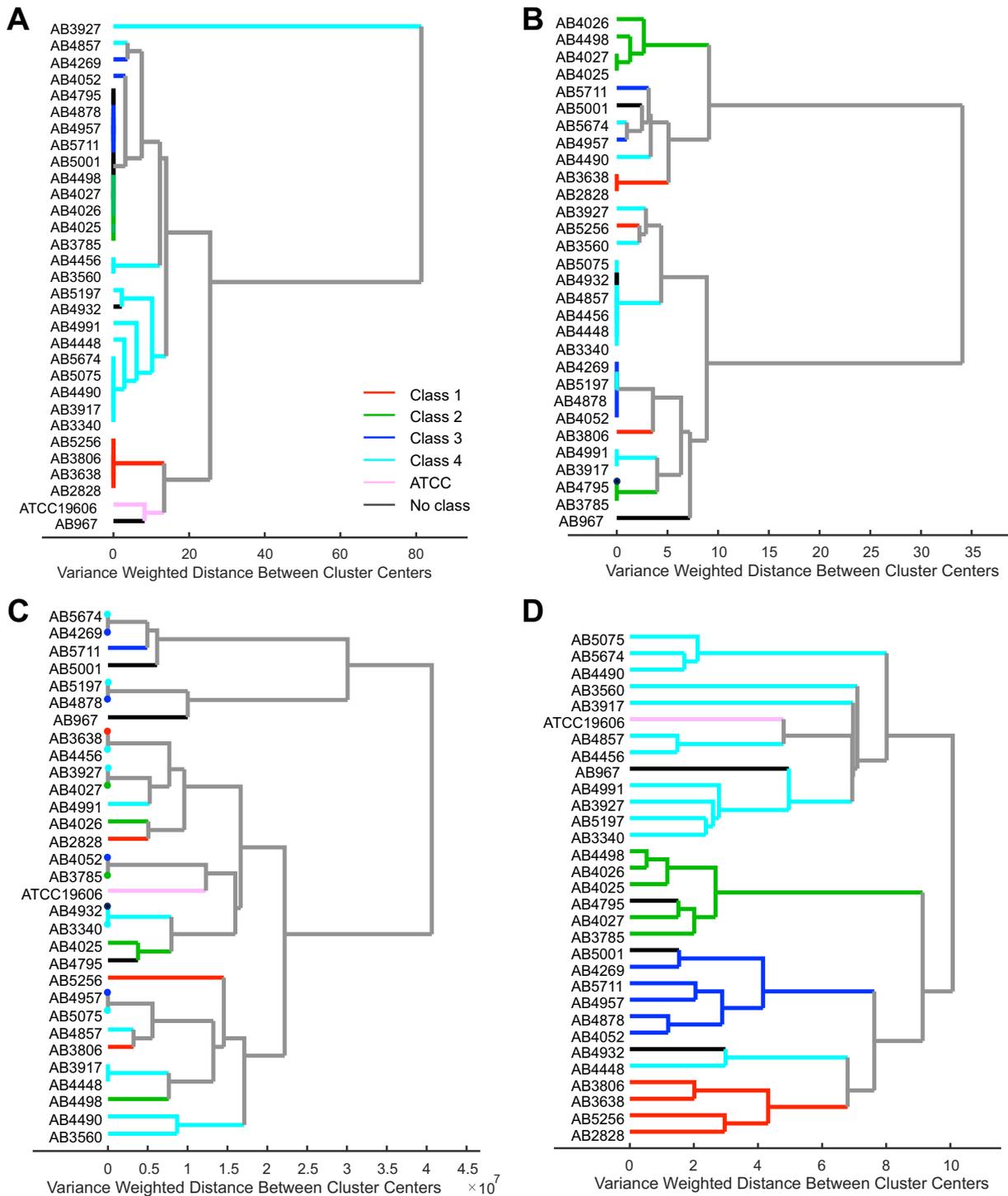
Due to the different types of data sets being analyzed here, it is expected that the hierarchical trees will demonstrate some variability in their structure. Therefore, the Rand index values were calculated by “cutting the dendrogram,” a method where the tree is scanned to detect the level possessing the optimal recovery of desired clusters



**FIG 1** PFGE dendrogram of 30 *A. baumannii* clinical isolates clustered using unweighted pair group method with arithmetic mean (UPGMA) implemented in Bionumerics 7.5 software (Applied Maths, Sint-Martens-Latem, Belgium). Clades on the resulting dendrogram were separated into PFGE classes 1, 2, 3, and 4. Some strains could not be assigned to a PFGE class.

(29). Each of the dendrograms displayed in Fig. 1 was individually evaluated, and a cutoff point was determined to reveal the inherent groupings of the data sets according to their PFGE classes. Those without a class label (strains AB967, AB4795, AB4932, AB5001, and ATCC 19606) were excluded for the Rand index calculations.

The dendrogram for the DNA sequences (Fig. 2A) matches fairly well with the PFGE classes with the exception of two misclustered isolates from class 4 and merging of isolates from classes 2 and 3 into one branch, giving a Rand index value of 0.78. AST successfully clustered class 2 isolates and the majority of class 4 isolates, giving a Rand index value of 0.70, while the rest of the isolates did not group well according to PFGE classes (Fig. 2B). The dendrogram for the MS data (Fig. 2C) shows the most disagreement with the PFGE classes, also resulting in the lowest Rand index value (0.58), compared to the rest of the data sets. Raman spectral data most accurately clustered the strains according to PFGE, forming four distinct clusters corresponding to the four PFGE classes, as depicted in Fig. 2D, and demonstrated a high Rand index value of 0.88. Leave-one-out cross-validation accuracy of Raman spectral clustering was calculated using the cutting the dendrogram method described above for each of the 26 models,



**FIG 2** Hierarchical cluster analysis (HCA) results of *A. baumannii* clinical isolates using DNA sequences aligned via MAFFT (multiple alignment using fast Fourier transform) and numerically coded to perform HCA using Ward's linkage and Euclidean distance (A), antimicrobial susceptibility testing data using Ward's linkage and Euclidean distance (B), preprocessed MS data performed using Ward's linkage and Euclidean distance (C), and preprocessed Raman spectroscopic data after performing principal-component analysis (D). HCA was done using Ward's linkage and Mahalanobis distance. All dendrograms are colored according to their resulting strain type (based on PFGE type, indicated by classes 1, 2, 3, and 4).

which resulted in 92.3% accuracy for strain differentiation, with the misclustering of samples AB3560 and AB4448.

**DISCUSSION**

Among bacterial strain typing methods, PFGE is considered to have the most discriminatory power in inferring relationships between strains and has been utilized as

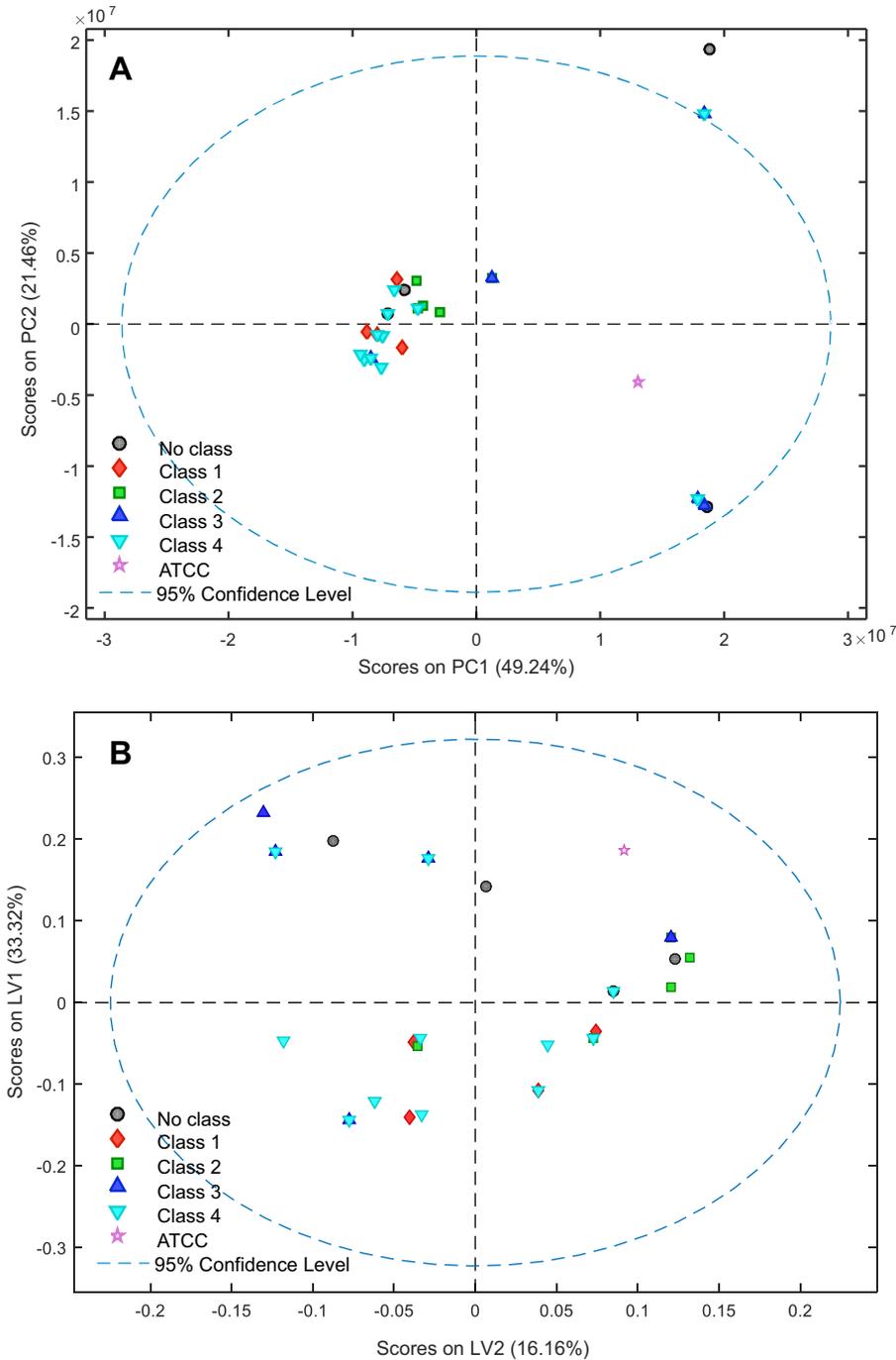
the gold standard for identifying microbial outbreaks (30–32). It provides high-resolution macrorestriction analysis, resolving microorganism identity using genomic data (32). Despite its advantages, however, PFGE is a time- and labor-intensive process with high consumable costs. Moreover, the technique's shortcomings include a lack of interlaboratory reproducibility and an inability to differentiate between bands nearly identical in size (32).

Therefore, many studies have explored alternative methods of microbial identification that are accurate, rapid, and cost-effective, such as MS and Raman spectroscopy (4, 10, 12–18, 33, 34). Other methods, such as DNA sequencing and AST, have also been used to differentiate among strains of microorganisms, including *A. baumannii* (35–37). AST, although not a primary bacterial typing method, was included in this study to examine its relationship with PFGE, Raman spectroscopy, and other typing methods. AST is implemented when antimicrobials cannot be predicted reliably based on knowledge of microbial identity (38); nonetheless, it has been shown to correlate with PFGE results in *A. baumannii* clinical isolates (36) and other bacterial species (35); therefore, it was added to this study for comparison.

In this study, Raman spectral data most accurately clustered *A. baumannii* strains according to their PFGE class via HCA, which was not only noticeable visually but was also reinforced by the high Rand index value. Similar results were observed by Maquelin et al. (18), where Raman spectra of a collection of well-characterized *Acinetobacter* species were analyzed via HCA and compared to HCA results of amplified fragment length polymorphism (AFLP) typing. Their findings, much like ours, showed highly similar groupings by both techniques. For the Raman data set, performance could be further improved with alternative spectral preprocessing and further exploration of more discriminative spectral regions. In our study, the agreement between PFGE and Raman spectroscopy results is not surprising; PFGE provides discrimination on the basis of differences in DNA sequences, and Raman spectroscopy provides discrimination on the basis entire cells, including DNA sequences.

In order to investigate the reasons behind successful clustering of Raman spectra, we examined the loading of the first few principal components (PCs) of the Raman spectra that accounted for the most variance in the data. The first PC, which accounted for 50.77% of variance, was predominantly composed of the region between  $1,200\text{ cm}^{-1}$  and  $1,400\text{ cm}^{-1}$ , with peaks at  $1,254\text{ cm}^{-1}$  and  $1,305\text{ cm}^{-1}$ ; these vibrational bands, associated with the amide III envelope, can be indicative of proteins (39). The second and third PCs, which accounted for 15.21% and 10.70% of variance, respectively, represented Raman spectral peaks at  $979\text{ cm}^{-1}$ ,  $1,129\text{ cm}^{-1}$ ,  $1,378\text{ cm}^{-1}$ , and  $1,411\text{ cm}^{-1}$ . The  $979\text{ cm}^{-1}$  band, which was prominent in PC2, was observed in specific *A. baumannii* strains in some of our previous studies; thus, PC2 is attributed to the bacterial capsule phenotype (40). Three of the *A. baumannii* strains in this study (AB5075, AB5674, and AB4490) exhibit this unique  $979\text{ cm}^{-1}$  band and, as a result, clustered distinctly with each other, as seen in Fig. 2D. The third PC (PC3) was composed of a mix of protein- and lipid-related vibrational bands; the amide III envelope and the  $1,179\text{ cm}^{-1}$  band are typical protein vibrational bands, while the spectral peaks at  $1,078\text{ cm}^{-1}$  and  $1,378\text{ cm}^{-1}$  are hallmarks of lipid vibrational bands, all of which can be associated with cell membrane components (41).

While much success has been achieved in the application of MS to identify *A. baumannii* isolates at the species level (5–7, 33, 34), difficulties seem to arise with MS when it is applied for strain-level identification of *A. baumannii* isolates (4, 8–11). In this study, to further evaluate the discriminative capability of MS, multiple chemometric methods were implemented aside from HCA, such as PCA and PLS-DA, along with various preprocessing methods (Fig. 3). However, no optimal separations were observed between the four PFGE classes. With the MS data, many of the samples from different PFGE classes show no clustering by class, further supporting the idea that MALDI-TOF MS has difficulty identifying closely related bacterial species and strains due to their close, if not identical, protein expressions (4, 12). This close relationship between MS spectra is also exhibited in the dendrogram, where many strains in the



**FIG 3** (A) PCA model of mass spectrometry data, preprocessed using a fourth-order Savitzky-Golay derivative and mean centering (spectral region used: 0 to 4,000 *m/z* values). (B) PLSDA model of mass spectrometry data, latent variable biplot of spectra preprocessed by normalizing to all variables and mean centering (spectral region used, 0 to 3,300 *m/z* values). Data are colored according to strain type (based on PFGE type indicated by classes 1, 2, 3, and 4).

same cluster, but belonging to different PFGE classes (such as AB5674 and AB4269, AB3638, AB4456, etc.), show a straight line on the distance axis, which indicates that there is no linkage distance between them.

While MALDI-TOF is rapid and relatively easy, this report, along with other studies (4, 8, 10, 11), revealed that there are significant disadvantages for its application in closely related microorganisms. For example, in a study done by Rim et al. (10), dendrograms were generated by MS data for *A. baumannii* strains collected from an intensive care

unit and were compared to a dendrogram generated using PFGE results. The study concluded that the dendrogram function of MALDI-TOF MS could not substitute for PFGE in MDR *A. baumannii* clonality analysis. Another study by Kishii and coworkers demonstrated only 72.4% accuracy for MALDI-TOF MS identification of *Acinetobacter* isolates from blood cultures (8). Finally, Alvarez-Buylla et al. concluded in a recent study that species within the genus *Acinetobacter* still require molecular techniques for identification (11).

In this study, however, we took one step further and compared MS spectral differentiation of strains with Raman spectroscopic differentiation of strains along with additional data types for *A. baumannii* isolates. To our knowledge, this the first study of this type to have been performed. The clear advantages of Raman spectroscopy shown here are that it is as rapid as MS and lower in cost while providing reliable and more highly discriminative power than MS for *A. baumannii* strain identification. Moreover, this study also demonstrates the advantages of Raman spectroscopy over both gene sequence and antimicrobial resistance data types for *A. baumannii* strain differentiation. The accuracy of the leave-one-out cross validation of the Raman spectral data model is an encouraging factor in improving and building on such models using additional data and validates using a separate test data set for future work. The enhanced performance of Raman spectroscopy for *A. baumannii* strain differentiation in this study may be attributed to the fact that the Raman spectra contain information about the whole cell, including cell morphology, rather than just genomic data.

In this study, hierarchical cluster analysis of Raman spectroscopic data, mass spectrometry, DNA sequence, and AST data sets of 31 *A. baumannii* clinical isolates was performed. As evidenced by the Rand indices of the different data sets, Raman spectroscopy more accurately differentiated among strains than the other typing methods. Furthermore, this study highlights the unreliability of MALDI-TOF MS, the current clinical standard for routine bacterial identification, in differentiating between closely related *A. baumannii* strains. This report suggests Raman spectroscopy as a promising new method for routine strain-level identification in *A. baumannii* clinical isolates. Future work will include developing and validating models using additional *A. baumannii* isolates and other MDR bacterial species using Raman spectroscopic data.

## MATERIALS AND METHODS

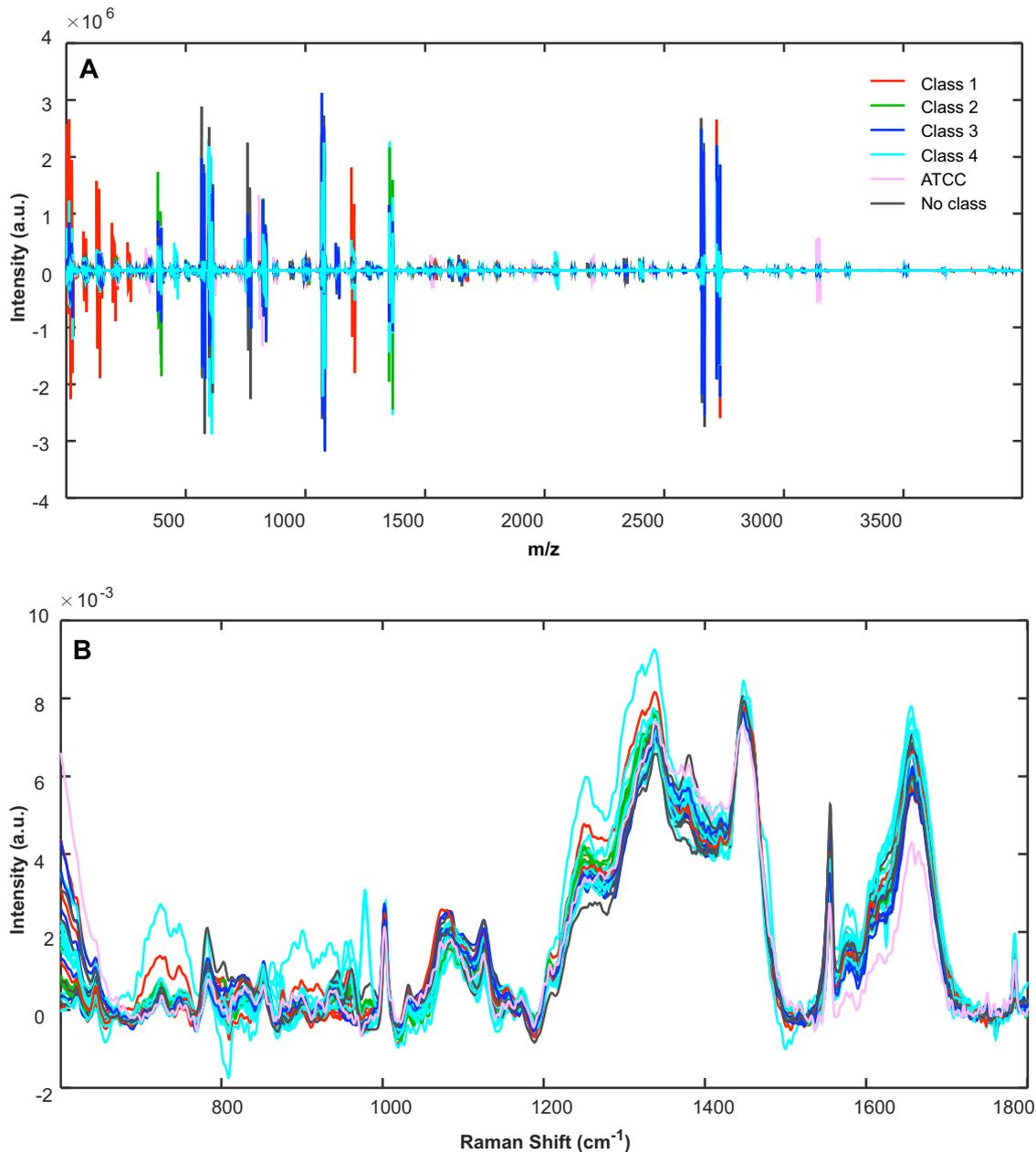
**Bacterial isolates.** 30 clinical isolates of *A. baumannii* were obtained from the Wound Infections Department at Walter Reed Army Institute of Research. This set of isolates has been previously published (42), and represents a diverse sample of MDR *A. baumannii* isolates recovered from military personnel under treatment at Walter Reed Army Medical Center in Washington, DC. A reference strain of *A. baumannii* (ATCC 19606) was obtained from the ATCC (Manassas, VA, USA).

**Pulsed-field gel electrophoresis.** Isolate strain identity was initially determined via restriction fragment length polymorphism (RFLP) followed by pulsed-field gel electrophoresis (PFGE). Briefly, genomic DNA was digested by an *Apal* restriction enzyme, and digested products were run on an agarose gel under an electric current to produce a DNA banding pattern. The banding pattern similarity coefficient was calculated using Dice's coefficient with 1% band-matching tolerance, and the results were clustered using unweighted pair group method with arithmetic mean (UPGMA) and implemented in Bionumerics 7.5 software (Applied Maths, Sint-Martens-Latem, Belgium).

**DNA sequencing.** Three genomic loci were sequenced for each isolate in the study. The name and primers for each locus are putative permease (5'-AAC ATG GGA TGG CTT GGT TTT-3' and 5'-CAG ATC TAC CCG TGC CTT GAT AA-3'), glutamyl tRNA ligase (5'-CAA ACC GCA TAG GAA AGA AAA GA-3' and 5'-CCT GAG AGG GAA TCA AAC TT-3'), and oxidoreductase (5'-CAA ACC GCA TAG GAA AGA AAA GA-3' and 5'-AAG TCC GCC CAG GTC AGC-3'). Primers were designed using Lasergene (DNASTAR, Inc., Madison, WI, USA) and synthesized by IDT (Integrated DNA Technologies, Coralville, IA, USA).

PCR was performed for each of the three loci in each isolate. The reactions were prepared using GoTaq Green master mix (Promega, Madison, WI, USA) in 50- $\mu$ L reactions according to the manufacturer's instructions. The reactions were run on a DNA Engine Dyad thermal cycler (Bio-Rad, Hercules, CA, USA) under the following conditions: 95°C for 2 min, followed by 35 cycles of 94°C for 1 min, 45°C for 1 min, and 72°C for 1.5 min. Final extension at 72°C was done for 3.5 min. A final volume of 15  $\mu$ L of amplified product was analyzed by agarose gel electrophoresis to confirm successful amplification. Reactions were purified and sequenced by Macrogen USA (Rockville, MD, USA). Resulting sequences were trimmed for quality using a Sequencher (Gene Codes, Ann Arbor, MI, USA).

**Antimicrobial susceptibility testing.** AST of isolates was performed on a BD Phoenix system (BD, Franklin Lakes, NJ, USA). The panels used for testing in this study were NMIC/ID-128 and NMIC/ID-132 (BD).



**FIG 4** Overlaid, preprocessed spectra of 31 *A. baumannii* clinical isolates using MS data, preprocessed by taking the third-order derivative with a filter width of 15, followed by mean centering (A), and Raman spectroscopic data preprocessed by truncation of data between the 600 and 1,800  $\text{cm}^{-1}$  Raman spectral region, baseline removal using a sixth-order polynomial, and normalization to the 1,445  $\text{cm}^{-1}$  Raman spectral band (B). Spectra are colored according to their resulting strain type (based on PFGE type indicated by classes 1, 2, 3, and 4).

**Mass spectrometry.** *A. baumannii* isolates were received from the Naval Medical Research Center (NMRC) on blood agar slants (catalog number R060310; Remel, Lenexa, KS, USA). The isolates were grown on Trypticase soy agar (TSA) with sheep blood agar (catalog number R01201; Remel) for 18 to 24 h at 37°C in a 5%  $\text{CO}_2$  incubator. MALDI-TOF MS data were collected on a Vitek MS system (bioMérieux SA, Marcy l'Étoile, France). A colony of interest was sampled using a 1- $\mu\text{l}$  inoculating loop. The collected sample was then smeared onto a MALDI slide (catalog number 410893; bioMérieux SA). This was done four times to make four spots using four different colonies per plate. A suspension of  $\alpha$ -cyano-4-hydroxycinnamic acid (Vitek MS CHCA, bioMérieux SA) was overlaid onto each sample (approximately 1  $\mu\text{l}$  dispensed by pipette). The matrix was allowed to dry before the slide was loaded into the mass spectrometer for analysis. Each group of 16 spots was tested along with a spot of *Escherichia Coli* (ATCC 8739), which served as a calibration target as well as a quality control indicator.

Mass spectra were generated in positive linear mode with a nitrogen laser ( $\lambda = 337 \text{ nm}$ ) operating with a pulse repetition frequency of 50 Hz. The ions were accelerated using a potential of 20 kV and an

extraction delay time of 200 ns. For each sample, 100 spectra were averaged and obtained from different positions of the target spot, operating in conjunction with the Acquisition Station software. A spectral range of 2,000 to 20,000 Da was collected. Peak lists and intensities for each isolate were then submitted for data analysis.

**Raman spectroscopy.** Isolates were grown on LB agar plates for 48 h at 37°C in 5% CO<sub>2</sub> and transferred into a 1.75-in.-width 0.38-in.-depth disposable plastic weigh dish covered with aluminum foil for spectral collection. A growth time of 48 h was used because it generated enough biomass to be transferred with a 10- $\mu$ l inoculating loop and smeared on the disposable weigh dish to make nine different spots (~3 mm in diameter each) to be measured by a particular Raman spectrometer. Raman spectra of the isolates were collected using a 785-nm Kaiser Rxn1 system (KOSI, Ann Arbor, MI, USA) equipped with a fiber-optic PhAT probe attachment with a 3-mm spot size lens. Each spectrum was the sum of 50 accumulations of 2 s and was collected for nine different spots of the same *A. baumannii* isolate; thus, nine Raman spectra were collected for each of the isolates.

**Data analysis.** All processing of data (aside from DNA sequence alignments) was performed in MATLAB using PLS Toolbox (Eigenvector Research, Inc., Wenatchee, WA, USA).

**DNA sequencing.** DNA sequences of the 31 *A. baumannii* isolates were aligned using multiple alignment using fast Fourier transform (MAFFT) (43). The sequences for each locus were aligned separately and then concatenated. The concatenated sequences were then numerically coded and imported into MATLAB to perform hierarchical cluster analysis (HCA).

**Antimicrobial susceptibility testing data.** Cluster analysis was also performed using the AST data for the 30 clinical isolates (excluding reference strain ATCC 19606 for lack of AST data). Data were numerically coded, and HCA was performed using Ward's linkage and Euclidean distance.

**Mass spectrometry data.** Four spectra per sample were collected. In order to account for spectral variability due to electronic variation in timing, a binning technique was implemented. Spectra were binned by identifying all the individual x-axis *m/z* values for each sample, finding the minimum and maximum values from this set, and building a mass axis of constant spacing with a desired input resolution (in this case, a resolution of two data points). Spectra were then averaged per sample and preprocessed by taking the third-order derivative with a filter width of 15 followed by mean centering (Fig. 4A). Finally, HCA was performed using Ward's linkage and Euclidean distance.

**Raman spectroscopy data.** Spectra were averaged per sample. Spectrum preprocessing included truncation of the Raman spectrum to a range of 600 to 1,800 cm<sup>-1</sup>, baseline removal using a sixth-order polynomial, and normalization to the 1,445 cm<sup>-1</sup> Raman spectral band (Fig. 4B). The region between 900 and 1,540 cm<sup>-1</sup> of the preprocessed data was then selected and mean centered before principal-component analysis (PCA) and hierarchical cluster analysis were performed. For HCA, nine principal components were used along with Ward's linkage and Mahalanobis distance. In order to validate this model, leave-one-out cross validation was carried out by performing PCA using 30 isolates and projecting the held-out test sample onto the PCA space. This was done 26 times (for the 26 samples with PFGE class labels). The first nine principal components of these models were used to create HCA models, along with Ward's linkage and Mahalanobis distance.

## ACKNOWLEDGMENTS

We thank the WRAIR Wound Infections Department, which provided the bacterial isolates used in this study.

This work was prepared as part of the authors' official duties. Title 17 U.S.C. §105 provides that "Copyright protection under this title is not available for any work of the United States Government." Title 17 U.S.C. §101 defines a U.S. Government work as a work prepared by a military service member or employee of the U.S. Government as part of that person's official duties. The views expressed in this article are those of the author and do not necessarily reflect the official policy or position of the Department of the Navy, Department of Defense, or the U.S. Government. This material has been reviewed by the Walter Reed Army Institute of Research. There is no objection to its presentation and/or publication. The opinions or assertions contained herein are the private views of the author and are not to be construed as official or as reflecting true views of the Department of the Army or the Department of Defense.

This study was performed under MUA 228 with the Walter Reed Army Institute of Research. We certify that all individuals who qualify as authors have been listed; each has participated in the conception and design of this work, the analysis of data (when applicable), the writing of the document, and the approval of the submission of this version; that the document represents valid work; that if we used information derived from another source, we obtained all necessary approvals to use it and made appropriate acknowledgments in the document; and that each takes public responsibility for it.

This effort was supported (in part) by the U.S. Navy Bureau of Medicine and Surgery under the Medical Development Program and Office of Naval Research work unit

number 602115HP.3720.001.A1015, USAMRMC Military Medical Research and Development award OR090136, and The Geneva Foundation (USAMRAA award W81XWH-13-2-0039).

## REFERENCES

- Magiorakos AP, Srinivasan A, Carey RB, Carmeli Y, Falagas ME, Giske CG, Harbarth S, Hindler JF, Kahlmeter G, Olsson-Liljequist B, Paterson DL, Rice LB, Stelling J, Struelens MJ, Vatopoulos A, Weber JT, Monnet DL. 2012. Multidrug-resistant, extensively drug-resistant and pandrug-resistant bacteria: an international expert proposal for interim standard definitions for acquired resistance. *Clin Microbiol Infect* 18:268–281. <https://doi.org/10.1111/j.1469-0691.2011.03570.x>.
- Mencacci A, Monari C, Leli C, Merlini L, De Carolis E, Vella A, Cacioni M, Buzi S, Nardelli E, Bistoni F, Sanguinetti M, Vecchiarelli A. 2013. Typing of nosocomial outbreaks of *Acinetobacter baumannii* by use of matrix-assisted laser desorption ionization-time of flight mass spectrometry. *J Clin Microbiol* 51:603–606. <https://doi.org/10.1128/JCM.01811-12>.
- Maatallah M, Bakhrouf A, Habeeb MA, Turlej-Rogacka A, Iversen A, Pourcel C, Sioud O, Giske CG. 2013. Four genotyping schemes for phylogenetic analysis of *Pseudomonas aeruginosa*: comparison of their congruence with multi-locus sequence typing. *PLoS One* 8:e82069. <https://doi.org/10.1371/journal.pone.0082069>.
- Sousa C, Botelho J, Grosso F, Silva L, Lopes J, Peixe L. 2015. Unsuitability of MALDI-TOF MS to discriminate *Acinetobacter baumannii* clones under routine experimental conditions. *Front Microbiol* 6:481. <https://doi.org/10.3389/fmicb.2015.00481>.
- Hsueh PR, Kuo LC, Chang TC, Lee TF, Teng SH, Chuang YC, Teng LJ, Sheng WH. 2014. Evaluation of the Bruker Biotyper matrix-assisted laser desorption ionization-time of flight mass spectrometry system for identification of blood isolates of *Acinetobacter* species. *J Clin Microbiol* 52:3095–3100. <https://doi.org/10.1128/JCM.01233-14>.
- Pailhories H, Daure S, Eveillard M, Joly-Guillou ML, Kempf M. 2015. Using Vitek MALDI-TOF mass spectrometry to identify species belonging to the *Acinetobacter calcoaceticus*-*Acinetobacter baumannii* complex: a relevant alternative to molecular biology? *Diagn Microbiol Infect Dis* 83: 99–104. <https://doi.org/10.1016/j.diagmicrobio.2015.06.009>.
- Jeong S, Hong JS, Kim JO, Kim KH, Lee W, Bae IK, Lee K, Jeong SH. 2016. Identification of *Acinetobacter* species using matrix-assisted laser desorption ionization-time of flight mass spectrometry. *Ann Lab Med* 36:325–334. <https://doi.org/10.3343/alm.2016.36.4.325>.
- Kishii K, Kikuchi K, Matsuda N, Yoshida A, Okuzumi K, Uetera Y, Yasuhara H, Moriya K. 2014. Evaluation of matrix-assisted laser desorption ionization-time of flight mass spectrometry for species identification of *Acinetobacter* strains isolated from blood cultures. *Clin Microbiol Infect* 20:424–430. <https://doi.org/10.1111/1469-0691.12376>.
- Sedo O, Nemeč A, Krizova L, Kacalova M, Zdrahal Z. 2013. Improvement of MALDI-TOF MS profiling for the differentiation of species within the *Acinetobacter calcoaceticus*-*Acinetobacter baumannii* complex. *Syst Appl Microbiol* 36:572–578. <https://doi.org/10.1016/j.syapm.2013.08.001>.
- Rim JH, Lee Y, Hong SK, Park Y, Kim M, D'Souza R, Park ES, Yong D, Lee K. 2015. Insufficient discriminatory power of matrix-assisted laser desorption ionization time-of-flight mass spectrometry dendrograms to determine the clonality of multi-drug-resistant *Acinetobacter baumannii* isolates from an intensive care unit. *Biomed Res Int* 2015:535027. <https://doi.org/10.1155/2015/535027>.
- Alvarez-Buylla A, Culebras E, Picazo JJ. 2012. Identification of *Acinetobacter* species: is Bruker biotyper MALDI-TOF mass spectrometry a good alternative to molecular techniques? *Infect Genet Evol* 12:345–349. <https://doi.org/10.1016/j.meegid.2012.01.002>.
- Cox CR, Jensen KR, Saichek NR, Voorhees KJ. 2015. Strain-level bacterial identification by CeO<sub>2</sub>-catalyzed MALDI-TOF MS fatty acid analysis and comparison to commercial protein-based methods. *Sci Rep* 5:10470. <https://doi.org/10.1038/srep10470>.
- Buijtelts PC, Willemsse-Erix HF, Petit PL, Endtz HP, Puppels GJ, Verbrugh HA, van Belkum A, van Soelingen D, Maquelin K. 2008. Rapid identification of mycobacteria by Raman spectroscopy. *J Clin Microbiol* 46: 961–965. <https://doi.org/10.1128/JCM.01763-07>.
- Zeiri L, Bronk BV, Shabtai Y, Czege J, Efrima S. 2002. Silver metal induced surface enhanced Raman of bacteria. *Colloids Surf A: Physicochem Eng Asp* 208:357–362. [https://doi.org/10.1016/S0927-7757\(02\)00162-0](https://doi.org/10.1016/S0927-7757(02)00162-0).
- Zeiri L, Bronk BV, Shabtai Y, Eichler J, Efrima S. 2004. Surface-enhanced Raman spectroscopy as a tool for probing specific biochemical components in bacteria. *Appl Spectrosc* 58:33–40. <https://doi.org/10.1366/000370204322729441>.
- Escoriza MF, VanBriesen JM, Stewart S, Maier J, Treado PJ. 2006. Raman spectroscopy and chemical imaging for quantification of filtered waterborne bacteria. *J Microbiol Methods* 66:63–72. <https://doi.org/10.1016/j.mimet.2005.10.013>.
- Kalaszinsky KS, Hadfield T, Shea AA, Kalaszinsky VF, Nelson MP, Neiss J, Drauch AJ, Vanni GS, Treado PJ. 2007. Raman chemical imaging spectroscopy reagentless detection and identification of pathogens: signature development and evaluation. *Anal Chem* 79:2658–2673. <https://doi.org/10.1021/ac0700575>.
- Maquelin K, Dijkshoorn L, van der Reijden TJ, Puppels GJ. 2006. Rapid epidemiological analysis of *Acinetobacter* strains by Raman spectroscopy. *J Microbiol Methods* 64:126–131. <https://doi.org/10.1016/j.mimet.2005.04.028>.
- Maquelin K, Kirschner C, Choo-Smith LP, van den Braak N, Endtz HP, Naumann D, Puppels GJ. 2002. Identification of medically relevant microorganisms by vibrational spectroscopy. *J Microbiol Methods* 51: 255–271. [https://doi.org/10.1016/S0167-7012\(02\)00127-6](https://doi.org/10.1016/S0167-7012(02)00127-6).
- Prucek R, Ranc V, Kvitik L, Panacek A, Zboril R, Kolar M. 2012. Reproducible discrimination between Gram-positive and Gram-negative bacteria using surface enhanced Raman spectroscopy with infrared excitation. *Analyst* 137:2866–2870. <https://doi.org/10.1039/c2an16310a>.
- Wu X, Huang YW, Park B, Tripp RA, Zhao Y. 2015. Differentiation and classification of bacteria using vancomycin functionalized silver nanorods array based surface-enhanced Raman spectroscopy and chemometric analysis. *Talanta* 139:96–103. <https://doi.org/10.1016/j.talanta.2015.02.045>.
- Jarvis RM, Goodacre R. 2004. Discrimination of bacteria using surface-enhanced Raman spectroscopy. *Anal Chem* 76:40–47. <https://doi.org/10.1021/ac034689c>.
- Neugebauer U, Rosch P, Popp J. 2015. Raman spectroscopy towards clinical application: drug monitoring and pathogen identification. *Int J Antimicrob Agents* 46 Suppl 1:S35–a39. <https://doi.org/10.1016/j.ijantimicag.2015.10.014>.
- Rosch P, Harz M, Schmitt M, Peschke KD, Ronneberger O, Burkhardt H, Motzkus HW, Lankers M, Hofer S, Thiele H, Popp J. 2005. Chemotaxonomic identification of single bacteria by micro-Raman spectroscopy: application to clean-room-relevant biological contaminations. *Appl Environ Microbiol* 71:1626–1637. <https://doi.org/10.1128/AEM.71.3.1626-1637.2005>.
- Ghebremedhin M, Yesupriya S, Luka J, Crane NJ. 2015. Validation of hierarchical cluster analysis for identification of bacterial species using 42 bacterial isolates. *Proc SPIE* 9318, Optical Biopsy XIII: Toward Real-Time Spectroscopic Imaging and Diagnosis, 9318W.
- Ashton L, Lau K, Winder CL, Goodacre R. 2011. Raman spectroscopy: lighting up the future of microbial identification. *Future Microbiol* 6:991–997. <https://doi.org/10.2217/fmb.11.89>.
- Yeung KY, Medvedovic M, Bumgarner RE. 2003. Clustering gene-expression data with repeated measurements. *Genome Biol* 4:R34. <https://doi.org/10.1186/gb-2003-4-5-r34>.
- Manning CD, Raghavan P, Schütze H. 2008. Introduction to Information Retrieval. Cambridge University Press, New York, NY.
- Milligan GW, Cooper MC. 1986. A study of the comparability of external criteria for hierarchical cluster analysis. *Multivariate Behav Res* 21: 441–458. [https://doi.org/10.1207/s15327906mbr2104\\_5](https://doi.org/10.1207/s15327906mbr2104_5).
- Santos-Sanches I, Chambel L, Tenreiro R. 2015. Pulsed-Field Gel Electrophoresis (PFGE): Application in population structure studies of bovine mastitis-causing *Streptococci*, p 323–334. In Cunha VM, Iñacio J (ed), *Veterinary Infection Biology: Molecular Diagnostics and High-Throughput Strategies*. Springer, New York, NY.
- Faria NA, Carrico JA, Oliveira DC, Ramirez M, de Lencastre H. 2008. Analysis of typing methods for epidemiological surveillance of both methicillin-resistant and methicillin-susceptible *Staphylococcus au-*

- reus strains. *J Clin Microbiol* 46:136–144. <https://doi.org/10.1128/JCM.01684-07>.
32. Li W, Raoult D, Fournier PE. 2009. Bacterial strain typing in the genomic era. *FEMS Microbiol Rev* 33:892–916. <https://doi.org/10.1111/j.1574-6976.2009.00182.x>.
  33. Toh B, Patterson DL, Kamolvit W, Zowawi H, Kvaskoff D, Sidjabat H, Wailan A, Peleg AY, Huber CA. 2015. Species identification within *Acinetobacter calcoaceticus*–*baumannii* complex using MALDI-TOF MS. *J Microbiol Methods* 118:128–132. <https://doi.org/10.1016/j.mimet.2015.09.006>.
  34. Sousa C, Botelho J, Silva L, Grosso F, Nemeč A, Lopes J, Peixe L. 2014. MALDI-TOF MS and chemometric based identification of the *Acinetobacter calcoaceticus*–*Acinetobacter baumannii* complex species. *Int J Med Microbiol* 304:669–677. <https://doi.org/10.1016/j.ijmm.2014.04.014>.
  35. Camerer A, Kohlenberg A, Stefanik D, Seifert H. 2011. Evaluation of quantitative antibiotic susceptibility testing by Vitek 2 as a routine method to predict strain relatedness of coagulase-negative staphylococci isolated from blood cultures. *J Clin Microbiol* 49:3355–3357. <https://doi.org/10.1128/JCM.05130-11>.
  36. Huang XZ, Frye JG, Chahine MA, Cash DM, Barber MG, Babel BS, Kasper MR, Whitman TJ, Lindler LE, Bowden RA, Nikolich MP. 2010. Genotypic and phenotypic correlations of multidrug-resistant *Acinetobacter baumannii*–*A. calcoaceticus* complex strains isolated from patients at the National Naval Medical Center. *J Clin Microbiol* 48:4333–4336. <https://doi.org/10.1128/JCM.01585-10>.
  37. Petti CA. 2007. Detection and identification of microorganisms by gene amplification and sequencing. *Clin Infect Dis* 44:1108–1114. <https://doi.org/10.1086/512818>.
  38. Jenkins SG, Schuetz AN. 2012. Current concepts in laboratory testing to guide antimicrobial therapy. *Mayo Clin Proc* 87:290–308. <https://doi.org/10.1016/j.mayocp.2012.01.007>.
  39. Uzunbajakava N, Lenferink A, Kraan Y, Willekens B, Vrensen G, Greve J, Otto C. 2003. Nonresonant Raman imaging of protein distribution in single human cells. *Biopolymers* 72:1–9. <https://doi.org/10.1002/bip.10246>.
  40. Regeimbal JM, Jacobs AC, Corey BW, Henry MS, Thompson MG, Pavlicek RL, Quinones J, Hannah RM, Ghebremedhin M, Crane NJ, Zurawski DV, Teneza-Mora NC, Biswas B, Hall ER. 2016. Personalized therapeutic cocktail of wild environmental phages rescues mice from *Acinetobacter baumannii* wound infections. *Antimicrob Agents Chemother* 60:5806–5816. <https://doi.org/10.1128/AAC.02877-15>.
  41. Athamneh AI, Senger RS. 2012. Peptide-guided surface-enhanced Raman scattering probes for localized cell composition analysis. *Appl Environ Microbiol* 78:7805–7808. <https://doi.org/10.1128/AEM.02000-12>.
  42. McQueary CN, Kirkup BC, Si Y, Barlow M, Actis LA, Craft DW, Zurawski DV. 2012. Extracellular stress and lipopolysaccharide modulate *Acinetobacter baumannii* surface-associated motility. *J Microbiol* 50:434–443. <https://doi.org/10.1007/s12275-012-1555-1>.
  43. Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* 30:772–780. <https://doi.org/10.1093/molbev/mst010>.