Differentiation of Environmental Bacteria Using Hyperspectral Imaging Technology And Multivariate Analysis

Aline Rodrigues Gama[#1], Venine Prado Saêta[#2], Raylane Pereira Gomes[#3], José Daniel Gonçalves Vieira[#4], Monica Santigo Barbosa[#5], Luiz Artur Mendes Bataus*[6], Arlindo Rodrigues Galvão Filho[+7], Clarimar José Coelho[+8], Lilian Carla Carneiro[#9]

[#]*Institute of Tropical Pathology and Public Health, Federal University of Goiás, 235 street, Goiânia, GO, Brazil*

[*]*Institute of Biological Sciences 2, Federal University of Goiás, Esperança avenue, Goiânia, GO, Brazil*
[+]*Engineering school, Pontifical Catholic University of Goiás, avenue Universitária, Goiânia, GO, Brazil.*

Abstract *Efficiency in the identification of bacterial isolates is of paramount importance in clinical microbiology for the correct diagnosis of infections and their subsequent treatment. It is also highly required in environmental microbiology to identify new isolates and the knowledge of the biodiversity of an environment. This raises the need for the development of identification methodologies, which make this practice accessible, safe, and at lower costs, with the emergence, in this sense, of the use of hyperspectral imaging technologies (HSI). In this study, we investigated the potential of HSI in differentiating different bacterial genera. For this, bacterial isolates obtained from water samples were previously sequenced and identified based on the 16S rRNA gene and then grown in Petri dishes containing culture medium. The HSI of each isolate was acquired using an short infrared wave imaging system (SWIR) at the SisuCHEMA workstation. Principal component analyzes showed a clear differentiation between genera and bacterial species. To transform these differentiation assays into an applicable identification methodology, it is necessary to build robust databases capable of covering the wide range of bacterial diversity. The primary results demonstrate that HSI is an objective tool and widely capable of promoting improvements in the identification of bacteria.*

Keywords — *Microbiology; SisuCHEMA; SWIR; biotechnology.*

I. INTRODUCTION

Infectious diseases are the major causes of death reported worldwide; for this reason, the development of rapid and accurate diagnostic methods is an urgent need [1]. The initial step for the execution of a definitive and correct antimicrobial chemotherapy is the microbiological diagnosis that determines the pathogen involved in the infection and the resistance profile observed in it [2,3].

In addition to the unquestionable relevance of identifying clinical isolates, the identification of bacteria of environmental origin is equally important, since it makes it possible to define new isolates, detect contaminants, analyze the quality of different environments and even knowledge of bacterial biodiversity in different niches; enabling the expansion of bioprospecting studies and consequent biotechnological applications [4].

The methods currently used for bacterial identification are based on culture. Although they are recognized as the gold standard, they have important deficiencies, such as the high demand for execution time and difficult work, Therefore, they are not suitable for the detection of the pathogen in time skillful [5]. In addition, manual screening for bacterial identification is susceptible to human errors, since genetically similar microorganisms can present important phenotypic differences [6]. The long incubation time required for bacterial culture prevents rapid and appropriate therapeutic decision making, thus contributing to the spread of infectious diseases and encouraging the misuse of antibiotics, culminating in the development of resistance [1].

Recent advances in DNA sequencing technologies allow the use of genomic sequences to accurately classify and identify members of Bacteria and Archaea [7]. The main molecular marker for bacterial classification/identification is the 16S rRNA gene. This 1500 base pair gene encodes the DNA that transcribes the catalytic RNA, which is part of the ribosomal 30S subunit. The use of the 16S rRNA sequence is a highly acceptable medium for bacterial identification. However, these molecular methods still have high costs and are only carried out using their equipment [2].

In view of this reality, the hyperspectral image (Hyperspectral Imaging - HSI), has been used as a new approach for the detection and identification of microorganisms present in different environments. The possibility of using HSI in the detection of genera and microbial species occurs through spectral and spatial assays, from a combination of spectroscopy and conventional images, based on the observed biochemical standards. This same tool has

been used successfully to analyze the microbiological quality of raw materials and products in the agricultural, food, and pharmaceutical industries [8-10].

The images obtained using a camera for the acquisition of HIS are collected according to the wavelength, resulting in a hypercube composed of a pile of formed images. At any wavelength, the image is formed by several pixels, and each pixel represents a spectrum containing specific chemical information. From the analysis of wavelengths, chemical differences between samples can be observed. To reduce the high dimensionality of the generated data and display differences in composition in an image, multivariate analysis tools such as Principal Component Analysis (PCA) are used [9].

The objective of this study was to investigate the potential of HSI generated in short infrared waves (Short Wave Infrared - SWIR) in the differentiation of various bacterial species.

II. MATERIALS AND METHODS

A. Bacterial samples

Bacterial isolates obtained in drinking water belonging to the genera *Acinetobacter*, *Klebsiella*, *Pseudomonas*, and *Serratia* were classified according to their biochemical characteristics, based on their generated HSI. These isolates were previously identified by sequencing the 16S rRNA gene, and the sequences were deposited in the GeneBank NCBI.

B. Sample preparation

After culturing for 24 hours on Nutrient Agar, all bacterial strains were resuspended in sterile saline (0.85% NaCl), standardized with turbidity compatible with the Mac Farland 0.5 scale (1x106 CFU / mL). A swab was moistened in this bacterial suspension, and it was sown on a Nutrient Agar plate by scanning technique. Two plates are being prepared for each bacterial strain described here. All the bacterial plaques were incubated in a bacteriological incubator for 24 hours at 30 ºC. After that, the images of the plates containing the cultures were acquired in their equipment.

C. Image acquisition and spectral extraction

HSI was obtained using a sisuCHEMA workstation (Specim, Spectral Imaging Ltd., Oulu, Finland), using a camera capable of capturing short wave infrared images (Short Wave Infrared - SWIR), using the Evince software 2.32.0. In addition to the images obtained in SWIR, the sisuCHEMA platform is also capable of capturing high-resolution images, which provide detailed information about the sample's chemical components.

D. SWIR image analysis

The principal component analysis (Principal Component Analysis - PCA) was applied with a centralized average over the entire image. The PCA can reduce the high dimensionality of data, decomposing interrelated variables into a new set of coordinates (PCs). In this study, wavelengths <1004, at the beginning of the spectrum, did not contain differentiating chemical information and were excluded from the data set, resulting in better classification of pixels.

E. Construction of phylogenetic tree

The phylogenetic tree for evaluating the evolutionary proximity relationship between species of the genus Acinetobacter was built using the software MEGA X 10.1. The distance matrix and the construction method were calculated by the algorithms: Jukes-Cantor and Neighbor-Joining, with a 500 repetition bootstrap.

III. RESULTS AND DISCUSSION

Acinetobacter spp., *Klebsiella* spp., *Pseudomonas* spp. And *Serratia* spp. were randomly selected for the classification analysis at the gender level (Fig 1). The PCA analysis of the images of the bacterial strains showed a clear separation between them. There is, however, an overlap of Acinetobacter spp., possibly not caused by the similarity between the biochemical characteristics, but by the "grooves" observed in the third plate produced during the sowing of this bacterium, leaving this area with a lower reflectance index.

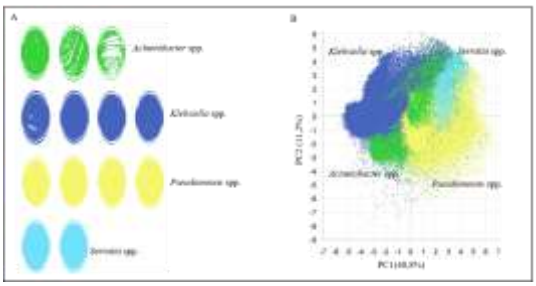

Fig 1 (a) Scoring graph of *Acinetobacter* spp., *Klebsiella* spp., *Pseudomonas* spp. And Serratia spp. showing the distinction between the four genres based on colors. (b) Corresponding score graph (PC1 x PC2) shows groups of distinct colored pixels according to the color scale in (a)

After evident differentiation between the aforementioned genera, an analysis was carried out to verify the separation between the species belonging to each genus. Fig 2 shows the analysis performed for *Acinetobacter* spp. The analysis of the images of the three species of *Acinetobacter* showed the separation of *A. lactucae* in relation to *A. calcoaceticus* and *A. vivianii* (Fig 2b; 2c).

However, the overlap between *A. calcoaceticus* and *A. vivianii* is noticeable. To justify this, a phylogenetic tree was built based on the sequences of the 16S rRNA gene and thus to verify the evolutionary proximity relationship between species (Fig 2d). Through this phylogeny, it was possible to observe that the species *A. calcoaceticus* and *A. vivianii* superimposed on the PCA graph, are separated in the same clade showing a higher degree of kinship and consequently more shared characteristics.

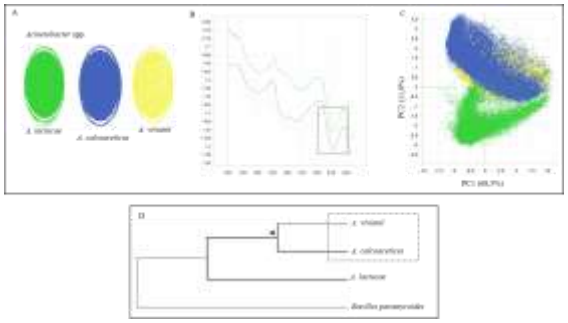

Fig 2 (a) Punctuation chart of three species of the genus *Acinetobacter* showing the distinction by colors. (b) Spectra for the images of the species of *Acinetobacter* showing overlapping regions. (c) Corresponding scoring graph showing the separation and overlap of clusters. (d) Phylogenetic tree showing the clade formed by *A. vivianii* and *A. calcoaceticus*

Based on the spectra and clusters formed from the PCA of *Klebsiella* spp. (Fig 3) it can be seen that the two isolates of *K. pneumoniae* and *K. variicola* are in a high degree of overlap while *K. michiganensis* forms a visibly separate cluster.

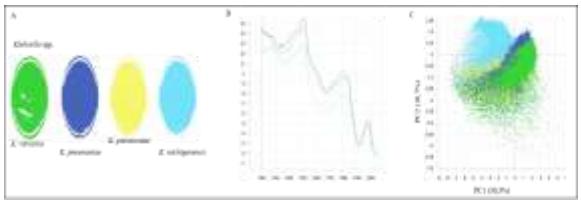

Fig 3 (a) Scoring graph of four species of the genus *Klebsiella*, showing the distinction by color. (b) Spectra for the images of the species of *Klebsiella*. (c) Corresponding scoring graph showing the separation and overlap of clusters

K. variicola and *K. pneumoniae* have overlapping biochemical and phenotypic characteristics, considering that they belong to the same complex called the pneumoniae complex. *K. variicola* until 2004 was identified as *K. pneumoniae*, and after numerous biochemical and genetic tests, *K. variicola* was proposed as a new bacterial species. As the bacterial differentiation from HSI is based on biochemical characteristics, it was not possible to establish significant differences that pointed out *K. variicola* as an organism distinct from *K. pneumoniae*, given these data, it is suggested that the correct differentiation of bacteria from *pneumoniae* complex, should be performed by associating HSI with different molecular, genomic and proteomic tools [11,12].

The HSI images of *Pseudomonas* spp. were analyzed. The results are shown in Fig 4, where the differentiation of *P. plecoglossicida*, present in a separate cluster, can be observed in relation to *P. entomophila* and *P. taiwanensis* presented in the overlap.

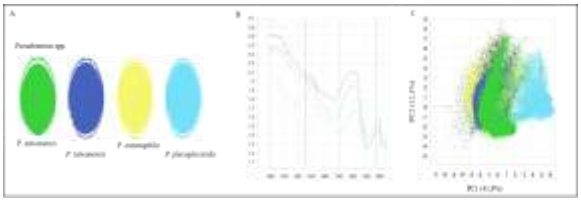

Fig 4 (a) Scoring graph of four species of the genus *Pseudomonas* showing the distinction by color. (b) Spectra for the images of the species of *Pseudomonas*. (c) Corresponding scoring graph showing the separation and overlap of clusters

When investigating the characteristics of *P. entomophila* and *P. taiwanensis* to understand the overlap between them, and consequent biochemical similarity, it was found that both *P. entomophila* and *P. taiwanensis* have a high degree of pathogenicity to flies of the genus *Drosophila*.

The insecticidal activity manifested by such bacteria is due to the expression of a gene known as the insecticidal toxin gene (tccC). The DNA sequence of the tccC gene encodes 980 amino acids. The alignment of the sequence of these amino acids from *P. taiwanensis* showed a high degree of identity with the insecticidal toxin from *P. entomophila*. The expression of this conclusively similar insecticide toxin between species may have contributed to the spectral similarity between *P. entomophila* and *P. taiwanensis* [13,14].

Two species of the genus *Serratia* were evaluated (Fig 5). Based on the generated spectra, as well as on the separation of the clusters, a noticeable differentiation between the species *S. marcescens* and *S. nematodiphila* is revealed. *S. nematodiphila* was described in 2009 as a new species, because despite having great similarity with species of the genus *Serratia*, it exhibits its own phenotypic, biochemical and molecular characteristics, not evidenced in other species, such as the ability to emit fluorescence and establish an association symbiotic with nematodes [15].

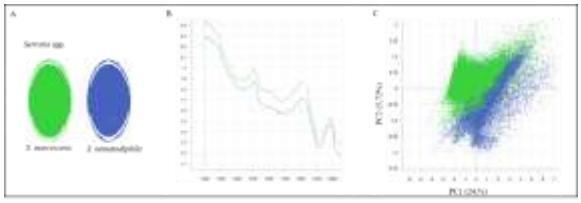

Fig 5 (a) Graph of scoring of two species of the genus *Serratia* shows the distinction by color. (b) Spectra for the images of the species of *Serratia* (c) Graph of the corresponding score showing the separation and overlap of clusters

The classification results obtained from the samples demonstrated high levels of precision in differentiation. In addition, it was possible to perceive that the hyperspectral analyzes performed, showed sensitivity to the general structure and the biochemical composition of the bacterial cells, being, therefore, possible to observe differences between bacteria with similar genotypes [16].

Different studies have been carried out with the proposal to differentiate and identify microbial isolates in different environments using HSI. Park et al. [17] performed the classification of five different *Salmonella* serotypes, obtaining an average 84% accuracy. In this work, the authors reinforce the importance of building spectral libraries with varied bacterial species to assist in detecting pathogens in the food industry for birds.

Seo et al. [18] developed classification models for *Staphylococcus five species*, using hyperspectral microscopic imaging, and highlighted the use of HSI in the presumptive screening of foodborne pathogenic bacteria. Foca et al. [19] evaluated spectral and hyperspectral techniques the bacterial contamination by *Lactobacillus curvatus* and *L. sakei* in ham and observed that the techniques used can be effective for recognizing bacterial contamination and still recognize species to which the bacteria belong.

Despite considerable applications, HSI microbial differentiation and identification present important challenges such as efficient standardization of culture conditions and sample preparation. Cultivation conditions that include growth time, type of culture medium, temperature, oxygen availability can have a strong impact on the manifested phenotypic characteristics and, therefore, on the spectra obtained.

Another factor seen as a difficulty with the HSI classification method is that it is an approach based on identification libraries and requires validated spectral databases, covering a wide range of bacterial taxonomies. However, such bases are not yet available, which constitutes the biggest challenge in transferring this technology to laboratory practice. However, with the improvement of this method, we hope to obtain solutions to these problems and possibly apply HSI in environmental, clinical, and food microbiology [16].

IV. CONCLUSIONS

In the present study, HSI was used to differentiate bacteria isolated from water. The proposed method was effective. It involved the culture of bacterial isolates followed by the pre-processing of the images at the SisuCHEMA station and subsequent application of multivariate analyzes to differentiate the HSI obtained.

The experiences acquired in this study suggest that the HSI analysis can enable the efficient differentiation of different bacterial species, however, to be able to identify bacteria based on this technology, it is necessary to standardize the cultures, and especially the construction of a database that can cover the wide existing bacterial variety. But supported by the level of differentiation achieved so far, we are confident that this study will initiate new research that may make it possible to conduct this technology in different microbiology areas.

ACKNOWLEDGMENT

We are grateful to the Scientific Computing Laboratory Pontifical Catholic University of Goiás, for carrying out the acquisition of hyperspectral images and enabling the entire analysis process. We would also like to thank the Coordination for the Improvement of Higher Education Personnel (CAPES) for providing a scholarship.

REFERENCES

[1] Noiphung J, Laiwattanapaisal W. "*Multifunctional Paper-Based Analytical Device for In Situ Cultivation and Screening of Escherichia coli Infections*". Sci Rep;9(1):1–10 (2019).

[2] Srinivasan R, Karaoz U, Volegova M, MacKichan J, Kato-Maeda M, Miller S, et al. "*Use of 16S rRNA gene to identify a broad range of clinically relevant bacterial pathogens*". PLoS One; 10(2):e0117617 (2015)

[3] Özenci V, Rossolini GM. "*Rapid microbial identification and antimicrobial susceptibility testing to drive better patient care: an evolving scenario*". J Antimicrob Chemother; 74:I2–5 (2019).

[4] Pandey A, Jain R, Sharma A, Dhakar K, Kaira GS, Rahi P, Dhyani A, Pandey N, Adhikari P, Shouche YS. 16S rRNA gene sequencing and MALDI-TOF mass spectrometry-based comparative assessment and bioprospection of psychrotolerant bacteria isolated from high altitudes under the mountain ecosystem. SN Appl Sci;1(3):1–12 (2019).

[5] Feng YZ, Yu W, Chen W, Peng KK, Jia GF. "*Invasive weed optimization for optimizing one-agar-for-all classification of bacterial colonies based on hyperspectral imaging. Sensors Actuators*," B Chem; 269:264–70 (2018).

[6] Tillman GE, Wasilenko JL, Simmons M, Lauze TA, Minicozzi J, Oakley BB, Narang N, Fratamico P, Cray WC. *"Isolation of Shiga Toxin-Producing Escherichia coli Serogroups* "O26, O45, O103, O111, O121, and O145 from Ground Beef Using Modified Rainbow Agar and Post–Immunomagnetic Separation Acid Treatment. J Food Prot; 75(9):1548–54 (2012).

[7] Yoon SH, Ha SM, Kwon S, Lim J, Kim Y, Seo H, Chun J. "*Introducing EzBioCloud: A taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies*." Int J Syst Evol Microbiol; 67(5):1613–7 (2017).

[8] Gowen AA, Feng Y, Gaston E, Valdramidis V. "*Recent applications of hyperspectral imaging in microbiology".* Talanta;137:43–54 (2015).

[9] Sandasi M, Vermaak I, Chen W, Viljoen AM. "*Hyperspectral imaging and chemometric modeling of Echinacea - A novel approach in the quality control of herbal medicines*". Molecules;19(9):13104–21 (2014).

[10] Yoon SC, Lawrence KC, Park B. "*Automatic Counting and Classification of Bacterial Colonies Using Hyperspectral Imaging*". Food Bioprocess Technol; 8(10):2047–65 (2015).

[11] Rosenblueth M, Martínez L, Silva J, Martínez-Romero E. Klebsiella variicola, A Novel Species with Clinical and Plant-Associated Isolates. Syst Appl Microbiol; 27(1):27–35 (2004).

[12] Rodríguez-Medina N, Barrios-Camacho H, Duran-Bedolla J, Garza-Ramos U. Klebsiella variicola: an emerging pathogen in humans. Emerg Microbes Infect; 8(1):973–88 (2019).

[13] Liu JR, Lin YD, Chang ST, Zeng YF, Wang SL. *"Molecular Cloning and Characterization of an Insecticidal Toxin from Pseudomonas taiwanensis"*. J Agric Food Chem; 8;58(23):12343–9 (2010).

[14] Chen WJ, Hsieh FC, Hsu FC, Tasy YF, Liu JR, Shih MC. "*Characterization of an Insecticidal Toxin and Pathogenicity of Pseudomonas taiwanensis against Insects*". Schneider DS, editor. PLoS Pathog; 21;10(8):e1004288 (2014).

[15] Zhang CX, Yang SY, Xu MX, Sun J, Liu H, Liu JR. Serratia nematodiphila sp. nov., associated symbiotically with the entomopathogenic nematode Heterorhabditidoides chongmingensis (Rhabditida: Rhabditidae). Int J Syst Evol Microbiol; 1;59(7):1603–8 (2009).

[16] Lasch P, Stämmler M, Zhang M, Baranska M, Bosch A, Majzner K. "*FT-IR Hyperspectral Imaging and Artificial Neural Network Analysis for Identification of Pathogenic Bacteria"*. Anal Chem; 90(15):8896–904 (2018).

[17] Park B, Seo Y, Eady M, Yoon S-C, Hinton, Jr. A, Lawrence KC. "*Classification of Salmonella Serotypes with Hyperspectral Microscope Imagery*". Ann Clin Pathol; 5(2):1108 (2017).

[18] Seo Y, Park B, Hinton A, Yoon SC, Lawrence KC. "*Identification of Staphylococcus species with hyperspectral microscope imaging and classification algorithms*". J Food Meas Charact; 10(2):253–63 (2016).

[19] Foca G, Ferrari C, Ulrici A, Sciutto G, Prati S, Morandi S, Brasca M, Lavermicocca P, Lanteri S. Oliveri P. "*The potential of spectral and hyperspectral imaging techniques for bacterial detection in food: A case study on lactic acid bacteria*". Talanta; 153:111–9 (2016).

[20] Dr. Badr Al ALI , Dr. Ahmed QARA ALI, Rami HAMMOD, *"Chemical Compounds And the Vital Importance of Organic Extract of Bacillus Polymyxa Marine Bacteria"* SSRG International Journal of Agriculture & Environmental Science 6.5 (2019): 1-3.