



**UNIVERSITY OF NAIROBI**

**School of Biological and Physical Sciences**

**LABEL-FREE RAMAN SPECTROMETRIC DETECTION AND DIFFERENTIATION  
BETWEEN BACTERIA *STAPHYLOCOCCUS AUREUS* AND *NEISSERIA  
GONORRHOEAE***

**MSc. Thesis**

**By**

**Mung'ei Consolata Kwamboka**

**B. ED (Sc) (Hons)**

**I56/7699/2017**

**Thesis submitted for examination in partial fulfillment of requirements for the award of  
Master of Science Degree in Physics of the University of Nairobi.**

**©August 2023**

# Declaration

I declare that this thesis is my original idea and has not been submitted elsewhere for examination, award of a degree, or publication. Where other people's work, or my work has been used, this has properly been acknowledged and referenced in accordance with the University of Nairobi's requirements.

Consolata Kwamboka Mung'ei

I56/7699/2017

Department of Physics

University of Nairobi

Signature:



Date: 12/09/2023

This thesis has been submitted with our approval as University supervisors:

**Supervisors;**

**Signature**

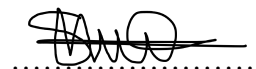
**Date**

Dr. Birech Zephania.  
Department of Physics,  
University of Nairobi,  
P.O BOX 30197-00100  
Nairobi, Kenya  
Email; birech@uonbi

  
.....

26/09/2023  
.....

Dr. Muuo Nzou.  
Principal Research Scientist,  
NUITM-KEMRI,  
P.O BOX 19993-00202,  
Nairobi, Kenya  
Email; [nzoumuuo@gmail.com](mailto:nzoumuuo@gmail.com)

  
.....

26/09/2023  
.....

## **Dedication**

*I dedicate this thesis to my two supervisors Dr. Z. Birech and Dr. Muuo who have taught me to hold up onto hope and hard work.*

*And most sincere dedication to my dear husband whom I look up to, and as Google Scholar says, 'am standing on the shoulders of the giant'.*

## **Acknowledgment**

My sincere appreciation to the Almighty God for the sufficient graces granted to me, the desire to work hard and achieve my goals, and most importantly, the good health that enabled me to trail through and seek knowledge up to this far.

My heartfelt gratitude to Dr. Zephania Birech of the Physics Department of the University of Nairobi for recommending this research area, he facilitated and familiarized me with the requirements of this research area, and made me understand all my expectations throughout the journey. I as well wish to recognize Prof. Muuo Nzou, principal research scientist at NUITM-KEMRI

for his endless efforts and support in ensuring the work was completed perfectly and on time. I cannot forget Mr. Omucheni, the Chief lab technician in the Laser lab in the Department of Physics who offered me the opportunity to learn the Raman spectroscopy machine and supported me in all the scheduled hours I could sit in the lab collecting data for my research.

I genuinely wish to thank the Kenya Medical Research Institute (NUITM-KEMRI) through my supervisor Prof. Muuo Nzou for freely providing me with the samples of the bacteria I needed and continuously following up for any assistance that I needed.

My most heartfelt acknowledgement goes to the Deutscher Akademischer Austauschdienst German Academic Exchange Programme (DAAD) for offering me the opportunity to pursue this Master of Science program through a scholarship.

Finally, I cannot forget my colleagues Ms. Annah Moraa, Ms. Teresa Mesoko, and Ms. Caroline Chepoo who rendered me great support and motivation. Also, my dad Mr. C. Mungei whose prayers have made me survive through the obstacles, my brothers and sisters who have been constantly tracking my progress, my loving and supportive husband Joshua, on whose shoulders I stood to see myself this far, and lastly my charming little daughters and son who gave me hope of waking up every day and push this dream ahead.

## Abstract

This work reports on the application of Raman spectroscopy in distinguishing rapidly between *Staphylococcus aureus* and *Neisseria gonorrhoeae* bacterial strains. The method was also explored in segregating between drug-responsive and drug-resistant *Staphylococcus aureus* strains. Currently, such tasks take several days and involve the use of labels in methods such as Polymerase Chain Reaction (PCR), Verigene Gram-positive blood culture (BC-GP) assay, and Alere's BinaxNOW. The samples were applied onto conductive silver paste-smeared glass slides then excited with a 785 nm laser and the Raman scattered radiation was collected, recorded, and analyzed. The Raman spectral data analyses were done using ANOVA and PCA which were able to reveal a distinction between *Staphylococcus aureus* and *Neisseria gonorrhoeae*. The bands that were uniquely identified to *Staphylococcus aureus* had their wavenumbers centered at 800 $\text{cm}^{-1}$ , 846 $\text{cm}^{-1}$ , and 1028 $\text{cm}^{-1}$ . These bands displayed high-intensity variations in ANOVA and high-loading signals in PCA. They are attributed to Tyrosine (800 $\text{cm}^{-1}$ ) and proteins (846, 1028 $\text{cm}^{-1}$ ), and their vibrational bands were tentatively assigned to C-N stretch, ring breathing, and C-C breathing respectively. Raman spectra of *Neisseria gonorrhoeae* exhibited unique bands centered at 635 $\text{cm}^{-1}$ , 710  $\text{cm}^{-1}$ , 941  $\text{cm}^{-1}$ , 1232  $\text{cm}^{-1}$  and 1320  $\text{cm}^{-1}$  which were attributed to Tyrosine(635  $\text{cm}^{-1}$ ), Polysaccharides (710  $\text{cm}^{-1}$ ), protein (941  $\text{cm}^{-1}$ ), Amide III, adenine, polyadenine (1232  $\text{cm}^{-1}$ ) and DNA (1320  $\text{cm}^{-1}$ ). Their vibrational assignments are the aromatic ring skeletal, CH<sub>2</sub> Rock, C-N ring, and NH<sub>2</sub> respectively. These two sets of bands can be used as the biomarker bands for *Staphylococcus aureus* and *Neisseria gonorrhoeae*. For the drug-resistant and the drug-responsive strains of *Staphylococcus aureus* bacteria, the peaks that were found to have high variations in terms of intensities in ANOVA and high loading signals in PCA and associated with the drug resistant strain were centered at 720 $\text{cm}^{-1}$ , 950 $\text{cm}^{-1}$ , and 1320 $\text{cm}^{-1}$ . These bands are attributed to polysaccharides (720 $\text{cm}^{-1}$ ), proteins (950 $\text{cm}^{-1}$ ), and DNA (1320 $\text{cm}^{-1}$ ), while their vibrational assignments are CH<sub>2</sub> Rock, C-C ring breathing, and NH<sub>2</sub> respectively. Similarly, the bands identified to the drug-responsive strain are 805 $\text{cm}^{-1}$  (proteins), 1035 $\text{cm}^{-1}$  (proteins), and 1443 $\text{cm}^{-1}$  (phospholipids, deformation mode of proteins), and vibrationally assigned to C-N stretch, C-C ring breathing, and CH<sub>2</sub> in that order. These results indicate that the Raman spectroscopy method is capable of rapidly identifying the biomarker bands that can be used in the diagnosis of bacteria strains in hospitals, hence the time used in diagnosis is significantly reduced.

# Table of Contents

|   |      |
|---|------|
| Declaration.....  | i    |
| Dedication.....   | ii   |
| Acknowledgment.....   | iii  |
| Table of Contents.....  | v    |
| List of Abbreviations and Acronyms.....                                   | vii  |
| List of Tables.....   | viii |
| List of Figures.....  | ix   |
| CHAPTER I.....  | 1    |
| 1 INTRODUCTION.....   | 1    |
| 1.1 Background.....   | 1    |
| 1.2 Statement of the Problem.....   | 3    |
| 1.3 Objectives.....   | 3    |
| 1.3.1 Main objective.....   | 3    |
| 1.3.2 Specific objectives.....  | 3    |
| 1.4 Justification and Significance of the Study.....                      | 4    |
| CHAPTER II.....   | 5    |
| 2 LITERATURE REVIEW.....  | 5    |
| 2.1 Overview.....   | 5    |
| 2.2 Raman Spectroscopy in Bacteria Detection and Identification.....      | 5    |
| 2.3 Raman Spectroscopy in Detection of Antibiotic Resistant Bacteria..... | 7    |
| 2.4 Chemometrics with Raman Spectroscopy in Analysis of Bacteria.....     | 8    |
| CHAPTER III.....  | 10   |
| 3 THEORETICAL BACKGROUND.....   | 10   |
| 3.1 Overview.....   | 10   |
| 3.2 Principles of Raman Scattering.....                                   | 10   |
| 3.2.1 The classical approach.....   | 10   |
| 3.2.2 Quantum description.....  | 12   |
| 3.3 Intensity of the Raman scattered light.....                           | 13   |
| 3.4 Chemometric techniques in Raman spectroscopy.....                     | 14   |
| 3.5 Spectral Data Processing.....   | 15   |
| 3.6 Supervised versus Unsupervised Learning.....                          | 15   |
| 3.7 Principal Component Analysis (PCA).....                               | 15   |

|                  |  |    |
|------------------|--|----|
| 3.8              | Analysis Of Variance (ANOVA).....  | 18 |
| CHAPTER IV.....  |  | 22 |
| 4                | MATERIALS AND METHODS .....  | 22 |
| 4.1              | Overview .....   | 22 |
| 4.2              | Raman Spectroscopy Instrumentation.....  | 22 |
| 4.3              | Preparation of Bacteria Samples .....  | 24 |
| 4.4              | Raman Analysis Procedure .....   | 24 |
| 4.5              | Statistical Analysis of Raman Spectra .....  | 24 |
| 5                | RESULTS AND DISCUSSION .....   | 25 |
| 5.1              | Overview .....   | 25 |
| 5.2              | Characteristic Raman spectra of <i>Staphylococcus aureus</i> and <i>Neisseria gonorrhoeae</i> bacteria ..... | 25 |
| 5.3              | Raman spectroscopy of antibiotic responsive and non-responsive <i>Staphylococcus aureus</i> ;<br>30          |    |
| CHAPTER VI.....  |  | 35 |
| 6                | CONCLUSION RECOMMENDATIONS AND PROSPECTS.....  | 35 |
| 6.1              | Summary and Conclusions.....   | 35 |
| 6.2              | Recommendations and Future Prospects .....   | 36 |
| REFERENCES.....  |  | 37 |
| APPENDICES ..... |  | 40 |

## **List of Abbreviations and Acronyms**

|      |                                     |
|------|-------------------------------------|
| ANN  | Artificial Neural Network           |
| AST  | Antibiotic Susceptibility Test      |
| CCD  | Coupled Channel Detector            |
| DNA  | Deoxyribonucleic acid               |
| EM   | Electro Magnetic                    |
| ET   | Exposure Time                       |
| FT   | Fourier Transform                   |
| LOD  | Limit of Detection                  |
| NAD  | Nicotinamide adenine dinucleotide   |
| NIR  | Near Infrared                       |
| PCA  | Principle Component Analysis        |
| PCR  | Polymerase Chain Reaction           |
| RNA  | Ribonucleic acid                    |
| SERS | Surface Enhanced Raman Spectroscopy |
| SVM  | Support Vector Machine              |



## List of Tables

|   |    |
|---|----|
| <b>Table 3.1.</b> A schematic data for a one-way ANOVA.....   | 19 |
| <b>Table 3.2.</b> A schematic data for a two-way ANOVA.....   | 20 |
| <b>Table 5.1:</b> Peak and vibrational assignment of Raman spectra of <b>Staphylococcus aureus</b> and <b>Neisseria gonorrhoeae</b> bacteria strains (González-Solís et al. 2014; Liu et al. 2017; Lu et al. 2013) ....                   | 29 |
| <b>Table 5.2:</b> <i>Peak and vibrational assignment of Raman spectra of the drug responsive Staphylococcus aureus</i> bacteria strain and its drug resistant strain. (González-Solís et al. 2014; Liu et al. 2017; Lu et al. 2013) ..... | 32 |

## List of Figures

|  |    |
|--|----|
| <b>Figure 3.2:1:</b> A model of the Rayleigh, Stokes-Raman and Anti-stokes-Raman scattering (Tarcea et al. 2008).....  | 13 |
| <b>Figure 3.7:1:</b> Fundamentals of PCA presented graphically.(Biancolillo and Marini 2018) .....   | 18 |
| <b>Figure 3.8:2:</b> ANOVA plot of variance showing the activity of fatty acid receptors in three categories of microfluidic devices. It shows the Raman peaks at 1002, 1078, and 1582cm <sup>-1</sup> as the variance peaks (Zhang et al. 2019). .....                | 21 |
| <b>Figure 4.2:1:</b> Raman spectrum of silicon wafer using 785nm laser.....  | 22 |
| <b>Figure 4.2:2:</b> Schematic diagram of the set-up of the Raman spectroscopy, Courtesy of: <a href="http://cnx.org/">http://cnx.org/</a> 23  |    |
| <b>Figure 5.2:1:</b> A graph of the Raman peaks of drug responsive bacteria <i>Staphylococcus aureus</i> (SA) and <i>Neisseria gonorrhoeae</i> (NG) together with a plot of variance showing variance peaks between the two bacteria strains. ....                     | 26 |
| <b>Figure 5.2:2:</b> 3-D score plot for the combined Raman spectral data sets from <i>Staphylococcus aureus</i> and <i>Neisseria gonorrhoeae</i> bacteria strains. A distinct segregation between the data sets was observed with an explained variance of 58.4%. .... | 27 |
| <b>Figure 5.2:3:</b> A loadings plot showing plots of PC1 and PC2 together with bacteria <i>Staphylococcus aureus</i> and <i>Neisseria gonorrhoeae</i> .....   | 28 |
| <b>Figure 5.3:1:</b> A graph of the Raman peaks of drug responsive bacteria <i>Staphylococcus aureus</i> (Resp) and its drug resistant strain(Rest) together with a plot of variance showing variance peaks between the two bacteria strains.....                      | 30 |
| <b>Figure 5.3:2:</b> A 3-D score plot for <i>Staphylococcus aureus</i> drug responsive and resistant strains .....   | 31 |
| <b>Figure 5.3:3:</b> A loadings plot showing plots of PC1 and PC2 together with the drug responsive and resistant strains of bacteria <i>Staphylococcus aureus</i> . ....  | 32 |

# CHAPTER I

## 1 INTRODUCTION

### 1.1 Background

Raman spectroscopy is a method used in the detection and differentiation of elements in a sample. It depends on Raman scattering whereby photons incident to a sample lose or gain energy after interacting with vibrating molecules in the sample (Gardiner 1989). The shifts in energy can then be used to acquire important information about the molecular composition of the sample with a very high degree of precision (Juma 2018). Raman spectroscopy has gained a lot of interest in biomedicine compared to other methods, because of its direct and rapid procedure of detection and minimal sample preparation, and this means that it will take less time to diagnose the type of bacterial infection in a patient. In addition, the method does not involve the use of labels, so the composition of the sample is not interfered with, for this reason, further sample tests can be carried out (Boardman et al., 2016). Also, this method is more applicable in the study of both bacteria and viruses since these are the key disease-causing agents that lead to mortality among patients if not quickly and accurately detected (Ferguson and Brown, 1996).

The utilization of Raman spectroscopy especially in the field of biomedicine has drawn a lot of interest to various researchers. The method has been tried in the detection and diagnosis of disease-causing organisms like bacteria and viruses, the classification of the organisms in various families, and also in the study of their cell structure (Galvan and Yu 2018). Raman spectroscopy has also been utilized in the detection and classification of 17 bacteria species into gram-positive and gram-negative categories (Boardman et al., 2016). The research involved the separation of 17 bacterial species from culture using an algorithm that was based on barcodes. All the spectra were classified into 4 groups using the Hierarchical cluster analysis (HCA) as a statistical tool, then a subsequent separation of all the bacteria into individual groups using the Partial least-squares discriminant analysis (PLS-DA) technique. They also discovered the fact that the method could give different spectra of drug-susceptible bacteria during the antibiotic susceptibility test (AST). This research work by Boardman and coworkers involved some strenuous sample preparation procedures using labels, however, this was improved by Liu et al who used the Raman spectroscopy method in both label and label-free methods in pathogen detection. In the research, it was pointed out that the label-based detection method with Raman spectroscopy allowed highly sensitive and multiplexing bacteria detection results. Moreover, when compared with the label-

free method, the label-based method had increased preparation steps, reactant volumes, and time of analysis. In addition, the biological information about the cells of the organism is lost (Liu et al., 2017).

Raman spectroscopy has also been used in the differentiation between *Chlamydia trachomatis* and *Neisseria gonorrhoeae* (Chen, Premasiri, and Ziegler 2018). Though these two bacteria strains are both gram-negative and both cause sexually transmitted diseases, the method was able to differentiate between them by giving two distinct spectra (Chen, Premasiri, and Ziegler 2018).

The Raman spectroscopy method is venturing into another area in bacteria diagnosis, this is the detection and differentiation between antibiotic-responsive and antibiotic-resistant bacteria strains (Novelli-Rousseau et al., 2018). Since most bacteria strains change their structure or develop resistance to specific drugs once poorly prescribed, it is therefore important for medics to determine the nature of antibiotic susceptibility of these strains.

This research work explored the performance of Raman spectroscopy in rapid testing and distinguishing between bacteria *Staphylococcus aureus* and *Neisseria gonorrhoeae*. Its performance was also applied in the differentiation between the drug-responsive and drug-resistant strains of *Staphylococcus aureus*. This method can be a great boost in the diagnosis of bacteria in hospitals since in the traditional blood culture method, one has to wait up to 24-48 hours for a positive blood culture before the testing is done. This is averagely 37-40 hours of diagnosis (Boardman et al., 2016). Other methods like Verigene Gram-positive blood culture (BC-GP) assay which uses microfluidic cassettes and has identified up to 37 microorganisms and the Alere's BinaxNOW which engages immunochromatographic assays has been used to identify *E. coli*, *Streptococcus pneumonia*, methicillin-resistant *Staphylococcus aureus* and others, have reduced the length of time for testing although they still depend on the positive blood culture results (Boardman et al., 2016).

Polymerase Chain Reaction (PCR), is another method that does not depend on the results of positive blood culture but uses labels that are placed in the sample whereby the test results depend on the reaction between the sample and the metal label used (Liu et al., 2017). The limitations of this method are that it interferes with the nature of the sample (by adding labels) and thus no further tests can be done on the sample. Also, it is expensive to obtain the labels and involves a lengthy sample preparation process (Boardman et al., 2016).

In this thesis, the literature on Raman spectroscopy is captured in the second chapter. The areas where the method has been utilized in detection, differentiation, and thus diagnosis of disease-causing

pathogens. Detection of antibiotic resistance of these bacteria has also been shown together with the chemometric analytical techniques that help in the classification of samples. The basic principles of Raman spectroscopy involving absorption and emission of laser light causing molecular vibration of atoms are discussed in the third chapter together with the processing and interpretation of Raman spectra by the statistical tools ANOVA and PCA. Chapter Four shapes the methodology of Raman spectrometry and the operating conditions behind it, the process of obtaining and preparing samples, the safety measures kept in mind, and also the collection and pre-processing of spectral data are discussed. Finally, in the fifth chapter of results and discussion, the results of the differentiation of bacteria *Neisseria gonorrhoeae* and *staphylococcus aureus* and its drug-resistant strain are obtained and discussed with the help of ANOVA and PCA statistical tools. Their biomarker bands are stated then the summary and conclusion are given in chapter six.

## **1.2 Statement of the Problem**

The current methods for pathogenic bacteria diagnosis are destructive, time-consuming, require tedious sample preparation and, to some extent, produce inaccurate results. Therefore, they are of limited utility in rapid bacteria diagnostic, especially in cases of severe infections that require immediate diagnosis and specific drug prescription. Raman spectroscopy coupled with multivariate chemometric techniques has the potential to overcome these shortcomings and develop a calibration model to achieve accurate diagnosis of bacterial infections.

## **1.3 Objectives**

### **1.3.1 Main objective**

The main aim of this research work is to utilize Raman spectroscopy in the detection and differentiation between non-drug resistant and resistant bacterial strains *staphylococcus aureus* and *Neisseria gonorrhoeae*.

### **1.3.2 Specific objectives**

- (i) To distinguish using Raman spectroscopy between *staphylococcus aureus* and *Neisseria gonorrhoeae* bacteria and identify and assign both component-wise and vibrationally their biomarker Raman bands.

- (ii) To distinguish, using Raman spectroscopy, between drug-responsive and drug-resistant *Staphylococcus aureus* strains and determine biomarker bands for each.

#### **1.4 Justification and Significance of the Study**

The blood culture method, for decades now, has been the gold method for bacteria diagnosis. Despite its well-known accuracy in detection, there is the most serious backdrop that seems more costly, the time it takes for a single result of a test to be found. This means that for serious infections it remains to be a challenge.

The bold step taken by physicians to prescribe general antimicrobials renders the bacteria chances of developing resistance to antibiotics. This is due to lack of specific diagnosis as the physician awaits the development of positive blood for specific detection. This is a major limitation since drug-susceptible and resistant bacteria cannot be treated the same way and a physician needs to perform an antibiotic susceptibility test (AST) to determine the nature of the pathogens. Therefore, this work aims to develop a technique that can perform accurate and timely bacteria diagnosis, and can be able to demonstrate the ability to discriminate between the susceptible and drug-resistant pathogens.

## CHAPTER II

### 2 LITERATURE REVIEW

#### 2.1 Overview

In this chapter, the literature relating to Raman spectroscopy and other vibrational techniques in bacteria detection and identification is reviewed. The chapter also points out the utility of Raman spectroscopy in the detection and differentiation of anti-biotic responsive and resistant bacteria. Finally, it critiques the application of chemometrics as a tool in the analysis of Raman spectra.

#### 2.2 Raman Spectroscopy in Bacteria Detection and Identification

The use of Raman spectroscopy has significantly advanced in the last decade due to its detection capabilities. This has been enhanced by its high sensitivity, speed, portability, and a comparable low cost. Moreover, when applied in the investigation of samples placed in special nano-resonators, the efficiency of SERS displayed an increased limit of detection (LOD) of up to the order of  $10^{-18} \text{mol dm}^{-3}$  was attained (Rebrošová et al., 2017). Due to the sensitivity of Raman spectroscopy, its utility has been widely incorporated in biological, medical, environmental, and industrial sample analysis (Kudelski 2008). The technique is mainly used to identify a sample giving different Raman spectra with different Raman intensities per sample which imply different compositions of molecules in the sample. For instance, two strains of Gram-positive *Staphylococcus aureus* bacteria strain were investigated on bionanocomposites on chitosan-silver nanoparticles (Potara et al., 2011). It was demonstrated that the silver nanoparticles can be used on antibacterial agents and on plasmonic substrates for bacteria identification and in monitoring the biochemical changes in the bacteria cell wall via Raman spectroscopy

In biomedical applications, Raman spectroscopy is more reliable in the diagnosis of disease-causative pathogens. For instance, it has demonstrated greater potential in detecting and discriminating bacteria belonging to various families. The families of *Enterobacteriaceae* and *Acinetobacter* genus have been found to have differences in the Raman peaks which are stated to be due to biological differences (Espagnon et al., 2014). Also, a study on the characterization of bacterial and yeast microcolonies using Raman spectroscopy showed characteristic peaks at 664, 781, 808, 1095 and  $1569 \text{cm}^{-1}$ . These peaks were associated with the position and tendency of the activities of the microcolonies, and they were assigned to DNA and RNA nucleic acid bands (Espagnon et al. 2014). In addition, the more intense

peaks at 745 and 1126 $\text{cm}^{-1}$  were associated with proteins and have their vibrations as C-N and C-C stretch respectively (Espagnon et al. 2014). In the diagnosis of bacteria *Chlamydia trachomatis* and *Neisseria gonorrhoeae*, SERS signatures of the two sexually transmitted infectious bacteria are completely different even though they are both gram-negative bacteria (Galvan and Yu 2018). The Raman spectra of the two bacteria were obtained for classification both on gold (Au) and silver (Ag) substrates (Chen, Premasiri, and Ziegler 2018). *Neisseria gonorrhoeae* bacteria had similar spectral bands on both substrates while *Chlamydia trachomatis* had the distinct bands 900, 1200 and 1550 $\text{cm}^{-1}$  on Au substrate and bands 936, 1002, 1035, 1385, 1446 and 1600 $\text{cm}^{-1}$  on Ag substrate (Chen, Premasiri, and Ziegler 2018). This difference was exclusively attributed to the structure of the metal surface of the substrate used. Chen and co-workers were able to demonstrate how the Raman spectral signatures distinguish these two bacteria based on Purine metabolites i.e. adenine, guanine, xanthine, hypoxanthine, uric acid, guanosine (Chen, Premasiri, and Ziegler 2018). The presence or absence of these enzymes in the purine metabolism of the organism plays a major role in the determination of the nature of the organism (W. R. Premasiri et al. 2005). For instance, *Neisseria gonorrhoeae* spectra would show strains of adenine, guanine and a few  $\text{NAD}^+$  as observed in the Raman spectra, their peak assignments are 734, 664 and 1030 $\text{cm}^{-1}$  respectively (W. Ranjith Premasiri et al. 2016). They all had their vibrational assignments as ring stretching, this indicates that the bacterium is associated with the cell membrane. *Chlamydia trachomatis*, that has no cell wall, therefore there are no purine metabolites in its Raman signature, instead the bacteria's biomarker bands are 900 $\text{cm}^{-1}$  for peptidoglycal layer, 1002 $\text{cm}^{-1}$  and 1550 $\text{cm}^{-1}$  both for proteins. Their vibrational assignments are C-C ring and PHE ring breathing in that order (Chen, Premasiri, and Ziegler 2018). This findings point out that Raman spectroscopy method is able to spectrally differentiate between *Neisseria gonorrhoeae* with cellular suspension and a dominant adenine, guanine and  $\text{NAD}^+$ , and *Chlamydia trachomatis* with no purine enzymes but proteins on the cell surface only.

Research has also been done to show the ability of Raman spectroscopy on label and label-free methods. The rich spectra and high sensitivity has made Raman spectroscopy an interesting technique in bacteria detection (Barhoumi and Halas 2010). The label based method uses plasmonic nanoparticles with specific Raman reporter molecules and modified target recognition elements to obtain strong Raman signals. However, rich bacterial cell information is lost, and the method is limited by its *in-situ* detection due to long steps of preparation, reactant volume and analytical time (Kneipp, and Bohr 2006). The label-free method is more preferred due to its direct technique of Raman peak assignment and easy



interpretation where it involves forming a thin film of nanoparticles on a substrate for signal enhancement (Liu et al. 2017).

The Raman spectroscopy has also been found to be efficient in differentiating between gram positive and gram negative bacteria (Prucek et al. 2012). In the study, NIR excitations of a laser light were directed on crystallized particles of identified gram positive bacteria *Enterococcus faecalis* and *Streptococcus pyrogenes*, and gram negative bacteria *Acinetobacter baumannii*, and *Klebsiella pneumonia*. Typical Raman spectra of these bacteria were obtained and recorded, then the bacterial cell wall compositions were compared. The variant peaks for the two categories of bacteria were detected at  $850\text{cm}^{-1}$  for Adenine with C-C breathing vibrational assignment,  $1556\text{cm}^{-1}$  for N-acetyl related bands and at  $1587\text{cm}^{-1}$  for Amide II with CN ring breathing vibrational assignment (Prucek et al. 2012). These bands were lacking in the gram negative bacteria due to the absence of peptidoglycans in their cell wall, i. e gram negative bacteria (Potara et al. 2011).

A deep interest on development of a rapid method of detection and diagnosis of bacteria is increasing among researchers. This is influenced by high rates of mortality caused by sepsis. Resistance to drugs by bacteria is caused by inappropriate drug prescription and delayed diagnosis which is due to complex and lengthy sample preparation methods. This is why SERS is gaining interest due to its molecular specificity and potential in pathogen species identification. Boardman and company identified that the Raman spectroscopy was able to accurately identify and differentiate samples of *Staphylococcus aureus* and *Escherichia coli* with 88% sensitivity and 97% specificity (Boardman et al. 2016). Also Raman spectroscopic analysis is also nondestructive, the samples can still be used in other tests.

### **2.3 Raman Spectroscopy in Detection of Antibiotic Resistant Bacteria**

Various studies are increasingly being directed into a deep understanding of the difference between antibiotic- susceptible and resistant bacteria (Fridkin et al. 2005). This is due to an increasing trend of resistance to antibiotics by a wide range of bacteria, posing a serious challenge in healthcare management. This implies that a new method for detection of bacterial resistance is required in order to enhance a breakthrough into antibiotic therapeutics. Raman spectroscopy has been identified as a promising technique for real-time bacteria detection and for the fundamental study of antibiotic-resistant mechanisms. This is because of its unique nondestructive sample analysis that also provides rich vibrational information, and enables chemical imaging/mapping of bacteria (Galvan and Yu 2018). Galvan and Yu, 2018 utilized SERS to study a wide range of bacteria and developed an understanding

of the bacterial cells' antibiotic degradation and so enabling the development of novel therapeutics. Their research had two of the bacteria *Staphylococcus aureus* and *Escherichia coli* characterized with peaks at 730 and 1095 $\text{cm}^{-1}$  which were associated with the glycosidic ring of peptidoglycan on their cell walls (Galvan and Yu 2018). The peaks significantly disappeared in the antibiotic-exposed bacteria cells because of the cell wall modification to develop resistance to drugs.

Raman spectroscopy has also been used in the rapid detection of antibiotic susceptibility of the bacterium *Escherichia coli* (Novelli-Rousseau et al. 2018). In this work by Novelli-Rousseau *et al*, the bacteria strains were incubated separately in three different antibiotics i.e. amoxicillin, gentamicin, and ciprofloxacin for two hours to allow for the development of resistance to the drugs. Raman spectra were then acquired on the three samples alongside susceptible bacteria and spectral data sets were isolated and processed for differentiation. Principle Component Analysis (PCA) was performed on the Raman spectra which evidenced a variation in Raman spectra between the bacteria exposed to antibiotics and the unexposed bacteria. (Novelli-Rousseau et al. 2018).

## **2.4 Chemometrics with Raman Spectroscopy in Analysis of Bacteria**

Currently, analytical instruments generate voluminous amounts of data that may be tedious to handle. Therefore, researchers use chemometric analytical tools as a means of dealing with this extensive information. Chemometric analytical tools range from those that do classification of data, preprocessing, projection, clustering, and prediction using pattern recognition methods. Principle component analysis (PCA) has been used for cluster classification of gram-positive and gram-negative bacteria (Prucek et al. 2012). Prucek et al used the method to present a definite potential of classifying gram-positive *Enterococcus faecalis*, *Streptococcus pyogenes* and gram-negative *Acinetobacter baumannii* and *Klebsiella pneumoniae* (Prucek et al. 2012).

During the process of studying the antibiotic susceptibility of bacteria, supervised machine learning techniques were used to track effects of antibiotics on a few bacterial strains using SERS (Novelli-Rousseau et al. 2018). The Raman spectroscopy was used to create a distinct database of the spectra of antibiotic susceptible bacteria and the spectra of bacteria with an antibiotic effect i.e. susceptible to drugs. The Support Vector Machine (SVM) algorithms together with the relative software were able to show a difference between bacteria that are antibiotic resistant and those that are susceptible, and so it was easy to make predictions of spectra from unknown samples. PCA was also used to categorize the

two groups of spectra even giving the distinction among bacteria at various concentrations of the antibiotic showing those that are antibiotic effect positive and antibiotic effect negative.

## CHAPTER III

### 3 THEORETICAL BACKGROUND

#### 3.1 Overview

This chapter reviews the basic principles of Raman with regard to the absorption and emission spectroscopy. It also presents the molecular vibrations of atoms, the processing and interpretation of Raman spectra. In addition, the statistical method ANOVA which is utilized in data analysis in this research, and a brief summary of unsupervised learning method i.e. PCA is discussed.

#### 3.2 Principles of Raman Scattering

##### 3.2.1 The classical approach

The laser light used in Raman spectroscopy is an electromagnetic wave. When the electric field  $\mathbf{E}$  is introduced into a sample, it causes a disturbance in the molecular electron orbits, hence an electrical dipole moment  $\tilde{\mathbf{p}}$  is induced.

The induced dipole moment is considered to be directly proportional to the electric field. Thus the induced dipole moment is given by (Vandenabeele 2013);

$$\tilde{\mathbf{p}} = \alpha \cdot \bar{\mathbf{E}} \quad (1)$$

Where  $\alpha$  is the electric polarizability tensor, which depends on the shape and dimensions of the chemical bond.

Since the Raman effect is caused by an EM radiation, and light is considered to be an oscillating electric field, then the electric field  $\mathbf{E}$  in time  $t$  is;

$$\bar{\mathbf{E}} = \bar{\mathbf{E}}_0 \cos \omega t = \bar{\mathbf{E}}_0 \cos (2\pi\nu_0 t) \quad (2)$$

Since  $\omega = 2\pi\nu_0$

Where  $\nu_0$  is the vibrational frequency of the radiation.

Since the polarizability tensor  $\alpha$  depends on the mode coordinates  $Q$  of the molecule, then the dependence relation can be expressed as a Taylor series;

$$\alpha = \alpha_0 + \sum_k \left( \frac{\partial \alpha}{\partial Q_k} \right)_0 \cdot Q_k + \frac{1}{2} \sum_{k,l} \left( \frac{\partial^2 \alpha}{\partial Q_k \partial Q_l} \right)_0 \cdot Q_k \cdot Q_l + \dots \quad (3)$$

Where;  $Q_k$  and  $Q_l$  are mode coordinates that correspond to the  $k^{th}$  and  $l^{th}$  vibrational modes with the vibrational frequencies  $\nu_k$  and  $\nu_l$ .

In approximation, only the first two terms in the equation 3 are maintained, hence the corresponding mode vibration equation reduces to;

$$\alpha_\nu = \alpha_0 + \alpha'_\nu \cdot Q_\nu \quad (4)$$

Where  $\alpha'_\nu = \sum_k \left( \frac{\partial \alpha}{\partial Q_k} \right)_0$  is the derivative of the polarizability tensor to the mode vibration  $Q_\nu$  at equilibrium.

From the previous approximation, the modes vibrate according to harmonic oscillator, hence the mode function  $Q_\nu$  varies with time as;

$$Q_\nu = Q_{\nu 0} \cdot \cos(2\pi \cdot \nu_\nu \cdot t + \varphi_\nu) \quad (5)$$

With  $Q_{\nu 0}$  being the amplitude of the mode vibration, and  $\varphi_\nu$  the phase angle.

Substituting equation 5 in equation 4, yields;

$$\alpha_\nu = \alpha_0 + \alpha'_\nu \cdot Q_{\nu 0} \cdot \cos(2\pi \cdot \nu_\nu \cdot t + \varphi_\nu) \quad (6)$$

This shows that the polarizability tensor undergoes a harmonic oscillation with frequency  $\nu_\nu$  which is equal to the vibrational frequency of the molecule's mode coordinate.

By substituting equation 2 and equation 6 for equation 1, then the induced electric dipole moment is given as;

$$p = [\alpha_0 + \alpha'_\nu \cdot Q_{\nu 0} \cdot \cos(2\pi \cdot \nu_\nu \cdot t + \varphi_\nu)] \bar{E}_0 \cos(2\pi\nu_0 t) \quad (7)$$

Simplifying further, equation (7) becomes;

$$p = \alpha_0 E_0 \cos(2\pi\nu_0 t) + \alpha'_\nu E_0 Q_{\nu 0} \cos(2\pi \nu_\nu t + \varphi_\nu) \cdot \cos(2\pi\nu_0 t) \quad (8)$$

Using the trigonometric expression;

$$\cos A \cdot \cos B = \frac{1}{2}[\cos(A + B) + \cos(A - B)] \quad (9)$$

Equation (8) is then modified and simplified to;

$$p = \alpha_0 E_0 \cos(2\pi\nu_0 t) + \frac{1}{2} \alpha'_\nu E_0 Q_{\nu_0} \cos[2\pi(\nu_0 + \nu_\nu) t + \varphi_\nu] + \frac{1}{2} \alpha'_\nu E_0 Q_{\nu_0} \cos[2\pi(\nu_0 - \nu_\nu) t - \varphi_\nu] \quad (10)$$

It can now be concluded that the induced dipole moment is a function of two frequencies, namely, the vibrational frequency of the molecules ( $\nu_\nu$ ) and the frequency of the incident radiation ( $\nu_0$ ), which is further simplified to;

$$p = p(\nu_0) + p(\nu_0 + \nu_\nu) + p(\nu_0 - \nu_\nu) \quad (11)$$

From the equation (11), the induced dipole moment consists of 3 components each with a particular dependence on frequency. The first term in the equation corresponds to elastic scattering of the incident radiation whereby the induced dipole moment has the same frequency as the incident radiation, hence the same energy. This is referred to as Rayleigh scattering, which was discovered by Lord Rayleigh (1842 - 1919). The second and third terms in the equation correspond to inelastic scattering whereby the induced dipole moment does not have the same frequency as the incident radiation; Raman scattering. The second term represents scattered photons with higher energy than the incident photons (Anti-stokes scattering), while the third term represents scattered photons with lower energy than the incident photons (Stokes scattering).

### 3.2.2 Quantum description

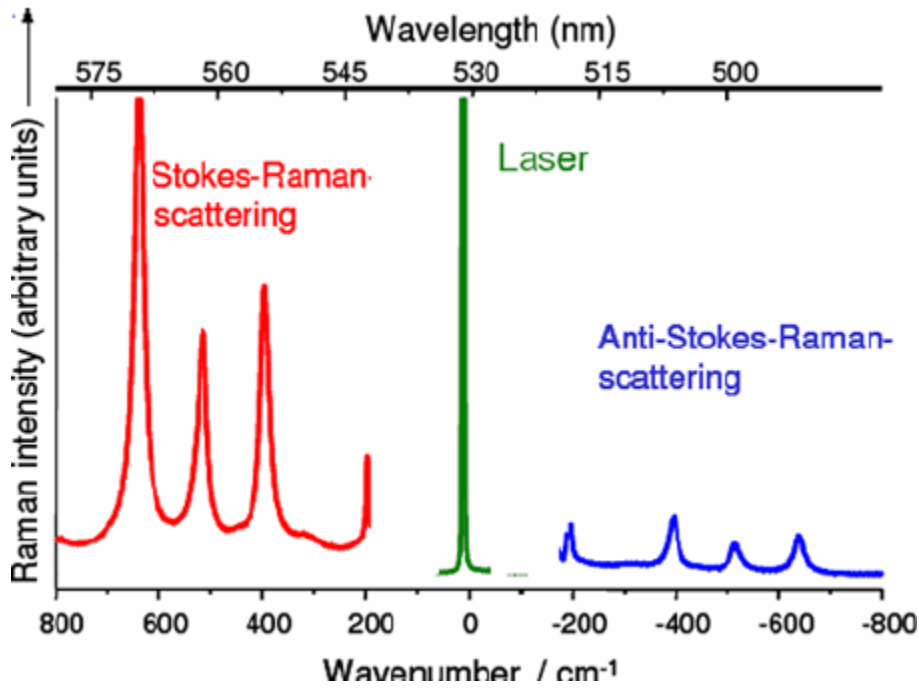
According to Boltzmann distribution function, molecules occupy various vibrational energy levels at room temperature, but according to Schrodinger equation for a particle in a box, there are only limited number of energy levels allowed with the lowest level being highly populated. Molecular bonds do confine the atoms to specific quantized vibrational modes where by vibrational energy levels are given as;

$$E_\nu = (v + \frac{1}{2}) h\nu_\nu \quad (v = 0, 1, 2, \dots) \quad (12)$$

Where  $\nu_\nu$  is the vibrational frequency,  $v$  is the vibrational quantum number of individual atoms,  $h$  is Planck's constant.

The laser light used in Raman spectroscopy is an electromagnetic wave. Once there is an incident EM radiation on the molecules, they absorb the energy and move to the higher energy states if the photon energy is equal to the energy difference  $\Delta E = E_2 - E_1$ , Where  $E_1$  and  $E_2$  are the energies of the initial and final levels. If this energy is not equal to  $\Delta E$ , the electron clouds may be displaced slightly from

equilibrium, a phenomenon called polarization. The polarized electrons are thought to be excited to a virtual level above the ground level. If the difference  $\Delta E = 0$ , then the Rayleigh scattering is considered, and the otherwise case is Stokes or Anti-stokes scattering is considered. For instance, when the scattered photons have higher energy than the incident photons it is considered to be Anti-stokes scattering, and when they have lower energy than the incident photons, then it is Stokes scattering. To determine the position of a Raman band in the spectrum, the difference in the ground state and the first vibrational state is used. **Figure 3.2:1** shows the model of dispersion of light by molecules.



*Figure 3.2:1: A model of the Rayleigh, Stokes-Raman and Anti-stokes-Raman scattering (Tarcea et al. 2008)*

### 3.3 Intensity of the Raman scattered light

Generally, the intensity of a Raman band is influenced by a number of factors which can be sample related or instrument related according to the equation;

$$I = \beta \cdot \gamma \tag{13}$$

Where  $\beta$  represents parameters that are related to the sample and  $\gamma$  those that relate to the instrument (Vandenabeele 2013). The instrument-related factors  $\gamma$  include the following; the intensity of the radiation source, the frequency of the exciting laser, the spectrometer's spectral response as well as the fourth power of the excitation energy (Juma 2018). The intensity of the Raman band as a function of sample-related factors is also dependent on a number of factors such as the size of the sample, the number of Raman active molecules in the sample and the molar scattering coefficient of the sample.

In the application of Raman spectroscopy, a laser (an intense beam of light) is focused on a sample, then the intensity of the scattered radiation as a function of its wavelength is measured. The intensity is plotted against the Raman wave number  $\omega$  ( $\text{cm}^{-1}$ ) which is the difference in frequency between the incident EM radiation and the scattered light;

$$\omega = \frac{1}{\lambda(\text{incident})} - \frac{1}{\lambda(\text{scattered})} = \frac{\nu_m}{c} - \frac{\nu_0}{c} \quad (14)$$

The wave number is clearly the difference in wavelengths of the incident radiation and the measured (scattered) radiation expressed in  $\text{cm}^{-1}$ ,  $\nu_m$ , and  $\nu_0$  are the frequencies of the respective radiations and  $c$  is the speed of light.

It is this knowledge that makes it possible for spectrophysicists to perform peak assignment for the purpose of identification of elements in a sample based on the Raman peaks. The Raman band position is influenced by the force of molecular bonds. For instance, the hydrogen atom has a single bond and the peaks lie between  $3650\text{-}2750 \text{ cm}^{-1}$ , this is the region where single bonds (O-H, N-H, C-H) absorb energy. Double bonds ( $\text{-C=C-}$ ,  $\text{-C=O}$ ) appear between  $2250\text{-}1500 \text{ cm}^{-1}$ , then the other groups that do have complex patterns with many molecules have their Raman bands below  $1500\text{cm}^{-1}$ . This information is therefore important during the interpretation of Raman spectra.

### 3.4 Chemometric techniques in Raman spectroscopy

Chemometrics is defined as a chemical discipline that selects and designs optimum measurement experiments and procedures, and provides fundamental chemical information by exploring the chemical data using mathematical and statistical methods (Krull 2012). This is according to the International Chemometric Society (Kudelski 2008).



### 3.5 Spectral Data Processing

Data is usually processed before the main analysis is done. This process is considered very crucial especially when dealing with huge amounts of data (Cozzolino et al. 2011). More importantly, preprocessing data before analysis can lead to the following;

- I. Transforming data to become more suitable for further analysis
- II. Reducing large amounts of data by eliminating any forms of variations
- III. Providing unique information within the data that can enlighten achievement of intended goals (Varmuza and Filzmoser 2016)

### 3.6 Supervised versus Unsupervised Learning

Generally, statistical data is categorized as either supervised or unsupervised. Supervised learning seeks to develop a characteristic model, based on the experimental data, which relates the outcome to the input data. This is done with an aim of accurately predicting the outcome for future predictions as well as creating a better understanding of the relationship between the outcome and the input data.

On the contrary, in unsupervised learning there is no specific response. Therefore, there is no outcome to fix a model to it and thus there are no outcomes to predict. The goal of unsupervised learning is to understand the relationship between the observations. This is achieved by clustering or classifying the data into groups. One statistical learning tool used in this research is Principle component analysis (PCA) (Colins 2017).

### 3.7 Principal Component Analysis (PCA)

PCA is an analytical tool that projects multivariate data progressively into data at a lower dimensional space. For instance, it provides the best of the data distribution. For exploratory data analysis, it is the most preferred analytical tool since from the definition, the desired number of dimensions (principle components) selected for the final exhibition and the subspace by the PCA is made up of the most faithful and accurate dimensional space of the original data (Biancolillo and Marini 2018). This implies that, in the process, there is minimal loss of crucial information in the data as well as it allows large dimensional data compression.

For instance, if we wish to visualize several observations say  $\mathbf{n}$ , with measurements on a set of  $\mathbf{t}$  features  $x_1, x_2, x_3, \dots, x_t$  as part of the data under exploration, visualization can be achieved by examining two-

dimensional scatter plots of the data, each containing  $n$  observations on two of the features, i.e. if there are  $t_2$  scatter plots, then  $t_2 = t(t-1)/2$ , for example, for  $t = 5$ , there are 10 plots.

However, if it is too large, then certainly this is not possible to look at each of them. Moreover, it is highly likely that none of them is informative since each contains just a small fraction of the total information present in the data. Clearly, it is impossible to visualize  $n$  observations when  $t$  is large, thus a method that can find a low dimensional representation of the data required, especially if it can as much of the data's important information as possible. This is exactly what PCA does (Satapathy et al. 2014).

It represents a low-dimensional data set containing variations as much as possible. These dimensions contain observations that vary along each dimension. Therefore each of them is a linear combination of the  $t$  features. These dimensions are therefore referred to as principle components. We now explore how the PCs are computed; the first PC  $Y_1$  of the set of features  $x_1, x_2, x_3, \dots, x_t$  is the linear combination of the features with the largest variance. It is given as;

$$Y_1 = \phi_{11}x_1 + \phi_{21}x_2 + \phi_{31}x_3 + \dots + \phi_{t1}x_t \quad (15)$$

The normalizing function here shows that  $\sum_{j=1}^t \phi_{j1}^2 = 1$ , and the elements  $\phi_{11}, \phi_{12}, \dots, \phi_{1t}$  are referred to as the loadings of the first principle component and summing them up makes the PC loading vector.

The loadings are constrained so that their sum of squares is equal to 1, otherwise setting them to be arbitrarily large could result in a large variance. To compute the first PC given an  $n \times t$  data set  $x$ , we take each of the variables in  $x$  to be centered to mean zero i.e. the column means of  $x$  are zero. Then we obtain a linear combination of the sample features with the largest sample variance owing to the constraint that  $\sum_{j=1}^t \phi_{j1}^2 = 1$ . The sample features are of the nature;

$$Y_{i1} = \phi_{11}x_{i1} + \phi_{21}x_{i2} + \dots + \phi_{t1}x_{it} \quad (16)$$

This means that the first PC loading vector solves the problem of optimization below;

$$\max_{\phi_{11}, \dots, \phi_{t1}} \left[ \frac{1}{n} \sum_{i=1}^n \left( \sum_{j=1}^t \phi_{j1} x_{ij} \right)^2 \right] \text{ subject to } \sum_{j=1}^t \phi_{j1}^2 = 1 \quad (17)$$

Thus the objective in (3) is  $\frac{1}{n} \sum_{i=1}^n Y_{i1}^2$ .

Since  $\frac{1}{n} \sum_{i=1}^n x_{ij}^2 = 0$ , then the average of  $Y_{11}, Y_{21}, \dots, Y_{n1}$  is also equal to zero, and they are referred to as scores of PC1.

The loading vector  $\emptyset_1$  containing elements  $\emptyset_{11}, \emptyset_{21} \dots \emptyset_{t1}$  shows the direction in the feature space along which data varies the most.

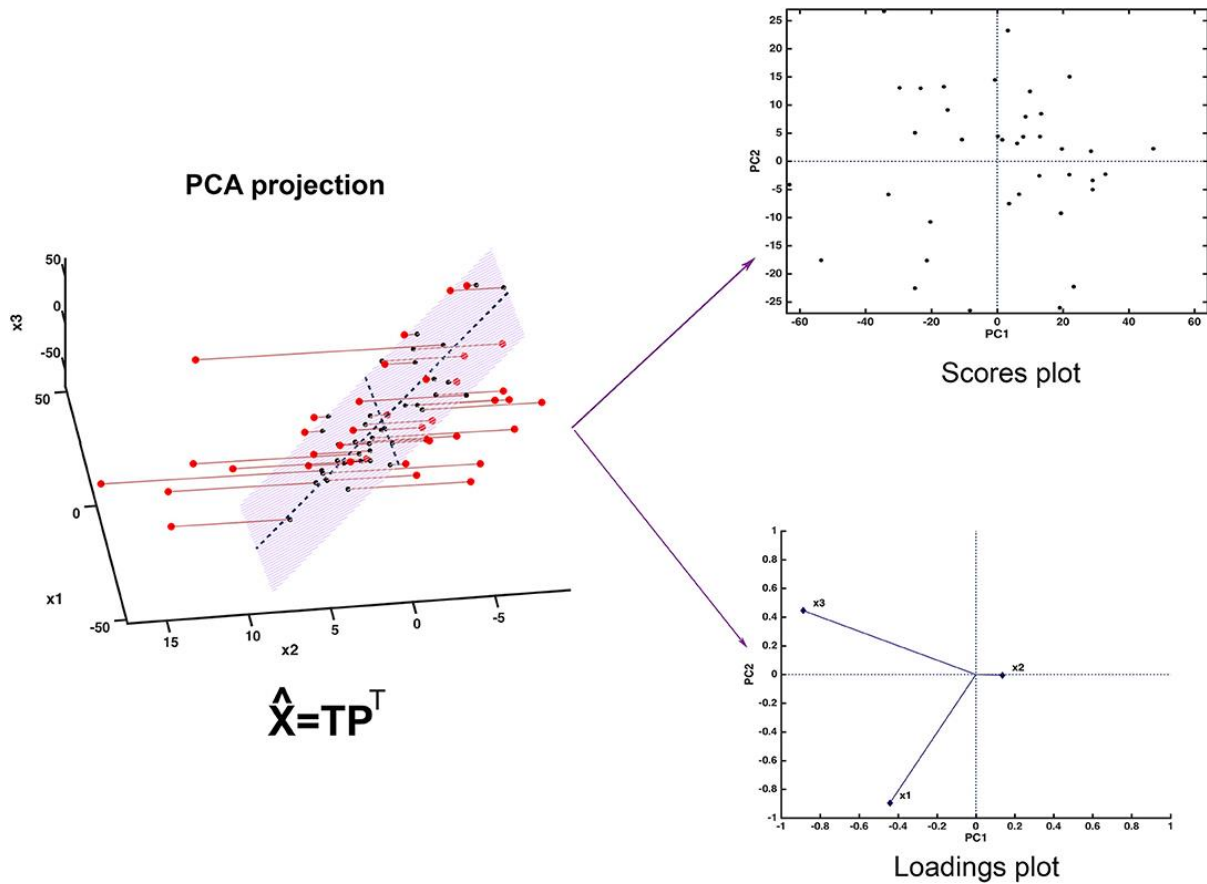
Projecting  $n$  data points  $x_1, x_2, \dots$  onto this direction, the values projected are themselves are PC scores  $Y_{11}, Y_{21}, \dots, Y_{n1}$ .

To compute the second PC,  $Y_2$ , which is a linear combination of  $x_1, x_2, x_3, \dots, x_t$  with maximal variance of all linear combinations that are not correlated with PC1. Therefore PC2 scores  $Y_{12}, Y_{22}, Y_{32}, \dots, Y_{n2}$  are given as;

$$Y_{12} = \emptyset_{11}x_{i1} + \emptyset_{i2}x_{22} + \dots + \emptyset_{t2}x_{it} \quad (18)$$

Where  $\emptyset_2$  is the loading vector of PC2.

Clearly, constraining  $Y_2$  to be uncorrelated to  $Y_1$  is like constraining the direction of  $\emptyset_2$  to be normal to the direction of  $\emptyset_1$ . In a larger data set where there are  $t > 2$  variables, multiple PCs are computed in a similar manner. Once the PCs are computed, they are then plotted against each other so as to have a low dimensional view of the data. For instance, we plot  $Y_1$  against  $Y_2$  or  $Y_2$  against  $Y_3$  or  $Y_1$  against  $Y_3$  and so on (Brereton 2009). The plots of the PCs against each other is called a PCA scores plot. They help in visualizing and inspecting deeply into the data, and observing the distribution of samples in the informative PC space. Similarly, loadings plots are developed from the scores plot. The loadings plots interpret the data by exploring the relative contribution of each variable(s) in a particular PC (Biancolillo and Marini 2018). The basics of the scores and loadings plots are shown in the **Figure 3.7:1**.



**Figure 3.7:1:** Fundamentals of PCA presented graphically.(Biancolillo and Marini 2018)

The first plot is a 3-dimensional space presentation which has been projected into a low dimensional subspace. The points highlighted in red in the far left diagram have been traversed into the first two PCs. The scores plot has four sub sections, where by each PC has a positive and a negative space. This means that a sample can fall in either of the four spaces according to their characteristics. Thus, samples with similar characteristics lie in the same space. In a similar way, the loadings plot simplifies further the information on the scores plot. For instance, the samples in the data lying in the negative of PC1 and positive of PC2 are seen to be having similar characteristics and can be identified in  $x_2$  as shown.

### 3.8 Analysis Of Variance (ANOVA)

Analysis of variance (ANOVA) is an analytical tool that is used to explore groups of data that depend on one or more factors, and determine the variations among the groups.

When considering a single factor, it is termed as one-way analysis of variance. This tests for an existence of systematic differences between the groups. For instance, to estimate the systematic differences between the groups, i.e. m and n, the measurement of repetitions m in n  $X_{mn}$  is given by;

$$X_{mn} = \bar{X}_{total} + (\bar{X}_m - \bar{X}_{total}) + t_{mn} \quad (19)$$

Where  $\bar{X}_{total}$  is the sum of total mean,  $\bar{X}_m$  is the group mean and  $t_{mn}$  random error.

A schematic data for a one-way ANOVA is shown in the **Table 3.1**.

**Table 3.1.** A schematic data for a one-way ANOVA

|                    |             |             |       |                   |
|--------------------|-------------|-------------|-------|-------------------|
| <b>Groups</b>      | 1           | 2           | ..... | n                 |
|                    | $x_{11}$    | $x_{12}$    | ..... | $x_{1n}$          |
|                    | $x_{21}$    | $x_{22}$    | ..... | $x_{2n}$          |
|                    | $x_{31}$    | $x_{32}$    | ..... | $x_{3p}$          |
|                    | .           | .           |       | .                 |
|                    | .           | .           |       | .                 |
|                    | .           | .           |       | .                 |
|                    | $x_{m1}^1$  | $x_{m2}^2$  |       | $x_{mn}^n$        |
| <b>Group means</b> | $\bar{x}_1$ | $\bar{x}_2$ | ..... | $\bar{x}_{total}$ |

A two way analysis of variance involves the investigation of two factors that influence the differences between groups. If the two factors are given as P and Q, then the measure of repetitions in row m and column n  $X_{mn}$  is given by;

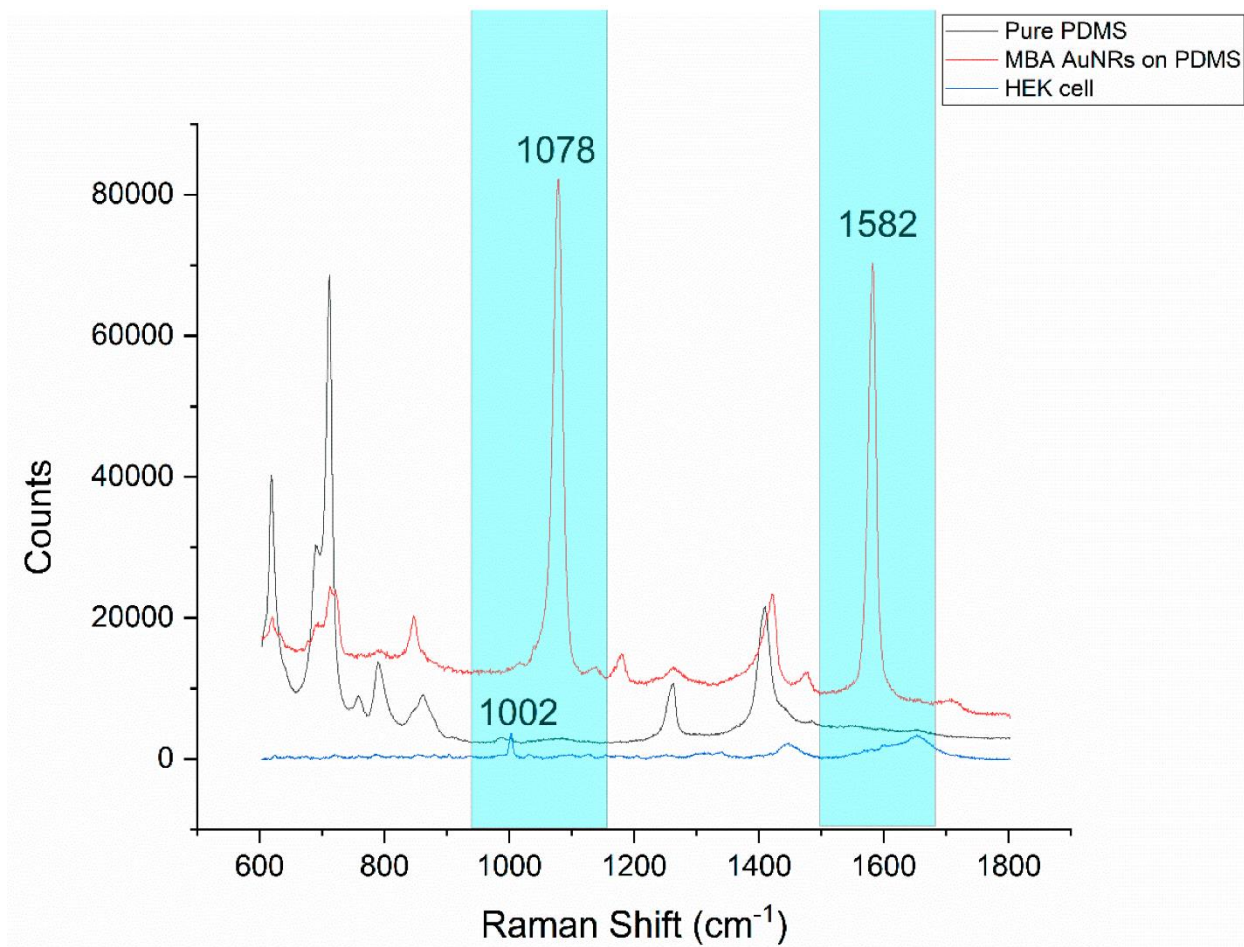
$$X_{mn} = \bar{X}_{total} + (\bar{x}_m^P - \bar{X}_{total}) + (\bar{x}_n^Q - \bar{X}_{total}) + t_{mn} \quad (20)$$

Where  $\bar{x}_m^P$  is the mean of factor P in row m,  $\bar{x}_n^Q$  is the mean of factor Q in column n, and  $t_{mn}$  is the random error. A schematic data for a two-way ANOVA is shown in the **Table 3.2**.

**Table 3.2.** A schematic data for a two-way ANOVA

| Factor P                | Factor Q           |                    |       |                    | Mean                   |
|-------------------------|--------------------|--------------------|-------|--------------------|------------------------|
|                         | 1                  | 2                  | ..... | N                  | $\overline{x^P_n}$     |
| 1                       | $X_{11}$           | $X_{12}$           | ..... | $X_{1n}$           | $\overline{x^P_1}$     |
| 2                       | $X_{21}$           | $X_{22}$           | ..... | $X_{2n}$           | $\overline{x^P_2}$     |
| .                       | .                  | .                  | .     | .                  | .                      |
| .                       | .                  | .                  | .     | .                  | .                      |
| .                       | .                  | .                  | .     | .                  | .                      |
| m                       | $X_{m1}$           | $X_{m2}$           | ..... | $X_{mn}$           | $\overline{x^P_n}$     |
| Mean $\overline{x^Q_m}$ | $\overline{x^Q_1}$ | $\overline{x^Q_2}$ | ..... | $\overline{x^Q_m}$ | $\overline{X_{total}}$ |

In the classification of samples using Raman spectroscopy, a plot of variance of the two or more samples is obtained from the mean columns. It is then plotted with the variances in the means against Raman shift. The main differences between the groups is then determined by the Raman peaks with the highest variance. An example of a variance peak is shown in the Figure 3.8:1. A study is conducted to compare the activity of fatty acid receptors in three categories of microfluidic devices. A plot of variance here shows the bands with the highest peaks are the peaks of variance among the groups under study together with their peak assignments (Zhang et al. 2019).



**Figure 3.8:1:** ANOVA plot of variance showing the activity of fatty acid receptors in three categories of microfluidic devices. It shows the Raman peaks at 1002, 1078, and 1582cm<sup>-1</sup> as the variance peaks (Zhang et al. 2019).

## CHAPTER IV

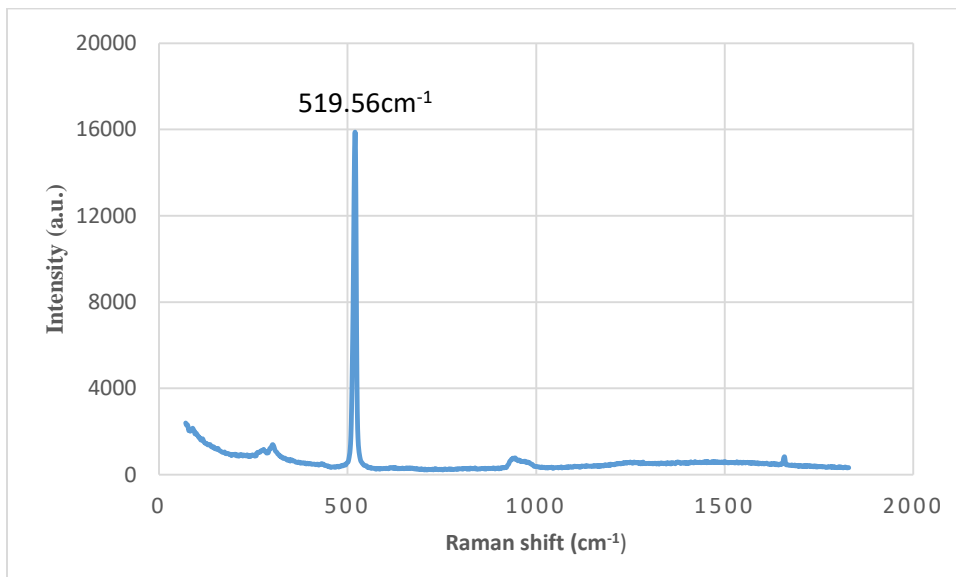
### 4 MATERIALS AND METHODS

#### 4.1 Overview

This chapter outlines the methodology used in accomplishing the objectives of this research. The Raman spectrometry used in performing the analysis is presented with its operating conditions. In addition, the process of obtaining and preparing the samples for spectra collection has been discussed, including the safety measures taken while handling the samples. Finally, the spectra pre-processing techniques used in this work have been clearly described.

#### 4.2 Raman Spectroscopy Instrumentation

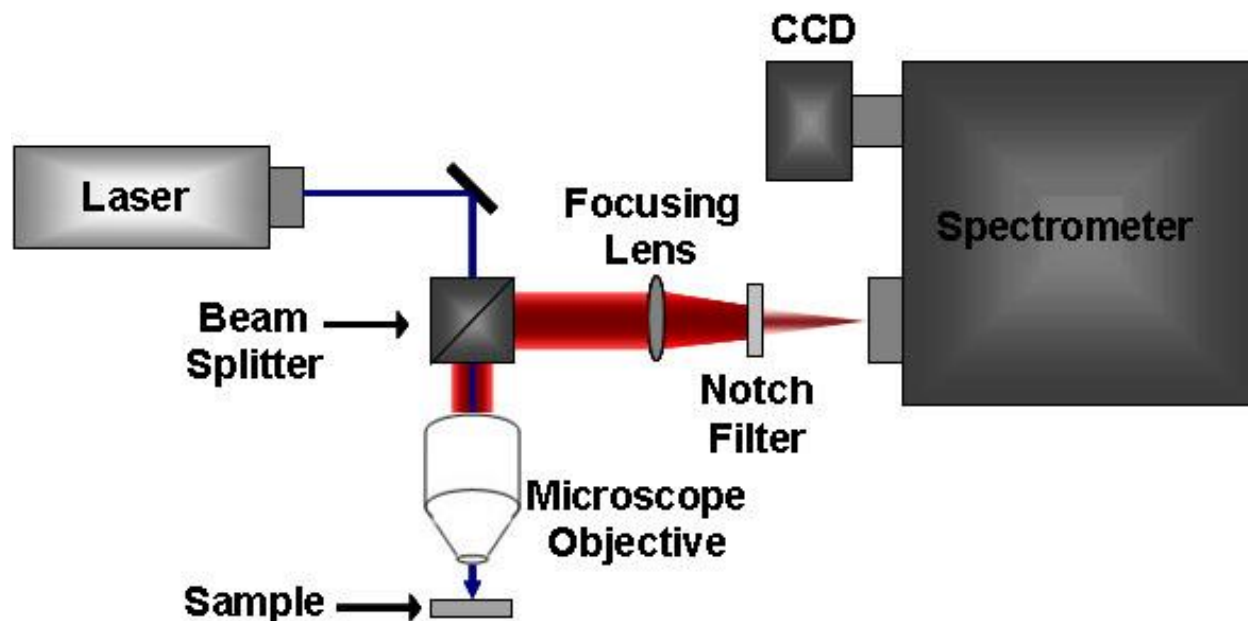
In this study, a confocal Raman spectroscopy (STR Corporation) was used, which was fitted with an in-built imaging monochromator spectrograph with two lasers that excite at 532nm and 785nm. The imaging spectrometer, (Princeton Instruments) equipped with 600, 1200, and 1800 lines/mm grating and a backscatter-illumination CCD camera purposely for acquiring spectra within the time frames optimized. (Juma 2018). The instrument was always calibrated before taking any measurement using a standard silicon wafer, and the Raman scatter bands were maintained at  $519.56\text{cm}^{-1}$ , as shown in Figure 4.2:1 (Murray et al. 1989).



**Figure 4.2:1:** Raman spectrum of silicon wafer using 785nm laser



During measurement, the laser light was conveyed to the Raman set up through the fiber optic cables that transmitted the beam under total internal reflection. It was then delivered to the sample to be measured by the shutter through the laser's band pass filter, **Figure 4.2:2**. The beam of light was then delivered by the band pass filter to the beam splitter that separated it into two equal parts; one that was reflected and another that passed on to the sample through the beam splitter. All these measurements were executed in a dark room to minimize background radiation due to fluorescence (Bhanot 2015).



*Figure 4.2:2: Schematic diagram of the set-up of the Raman spectroscopy, Courtesy of: <http://cnx.org/>*

Selection of the Raman spectral parameters was done depending on the nature of the sample and hence the region of occurrence of the spectral fingerprint for biological samples. The chosen spectral window for analysis for this work was  $300 - 1800\text{cm}^{-1}$ . This is usually the finger-print spectral region for biological samples (Potara et al. 2011). For this study, the chosen grating was 600 lines/mm which provided a wider spectral window, and also because it is a high frequency that can improve spectral resolution (Butler et al. 2016). In addition, the laser light intensity for excitation needed was controlled to avoid photo-thermal damaging sample cells as the measurement was done. This is done by choosing an objective lens of lower magnification of x10 and a laser power of 18.20mW.

The numerical aperture was also set at 0.3, and a levered stage that enabled and controlled the movement of the laser spot to be focused on the screen, the laser spot size used was  $68.5\mu\text{m}$ . Fine focus

on the sample was done by the movable stage. On exciting the sample, the scattered weak Raman signal was measured by the spectrometer, the computer with STR software was used for visualization purposes.

### **4.3 Preparation of Bacteria Samples**

The samples from infected patients were obtained from the KEMRI lab where they were stored at temperatures below  $-80^{\circ}\text{C}$ . The bacteria were not analyzed directly on a glass slide, a silver paste was smeared on the glass which acted as a substrate, and was left to air dry for 20-30min before the sample was placed on the silver paste which was also allowed to air dry for 5-8min before the Raman spectra were obtained.

### **4.4 Raman Analysis Procedure**

Well prepared samples of each bacteria were analyzed by the Raman spectroscopy method where a 785 nm laser was used to excite the samples and the Stokes Raman scattered radiation collected, recorded and analyzed. At least five samples of each bacteria were excited, this was depending on the amount of bacteria available for measurement, for each sample, five random points were excited and 10 different Raman spectra were obtained from each point making a total of 50 per bacterial sample. The massive amount of data obtained implies that more accurate results were expected from this procedure. The other experimental parameters were: an exposure time of 10 seconds, spectral accumulation of 10, a  $1000\text{cm}^{-1}$  center wavelength.

### **4.5 Statistical Analysis of Raman Spectra**

The raw data usually measured directly from the Raman spectrometer is made up of pure Raman signals superimposed on background auto fluorescence signals(Gervini 2006). Therefore, data pre-processing was done to extract pure Raman signals from the measured raw data.

The Vancouver Raman Algorithm was used in smoothing the data and eliminate the unwanted spectral peaks(León-Bejarano et al. 2019), while data analysis was performed using PCA and ANOVA as statistical tools.

## CHAPTER V

### 5 RESULTS AND DISCUSSION

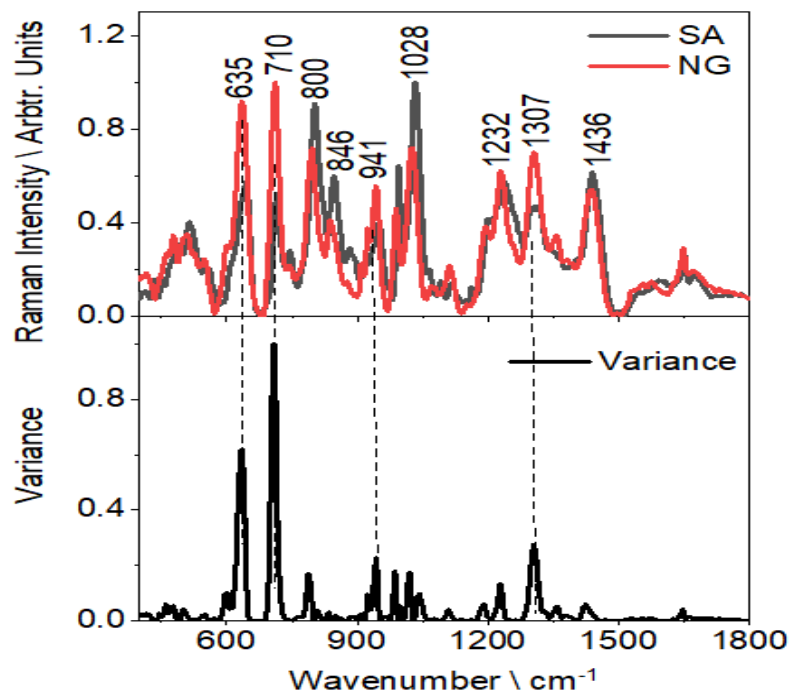
#### 5.1 Overview

This chapter presents the results obtained from the Raman spectrometric detection and differentiation of two bacterial strains: *Staphylococcus aureus* and *Neisseria gonorrhoea*, and also between the drug responsive and drug resistant strains of the bacteria *Staphylococcus aureus*. The characteristic spectra of each of the strains were collected and compared. Common and unique bands to each were identified and assignments made. Principal component analysis and ANOVA were used to aid in spectral data differentiation.

#### 5.2 Characteristic Raman spectra of *Staphylococcus aureus* and *Neisseria gonorrhoeae* bacteria

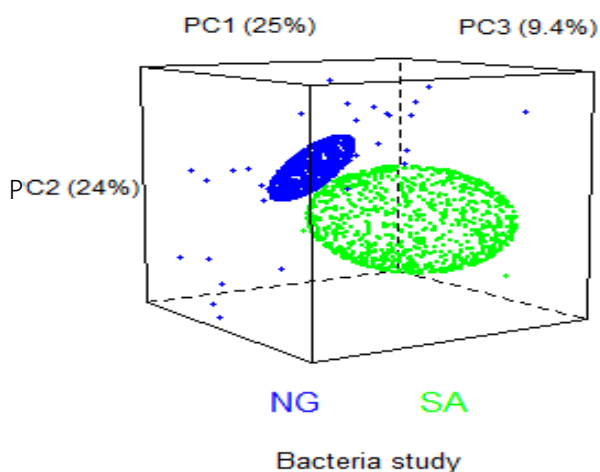
The **Figure 5.2:1** displays Raman spectra of bacteria *Staphylococcus aureus* (SA, black line) and *Neisseria gonorrhoeae* (NG, red line) together with the variance plot obtained from ANOVA. The Raman spectral profiles of the two bacterial strains (SA and NG) were similar indicating presence of similar Raman active molecules in each. Deviations were only observed in the intensities of bands centered at  $635\text{cm}^{-1}$ ,  $710\text{cm}^{-1}$ ,  $800\text{cm}^{-1}$ ,  $846\text{cm}^{-1}$ ,  $941\text{cm}^{-1}$ ,  $1028\text{cm}^{-1}$ ,  $1232\text{cm}^{-1}$  and  $1307\text{cm}^{-1}$ . These bands can be attributed to Tyrosine ( $635\text{cm}^{-1}$ ), Polysaccharides ( $710\text{cm}^{-1}$ ), proteins ( $846$ ,  $941$ ,  $1028\text{cm}^{-1}$ ), Amide III, adenine and polyadenine ( $1232\text{cm}^{-1}$ ), and DNA ( $1307\text{cm}^{-1}$ ) (Lu et al. 2013). Since Raman intensity is proportional to the number of molecules scattering the radiation (Kudelski 2008), it may be inferred that levels of tyrosine ( $800\text{cm}^{-1}$ ), and proteins ( $846$ ,  $1028\text{cm}^{-1}$ ) (Mungroo, Oliveira, and Neethirajan 2016) were elevated in *Staphylococcus aureus* bacteria strain as compared to those in *Neisseria gonorrhoeae* strain.

These variant bands are vibrationally assigned to C-N stretch, ring breathing and C-C ring breathing for the bands centered at wavenumbers  $800$ ,  $846$  and  $1028\text{cm}^{-1}$  respectively (Lu et al., 2013). Similarly, levels of Tyrosine ( $635\text{cm}^{-1}$ ), Polysaccharides ( $710\text{cm}^{-1}$ ), proteins ( $941\text{cm}^{-1}$ ), Amide III, adenine and polyadenine ( $1232\text{cm}^{-1}$ ) and DNA ( $1307\text{cm}^{-1}$ ) are elevated in *Neisseria gonorrhoeae* strain as compared to those in *Staphylococcus aureus* strain. They are vibrationally assigned to aromatic ring skeletal ( $635\text{cm}^{-1}$ ),  $\text{CH}_2$  Rock ( $710\text{cm}^{-1}$ ), C-N ring ( $941\text{cm}^{-1}$ ), and  $\text{NH}_2$  ring ( $1232\text{cm}^{-1}$ ) (Liu et al., 2017).



**Figure 5.2:1:** A graph of the Raman peaks of drug responsive bacteria *Staphylococcus aureus* (SA) and *Neisseria gonorrhoeae* (NG) together with a plot of variance showing variance peaks between the two bacteria strains.

PCA was performed on combined Raman spectral data sets of *Staphylococcus aureus* and *Neisseria gonorrhoeae* bacteria to evaluate further the power of Raman spectroscopic technique in detection and differentiation between the two bacteria strains. The intention was to further identify with the aid of PC loadings plots, the biomarker Raman bands, and compare them with those identified using ANOVA. These bands can potentially facilitate rapid diagnosis of bacterial infection using Raman spectroscopy. The PCA 3-D scores plots done using the R software, is shown in **Figure 5.2:2**. As can be seen in the plot (score plot) the Raman data sets from NG and SA were clearly segregated implying that their spectral patterns were different.



**Figure 5.2:2:** 3-D score plot for the combined Raman spectral data sets from *Staphylococcus aureus* and *Neisseria gonorrhoeae* bacteria strains. A distinct segregation between the data sets was observed with an explained variance of 58.4%.

The scores plot in Appendix 1 showing a plot of PC1 against PC2 also indicate that the bacteria strains *Staphylococcus aureus*, and *Neisseria gonorrhoea* were clearly differentiated using PCA. The bacteria *Staphylococcus aureus* whose biomarker bands lie at wave numbers  $800\text{cm}^{-1}$ ,  $846\text{cm}^{-1}$  and  $1028\text{cm}^{-1}$  according to the results of ANOVA has intense loadings in **Figure 5.2:3** being influenced by the same wave numbers. They are attributed to Tyrosine ( $800\text{cm}^{-1}$ ) and proteins ( $846$ ,  $1028\text{cm}^{-1}$ ) and their vibrational bands are assigned to C-N stretch, ring breathing and C-C breathing respectively (Lu et al. 2013).

Similarly, the loadings plot for bacteria *Neisseria gonorrhoea* is also shown in **Figure 5.2:3**. The negative of PC2 are represented by the peaks  $635\text{cm}^{-1}$  and  $710\text{cm}^{-1}$ ,  $941\text{cm}^{-1}$ ,  $1232\text{cm}^{-1}$ , and  $1307\text{cm}^{-1}$ . Likewise, these bands have a higher intensity variation in the ANOVA analysis from the **Figure 5.2:1** and also large loading signals in the loadings plot in **Figure 5.2:3**. The peaks are attributed to Tyrosine ( $635\text{cm}^{-1}$ ), Polysaccharides ( $710\text{cm}^{-1}$ ), protein ( $941\text{cm}^{-1}$ ), Amide III, adenine, polyadenine ( $1232\text{cm}^{-1}$ ) and DNA ( $1320\text{cm}^{-1}$ ) (Lu et al. 2013). They can be used also as the biomarker bands for bacteria *Neisseria gonorrhoeae* for Raman spectroscopy method. Their vibrational assignments are the aromatic ring skeletal,  $\text{CH}_2$  Rock, C-N ring and  $\text{NH}_2$  respectively (Mungroo, Oliveira, and Neethirajan 2016), as shown in table 5.1.

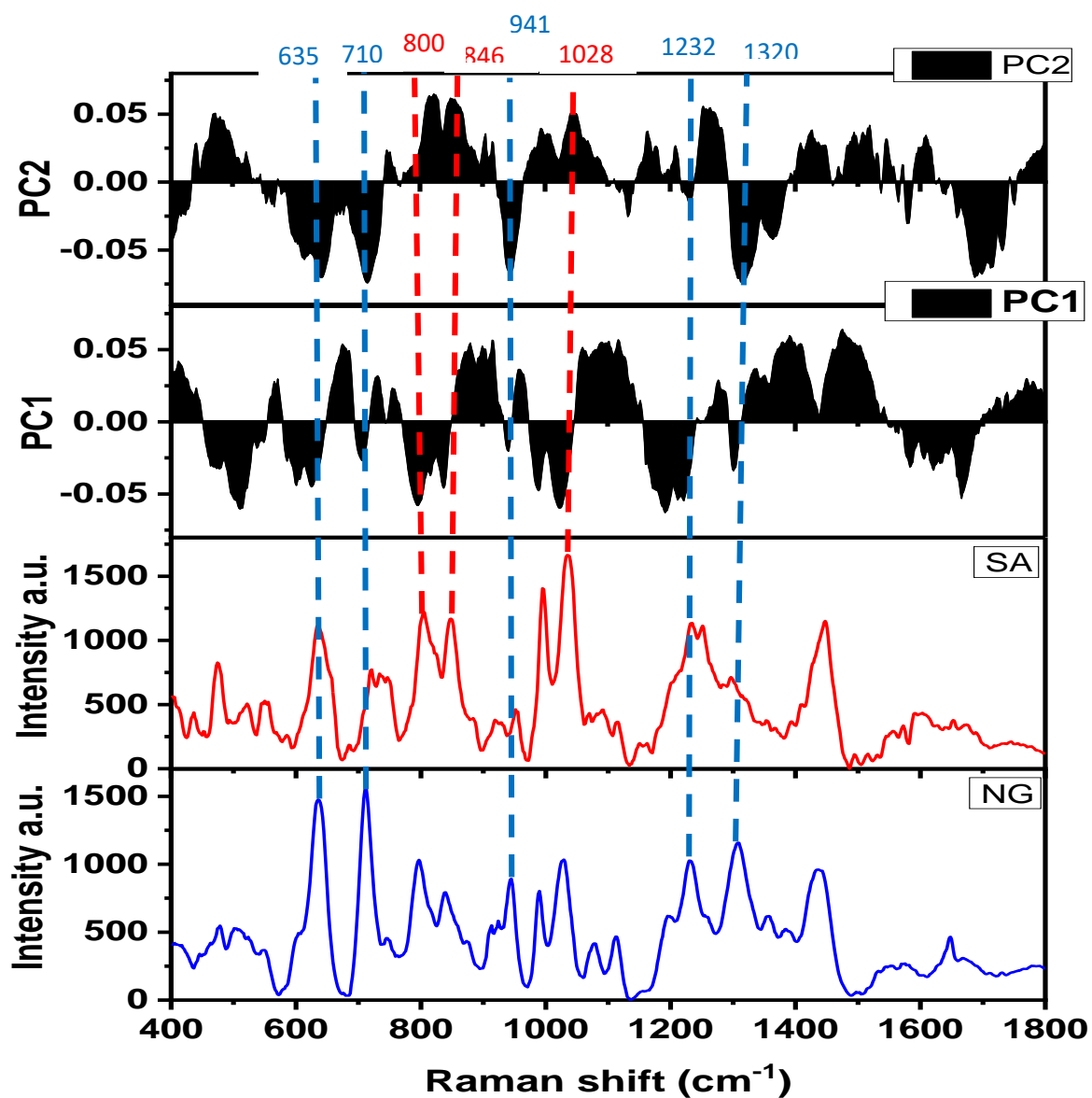


Figure 5.2:3: A loadings plot showing plots of PC1 and PC2 together with bacteria *Staphylococcus aureus* and *Neisseria gonorrhoeae*

In the comparison of the two bacteria strains *Staphylococcus aureus* and *Neisseria gonorrhoeae* in this chapter, it can be concluded that Raman spectroscopy has the ability to detect and differentiate the two strains. For instance, this method, Raman spectroscopy, was able to show that the spectral profiles of the two bacterial strains differed in terms of intensities of bands attributed proteins, tyrosine,

polysaccharides, Amide III, adenine and polyadenine (Wang et al. 2003). It may be inferred that the levels of these biomolecules differ depending on the bacterial strain (Liu et al. 2017).

A summary of the Raman bands and their peak and vibrational assignment is shown in **Table 5.1**, including their reference from literature.

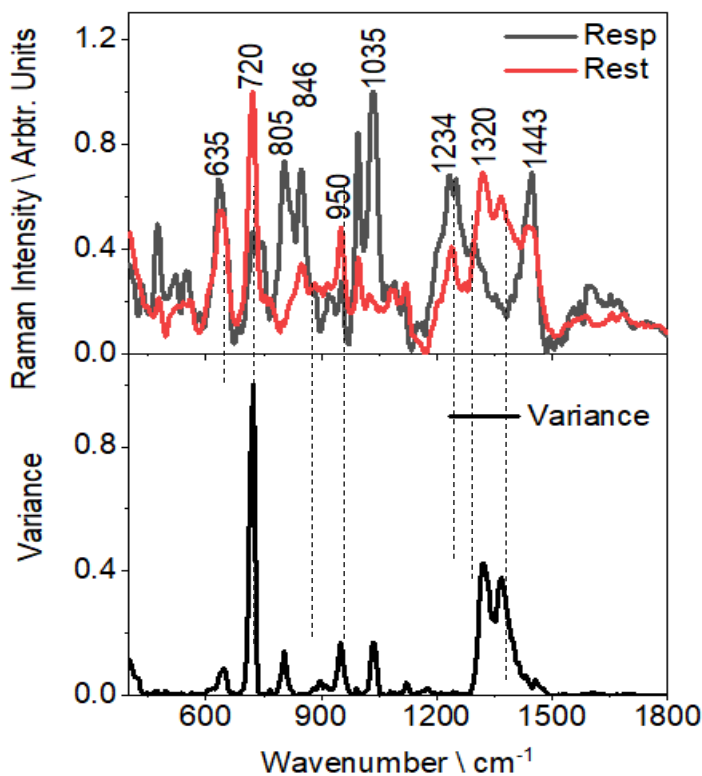
**Table 5.1:** Peak and vibrational assignment of Raman spectra of *Staphylococcus aureus* and *Neisseria gonorrhoeae* bacteria strains (González-Solís et al. 2014; Liu et al. 2017; Lu et al. 2013)

| <i>Staphylococcus aureus</i> (cm <sup>-1</sup> ) | <i>Neisseria gonorrhoeae</i> (cm <sup>-1</sup> ) | Literature (cm <sup>-1</sup> ) | Peak component assignment       | Vibrational assignment | Reference                                 |
|--|--|--------------------------------|---------------------------------|------------------------|---|
| -  | 635  | 624                            | Tyrosine                        | Aromatic ring skeletal | (Liu et al. 2017)                         |
| -  | 710  | 710                            | Polysaccharides                 | CH <sub>2</sub> Rock   | (Liu et al. 2017)                         |
| 800  | -  | 808                            | Tyrosine                        | C-N stretch            | (Liu et al. 2017)                         |
| 846  | -  | 855                            | Protein                         |                        | (Lu et al. 2013)                          |
| -  | 941  | 946                            | protein                         | C-N ring               | (González-Solís et al. 2014)              |
| 1028   | -  | 1028                           | Proteins and Tyrosine           | C-C ring breathing     | (González-Solís et al. 2014)              |
| -  | 1232   | 1220                           | Amide III, Adenine, Polyadenine |                        | (González-Solís et al. 2014)              |
| -  | 1307   | 1320                           | DNA                             | NH <sub>2</sub> ring   | (Mungroo, Oliveira, and Neethirajan 2016) |

### 5.3 Raman spectroscopy of antibiotic responsive and non-responsive

#### *Staphylococcus aureus*;

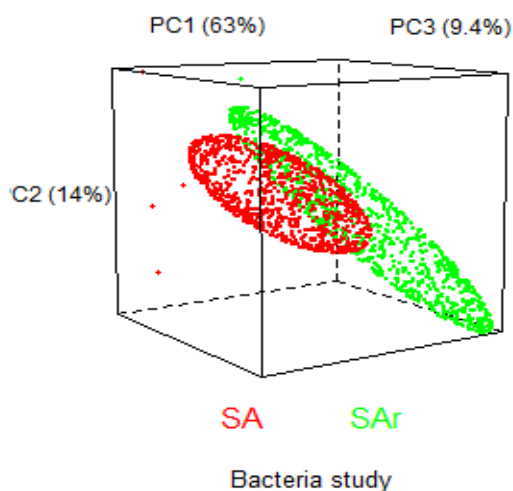
This work also explored the potential of Raman spectroscopy in distinguishing between drug responsive and drug resistant strains of *Staphylococcus aureus* bacteria. **Figure 5.3:1** displays the spectra of the two strains of SA (drug responsive, Resp; drug resistant, Rest). Evidently, the profiles indicated higher intensities of bands centered at 805, 1035, and 1443 $\text{cm}^{-1}$  in drug responsive and at 720, 950 and 1320  $\text{cm}^{-1}$  in drug resistant strain. These bands for the drug responsive strain can be attributed to tyrosine (805 $\text{cm}^{-1}$ ), proteins (1035  $\text{cm}^{-1}$ ), and phospholipids, deformation of proteins (1443 $\text{cm}^{-1}$ ) (Lu et al. 2013). These bands can be tentatively vibrationally assigned to C-N stretch, C-C ring breathing and  $\text{CH}_2$  respectively. The bands for the drug resistant strain are also attributed to polysaccharides (720  $\text{cm}^{-1}$ ), proteins (950 $\text{cm}^{-1}$ ) and DNA (1320  $\text{cm}^{-1}$ ) (Lu et al. 2013). The bands can also be tentatively vibrationally assigned to  $\text{CH}_2$  Rock, C-N ring and  $\text{NH}_2$  in that order (Mungroo, Oliveira, and Neethirajan 2016).



**Figure 5.3:1:** A graph of the Raman peaks of drug responsive bacteria *Staphylococcus aureus* (Resp) and its drug resistant strain (Rest) together with a plot of variance showing variance peaks between the two bacteria strains.



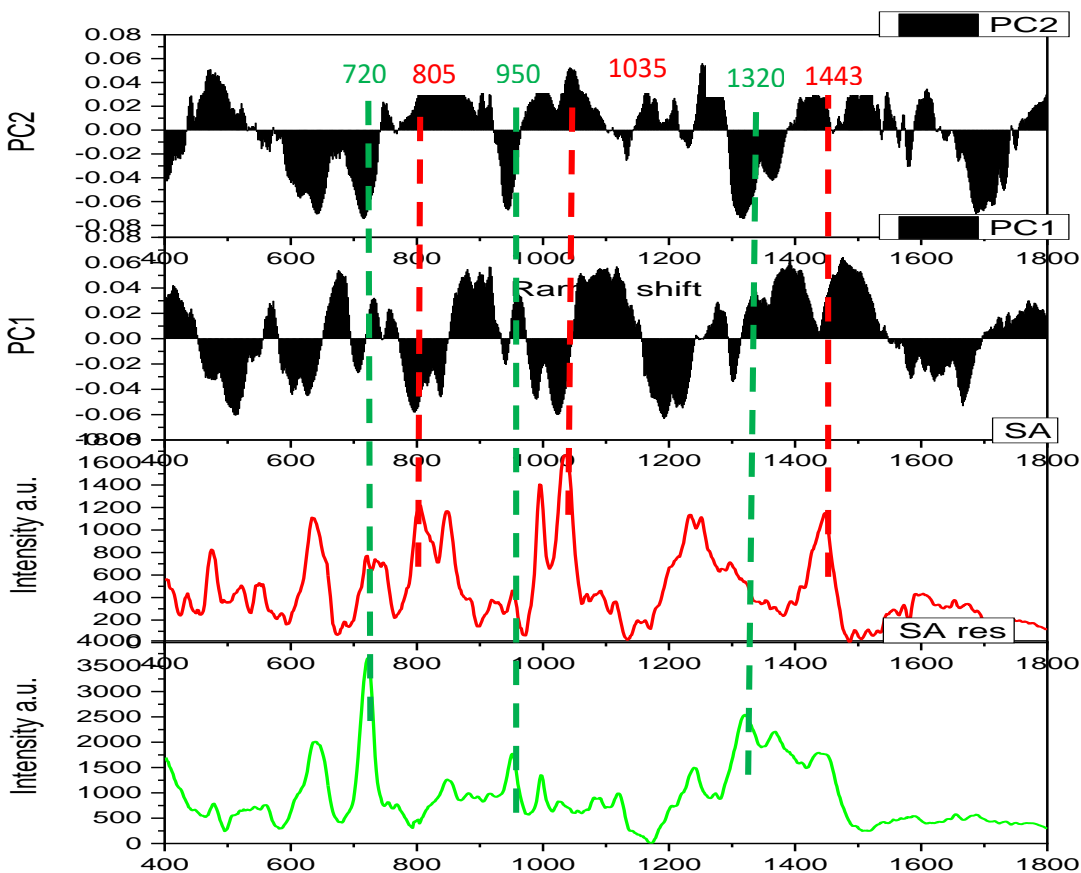
The results obtained from ANOVA above were compared to those obtained from PCA so as to a certain that the results are the same. A 3-D score plot in Figure 5.3:2 shows 86.4% segregation of the two strains, similarly, a 2-D scores plot in appendix 2 shows a clear segregation whereby the *Staphylococcus aureus* resistant bacteria strain lies in the positive of PC1 while *Staphylococcus aureus* responsive strain lies in the negative of PC1. In addition, the loadings plot in Figure 5.3:3 also show large loadings being influenced by the same wave numbers as in the results obtained from ANOVA.



**Figure 5.3:2:** A 3-D score plot for *Staphylococcus aureus* drug responsive and resistant strains

The unique bands that identify the drug resistant *Staphylococcus aureus* bacteria strain are centered at  $720\text{cm}^{-1}$ ,  $950\text{cm}^{-1}$  and  $1320\text{cm}^{-1}$ . These are the bands that have high intensity variance in the ANOVA analysis, and now they have high loading signals in the PCA analysis as compared to other bands as shown in **Figure 5.3:3**. They are attributed to polysaccharides ( $720\text{cm}^{-1}$ ), proteins ( $950\text{cm}^{-1}$ ) and DNA ( $1320\text{cm}^{-1}$ ), while their vibrational assignments are  $\text{CH}_2$  Rock, C-C ring breathing and  $\text{NH}_2$  respectively (Lu et al. 2013). Likewise, the drug responsive bacteria *Staphylococcus aureus* has its peaks at  $805\text{cm}^{-1}$  (proteins),  $1035\text{cm}^{-1}$  (proteins) and  $1443\text{cm}^{-1}$  (phospholipids, deformation mode of proteins). They are also vibrationally assigned to C-N stretch, C-C ring breathing and  $\text{CH}_2$  respectively (Lu et al. 2013).

Therefore these two sets of Raman peaks are said to be the biomarker bands for the two bacteria strains since they have high loading signals in PCA and high variation in intensity from the ANOVA analysis.



**Figure 5.3.3:** A loadings plot showing plots of PC1 and PC2 together with the drug responsive and resistant strains of bacteria *Staphylococcus aureus*.

**Table 5.2** shows a summary of the Raman peaks and their peak and vibrational assignments for the bacteria *Staphylococcus aureus* that is drug responsive and its drug resistant strain.

**Table 5.2:** Peak and vibrational assignment of Raman spectra of the drug responsive *Staphylococcus aureus* bacteria strain and its drug resistant strain. (González-Solís et al. 2014; Liu et al. 2017; Lu et al. 2013)

| <i>Staphylococcus aureus</i> responsive (cm <sup>-1</sup> ) | <i>Staphylococcus aureus</i> resistant (cm <sup>-1</sup> ) | Literature (cm <sup>-1</sup> ) | component assignment                        | Vibration assignment | Reference  |
|---|--|--------------------------------|---|----------------------|--|
| -   | 720  | 714                            | polysaccharides                             | CH <sub>2</sub> Rock | <i>Liu et al. 2017</i>                           |
| 805   | -  | 808                            | tyrosine                                    | C-N stretch          | <i>Lu et al. 2013)</i>                           |
|   | 950  | 945                            | proteins                                    | C-N stretch          | <i>Liu et al. 2017</i>                           |
| 1035  | -  | 1028                           | proteins                                    | C-C ring breathing   | <i>Liu et al. 2017</i>                           |
|   | 1320   | 1328                           | Adenine, polyadenine, DNA                   | NH <sub>2</sub>      | <i>(Mungroo, Oliveira, and Neethirajan 2016)</i> |
| 1443  | -  | 1447                           | Phospholipids, deformation mode of proteins | CH <sub>2</sub>      | <i>Liu et al. 2017</i>                           |

These results have a great significance in the diagnosis of bacteria in hospitals. This is because, a rapid, accurate and direct method of diagnosis means that the doctor shall make timely decisions in drug prescription which shall save lives, and more specifically if it is possible to determine whether the diagnosed bacteria are resistant to drugs makes it more easy to prescribe the correct drugs. The Raman spectroscopy is the method that has proved to perform this task more precisely and direct when coupled with the two statistical tools.

From literature, a cell gains resistance to drugs by modifying its structure and functions in a number of ways (Boardman et al. 2016). For example, if a bacterial cell has encountered with an antibiotic, it will either destroy the antibiotic, alter the antibiotic target or inhibit antibiotic permeability into the cell by changing the components of the cell wall. The biomarker bands in the two sets of bacteria show a shift in the peaks assigned to proteins. For instance, the drug responsive strain has the proteins assigned at the Raman wave number 1035cm<sup>-1</sup>, and this is attributed to a semi permeable membrane, while the drug resistant strain, the protein components in the cell wall have their wave number shifted to 950cm<sup>-1</sup> (Boardman et al. 2016). This can be attributed to a modified cell wall to possibly block the antibiotics

from the cell cytoplasm. The DNA of the cells is also showing disparities since the Raman peak at  $1320\text{cm}^{-1}$  which is attributed to DNA is elevated in the drug responsive strain than in the drug resistant one. An indication that there are changes in the DNA of the cells after developing resistance traits (Waters and Bassler 2005b).

## CHAPTER VI

### 6 CONCLUSION RECOMMENDATIONS AND PROSPECTS

#### 6.1 Summary and Conclusions

This research work utilized the Raman spectroscopy in obtaining Raman spectra of the drug responsive and drug resistant strains of bacteria *Staphylococcus aureus* and *Neisseria gonorrhoeae*. The spectra were then analyzed with ANOVA and PCA in distinguishing between *Staphylococcus aureus* and *Neisseria gonorrhoeae* bacteria and between the drug responsive and drug resistant strains of bacteria *Staphylococcus aureus*. The analyses obtained from the ANOVA and PCA were then used to obtain the biomarker bands that uniquely identified the bacteria.

Differentiating between *Staphylococcus aureus* and *Neisseria gonorrhoeae*, the Raman spectra obtained were first analyzed using ANOVA. The bands that had high variance in terms of intensities were the bands that distinguish between the two strains. These bands were compared to the PCA loading plots which showed that the particular bands have high loading signals. These bands are centered at  $800\text{cm}^{-1}$  (Tyrosine),  $846\text{cm}^{-1}$  and  $1028\text{cm}^{-1}$  (proteins) for bacteria *Staphylococcus aureus*, whereby their vibrational assignments are C-N stretch, ring breathing and C-C breathing respectively (Lu et al. 2013). The bands associated with *Neisseria gonorrhoeae* are centered at  $635\text{cm}^{-1}$  (tyrosine),  $710\text{cm}^{-1}$  (polysaccharides),  $941\text{cm}^{-1}$  (proteins),  $1232\text{cm}^{-1}$  (Amide III, adenine, polyadenine) and  $1320\text{cm}^{-1}$  (DNA) (Lu et al. 2013). Their vibrational assignments are the aromatic ring skeletal,  $\text{CH}_2$  Rock, C-N ring and  $\text{NH}_2$  respectively. These bands can be used as the biomarker bands for identifying the two bacteria strains.

Similarly, the Raman spectra obtained from the drug resistant and responsive strains of bacteria *Staphylococcus aureus* had been analyzed using ANOVA and PCA and the biomarker bands obtained were centered at  $720\text{cm}^{-1}$  (polysaccharides),  $950\text{cm}^{-1}$  (proteins) and  $1320\text{cm}^{-1}$  (DNA) for the drug resistant strain. Their vibrational assignments are  $\text{CH}_2$  Rock, C-C ring breathing and  $\text{NH}_2$  respectively (Lu et al. 2013). The drug responsive one also was influenced by the bands centered at  $805\text{cm}^{-1}$  (proteins),  $1035\text{cm}^{-1}$  (proteins) and  $1443\text{cm}^{-1}$  (phospholipids, deformation mode of proteins), with vibrational assignments being  $\text{CH}_2$  Rock, C-C ring breathing and  $\text{NH}_2$  respectively (Lu et al. 2013).

The results above can be significant in the current diagnostic procedures needed by clinical officers when trying to find out the specific bacteria infection for a patient and more precisely whether the bacteria is drug responsive or drug resistant. This information can be very crucial in bacteria diagnosis as it reduces the time a patient has to wait before being diagnosed and thus reduce mortality and morbidity rates in case of serious infection. In addition, instead of medical officers prescribing general antibiotics or using patient symptoms in determining the patient's illness, the Raman spectroscopy can be used in bacteria diagnosis. Furthermore, the clinical officer can directly determine whether the bacteria is responsive to drugs or not.

## **6.2 Recommendations and Future Prospects**

There are several other kinds of bacteria that cause severe infections that even lead to serious infection and death apart from *Staphylococcus aureus* and *Neisseria gonorrhoeae*. These include *Escherichia coli* and *Salmonella* that cause food poisoning, *Neisseria meningitides* that causes meningitis, *Helicobacter pylori* causing gastritis and ulcers, *Streptococcal* bacteria that cause a wide range of infections like pneumonia and ear infection (Bacterial Diseases 2014), just to mention a few. Several other types of bacteria do cause urinary tract and genital infections can be so severe and can lead to mortality if not well diagnosed and treated with specific drugs.

This means a lot of work needs to be done in utilising the Raman spectroscopy in rapid diagnosis of these bacteria so as to avoid severity of infection and to be precise, the medical officers need to determine the drug responsive status of each of them. Furthermore, since we have proven that this novel technology can detect and differentiate between a drug responsive and resistant bacteria, more research is needed to determine the specific drug that the bacteria is resistant to. This means that the medical officer needs to know the specific resistance of the bacteria and avoid prescribing the same drug.

## REFERENCES

- “Bacterial Diseases.” 2014. *Healthgrades*. <https://www.healthgrades.com/right-care/infections-and-contagious-diseases/bacterial-diseases> (May 20, 2020).
- Barhoumi, Aoune, and Naomi J. Halas. 2010. “Label-Free Detection of DNA Hybridization Using Surface Enhanced Raman Spectroscopy.” *Journal of the American Chemical Society* 132(37): 12792–93.
- Bhanot, Dr Deepak. 2015. “What Are the Differences between Raman and IR Spectroscopy?” *Lab-Training.com*. <https://lab-training.com/2015/06/26/what-are-the-differences-between-raman-and-ir-spectroscopy/> (March 18, 2020).
- Biancolillo, Alessandra, and Federico Marini. 2018. “Chemometric Methods for Spectroscopy-Based Pharmaceutical Analysis.” *Frontiers in Chemistry* 6. <https://www.frontiersin.org/articles/10.3389/fchem.2018.00576/full> (June 30, 2021).
- Boardman, Anna K. et al. 2016. “Rapid Detection of Bacteria from Blood with Surface-Enhanced Raman Spectroscopy.” *Analytical chemistry* 88(16): 8026–35.
- Brereton, Richard G. 2009. *Chemometrics for Pattern Recognition*. John Wiley & Sons.
- Butler, Holly J et al. 2016. “Using Raman Spectroscopy to Characterize Biological Materials.” *Nature Protocols* 11(4): 664–87.
- Chen, Y., W. R. Premasiri, and L. D. Ziegler. 2018. “Surface Enhanced Raman Spectroscopy of Chlamydia Trachomatis and Neisseria Gonorrhoeae for Diagnostics, and Extra-Cellular Metabolomics and Biochemical Monitoring.” *Scientific Reports* 8(1): 1–12.
- Colins, Michael. 2017. *Machine Learning: An Introduction to Supervised and Unsupervised Learning Algorithms*. CreateSpace Independent Publishing Platform.
- Cozzolino, D., W. U. Cynkar, N. Shah, and P. Smith. 2011. “Multivariate Data Analysis Applied to Spectroscopy: Potential Application to Juice and Fruit Quality.” *Food Research International* 44(7): 1888–96.
- Espagnon, Isabelle et al. 2014. “Direct Identification of Clinically Relevant Bacterial and Yeast Microcolonies and Macrocolonies on Solid Culture Media by Raman Spectroscopy.” *Journal of Biomedical Optics* 19(2): 027004.
- Ferguson, Kevin L., and Lanying Brown. 1996. “BACTEREMIA AND SEPSIS.” *Emergency Medicine Clinics* 14(1): 185–95.

- Fridkin, Scott K. et al. 2005. "Methicillin-Resistant Staphylococcus Aureus Disease in Three Communities." *New England Journal of Medicine* 352(14): 1436–44.
- Galvan, Daniel D., and Qiuming Yu. 2018. "Surface-Enhanced Raman Scattering for Rapid Detection and Characterization of Antibiotic-Resistant Bacteria." *Advanced healthcare materials*: 1701335.
- Gardiner, Derek J. 1989. "Introduction to Raman Scattering." In *Practical Raman Spectroscopy*, eds. Derek J. Gardiner and Pierre R. Graves. Berlin, Heidelberg: Springer, 1–12. [https://doi.org/10.1007/978-3-642-74040-4\\_1](https://doi.org/10.1007/978-3-642-74040-4_1) (October 13, 2021).
- Gervini, Daniel. 2006. "Free-Knot Spline Smoothing for Functional Data." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 68(4): 671–87.
- González-Solís, José Luis et al. 2014. "Cervical Cancer Detection Based on Serum Sample Raman Spectroscopy." *Lasers in Medical Science* 29(3): 979–85.
- Juma, Moses W. 2018. "Laser Raman Microspectrometric Assessment Of Uranium Forensics Signature In Aerosols Over A Model Nuclear Atmosphere." Thesis. University of Nairobi. <http://erepository.uonbi.ac.ke/handle/11295/105393> (February 12, 2021).
- Kneipp, Katrin, Harald Kneipp, and Henrik G. Bohr. 2006. "Single-Molecule SERS Spectroscopy." In *Surface-Enhanced Raman Scattering*, Springer, 261–77.
- Krull, Ira S. 2012. *Analytical Chemistry*. BoD – Books on Demand.
- Kudelski, Andrzej. 2008. "Analytical Applications of Raman Spectroscopy." *Talanta* 76(1): 1–8.
- León-Bejarano, Fabiola, Martin O. Méndez, Miguel G. Ramírez-Elías, and Alfonso Alba. 2019. "Improved Vancouver Raman Algorithm Based on Empirical Mode Decomposition for Denoising Biological Samples." *Applied Spectroscopy* 73(12): 1436–50.
- Liu, Yu et al. 2017. "Label and Label-Free Based Surface-Enhanced Raman Scattering for Pathogen Bacteria Detection: A Review." *Biosensors and Bioelectronics* 94: 131–40.
- Lu, Xiaonan et al. 2013. "Detecting and Tracking Nosocomial Methicillin-Resistant Staphylococcus Aureus Using a Microfluidic SERS Biosensor." *Analytical chemistry* 85(4): 2320–27.
- Mungroo, Nawfal Adam, Gustavo Oliveira, and Suresh Neethirajan. 2016. "SERS Based Point-of-Care Detection of Food-Borne Pathogens." *Microchimica Acta* 183(2): 697–707.
- Murray, R. et al. 1989. "The Calibration of the Strength of the Localized Vibrational Modes of Silicon Impurities in Epitaxial GaAs Revealed by Infrared Absorption and Raman Scattering." *Journal of Applied Physics* 66(6): 2589–96.



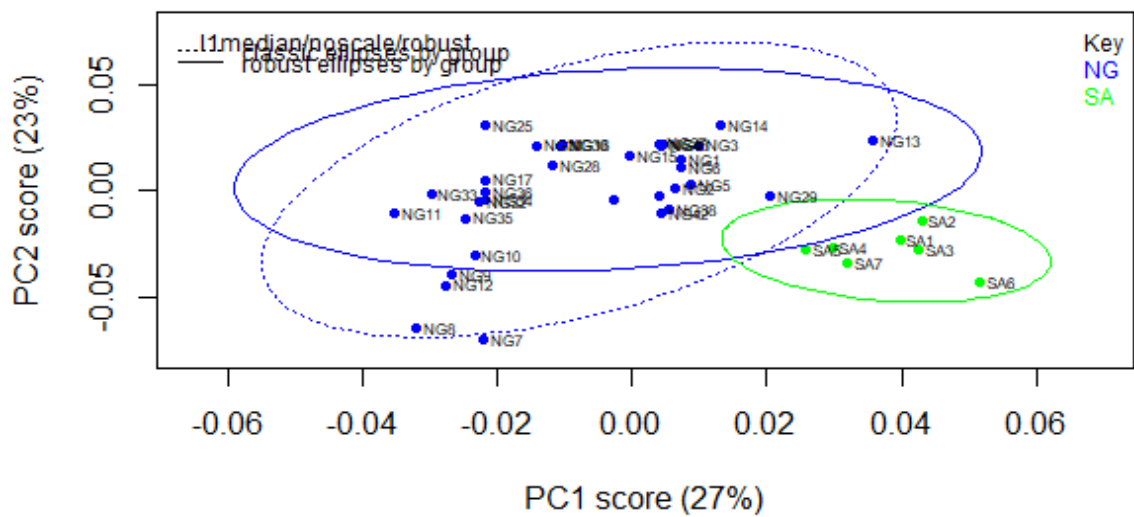
- Novelli-Rousseau, A. et al. 2018. "Culture-Free Antibiotic-Susceptibility Determination From Single-Bacterium Raman Spectra." *Scientific reports* 8(1): 3957.
- Potara, Monica et al. 2011. "Synergistic Antibacterial Activity of Chitosan–Silver Nanocomposites on Staphylococcus Aureus." *Nanotechnology* 22(13): 135101.
- Premasiri, W. R. et al. 2005. "Characterization of the Surface Enhanced Raman Scattering (SERS) of Bacteria." *The Journal of Physical Chemistry B* 109(1): 312–20.
- Premasiri, W. Ranjith et al. 2016. "The Biochemical Origins of the Surface-Enhanced Raman Spectra of Bacteria: A Metabolomics Profiling by SERS." *Analytical and bioanalytical chemistry* 408(17): 4631–47.
- Prucek, Robert et al. 2012. "Reproducible Discrimination between Gram-Positive and Gram-Negative Bacteria Using Surface Enhanced Raman Spectroscopy with Infrared Excitation." *Analyst* 137(12): 2866–70.
- Rebrošová, Katarína et al. 2017. "Rapid Identification of Staphylococci by Raman Spectroscopy." *Scientific Reports* 7. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5665888/> (July 24, 2020).
- Satapathy, Suresh Chandra, Bhabendra Narayan Biswal, Siba K. Udgata, and J. K. Mandal. 2014. *Proceedings of the 3rd International Conference on Frontiers of Intelligent Computing: Theory and Applications (FICTA) 2014: Volume 1*. Springer.
- Tarcea, Nicolae et al. 2008. "Raman Spectroscopy A Powerful Tool for in Situ Planetary Science." *Space Science Review* (2007) 135.
- Vandenabeele, Peter. 2013. *Practical Raman Spectroscopy: An Introduction*. John Wiley & Sons.
- Varmuza, Kurt, and Peter Filzmoser. 2016. *Introduction to Multivariate Statistical Analysis in Chemometrics*. CRC Press.
- Wang, Susan A. et al. 2003. "Multidrug-Resistant Neisseria Gonorrhoeae with Decreased Susceptibility to Cefixime—Hawaii, 2001." *Clinical Infectious Diseases* 37(6): 849–52.
- Waters, Christopher M., and Bonnie L. Bassler. 2005a. "QUORUM SENSING: Cell-to-Cell Communication in Bacteria." *Annual Review of Cell and Developmental Biology* 21(1): 319–46.
- . 2005b. "QUORUM SENSING: Cell-to-Cell Communication in Bacteria." *Annual Review of Cell and Developmental Biology* 21(1): 319–46.

Zhang, Han et al. 2019. "Use of Surface-Enhanced Raman Scattering (SERS) Probes to Detect Fatty Acid Receptor Activity in a Microfluidic Device." *Sensors* 19(7): 1663.

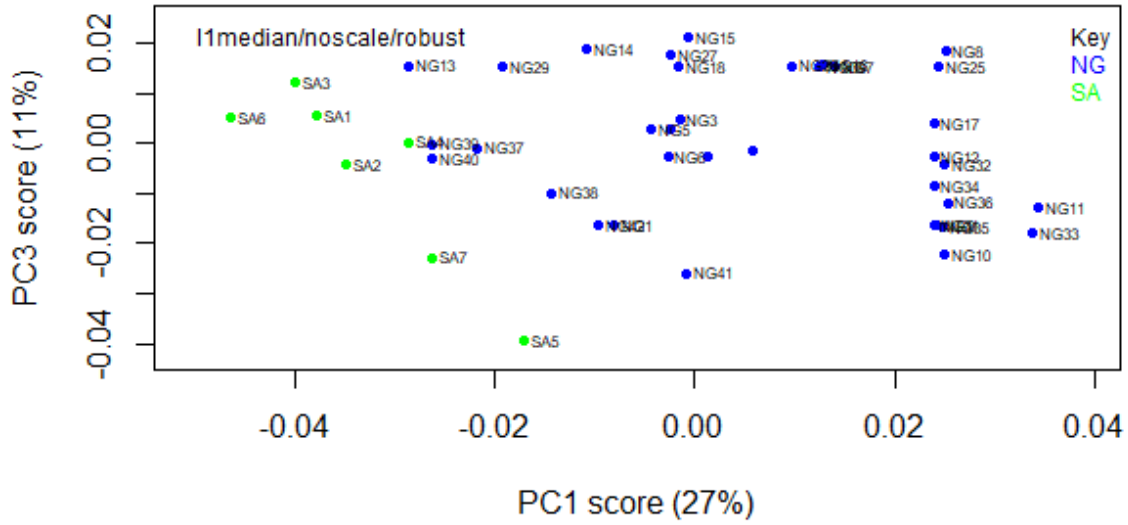
## APPENDICES

### Appendix 1: PCA scores plots of drug responsive *Staphylococcus aureus*(SA) and *Neisseria gonorrhoea* (NG).

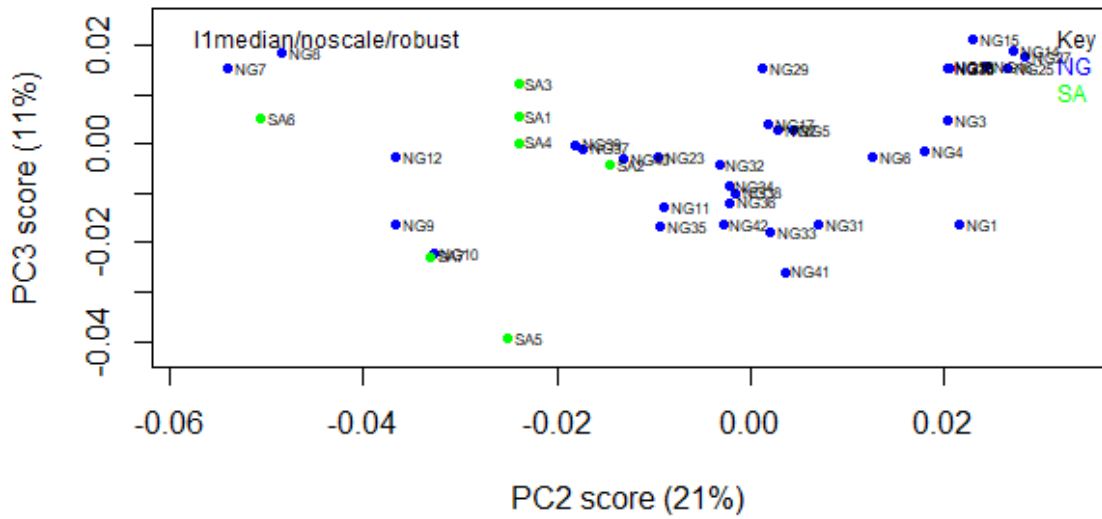
Figure 1 A: PCA scores plot of PC1 versus PC2.



**Figure 1 B:** PCA scores plot of PC1 versus PC3

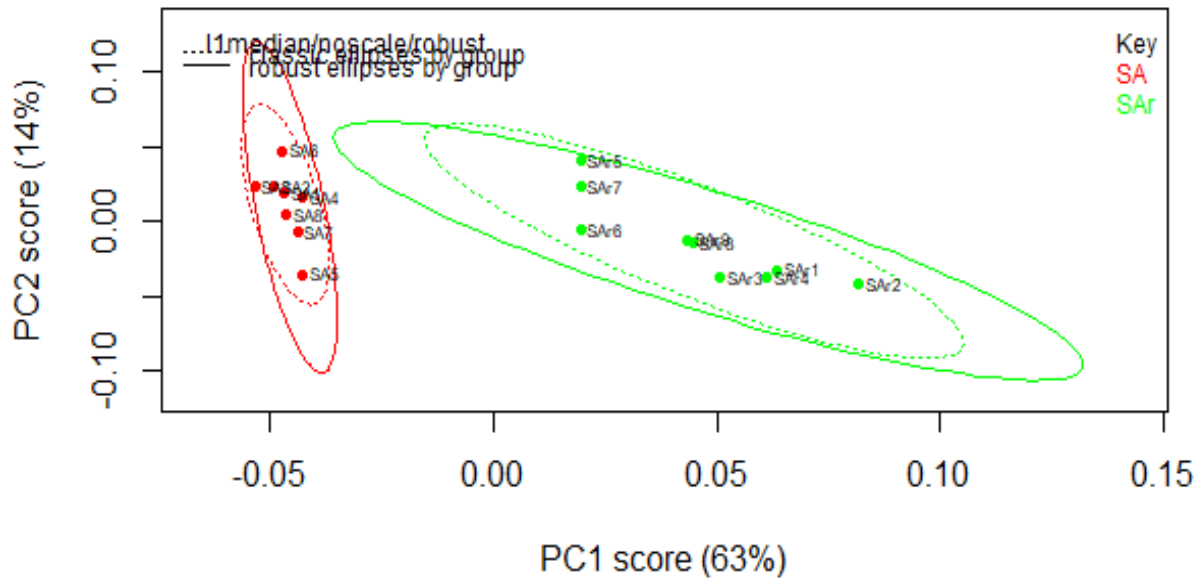


**Figure 1 C:** PCA scores plot of PC2 versus PC3

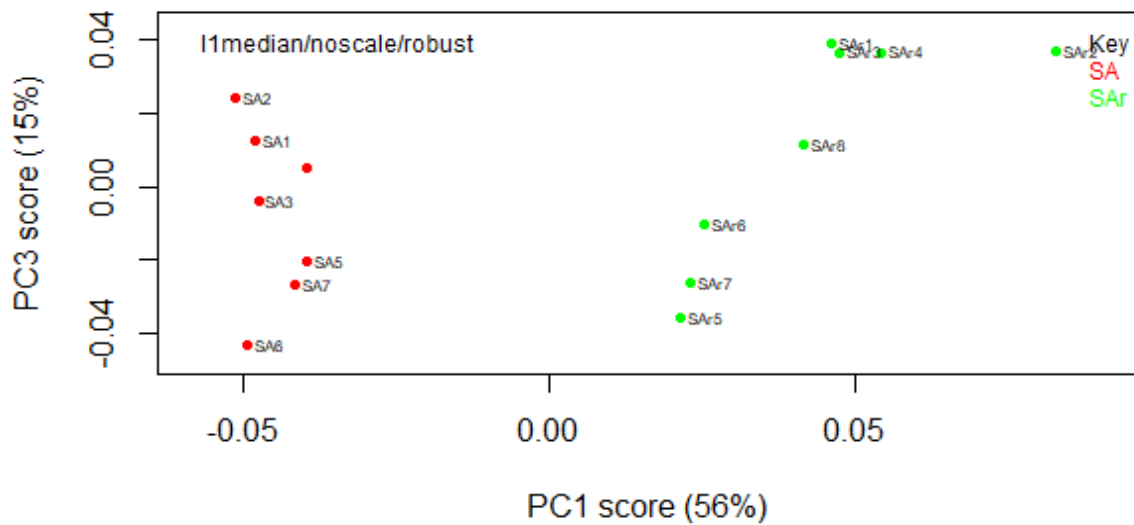


**Appendix 1: PCA scores plots of drug responsive Staphylococcus aureus(SA) and its drug resistant strain.**

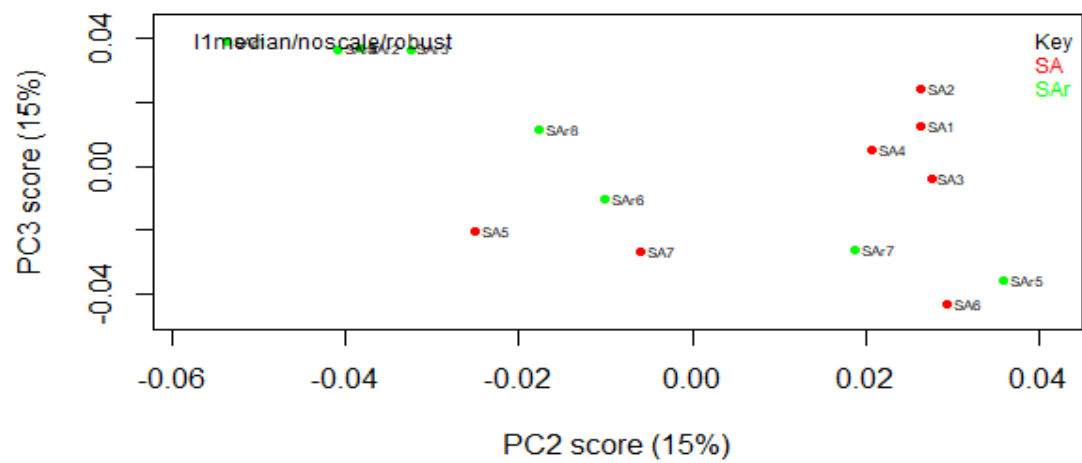
**Figure 2 A:** PCA scores plot of PC1 versus PC2.



**Figure 2 B:** PCA scores plot of PC1 versus PC3.



**Figure 2 C:** PCA scores plot of PC2 versus PC3.



Ms

Checked and approved.  
Dr. Birech Zephania  
12/09/2023



ORIGINALITY REPORT

9%

SIMILARITY INDEX

9%

INTERNET SOURCES

2%

PUBLICATIONS

0%

STUDENT PAPERS

PRIMARY SOURCES

1

[erepository.uonbi.ac.ke:8080](http://erepository.uonbi.ac.ke:8080)

Internet Source

6%

2

[erepository.uonbi.ac.ke](http://erepository.uonbi.ac.ke)

Internet Source

3%

Certified.



4/10/2023

CHAIRMAN, DEPARTMENT OF PHYSICS

Exclude quotes  Off

Exclude bibliography  On

Exclude matches  < 2%