Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy 189 (2018) 282–290

Contents lists available at ScienceDirect

Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy

journal homepage: www.elsevier.com/locate/saa

Rapid classification of heavy metal-exposed freshwater bacteria by infrared spectroscopy coupled with chemometrics using supervised method

Rafig Gurbanov [a,b], Ayse Gul Gozen [c], Feride Severcan [c,d,*]

[a] Department of Molecular Biology and Genetics, Bilecik S.E. University, 11230 Bilecik, Turkey
[b] Department of Biochemistry, Middle East Technical University, 06800 Ankara, Turkey
[c] Department of Biological Sciences, Middle East Technical University, 06800 Ankara, Turkey
[d] Department of Biophysics, Faculty of Medicine, Altinbas University, 34217 Istanbul, Turkey

ARTICLE INFO

Article history:
Received 2 June 2017
Received in revised form 8 August 2017
Accepted 13 August 2017
Available online 15 August 2017

Keywords:
Bacteria classification
Chemometric methods
Soft independent modeling of class analogy (SIMCA)
IR spectroscopy

ABSTRACT

Rapid, cost-effective, sensitive and accurate methodologies to classify bacteria are still in the process of development. The major drawbacks of standard microbiological, molecular and immunological techniques call for the possible usage of infrared (IR) spectroscopy based supervised chemometric techniques. Previous applications of IR based chemometric methods have demonstrated outstanding findings in the classification of bacteria. Therefore, we have exploited an IR spectroscopy based chemometrics using supervised method namely Soft Independent Modeling of Class Analogy (SIMCA) technique for the first time to classify heavy metal-exposed bacteria to be used in the selection of suitable bacteria to evaluate their potential for environmental cleanup applications. Herein, we present the powerful differentiation and classification of laboratory strains (*Escherichia coli* and *Staphylococcus aureus*) and environmental isolates (*Gordonia* sp. and *Microbacterium oxydans*) of bacteria exposed to growth inhibitory concentrations of silver (Ag), cadmium (Cd) and lead (Pb). Our results demonstrated that SIMCA was able to differentiate all heavy metal-exposed and control groups from each other with 95% confidence level. Correct identification of randomly chosen test samples in their corresponding groups and high model distances between the classes were also achieved. We report, for the first time, the success of IR spectroscopy coupled with supervised chemometric technique SIMCA in classification of different bacteria under a given treatment.

1. Introduction

Various industrial activities spoil the quality of natural freshwaters by releasing different heavy metals [1]. Overwhelming metal contamination poses great danger to public health since they generally cannot be transformed into the safe forms in nature [2]. To prevent these dangers, the contaminants should be removed efficiently. In many studies, bacteria species have been employed for the decontamination of environmental metal pollutants since in their natural habitat, they adapt themselves to survive and grow under the toxic effects of pollutants including silver (Ag), cadmium (Cd) and lead (Pb) [3]. However, until now none of these studies recommends selection strategies for bacteria to be used in general bioremediation applications [4]. The major drawback for biological decontamination is the efficient and accurate selection, identification and characterization of the appropriate microorganisms among numerous species or strains. In addition, low decontamination capacity of some microorganisms due to the difficulties during the adaptation is another disadvantage [2]. A rapid and an efficient selection of specific microbial strains exhibiting strong heavy metal adaptation ability appear to be the key point for the removal and/or degradation of target contaminants from the ecosystem [5]. Therefore, approaches for screening, characterization and classification of the bacteria are essential. To achieve these processes, the bulk chemical data obtained from analyses of microorganisms should be described and interpreted in a qualitative and quantitative manner [6]. To identify, discriminate and classify bacteria, traditional culture-based, brand-new molecular (PCR based) and immunological techniques have been employed as standard procedures in microbiology [7–9]. However, these techniques do not provide quick and routine applications especially for the fieldwork. In this context, advances in spectroscopy-based fast, easy to use and low-cost techniques provide an opportunity [7,9]. Infrared (IR) spectroscopy elucidates chemical properties of the biomolecules in detail resulting in establishing chemical fingerprint. It is a rapid, versatile, reliable, accurate and automated technique. Low cost and reduced sample preparation steps make this technique a very convenient tool for analytical measurements. Abundant data can be

* Corresponding author at: Department of Biological Sciences, Middle East Technical University, 06800 Ankara, Turkey.
E-mail address: feride@metu.edu.tr (F. Severcan).

http://dx.doi.org/10.1016/j.saa.2017.08.038
1386-1425/

obtained from the specimen through the interpretation of its IR spectrum. Furthermore, definite wavenumber detections, superior signal-to-noise ratios and multiplex advantage, put IR spectroscopy in a favorable position among other analytical techniques [10].

As indicated by Naumann in 2006, it is possible to detect, enumerate, classify and identify microorganisms in one single IR apparatus and interpret the outcomes of the experiment in 24 h followed by their isolation and growth [11]. The technique is exceptionally specific to differentiate the strains, species and genus compared to common identification methods. In addition, it is used in process controls in biotechnological operations and to diagnose microbial accumulations in different industrial facilities [11]. Furthermore, IR spectroscopy generates qualitative and quantitative information about the composition, function and structure of the molecules which is unique for the studied system [12,13]. Due to these advantages, this technique has been extensively used in the determination of bacterial components [14–17]. Moreover, it is a fast, sensitive, easy to use, nondestructive, relatively inexpensive and comprehensive method [18]. Along the same line, we have previously determined gross biomolecular alterations happening in cobalt-acclimated as well as Cd and Pb-adapted bacteria by taking advantage of qualitative and quantitative molecular discrimination power of IR spectroscopy that measures the vibrations of atoms, enabling the identification of functional groups of molecules [15,16].

Since the IR spectra of microorganisms contain complex chemical fingerprint information, their analyses and interpretation require chemometric tools. This is necessary to establish an adequate generality in the accurate and rapid classification of the microbes. In other words, proper and reliable data mining assays should be selected for the efficient classification and discrimination of microorganisms [19]. Chemometric methods use multivariate analysis tools and mathematical modeling to the analytical data [20]. The meaningful information can be gathered quickly and efficiently from the complex and large chemical data sets by using these techniques [21]. Although the previous studies showed the applicability of IR based unsupervised and supervised methods for the discrimination of bacteria [11,15,22–25], supervised ones have not been applied yet to select and differentiate the heavy metal-exposed bacteria. In our previous work, although, we have used unsupervised hierarchical cluster analysis (HCA) and principal component analysis (PCA) for differentiation of the Cd and Pb-exposed bacteria, the development of powerful supervised chemometric models with SIMCA and their thorough testing are necessary to classify the heavy metal-exposed bacterial classes properly with high assurance. Therefore, in the current study, for the first time, we aimed to develop an accurate and rapid differentiation and most importantly classification method for Ag, Cd and Pb-exposed laboratory strains and bacterial isolates from a freshwater environment. In this study, bacterial strains such as *Escherichia coli*, *Staphylococcus aureus*, *Gordonia* sp. FS18 and *Microbacterium oxydans* FS45 [26] were tested because of their capability to survive and grow at growth inhibitory concentrations of Ag, Cd and Pb. In this context, we first performed unsupervised chemometric tool- PCA, which is a mandatory step, and then, we performed SIMCA on IR spectral data of the bacteria.

The identification and classification of heavy metal-exposed bacteria would help to determine the most appropriate strains with sufficient biosorption or transformation capacities and eliminate the inapplicable strains in accurate, fast and productive manner. Thereupon, our approach would contribute for the establishment of sustainable, green and worthwhile biogeotechnological operations to rehabilitate the contaminated soil and water.

2. Material and Methods

2.1. Bacterial Growth Conditions

Laboratory (*Escherichia coli* ATCC 8739 (*E. coli*) and *Staphylococcus aureus* ATCC 6538 (*S. aureus*)) strains and environmental (freshwater) bacterial isolates (*Gordonia* sp. FS18 and *Microbacterium oxydans* FS45 (*M. oxydans*)) were grown respectively at 37 °C and 28 °C under aerobic conditions, in an orbital shaker at 200 rpm. The freshwater bacteria were previously isolated from Lake Mogan and identified via both microbiological and genetic (16sRNA) techniques by our group [26]. Nutrient broth (5 g peptone from meat and 3 g meat extract per liter, *Merck*, US) and nutrient agar (5 g pancreatic digest of gelatin, 3 g beef extract, and 15 g agar per liter, *Becton Dickinson*, US) were used as culture media. Cadmium chloride and silver nitrate (*Sigma*, US) stock solutions were prepared by dissolving in dH_2O and lead nitrate was dissolved in 1:10 diluted nitric acid (in dH_2O). For the sterilization of the stock solutions, 0.22 μm filters (*Pall*, US) were used. The standard metal solutions were added to the growth media after they were autoclaved and cooled to 45–50 °C. All bacteriological procedures were carried out under sterile conditions, in a laminar flow hood (*Esco*, US).

2.2. Sample Preparation for Attenuated Total Reflectance (ATR)-Fourier Transform Infrared (FTIR) Spectroscopy Measurements

The bacterial concentrations were adjusted as 0.5 OD at 600 nm using sterilized distilled water (*UV-2600/2700*, *Shimadzu*, JP). The bacteria were collected by centrifugation at 10,000 g (*Sigma 1–14 Microfuge*, *SciQuip*, UK) for 10 min. After the supernatant was removed, the pellets were gently resuspended in sterilized distilled water. The sample preparation steps were performed according to the established nondestructive routine laboratory procedures for obtaining the IR spectra of intact bacterial cells [11,13,15,27].

2.3. ATR-FTIR Spectroscopic Measurements and Data Preprocessing

Spectrum 100, FTIR spectrometer (*PerkinElmer*, US) equipped with a Universal ATR accessory, was used to collect the IR spectra of control and heavy metal-exposed bacteria. The spectrum of air was used as a reference. Each sample (5 μl) was placed on a diamond/ZnSe crystal plate (*PerkinElmer*, US) and mildly purged and dried under inert nitrogen (N_2) gas flux for 2 min. Purging samples with noninvasive N_2 was applied in order to remove the overlapping free water bands from the samples (while keeping the bound water in the system) as a common procedure in ATR-FTIR studies. The samples were scanned over the spectral range 4000 to 650 cm^{-1}, at room temperature, 100 times per sample and with a resolution of 4 cm^{-1}. For each bacteria group, 15 independent spectra were collected in triplicate and average spectra of these triplicates were used in all data analyses. Spectrum 100 software (*PerkinElmer*, US) was used for data collection and analyses, which included averaging of spectra within samples, smoothing and baseline corrections. In data preprocessing, the second derivative and vector-normalized IR spectra were obtained using *Opus 5.5* software (*Bruker*, US). Vector normalization is carried out in the following way: spectra are first mean-centered, i.e. the average value of the absorbances is calculated for the spectral region indicated. This value is then subtracted from the spectrum. Then, the spectra are scaled such that the sum squared deviation over the indicated wavelengths equals one.

2.4. Chemometric Methods

2.4.1. PCA

PCA is an extensively employed versatile unsupervised pattern-recognition method [28] for the differentiation of microorganisms using large and complex IR spectral data [29]. PCA illustrates the items through the variables generated after the linear integrations of authentic variables. These linear integrations termed principal components are estimated along the course of most extreme change and opposite to one another. In spite of the fact that PCA grants an outlined representation of the information set, its fundamental utility is to show structure in the information [30]. The peculiar scores obtained for each information set can be utilized to group the information in the PC-based coordinate

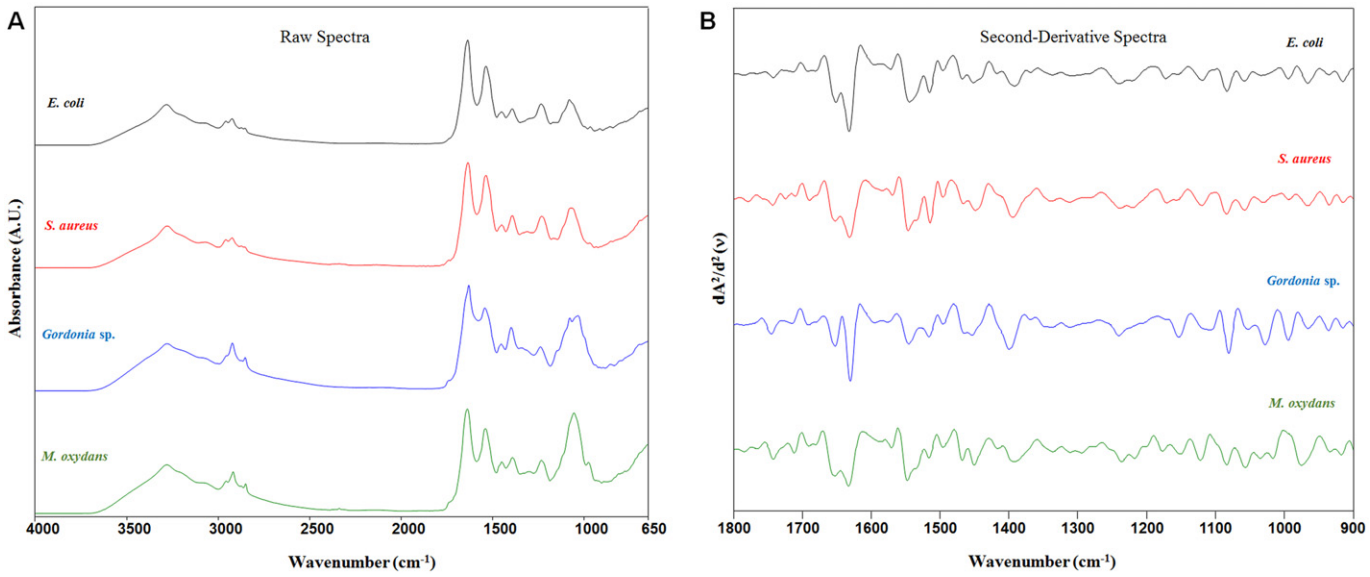

Fig. 1. Representative raw (A) and second derivative and vector-normalized (B) IR spectra of *E. coli* (black), *S. aureus* (red), *M. oxydans* (blue) and *Gordonia* sp. (green) in the whole (4000–650 cm^{-1}) and fingerprint (1800–900 cm^{-1}) spectral regions, respectively.

system, or to relapse back against the pre-established concentration matrix for quantitative assessment. On the other hand, PCA is just equipped for perceiving total variance in regards to an entire information set and not suitable for distinguishing intra-group changes [31]. These scores plots can be utilized to translate the similitudes and contrasts between microscopic organisms (bacteria). The closer the specimens are inside of a scores plot, the more comparable they are regarding the principal component scores assessed [32].

PCA was applied to ATR-FTIR spectral data of studied bacterial groups [33,34]. In this chemometric approach, the spectra of each group (sample size: 15) were imported into *Unscrambler* × 10.3 (*Camo Software*, NO) multivariate analysis (MVA) software. The PCA

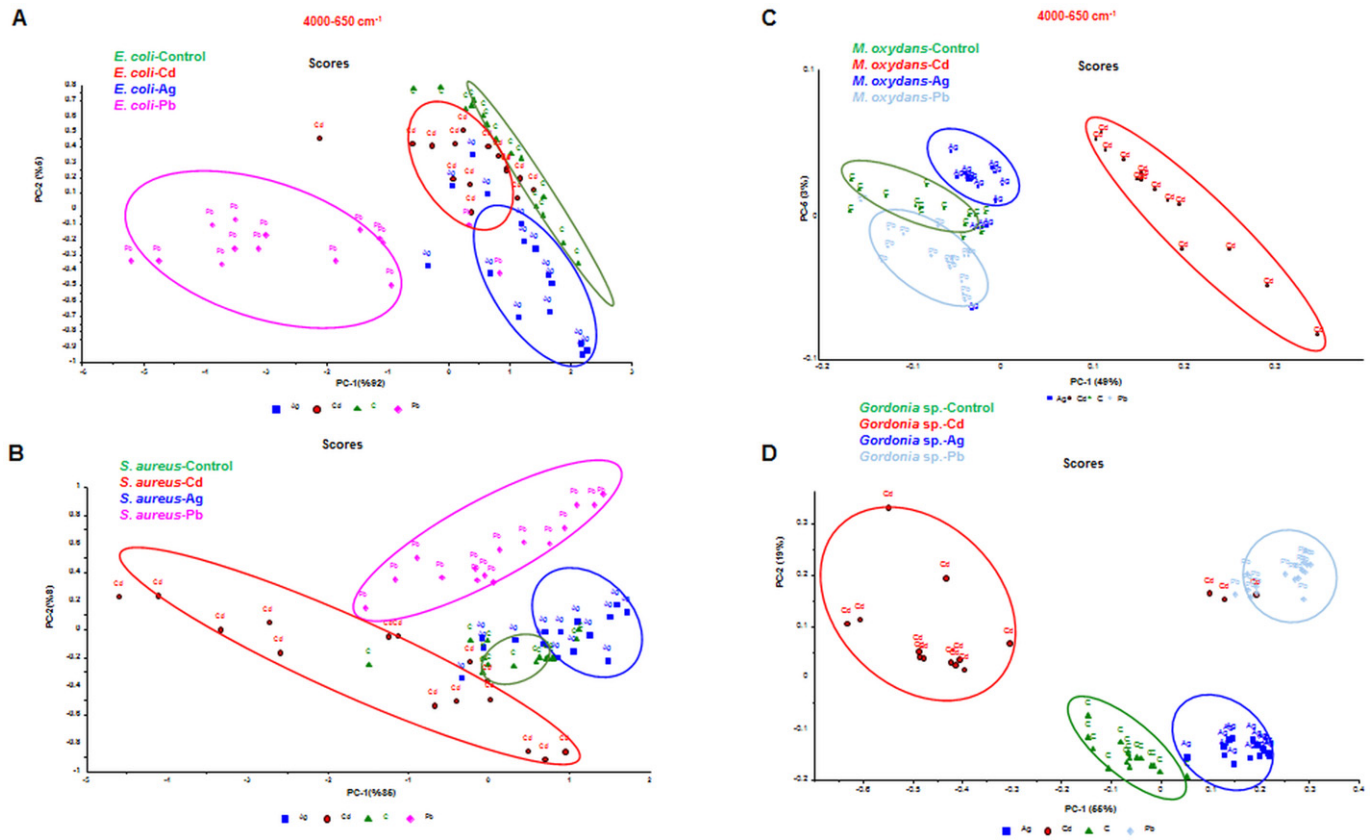

Fig. 2. PCA scores plots for the control and Ag, Cd and Pb-exposed *E. coli* (A), *S. aureus* (B), *M. oxydans* (C) and *Gordonia* sp. (D) samples in the 4000–650 cm^{-1} spectral region.

software was administered to our spectral data in the whole IR region (4000–650 cm^{-1}). The results were presented as scores plots.

2.4.2. SIMCA

Construction of confidence intervals for each experimental group is a principal mission of SIMCA, in which unclassified samples are placed into the PC margins and organized according to the most appropriate class [35]. For that reason, conduction of PCA is a mandatory step for SIMCA classification [36]. To clarify, SIMCA computes PC models for every class of the data set and classifies the unknown samples. Unknowns are identified and discriminated by contrasting the residual variance of the modeled class with the residual variance of the unknown specimen [37–39]. Biggest interval or distance among the classes means the excellent classification for SIMCA [38].

In the present study, supervised pattern recognitions of control and heavy metal-exposed bacterial groups were developed using SIMCA method. SIMCA was built for the whole IR (4000–650 cm^{-1}) region of spectral data via *Unscrambler* × 10. 3 (*Camo Software*, NO). In this supervised classification method, each class of data set was mathematically modeled via PCA as a first step to generate training sets. The number of samples used were 15 for each group and 2 of these samples from each group were randomly chosen and tested as unknowns (shown in green in all Coomans plots). Subsequently, samples were tested for each class set to build an appropriate discrimination model for the controls and heavy metal-exposed bacterial groups [34,38,39]. In the current study, four different bacteria species were modeled independently. Each heavy-metal exposed bacteria class ($n = 15$) was tested versus their untreated control class ($n = 15$).

3. Results

The raw and the second-derivative and vector-normalized representative IR spectra in whole (4000–650 cm^{-1}) and fingerprint (1800–900 cm^{-1}) regions for studied bacterial isolates are shown in Fig. 1A and B, respectively. As can be seen from the figure, the different bacterial species reflect unique and visible distinct IR profile especially in the second-derivative spectra, proving again the sensitivity of IR spectroscopy in microbiological research. In order to build SIMCA models, first, PCA was performed for the control, Cd, Ag and Pb groups of *E. coli, S. aureus, M. oxydans* and *Gordonia* sp. in the whole (4000–650 cm^{-1}) IR regions. The scores plot results for *E. coli and S. aureus* in this region are shown in Fig. 2A and B, respectively. As shown in Fig. 2A, the maximum variation value was quite high for *E. coli* (PC1 + PC2 = 97%). The scores plot showed that Pb and Ag-exposed samples were clearly segregated from the control and Cd-exposed samples. The plot furthermore revealed that the Pb-exposed samples were aggregated more clearly from the control samples in comparison to the Ag-exposed samples. Similarly, the maximum variation value was also high for *S. aureus* (PC1 + PC2 = 93%) (Fig.2B). The scores plot showed that Pb and Cd-exposed samples were more successfully segregated from the control and Ag-exposed samples in this region.

The scores plots of *M. oxydans* and *Gordonia* sp. are shown in Fig. 2C and D, respectively. The maximum variation value for *M. oxydans* was 52% (PC1 + PC5); whilst it was 74% (PC1 + PC2) for *Gordonia* sp. All the groups were clearly differentiated from each other in these environmental bacterial isolates.

PCA shows the common relationship between all variables and differentiates the groups instinctively without prior information. In other

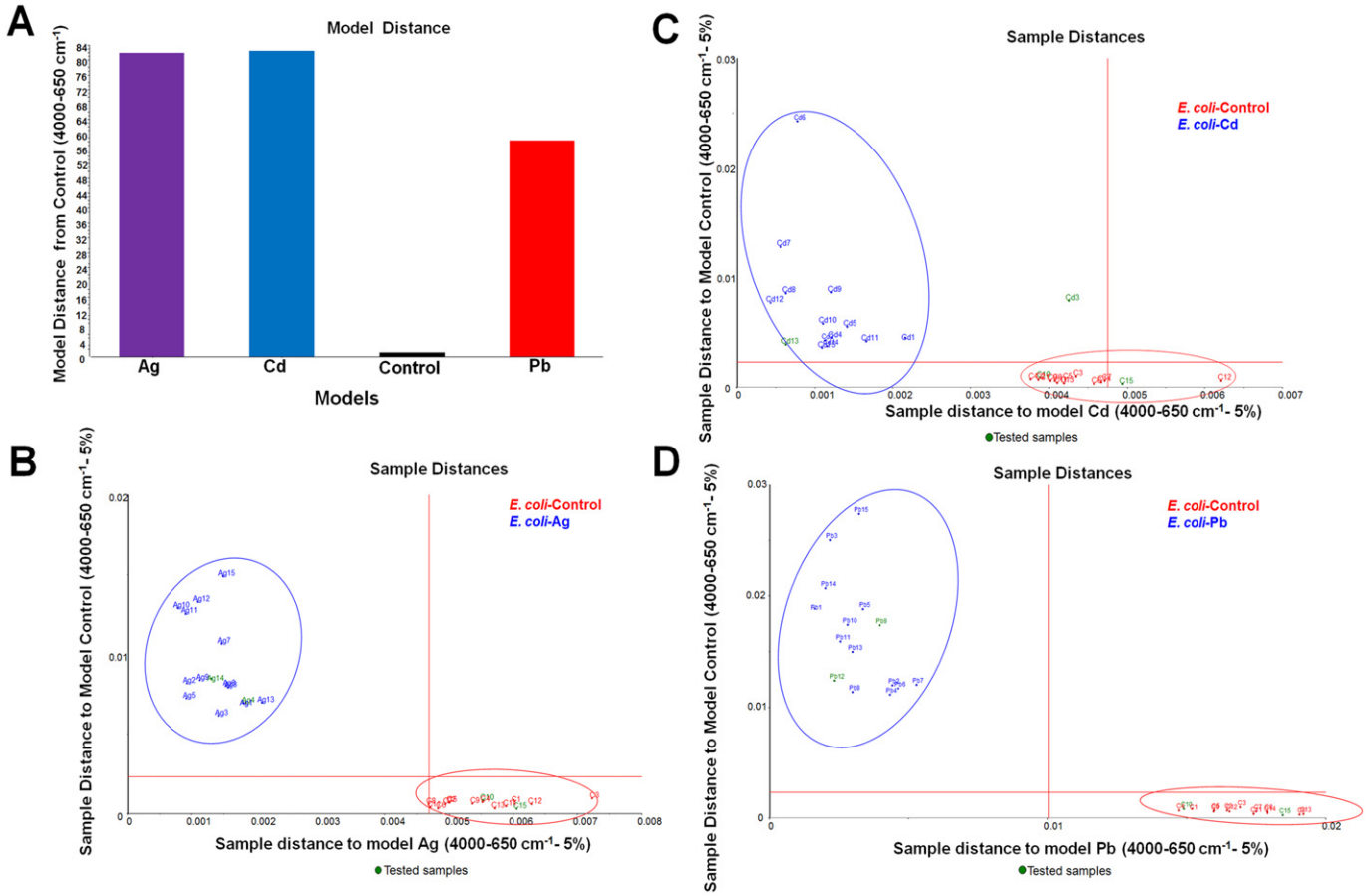

Fig. 3. The SIMCA results for *E. coli*. (A) Model distances between the control, Ag, Cd and Pb-exposed *E. coli*. (B), (C) and (D) Coomans plots for *E. coli* control-Ag, control-Cd and control-Pb models, respectively.

words, variables providing identical information are clustered together. Testing and comparison of the results with correct input models are not applicable in PCA. However, SIMCA confidentially deals with each sample but within each class separately, and reflects the distance between every data point to determine whether or not the sample belongs to the corresponding class that enhances the classification power and accuracy [40]. Therefore, to ascertain the success of the PCA differentiation and to perform classification, we further proceeded SIMCA analysis.

Based on the PCA scores, SIMCA was built for the control and heavy metal-exposed groups of *E. coli, S. aureus, M. oxydans* and *Gordonia* sp. in the 4000–650 cm^{-1} spectral region. The results are summarized in Figs. 3, 4, 5 and 6, respectively.

SIMCA results for *E. coli* are summarized in Fig. 3. The distance in PC space of the Ag, Cd, and Pb calibration models from the control model are shown in Fig. 3A. In general, a distance of 3–4 in the PCA space indicates a clear segregation of the models from each other [41,42]. As shown in the figure, the distances of the Ag and Cd calibration models from the control model lied between 80 and 84, while the distance of the Pb calibration model from the control model lied between 58 and 60. The large distances indicated that there is a large difference between the Ag, Cd and Pb calibration models and the control model. The results from the SIMCA analysis were also presented in a plot called the Coomans plot for *E. coli* control-Ag, control-Cd and control-Pb models, respectively, where distances between two classes were plotted against each other in a scores plot (Fig. 3B, C and D). The Coomans plots also indicated the large distances between the control and heavy metal-exposed classes of *E. coli* with 95% confidence level, supporting calibration model results. Furthermore, the tested samples chosen randomly from each *E. coli* group (shown in green in the Coomans plot) were identified correctly in their corresponding *E. coli* groups.

SIMCA results for *S. aureus* are illustrated in Fig. 4. Similar to *E. coli*, the large distances (from 38 to 77) were obtained between the Ag, Cd, and Pb calibration models and the control model in *S. aureus* (Fig. 4A). The Coomans plots for *S. aureus* control-Ag, control-Cd and control-Pb models are shown in Figs. 4B, C and D, respectively. The big intervals between the control and heavy metal-exposed classes of *S. aureus* with 95% confidence level in the Coomans plots confirmed calibration model findings. Randomly chosen *S. aureus* samples were also correctly identified (shown in green) in their corresponding *S. aureus* groups.

The model distances between the control, Ag, Cd and Pb-exposed environmental *M. oxydans* obtained in SIMCA are shown in Fig. 5A. As shown in the figure, the distance between the control and Cd models was 68, whilst the distances between the control-Ag and control-Pb models were 13 and 6, respectively. This result indicated the good classification between control and heavy metal-exposed *M. oxydans* classes. Fig. 5B, C and D represent the Coomans plots for Ag, Cd and Pb-exposed *M. oxydans*, respectively. The Coomans plots indicated the large distance between the control and Cd-exposed classes of *M. oxydans*. Although, the distances between the control-Ag and control-Pb classes were not as large as for control-Cd class, though adequate, supporting the calibration model results. All of the Coomans plot for *M. oxydans* were obtained at 95% confidence level, again proving good differentiation between control and heavy metal-exposed *M. oxydans*. Randomly chosen *M. oxydans* samples were also correctly identified (shown in green) in their corresponding *M. oxydans* groups.

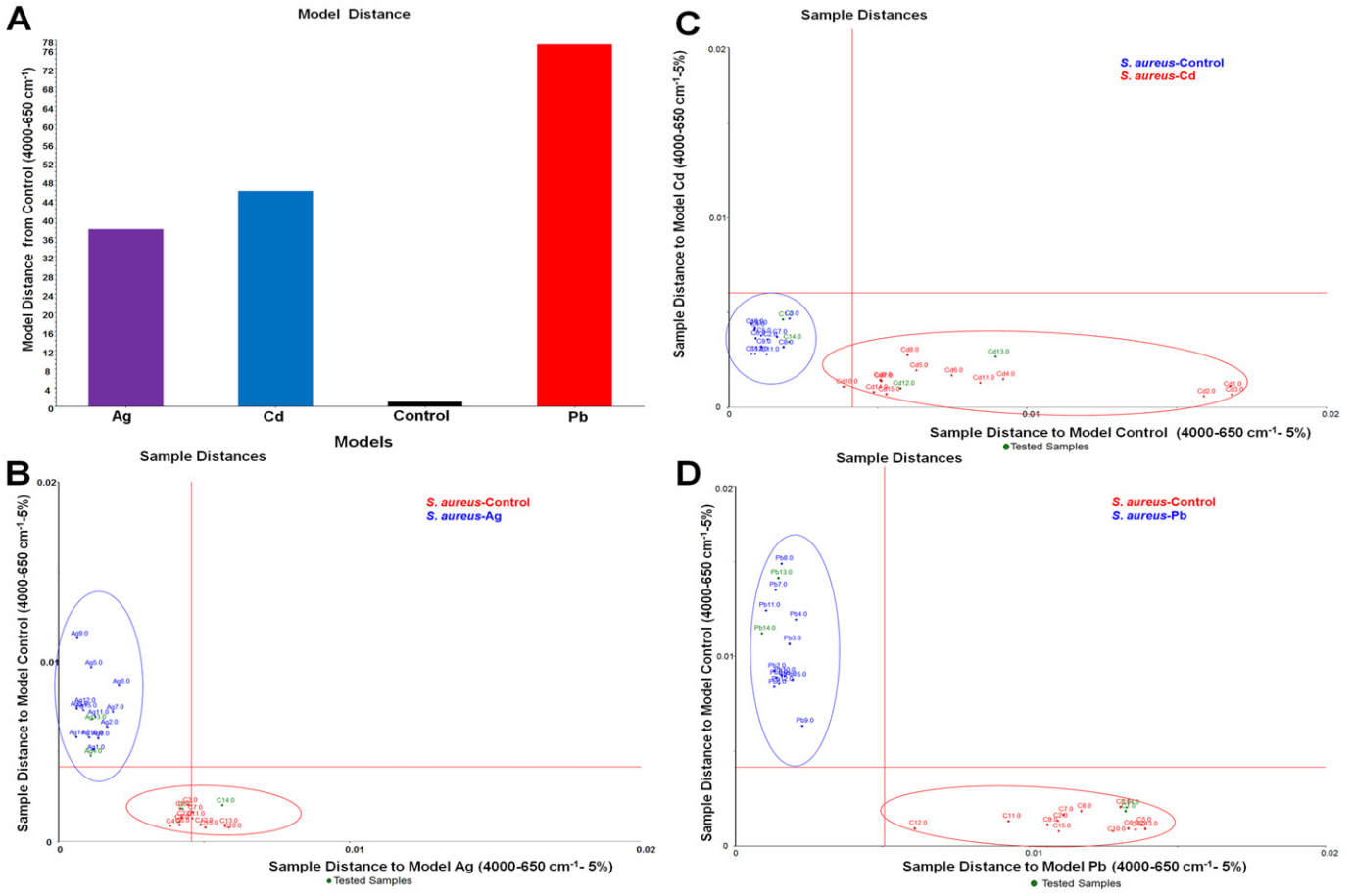

Fig. 4. The SIMCA results for *S. aureus*. (A) Model distances between the control, Ag, Cd and Pb-exposed *S. aureus*. (B), (C) and (D) Coomans plots for *S. aureus* control-Ag, control-Cd and control-Pb models, respectively.

The model distances for the control, Ag, Cd and Pb-exposed environmental *Gordonia* sp. are demonstrated in Fig. 6A. The distances between the control-Ag, control-Cd and control-Pb models were 62, 115 and 93, respectively. This result indicated the good classification between control and heavy metal-exposed *Gordonia* sp. classes. The Coomans plots for Ag, Cd and Pb-exposed *Gordonia* sp. are shown in Fig. 6B, C and D, respectively. The Coomans plots revealed the big interval between the control-Pb, and intervals between the control-Ag and control-Cd classes were satisfactory. All the Coomans plot for *Gordonia* sp. were obtained at 95% confidence level. Randomly chosen *Gordonia* sp. samples were also correctly identified (shown in green) in their corresponding *Gordonia* sp. groups.

4. Discussion

To the best of our knowledge, neither the differentiation of heavy metal-exposed bacteria nor their classification using IR spectroscopy coupled with chemometrics using supervised methods have been previously presented in the literature. We previously obtained noteworthy results on this topic but using unsupervised methods applied to IR spectral data. However, the large phenotypic variation between the same species of bacteria under stress conditions requires robust classification methods for the detection of molecular features in the IR spectra. In our study, intra-class members did not show variations in their phenotype, however inter-class members, i.e. untreated and heavy metal-exposed classes exhibited different phenotypic variations as revealed by IR spectral data. The inter-class phenotypic variations are presented as PCA scores plots. The results of this study also show that the requirement can be successfully met in the case of heavy metal-exposed bacteria's classification by applying ATR-FTIR spectroscopy based SIMCA technique. In chemometrics there are two main classes of pattern-recognition techniques; unsupervised and supervised. As mentioned in experimental section, unsupervised techniques are exploratory methods analyzing the natural relationships of specimens without any information on their training sets. Accordingly, the specimens are unpredictably sorted in several different categories beyond their primary quality and grouping positions [31]. On the contrary, primary information about the classes is needed in supervised techniques. Since the data are grouped into predefined classes over training procedures, they provide more accurate classification of samples [31]. In supervised methods, it is obligatory to validate the already trained model. In other words, independent information sets of bacteria, which were not used to calibrate the model, should be ascertained by simulating the ordinary operations. This predicts the accuracy of a model in the line of routine circumstances to estimate the percentage of correct assignments [19]. SIMCA method has some advantages with respect to other classification methods. These advantages include, placing the unknown sample to a class, for which it has a high probability. Other advantage is its sensitivity toward the quality of the data employed to compose the principal components. Therefore, the low quality variables causing noise to the principal components are generally omitted from the dataset. Furthermore, the capacity to work with as little as 10 samples per class and calculating the number of variables without restriction is a crucial advantage of SIMCA. Majority of routine classification methods fail in these circumstances due to the obstacles emerging from collinearity and chance classification [43]. In our work, all of the Coomans plots showed the discriminations of the Ag, Pb and Cd-exposed laboratory and environmental bacterial groups from the control groups with 5%

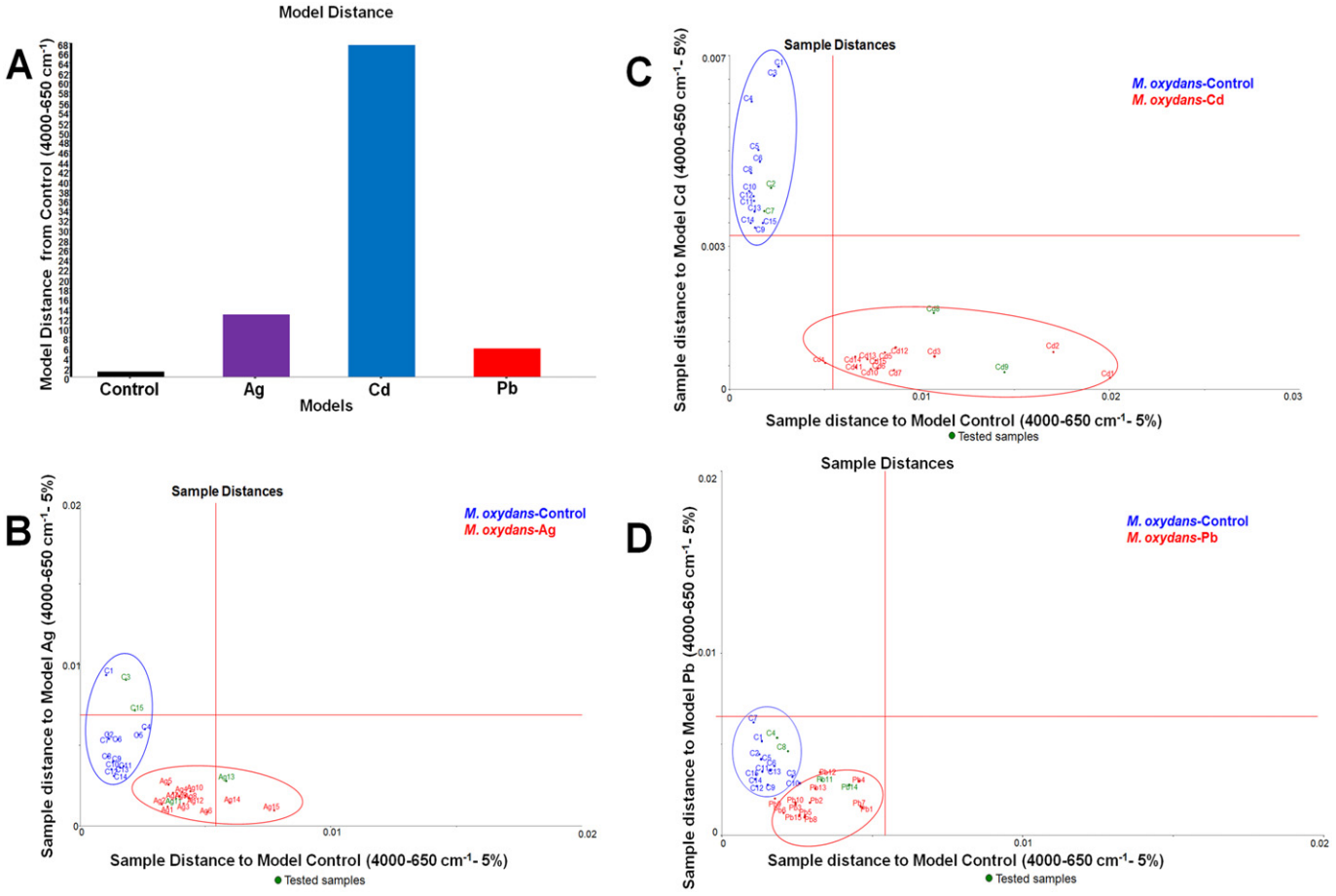

Fig. 5. The SIMCA results for *M. oxydans*. (A) Model distances between the control, Ag, Cd and Pb-exposed *M. oxydans*. (B), (C) and (D) Coomans plots for *M. oxydans* control-Ag, control-Cd and control-Pb models, respectively.

significance or 95% confidence level. In addition, the tested samples chosen randomly from each group (shown in green in the Coomans plot, Figs. 3, 4, 5 and 6) were identified correctly in their corresponding groups. Our Coomans plot results together with high model distances between bacterial groups that we obtained support that, FTIR spectroscopy coupled to proper chemometric tools are good substitutes for molecular/immunologic bacterial identification and strain typing methods by displaying rapidity and superior robustness [19]. Recently, high discrimination capacity of FTIR spectroscopy was shown to be equipotential to *spa* gene typing and pulsed field gel electrophoresis (PFGE) results for bacteria even at strain level [44]. Furthermore, these high-throughput IR based applications serve for achieving the results rapidly and they complement modern molecular and proteomic strain typing methods [44].

Numerous studies were performed over the past decade, which used FTIR spectroscopy and multivariate pattern-recognition methods including PCA and SIMCA toward the differentiation, classification and identification of a wide range of microorganisms including probiotic and foodborne bacteria [23,25,32,45–50]. In these studies differentiation was achieved in terms of taxonomical classification, Gram staining, susceptibility to antibiotics, growth medium, and contamination level of meat and apple juice and biodegradation capacities. These techniques were suggested as an effective tool to be further developed for the direct identification of bacteria [47].

Environmental microbiology also employs FTIR spectroscopy together with chemometric methods for the identification and classification of microorganisms as reported in several studies [22,24,51–55]. As an example, different strains of cyanobacteria living in freshwater, marine and terrestrial habitats were examined via FTIR spectroscopy together with HCA, PCA, *k*-Nearest Neighbors algorithm (*k*-NN) and SIMCA techniques. The study recommended FTIR spectroscopy coupled with chemometric methods as an alternative distinguishing approach for cyanobacteria based on their correct classifications in *k*-NN and SIMCA analyses [54]. Likewise, saline-tolerant strains of terrestrial cyanobacteria have been differentiated and classified via ATR-FTIR spectroscopy in combination with PCA and SIMCA [22]. FTIR spectroscopy and HCA were applied to categorize a number of environmental bacteria in faster and more economical way than any molecular approach including FAME (a microbial fingerprinting technique) [55]. In addition to bacteria, different *Aspergillus* (microfungi) species isolated from feed and bioaerosols in agricultural environment were discriminated and aflatoxin producing *A. flavus* and *A. parasiticus* were differentiated from controls, based on their FTIR spectra coupled to HCA and Discriminant Analysis (DA) [53].

Our results clearly indicated applicability of SIMCA for the differentiation and classification of heavy metal-exposed bacteria. In the previous study we obtained similar results for Cd and Pb-exposed *E. coli* and *S. aureus* where we applied unsupervised methods, namely HCA and PCA [15]. Although these techniques have been proved to be sensitive and reliable discriminatory techniques, in this study, we tested and confirmed their accuracy using powerful supervised chemometric method-SIMCA with its superior classification potential.

5. Conclusions

In this study, we propose a rapid differentiation and classification method for heavy metal-exposed bacterial strains. Considering their potential to grow at heavy metal polluted areas, they can be further

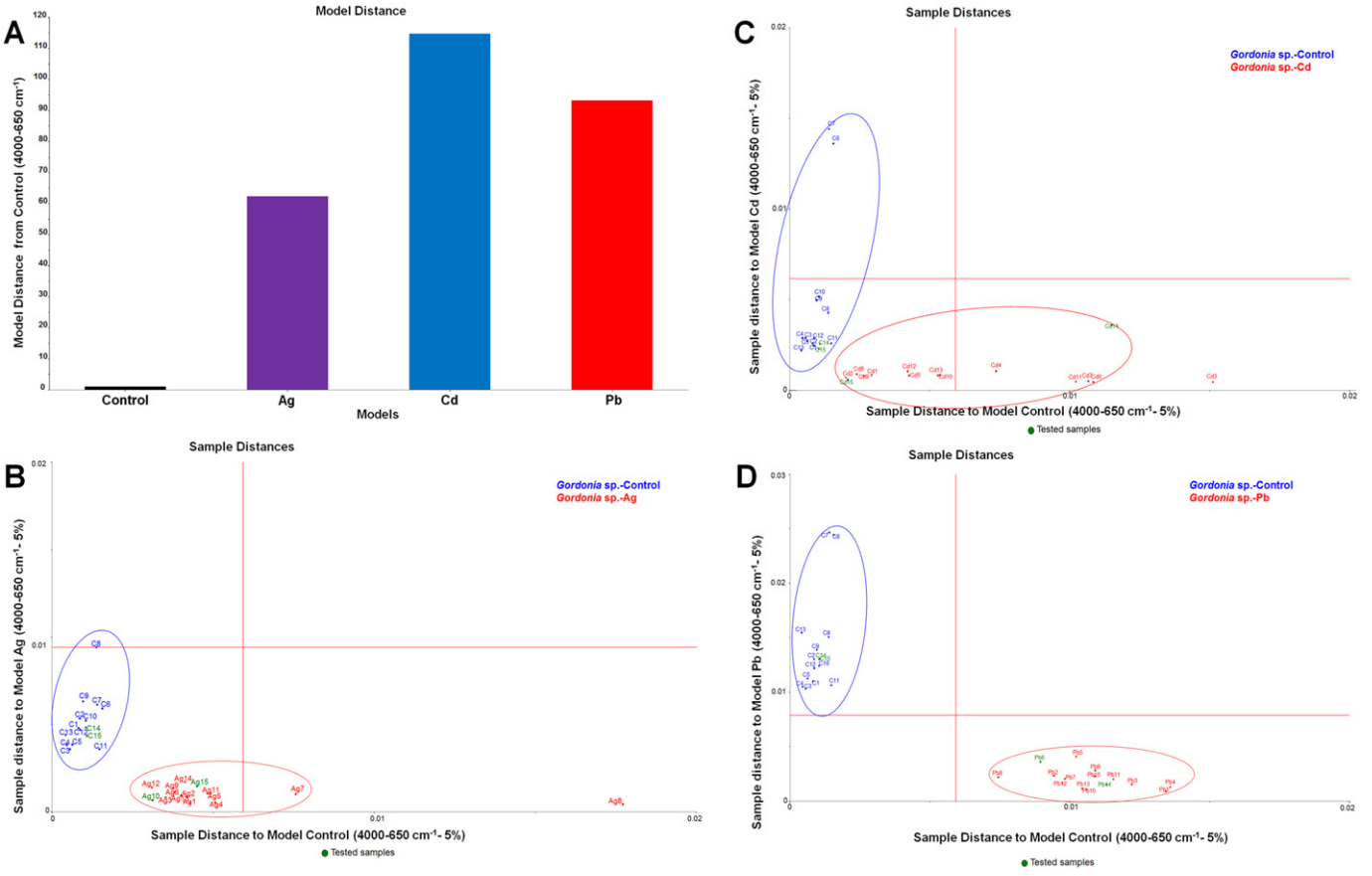

Fig. 6. The SIMCA results for *Gordonia* sp. (A) Model distances between the control, Ag, Cd and Pb-exposed *Gordonia* sp. (B), (C) and (D) Coomans plots for *Gordonia* sp. control-Ag, control-Cd and control-Pb models, respectively.

employed for the development of environmental decontamination strategies. PCA and SIMCA are commonly used to analyze chemical data by displaying the predominant components to provide accurate evaluation. Within this work, we confirmed high differentiation capacity of PCA and superior classification potential of SIMCA applied to heavy metal-exposed bacterial isolates. Promising as they look, these chemometric tools need to be further improved through the abundant FTIR databases of heavy metal-exposed bacteria prior to the direct field application. Classification of bacteria is chiefly important in microbiology, therefore, proposal of an alternative and/or complementing method to time-consuming and high-cost procedures can be beneficial for bioremediation of environmental pollutants. In this context, the identification and classification of heavy metal-exposed microorganisms in accurate, fast and productive manner, would contribute for the establishment of sustainable, green and worthwhile biogeotechnological operations to rehabilitate the soil and water.

Conflict of Interest

The authors declare no competing financial interest

Acknowledgements

The Scientific and Technological Research Council of Turkey (Grant 113Y515) and The Scientific Research Projects Fund (BAP) of the Middle East Technical University (Grant BAP-01-08-2016-005) supported this work. The authors would like to thank Dr. Nihal Simsek Ozek for her valuable comments on the manuscript.

References

[1] P.K. Rai, Heavy metal pollution in aquatic ecosystems and its phytoremediation using wetland plants: an ecosustainable approach, Int. J. Phytoremediation 10 (2008) 131–158.
[2] H. Guo, S. Luo, L. Chen, X. Xiao, Q. Xi, W. Wei, G. Zeng, C. Liu, Y. Wan, J. Chen, Y. He, Bioremediation of heavy metals by growing hyperaccumulator endophytic bacterium Bacillus sp. L14, Bioresour. Technol. 101 (2010) 8599–8605.
[3] V.V. Umrania, Bioremediation of toxic heavy metals using acidothermophilic autotrophes, Bioresour. Technol. 97 (2006) 1237–1242.
[4] R. Dixit, D. Malaviya, K. Pandiyan, U.B. Singh, A. Sahu, R. Shukla, B.P. Singh, J.P. Rai, P.K. Sharma, H. Lade, Bioremediation of heavy metals from soil and aquatic environment: an overview of principles and criteria of fundamental processes, Sustainability 7 (2015) 2189–2212.
[5] P.J. Alvarez, W.A. Illman, Bioremediation and Natural Attenuation: Process Fundamentals and Mathematical Models, John Wiley & Sons, 2005.
[6] P.K. Hopke, Chemometrics applied to environmental systems, Chemometr. Intell. Lab. 149 (2015) 205–214.
[7] Y.-Z. Feng, G. Downey, D.-W. Sun, D. Walsh, J.-L. Xu, Towards improvement in classification of *Escherichia coli*, *Listeria innocua* and their strains in isolated systems based on chemometric analysis of visible and near-infrared spectroscopic data, J. Food Eng. 149 (2015) 87–96.
[8] P. Leonard, S. Hearty, J. Brennan, L. Dunne, J. Quinn, T. Chakraborty, R. O'Kennedy, Advances in biosensors for detection of pathogens in food and water, Enzym. Microb. Technol. 32 (2003) 3–13.
[9] S. Neethirajan, D.S. Jayas, Nanotechnology for the food and bioprocessing industries, Food Bioprocess Technol. 4 (2011) 39–47.
[10] B.C. Smith, Fundamentals of Fourier Transform Infrared Spectroscopy, CRC press, 2011.
[11] D. Naumann, Infrared Spectroscopy in Microbiology, in: Encyclopedia of Analytical Chemistry, John Wiley & Sons, Ltd, 2006.
[12] K. Maquelin, C. Kirschner, L.P. Choo-Smith, N.A. Ngo-Thi, T. van Vreeswijk, M. Stammler, H.P. Endtz, H.A. Bruining, D. Naumann, G.J. Puppels, Prospective study of the performance of vibrational spectroscopies for rapid identification of bacterial and fungal pathogens recovered from blood cultures, J. Clin. Microbiol. 41 (2003) 324–329.
[13] D. Naumann, D. Helm, H. Labischinski, Microbiological characterizations by FT-IR spectroscopy, Nature 351 (1991) 81–82.
[14] H.A. Bullen, S.A. Oehrle, A.F. Bennett, N.M. Taylor, H.A. Barton, Use of attenuated total reflectance Fourier transform infrared spectroscopy to identify microbial metabolic products on carbonate mineral surfaces, Appl. Environ. Microbiol. 74 (2008) 4553–4559.
[15] R. Gurbanov, N. Simsek Ozek, A.G. Gozen, F. Severcan, Quick discrimination of heavy metal resistant bacterial populations using infrared spectroscopy coupled with chemometrics, Anal. Chem. 87 (2015) 9653–9661.
[16] M. Kardas, A.G. Gozen, F. Severcan, FTIR spectroscopy offers hints towards widespread molecular changes in cobalt-acclimated freshwater bacteria, Aquat. Toxicol. 155 (2014) 15–23.
[17] M. Wenning, F. Breitenwieser, R. Konrad, I. Huber, U. Busch, S. Scherer, Identification and differentiation of food-related bacteria: a comparison of FTIR spectroscopy and MALDI-TOF mass spectrometry, J. Microbiol. Methods 103 (2014) 44–52.
[18] F. Severcan, P.I. Haris, Vibrational Spectroscopy in Diagnosis and Screening, IOS Press, 2012.
[19] M. Wenning, S. Scherer, Identification of microorganisms by FTIR spectroscopy: perspectives and limitations of the method, Appl. Microbiol. Biotechnol. 97 (2013) 7111–7120.
[20] S.S. Souza, A.G. Cruz, E.H. Walter, J.A. Faria, R.M. Celeghini, M.M. Ferreira, D. Granato, A.d.S. Sant'Ana, Monitoring the authenticity of Brazilian UHT milk: a chemometric approach, Food Chem. 124 (2011) 692–695.
[21] P. Gemperline, Practical Guide to Chemometrics, CRC press, 2006.
[22] S. Bounphanmy, S. Thammathaworn, N. Thanee, K. Pirapathrungsuriya, J. Beardall, D. McNaughton, P. Heraud, Discrimination of cyanobacterial strains isolated from saline soils in Nakhon Ratchasima, Thailand using attenuated total reflectance FTIR spectroscopy, J. Biophotonics 3 (2010) 534–541.
[23] A. Oust, T. Møretrø, C. Kirschner, J.A. Narvhus, A. Kohler, FT-IR spectroscopy for identification of closely related lactobacilli, J. Microbiol. Methods 59 (2004) 149–162.
[24] C. Winder, E. Carr, R. Goodacre, R. Seviour, The rapid identification of Acinetobacter species using Fourier transform infrared spectroscopy, J. Appl. Microbiol. 96 (2004) 328–339.
[25] Y. Xie, S. Xu, Y. Hu, W. Chen, Y. He, X. Shi, Rapid identification and classification of staphylococcus aureus by attenuated total reflectance fourier transform infrared spectroscopy, J. Food Saf. 32 (2012) 176–183.
[26] T. Ozaktas, B. Taskin, A.G. Gozen, High level multiple antibiotic resistance among fish surface associated bacterial populations in non-aquaculture freshwater environment, Water Res. 46 (2012) 6382–6390.
[27] J.J. Ojeda, M. Dittrich, Fourier transform infrared spectroscopy for molecular analysis of microbial cells, in: A. Navid (Ed.), Microbial Systems Biology: Methods and Protocols, Humana Press, Totowa, NJ 2012, pp. 187–211.
[28] R. Bro, A.K. Smilde, Principal component analysis, Anal. Methods 6 (2014) 2812–2831.
[29] O. Preisner, J.A. Lopes, J.C. Menezes, Uncertainty assessment in FT-IR spectroscopy based bacteria classification models, Chemometr. Intell. Lab. 94 (2008) 33–42.
[30] C. Muehlethaler, G. Massonnet, P. Esseiva, The application of chemometrics on infrared and Raman spectra as a tool for the forensic analysis of paints, Forensic Sci. Int. 209 (2011) 173–182.
[31] L. Wang, B. Mizaikoff, Application of multivariate data-analysis techniques to biomedical diagnostics based on mid-infrared spectroscopy, Anal. Bioanal. Chem. 391 (2008) 1641–1654.
[32] R. Davis, L. Mauer, Fourier Transform Infrared (FT-IR) Spectroscopy: A Rapid Tool for Detection and Analysis of Foodborne Pathogenic Bacteria, 2010.
[33] R.G. Brereton, Applied Chemometrics for Scientists, John Wiley & Sons, 2007.
[34] Q. Godoi, F.O. Leme, L.C. Trevizan, E.R. Pereira Filho, I.A. Rufini, D. Santos, F.J. Krug, Laser-induced breakdown spectroscopy and chemometrics for classification of toys relying on toxic elements, Spectrochim. Acta B 66 (2011) 138–143.
[35] I. Stanimirova, B. Üstün, T. Cajka, K. Riddelova, J. Hajslova, L. Buydens, B. Walczak, Tracing the geographical origin of honeys based on volatile compounds profiles assessment using pattern recognition techniques, Food Chem. 118 (2010) 171–176.
[36] A. Cruz, R. Cadena, M. Alvaro, A. Sant'Ana, C. Oliveira, J. Faria, H. Bolini, M. Ferreira, Assessing the use of different chemometric techniques to discriminate low-fat and full-fat yogurts, LWT-Food Sci. Technol. 50 (2013) 210–214.
[37] M. Hernández-Martínez, T. Gallardo-Velázquez, G. Osorio-Revilla, Rapid characterization and identification of fatty acids in margarines using horizontal attenuate total reflectance Fourier transform infrared spectroscopy (HATR-FTIR), Eur. Food Res. Technol. 231 (2010) 321–329.
[38] A.S. Luna, A.P. da Silva, J.S. Pinho, J. Ferre, R. Boque, Rapid characterization of transgenic and non-transgenic soybean oils by chemometric methods using NIR spectroscopy, Spectrochim. Acta A Mol. Biomol. Spectrosc. 100 (2013) 115–119.
[39] J.G. Sabin, M.F. Ferrão, J.C. Furtado, Análise multivariada aplicada na identificação de fármacos antidepressivos. Parte II: Análise por componentes principais (PCA) e o método de classificação SIMCA, Braz. J. Pharm. Sci. 40 (2004).
[40] I. Nejadgholi, M. Bolic, A comparative study of PCA, SIMCA and Cole model for classification of bioimpedance spectroscopy measurements, Comput. Biol. Med. 63 (2015) 42–51.
[41] N.B. Lozano, R.F. Oliveira, K.C. Weber, K.M. Honorio, R.V. Guido, A.D. Andricopulo, A.G. de Sousa, A.B. da Silva, Pattern recognition techniques applied to the study of leishmanial glyceraldehyde-3-phosphate dehydrogenase inhibition, Int. J. Mol. Sci. 15 (2014) 3186–3203.
[42] N. Vogt, H. Knutsen, SIMCA pattern recognition classification of five infauna taxonomic groups using non-polar compounds analysed by high resolution gas chromatography, Mar. Ecol. Prog. Ser. (1985) 145–156.
[43] CAMO, SIMCA - Soft Independent Modeling of Class Analogy. CAMO Software AS, in, http://www.camo.com/resources/simca.html (accessed 8.9.16), 2016.
[44] S. Johler, R. Stephan, D. Althaus, M. Ehling-Schulz, T. Grunert, High-resolution subtyping of *Staphylococcus aureus* strains by means of Fourier-transform infrared spectroscopy, Syst. Appl. Microbiol. 39 (2016) 189–194.
[45] M.A. Al-Holy, M. Lin, A.G. Cavinato, B.A. Rasco, The use of Fourier transform infrared spectroscopy to differentiate *Escherichia coli* O157:H7 from other bacteria inoculated into apple juice, Food Microbiol. 23 (2006) 162–168.
[46] F. Chaillan, A. Le Fleche, E. Bury, Y.H. Phantavong, P. Grimont, A. Saliot, J. Oudot, Identification and biodegradation potential of tropical aerobic hydrocarbon-degrading microorganisms, Res. Microbiol. 155 (2004) 587–595.

[47] M.K. Grewal, P. Jaiswal, S. Jha, Detection of poultry meat specific bacteria using FTIR spectroscopy and chemometrics, Food Sci. Technol. 52 (2014) 3859–3869.
[48] L. Mariey, J. Signolle, C. Amiel, J. Travert, Discrimination, classification, identification of microorganisms using FTIR spectroscopy and chemometrics, Vib. Spectrosc. 26 (2001) 151–159.
[49] A. Slavchev, Z. Kovacs, H. Koshiba, A. Nagai, G. Bázár, A. Krastanov, Y. Kubota, R. Tsenkova, Monitoring of water spectral pattern reveals differences in probiotics growth when used for rapid bacteria selection, PLoS One 10 (2015), e0130698.
[50] A. Subramanian, J. Ahn, V.M. Balasubramaniam, L. Rodriguez-Saona, Monitoring biochemical changes in bacterial spore during thermal and pressure-assisted thermal processing using FT-IR spectroscopy, J. Agric. Food Chem. 55 (2007) 9311–9317.
[51] U. Behrendt, A. Ulrich, P. Schumann, D. Naumann, K.-i. Suzuki, Diversity of grass-associated Microbacteriaceae isolated from the phyllosphere and litter layer after mulching the sward; polyphasic characterization of *Subtercola pratensis* sp. nov., *Curtobacterium herbarum* sp. nov. and *Plantibacter flavus* gen. nov., sp. nov, Int. J. Syst. Evol. Microbiol. 52 (2002) 1441–1454.
[52] G. Fischer, S. Braun, R. Thissen, W. Dott, FT-IR spectroscopy as a tool for rapid identification and intra-species characterization of airborne filamentous fungi, J. Microbiol. Methods 64 (2006) 63–77.
[53] D. Garon, A. El Kaddoumi, A. Carayon, C. Amiel, FT-IR spectroscopy for rapid differentiation of *Aspergillus flavus*, *Aspergillus fumigatus*, *Aspergillus parasiticus* and characterization of aflatoxigenic isolates collected from agricultural environments, Mycopathologia 170 (2010) 131–142.
[54] M. Kansiz, P. Heraud, B. Wood, F. Burden, J. Beardall, D. McNaughton, Fourier transform infrared microspectroscopy and chemometrics as a tool for the discrimination of cyanobacterial strains, Phytochemistry 52 (1999) 407–417.
[55] B.J. Tindall, E. Brambilla, M. Steffen, R. Neumann, R. Pukall, R.M. Kroppenstedt, E. Stackebrandt, Cultivatable microbial biodiversity: gnawing at the Gordian knot, Environ. Microbiol. 2 (2000) 310–318.