Talanta 121 (2014) 65–70

Contents lists available at ScienceDirect

Talanta

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

# Rapid identification and discrimination of bacterial strains by laser induced breakdown spectroscopy and neural networks

S. Manzoor [a], S. Moncayo [a], F. Navarro-Villoslada [a], J.A. Ayala [b], R. Izquierdo-Hornillos [a], F.J. Manuel de Villena [a], J.O. Caceres [a,*]

[a] Departamento de Química Analítica, Facultad de Ciencias Químicas, Universidad Complutense, 28040 Madrid, Spain
[b] Centro de Biología Molecular "Severo Ochoa", CSIC. C/Nicolás Cabrera, 1, Cantoblanco, 28049 Madrid, Spain

ARTICLE INFO

*Article history:*
Received 18 October 2013
Received in revised form
16 December 2013
Accepted 24 December 2013
Available online 4 January 2014

*Keywords:*
Bacterial strain discrimination
Laser induced breakdown spectroscopy (LIBS)
Neural networks (NN)
Bacteria
Hospital acquired infections (HAI)

ABSTRACT

Identification and discrimination of bacterial strains of same species exhibiting resistance to antibiotics using laser induced breakdown spectroscopy (LIBS) and neural networks (NN) algorithm is reported. The method has been applied to identify 40 bacterial strains causing hospital acquired infections (HAI), i.e. *Pseudomonas aeruginosa*, *Escherichia coli, Klebsiella pneumoniae, Salmonella typhimurium*, *Salmonella pullurum* and *Salmonella salamae*. The strains analyzed included both isolated from clinical samples and constructed in laboratory that differ in mutations as a result of their resistance to one or more antibiotics. Small changes in the atomic composition of the bacterial strains, as a result of their mutations and genetic variations, were detected by the LIBS–NN methodology and led to their identification and classification. This is of utmost importance because solely identification of bacterial species is not sufficient for disease diagnosis and identification of the actual strain is also required. The proposed method was successfully able to discriminate strains of the same bacterial species. The optimized NN models provided reliable bacterial strain identification with an index of spectral correlation higher than 95% for the samples analyzed, showing the potential and effectiveness of the method to address the safety and social-cost HAI-related issue.



## 1. Introduction

Hospital acquired infections (HAI) have been widely studied in the past 30 years and have been raised to top-priority issue due to the associated economic and social costs [1]. Therefore, many preventive campaigns as well as new protocols have been implemented [2,3]. On average 5–7% of hospitalized patients are affected by HAI, and 1% of such unwanted events result in the patient′s death [4]. Bacteria are responsible for 95% of HAI, *Escherichia coli* (18.2%), *Staphylococcus* (18.1%), *Pseudomonas* (6.0%), *Enterococcus* (15.4%), *Klebsiella* (3.7%), *Acinetobacter* (0.8%), and *Salmonella* (2.8%) being the most relevant ones.

An important issue highlighted in recent years has been the increasing emergence of bacteria that are resistant to many antimicrobial therapies, sometimes resulting in multidrug-resistant strains or "super bugs." One of the overriding reasons for this is the widespread indiscriminate use of antibiotics to treat infections [5]. This antibiotic resistance is evolved under the treatment regimens of single or multidrug combinations as a result of the mutations [6].

During the past decades several methods have been proposed to optimize the identification of bacterial strain, which are based on molecular techniques such as fluorescent probes [7], microarray assemblies [8,9] and polymerase chain reaction [10,11]. However, these methodologies present some difficulties and drawbacks such as use of consumables, primer, probes or fluorescently labeled RNA antibodies [12]. Moreover, sometimes the sequences in the database are not accurate or up-to-date and micro-heterogeneity is also found common in 16S rRNA gene sequence within a species [13,14]. The phenotypic similarities between the strains of the same bacterial species restrict their identification using routine diagnostic methods [15]. Although these methods provide a reliable and accurate bacterial identification, special sample treatment methods, the high costs and low speed to perform such analysis limit their use as rapid diagnostic methods in common laboratories in order to provide quick results which leads to an increase in the rates of infectious diseases in clinical settings. Further the direct handling of these potentially pathogenic bacterial samples poses health-associated security risks. At present, clinical safety procedures and cost-related considerations do not allow an easy routine analysis of highly dangerous pathogenic bacterial specimens causing hospital acquired infections. Nevertheless, bacterial identification within the first 24 h of infection allows the use of a more effective and less risky targeted-therapy decreasing unnecessary hospitalization days and costs.

* Corresponding author. Tel.: +34 91 3944322; fax: +34 91 3944329.
*E-mail address:* jcaceres@ucm.es (J.O. Caceres).

0039-9140/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.talanta.2013.12.057

It is also necessary to point out the importance of sample preparation, which is an important step to achieve a significant result within a reasonable amount of time, thereby, avoiding or reducing the need for time-consuming culture enrichment steps. Thus, the increasing need for high speed and precision illustrates the importance of sophisticated methods for sampling and sample preparation within the overall process. The proper development and adaptation of sample preparation towards the detection method, is essential for exploiting the whole potential of the complete workflow of any diagnostic method.

Detection and identification of biological samples and, in particular bacteria, using laser induced breakdown spectroscopy (LIBS) analysis has been studied by several research groups [12,16–22]. The motivation of these studies was to evaluate the ability of LIBS to provide fast identification compared to traditional bioanalytical methods, benefitting from the possibility to combine it with chemometric methods in order to increase the performance of the technique. Morel et al. [18] investigated the detection of six bacteria by time resolved LIBS. They placed particular emphasis on *Bacillus globigii*, which acts as a non-pathogenic surrogate for *Bacillus anthracis* (anthrax), demonstrating the ability of LIBS to detect bacteria. Baudelet et al. [20] showed an unambiguous discrimination of different bacteria based on the concentration profile of trace elements. Recently, Rehse et al. [12,17] studied the effect of different experimental conditions (e.g. bacteria dilution and nutrient deprivation) on bacteria identification by discriminant functional analysis showing successful bacterial classification. Multari et al. [19] applied partial least squares regression analysis to differentiate *E. coli* from *Staphylococcus aureus* strains. Although these studies present good results, in some cases correct identification rate of bacterial strains or correlation falls below 85%. Thus, there is a clear need for more thorough and systematic studies that include new approaches making it possible to take this methodology in clinical setting for diagnosis of diseases and public health. Therefore, the motivation behind this study is to use a classification model for bacterial identification using artificial intelligence algorithms like neural networks (NN) to improve the accuracy and precision of the classification process. In a previous study by our group [23] artificial intelligence algorithms like neural networks (NN) have been used that have shown to be a promising chemometric methodology to classify and predict bacterial samples at genus level with a high degree of precision and accuracy. The full sets of variables (intensities at each wavelength) that constitute the sample spectrum are important in the process of comparison performed by the NN, which constitutes the basis of their ability to carry out discrimination. The NN is able to compute internal parameters (weights and bias) in the learning process for classifying a given set of input variables as belonging to particular sample with a high tolerance for noise and the presence of outliers [24].

In this paper we have intended to extend the previous study to investigate the application of LIBS–NN to discriminate different antibiotic resistant strains of same bacterial species and address its use as a rapid potential diagnostic methodology. The aim is to determine if genetic variations between bacterial strains of the same bacterial species, even when there is a difference in only one gene, generate sufficient or significant changes in their atomic composition which can be detected by LIBS–NN method in order to achieve their discrimination and identification.

## 2. Material and methods

### 2.1. LIBS set-up

The LIBS technique and the methodology used in the present work together with the most significant experimental conditions have been previously described [23]. Briefly, LIBS measurements were obtained using a *Q*-switched Nd:YAG laser (Quantel, Brio model) operating at 1064 nm, with a pulse duration of 4 ns full width at half maximum (FWHM), 4 mm beam diameter and 0.6 mrad divergence. Samples were placed over an X–Y–Z manual micro-metric positionator with a 0.5 μm stage of travel at every coordinate to ensure that each laser pulse impinged on a fresh sample. The laser beam was focused onto the sample surface with a 100 mm focal-distance lens, producing a spot of 100 μm in diameter. The laser fluence was fixed to 20 J/cm$^2$ and the repetition rate was 1 Hz. Emission from the plasma was collected with a 4 mm aperture, and 7 mm focus fused silica collimator placed at 3 cm from the sample, and then focused into an optical fiber (1000 μm core diameter, 0.22 numerical aperture), coupled to a spectrometer. The spectrometer system was a user-configured miniature single-fiber system EPP2000, StellarNet (Tampa, FL, U.S.A.) with a CCD detector. A grating of 300 l/mm was selected; a spectral resolution of 0.5 nm was achieved with a 7 μm entrance slit. The wavelength range used was from 200 to 1000 nm. Therefore, 2048 data points were recorded for each sample. The detector integration time was set to 1 ms. In order to prevent the detection of bremsstrahlung, the detector was triggered with a 5 μs delay time between the laser pulse and the acquired plasma radiation using a digital delay generator (Stanford model DG535). The spectrometer was computer-controlled using an interface developed with Matlab, which allowed for data processing and real-time analysis.

### 2.2. Bacterial samples

A total of 40 strains of different bacterial species i.e. *E. coli* (Ec) [25], *Pseudomonas aeruginosa* (Pa) [26], *Klebsiella pneumoniae* (Kp) [27], *Salmonella typhimurium* (St) [28], *Salmonella pullorum* (Sp) [28] and *Salmonella salamae* (Ss) [28] were included in the study. Kp, Ec and Pa strains showed multidrug antibiotic resistance and multiple genes mutations (Table 1), whereas St, Sp and Ss were resistant to kanamycin and differed in only one gene (Table 2). Two bacterial strains of the *Salmonella* species (strains 2 and 3 of Table 2) were constructed by inactivation by directed mutagenesis of specific gene in the wild type background (strains 1 of Table 2). All the bacterial strains were cultivated in LB agar (Difco Microbiology, Lawrence, KS, U.S.A) at 37 °C for 12 h in three Petri dishes (8.9 cm in diameter).

Sample ID has been represented in XYZ format, where X is the genus, Y the bacterial species and Z the type of strain. Thus, for example, Kp1 refers to the first strain of *K. pneumoniae* (K21P).

### 2.3. LIBS measurements and spectral libraries

Bacterial samples were measured directly in the Petri dish at room experimental conditions. The water content of the bacterial samples was reduced by flowing air before LIBS spectra acquisition. For each bacterial strain four replicate Petri dishes were measured. Eighty single-laser-shot spectra from one Petri dish and twenty single-laser-shot spectra each from the remaining three Petri dishes were acquired. The acquisition time of these 140 spectra was approximately 2 min. Because the emission intensity signal may vary with laser pulse, spectra were normalized by the most intense emission line, i.e. Na(I), to avoid instrumental variations (Fig. 1). The 80 spectra from the first Petri dish were used to train the NN model for each bacterial strain (spectral training library), whereas the 20 spectra from the remaining three Petri dishes each were used to test the model (spectral test library). Although the matrix dataset was considerably large, the computation time for training each NN model was below 10 s.

**Table 1**
Description of bacterial strains with multidrug antibiotic resistance.

| Bacterial species | Strain | Antibiotic resistance phenotypes diffusion disk[ab] | Sample ID |
|---|---|---|---|
| ***Klebsiella pneumoniae*** | K21P[c] | Amp, Amo, Amc, Tet, Nal, Fox, Cip, Tsu, Ctx, Caz, Azt | Kp1 |
| | K18P[c] | Amp, Amo, Amc, Clo, Tet, Nal, Cip, Stm, Tsu, Caz, Azt | Kp2 |
| | K17P[c] | Amp, Amo, Amc, Tet, Nal, Cip, Tsu, Ctx, Caz, Azt | Kp3 |
| | K16R[c] | Amp, Amo, Amc, Clo, Stm, Tsu, Ctx, Caz | Kp4 |
| | K11CM[c] | Amp, Amo, Amc, Clo, Tet, Nal, Fox, Cip, Tsu, Ctx, Caz, Azt | Kp5 |
| | K11P[c] | Amp, Amo, Amc, Clo, Tet, Nal, Cip, Stm, Tsu, Ctx, Caz, Azt | Kp6 |
| | K7P[c] | Amp, Amo, Amc, Clo, Tet, Nal, Fox, Cip, Tsu, Ctx, Caz, Azt | Kp7 |
| | K6P[c] | Amp, Amo, Amc, Clo, Tsu, Ctx, Caz, Azt | Kp8 |
| | K3C[c] | Amp, Amo, Tet, Ctx (Int.) | Kp9 |
| | K2P[c] | Amp, Amo, Tet, Nal, Tsu, Caz, Azt | Kp10 |
| ***Escherichia coli*** | MC6-RP11 | MC6RP1 *leu*+*ftsA3*(Ts) | Ec1 |
| | QCB1 | MC6RP1 *ponB*::*Spc*$^r$ | Ec2 |
| ***Pseudomonas aeruginosa*** | PA1 | Serotype 010 | Pa1 |
| | PA2 | Serotype 014, Gen, Amk, Tob, Azt, Tic, Pip, Clo, Car, Eri | Pa2 |
| | PA3 | Serotype 011 | Pa3 |
| | PA4 | Serotype N/A | Pa4 |
| | PA5 | Serotype 04, Azt, Tic, Pip, Cip, Clo, Car, Eri | Pa5 |
| | PA6 | Serotype 011 | Pa6 |
| | PA7 | Serotype 014 | Pa7 |
| | PA8 | Serotype 01, Azt, Tic, Pip, Cip, Clo, Car, Eri | Pa8 |
| | PA9 | Serotype 09 | Pa9 |
| | PA10 | Serotype 06, Azt, Tic, Pip, Cip, Clo, Car, Eri | Pa10 |
| | PA11 | Serotype 010 | Pa11 |
| | PA12 | Serotype 015 | Pa12 |
| | PA13 | Serotype 09 | Pa13 |
| | PA14 | Serotype 011 | Pa14 |
| | PA15 | Serotype 011 | Pa15 |
| | PA16 | Serotype 011, Gen, Amk, Tob, Azt, Tic, Pip, Imp, Cip | Pa16 |
| | PA17 | Serotype 010, Gen, Amk, Tob, Azt, Tic, Pip, Cip, Imp, Clo, Car, Eri | Pa17 |
| | PA18 | Serotype 015 | Pa18 |
| | PA19 | Serotype 011 | Pa19 |

[a] Amp, ampicillin; Amo, amoxicillin; Amc, amoxicillin-clavulanate; Clo, chloramphenicol; Tet, tetracycline; Nal, nalidixic acid; Fox, cefoxitin; Cip, ciplofloxacin; Stm, streptomycin; Tsu, trimethoprim/sulfamethoxazole; Ctx, cefotaxime; Caz, ceftazimide; Azt, aztreonam; (Int.), Intermediate.
[b] Gen, gentamicin; Amk, amikacin; Tob, tobramicin; Azt, aztreonam; Tic, ticarcilin; Pip, piperacillin; Cip, ciplofloxacin; Imp, imipenem; Clo, chloramphenicol; Car, carbenicillin; Caz, ceftazimide; Eri, erithromicin.
[c] *K. pneumoniae* isolates resistant or intermediate to cephalosporin by the disk diffusion method.

**Table 2**
*Salmonella* strains that differ in a single gene.

| Bacterial species | Strain | | | Sample ID |
|---|---|---|---|---|
| | 1 | 2 | 3 | |
| *Salmonella pollorum* | 1JVC | 1/1Km | 2/1Km | Sp |
| *Salmonella salamae* | 2JVC | 1/2Km | 2/2Km | Ss |
| *Salmonella typhimurium* | SL-1344 | 1/22Km | 2/22Km | St |

JVC: Wild type codification.
Km: Mutant in a geneby insertion of a Kanamycin resistance casette.

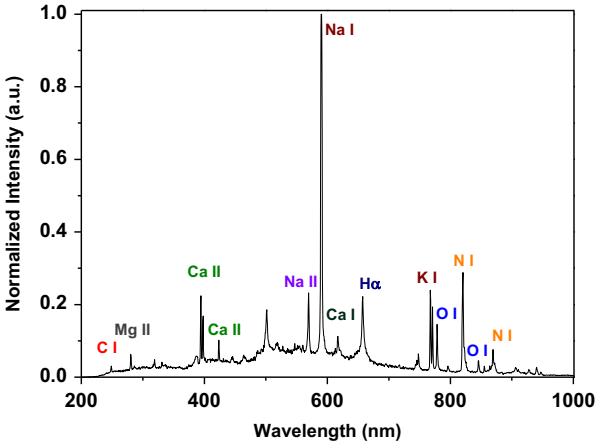


**Fig. 1.** Normalized LIBS spectrum of *Escherichia coli.*

### 2.4. Neural network model

Home-made neural network software specifically designed to deal with the problem of bacterial classification was developed. The NN models were based on a multilayer perceptron, feedforward, supervised network that consisted of several neurons (information processing units) arranged in two or more layers receiving information from all of the neurons of the previous layer. The connections are controlled by a weight that modulates the output from the neuron before inputting its numerical content into a neuron in the next layer.

The NN topology consists of three layers (input, hidden and output), which is widely used to model systems with a similar level of complexity [29]. In particular, the input layer consisted of 2048 nodes (spectral response in the 200–1000 nm wavelength range). The number of neurons in the hidden layer was 10. The output layer (classification result) was comprised of $J$ neurons (where $J$= number of reference samples used) for estimating the similarity between the reference sample spectra and the testing sample spectrum.

The process that optimizes the weights, i.e. the learning or training process was based on a back-propagation (BP) algorithm [30,31]. The inputs from each neuron are added by an activation function, and the result is transformed by a transfer function that limits the amplitude of the neuron output. In this work, the hyperbolic tangent sigmoid function was used as the NN transfer function. Every NN model was estimated using Matlab software (Mathworks, 2010a).

#### 2.4.1. Neural network model training

Because the NN is a supervised method, in order to optimize the weight matrix it is necessary to use input and output data that

adequately characterized the system to be modeled. The spectral data of the training library was randomly divided as a part of the training process into two subsets: 80% (64 spectra) for training and 20% (16 spectra) for self-validation of the model. Once the training and self-validation process was carried out, the models were validated by testing the 60 spectra from the remaining three replicas of the bacterial strains.

The identification process was based on the ability of the NN to detect the degree of similarity between the new spectrum and each of the reference spectra used in the training process. During the training process, each bacterial strain was associated with an identification number in the output layer. Thus, a perfect identification is obtained if the output from the NN model for the test samples of the same bacterial strain matched the identification number assigned to the bacterial strain used to train the model. Nevertheless, it is possible to use more than two bacterial strains simultaneously to train the NN to speed up the sample analysis but an increase in data manipulation occurs. Zero identification number was always used to indicate no match at all.

NN training was achieved by applying the BP algorithm based on the conjugate gradient method [32], one of the general-purpose second-order techniques that helps minimize the goal functions of several variables. Second order indicates that such methods use the second derivatives of the error function, whereas a first-order technique, such as standard back-propagation, uses only the first derivatives. To determine when the training should be stopped, an early stopping criteria based on performance improving (error rate) of the validation set [33]. The number of epochs was not relevant in this case. To avoid an overfitting of the NN model, the learning process was repeated until a minimum of the mean square error (MSE) of the verification data, defined in Eq. (1), was reached:

$$MSE = \frac{1}{N}\sum_{k}^{N}(r_k - y_k)^2 \quad (1)$$

where $N$, $y_k$, and $r_k$ are the number of input data, the response from each output neuron, and the observed output response, respectively. A detailed description of the calculation process is provided in the literature [30,33].

#### 2.4.2. Neural network model validation

Three validation tests were carried out for each NN model. The first validation test performed was inter-bacterial species discrimination where the ability of NN models was evaluated to classify the bacterial strains on the basis of their species. The second and third validation tests were performed to evaluate the capacity of the LIBS–NN methodology to discriminate between strains of the same bacterial species with small changes in the atomic composition as a result of their resistance, mutations and genetic variations. The second validation test included strains which differed in resistance to different antibiotics, whereas, the third validation test included strains which differed in only one gene.

The model performance was evaluated by its accuracy, i.e. the rate of correct bacterial classification or correlation within the classified spectra. The model accuracy was estimated by the parameter "spectral correlation", which is expressed as the percentage of test-set spectra classified correctly. The higher the index of spectral correlation of a test set, the better the capacity of the NN model for discriminating a sample. A spectral correlation value higher than 90% was considered as correct identification of the sample.

## 3. Results and discussion

LIBS experiments were performed for some typical HAI-causing bacteria and therefore important pathogens from medical point of view. Fig. 1 shows a normalized LIBS spectrum of an Ec bacterial sample, identifying the elements with the most intense emission lines.

It has been shown that elements such as P, C, Mg, Ca and Na provide sufficient information to achieve bacterial identification [12,20,34]. However, due to the high similarity between the spectra of the bacterial strain samples (especially in the case of the differences is only in one gene), a broad spectral range (200–1000 nm) was used in order to cover the greater number of spectral characteristics and improve the performance of the NN model. Several papers have shown [35] that reducing the number of variables in the training of NN using spectral zones with few peaks, selected by PCA, decreases the performance of the NN model. Therefore, the selection of certain variables does not imply that the other variables are redundant with respect to selected. Thus, the larger the information (representative data) is used in the training of the NN model, the better is the predictive capability of the model. On the other hand, neural networks have a much better noise performance.

Despite that there are not significant variations in the spectra to easily discriminate the bacterial strains, from the mathematical point of view each bacterial strain can be discriminated based on its complete spectral fingerprint [24]. Therefore, the whole set of variables that constitute the sample spectrum is important in the classification process performed by the NN model by computing internal parameters (weights and bias) in the training step. This fact constitutes the basis of their ability to carry out the discrimination between the bacterial strains with high tolerance for noise and the presence of outliers.

Because samples were measured at room conditions, emission signals from air were also observed in the LIBS spectra. However, as it has been demonstrated, these signals do not significantly interfere the relative spectral contribution of the trace elements of the samples [36]. In order to improve the signal to noise ratio of the emission lines, the water content of the samples was reduced by flowing air before LIBS spectra acquisition. On the other hand, a continuum background emission was also observed as a typical component of LIBS spectra.

### 3.1. Inter-bacterial species discrimination

Inter-bacterial species discrimination test involved the classification of bacterial strains based on their species. The capacity of LIBS–NN method to discriminate bacterial species, independent of the strain of the bacterial species used to train the NN model, was evaluated. For carrying out this validation test, the NN was trained with two libraries of different bacterial species, and tested with the libraries of the bacterial strains of all species. Even though it is possible to include the spectral datasets of all the bacterial strains in the training phase of the neural network, we have used only two bacterial libraries to train the neural network as a first approach to test whether there are sufficient significant spectral differences between different bacterial strains that allow for discrimination. Fig. 2 shows that the LIBS–NN methodology developed was able to correctly classify all the bacterial strains to their corresponding bacterial species, achieving accurate and reliable results. All bacterial strains were assigned to their species (95% spectral correlation) independent of the bacterial strains used to train the NN model, demonstrating the robustness of the methodology. Hence, classification is based on the major differences in the spectral fingerprints of the bacteria at species level.

### 3.2. Intra-bacterial strain discrimination of multidrug resistant bacteria

Strains of Kp, Ec and Pa which acquired resistance to multiple antibiotics were selected for LIBS analysis. These strains exhibited

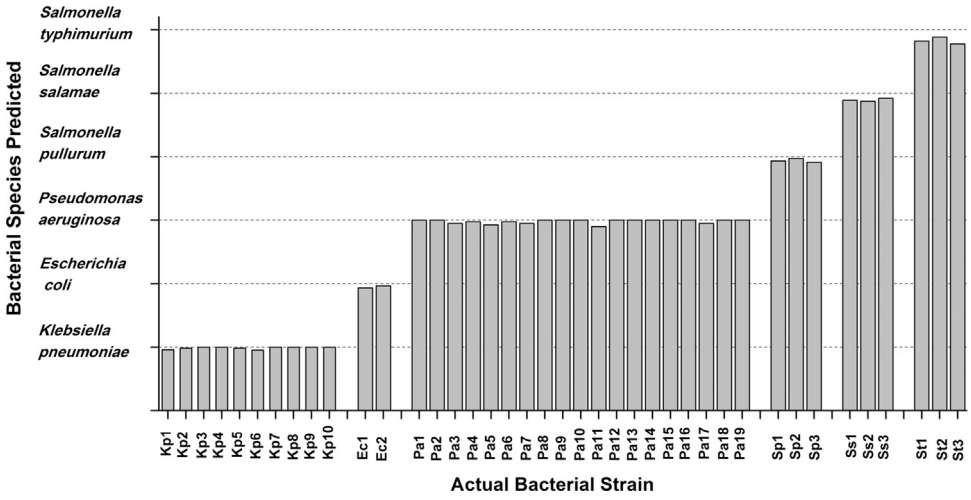


**Fig. 2.** Inter-bacterial species classification.

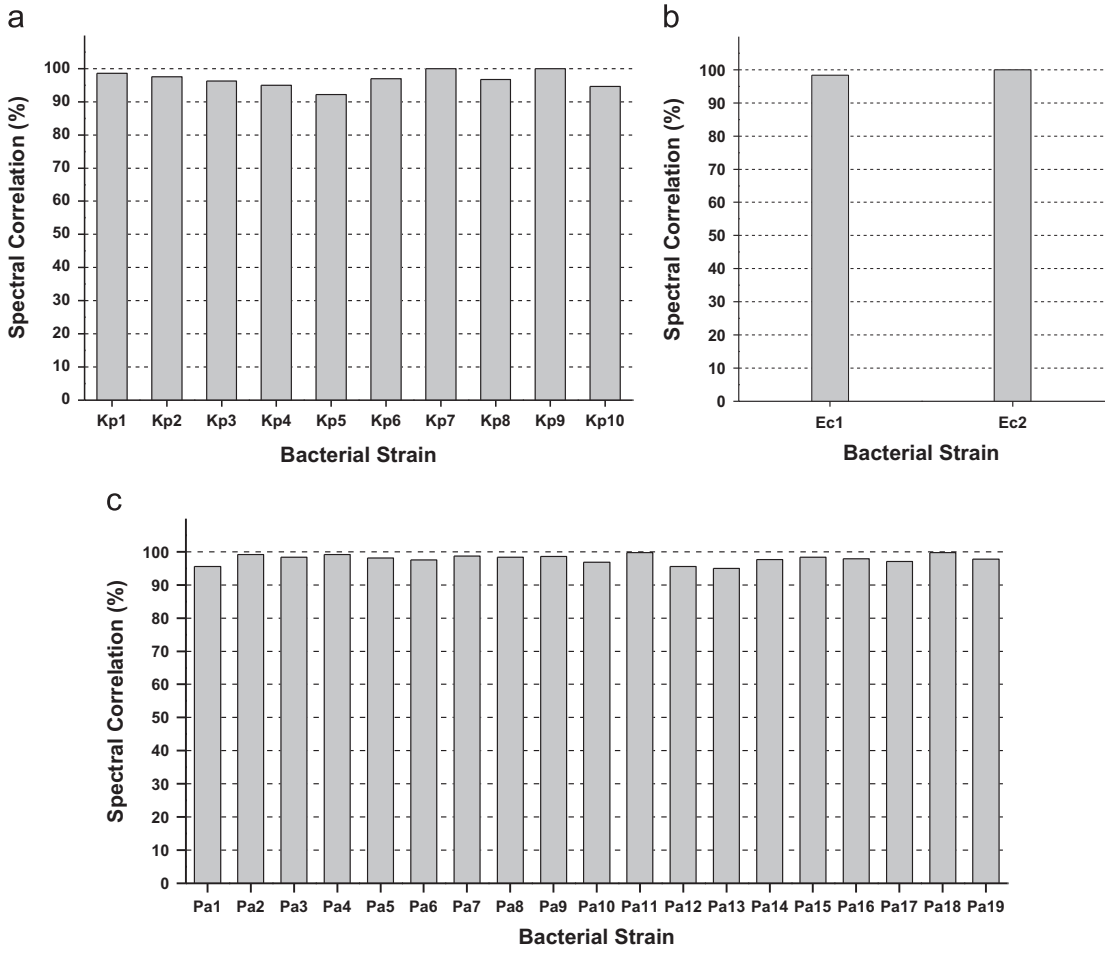


**Fig. 3.** Intra-bacterial strain classification of multidrug resistant strains of (a) *K. pneumoniae* (b) *E. coli* (c) *P. aeruginosa*.

different antibiotic resistance patterns and, hence, different mutations. In this validation, NN models were estimated training the NN with two libraries of different strains of the same bacterial species and validated with all the test libraries of the strains of that species. An average spectral correlation for each bacterial strain and estimated model was calculated. Fig. 3 shows the classification results (spectral correlation) obtained for all the models and bacterial species. A spectral correlation higher than 92%, 98.3% and 95% for *K. pneumoniae*, *E. Coli* and *P. aeruginosa* was achieved, respectively, demonstrating the good performance of the LIBS–NN methodology to classify the strain samples for each of the bacterial species.

The fact that 100% spectral correlation was not achieved could be attributed to that each spectrum was recorded with a single laser shot. Nevertheless, this lack of complete accuracy was considered no significant in the overall performance of the LIBS–NN method

**Table 3**
Single-gene intra-bacterial strain classification results.

| Bacterial species | Spectral correlation % Strain | | |
|---|---|---|---|
| | 1 | 2 | 3 |
| *S. pollorum* | 97.3 | 95.0 | 97.0 |
| *S. salamae* | 95.0 | 96.0 | 97.2 |
| *S. typhimurium* | 96.6 | 97.5 | 97.5 |

developed to discriminate strains of different bacterial species. Therefore, these results showed that NN trained with datasets of different bacterial strains allowed to discriminate strains of the same bacterial species with multidrug resistance and genetic variations that imparted changes in their elemental composition providing a characteristic fingerprint and hence, enabling LIBS-based bacterial identification.

### 3.3. Single-gene intra-bacterial strain discrimination

A single-gene intra-bacterial strain discrimination of *Salmonella* species was carried out. Three different strains for each species with genetic variation of only one gene were selected. This minor variation in the genetic code of the bacterial strains made the training of the NN to take more time (ca. 3 s) in order to reach the minimum MSE. In this validation, NN models were estimated training the NN with two libraries of different strains of the same bacterial species and validated with all the strains' libraries of the same species. An average spectral correlation for the three estimated models for each bacterial strain and bacterial species was calculated. Table 3 shows that a spectral correlation higher than 95%, 95% and 96.6% for *S. pollorum*, *S. salamae* and *S. typhimurium* was achieved, respectively, demonstrating the capacity of the LIBS–NN methodology to classify bacterial strain samples that differed in a single gene.

These results demonstrate that when the NN models were trained with libraries of different species, they were able to correctly identify strains on the basis of their species. Moreover, the discrimination capability of the NN models is not only based on their ability to find similarities between spectra but also due to their capacity to tolerate small spectral variations among them. This feature led to classify the strains to their corresponding species and also the intra-bacterial strain discrimination.

## 4. Conclusions

A method based on laser induced breakdown spectroscopy (LIBS) and neural networks (NN) algorithms was developed and applied to achieve rapid identification and discrimination of different bacterial species and strains causing hospital acquired infections, including multidrug resistance and single gene variation bacterial strains based on their characteristic spectral fingerprint. Single shot LIBS measurements combined with a supervised neural network method were sufficient for a clear identification and classification of bacterial strains of different species due to the reliability and robustness of the estimated non-linear classification models.

The results show that the LIBS–NN methodology proposed is able to discriminate bacterial species and strains with high accuracy. Therefore, it may be considered a quick, simple and cost-effective alternative for the slower and more expensive biological methods to discriminate strains of the same bacterial species. The three validation tests carried out show the ability of the NN models for rapid identification of bacterial samples from species to strain level with minor genetic variations. Moreover, single gene variations are enough to discriminate strain of the same bacterial specie. From the medical point of view, these capabilities would allow an early diagnosis of the bacterial infections and their treatment which may reduce the recurrence of HAI.

## Acknowledgements

The authors gratefully acknowledge financial support from Complutense University of Madrid (CCG10-UCM/PPQ-4713) and the Ministry of Science and Innovation of Spain (BFU2009-09200).

## References

[1] K. Kaier, Value in Health 15 (2012) 87–93.
[2] J. Vaqué, J. Rosselló, J.L. Arribas, J. Hosp. Infect. 43 (Suppl. 1) (1999) S105–S111.
[3] K.W. Lobdell, S. Stamou, J.A. Sanchez, Surg. Clin. North Am. 92 (2012) 65–77.
[4] The European Healthcare-Associated Infections. HAI-Net Annual Reports. ⟨http://www.ecdc.europa.eu/en/healthtopics/healthcare-associated_infections/epidemiological_data/pages/hai-net_annual_reports.aspx⟩, (2012).
[5] M. Zourob, S. Elwary, A. Turner (Eds.), Principles of Bacterial Detection: Biosensors, Recognition Receptors and Microsystems, Springer, New York, 2008.
[6] A.C. Palmer, R. Kishony, Nat. Rev. Genet. 14 (2013) 243–248.
[7] K. Zwirglmaier, FEMS Microbiol. Lett. 246 (2005) 151–158.
[8] D. Wang, L. Coscoy, M. Zylberberg, P.C. Avila, H.A. Boushey, D. Ganem, J.L. DeRisi, Proc. Nat. Acad. Sci. U.S.A. 99 (2002) 15687–15692.
[9] T. Kostic, P. Butaye, J. Schrenzel, Detection of Highly Dangerous Pathogens: Microarray Methods for the Detection of BSL 3 and BSL 4 Agents, John Wiley & Sons, 2009.
[10] L.S. Garcia, Clinical Microbiology Procedures Handbook, ASM Press, 2010 (3 Lslf).
[11] American Society for Microbiology (Washington), Manual of Clinical Microbiology, eighth ed., ASM Press, Washington, D.C, 2003.
[12] S.J. Rehse, Q.I. Mohaidat, S. Palchaudhuri, Appl. Opt. 49 (2010) C27–C35.
[13] J.E. Clarridge, Clin. Microbiol. Rev. 17 (2004) 840–862.
[14] J.L. Cloud, D. Harmsen, P.C. Iwen, J.J. Dunn, G. Hall, P.R. LaSala, K. Hoggan, D. Wilson, G.L. Woods, A. Mellmann, J. Clin. Microbiol. 48 (2010) 1442–1444.
[15] A. Álvarez-Buylla, E. Culebras, J.J. Picazo, Infect. Genet. Evol. 12 (2012) 345–349.
[16] S.J. Rehse, N. Jeyasingham, J. Diedrich, S. Palchaudhuri, J. Appl. Phys. 105 (2009). (102034/1-102034/13).
[17] Q. Mohaidat, S. Palchaudhuri, S.J. Rehse, Appl. Spectrosc. 65 (2011) 386–392.
[18] S. Morel, N. Leone, P. Adam, J. Amouroux, Appl. Opt. 42 (2003) 6184–6191.
[19] R.A. Multari, D.A. Cremers, J.M. Dupre, J.E. Gustafson, Appl. Spectrosc. 64 (2010) 750–759.
[20] M. Baudelet, J. Yu, M. Bossu, J. Jovelet, J.-P. Wolf, T. Amodeo, E. Frejafon, P. Laloi, Appl. Phys. Lett. 89 (2006). (163903/1-163903/3).
[21] M. Baudelet, L. Guyon, J. Yu, J.-P. Wolf, T. Amodeo, E. Frejafon, P. Laloi, J. Appl. Phys. 99 (2006). (084701/1-084701/9).
[22] R.A. Putnam, Q.I. Mohaidat, A. Daabous, S.J. Rehse, Spectrochim. Acta, Part B 87 (2013) 161–167.
[23] D. Marcos-Martinez, J.A. Ayala, R.C. Izquierdo-Hornillos, F.J.M. de Villena, J.O. Caceres, Talanta 84 (2011) 730–737.
[24] J.O. Caceres, S. Moncayo, J.D. Rosales, F.J.M. de Villena, F.C. Alvira, G.M. Bilmes, Appl. Spectrosc. 67 (2013) 1064–1072.
[25] F. Garcia del Portillo, M.A. de Pedro, J. Bacteriol. 172 (1990) 5863–5870.
[26] C.P. Guillermo, Ph.D. Thesis – Universidad Autónoma de Madrid – Spain, 2003.
[27] A.C.S. Lopes, D.L. Veras, A.M.S. Lima, R. de C.A. Melo, J.A. Ayala, Mem. Inst. Oswaldo Cruz 105 (2010) 163–167.
[28] J. Velasco, M. Araque, J.A. Ayala, Colombiana de Microbiología 12 (2010) 55–66.
[29] J.-B. Sirven, B. Bousquet, L. Canioni, L. Sarger, Anal. Chem. 78 (2006) 1462–1469.
[30] H.B. Demuth, M.H. Beale, M.T. Hagan, Neural Network Toolbox for Use with MATLAB: User's Guide Ninth for Version 6.0 (Release 2008a), Math Works, 2007.
[31] A.J. Maren, C.T. Harston, Handbook of Neural Computing Applications, Academic Press, San Diego, USA, 1990.
[32] M.F. Møller, Neural Networks 6 (1993) 525–533.
[33] C.M. Bishop, Neural Networks for Pattern Recognition, Oxford University Press, USA, 1996.
[34] S.J. Rehse, J. Diedrich, S. Palchaudhuri, Spectrochim. Acta, Part B 62 (2007) 1169–1176.
[35] A. Ramil, A.J. López, A. Yáñez, Appl. Phys. A 92 (2008) 197–202.
[36] R.J. Nordstrom, Appl. Spectrosc. 49 (1995) 1490–1499.