Analytical Methods

PAPER

Cite this: *Anal. Methods*, 2016, **8**, 7603

Rapid discrimination of *Enterococcus faecium* strains using phenotypic analytical techniques†

Najla AlMasoud,[a] Yun Xu,[a] David I. Ellis,[a] Paul Rooney,[b] Jane F. Turton[c] and Royston Goodacre*[a]

Clinical isolates of glycopeptide resistant enterococci (GRE) were used to compare three rapid phenotyping and analytical techniques. Fourier transform infrared (FT-IR) spectroscopy, Raman spectroscopy and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS) were used to classify 35 isolates of *Enterococcus faecium* representing 12 distinct pulsed-field gel electrophoresis (PFGE) types. The results show that the three analytical techniques provide clear discrimination among enterococci at both the strain and isolate levels. FT-IR and Raman spectroscopic data produced very similar bacterial discrimination, reflected in the Procrustes distance between the datasets (0.2125–0.2411, $p < 0.001$); however, FT-IR data provided superior prediction accuracy to Raman data with correct classification rates (CCR) of 89% and 69% at the strain level, respectively. MALDI-TOF-MS produced slightly different classification of these enterococci strains also with high CCR (78%). Classification data from the three analytical techniques were consistent with PFGE data especially in the case of isolates identified as unique by PFGE. This study presents phenotypic techniques as a complementary approach to current methods with a potential for high-throughput point-of-care screening enabling rapid and reproducible classification of clinically relevant enterococci.

Received 18th August 2016
Accepted 1st October 2016

DOI: 10.1039/c6ay02326f

www.rsc.org/methods

1. Introduction

Enterococcus is a highly significant genus of bacteria, which causes important clinical infections including urinary tract infections (UTIs), endocarditis, meningitis, catheter-related infections, bacteremia, wound infections, pelvic and intra-abdominal infections amongst others. Some of these Gram-positive cocci were originally classified as *Streptococcus* spp. until genomic analysis by Schleifer and Kilpper-Balz in 1984 demonstrated the requirement for a separate genus classification.[1] This well-known genus is part of the normal intestinal microflora of humans and other animals.[2] *Enterococcus* are also part of the lactic acid bacteria (LAB) group present in foods, and whilst they are able to spoil fresh meats,[3] they are important in ripening and development of certain foods (*i.e.* dairy products), as well as being used as probiotics in humans.[4]

The majority of human clinical isolates of enterococci belong to two species, *Enterococcus faecalis* and *Enterococcus faecium*.[5] In addition to their prevalence and pathogenicity, another very important factor associated with *Enterococcus* is the high level of antimicrobial resistance, particularly resistance to glycopeptide antibiotics (such as vancomycin, teicoplanin and telavancin); resistant strains are referred to as GRE (glycopeptide-resistant enterococci).[6,7]

There is a constant requirement to develop analytical methods for the discrimination of bacteria, which can be used in clinical diagnostics and food quality control. These methods should ideally be rapid, reproducible, easy to use and automated, in addition to having high resolution and sensitivity.[8] Over a decade ago, it was common to use methods, such as polymerase chain reaction (PCR) for identification of specific DNA sequences and recognition by antibodies *via* enzyme-linked immunosorbent assay (ELISA), to characterize bacteria. Although these techniques are sensitive and specific, and carried out using relatively inexpensive equipment, their use is limited by the complexity of preparation procedures and the requirement for specific primers and antibodies.[9–12] Nowadays, modern analytical techniques, such as matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS),[13–16] Fourier transform infrared (FT-IR) spectroscopy[17–21] and Raman spectroscopy[22–24] are also used for the characterization of bacteria. High dimensional and information rich datasets are produced from these techniques, which has also directly led to the requirement of robust and reliable chemometric methods to assist with data deconvolution and in-depth analysis.[25] This saw the introduction, acceptance and use

[a]*School of Chemistry, Manchester Institute of Biotechnology, University of Manchester, 131 Princess Street, Manchester, M1 7DN, UK. E-mail: roy.goodacre@manchester.ac.uk*

[b]*Belfast City Hospital, 51 Lisburn Rd, Belfast, BT9 7AB, UK*

[c]*Antimicrobial Resistance and Healthcare Associated Infection Reference Unit, National Infection Service, Public Health England, London, NW9 5EQ, UK*

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c6ay02326f

of chemometrics, such as discriminant function analysis (DFA)[22] and hierarchical cluster analyses (HCA).[26–28]

Previously, MALDI-TOF-MS has shown promising results for bacterial characterization.[13] FT-IR and Raman spectroscopy complement each other for bacterial classification; both are robust metabolic fingerprinting techniques and need little sample preparation.[29,30] FT-IR spectroscopy is used by many researchers since it is not only rapid but also offers a high-throughput and non-destructive method, allowing the analysis of intact bacteria and producing unique, reproducible and distinct biochemical fingerprints.[31] Raman spectroscopy shares similar advantages to FT-IR spectroscopy and also has the additional advantage of water being a very weak Raman scatter[32] so that samples do not need to be dried.

Here, the aim was to use these three distinct phenotypic approaches (namely MALDI-TOF-MS, FT-IR and Raman spectroscopies) in combination with rigorous chemometric analysis of the resultant datasets to classify 35 clinically relevant isolates of enterococci, which had been previously typed by pulsed-field gel electrophoresis (PFGE). This was carried out in order to compare the results from, and determine the efficiency of, these analytical techniques for the rapid differentiation of *E. faecium* strains. In future, this may allow clinical diagnostic laboratories to analyze multiple bacterial samples rapidly for infection control purposes in point-of-care setting within hospitals, clinics, or GP surgeries which would significantly accelerate diagnosis, and potentially ensure that the correct antimicrobial therapies were used if required.

2. Experimental

2.1 General chemicals

Trifluoroacetic acid (TFA), HPLC grade water, acetonitrile, sinapinic acid (SA), α-cyano-4-hydroxycinnamic acid (CHCA), and ferulic acid (FA) were purchased from Sigma-Aldrich (Dorset, UK).

2.2 Media

Two different types of media were used to culture the enterococci: lysogeny broth (LB) and nutrient agar (NA). LB was prepared by mixing 5 g of yeast extract (Amersham Life Sciences, Cleveland, USA), 10 g of tryptone (Formedia, Hunstanton, UK) and 10 g of NaCl dissolved in 1 L of distilled water and the broth was then autoclaved (at 121 °C and 15 psi for 45 min). NA was prepared from a preparatory mixture (beef extract 3 g L^{-1}, peptone 5 g L^{-1}, NaCl 8 g L^{-1} and agar 2 at 12 g L^{-1}) (Lab-M, Bury, UK) following the manufacturer's instructions (28 g in 1 L of deionised water) and the broth was autoclaved (at 121 °C and 15 psi for 15 min).

2.3 Enterococci strains

Isolates were from faecal samples from patients in a surgical ward in a hospital in Belfast, UK and were collected following an increase in enterococcal infections on the ward. Faecal material samples were screened onto Brilliance VRE (vancomycin-resistant enterococci) agar (Oxoid, Basingstoke, UK). This agar contains antibiotics, which eradicate all Gram-negative bacteria, and a high concentration of vancomycin, and therefore, only vancomycin-resistant Gram-positive bacteria can grow, which leads to selection of vancomycin-resistant enterococci strains. They were identified as *E. faecium* by a VITEK® system (bioMérieux) and their identity confirmed by MALDI-TOF analysis using a Bruker microflex instrument. The 35 isolates were typed using pulsed-field gel electrophoresis (PFGE) of *Sma*I-digested genomic DNA by Public Health England's National Reference Laboratory as described previously.[33] Table S1† summarizes information on the 35 clinical isolates, which were classified into 12 groups (12 PFGE-defined types) named: EC04, EC09, EC10, EC13, EC14, EC15, EC19, EC20, UNI 156, UNI 178, UNI 191 and UNI 214, where 'UNI' types describe isolates that were unique within the set. All samples were collected with ethical approval from the Northern Ireland Research Ethics Committee, reference number "10/NIR01/20". This work did not involve any experimentation on human subjects.

2.4 Bacterial isolates

The samples analyzed by the three techniques (*viz.* MALDI-TOF-MS, FT-IR and Raman) were collected from the same flask to avoid any variations between different preparations that may affect results obtained using the different analytical platforms. First, enterococci were cultured on nutrient agar (NA) plates for 24 h at 37 °C. A single colony from the agar culture was used to inoculate 50 mL of lysogeny broth (LB) in a 250 mL flask which was incubated overnight at 37 °C with shaking at 200 rpm. This was followed by measuring the optical density (OD) at 600 nm using a Biomate 5 spectrophotometer (Thermo, Hemel Hempstead, UK) for each isolate. The volume of analyzed bacterial suspension was then normalized to account for variation in cell biomass in the different replicate cultures (4 biological replicates were prepared for each isolate) and used to inoculate a fresh flask of broth, which was incubated at 37 °C for 11 h. This isolate enrichment step is required to reduce interference from mixtures of different strains, which can introduce a significant level of noise to readings from analytical methods. Subsequently, 10 mL from each flask was collected and centrifuged at 4800 *g* for 10 min and the pellet washed three times with sterile deionized water. Fig. S1† illustrates the preparation process.

For vibrational spectroscopic analysis, the collected pellets were suspended in suitable volumes of saline (0.9% (w/v) NaCl) depending on the OD (all isolates had approximately the same cell density). Then, 15 μL was spotted onto a silicon plate (Bruker Ltd., Coventry, UK) and was allowed to dry at 40 °C for 45 min before analysis with FT-IR spectroscopy. For Raman spectroscopy, 4 μL of each sample was spotted onto a stainless steel plate and then allowed to dry at 40 °C for 45 min.

For MALDI-TOF-MS, three different matrices were tested to find the most compatible matrix with enterococci; these matrices were: FA, SA and CHCA. In addition, 3 different deposition methods (sample-matrix) were tested as described previously[16] to find the best method for depositing the samples: mix, overlay and underlay (data not shown). SA matrix and the mix deposition method were found to be the optimal

 This journal is © The Royal Society of Chemistry 2016

combination for MALDI-TOF-MS analysis for these samples. On the day of analysis of the samples, the biomass was suspended in 1000 μL of 2% TFA then vortexed for 3 min. An equal volume of 1 μL of bacterial suspension and matrix were vortexed for 2 s and 2 μL of this mixture spotted onto a MALDI stainless steel plate and allowed to dry at ambient temperature.

2.5 Fourier transform infrared (FT-IR) spectroscopy

A FT-IR silicon sampling plate (Bruker Ltd., Coventry, UK) which contained 96 locations/spots was washed using 5% sodium dodecyl sulfate (SDS) solution. This was followed by washing the plate using deionized water and allowing it to dry at room temperature.[34] High-throughput screening (HTS) was carried out using a Bruker Equinox 55 FT-IR spectrometer. The HTX™ module described by Winder *et al.*[35] was used with this instrument. Transmission mode was used to analyze the dried biomass to produce FT-IR spectra. The parameters used for FT-IR analysis included the following: spectra were collected in the wavenumber range 4000 and 600 cm^{-1}, resolution was 4 cm^{-1} and each spectrum was the average of 64 co-adds. Spectral acquisition and subtracting the background were achieved using Opus software (Bruker Ltd.). Four biological replicates, each in four analytical replicates, were analyzed and analysis was performed in three machine runs, resulting in 1680 FT-IR spectra.

2.6 Raman spectroscopy

This was carried out using a confocal Raman system (inVia, Renishaw plc., Wotton-Under-Edge, UK) coupled with a 785 nm wavelength laser. A power intensity of ~30 mW was applied on the samples at an exposure time of 20 s. Four biological replicates and seven different locations within each sample spot were analyzed, resulting in a total of 980 Raman spectra.

2.7 MALDI-TOF-MS

The enterococci isolates were analyzed using an AXIMA-Confidence MALDI-TOF-MS (Shimadzu Biotech, Manchester, UK), equipped with a nitrogen pulsed UV laser with a wavelength of 337 nm. The parameters of this device were set as follows: 90 mV laser power, 91 acquired profiles with each profile containing 20 shots, linear TOF, positive ionization mode, and mass-to-charge (m/z) range of 1000–18 000. The spectra were collected using a circular raster pattern. The MALDI-TOF-MS device was calibrated using a protein mixture: insulin (5735), cytochrome c (12 362), and apomyoglobin (16 952) (Sigma-Aldrich). Each of 4 biological replicates from the 35 isolates was analyzed in four technical replicates on four different days; this led to the generation of a total of 560 MALDI-TOF-MS spectra (35 isolates × 4 biological replicates × 4 analytical replicates).

2.8 Data analysis

2.8.1 Data pre-processing. Opus software was used to export FT-IR data into ASCII format; data were then transferred into MATLAB 2012a (The Mathworks Inc., MA, US). All FT-IR spectra were baseline corrected using standard normal variate (SNV) to remove any light scattering effects. The analytical replicates were then averaged to reduce the number of redundant samples. Due to the large number of samples, 8 separate (96 spot silicon) sampling plates were used; therefore, it was necessary to correct for the subtle differences in signals from different silicon plates. This was achieved using a piece-wise direct standardization (PDS) model.[36] The PDS model was built on two different 'reference' isolates which were spotted on every plate. The pre-processed FT-IR spectra were then subjected to multivariate analysis (MVA, see below). Raman spectra were also normalized using standard normal variate (SNV) and then subjected to MVA.

MALDI-TOF-MS data were pre-processed as follows: (i) the baseline was corrected using asymmetric least squares (AsLS),[37] and (ii) spectra were normalized by dividing each individual baseline corrected spectrum by the square root of the sum of squares of the spectrum.[38] The pre-processed MALDI-TOF-MS data were subjected to the same data analysis flow as Raman and FT-IR spectral data.

2.8.2 Multivariate data analysis. A flowchart of multivariate data analysis is provided in Fig. 1. For all three datasets, two types of classification were performed: one at the strain level (*i.e.* 12 classes) defined by PFGE, and the other at the isolate level (*i.e.* 35 classes, one for each isolate).

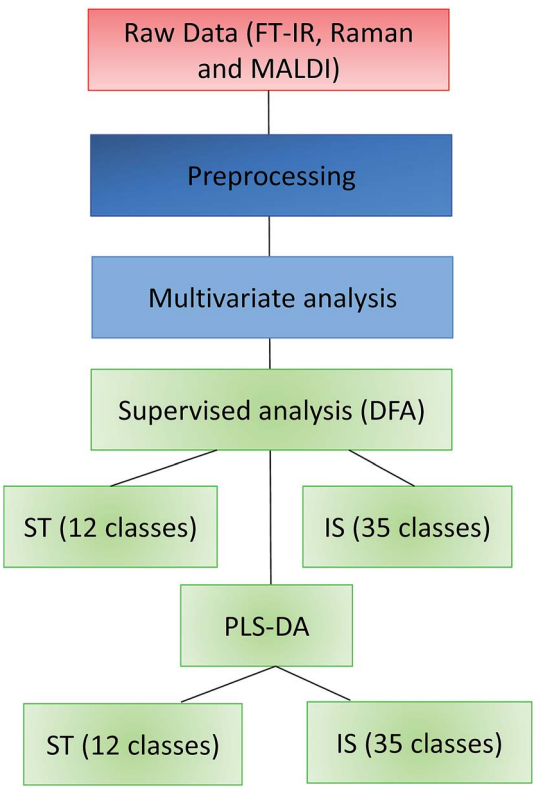

Fig. 1 Workflow of data analysis undertaken for FT-IR spectroscopy, Raman spectroscopy and MALDI-TOF-MS. The data were first pre-processed then multivariate analysis MVA was applied using PC-DFA at both the (ST) strain (12 classes) and (IS) isolate (35 classes) levels. This was followed by PLS-DA.

For cluster analyses, principal components-discriminant function analysis (PC-DFA)[39–41] was applied to reduce the dimensionality of the data and to discriminate samples from the designated classes. The PC-DFA scores of each class were then averaged and subjected to hierarchical cluster analysis (HCA).[42] Dendrograms from each analysis were generated to illustrate the relative relatedness of these bacteria.

Partial least squares-discriminant analysis (PLS-DA),[43] with 1000 bootstrapping validations,[44] was also applied to obtain a validated supervised classification model for discriminating different strains or isolates. In each bootstrapping process, the data were randomly split into two different sets: a training set and a test set. A PLS-DA model was trained using the training set and then applied to the test set to predict the class membership of the samples in the test set. This process was repeated 1000 times and the results were recorded and averaged to produce a $c \times c$ confusion matrix (c is the number of designated classes, either 12 (strains) or 35 (isolates)), in which the element at the i^{th} row, j^{th} column is the percentage of samples in class i being predicted as class j on average. In order to assess the statistical significance of the predictive performance of the PLS-DA models, a corresponding permutation test within each bootstrapping resampling was also performed. This means that in addition to building the PLS-DA model using the known class membership, another model (called the 'null' model) was also built using a randomly permuted class membership. The results of the null model were also recorded and from this the null distribution was obtained. An empirical p-value was calculated by counting the number of cases where the null model had obtained better predictive accuracy than the real model and dividing the obtained number by the total number of bootstrapping resampling (*i.e.* 1000 in this study).

Finally, similarities between the cluster analyses from the three different datasets (FT-IR, Raman and MALDI-TOF-MS data) were measured using Procrustes analysis.[45] Procrustes analysis is an excellent approach for assessing the differences and similarities between different ordination space from cluster analyses and has been used previously for the assessment of different analytical techniques.[46] The distances were calculated based on the averaged PC-DFA scores for the biological replicates.

3. Results and discussion

Table S1† shows all 35 isolates belonging to 12 strains (PFGE-defined 12 types) including: EC04, EC09, EC10, EC13, EC14, EC15, EC19, EC20, UNI 156, UNI 178, UNI 191 and UNI 214. These strains were previously confirmed to belong to *E. faecium* using a VITEK® system and by MALDI-TOF analysis using a Bruker Microflex system (data not shown). The PFGE results (Fig. S2†) were compared to results obtained in this study using FT-IR spectroscopy,[17,30,46–49] Raman spectroscopy[25,30,50,51] and MALDI-TOF-MS.[13,14,16,52–54] We believe that these analytical techniques in combination with chemometrics offer an improvement in the classification of bacteria due to their higher biochemical resolution.

3.1 Classification using FT-IR spectroscopy

In this study, four biological replicates of bacterial isolates were analyzed in four analytical replicates and analysis was performed in three machine runs, resulting in a total of 1680 FT-IR spectra. The three machine replicate measurements were performed in order to evaluate the reproducibility of the FT-IR technique. Typical spectra based on four biological replicates of representatives of 12 strains from *Enterococcus* (EC04, EC09, EC10, EC13, EC14, EC15, EC19, EC20, UNI 156, UNI 178, UNI 191 and UNI 214) are provided in Fig. S3A.† The infrared spectra contain distinct regions that can be used to characterize bacterial samples. These have been well documented previously and include: wavenumbers between 3400–2850 cm^{-1} corresponding to fatty acids, at 1705–1454 cm^{-1} related to amide I and II regions attributed to peptides and proteins, and between 1085–1052 cm^{-1} corresponding to polysaccharides.[19,55,56]

Discrimination between the strains based on visual inspection of the spectra was difficult[17] because these strains are very similar phenotypically. Therefore, in order to develop a classification model to distinguish between bacterial samples based on similarities in the spectral data, multivariate analysis was used to reduce the high dimensionality of the data. First, PC-DFA was applied using 40 principal components (PC) to the 12 strains (*i.e.* 12 classes) and 35 isolates (*i.e.* 35 classes) using the pre-processed FT-IR spectra (Fig. 2A and 3A, respectively). Fig. 2A shows a clear separation between the 12 strains, displaying 4 main clusters; Cluster 1 is a single-member cluster (SMC) containing only (EC10), Cluster 2 includes (EC20 and UNI 156), Cluster 3 (UNI 191, EC04 and EC15) and Cluster 4 formed a large group and is a combination of (EC13, EC19, EC14, EC09, UNI 214 and UNI 178). Each cluster is represented by a different color in the figure. As described above, HCA was undertaken using spectral data in order to simplify the DFA plot and to illustrate the related strains. Cluster analysis was based on averaged DFA scores (12 classes/strains), using Ward's linkage as shown in Fig. 2B. Clusters seen in Fig. 2A are reflected in the HCA dendrogram plot (Fig. 2B).

PC-DFA was subsequently performed for all the 35 isolates and the results are provided in Fig. 3. Clear separation between all 35 isolates was observed despite the fact that there were a much higher number of classes to be separated than the number of strains. For example, clear separation was observed between the two representatives of EC10 (139 and 151). Furthermore, results generated using PFGE correlated well with FT-IR spectroscopic data. For example, the UNI 156 and UNI 178 were seen as unique by both techniques. In addition, the three EC20 isolates (192, 198 and 204) and EC19 isolates (173, 174 and 175) clustered together and were not differentiated using FT-IR spectroscopy, which was also observed in the PFGE results, where the bands were quite similar (Fig. 3B). This implies that the isolates within each of these groups are highly similar to each other phenotypically and genetically. Finally, two more clusters were observed, with one cluster containing all the EC04, EC15 and UNI 191 strains and the remainder of the isolates forming another cluster.

 This journal is © The Royal Society of Chemistry 2016

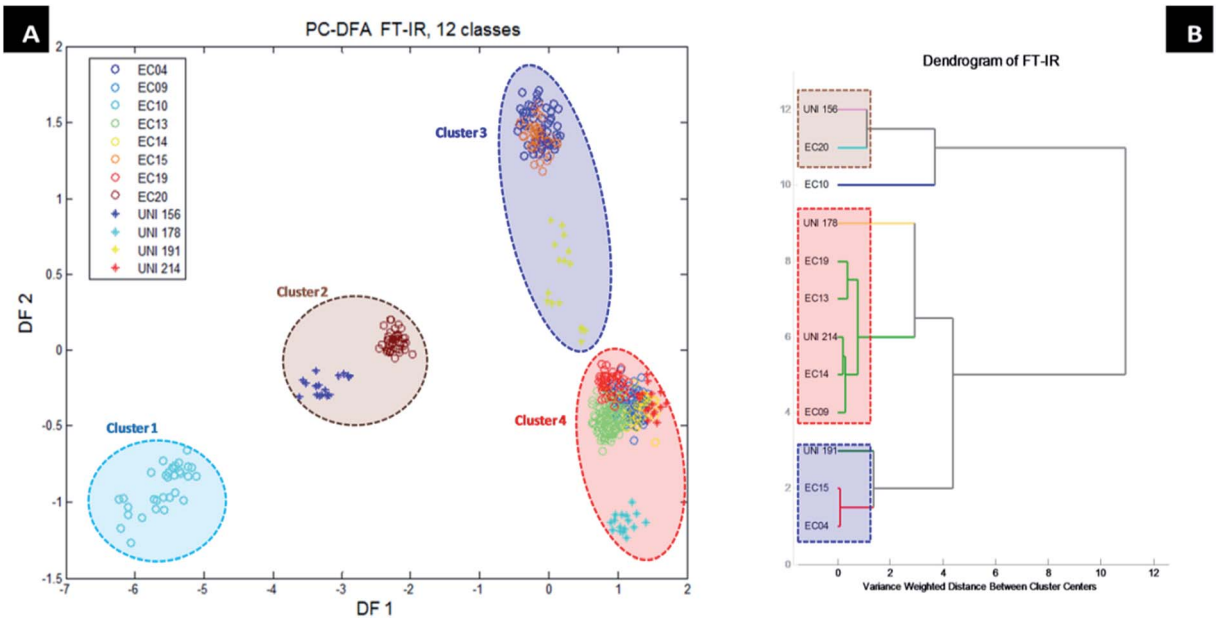

Fig. 2 (A) Discriminant function analysis (DFA) scores plot from FT-IR data after pre-processing, illustrating the relationship between the 12 enterococci. (B) Cluster analysis on averaged PC-DFA scores (12 classes/strains) using Ward's linkage.

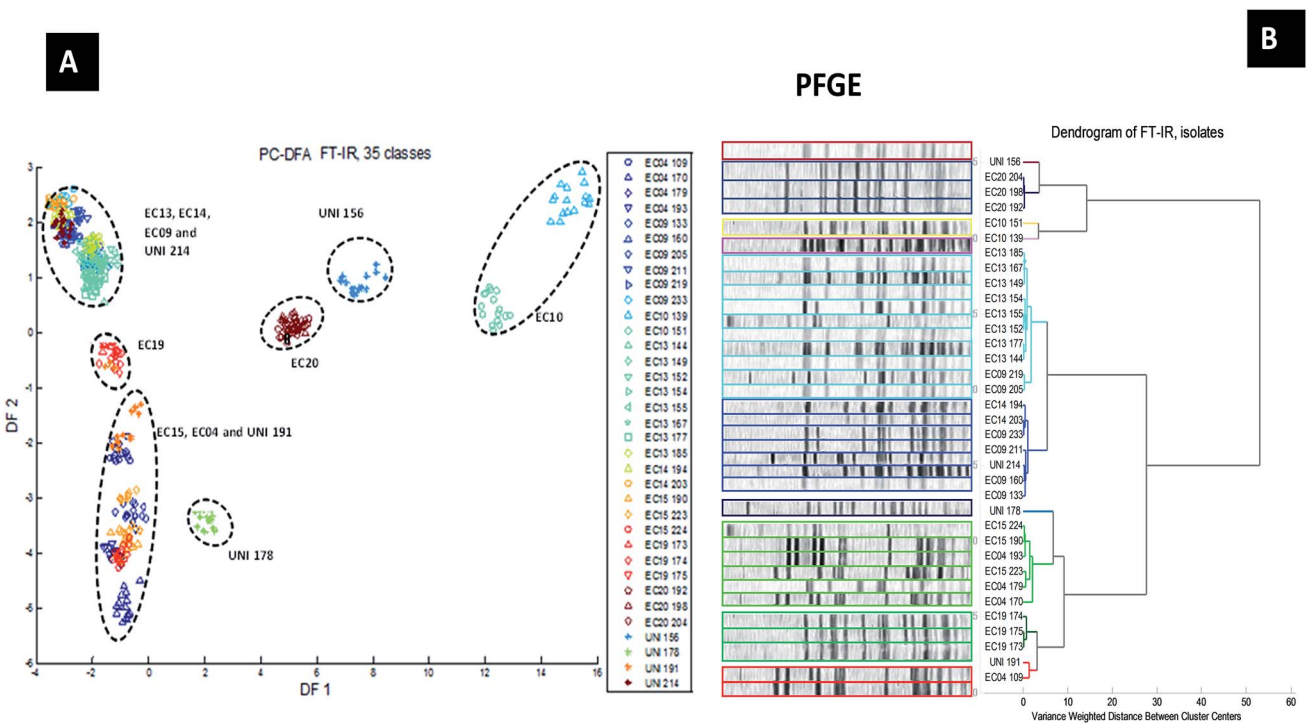

Fig. 3 (A) PC-DFA plot from FT-IR data after pre-processing which illustrates the relationship between the 35 *Enterococcus* isolates. (B) Hierarchical cluster analysis on averaged PC-DFA scores (35 classes/isolates) using Ward's linkage (right) and PFGE results (left). Each isolate is represented by the same color in both the boxes around the PFGE images and the FT-IR dendrogram.

The PLS-DA classification using FT-IR spectral data achieved an average correct classification rate (CCR) of 89.4% at the strain level and 54.3% at the isolate level, both with an empirical p-value of <0.001, *i.e.* not a single case where the null model obtained better results, indicating that the predictive accuracies were highly significant. The null distributions are provided in Fig. S4A and B† at the two different levels.

The confusion matrices of strains and isolates classification are presented in Tables 1 and S3,† respectively. Most of the 12 strains showed high prediction accuracies, for example EC04,

Table 1 The prediction accuracies of the 12 enterococci strains using FT-IR spectroscopy data

Class known/predicted	EC04	EC09	EC10	EC13	EC14	EC15	EC19	EC20	UNI 156	UNI 178	UNI 191	UNI 214
EC04	89.9%	0.5%	0.0%	0.0%	0.4%	8.3%	0.1%	0.0%	0.0%	0.0%	0.7%	0.1%
EC09	0.1%	90.3%	0.0%	1.3%	4.8%	0.0%	3.5%	0.0%	0.0%	0.0%	0.0%	0.0%
EC10	0.0%	0.1%	99.7%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.1%
EC13	0.0%	0.0%	0.0%	99.8%	0.0%	0.0%	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%
EC14	0.1%	48.9%	0.0%	1.1%	47.3%	1.0%	1.4%	0.1%	0.0%	0.0%	0.1%	0.0%
EC15	6.8%	1.4%	0.0%	0.0%	0.5%	91.1%	0.0%	0.0%	0.0%	0.0%	0.1%	0.0%
EC19	1.6%	9.3%	0.0%	0.2%	3.6%	0.0%	83.5%	0.0%	0.0%	0.0%	0.0%	1.8%
EC20	0.0%	0.1%	0.0%	0.0%	0.0%	0.7%	0.0%	99.2%	0.0%	0.0%	0.0%	0.0%
UNI 156	0.4%	0.0%	0.0%	0.5%	0.0%	0.0%	0.1%	0.9%	98.1%	0.0%	0.0%	0.0%
UNI 178	0.0%	5.3%	0.0%	0.1%	0.0%	0.0%	0.4%	0.0%	0.0%	93.9%	0.2%	0.0%
UNI 191	6.5%	0.9%	0.0%	25.2%	0.0%	1.3%	0.0%	0.0%	0.0%	0.0%	66.1%	0.1%
UNI 214	1.9%	13.4%	0.0%	1.0%	0.1%	0.0%	20.4%	0.0%	0.0%	0.0%	4.2%	58.9%

EC10, EC13 and EC20 had accuracies of 89.9%, 99.7%, 99.8% and 99.2%, respectively. However, EC14 and UNI 214 had lower prediction accuracies of 47.3% and 58.9%, respectively. The confusion matrix showed that there was a certain level of overlap between (EC14 and EC09) and (UNI 214 and EC19).

Furthermore, in-depth analysis of the confusion matrix (Fig. 4) showed that classification of unique strains was generally in line with PFGE results. In Fig. 4, high percentage class membership assignments are represented by warm colors (*e.g.* red), indicating agreement between predicted classes and known classes. It is also interesting to see that representatives from EC19 and EC20 formed two "squares" of "tiles" on the diagonal line, in which the colors were similar to each other. Results from Fig. 4 suggest that the PLS-DA model was not able to differentiate the isolates within EC19 and EC20, yet another observation that is consistent with PFGE results. On the other hand, all representatives of EC04 and EC09 (160 and 133) were unique in the FT-IR spectroscopy profile using the PLS-DA model but had visually similar PFGE profiles (Fig. S2†). This is most likely due to PFGE providing genetic information[57,58] while FT-IR spectroscopy describes phenotypes.[27,59] This implies that isolates from EC19 and EC20 are highly conserved phenotypically, whereas those from EC04 and EC09 are not, and such subtle differences in phenotypes were detected by FT-IR spectroscopy. Our observations showed that FT-IR spectroscopy appears to be a very promising analytical approach for discrimination of enterococci at different levels. In line with the results presented in this study, work carried out by Guibet *et al.*

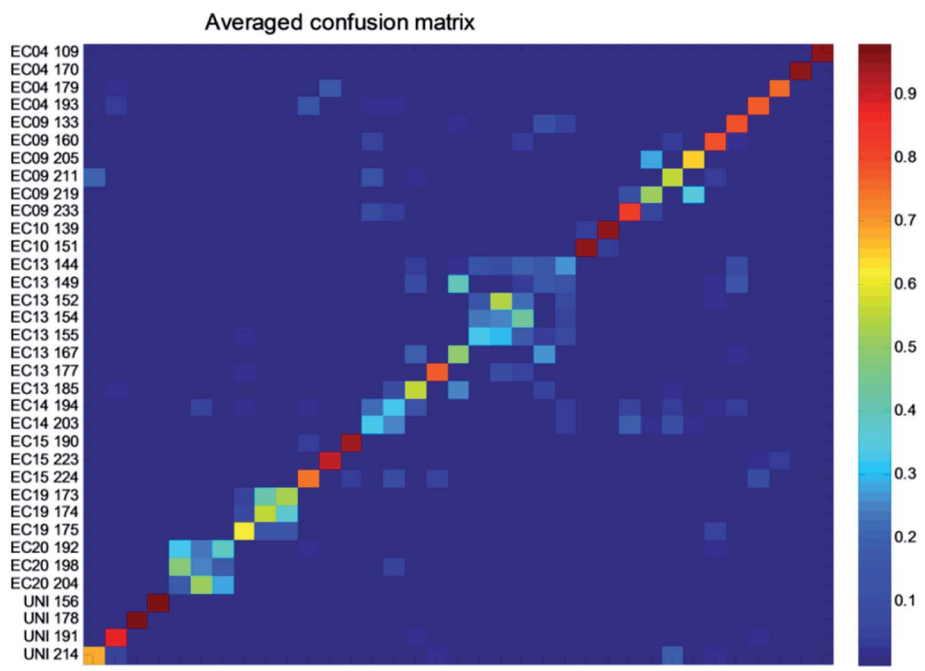

Fig. 4 PLS-DA trained on 35 classes (*i.e.* 35 isolates) from FT-IR spectral data. High percentage class membership assignments are represented by warm colors (*e.g.* red) whilst the cold colors (*e.g.* blue) represent low percentage class membership assignments. The diagonal "tiles" are much warmer than off-diagonal "tiles", which indicates agreement between predicted classes and known classes.

This journal is © The Royal Society of Chemistry 2016

showed that clear discrimination and classification of enterococci strains can be achieved using FT-IR spectroscopy.[60,61]

3.2 Classification using Raman spectroscopy

In addition to the FT-IR spectroscopy technique used in this study, Raman spectroscopy was used as a complementary vibrational spectroscopic technique.[17,61–63] As expected, the two techniques generated different spectra. These two approaches are complementary due to the selection rules, whereby infrared causes a change in the net dipole moment in a particular functional group, induced by molecular vibrations, whereas Raman causes a change in the polarization of bonds within a molecule. Therefore, bonds within a molecule are generally infrared or Raman active with the result being that the two techniques can provide complementary (bio) chemical information.[29,64]

Raman spectra of the 12 *E. faecium* strains are shown in Fig. S3B.† Raman spectra for these types appeared almost indistinguishable and no differences were detected on visual inspection. Moreover, some specific peaks which were identified in these spectra included: peaks at around 722 cm^{-1}, 783 cm^{-1}, 854 cm^{-1}, 1004 cm^{-1}, 1098 cm^{-1}, 1334 cm^{-1}, 1451 cm^{-1} and 1664 cm^{-1}, which correspond to adenine, cytosine/uracil, tyrosine, phenylalanine, phosphate, guanine, protein and amide I, respectively.[65–67]

PC-DFA scores plot of pre-processed Raman spectra for the 12 PFGE-defined types is shown in Fig. S5A.† The figure shows classification results similar to those seen with FT-IR data. There was an obvious overlap between the two spectroscopic techniques, especially with representatives of EC10. However, EC20 overlapped with UNI 156 in FT-IR data, whereas EC20 was closer to UNI 178 based on Raman data. These observations can be seen in the HCA dendrogram based on Raman data (Fig. S5B†), which was quite similar to the HCA results generated from FT-IR data. Looking back at the dendrogram in Fig. S2† based on PFGE data, visual inspection showed that there were some similarities between results generated *via* spectroscopic techniques and those based on PFGE; for example, EC04 and EC15 were shown to overlap in both sets of results (Fig. S2†).

As with FT-IR data, Raman data on the 35 isolates were also analyzed using to PC-DFA and HCA (Fig. S6A and B,† respectively). The results suggested that Raman was also successful in discriminating the two representatives of EC10 (139 and 151), which was also the case using FT-IR analysis (Fig. 3). Furthermore, in order to ensure the classification is robust, the data were visualised and analysed using a heat map based on PLS-DA (Fig. S6C†). The results suggested that all the isolates indicated as unique (UNI) by PFGE were also unique in the PLS-DA model generated using Raman data.

In addition, chemometric-based identification was carried out using PLS-DA at both the strain and isolate levels and the predictive accuracies were calculated based on 1000 bootstrapping resampling using Raman spectral data. The null distribution was obtained (Fig. S4C and D†) at both the strain (12 classes) and isolate levels (35 classes) resulting in average CCR of 69.3% ($p < 0.001$) and 21.1% ($p < 0.001$), respectively. The CCR from FT-IR data was higher at both levels compared to Raman data possibly due to the higher reproducibility of FT-IR data. Confusion matrices were also generated at both the strain level (Table S2A†) and the isolate level (data not shown); these results suggested that Raman can also be used as a robust technique for bacterial discrimination. In-depth analysis showed that Raman generated around 70% prediction accuracy at the strain level which is lower than that of FT-IR (nearly 90%). This is most likely due to the low concentration of cells used for analysis: the infrared interrogation beam used was *ca.* 1 mm and passes completely through the dried bacterial film; while the Raman microscope delivers a highly focussed laser beam with an interrogation volume of ~1 pL and therefore measures very few bacteria. To overcome this limitation with Raman, bacteria can be analyzed directly from the agar plates or surface-enhanced Raman scattering (SERS) as an alternative technique,[68–70] but this is an area for future study.

3.3 Classification using MALDI-TOF-MS

As described in the Materials and Methods section, four biological replicates were analyzed in four analytical replicates for each bacterial strain, resulting in 560 MALDI-TOF-MS spectra; both the biological and technical replicates clustered closely together ensuring good bioanalytical reproducibility (data not shown). The spectra for the 35 enterococci isolates were pre-processed before data analysis. The typical pre-processed positive ion mode MALDI-TOF-MS spectra for all 12 *Enterococcus* strains (EC04, EC09, EC10, EC13, EC14, EC15, EC19, EC20, UNI 156, UNI 178, UNI 191 and UNI 214) are provided in Fig. S3C.† In general, the MALDI-TOF-MS spectra were of high quality with high signal-to-noise ratios in the acquisition m/z range 1000–18 000 and a high number of peaks for each studied strain were detected. There are many factors that can affect MALDI-TOF-MS results and some of these can differ from one lab to another, such as the type of medium used,[71] sample handling, type of matrix,[72] sample deposition method,[73] solvents, instrument settings[74,75] and the type of data analysis chosen.[41,76] These can inadvertently affect MALDI-TOF-MS results and subsequent PC-DFA and HCA.

MALDI-TOF-MS spectra are not readily interpretable from the 35 isolates as they are similar phenotypically and MALDI-TOF-MS spectra show only two dimensions (m/z × intensity). Therefore, as is the case for the vibrational spectroscopy techniques, robust multivariate analysis methods were employed for this purpose. The results of PC-DFA using 12 classes (12 strains) in a three-dimensional plot of DF1 *vs.* DF2 *vs.* DF3 and a two-dimensional plot of DF2 *vs.* DF3 are shown in Fig. S7A and B,† respectively. Four main clusters were observed in the PC-DFA plots; SMC (Cluster) 1 contains only UNI 178; Cluster 2 contains EC20; Cluster 3 consists of EC04, EC10, EC15 and UNI 191; and Cluster 4 formed a large group of (EC13, EC19, EC14, EC09, UNI 214 and UNI 156). Results from the HCA dendrogram (Fig. S7C†) confirmed the separation between the 12 classes (*i.e.* 12 strains). This indicated that UNI 178 is phenotypically very different from the other strains based on MALDI-TOF-MS data.

PC-DFA was also applied to data from the 35 isolates; the results showed that isolates number 160 and 219 (both from EC09) were very different from the other isolates. Therefore, another PC-DFA was carried out with these two outliers removed and the HCA results are shown in Fig. S8D.† It appears that all representatives of EC20 (204, 198 and 192) overlap with each other, which was also observed in FT-IR and Raman data, with the exception that isolate 192 slightly differed from the other two representatives (204 and 198) in the HCA dendrogram when using Raman data (Fig. S6B†). However, analysis by PFGE showed that isolates 192 and 198 clustered more closely with each other than with isolate 204.

Furthermore, PLS-DA model applied to MALDI-TOF-MS data achieved an average CCR of 78.2% ($p < 0.001$) and 35.7% ($p < 0.001$) for the 12 (strains) and 35 (isolates) classes, respectively. When PLS-DA was undertaken with 33 isolates (with isolates 160 and 219 removed), the average CCR for the isolates increased to 53.95% ($p < 0.001$). The prediction accuracies for the 12 classes are shown in Table S2B† and those for the 35 classes (isolates) are shown in Table S4.† Table S2B† shows that discrimination between most of the strains (12 classes) using MALDI-TOF-MS data achieved high correct classification rates, except for EC14 and UNI 191, which had rather low classification rates. Confusion matrices for the 35 classes and the 33 classes (160 and 219 isolates removed) are shown in Fig. S8A and C,† respectively. From these matrices, it can be seen that all the isolates identified by the reference laboratory as unique (UNI), which included isolates 156, 178, 191 and 214, were also classified as unique based on MALDI-TOF-MS data. Moreover, EC20 and EC19 were assigned the same classification in PFGE typing, and this was in agreement with MALDI-TOF-MS, FT-IR and Raman data. In addition, based on MALDI-TOF-MS data (Fig. S8A and C†), representatives of EC13 (152, 154 and 155) belonged to the same cluster, and isolates 177 from EC13 was significantly different from the remaining EC13 strains; this was also observed in FT-IR and PFGE data. Looking back at Fig. S8C,† it can be seen that all the strains from EC04 were unique in MALDI-TOF-MS and FT-IR profiles when using PLS-DA modelling.

3.4 Procrustes distance test of the three analytical techniques

Analytical techniques such as FT-IR, Raman and MALDI-TOF-MS are currently used in clinical research studies worldwide and many reports have been published showing advantages of using such techniques.[24,54,77,78]

Compared to PCR, the 'gold standard' technique for enterococci identification, no conclusive evidence was identifiable in the literature. Application of PCR to the classification of *E. coli* from five different sources recorded average correct classification of 84% (Seurinck *et al.*, 2003), a level comparable to that obtained with FT-IR spectroscopy in this study (89.4%) albeit with a different microorganism. Kirschner *et al.*[61] demonstrated accurate identification and classification of 18 strains from 6 different species belonging to enterococci using vibrational spectroscopic techniques in combination with chemometrics. This study suggested that FT-IR and Raman spectroscopies can offer potential alternatives to the conventional typing tests, based on PCR, due to their speed and ease of use, demonstrating high consistency between classifications based on FT-IR and Raman methods. Based on comparison with classification using PCR, they advocated the use of FT-IR and Raman techniques due to the limited number of enterococci species that could be analysed by PCR and the requirement for very specific procedures, which makes these analytical techniques more suitable for routine use. Our results are in agreement with their findings; however, Kirschner *et al.*[61] did not report comparative analysis of correct classification rates from these techniques. Oliveira *et al.*[51] showed that Raman spectroscopy, in combination with a chemometric algorithm, can be used to discriminate between seven different colonies of Gram-positive and Gram-negative bacteria. In another previous study, it was also shown that 59 clinical bacterial strains associated with urinary tract infections (UTIs) could be identified using FT-IR and Raman spectroscopy.[17] As an alternative to vibrational spectroscopic techniques, MALDI-TOF-MS is a relatively new technique which has shown very promising results for identification in agreement with methodologies carried out in microbiological laboratories, and therefore has been used for the identification and classification of bacterial species[15,79,80] and is appearing in many clinical microbiology testing laboratories.[54,81,82]

Previous studies have generally focussed on the application of just one or two analytical techniques for the classification of *Enterococcus* spp. However, to generate complementary data and more comprehensive analysis, this study combines three different analytical techniques – FT-IR, Raman and MALDI-TOF-MS – to analyze whole bacterial cells. Successful classification was demonstrated at the strain (*i.e.* 12 classes) and isolate (*i.e.* 35 classes) levels based on data generated by the three analytical platforms. In order to assess the overall information content in the spectra that has been revealed by the cluster analysis from the scores plots, Procrustes analysis was employed to assess the overall similarity between the patterns detected by these three platforms. The results are presented in terms of Procrustes distance (Table 2A and B), where the Procrustes distance varies from 0 to 1; the lower the distance, the higher the similarity between the results. The comparisons were made using averaged PC-DFA scores. For each dataset, there were two sets of PC-DFA scores, one at the strain level (12 classes) and another for isolates classification (35 classes). For each set of PC-DFA scores, the scores were then averaged according to their strain label and isolate label to give two sets of *averaged* PC-DFA scores.

The findings in Table 2 can be summarized as follows:

(i) The patterns in the PC-DFA scores at strain and isolate levels were highly similar to each other for all the three analytical platforms. The Procrustes distances varied from 0.0681 to 0.1812. This suggested that the variation originating from different bacteria is the main factor in PC-DFA, *i.e.* the differences between different strains were significantly higher than those between different isolates.

 This journal is © The Royal Society of Chemistry 2016

Table 2 The similarity between three different datasets using Procrustes distance[a]

(A) PC-DFA at the strain level

Averaging on ST level	FT-IR (IS)	FT-IR (ST)	Raman (IS)	Raman (ST)	MALDI (IS)	MALDI (ST)
FT-IR (IS)	—					
FT-IR (ST)	0.0858	—				
Raman (IS)	0.2125	0.2933	—			
Raman (ST)	0.2314	0.3187	0.1502	—		
MALDI (IS)	0.8602	0.889	0.899	0.8202	—	
MALDI (ST)	0.9125	0.8846	0.9149	0.8988	0.1812	—

(B) PC-DFA at the isolate level

Averaging on IS level	FT-IR (IS)	FT-IR (ST)	Raman (IS)	Raman (ST)	MALDI (IS)	MALDI (ST)
FT-IR (IS)	—					
FT-IR (ST)	0.1085	—				
Raman (IS)	0.2112	0.2446	—			
Raman (ST)	0.2411	0.3168	0.1132	—		
MALDI (IS)	0.8593	0.8719	0.8196	0.8001	—	
MALDI (ST)	0.8975	0.8608	0.8841	0.8703	0.0681	—

[a] (ST) and (IS) indicate the PC-DFA was calculated at the strain (12 classes, PFGE-defined 12 types) and isolate (33 classes) levels, respectively.

(ii) The two vibrational spectroscopic techniques (FT-IR and Raman) generated highly similar results both at the strain and isolate classification levels, with the corresponding Procrustes distances varying from 0.2112 to 0.3187.

(iii) However, the results generated by MALDI-TOF-MS were significantly different from those generated by the two spectroscopic techniques, and the corresponding Procrustes distances were all above 0.8. Such differences can be mainly attributed to data from isolate UNI 178, which appeared to be very different to other strains in the MALDI-TOF-MS dataset.

Table S5† shows a summative comparison of the 4 main clusters identified based on the three analytical techniques using PC-DFA plots of the 12 *E. faecium* strains (12 classes). It can be seen from this table that despite the large Procrustes distances between data generated by MALDI-TOF-MS and those generated by the other two techniques, the main identified clusters patterns observed in all three datasets were still largely consistent.

4. Conclusions

The results obtained from the three analytical techniques (mass spectrometry and two vibrational spectroscopies) demonstrated that good discrimination between *E. faecium* bacteria can be achieved at both the strain and isolate levels and the detected patterns from these techniques were highly similar. However, UNI 178 was observed to be different in MALDI-TOF-MS data, which differed from the two vibrational spectroscopy techniques employed in this study.

The results obtained using these spectroscopic phenotyping approaches were mostly consistent with previous results obtained from experiments carried out using the genotypic classification method, PFGE. Some of the results differed when directly comparing our analytical approach with results from the molecular approach and these differences may be due to comparing phenotypic differences from whole-organism fingerprinting with genotypic differences using PFGE.

In conclusion, we have assessed multiple analytical phenotypic as complementary approaches to current molecular methods. All methods provided excellent clustering which was in general agreement with genotypic baseline methods, as well as allowing excellent discrimination to the strain level and good resolution at the sub-strain level. We believe that these three different physicochemical techniques have excellent potential as high-throughput point-of-care screening tools, and for the rapid and reproducible classification of clinically relevant bacteria, such as *E. faecium*. Further method development may be required to optimise these methods for reliable analysis of bacterial mixtures.

Acknowledgements

NM thanks The Saudi Ministry of Higher Education and Princess Nora bint Abdul Rahman University for funding. YX thanks Cancer Research UK (including Experimental Cancer Medicine Centre award) and Wolfson Foundation, DIE and RG thank BBSRC (BB/L014823/1) for support for Raman spectroscopy.

References

1 K. H. Schleifer and R. Kilpper-Bälz, *Int. J. Syst. Evol. Microbiol.*, 1984, **34**, 31–34.
2 F. H. Kayser, K. A. Bienz and J. Eckert, *Medical microbiology*, Georg Thieme Verlag, 2011.

3 J. R. Hayes, L. L. English, P. J. Carter, T. Proescholdt, K. Y. Lee, D. D. Wagner and D. G. White, *Appl. Environ. Microbiol.*, 2003, **69**, 7153–7160.
4 C. M. Franz, M. E. Stiles, K. H. Schleifer and W. H. Holzapfel, *Int. J. Food Microbiol.*, 2003, **88**, 105–122.
5 M. McCracken, A. Wong, R. Mitchell, D. Gravel, J. Conly, J. Embil, L. Johnston, A. Matlow, D. Ormiston and A. Simor, *J. Antimicrob. Chemother.*, 2013, **68**, 1505–1509.
6 N. Woodford, *J. Med. Microbiol.*, 1998, **47**, 849–862.
7 C. A. Arias and B. E. Murray, *Nat. Rev. Microbiol.*, 2012, **10**, 266–278.
8 S. Altekruse, M. Cohen and D. Swerdlow, *Emerging Infect. Dis.*, 1997, **3**, 285–294.
9 E. Engvall, *Med. Biol.*, 1977, **55**, 193.
10 R. H. Yolken, *Yale J. Biol. Med.*, 1980, **53**, 85.
11 D. Ke, F. J. Picard, F. Martineau, C. Ménard, P. H. Roy, M. Ouellette and M. G. Bergeron, *J. Clin. Microbiol.*, 1999, **37**, 3497–3503.
12 D. J. Reen, *Basic Protein and Peptide Protocols*, 1994, pp. 461–466.
13 P. Lasch, H. Nattermann, M. Erhard, M. Stämmler, R. Grunow, N. Bannert, B. Appel and D. Naumann, *Anal. Chem.*, 2008, **80**, 2026–2034.
14 M. Quintela-Baluja, K. Böhme, I. C. Fernández-No, S. Morandi, M. E. Alnakip, S. Caamaño-Antelo, J. Barros-Velázquez and P. Calo-Mata, *Electrophoresis*, 2013, **34**, 2240–2250.
15 P. Lasch, C. Fleige, M. Stämmler, F. Layer, U. Nübel, W. Witte and G. Werner, *J. Microbiol. Methods*, 2014, **100**, 58–69.
16 N. AlMasoud, Y. Xu, N. Nicolaou and R. Goodacre, *Anal. Chim. Acta*, 2014, **840**, 49–57.
17 R. Goodacre, R. Burton, N. Kaderbhai, A. M. Woodward, D. B. Kell and P. J. Rooney, *Microbiology*, 1998, **144**, 1157–1170.
18 D. Helm, H. Labischinski, G. Schallehn and D. Naumann, *J. Gen. Microbiol.*, 1991, **137**, 69–79.
19 D. Naumann, D. Helm and H. Labischinski, *Nature*, 1991, **351**, 81–82.
20 C. X. Lu, Y. A. Liu, X. H. Sun, Y. J. Pan and Y. Zhao, *Acta Chim. Sin.*, 2011, **69**, 101–105.
21 Y. Burgula, D. Khali, S. Kim, S. S. Krishnan, M. A. Cousin, J. P. Gore, B. L. Reuhs and L. J. Mauer, *J. Rapid Methods Autom. Microbiol.*, 2007, **15**, 146–175.
22 E. C. López-Díez and R. Goodacre, *Anal. Chem.*, 2004, **76**, 585–591.
23 K. Maquelin, C. Kirschner, L. P. Choo-Smith, N. van den Braak, H. P. Endtz, D. Naumann and G. J. Puppels, *J. Microbiol. Methods*, 2002, **51**, 255–271.
24 M. Beekes, P. Lasch and D. Naumann, *Vet. Microbiol.*, 2007, **123**, 305–319.
25 D. I. Ellis, D. P. Cowcher, L. Ashton, S. O'Hagan and R. Goodacre, *Analyst*, 2013, **138**, 3871–3884.
26 C. S. Gutteridge, L. Valus and H. J. H. Macfie, in *Computer-Assisted Bacterial Systematics*, ed. M. G. J. G. Priest, Academic Press, London, 1985, pp. 369–401.
27 R. Davis and L. Mauer, *Current Research, Technology and Education Topics in Applied Microbiology and Microbial Biotechnology*, 2010, vol. 2, pp. 1582–1594.
28 J. P. Dworzanski, S. V. Deshpande, R. Chen, R. E. Jabbour, A. P. Snyder, C. H. Wick and L. Li, *J. Proteome Res.*, 2006, **5**, 76–87.
29 R. Goodacre, B. S. Radovic and E. Anklam, *Appl. Spectrosc.*, 2002, **56**, 521–527.
30 D. I. Ellis and R. Goodacre, *Analyst*, 2006, **131**, 875–885.
31 A. A. Argyri, R. M. Jarvis, D. Wedge, Y. Xu, E. Z. Panagou, R. Goodacre and G.-J. E. Nychas, *Food Control*, 2013, **29**, 461–470.
32 R. Goodacre, E. M. Timmins, P. J. Rooney, J. J. Rowland and D. B. Kell, *FEMS Microbiol. Lett.*, 1996, **140**, 233–239.
33 N. Woodford, R. Reynolds, J. Turton, F. Scott, A. Sinclair, A. Williams and D. Livermore, *J. Antimicrob. Chemother.*, 2003, **52**, 711–714.
34 S. A. Patel, F. Currie, N. Thakker and R. Goodacre, *Analyst*, 2008, **133**, 1707–1713.
35 C. L. Winder, S. V. Gordon, J. Dale, R. G. Hewinson and R. Goodacre, *Microbiology*, 2006, **152**, 2757–2765.
36 Y. Wang, D. J. Veltkamp and B. R. Kowalski, *Anal. Chem.*, 1991, **63**, 2750–2756.
37 P. H. C. Eilers, *Anal. Chem.*, 2004, **76**, 404–411.
38 R. G. Brereton, *Chemometrics: data analysis for the laboratory and chemical plant*, John Wiley & Sons, 2003.
39 B. F. Manly, *Multivariate statistical methods: a primer*, CRC Press, 2004.
40 G. G. Harrigan, R. H. LaPlante, G. N. Cosma, G. Cockerell, R. Goodacre, J. F. Maddox, J. P. Luyendyk, P. E. Ganey and R. A. Roth, *Toxicol. Lett.*, 2004, **146**, 197–205.
41 P. S. Gromski, H. Muhamadali, D. I. Ellis, Y. Xu, E. Correa, M. L. Turner and R. Goodacre, *Anal. Chim. Acta*, 2015, **879**, 10–23.
42 T. Hastie, R. Tibshirani and J. Friedman, *The Elements of Statistical Learning*, Springer, New York, 2009.
43 M. Barker and W. Rayens, *J. Chemom.*, 2003, **17**, 166–173.
44 B. Efron and R. J. Tibshirani, *An introduction to the bootstrap*, CRC press, 1994.
45 J. C. Gower and G. B. Dijksterhuis, *Procrustes problems*, Oxford University Press, Oxford, 2004.
46 P. R. Peres-Neto and D. A. Jackson, *Oecologia*, 2001, **129**, 169–178.
47 H. AlRabiah, Y. Xu, N. J. Rattray, A. A. Vaughan, T. Gibreel, A. Sayqal, M. Upton, J. W. Allwood and R. Goodacre, *Analyst*, 2014, **139**, 4193–4199.
48 D. Naumann, *Infrared Phys.*, 1984, **24**, 233–238.
49 D. Naumann, V. Fijala, H. Labischinski and P. Giesbrecht, *J. Mol. Struct.*, 1988, **174**, 165–170.
50 L. Mariey, J. P. Signolle, C. Amiel and J. Travert, *Vib. Spectrosc.*, 2001, **26**, 151–159.
51 F. S. de Siqueira e Oliveira, H. E. Giana and L. Silveira, Jr., *J. Biomed. Opt.*, 2012, **17**, 107004.
52 H. Muhamadali, M. Chisanga, A. Subaihi and R. Goodacre, *Anal. Chem.*, 2015, **87**, 4578–4586.
53 C. Benagli, V. Rossi, M. Dolina, M. Tonolla and O. Petrini, *PLoS One*, 2011, **6**, e16424.

This journal is © The Royal Society of Chemistry 2016

54 R. Cramer, J. Gobom and E. Nordhoff, *Expert Rev. Proteomics*, 2005, **2**, 407–420.
55 D. I. Ellis, W. B. Dunn, J. L. Griffin, J. W. Allwood and R. Goodacre, *Analyst*, 2007, **131**, 875–885.
56 E. Carbonnelle, C. Mesquita, E. Bille, N. Day, B. Dauphin, J.-L. Beretti, A. Ferroni, L. Gutmann and X. Nassif, *Clin. Biochem.*, 2011, **44**, 104–109.
57 S. Kim, B. L. Reuhs and L. J. Mauer, *J. Appl. Microbiol.*, 2005, **99**, 411–417.
58 D. I. Ellis, G. G. Harrigan and R. Goodacre, in *Metabolic Profiling: its role in biomarker discovery and gene function analysis*, Springer, 2003, pp. 111–124.
59 D. Turabelidze, M. Kotetishvili, A. Kreger, J. G. Morris and A. Sulakvelidze, *J. Clin. Microbiol.*, 2000, **38**, 4242–4245.
60 F. Guibet, C. Amiel, P. Cadot, C. Cordevant, M. H. Desmonts, M. Lange, A. Marecat, J. Travert, C. Denis and L. Mariey, *Vib. Spectrosc.*, 2003, **33**, 133–142.
61 C. Kirschner, K. Maquelin, P. Pina, N. N. Thi, L.-P. Choo-Smith, G. Sockalingum, C. Sandt, D. Ami, F. Orsini and S. Doglia, *J. Clin. Microbiol.*, 2001, **39**, 1763–1770.
62 H. Muhamadali, A. Subaih, M. Mohammadtaheri, Y. Xu, D. Ellis, R. Ramanathan, V. Bansal and R. Goodacre, *Analyst*, 2016, **17**, 5127–5136.
63 D. Schiff, H. Aviv, E. Rosenbaum and Y. R. Tischler, *Anal. Chem.*, 2016, **88**, 2164–2169.
64 J. van de Vossenberg, H. Tervahauta, K. Maquelin, C. H. Blokker-Koopmans, M. Uytewaal-Aarts, D. van der Kooij, A. P. van Wezel and B. van der Gaag, *Anal. Methods*, 2013, **5**, 2679–2687.
65 N. Colthup, *Introduction to infrared and Raman spectroscopy*, Elsevier, 2012.
66 N. Uzunbajakava, A. Lenferink, Y. Kraan, B. Willekens, G. Vrensen, J. Greve and C. Otto, *Biopolymers*, 2003, **72**, 1–9.
67 K. Maquelin, C. Kirschner, L.-P. Choo-Smith, N. van den Braak, H. P. Endtz, D. Naumann and G. Puppels, *J. Microbiol. Methods*, 2002, **51**, 255–271.
68 W. E. Huang, M. Li, R. M. Jarvis, R. Goodacre and S. A. Banwart, in *Advances in Applied Microbiology*, ed. I. L. Allen, S. Sima and M. G. Geoffrey, Academic Press, 2010, vol. 70, pp. 153–186.
69 T. M. Cotton, J. H. Kim and G. D. Chumanov, *J. Raman Spectrosc.*, 1991, **22**, 729–742.
70 I. Nabiev, I. Chourpa and M. Manfait, *J. Raman Spectrosc.*, 1994, **25**, 13–23.
71 R. M. Jarvis and R. Goodacre, *Chem. Soc. Rev.*, 2008, **37**, 931–936.
72 X. Shu, Y. Li, M. Liang, B. Yang, C. Liu, Y. Wang and J. Shu, *Int. J. Mass Spectrom.*, 2012, **321–322**, 71–76.
73 R. Giebel, C. Worden, S. Rust, G. Kleinheinz, M. Robbins and T. Sandrin, *Adv. Appl. Microbiol.*, 2010, **71**, 149–184.
74 K. Dreisewerd, *Chem. Rev.*, 2003, **103**, 395–426.
75 T. L. Williams, D. Andrzejewski, J. O. Lay and S. M. Musser, *J. Am. Soc. Mass Spectrom.*, 2003, **14**, 342–351.
76 A. Freiwald and S. Sauer, *Nat. Protoc.*, 2009, **4**, 732–742.
77 P. S. Gromski, Y. Xu, E. Correa, D. I. Ellis, M. L. Turner and R. Goodacre, *Anal. Chim. Acta*, 2014, **829**, 1–8.
78 E. De Carolis, B. Posteraro, C. Lass-Flörl, A. Vella, A. R. Florio, R. Torelli, C. Girmenia, C. Colozza, A. M. Tortorano, M. Sanguinetti and G. Fadda, *Clin. Microbiol. Infect.*, 2012, **18**, 475–484.
79 M. Risch, D. Radjenovic, J. N. Han, M. Wydler, U. Nydegger and L. Risch, *Swiss Med. Wkly.*, 2010, **140**, w13095.
80 C. Benagli, V. Rossi, M. Dolina, M. Tonolla and O. Petrini, *PLoS One*, 2011, **6**, e16424.
81 T. C. Dingle and S. M. Butler-Wu, *Clin. Lab. Med.*, 2013, **33**, 589–609.
82 S. Sauer and M. Kliem, *Nat. Rev. Microbiol.*, 2010, **8**, 74–82.