Microchemical Journal 124 (2016) 306–310

Contents lists available at ScienceDirect

Microchemical Journal

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

# Near-infrared spectroscopy and variable selection techniques to discriminate *Pseudomonas aeruginosa* strains in clinical samples

Aline S. Marques [a], Jenielly N.F. Castro [b], Fagner J.M.D. Costa [b], Renato M. Neto [b], Kássio M.G. Lima [a,*]

[a] Biological Chemistry and Chemometrics, Institute of Chemistry, Federal University of Rio Grande of Norte, Natal 59072-970, RN-Brazil
[b] Laboratory of Mycobateria, Department of Microbiology and Parasitology, Federal University of Rio Grande do Norte, Natal 59072-970, RN-Brazil

ARTICLE INFO

*Article history:*
Received 16 July 2015
Received in revised form 9 September 2015
Accepted 10 September 2015
Available online 21 September 2015

*Keywords:*
*Pseudomonas aeruginosa*
Multi-resistant
Sensitive
Near-infrared spectroscopy
SPA-LDA
GA-LDA

ABSTRACT

*Pseudomonas aeruginosa* is a leading cause of nosocomial infections, ranking second among the negative Gram pathogens reported to the National Nosocomial Infection Surveillance System. *P. aeruginosa* may develop resistant during prolonged therapy with all antimicrobial agents. Therefore, isolates that are initially susceptible may become resistant within 3–4 days after initiation of therapy. Testing of repeat isolates may be warranted. There is a need for sensitive and specific tests. We set out to determine whether near-infrared spectroscopy (NIR) combined with variable selection techniques employing successive projection algorithm – linear discriminant analysis (SPA-LDA) or genetic algorithm – (GA-LDA) could discriminate *P. aeruginosa* strains according to resistant vs. sensitive. The variables selected were then used for discriminating the strains. The influence of various spectral pre-treatments (Savitzky–Golay smoothing, multiplicative scatter correction (MSC), and Savitzky–Golay derivatives) was calculated. In addition, accuracy test results including sensitivity and specificity were determined. Sensitivity in the resistant category was 95% using a SPA-LDA model with 70 wavelengths. Sensitivity and specificity in both categories was 93% using a GA-LDA model with 32 wavelengths. We show that NIR spectroscopy of *P. aeruginosa* combined with variable selection techniques is a powerful tool for resistant vs. sensitive strains based on the unique spectral "fingerprints" of their biochemical microbial identification, emerging as an alternative for rapid and cost-effective identification of strains.



## 1. Introduction

Some bacteria are intrinsically resistant to antimicrobial agents because they lack the target site for that drug, or the drug is unable to reach the site of action, or the organism contains a chromosomally encoded resistant mechanism. For example, isolates of *Pseudomonas aeruginosa* may become resistant to the antibiotic being used to treat the infection [1] and that prior use of a particular antibiotic predicts that *P. aeruginosa* will develop resistant to that antibiotic [2]. *P. aeruginosa* is the most prevalent species of non-fermentative Gram-negative bacilli (NFGNB) in clinical isolates of nosocomial nature, the leading cause of nosocomial bloodstream infections, ranking third among Gram-negative bacteria after *Escherichia coli* and *Klebsiella* species [3].

Detection of *P.aeruginosa* in the clinical laboratory is of major importance for the determination of appropriate therapeutic schemes and the implementation of infection control measures. In order to detect the presence of *P.aeruginosa* in clinical samples, the polymerase chain reaction (PCR), [4] disk diffusion or dilution methods, [5] broth microdilution (BMD), [6] and enzyme block test [7] are standard microbiological assays. However, these methods are time consuming, expensive, and involve numerical preparation steps like indispensable selective pre-enrichment. There is a need for sensitive and specific tests which are faster, non-destructive, and have/are specific means of detecting and differentiating *P. aeruginosa* strains.

Alternatively, near-infrared spectroscopy (NIRS) can be used on various microbiologic species for direct detection and identification [8–10], emerging as an interesting alternative for quick and cost-effective identification of living specimens. NIRS is also characterized by a minimum of sample handling. It requires no extractions and is non-destructive. Each bacterial species has a complex cell wall/membrane composition which gives a unique NIR fingerprint, due to the stretching and bending vibrations of molecular bonds or functional groups present in its proteins, nucleic acids, lipids, sugars, and lipopolysaccharides, among others. Therefore, each bacterium will have a unique and characteristic spectrum, and single microorganisms could be identified from an NIR spectrum. For example, Marques et al. [9] recently determined the biochemical intra-individual differences or "fingerprint" features between *Klebsiella pneumoniae* Carbapenemase (KPC–2)-producing and non-producing *Klebsiella pneumoniae* (KP) using NIR spectroscopy with subsequent variable selection methods. The authors concluded that NIRS

* Corresponding author. Tel.: +55 84 3342 2323.
*E-mail address:* kassiolima@gmail.com (K.M.G. Lima).

http://dx.doi.org/10.1016/j.microc.2015.09.006
0026-265X/

may provide a novel approach to identify potential biochemical markers that may have a relation with microbial functional roles between KPC-2 producing and non-producing KP strains.

Several circumstances have contributed to the successful development of NIR spectroscopy of microbiologic species. The use of multivariate statistical approaches, which allow for the extraction of qualitative and quantitative information from complex spectra for bacterial characterization and are largely responsible for advancing the NIR technique. For example, principal component analysis (PCA) for reducing the multidimensionality of the data set into its most dominant components or scores while maintaining the relevant variation between the data points [8], hierarchical cluster analysis (HCA) for identifying similarities between the spectra of microorganisms using the distances between spectra [11], and linear discriminant analysis (LDA) for the classification of objects into groups or clusters by determining the similarity of a set of values from an unknown sample to a set of values measured from a set of known samples [12]. In addition, variable selection methods, more specifically, successive projections algorithm (SPA) [13] and genetic algorithm (GA) [14] in conjunction with LDA have improved the model performance compared with the full spectrum model. On the other hand, a standardized experimental protocol in relation to media preparation, incubation time and temperature, cell harvesting conditions, sample preparations, and NIR measurement should be followed to obtain reproducible data.

The present paper investigates the use of NIR spectroscopy and chemometric techniques as a quick and non-destructive method for classification of *P. aeruginosa* strains into two different categories: species of standard multidrug-resistant phenotype (resistant to more than one class of antibacterial agent), and sensitive to all classes of antibacterial agents tested. This study is the first to apply NIR spectroscopy coupled with variable selection techniques to identify *P. aeruginosa* with multidrug-resistant isolated from clinical material profile. We employed SPA and GA to select an appropriate subset of wavelengths for LDA that reflects a specific biochemical feature of each category.

## 2. Material and methods

### 2.1. Bacteria strains

Sixty-three strains of *P. aeruginosa*, with varied susceptibility profile of the different classes of antibacterial agents (penicillins, β-lactam, β-lactamase inhibitor combinations, cephems, monobactams, carbapenems, lipopeptides, aminoglycosides, and fluoroquinolones) from different clinical specimens (various secretions, urine culture, blood culture, and catheter tip) were used in this study. From these samples, 54 (class 1) species were confirmed of being standard multidrug-resistant phenotype (resistant to more than one class of antibacterial agent), and 44 (class 2) were sensitive to all classes of antibacterial agents tested.

### 2.2. Maintenance of strains

After confirmation of the etiological and phenotypic resistant patterns of species, their identifications were stored for 18 months (January 2013 to July 2014) in tubes of nutrient agar with eppendorffs (HIMEDIA) and kept in the dark during the experiments. The cultures were renewed every 90 days. For the study, species were later removed from the culture collection of the Laboratory of Mycobacteria, Federal University of Rio Grande of Norte, Brazil.

### 2.3. Etiologic identification of strains

The samples were phenotypically identified using methods published by the American Society of Microbiology [15].

### 2.4. Phenotypic identification of bacterial resistant

The agar disk diffusion technique was used to assess susceptibility to different classes of antimicrobials for confirmation. Test-resistant patterns such as combined disk and enzyme block were used for strains that showed reduced susceptibility halo of the third-generation cephalosporins or carbapenems. *P. aeruginosa* ATCC27853 was used for control and validation tests.

### 2.5. NIR spectroscopy

Bacterial species were transplanted in studies of stock for nutrient Agar (HIMEDIA) plate and incubated for 24 h in bacteriological incubator (35 °C). Each NIR (spectral resolution of 8 $cm^{-1}$,in duplicate) spectra were directly acquired in the reflectance mode, using a miniature Fourier-transform scanning spectrometer from ARCspectro ANIR (Neuchâtel, Switzerland). The portable NIR device uses an InGaAs photodiode (900–2600 nm) and the reflected light was directed to the spectrometer via a bundle of optical fibers (model R600-7-VIS-125 F, Ocean Optics, USA) linked to the probe end and the data acquisition and analyses were carried out by ARCspectro ANIR 1.64 software. Absorbance spectra of samples of *P. aeruginosa* were obtained from the spectrum of the reflectance standard (Labsphere, 8750) which was used in the background. disposable syringes (1 mL) were used to place the sample on an aluminum plate (0.1 mm thickness) and then the reflectance probe was positioned under each sample, which was in a petri dish containing nutrient agar. The probe reflectance was washed with ethanol (70% v/v) and dried with paper after each sample. Spectral measurements were done in an acclimatized room under a controlled temperature of 22 °C, 60% relative air humidity, and samples were allowed to equilibrate to this temperature before analysis.

### 2.6. Multivariate analysis

The data import, pre-treatment, and construction of chemometric classification models (SPA-LDA and GA-LDA) were implemented in MATLAB R2014a software (Mathworks Inc, Natick, MA, USA). NIR spectra were pre-processed by Savitzky–Golay smoothing with different windows (3, 5, 7, and 15), first polynomial order, derivatisation of first/second derivatives, and multiplicative scatter correction (MSC). Mean centering was applied to all spectra before performing variable subset selection and calibration. For SPA-LDA and GA-LDA model, the samples were divided into training, validation, and prediction sets by applying the classic Kennard–Stone (KS) uniform sampling algorithm to the NIR spectra [16]. Sample numbers in each set are presented in Table 1. Training samples were used in the modeling procedure (including variable selection for LDA), whereas the prediction set was only used in the final evaluation of the classification. The optimum number of variables for SPA-LDA and GA-LDA was determined from the minimum cost function G calculated for a given validation dataset:

$$G = \frac{1}{N_V}\sum_{n=1}^{N_V} g_{n,} \quad (1)$$

**Table 1**
Number of training, validation, and test spectra in each category.

| Category | Set training | Validation | Test |
|---|---|---|---|
| (1) *P. aeruginosa* resistant | 34 | 10 | 10 |
| (2) *P. aeruginosa* sensitive | 24 | 10 | 10 |
| Total | 58 | 20 | 20 |