# Detection of Environmental Contaminants in Water Utilizing Raman Scanning for *E. coli* Phenotype Changes.

**Hunter Flick**

Thesis submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of Master of Science in Biological Systems Engineering

Ryan S. Senger, Chair

Chenming “Mike” Zhang

Zhen “Jason” He

April 30th, 2019

Blacksburg, VA

Keywords: *E. coli*, biosensors, Raman Spectroscopy, water testing, antibiotics, nanoparticles

**Detection of Environmental Contaminants in Water Utilizing Raman Scanning for *E. coli* Phenotype Changes.**

**Hunter Flick**

# Academic Abstract

Raman spectroscopy and its counterpart surface-enhanced Raman scattering (SERS) have proven to be effective methods for detecting miniscule changes in the phenotypes of *E. coli* and other single-celled organisms to aid in the detection of new strains for industrial use and discovery of new antibiotics. The purpose of this study is to develop a method to quickly and accurately detect contaminants in water samples through phenotype changes in *E. coli* measured through SERS. Contaminated Luria-Bertani (LB) media was inoculated with LB with an $OD_{600}$ of 1, grown for two hours, and then dried on a flat piece of aluminum foil. These samples were then Raman scanned and processed to determine contaminant-induced changes to the phenotypes of the *E. coli*. Three types of tests were run to show the effectiveness of this method: single-component, multicomponent, and impure water sources. In single-component tests, it was found that differences due to NaCl contamination could be detected to 5.0E-9 weight percent (wt %), ethanol (EtOH) to 5.0E-7 volumetric percent (% v/v), citric acid (CA) to 2.8E-4 wt %, acetic acid (AA) to 2.6E-4 wt %, kanamycin to 2.5E-11 wt %, ampicillin to 2.5E-10 wt %, $CoCl_2$ to trace amounts, and silver nanoparticles (AgNP) to 5.2E-7 wt %. Many of these are below the detection limits of analytical instrumentation, but their effects on *E. coli* phenotypes were detectable by Raman spectroscopy. Multicomponent tests showed that in a mixture, the most toxic or most concentrated contaminants have the most effect on cell phenotype. However, it was shown that similar concentrations of similar contaminants may be difficult to discern with current methods. This behavior was also seen in the impure water samples, showing that tap water behaves the closest to a DI control, followed by running water, and finally stagnant bodies. This new method of monitoring *E. coli* phenotypes with Raman spectroscopy as a biosensor shows promise for the fast, portable, and accurate determination of environmental contaminants with a broad-spectrum and very low detection limits.

**Detection of Environmental Contaminants in Water Utilizing Raman Scanning for *E. coli* Phenotype Changes.**

**Hunter Flick**

# General Audience Abstract

Recently, Raman spectroscopy and an enhanced version called surface-enhanced Raman scattering (SERS) have shown promise in showing the effects of a cell's environment on how it expresses genes and what chemical compounds it produces to survive. It does this through reading the chemical signature it gives off due to these changes, and these readings have been used for showing the effects of antibiotics, finding varieties that are resistant to toxic byproducts of their activities, and as biosensors. The study outlined in this thesis aims to develop a test utilizing SERS to see the reactions of a non-pathogenic strain of *E. coli* to contaminants in their media and determine their identity. This test was run for three types of contaminated samples: one contaminant, three contaminants, and tests using impure water from a sink tap, a pond, and a stream. For the single contaminants, eight types were tested; NaCl, ethanol, citric acid, acetic acid (AA), kanamycin, ampicillin, $CoCl_2$, and silver nanoparticles. Detection limits for all contaminants were found, with the lowest detectable concentrations all falling below or matching detection limits of common methods. The lowest detectable concentrations came from kanamycin and $CoCl_2$, at 2.5E-11 weight percent and in amounts which are considered beneficial to the environment, respectively. The three-contaminant test shows that it is possible to pinpoint which contaminants are having the highest effect, though if contaminants are similar in nature and in similar concentrations, it may be difficult to pinpoint which is causing the effect. In the final test, similarity of water sources to pure water were found, with tap water being closest, followed by stream water and the most different being pond water. It was also found that pond water and stream water were closest in behavior, showing the power of this method in differentiating sources from each other.

# Acknowledgements

I would like to thank the following people for their support through my journey to my MS:

First, I would like to thank my PI Dr. Ryan Senger for his advice and guidance through this project, which provided me with the necessary skills and resources to complete this project and the requirements for my degree. I would also like to thank him for allowing myself and multiple other members of my previous lab to join his group and take on projects after the resignation of our original advisors, which I will forever be grateful for.

Second, I would like to thank the members of my advisory committee; Dr. Mike Zhang and Dr. Jason He. Their recommendations and feedback from our meetings helped me think of options that otherwise would have passed me by and have improved the quality of my work.

Third, I would like to thank my colleagues within the Metabolic Engineering and Systems Biology Laboratory; Imen Tanniche, Bert Huttanus, David Sherr, and Emily Berg for the support, collaboration, and camaraderie over the course of my project and degree. The assistance provided was crucial to my success.

Fourth, I would like to thank my family for their unconditional love and support over the years. Their encouragement has helped me overcome multiple hurdles and challenges through life, and I couldn't have done this without them.

Fifth and finally, I would like to thank the wonderful friends I have made at Virginia Tech and from my home in Roanoke for the support and fantastic experiences over the years. I look forward to sharing many more.

# Table of Contents

# Abbreviations

% v/v- Volumetric Percent

ANOVA- Analysis of Variance

AU- Arbitrary Unit

DAPC- Discriminant Analysis of Principal Components

DI- Deionized

HPLC- High-Performance Liquid Chromatography

LB- Luria-Bertani (Lennox)

PC- Principal Component

PCA- Principal Component Analysis

RID- Refractive Index Detector

SERS- Surface Enhanced Raman Scattering

Wt %- Weight Percent

# List of Figures

# List of Tables

# List of Equations

# Chapter 1

## 1.1: Introduction

In recent years, Raman spectroscopy, specifically surface-enhanced Raman scattering (SERS), has shown the ability to detect small changes in the chemical makeup of microorganisms such as *Escherichia coli* (*E. coli*). These data have been processed by multivariate statistics to produce "chemometric fingerprints" of microbial phenotypes. Changes in these fingerprints, induced by chemical or biological contaminants, allows for accurate detection of organic and inorganic components added to the media, as well as their concentration[1–4]. Traditionally, Raman spectra have been difficult to read, as well as being easy to interfere with due to the low intensity of the scattering light. Recent advances in data collection and processing have made the scattered light more intense and easier for the spectrometer to detect, while improvements in processing methods have allow for more accurate readings and the ability to only look at certain components such as concentration, points where bands show up, and with a higher degree of accuracy. While other analytical methods show the ability to detect low concentrations of contaminants in water, they are not as portable, noninvasive, have higher specificity, and will take a longer time to run. The goal of this study was to develop a methods utilizing a form of SERS to read changes in *E. coli* phenotypes induced due to contaminants added to liquid sources, and determine detection limits based on when the changes in cellular phenotypes are no longer statistically different from those induced by higher concentrations. After finding a baseline for singular contaminants, multicomponent tests and measurement of phenotype changes induced

by impure water samples from three sources were performed to determine the robustness of the method and how their behavior will look.

This thesis is divided into four parts: Chapter 1 shows the background for the research completed and shows a summary of related literature used for this project. Chapters 2 and 3 outline the methods and results shown for the upcoming manuscript, "Detection of Environmental Contaminants in Water Utilizing Raman Scanning for *E. coli* Phenotype Changes." Chapter 4 discusses the implications of these findings, as well as showing how the basic conclusions came about. Chapter 5 summarizes the work and conclusions drawn from it, as well as the future directions in which the work may go.

# 1.2: Literature Review

## 1.2.1: Toxicity and Dangers of Contaminants Used

### 1.2.1.1: Sodium chloride

Sodium chloride (NaCl) is found in almost all life and in many aquatic environments, ranging from fresh to saline[5–7]. Even so, one of its first major uses was for preservation due to its antiseptic properties[6,8–10], making it an excellent candidate for inducing phenotype changes within *E. coli*. The main bactericidal property of NaCl is when it is used in tandem with drying. This draws moisture from cells which attempt to gain a foothold onto solids, such as dried meat, killing them[6,9]. Toxicity due to cation buildup is also a possibility, namely if the chloride ion is taken up by another compound or bodily function[11]. NaCl buildup in water has the possibility of

depleting suitable drinking water reservoirs for communities and hurting local flora and fauna by creating a saline environment.

#### 1.2.1.2: Ethanol

Ethanol (EtOH) is a common byproduct of organisms like yeast, and humans consume and dispose of EtOH in large quantities regularly. It is also a common sterilizing agent, being lethal to *E. coli* at relatively low concentrations[12–16]. EtOH kills bacteria such as *E. coli* through multiple methods, with the two most common being protein denaturation and damage to cell membranes. Denaturation can occur directly or by inhibiting both transcription of DNA and RNA as well as translation into proteins, while membrane damage come from reduction peptidoglycan cross-linking and altering membrane lipid composition[12,15]. While the relatively low concentration of EtOH which can be found in water will not have much of an effect on humans, it could drastically affect the flora and fauna are areas where EtOH is dumped. Many strains of bacteria are able to become highly-resistant to EtOH, and constant exposure to small amounts within wastewater treatment plants or direct dumping could create strains of bacteria which are more resistant to current sanitizing methods involving various alcohol types.

#### 1.2.1.3: Citric Acid

Similar to EtOH and NaCl, citric acid (CA) is common in food, whether it comes from natural sources such as citrus fruits or synthetically produced and added into consumer goods for flavor and preservation[17–19]. Industry has uses for CA as well, namely as a cleaner[20–22]. These reasons, as well as the fact that large concentrations can drastically alter the pH of a solution, are

why it was chosen as one of the organic acids used for this experiment. While CA is not toxic to most life, as it is utilized as a food source and is a key intermediate and component of the Krebs cycle, high concentrations in areas where waste is dumped will lower the pH, potentially affecting the environment around and what happens downstream. This change in the water's pH could also harm and human activities in the area, should the water not be purified and pH-balanced before use. CA is also a known metal chelator, like ethylenediaminetetraacetic acid[21,22]. High concentrations of the compound could speed up corrosion of metal within the local water ecosystem, as well as bring metal ions and atoms into an environment not prepared for them.

#### 1.2.1.4: Acetic Acid

Acetic acid (AA) is a "weak" acid which is commonly used in industry as a solvent and is also used in food products to impart flavor or help with preservation of foods, similar to CA[18,19,23]. While the exact mechanisms of the antibacterial properties of AA are not known, it is known to be antibacterial and induce apoptosis in yeast. In yeast, it seems as though it affects amino acid synthesis and stress responses, among other mechanisms[24]. Potential effects of AA entering the environment are similar to that of CA, in that it could raise the pH and cause an imbalance. This, along with its bactericidal properties, can disrupt the local ecosystems and potentially affect piping that it contacts.

#### 1.2.1.5: Kanamycin

Kanamycin A, known simply as kanamycin, is an aminoglycoside antibiotic formally discovered and described in 1957[25]. It is used to treat infections caused by both Gram-positive

and Gram-negative bacteria[26]. It is mainly used in agriculture, but can be used for a number of infections in humans, including *E. coli*[25–28]. For bacteria, kanamycin works by interfering with the 30S ribosomal unit, which causes mistranslation and inhibition of protein translocation[25,27–29]. In humans, kanamycin can have multiple direct side-effects, including nephro- and ototoxicity, as well as allergic reactions or damaging intestinal flora[27,28]. The use of kanamycin contributes to the rise of antibiotic resistance in microbes as well[27,30,31]. The methods for the spread of resistance include direct evolution from trace-amounts of kanamycin entering wastewater systems and direct transfer of resistance genes through horizontal gene transfer[30–32].

**1.2.1.6: Ampicillin**

Ampicillin, like kanamycin, is a commonly-used antibiotic. However, ampicillin belongs to the $\beta$-lactam group, with a mechanism of action similar to penicillin[33–35]. Unlike most other penicillins, it is able to be adequately absorbed with oral doses while maintaining low-toxicity in humans and rats[33]. The mechanism of action of ampicillin and related $\beta$-lactam antibiotics involves blocking channels and the OmpF porin in the cell membrane, preventing cell transport[36]. Like with kanamycin, concentrations of ampicillin in the environment can convey antibiotic resistance onto otherwise susceptible bacteria, whether it's by direct exposure or horizontal gene transfer with a bacteria that gained resistance[32]. While ampicillin is relatively nontoxic, there are a fair amount of penicillin-sensitive individuals who could react to subtherapeutic doses in water or the environment, which also raises concerns about its presence[35].

#### 1.2.1.7: Cobalt and Cobalt Chloride

Cobalt is a metallic element which falls into the transition metal category[37–39]. It is required by many organisms, both eukaryotic and prokaryotic. Its main use is within vitamin B12[37–41], though other cofactors and even enzymes utilizing it exist[37]. A common form of cobalt used by humans is its Co(II), or $Co^{2+}$ form bound to chlorine to make cobalt chloride ($CoCl_2$) [38]. When it or elemental cobalt is ingested, it tends to affect many iron-containing proteins, namely iron-sulfur enzymes[38] and other hemes[40]. The Co(II) tends to outcompete the iron in these enzymes and severely affects their functions, including hindering electron transport and respiration[38–40]. While trace amounts of cobalt in the environment are normal, an increase in either elemental or combined cobalt can pose a severe danger not only for the local flora and fauna, but to anyone utilizing the water source or land for food and water.

#### 1.2.1.8: Silver Nanoparticles

Similar to antibiotics like ampicillin and kanamycin, nanoparticles are a recently emerging contaminant. One of the most common types of nanoparticles found in many facets of science is silver. Its uses are wide, including drug delivery, spectroscopic research, biosensors and detectors, and even environmental monitoring[27,42–44]. Silver nanoparticles (AgNPs) and other nanoparticles, however, are not regulated as tightly as other metals that are used, namely due to a lack of understanding of their long-term effects[45–48]. This means that, while waste does tend to be collected, more is released than is healthy.

In its elemental and ionic forms, silver has long been used for antibacterial purposes[49]. It tends to alter proteins, inhibiting their structure, as well as metabolic processes, such as the

electron transport chain and cell membrane permeability[49,50]. AgNPs have similar effects, as well as disrupting ATP production and DNA replication[48]. The size of nanoparticles has a large effect on their toxicity as well, as they have large surface to volume ratios, which can affect permeability, stability, and how dispersed they can become. In humans, silver tends to be relatively safe, though nanoparticles have been found to induce toxicity in lung cells with a release of silver. These nanoparticles are also able to induce many types of toxicity symptoms, including cytotoxicity and genotoxicity depending on their size and concentration[45].

## 1.2.2: Methods for Water Quality Testing

### 1.2.2.1: Salts and Cobalt

Two very common contaminants in water were various salt compounds and the metal ions which make them up, specifically Co. While many of these ions are present in trace amounts in both trace and higher amounts in fresh water and the land, too many of them will begin to cause problems and create toxic environments[7,51–53]. Thus, multiple methods have been developed to test for them.

One of the most common methods for detecting common metal ions in salts (Li, K, Na) and anions such as Cl is through conductivity[7,52,54]. This is because the conductivity of water is highly dependent on the charged ions of dissolved compounds such as salts. Other potential tests include the use of hydrometers to compare density[7], various forms of titration with various indicators[52] and atomic absorption spctrophotometry[51,55]. These tests have a variety of ranges, and the most precise can go as low as 2 mg/L for chloride conductivity[52], 5 mg/L for titration[52], and 0.1 mg/L weight percent (1E-5 wt %) for common spectrophotometry methods at 589.6 nm,

though reports of values down to 5 µg/L (5E-7 wt %) have been reported, though these are under very specific conditions[55].

While it is possible to measure Co with the above methods, the standard methods for measuring Co concentration are various forms of atomic absorption spectrophotometry. Ranges for the method can be similar to with salts with the least-sensitive tests, going as low as 0.1 mg/L. The lowest detectable concentration through the most sensitive method utilizing a graphite furnace is 5 µg/L (5E-7 wt %) under normal conditions[53].

#### 1.2.2.2: Volatile Solvents

Solvents such as EtOH are commonly used in cleaning and food products worldwide. However, in high concentrations, and low concentrations in the cases of related solvents, they are harmful to the environment and the organisms within[13,15,56]. A common trait of many of their solvents are their relatively low boiling points when compared to water, allowing for methods involving heat and vaporization to be used for detection, namely with gas chromatography (GC)[57], though methods involving high-performance liquid chromatography (HPLC) can be used for detection as well. HPLC methods tend to have higher detection limits than GC, as well as tending to require more complicated systems and couplings, as well as more optimization than GC [58,59]. For GC, the detection limit can be as low as 1 mg/L for most solvents, including EtOH[57], while HPLC goes down to 6.2E-4 mg/mL (6.2E-5 wt %)[58]. While accurate and sensitive, these methods require laboratory equipment and are not portable.

#### 1.2.2.3: Organic Acids

While many organic acids have similar properties to the organic solvents mentioned above, testing conditions for the two categories are different enough that many methods are not completely interchangeable[57]. For detection of acids, one of the most common methods is finding the pH of a solution, which may be done with titration, pH meters or strips of pH paper[18,60,61]. While these methods are able to detect the general category of contaminants in water and are portable, it has little-to-no specificity and serves solely as an indicator of presence. For specificity, methods such as variations of HPLC and adsorption chromatography are able to determine what kinds of acids are present in water samples[62,63]. While accurate to fairly low concentrations were found for organic acids with these concentrations (200 µg/mL or 0.02 wt %) could be found, these methods require a fair amount of preparation and require laboratory equipment, so field and near-real time tests are not possible.

#### 1.2.2.4: Antibiotics

As an emerging contaminant, the detection antibiotics in wastewater and river water has been a common subject of research. Two common types of antibiotics searched for are the $\beta$-lactams and the aminoglycoside families, due to their widespread use in prescription medicine and in agriculture[34,64–66]. Methods for detection vary, but most involve low-temperature chromatography and different types of spectroscopic measurement[26,27,65–69]. Enzyme assays and indicator tests are also available with varying degrees of accuracy[28,34,64,70,71]. While these tests

are able to detect concentrations of the antibiotics to low concentrations (9 ng/L or 9E-10 wt % in water, 2 pg with SERS, and in the μM range over all other media), most of the tests are not portable, and the ones that are either require tedious preparation or can become costly.

#### 1.2.2.5: Nanoparticles

Like antibiotics, nanoparticles, specifically metal ones like AgNP, are emerging contaminants and a common subject of research[45,47,72]. Detection of these nanoparticles can be similar to metals and salts, such as spectroscopy and measuring conductivity and electroanalysis. It also includes multiple types of microscopy (transmission electron, scanning electron, and atomic force, as examples). These methods are able to detect both small sizes and low concentrations (<1 nm for TEM and 0.2 μg/L or 2E-8 wt % with capillary electrophoresis), many of these methods are still emerging and require robust laboratory equipment[72].

### 1.2.3: Biosensors and Biosensing Techniques

One of the first biosensors created was by Dr. Leland C. Clark, where creating the enzymatic Clark Electrode[73]. One of the most common definitions of a biosensor is "an analytical device incorporating a deliberate and intimate combination of a specific biological element (that creates a recognition event) and a physical element (that transduces the recognition event)."[74]. What this translates to is that a biologically-derived element, most commonly an enzyme, recognizes a specific analyte and is attached to a senor, creating an electrical signal which can be interpreted[74,75]. If a microorganism is used, rather than a specific enzyme, it can be used for a variety of analytes, though it can only be used for one at a time, in most cases[76] due to changes

within a certain enzyme or output of the organism being the signal, which may be muddled by multiple analytes.

### 1.2.4: Raman Spectroscopy

The Raman effect and Raman scattering were first discovered in 1928 by Sir C.V. Raman[77], and this type of scattering has been utilized in analytical work to detect one or multiple analytes. Raman scattering is the inelastic scattering of light from interactions with a molecule's vibrational modes, which alters a photon's energy and wavelength, generally lengthening the wavelength, or lowering the energy. This allows for unique spectral signals to be found for any chemical analyzed by this technique[77,78]. Raman spectroscopy has four main advantages over other forms of spectroscopic analysis: short integration time (near-real time), minimal sample preparation, no spectral interference from water, and its non-destructive and non-invasive properties[79–83].

Though standard Raman spectroscopy is a powerful technique which allows for fast and accurate measurement of samples, it has its limitations. One of the biggest limitations is that the Raman scattering utilized is relatively weak compared to other light sources, and can be blocked out by phenomena such as fluorescence[42,43]. The signal can be enhanced by using a smooth metal surface such as aluminum foil[1,82] roughening metal surfaces and probes[43,83,84], and even utilizing nanoparticles and nanoparticle ligands[27,42–44,83]. All of these methods increase the surface area to either adsorb more chemical or allow for a greater scattering of the light, creating a much more intense signal.

## 1.2.5: Raman Spectroscopy as a Biosensing Technique

### 1.2.5.1: General

As mentioned in the previous section, Raman spectroscopy and SERS have both been recognized as powerful analytical tools with a large range of use and high sensitivity. Coupled with biosensing techniques, these methods have proven to be excellent candidates for becoming broad-spectrum tests that have the power of laboratory-based tests and the portability and ease-of-use of field tests[2]. Examples of this include monitoring of phenotype changes to 1-butanol[3,4], finding new strains of cells for production of value-added chemicals[81], and detection of new antibiotic mechanisms and antibiotic-resistant phenotypes[1].

### 1.2.5.2: Screening Techniques

One of the advantages of Raman technologies are that they are nondestructive and noninvasive. This makes it an excellent candidate for disease and mutation screening. Uses include monitoring of perfused livers to ensure safe transplantation[80], screening for antibiotic resistance and antibacterial mechanisms[1], and finding biomarkers in samples of urine and kidneys to screen for chronic kidney disease and general health[85–87]. These tests are able to be performed in near-real time and are able to continuously occur with little to no damage done to the samples, since it is a low-frequency laser and the samples require little preparation.

### 1.2.5.3: The Rametrix™ LITE Toolbox

While Raman spectra and the data they give have many strengths by themselves, its use within biological materials can be cumbersome due to a number of variables. These include interference from fluorescence, the overlap of bands and band appearing at multiple wavenumber values, and variable concentrations from sample-to-sample, or even within a sample. Multiple methods have been applied to try and solve this, including using chemometric analyses like principal component analysis (PCA) clustering, applying multivariate analysis of variance (MANOVA) to the principal components (PCs) of PCA analysis results for further characterization of the clusters, and discriminant analysis of PCs (DAPC) to provide a visual and logical grouping for clusters and groups. To make interpretation of Raman data easy, yet still robust, the Rametrix™ LITE Toolbox was developed for MATLAB[79]. This software simplifies Raman readings by allowing for spectral comparison and trimming, baseline correction, vector normalization to account for concentration variability, PCA, and DAPC; among other functions. This data can be further analyzed through analysis of variance (ANOVA), MANOVA, and finding distances between cluster groups if so desired.

# Chapter 2: Materials and Methods

## 2.1: Raman Spectroscopy

### 2.1.1: Preparation of cultures

A colony of the NEB 10-beta strain of *E. coli* (New England Biolabs) containing an empty pUC19 vector (New England Biolabs) was picked from a Luria-Bertani (LB) medium (Sigma-Aldrich, Fisher Scientific) agar plate and allowed to grow overnight (12-15 hr) at 37 $^{o}$C in liquid LB medium. The pUC19 vector provided resistance to ampicillin (Fisher Scientific) up to 100 µg/mL. Once grown, 50 mL of LB containing the same concentration of ampicillin was inoculated with 1 mL of the overnight culture. The culture was grown to an $OD_{600}$ value of 1.00 ± 0.05, which generally took around 3 hours in the same shaker. This value was chosen due to it being within the exponential (growth) phase of *E. coli* and to ensure that results were not skewed due to large differences in cell concentration between experiments.

To create the contaminated media, an appropriate amount of contaminant solution was mixed with pure deionized (DI) water and added to 4 mL of LB. The exceptions to this rule were the highest concentration used for a contaminant (e.g., 1 wt % NaCl) and the controls for each run (0 wt % contaminant). The highest concentration represents the highest amount of added liquid. At 0%, the same amount of DI water was added instead of the contaminant to keep the LB/additive ratio constant. This same concept was utilized for values in-between the highest and

control, adding a mixture of contaminated media and sterile DI water to 4 mL of media. An example of this is shown below in table 2-1, and the setup is shown in figure 2-1.

**Table 2-1: Example of Contaminant Concentration Scheme**

Example of how the amount of NaCl, and by extension all contaminants, to add was planned. Subsequent runs were done by making a volume equal to 1 wt %.

| Wt % | 4 M NaCl (mL) | $H_2O$ (mL) |
|---|---|---|
| 1 | 0.1712 | 0 |
| 0.75 | 0.1284 | 0.0428 |
| 0.5 | 0.0856 | 0.0856 |
| 0.25 | 0.0428 | 0.1284 |
| 0.1 | 0.0171 | 0.01541 |
| 0.075 | 0.0128 | 0.1584 |
| 0.05 | 0.0086 | 0.1626 |
| 0.025 | 0.0043 | 0.1669 |
| 0.01 | 0.0017 | 0.1695 |
| 0 | 0 | 0.1712 |

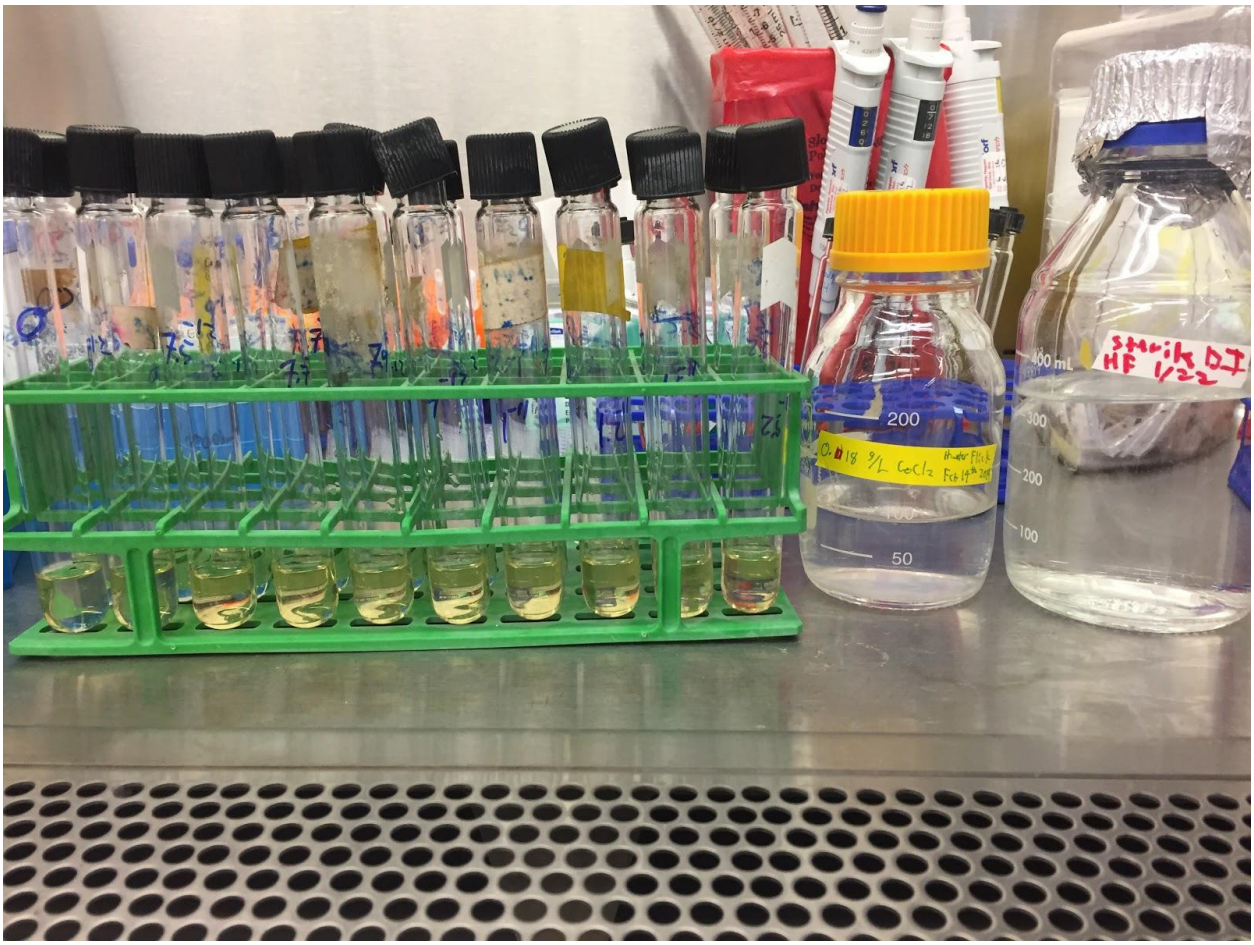


**Figure 2-1: Layout of the tubes used.**

4 mL of LB media are within the tubes, and the appropriate amounts of water and contaminant ($CoCl_2$ in this case) are added based on their concentration labels.

Once the 2 hours passed, aliquots of 0.5 mL were taken and placed into 0.6 mL microcentrifuge tubes. The samples were spun at 5000 RPM for 1.5 minutes to create pellets of the cells. Most of the leftover liquid was decanted into a proper waste container. The leftover cell debris was resuspended in the media using a vortex mixer (VWR) at the maximum setting. After mixing, 5 µL of cells were placed onto a flat piece of cardboard wrapped in smoothed-out

aluminum foil with the “shiny” side up, as shown in figure 2-2 below. The samples were dried at 37 $^{o}$C before Raman scanning.

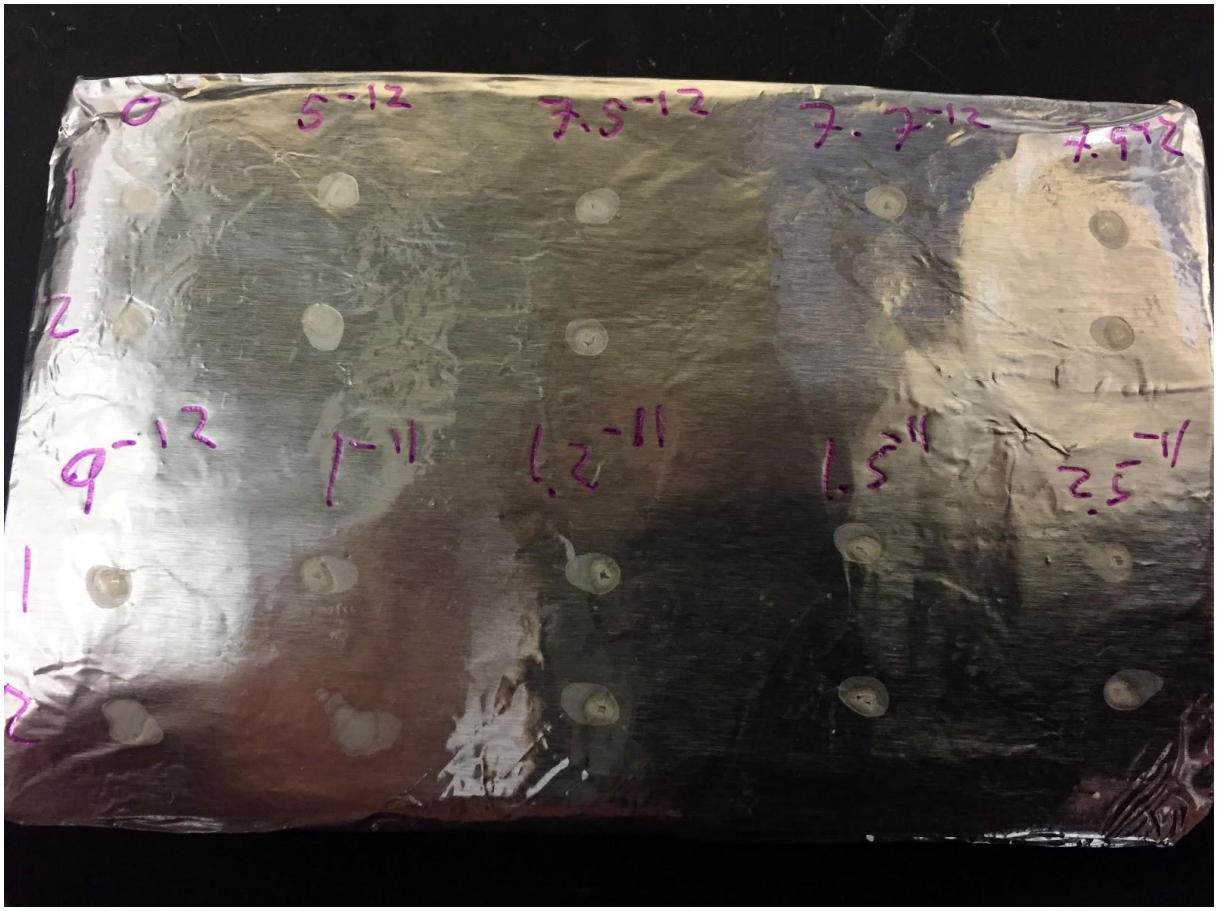


**Figure 2-2: Example of SERS template**

Example of aluminum foil wrapped around cardboard “shiny-side up”, this example being $CoCl_2$. There are 20 samples in total, with each biological replicate set under a concentration label and in rows labeled for which replicate they represent.

## 2.1.2: Raman Readings and Data Processing

When the samples were dried, they were placed under a confocal microscope (AmScope) at the 60X zoom setting with the setup shown below in figure 2-3. This microscope was hooked

up to a Raman spectrometer (Agiltron Peakseeker) which utilizes the software RSIQ (Agilent). Each sample was scanned 5 times for 10 seconds by exposure to a 785 nm laser with a 5 second rest between scans. Since 2 biological replicates were used with each concentration, they were scanned 10 times, giving a total of 100 spectra for the group.

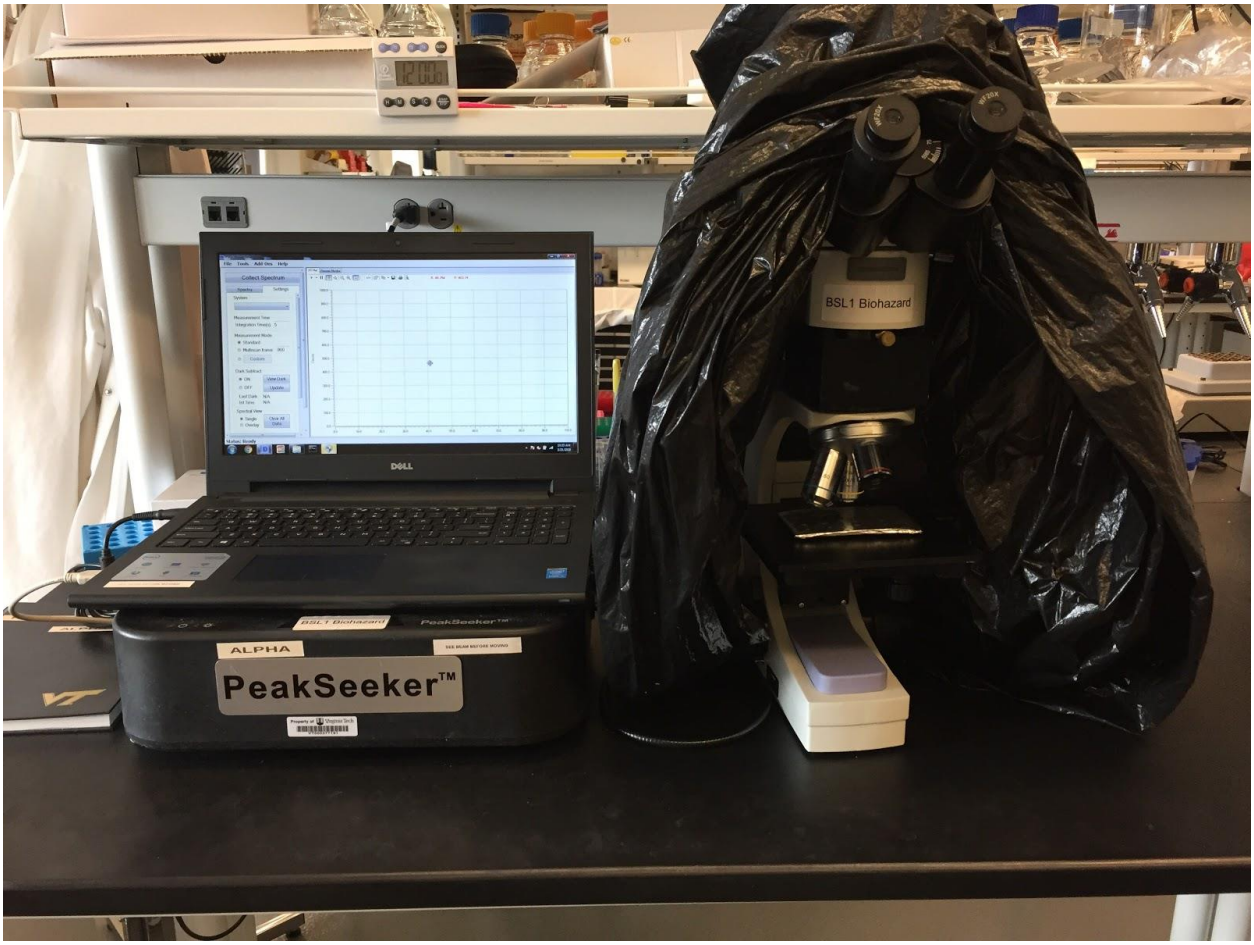


**Figure 2-3: Raman spectrometer setup**

The laser on the Peakseeker (left) is hooked up to the confocal microscope (right). After it is covered, a dark reading is taken to filter out any ambient light pollution, and five scans are taken. The data is then sent to the computer and processed by the program to show peaks, creating a .spc file for use with the Rametrix™ Toolbox.

Once the data was gathered, it was input into a Google Sheet or .csv file to sort and show what each data point means (Example of labels shown in Appendix B). This sheet and the .spc files with the scans were read by the Rametrix™ Toolbox[79]. The spectra (Figure 2-4) were baselined and normalized according to the methods outlined by Fisher et al.[79] in order to account for any discrepancies found in cell concentration. Once run through this process, the window for analysis was truncated to 400-1800 $cm^{-1}$ (Figure 2-5). This is what is known as the biological range, where the spectra of many organic and bio-based compounds are found, and is where all of the phenotypic changes will appear[1,3,4]. For compounds like EtOH, the truncation range of 700-1600 $cm^{-1}$ was used, since the normalized spectra deviates from the baseline too often. The reasoning behind this is laid out in the "Challenges and Solutions" section. Another exception was looking at where samples of water from the field fall. Only the baseline corrections are included, which allowed for concentrations of components within the biological range to be taken into account. This showed where the clusters fell based on what is in the environment, rather than solely based on the chemometric fingerprint.

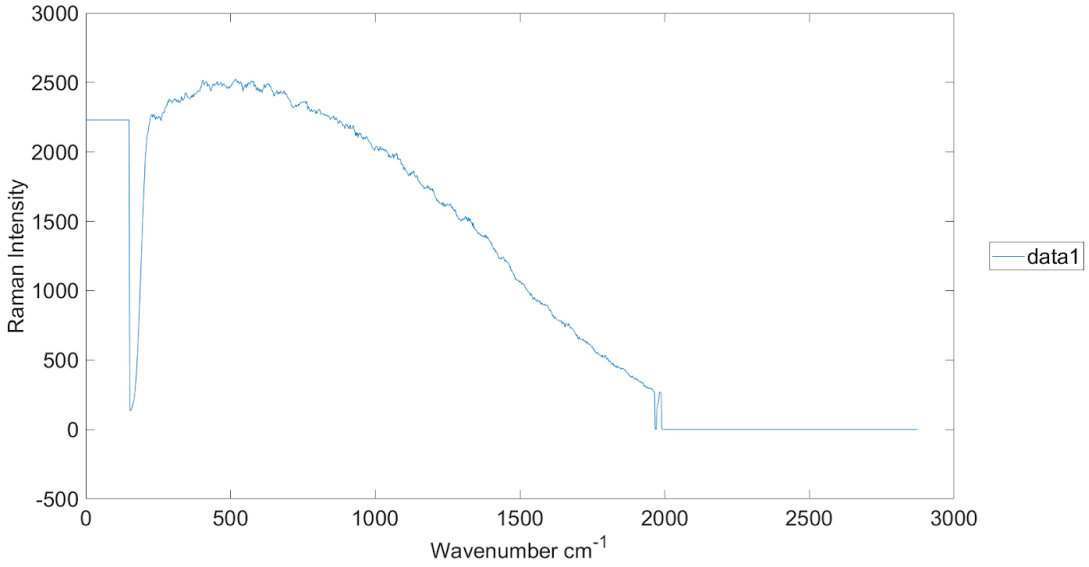


**Figure 2-4: Overall Raman Spectra**

Image of the overall Raman spectra of the control (0 wt %) for NaCl. The full range of the spectrometer is from 200 to 3000 $cm^{-1}$, though most of the bands and activity appear in the range of ~300-2000 $cm^{-1}$.

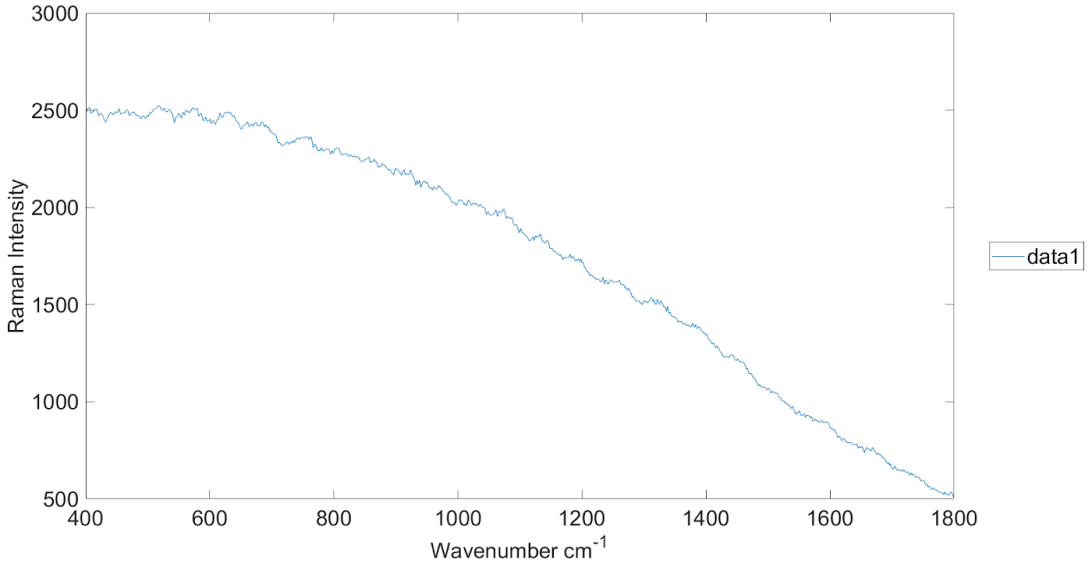


**Figure 2-5: Truncated Raman Spectra**

This is the same spectra as Figure 2-4, but truncated to the "biological range" of 400-1800 $cm^{-1}$. This truncation occurs prior to the baseline and vector normalization step so that only peaks within this range are taken into account.

After normalization, baselining, and truncation (Shown below in figure 2-6), the data was then run through three statistical analyses; PCA, DAPC, and a one-way ANOVA paired with both a Dunnett's comparison test and Tukey's test. The first of these programs involved running a PCA. In short, this analysis allows for easy and powerful analysis of any group of spectra, no matter their relation or how much is known of them. The analysis compared the groups based on their identifiers (chemical concentrations, in this case) and gives an idea of how closely-related they are. The PCA then generated a plot to visually show a general trend of clustering, and where the phenotypes from each scan fall (seen in figure 2-7). When reading the data, only the first two principal component (PC) groups (PC1 and PC2) were utilized, as these accounted for most of the

changes in phenotype. As with the PCA, each dot within the plot represents an entire Raman spectrum, and the clusters show how the average of each concentration's phenotype compares to the others. When clusters group together and overlap, most importantly over PC1, it showed that the signals given by the spectra, or their phenotypes, were statistically the same.

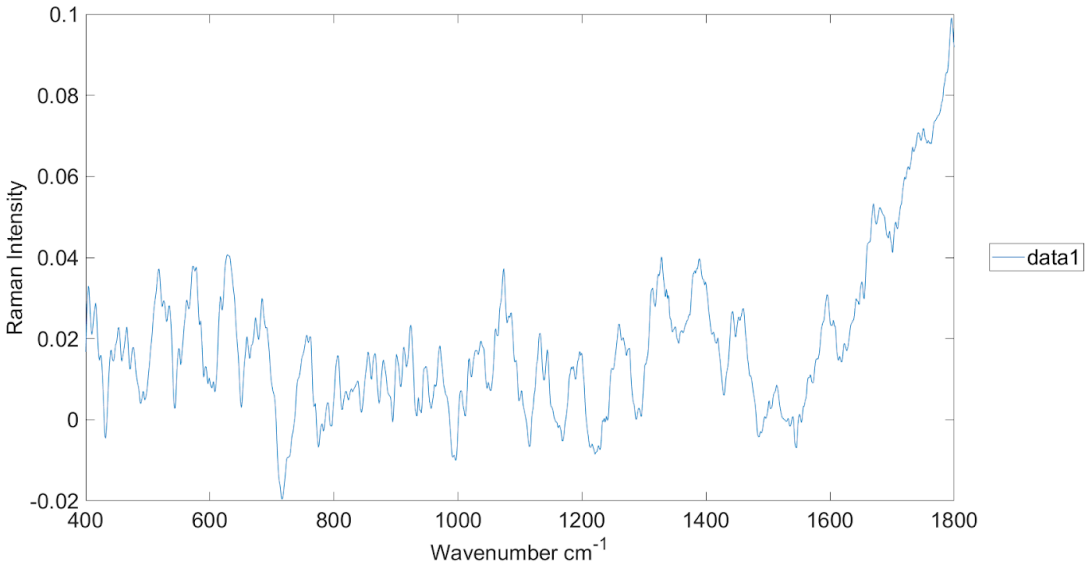


**Figure 2-6: Baseline-Corrected and Vector Normalized Spectra**

The spectra shown in the previous two figures, baselined and vector normalized. This action accounts for any variation in concentration between samples, as it sets a general baseline and trend for each spectra to follow.

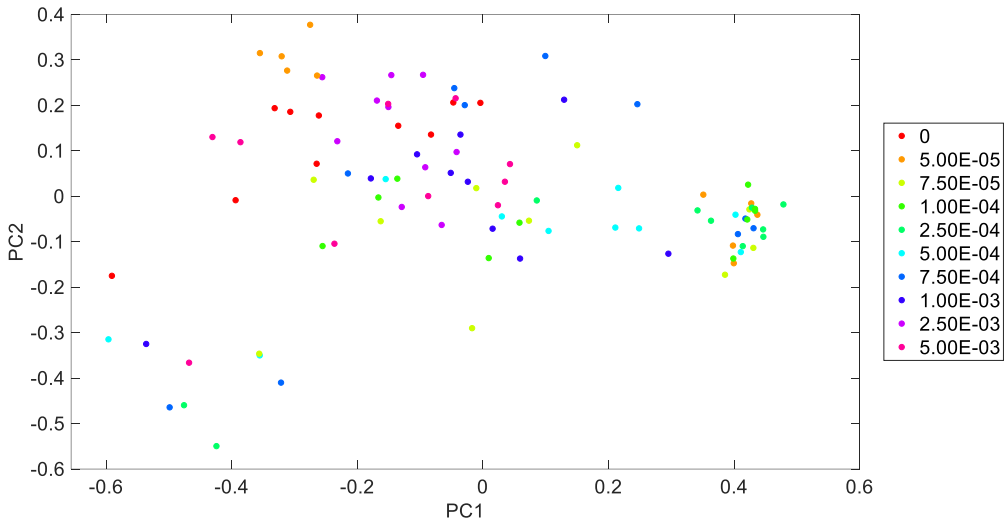


**Figure 2-7: Example PCA**

An example of a PCA plot, showing the results from NaCl with a range from 5.0E-5 to 5.0E-3 wt %, baselined and vector normalized and truncated to 400-1800 $cm^{-1}$. Each dot represents an individual Raman scan, and these scans group together to form clusters. The close proximity of spectra clusters along PC1 and PC2 shows their similarity.

The second analysis utilized was DAPC. As with PCA, this is an analysis method that allows for visualization of groupings, with each singular dot representing a Raman scan. An example is shown below in figure 2-8. This analysis method takes factors, PCs from the PCA in this case, and creates *a priori* groups based on them, giving another visualization of the data and phenotype variance. For the scans performed, half of the total number of scans read (50 for single-component tests, 20 for multicomponent and impure tests) were set to be the number of significant PCs, which accounted for variability within the data. The advantage of using DAPC after the PCA is that separation was maximized, allowing for a clearer picture of what concentrations induced noticeable phenotype changes. For single-component tests, as with the PCA, the first two canonicals were the only ones taken into account for this measurement, as

they show the majority of changes within the phenotype. For the multicomponent and impure water tests, all three were utilized to show the total canonical distance (TCD), which is detailed within Chapter 3.1.9. When a concentration cluster overlapped, specifically along canonical 1, it signifies that there is no significant change in phenotypes.

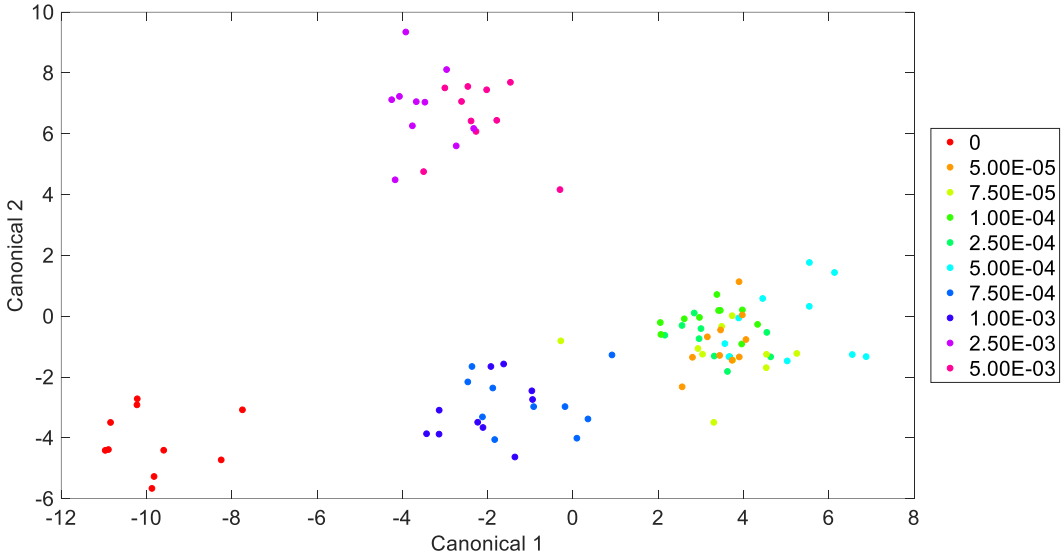


**Figure 2-8: Example DAPC**

A DAPC plot made from the PCA shown in figure 2-7, utilizing 50 PCs to account for 98.69 % of the variability. As with the PCA, each dot represents an individual Raman scan, and scans of the same kind cluster together to form groups. Similarity, is shown by different clusters grouping together, particularly along canonical 1.

The final analyses used were one-way ANOVA and comparison tests along canonical 1 of the DAPC data. Two of these tests were performed, with the first being an overall ANOVA paired with a Dunnett's many-to-one comparison test. This test compares the averages of all clusters to one, specifically to the control. The other test performed was one-way ANOVA on each individual concentration compared to the control, paired with a Tukey's comparison test, with the code for

it shown in Appendix A. The main purpose of these tests was to confirm results inferred from the DAPC and PCA performed earlier. The robustness of ANOVA and the power the two comparison tests added gave greater confidence in showing the significance of difference between the control and lower phenotypes.

# 2.2: Chemical Preparation

## 2.2.1: Choosing Chemical Contaminants and Starting Concentrations

When choosing the contaminants used, multiple factors were taken into account; including where they are used, how commonly used they are, and the possibility of its existing as an environmental contaminant. These factors led to a baseline containing 5 categories: common household food items (NaCl and EtOH) as a "proof-of-concept", common organic acids (CA and AA), antibiotics (kanamycin and ampicillin), heavy metals ($CoCl_2$), and nanoparticles (100 nm AgNP).

Once the contaminants were chosen, the next task was to determine the highest value to start at. This involved looking up the minimum inhibitory concentration (MIC) values of each contaminant within *E. coli* and related prokaryotic organisms. If the values were equal to or below 10 wt % in media, 5 % of the MIC was the starting point. This number was chosen to ensure that the phenotype would still change, but cell growth and replication would not be hindered to a point where no changes occur. If the MIC was greater than 10 wt % or could not be found, 0.5 wt % was utilized as a starting point, as it gives a sufficient amount of inhibition as to alter the phenotype, but will not inhibit growth to a point of no major changes occurring.

### 2.2.2: NaCl

The initial stock solution of NaCl (Fisher Scientific) had a concentration of 4 M in DI water, which led to adding 171 $\mu$L to 4 mL of LB media to make a 0.958 wt % solution. The concentration of the stock was halved as needed for further experiments.

### 2.2.3: Ethanol

The initial stock was a 70 volumetric percent (% v/v) solution of EtOH (Decon Laboratories, Inc.) in DI water, which led to adding 28.57 $\mu$L to 4 mL of LB media to make a 4.6E-3 wt % solution. The concentration of the stock was halved as needed for further experiments.

### 2.2.4: Citric Acid

The initial stock was a 0.5 M solution of CA (Sigma-Aldrich) in DI water, which led to adding 416 $\mu$L to 4 mL of LB to make a 0.905 wt % solution. The concentration of the stock was halved as needed for further experiments.

### 2.2.5: Acetic Acid

The initial stock was a 1 M solution of AA (Fisher Scientific) in DI water, which led to adding 666 $\mu$L to 4 mL of LB to make a 0.857 wt % solution. The concentration of the stock was halved as needed for further experiments.

### 2.2.6: Kanamycin

The initial stock was a filtered 12.5 $\mu$g/mL solution of kanamycin (Fisher BioReagents) in DI water, which led to adding 16 $\mu$L to 4 mL of LB to make a 5.0E-3 wt % solution. The concentration of the stock was halved as needed for further experiments.

### 2.2.7: Ampicillin

The initial stock was a filtered 100 $\mu$g/mL solution of ampicillin in DI water, which led to adding 40 $\mu$L to 4 mL of LB to make a 9.9E-5 wt % solution. The concentration of the stock was halved as needed for further experiments.

### 2.2.8: $CoCl_2$

The initial stock was 180 mg/L solution of $CoCl_2$ (Acros Organics) in DI water, which led to adding 22.22 $\mu$L to 4 mL of LB to make a 9.9E-5 wt % solution. The concentration of the stock was halved as needed for further experiments.

### 2.2.9: Silver Nanoparticles

The initial stock was a bottle of 100 nm AgNP (Sigma-Aldrich) at a concentration of 0.2 mg/mL. This led to adding 200 $\mu$L to 4 mL of LB to make a 1.0E-3 wt % solution. The concentration of the stock was halved as needed for further experiments.

### 2.2.10: Multicomponent Tests

Using the aforementioned stocks of kanamycin, $CoCl_2$, and AA, the same volume of solutions used for 5E-6, 5E-6, 2.5E-2 wt %, respectively, were added. These values were chosen because they fell evenly between the highest and lowest concentrations of their tested values. This ensured that a phenotype change will be seen without the risk of inhibition due to three contaminants at once.

For the second set of multicomponent tests, the previous stock of $CoCl_2$ was utilized. Along with that, $ZnCl_2$ (Fisher) and $MgCl_2$ (Fisher) stocks of 0.025 and 0.038 g/L were made in DI water, respectively. The concentrations used were 5.0E-6 wt % for all and 5.0E-5 wt % $CoCl_2$ with 5.0E-5 wt % for $ZnCl_2$ and $MgCl_2$, chosen for similar reasons to the previous multicomponent test.

### 2.2.11: Impure Water Test

#### 2.2.11.1: Stagnant Water Source

For the stagnant water source, the Virginia Tech Duck Pond was chosen for use. The reasoning for this was due to how popular it is for walking and outdoor activities, meaning that it and those in the area are regularly exposed to any waste brought in by people and animals. The location, shown in figure 2-7 below, was chosen due to being close to the midpoint of the pond, as well as being near a viewing platform, which placed it in proximity to an area where waste may enter the pond. The sample was also taken from shore, due to its proximity to any contamination and being relatively stagnant compared to the mouths of streams coming in and out.

**Figure 2-9: Stagnant Water Location**

The location on Virginia Tech's Duck Pond where the water sample was taken. The sample was taken by the water's edge.

To collect the sample from the Duck Pond, a sterile 500 mL glass bottle (Schott) was filled with water near the shoreline of the pond. The bottle was rinsed with pond water and emptied away from the site twice to ensure that any material originates from the pond. Once filled, the sample was transported to the lab on ice and vacuum-filtered it through a 0.45 nm membrane (VWR). After filtration, 20 mL of the water was transferred to a 30 mL test tube while the rest was stored at 4 $^{o}$C. LB powder (Fisher Scientific) was added to it and it was autoclaved at 121 $^{o}$C

for 20 minutes. The autoclaved media was aliquoted into 4 mL samples and tests were run in a similar manner to prior contaminants.

#### 2.2.11.2: Running Water Source

Stroubles Creek was the location chosen for the running water sample, taken in the location shown in figure 2-8. The reasoning behind the stream was due to how well-studied it is within Virginia Tech, being a local watershed and source of water. The sampling location was chosen due to ease of access and the surrounding areas. The location was also in close proximity to a main road overpass, a backroads area, and farms for horse and cattle. From these, it was concluded that there was a fair amount of contamination from various sources.

**Figure 2-10: Running Water Location**

The site from which the sample was taken. The sample was taken from the center of the creek, from upstream.

To collect the sample, a sterile 500 mL bottle (Kimax) was filled with water from the stream. The bottle was rinsed with stream water and emptied downstream twice to ensure that all material originates from Stroubles Creek. After filling post-rinse, the bottle was kept on ice to return to the laboratory. This prevented any growth or chemical changes due to a change in temperature. Once at the laboratory, the water was vacuum-filtered through a 0.45 nm membrane. After filtration, 20 mL of the water was transferred to a 30 mL test tube while the

rest was stored at 4 $^{o}$C. LB powder was added to it and it was autoclaved at 121 $^{o}$C for 20 minutes. The autoclaved media was then aliquoted into 4 mL samples and tests were run in a similar manner to prior contaminants.

#### 2.2.11.3: Tap Water

To create the tap water samples, 20 mL of tap water was filtered through a 0.45 nm membrane and transferred into a 30 mL test tube. LB powder was added and it was autoclaved at 121 $^{o}$C for 20 minutes. The sample was stored at 4 $^{o}$C until testing, with 4 mL being aliquoted to sterile test tubes as with the other two impure samples.

## 2.3: Challenges and Solutions for Sample Preparation and Reading

### 2.3.1: Problems with Baselining using the Rametrix™ LITE Toolbox

When comparing the phenotypes within NaCl and EtOH, the PCA showed a high number of outliers (seen in figure 2-7), as well as strangely erratic DAPC charts (figure 2-8). Due to consistency with varying methods of sampling, the data processing became the next focus of the error analysis. There, the data for many of the samples showed some form of deviation from the baseline, as is shown in figure 2-9. This became a concern, as the data should fit as it does in figure 2-10. While some deviation and some outliers do not pose a problem, as can be seen in further trials of NaCl, the trials for EtOH show that the variations from the baseline are too great and too frequent to show good results.

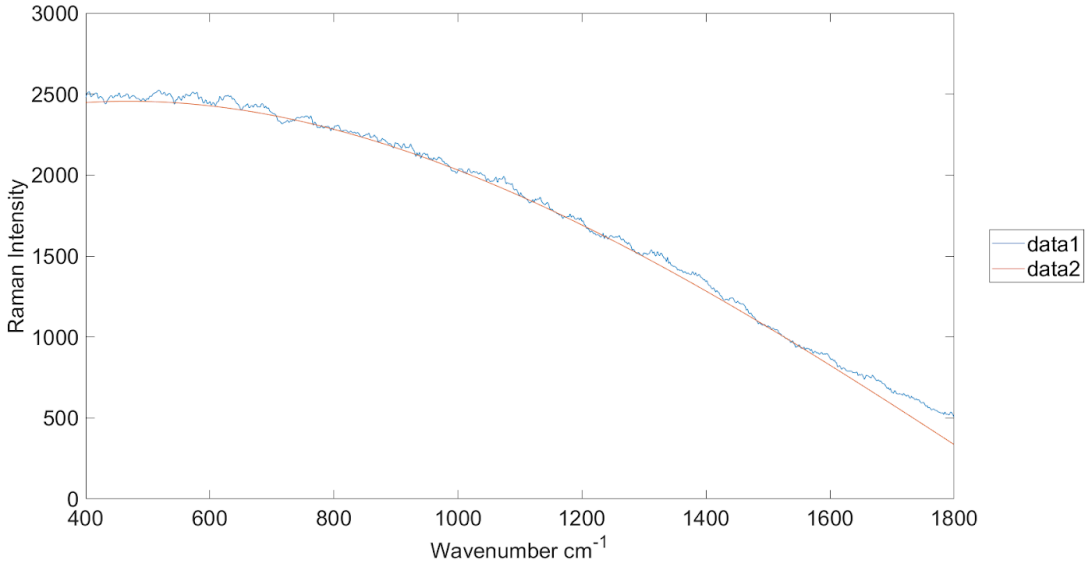


**Figure 2-11: Spectral Deviation from the Baseline**

An example of the baseline not fitting the curve from NaCl at 0 wt %. This is the most common variation, where it trails off on the right-hand side.

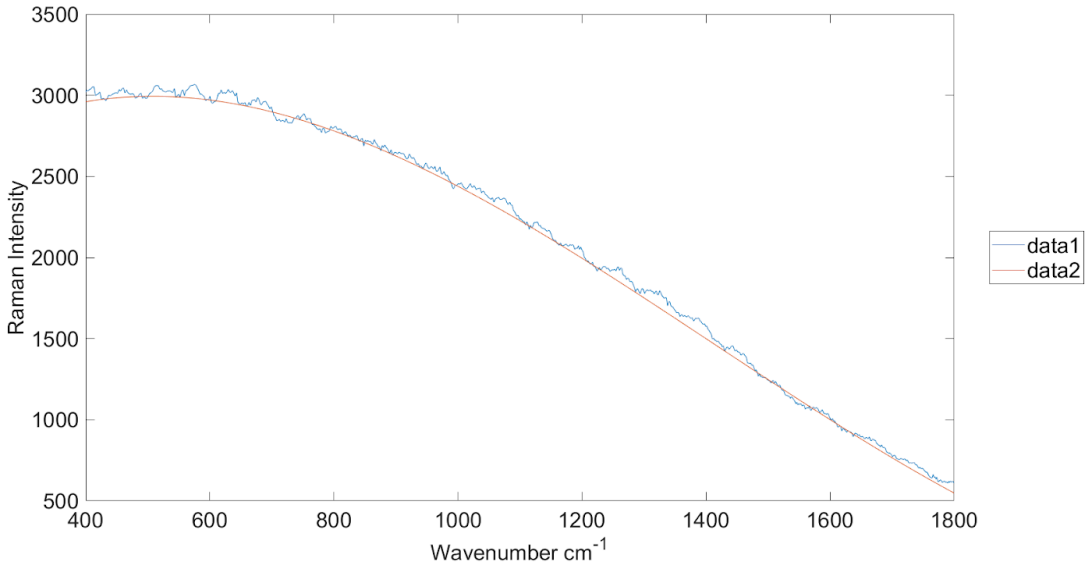


**Figure 2-12: Spectra Fitting the Baseline**

Example of a normalized spectra which follows the baseline closely. Spectra from NaCl at 5.0E-9 wt %.

In an attempt to remedy the problem, the wavenumber was truncated further, as most of the variation appeared prior to 700 $cm^{-1}$ and after 1600 $cm^{-1}$. An example of a truncated baseline is shown in figure 2-11. This seemed to solve the issue of data not acting properly and lining up with what is perceived from the PCA.

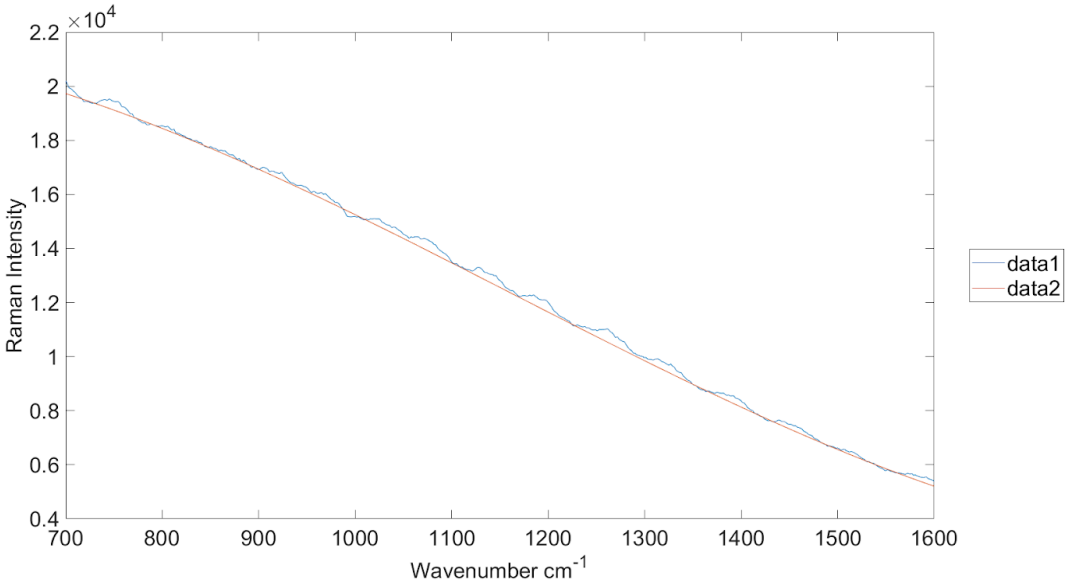


**Figure 2-13: Baseline of Spectra Truncated Further**

Example of the baseline of EtOH at a concentration of 5.0E-6 % v/v truncated to only show the results from 700 $cm^{-1}$ to 1600 $cm^{-1}$.

## 2.3.2: Effect of Inoculation Time on Phenotype

The next issue noticed was the time taken to inoculate the contaminated media with *E. coli*. This became a problem due to the phenotypes being different at a far lower wt % than should be possible, given that there was less than 1 molecule in the entire sample. The lowest percentage found was 2.5E-22 wt % kanamycin, which corresponded to 1.2E-2 molecules within the solution.

To find the source of the problem, “dummy” tests were performed. These tests involved using the same protocol as the real tests, but adding only DI water to the contaminated media. This would allow for visualization of data being skewed or phenotypes changing during the process, since the phenotypes should remain the same with only water. Upon multiple “dummy” tests being completed, it was found that the phenotypes changed according to the order in which they were added to the tubes. An example of this pattern is shown below in figure 2-12.

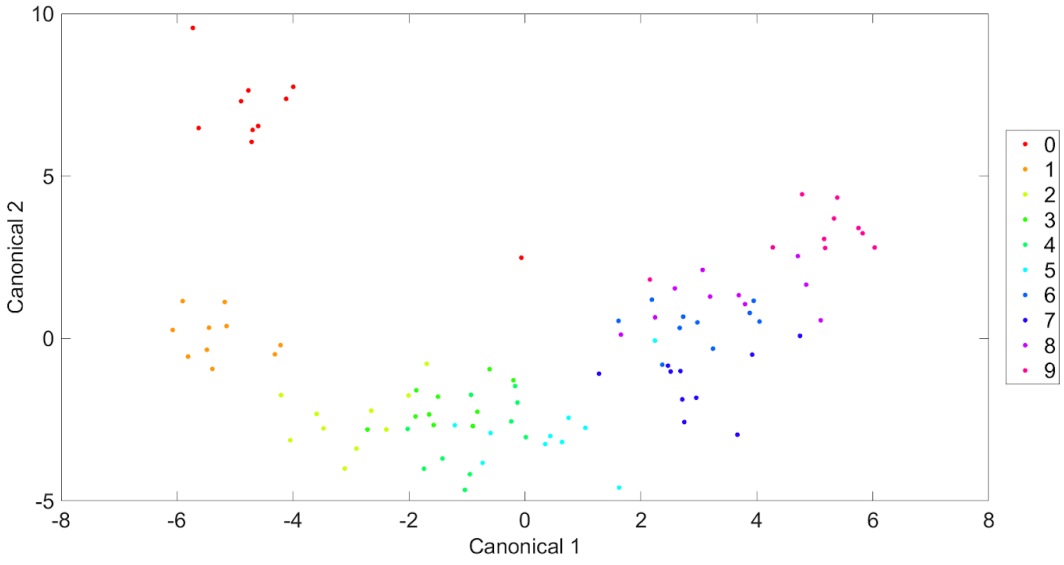


**Figure 2-14: Example of the DAPC in a Slower Inoculation Method**

The general DAPC pattern seen on the initial “dummy” tests. The pattern can be seen starting with “0”, representing the control, and a clear path of phenotypes going up to “9”, representing the highest concentration in the batch.

In an attempt to remedy the problem, changes to the inoculation method were made. Initially, a 100-1000 $\mu$L pipette (Eppendorf) was used to inoculate 4 mL of the flask culture into the contaminated media. Due to the trend seen in figure 2-11, it was thought that the inoculation

of the media may be too slow, and so further tests were made using a serological pipet (Fisherbrand). Upon using this method, it was found that the tests showed better results, shown in figure 2-12 below.

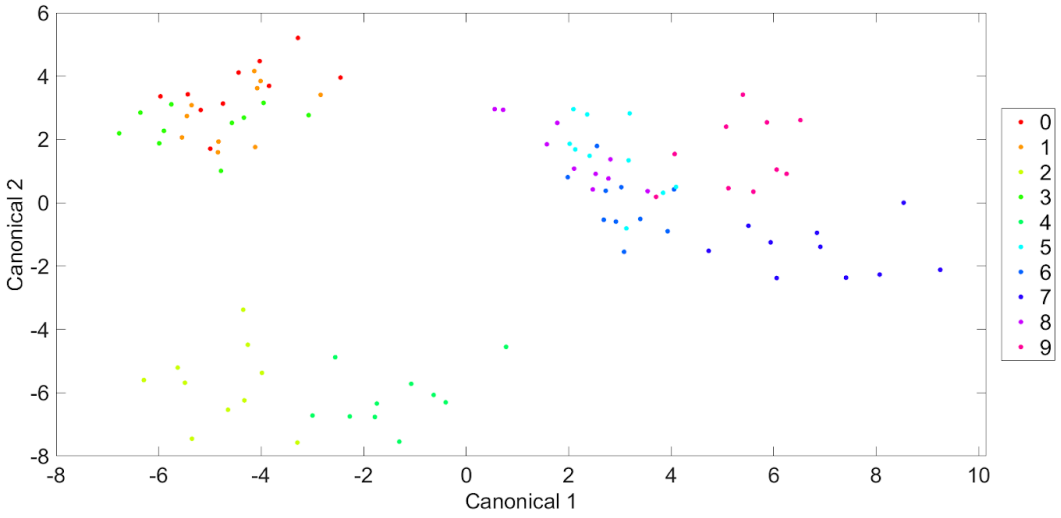


**Figure 2-15: DAPC of a Faster Inoculation Method**

DAPC of water dummy after fixing the issue of inoculation time. The clustering is much better, and while variation does exist with time, it can be seen that a faster method of inoculation allows for less change.

### 2.3.3: Implementation of One-Way ANOVA and Comparison Tests

While looking at the PCA and DAPC data, it was found that differences and similarities could not always be easily distinguished, due to outliers within a sample group or placement of groups on the plot, such as the PCA shown in figure 2-7. To remedy this, it was decided that additional statistical analyses needed to be utilized. One-way ANOVA was chosen due to its robustness when assumptions such as normal distributions are violated, as well as the kinds of follow-up comparison tests which may be utilized. The first ANOVA used compared individual

concentrations to the control (e.g. 0 to 5.0E-5 wt %, shown in table 2-2) and was followed by a Tukey comparison test (figure 2-16). This test proves to be useful for pairwise comparisons and allows for a more accurate determination of the similarity between two groups, and can easily be visualized[88].

**Table 2-2: Individual ANOVA Output**

Utilized on the same data as figures 2-7 and 2-8 to compare the control (0 wt %) to 5.0E-5 wt %. This output shows the sum of squares (SS), degrees of freedom (df), mean square (MS), F-statistic, and the probability of difference (p-value). The overall group numbers, error, and total values of all are shown.

| Source | SS | df | MS | F | Prob>F |
|---|---|---|---|---|---|
| Groups | 890.088 | 1 | 890.088 | 1288.76 | 5.08754E-18 |
| Error | 13.039 | 18 | 0.724 | N/A | N/A |
| Total | 903.127 | 19 | N/A | N/A | N/A |

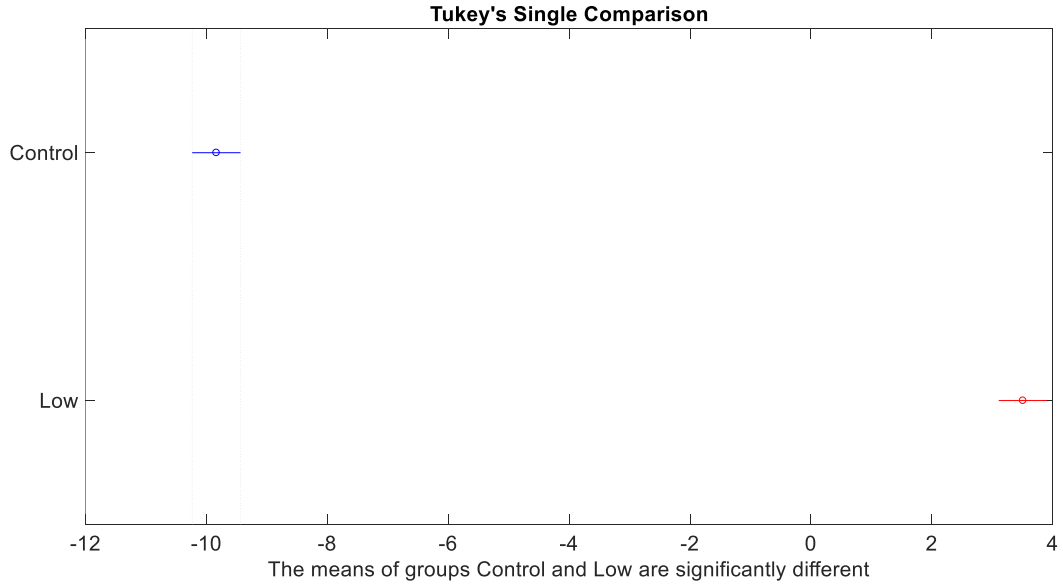


**Figure 2-16: Example of Tukey Comparison**

This test was performed on the one-way ANOVA to verify and visualize the difference between the control and 5.0E-5 wt % NaCl. The control is colored blue while 5.0E-5 wt % is colored red.

Along with the pairwise comparison of individual groups, an overall comparison of every group to the control proved to be useful as well. The overall comparison involves running a one-way ANOVA over the entire data range (output shown in table 2-3), followed by a Dunnett's many-to-one comparison test (output shown in table 2-4). This test works well when comparing multiple sample means to one (e.g. 9 concentrations compared to a single control), and outperforms the pairwise Tukey test. Tukey is a powerful analytical tool for comparing individual groups to each other; however, comparing many samples to one with this test inflates the chance of false positives, known as type I error[88–90].

**Table 2-3: Overall ANOVA Output**

An example ANOVA performed on the range of data shown in figure 2-7 and 2-8. This covers the entire range of data. The output labels are the same as in table 2-2.

| Source | SS | df | MS | F | Prob>F |
|---|---|---|---|---|---|
| Groups | 1913.2 | 9 | 212.578 | 212.58 | 1.2286E-56 |
| Error | 90 | 90 | 1 | N/A | N/A |
| Total | 2003.2 | 99 | N/A | N/A | N/A |

**Table 2-4: Output of the Dunnett's Function**

An example of the Dunnett's comparison of many to one output on the data from table 2-3. The control group shows up as N/A due to it being unable to compare to itself.

| Concentration (wt %) | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 0 | 5.0E-5 | 7.5E-5 | 1.0E-4 | 2.5E-4 | 5.0E-4 | 7.5E-4 | 1.0E-3 | 2.5E-3 | 5.0E-3 |
| P-Value | N/A | 0.27E-5 | 0.27E-5 | 0.27E-5 | 0.27E-5 | 0.27E-5 | 0.27E-5 | 0.27E-5 | 0.27E-5 | 0.27E-5 |

While the Dunnett's test takes less computing power and only requires one overall ANOVA rather than many individual for each Tukey test[88], it was deemed necessary to use both tests to further analyze the data. This was decided in order to further confirm results seen in the PCA and DAPC. To run both tests, a script was developed in MATLAB (shown in Appendix A) which utilizes an overall one-way ANOVA with a Dunnett's test as well as nine one-way ANOVA tests with a Tukey test. Both ANOVA tests utilized the function "anova1" with groups of 10 defined, the Dunnett's T-test utilized a custom function developed by Navin Pokala in 2012[90] due to no built-in functions existing in MATLAB, while the Tukey test utilized the function "multcompare". The tests were run solely along canonical 1 of the DAPC data due to it showing the most significant changes. Once run, the results can be observed to see if all tests agree on the significance of the changes to phenotype. If the results of the statistical tests are inconsistently insignificant or consistently significant differences, further tests were run to find the lower limit. When all of the tests show insignificant differences from the control (0 wt %), the highest concentration showing this is considered the limit for detection.

# 2.4: HPLC

## 2.4.1: Parts

The units on the HPLC included a LC-10AD VP pump, a SCL-10A VP controller, a SIL-20AC autosampler, a CTO-10A VP oven with an Aminex HPX-87H ion exclusion column (Bio-Rad), and a RID-10A refracrive index detector (RID). All parts aside from the column are made by Shimadzu. The overall setup is shown below in figure 2-13.

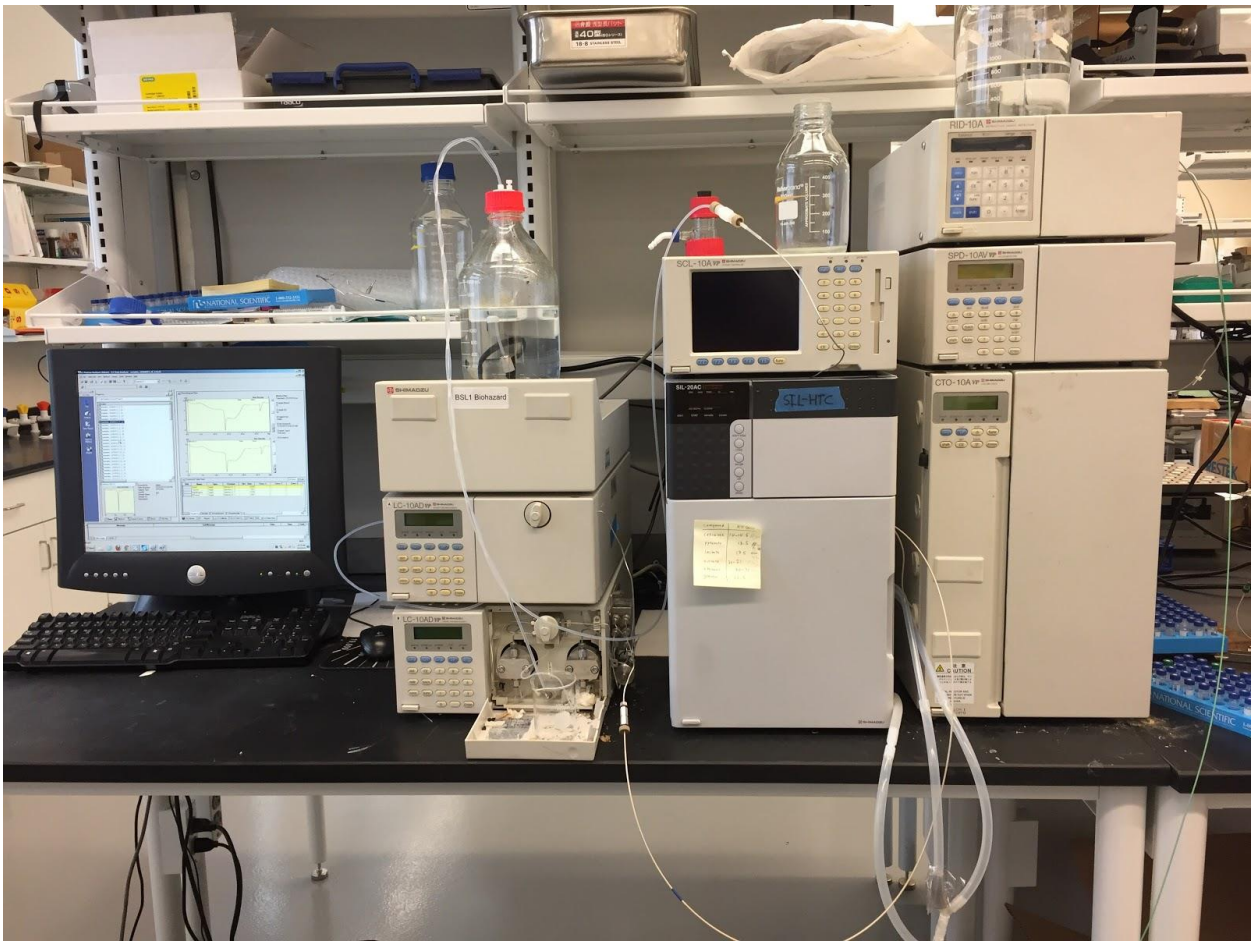


**Figure 2-17: The overall layout of the HPLC machine**

Clockwise from top-right: RID, UV-Vis detector (not used for this study), column oven, autosampler, pump, controller. The controller is hooked up to the desktop computer on the far-left side of the image and the column is inside of the oven.

## 2.4.2: Preparation of Carrier Liquid

The carrier liquid used was 5 mM $H_2SO_4$ (Spectrum Chemical). The solution was made by adding 1 mL of 10 N (5 M) $H_2SO_4$ to 1 L of DI water that has been vacuum-filtered through a 0.45 nm membrane. The liquid was degassed by placing it under a vacuum on top of a hot plate at 70

$^{o}$C and stirred with a bar at ~1100 RPM anytime from overnight to 2 days. This was done to ensure that no air bubbles are left in the liquid to interfere with the column packing or RID.

### 2.4.3: Running the HPLC for Detection of Organic Materials

After turning the machine on, the pump was purged for 3 minutes to ensure that air bubbles were removed and replacing old carrier with new. Once purged, the autosampler was purged for 5 minutes to replace the old carrier with new, remove any trapped air, and to ensure that any remaining sample from the last run was completely cleared from the HPLC. These steps were performed in this order to ensure that bubbles were completely cleared from each unit in-line. After purging the autosampler, pure carrier was allowed to flow through the reference cell of the RID for 15 minutes in order to purge any air bubbles that may have formed, as well as update the reference for the RID in case any changes occurred between runs.

After all parts were purged, the system was turned on so that flow is 0.1 mL/min. When carrier is flowing through, the outlet from the sampler was attached to the bottom of the column. Once liquid was flowing out of the top of the column, it was attached to the inlet of the RID. Over the course of 30 minutes, the flow through the entire machine was increased to 0.5 mL/min and the temperature of the oven was increased to 60 $^{o}$C. Each sample was run through a 30-minute cycle to allow for any sample to completely flow through the column and detector.

# Chapter 3: Results

## 3.1: NaCl

It was found that *E. coli* phenotype changed due to NaCl within the media were detectable to a concentration of 5E-9 wt %, shown in the PCA (3-1 A) and DAPC (3-1 B) results. For the DAPC, 50 PCs were used to account for any variability within the system, translating to 98.44% dataset variability explained by the PCs. This amount of variability showed a clear difference between groups of phenotypes along canonical 1 (Figure 3-1 B) while still allowing for similarities within the groups to be found. The results seen with the DAPC were also shown through ANOVA and a Dunnett's comparison test along canonical 1, with table 3-1 showing that there was no significant difference ($p > 0.05$) between the control and concentrations ranging from 1.0E-9 to 5.0E-9 wt %. This put the detection limit for NaCl at 5.0E-9 wt % when measuring changes in *E. coli* phenotypes by Raman spectroscopy and processing the data with Rametrix™.

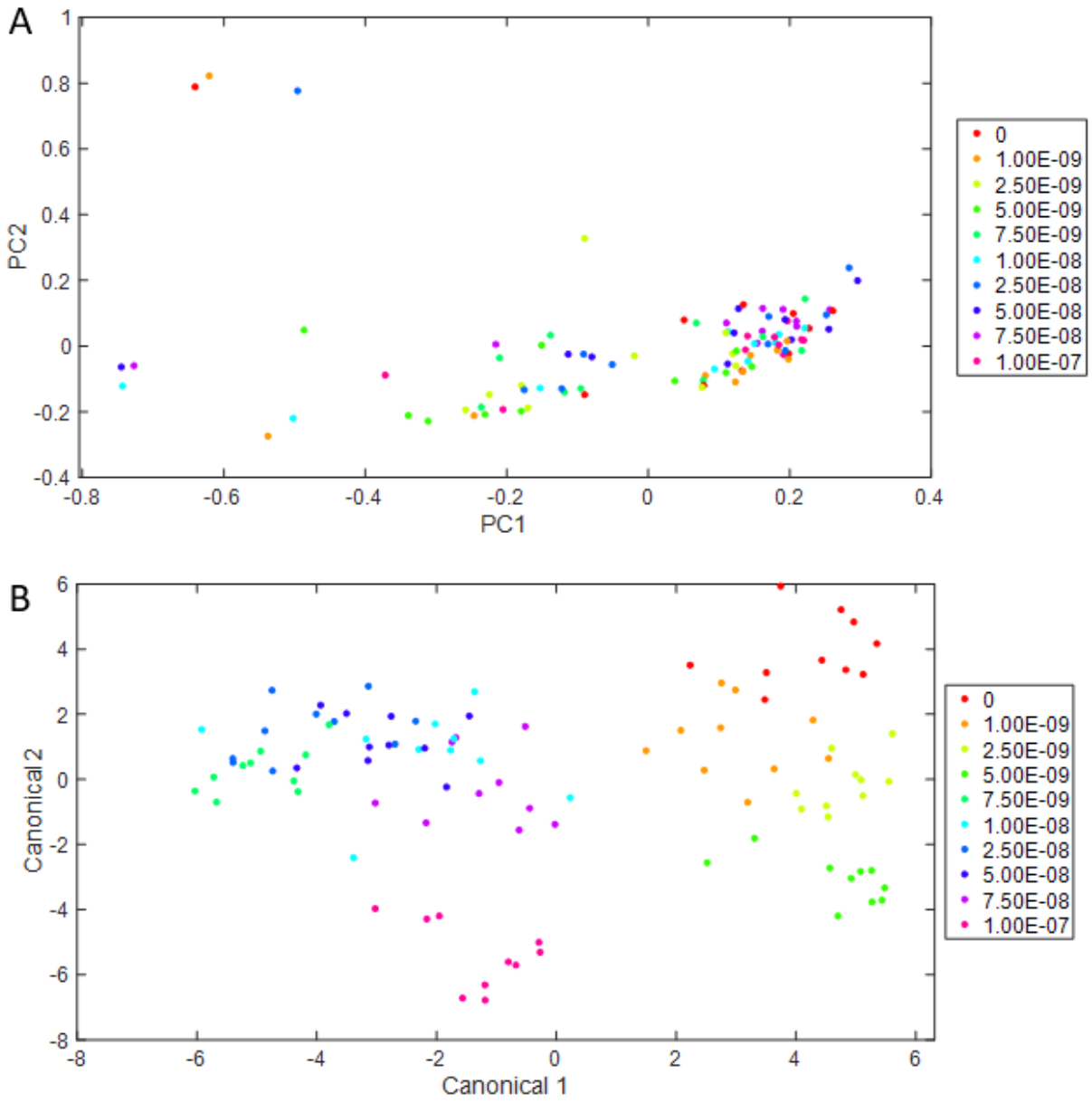


**Figure 3-1: PCA and DAPC of NaCl to 1.0E-9 wt %**

**A.** PCA of NaCl at a concentration range from 1.0E-7 to 1.0E-9 wt %. **B.** DAPC of the same NaCl data, utilizing 50 PCs to account for variability within the system.

**Table 3-1: NaCl P-Values**

P-values from the Dunnett's test showing similarities in the phenotypes of NaCl concentrations ranging from 1.0E-7 to 1.0E-9 wt %.

| Concentration (wt %) | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 0 | 1.0E-9 | 2.5E-9 | 5.0E-9 | 7.5E-9 | 1.0E-8 | 2.5E-8 | 5.0E-8 | 7.5E-8 | 1.0E-7 |
| P-Value | N/A | 0.0509 | 0.7449 | 0.9385 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |

# 3.2: Ethanol

Due to the variation of EtOH and its regular deviation from the baseline shown by the software, the wavenumber range was truncated from 700-1600 $cm^{-1}$ before the PCA and DAPC were run (Figure 3-2 A and B). Within this range, the lowest detectable concentration of EtOH was 7.5E-7 % v/v (5.9E-8 wt %). For this final test, 50 PCs accounted for 97.74 % of the variability within the PCs. This, along with the ANOVA and Dunnett's comparison in table 3-2, show that there was no statistical difference ($p > 0.05$) between the control and a concentration of 5.0E-7 % v/v. This put the detection limit for EtOH at 5.0E-7 % v/v when measuring changes in phenotypes by Raman spectroscopy and processing the data with Rametrix™.

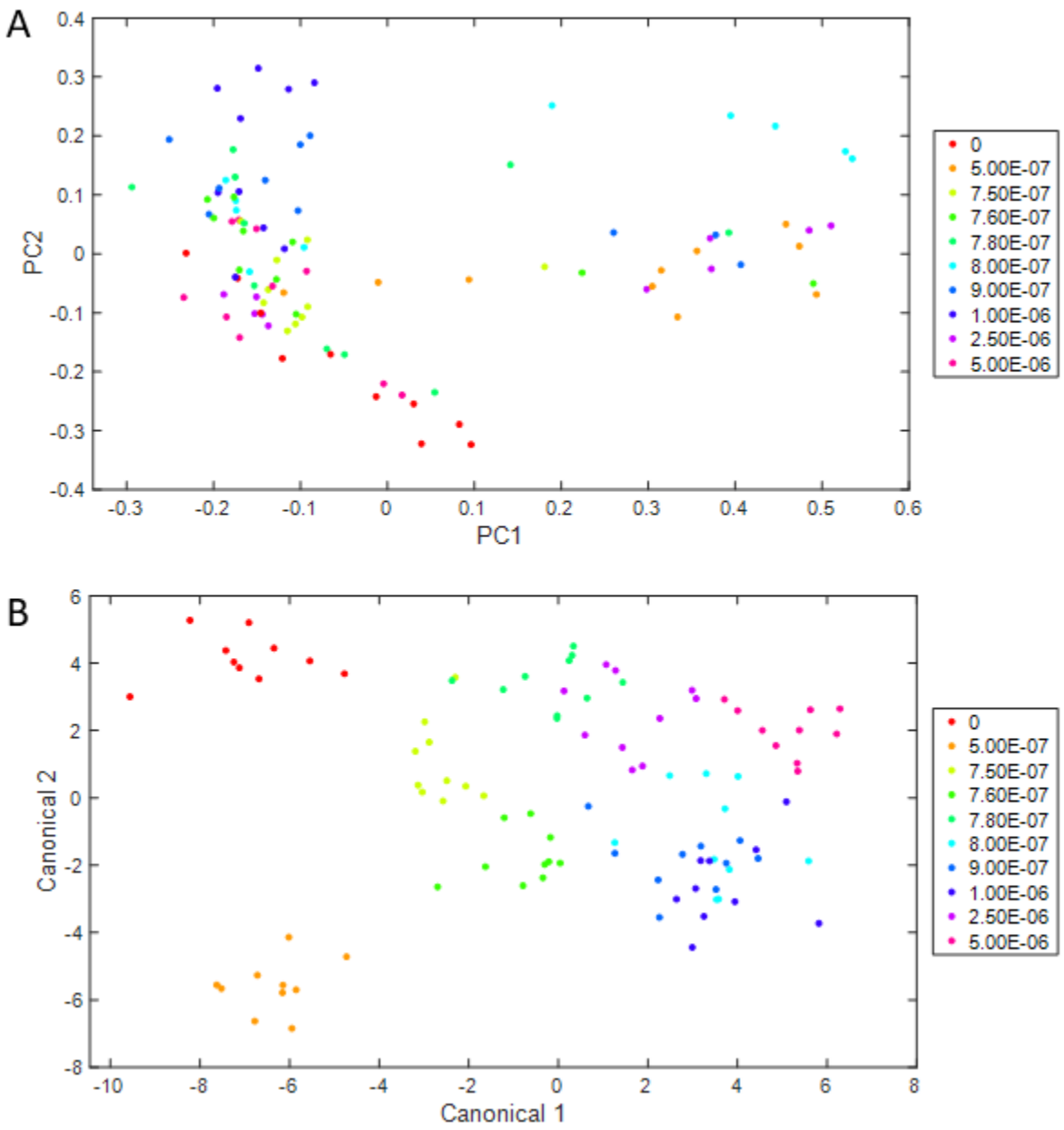


**Figure 3-2: PCA and DAPC of EtOH to 5.0E-7 % v/v**

**A.** PCA of EtOH at a concentration range from 5.0E-7 to 5.0E-6 % v/v. **B.** DAPC of the same EtOH data, utilizing 50 PCs to account for variability within the system.

**Table 3-2: EtOH P-Values**

P-values from the Dunnett's test showing similarities in the phenotypes of EtOH concentrations ranging from 5.0E-7 to 5.0E-6 % v/v.

| Concentration (% v/v) | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 0 | 5.0E-7 | 7.5E-7 | 7.6E-7 | 7.8E-7 | 8.0E-7 | 9.0E-7 | 1.0E-6 | 2.5E-6 | 5.0E-6 |
| P-Value | N/A | 0.6361 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |

To show the effectiveness of the Raman spectroscopy method compared to traditional methods, HPLC analysis was performed on EtOH utilizing a column for organic acids and solvents. The lowest detectable concentration was 9.5E-4 % v/v (5E-4 wt %) in filtered DI water, as is shown in figure 3-3. This shows that the Raman methods outlined here are similar, if not better, than a standard laboratory-based detection method such as HPLC analysis.

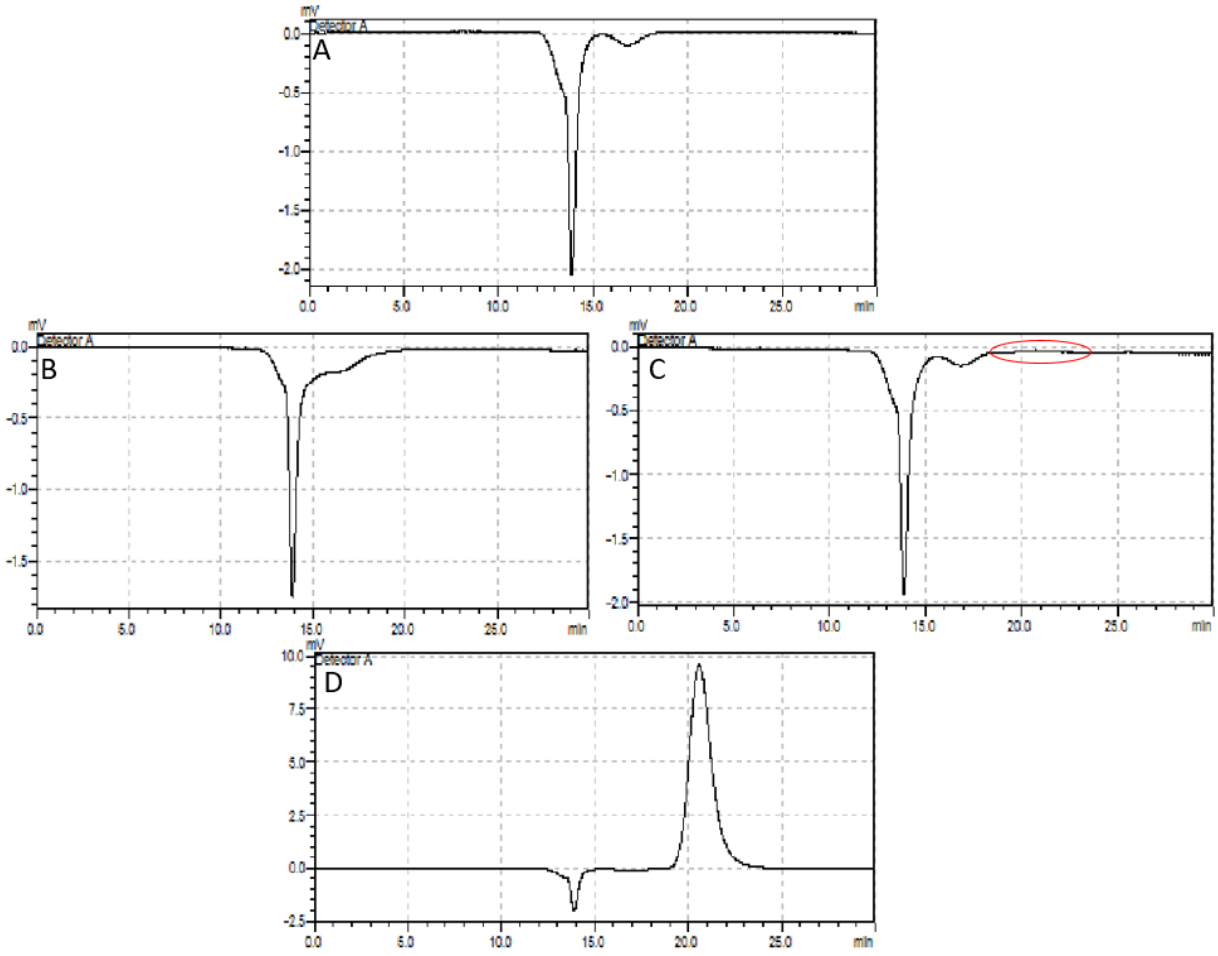


**Figure 3-3: HPLC results for EtOH**

The initial stock for the HPLC is 95 % v/v EtOH, diluted from 100% with water. Further concentrations are diluted with DI Water. **A**: A blank using only filtered DI water. **B**: 9.5E-5 % v/v EtOH. **C**: 9.5E-4 % v/v EtOH. The red circle indicates the location of the peak. **D**: 0.95 wt % EtOH.

# 3.3: Citric Acid

The PCA (Figure 3-4 A) and DAPC (Figure 3-4 B) showed that CA is detectable down to 2.8E-4 wt % based on the data from canonical 1 in DAPC analysis. As before, 50 PCs were used, accounting for 98.74 % of the variability within the PCs. This lack of statistical difference ($p > 0.05$)

was further confirmed by ANOVA analysis and a Dunnett's comparison test, shown in table 3-3. This puts the detection limit for CA at 2.8E-4 wt % when measuring changes in phenotypes by Raman spectroscopy and processing the data with Rametrix™.

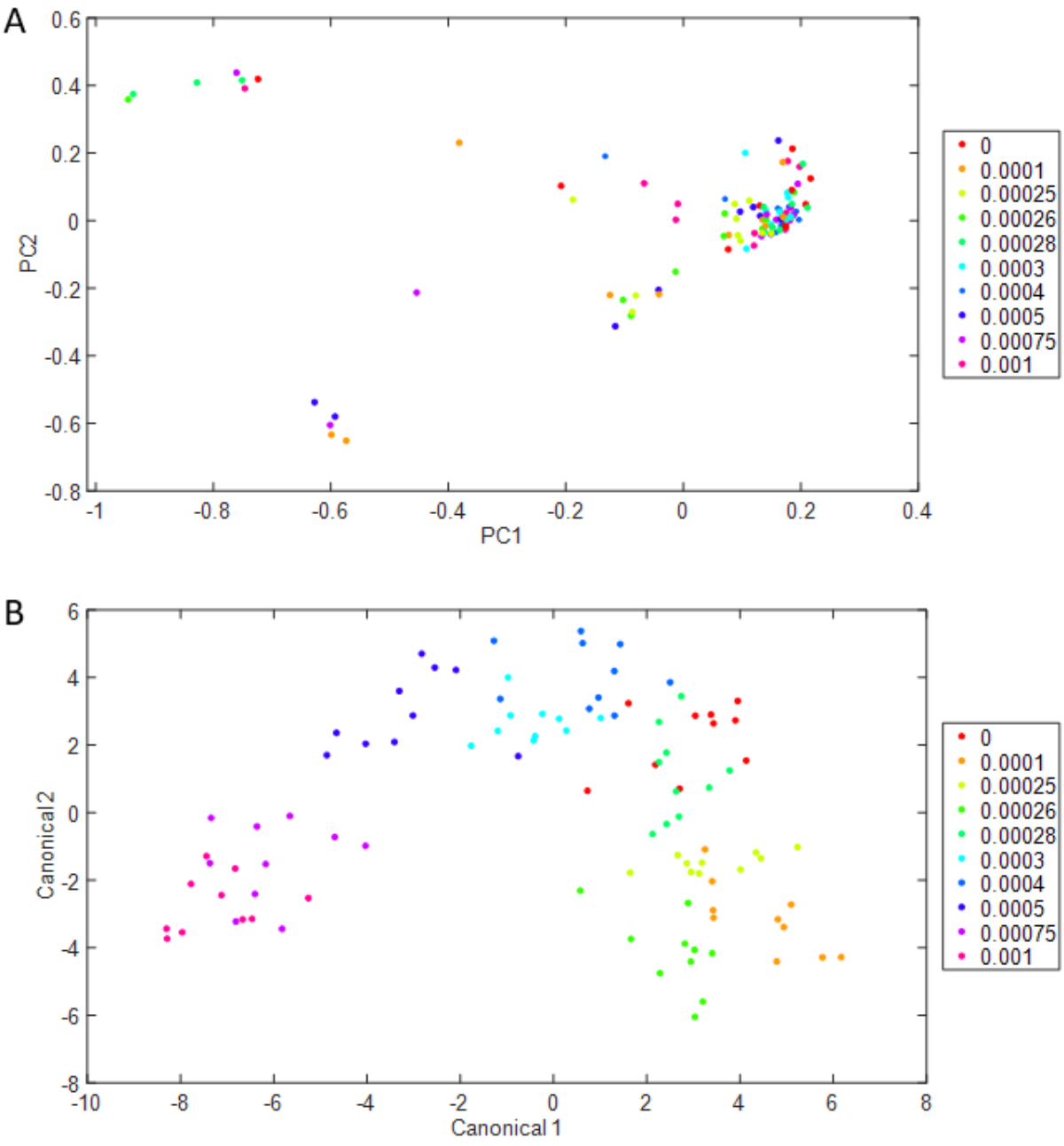


**Figure 3-4: PCA and DAPC of CA to 1.0E-4 wt %**

**A.** PCA of CA. **B.** DAPC of the same data, utilizing 50 PCs to explain 98.74 % of the variability in the data.

**Table 3-3: CA P-Values**

P-values from the Dunnett's test showing similarities in the phenotypes of CA concentrations ranging from 1.0E-4 to 1.0E-3 wt %.

| Concentration (wt %) | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 0 | 1.0E-4 | 2.5E-4 | 2.6E-4 | 2.8E-4 | 3.0E-4 | 4.0E-4 | 5.0E-4 | 7.5E-4 | 1.0E-3 |
| P-Value | N/A | 0.0043 | 0.7885 | 0.9852 | 0.9982 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |

As with EtOH, a direct comparison to what a standard procedure in a standard water quality laboratory would provide for CA detection is desired to show where the Raman method stands. One such test involves a HPLC utilizing a column created for organic acids and solvents. With dilutions using only filtered DI water, the lowest detectable concentration of CA was 5.0E-4 M (9.6E-3 wt %), as is shown in figure 3-5. These results show that detection with Raman scanning is comparable with standard methods for detection compounds like CA.

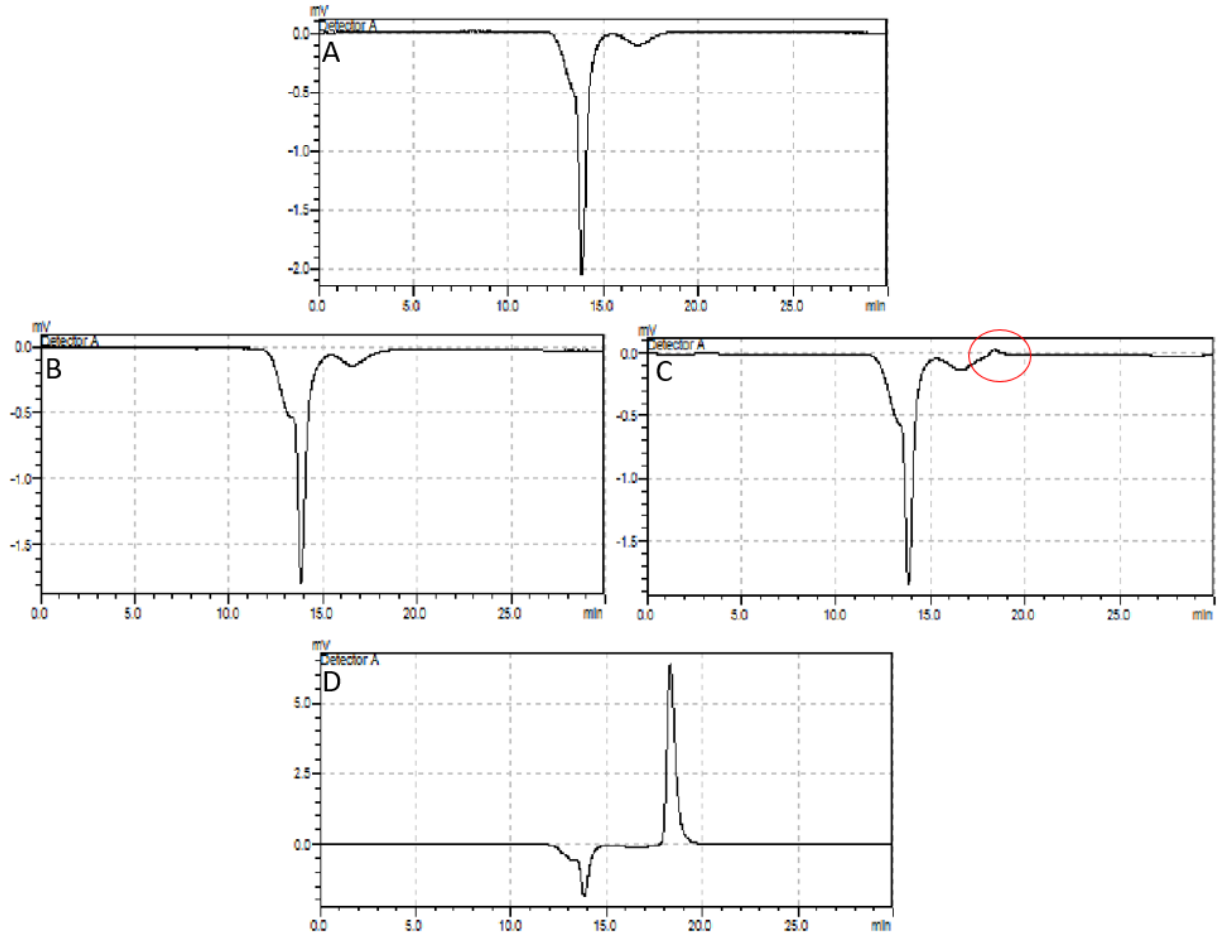


**Figure 3-5: HPLC Runs for CA**

The initial stock for the HPLC is 0.5 M CA. Further concentrations are diluted with filtered DI water. **A**: A blank using only filtered DI water. **B**: 5.0E-5 M CA. **C**: 5.0E-4 M CA. The red circle indicates the location of the peak. **D**: 0.05 M CA.

# 3.4: Acetic Acid

As with CA, AA was measured utilizing PCA (figure 3-6A) and DAPC (figure 3-6B) to find the lower limit of detection. As shown by both charts, the phenotypes of *E. coli* became significantly different at 2.6E-4 wt %. The DAPC utilized 50 PCs, accounting for 98.06 % of

variability within the data. These results were confirmed through ANOVA and a Dunnett's comparison test, which showed no significant difference ($p>0.05$) between the control and 2.6E-4 wt %. This put the detection limit for AA at 2.6E-4 wt % when measuring changes in phenotypes by Raman spectroscopy and processing the data with Rametrix™.

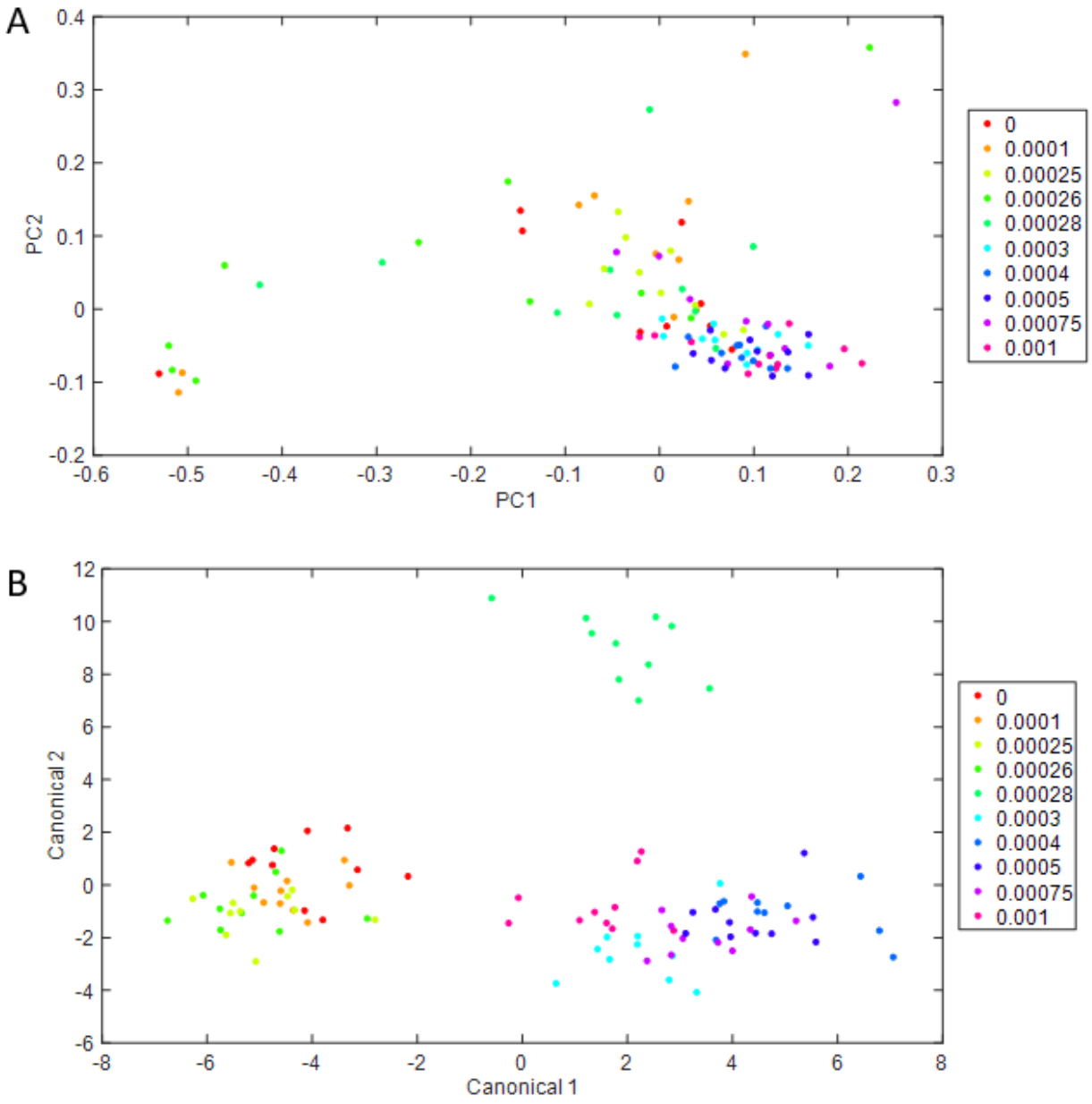


**Figure 3-6: PCA and DAPC of AA to 1.0E-4 wt %**

**A.** PCA of AA at a concentration range from 1.0E-7 to 1.0E-9 wt %. **B.** DAPC of the same AA data, utilizing 50 PCs to account for variability within the system.

**Table 3-4: AA P-Values**

P-values from the Dunnett's test showing similarities in the phenotypes of AA concentrations ranging from 1.0E-4 to 1.0E-3 M.

| Concentration (wt %) | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 0 | 1.0E-4 | 2.5E-4 | 2.6E-4 | 2.8E-4 | 3.0E-4 | 4.0E-4 | 5.0E-4 | 7.5E-4 | 1.0E-3 |
| P-Value | N/A | 0.9556 | 0.2627 | 0.0921 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |

As with CA and EtOH, the Raman method shown was compared to the standard practice of HPLC for detection of organic acids and solvents for AA standards. With dilutions using only filtered DI water, the lowest detectable concentration of AA was 1.0E-4 M (6.0E-4 wt %), as is shown in figure 3-7. A similar number to CA, this is lower than what was detected with Raman spectroscopy. As with CA and EtOH, however, this does show that the methods outlined within are comparable to common laboratory methods.

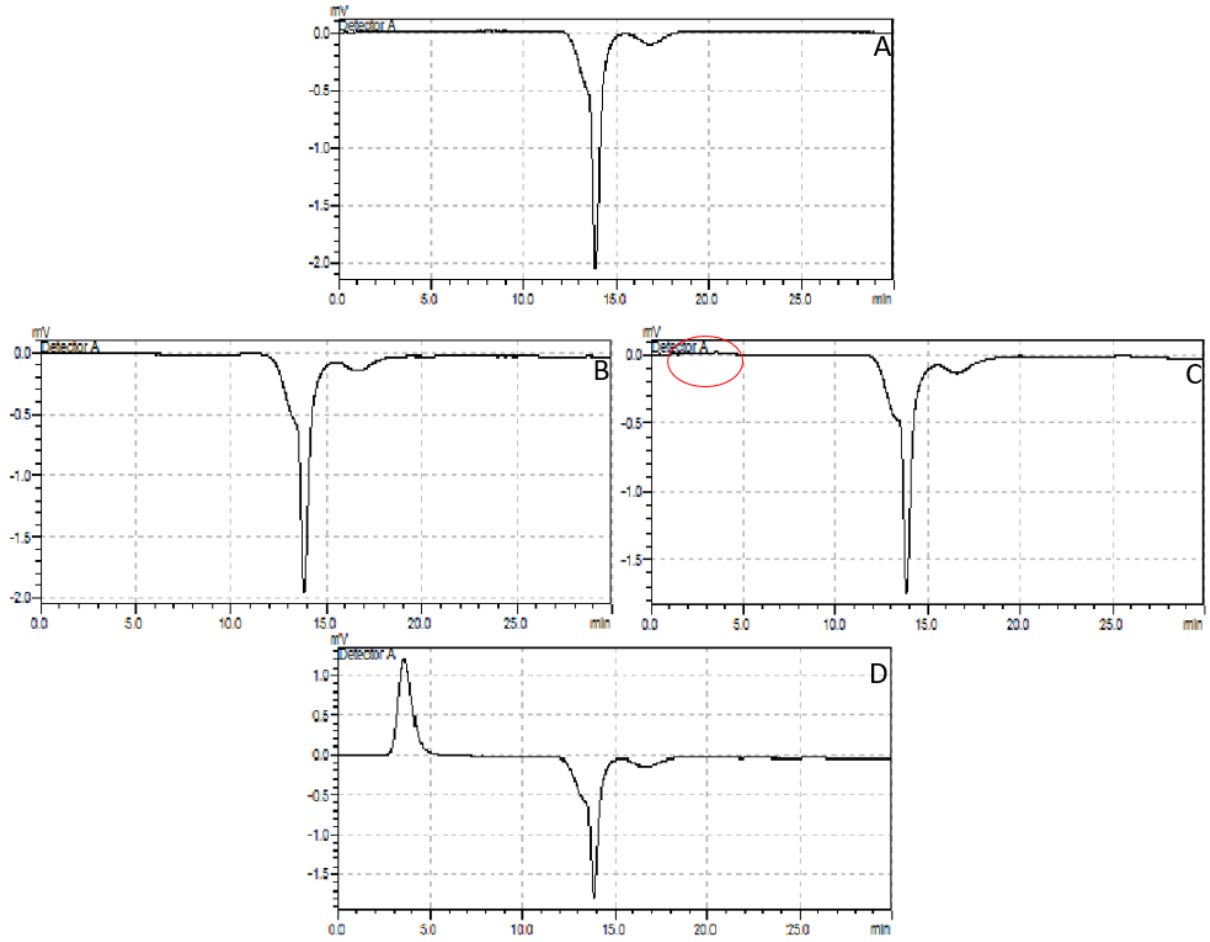


**Figure 3-7: HPLC Runs for AA**

The initial stock for the HPLC is 1 M AA. Further concentrations are diluted with filtered DI water. **A**: A blank using only filtered DI water. **B**: 4E-5 M CA. **C**: 1E-4 M CA. The red circle indicates the location of the peak. **D**: 0.01 M CA.

## 3.5: Kanamycin

From the results shown, PCA (figure 3-8A) and DAPC (figure 3-8B) set the detection limit of kanamycin to be 2.5E-11 wt %. The DAPC utilized 50 PCs to account for 99.09 % of the variability seen in the data. Along with this, ANOVA and a Dunnett's comparison test were applied to the DAPC results for canonical 1, showing that there is no significant difference ($p > 0.05$) was

found between 2.5E-11 wt % and the control. This put the detection limit for kanamycin at 2.5E-11 wt % when measuring changes in *E> coli* phenotypes by Raman spectroscopy and processing the data with Rametrix™.

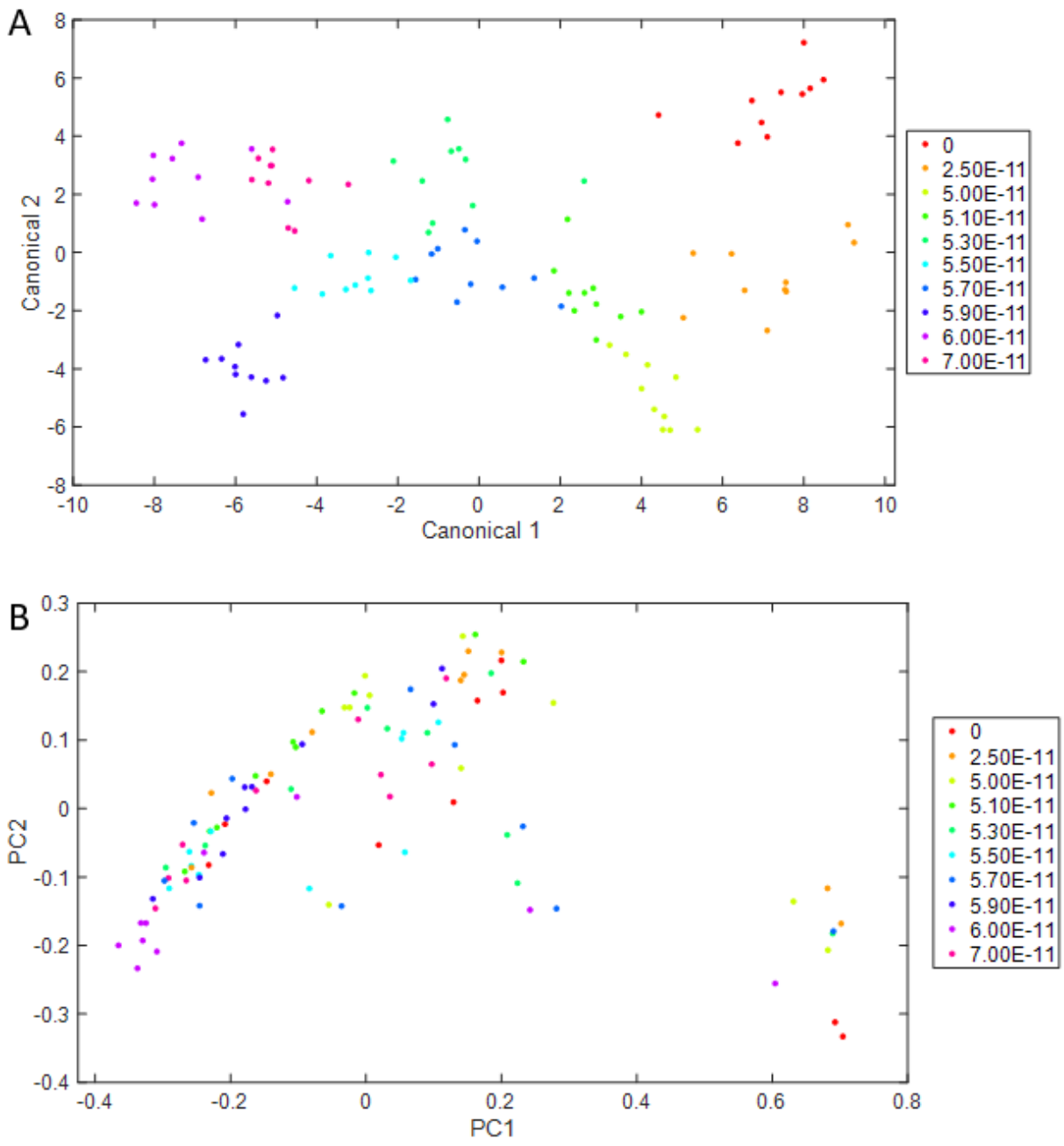


**Figure 3-8: PCA and DAPC of Kanamycin to 2.5E-11 wt %**

**A.** PCA of kanamycin at a concentration range from 1.0E-7 to 1.0E-9 wt %. **B.** DAPC of the same kanamycin data, utilizing 50 PCs to account for variability within the system.

**Table 3-5: Kanamycin P-Values**

P-values from the Dunnett's test showing similarities in the phenotypes of kanamycin concentrations from 2.5E-11 to 7.0E-11 wt %.

| Concentration (wt %) | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 0 | 2.5E-11 | 5.0E-11 | 5.1E-11 | 5.3E-11 | 5.5E-11 | 5.7E-11 | 5.9E-11 | 6.0E-11 | 7.0E-11 |
| P-Value | N/A | 1.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |

# 3.6: Ampicillin

For ampicillin, the PCA (figure 3-8A) and DAPC (figure 3-8B) set the detection to be 2.5E-10 wt %. The DAPC utilized 50 PCs to account for 99.09 % of the variability seen in the data. Along with this, ANOVA and a Dunnett's comparison test were applied to the DAPC results for canonical 1, shown in table 3-6. This confirmed that no significant difference ($p > 0.05$) was found between 2.5E-11 wt % and the control. This put the detection limit for ampicillin at 2.5E-10 wt % when measuring changes in *E. coli* phenotypes by Raman spectroscopy and processing the data with Rametrix™.

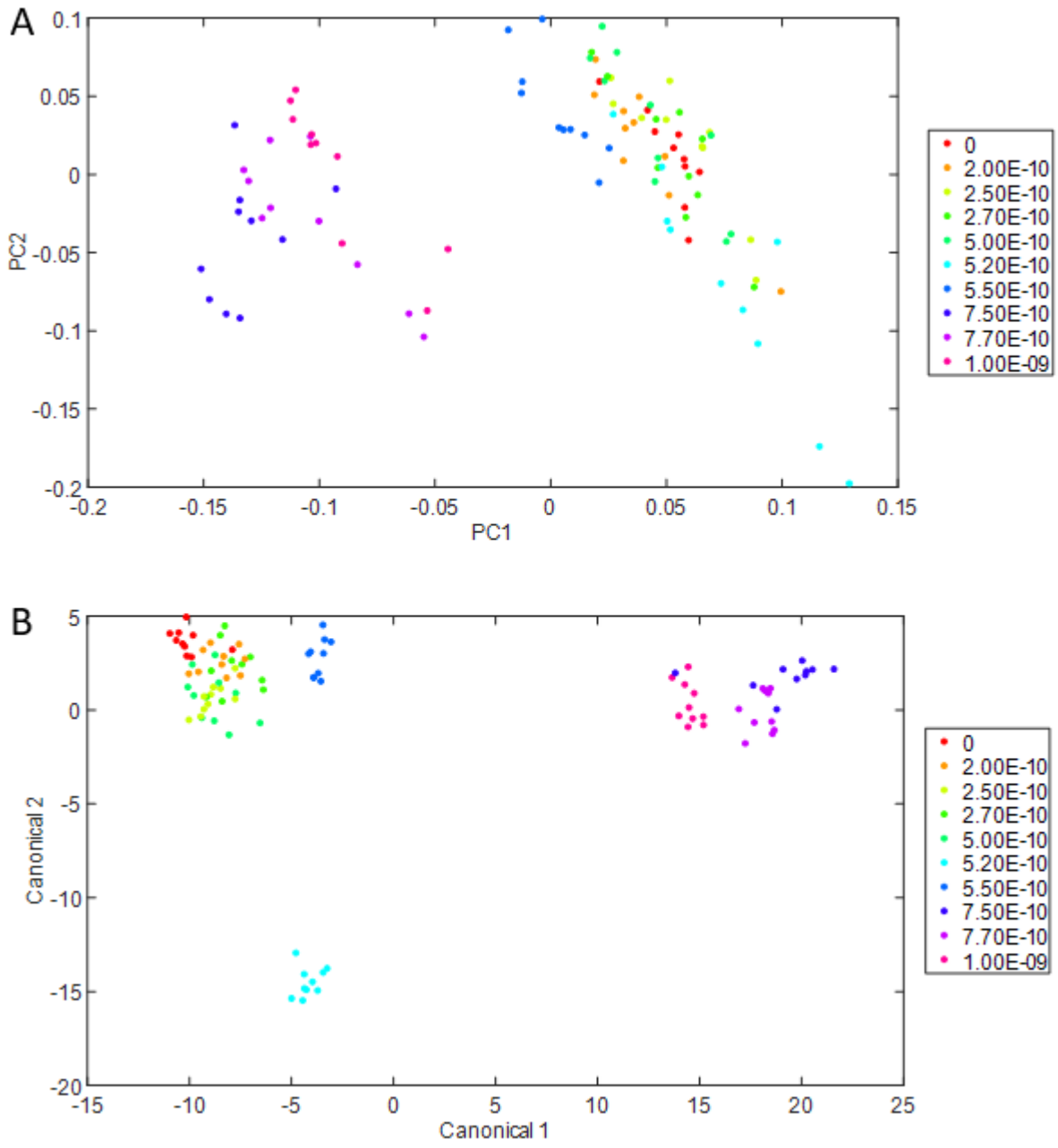


**Figure 3-9: PCA and DAPC of Ampicillin to 2.0E-10 wt %**

**A.** PCA of ampicillin at a concentration range from 1.0E-7 to 1.0E-9 wt %. **B.** DAPC of the same ampicillin data, utilizing 50 PCs to account for variability within the system.

**Table 3-6: Ampicillin P-Values**

P-values from the Dunnett's test showing similarities in the phenotypes of ampicillin concentrations from 2.0E-10 to 1.0E-9 wt %.

| Concentration (wt %) | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 0 | 2.0E-10 | 2.5E-10 | 2.7E-10 | 5.0E-10 | 5.2E-10 | 5.5E-10 | 7.5E-10 | 7.7E-10 | 1.0E-9 |
| P-Value | N/A | 0.0066 | 0.0693 | 0.0000 | 0.0311 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |

# 3.7: $CoCl_2$

$CoCl_2$ turned out to be surprising with how low of a concentration was detectable through changes in *E. coli* phenotypes. Experiments showed that phenotype changes for concentrations as low as 5.0E-12 could be detected, as shown in figure 3-10B, though the PCA analysis in figure 3-10A implied that the lower concentrations are similar, if not the same, as the control. As with previous components, 50 PCs were utilized in order to account for 96.75 % of variability within the PCs, in this case. The differences ($p < 0.05$) were confirmed via ANOVA and a Dunnett's comparison test, shown in table 3-7 below. While further experiments could be performed to find the lower limit of $CoCl_2$ detection, the amount of $CoCl_2$ was within the threshold of a trace element, which is considered to be a normal occurrence and even beneficial for environmental organisms[37,91]. This makes detection unnecessary at the low concentrations found. Of course, toxic levels will be easily detectable by this methodology.

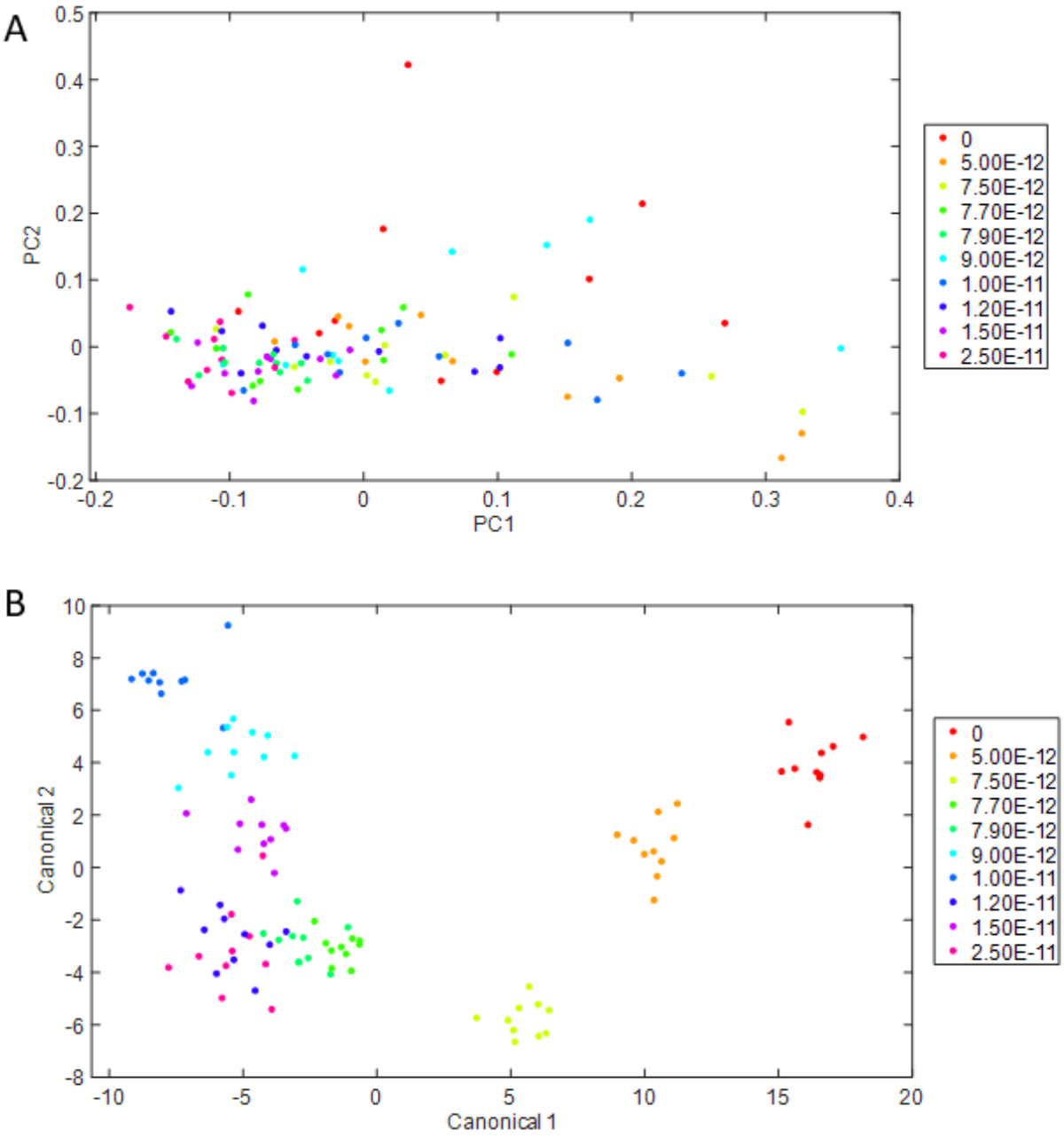


**Figure 3-10: PCA and DAPC of $CoCl_2$ to 5.0E-12 wt %**

**A.** PCA of $CoCl_2$ at a concentration range from 5.0E-12 to 2.5E-11 wt %. **B.** DAPC of the same $CoCl_2$ data, utilizing 50 PCs to account for variability within the system.

**Table 3-7: $CoCl_2$ P-Values**

P-values from the Dunnett's test showing similarities in the phenotypes of $CoCl_2$ concentrations from 5.0E-12 to 2.5E-11 wt %.

| Concentration (wt %) | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 0 | 5.0E-12 | 7.5E-12 | 2.7E-12 | 7.9E-12 | 9.0E-12 | 1.0E-11 | 2.2E-11 | 1.5E-11 | 2.5E-11 |
| P-Value | N/A | 0.27E-5 | 0.27E-5 | 0.27E-5 | 0.27E-5 | 0.27E-5 | 0.27E-5 | 0.27E-5 | 0.27E-5 | 0.27E-5 |

# 3.8: Silver Nanoparticles

As emerging contaminants, AgNP were an interesting component to test. The PCA (figure 3-11A) and DAPC (3-11B) show that AgNP affected the phenotype of *E. coli* to a concentration of 5.2E-7 wt %. The DAPC utilized 50 PCs to account for 99.12 % of the variability within the data. These results were confirmed through ANOVA and a Dunnett's comparison test (table 3-8), showing that there was no significant difference ($p > 0.05$) between the control and 5.2E-7 wt % along canonical 1. This put the detection limit for AgNP at 5.2E-7 wt % when measuring changes in *E. coli* phenotypes by Raman spectroscopy and processing the data with Rametrix™.

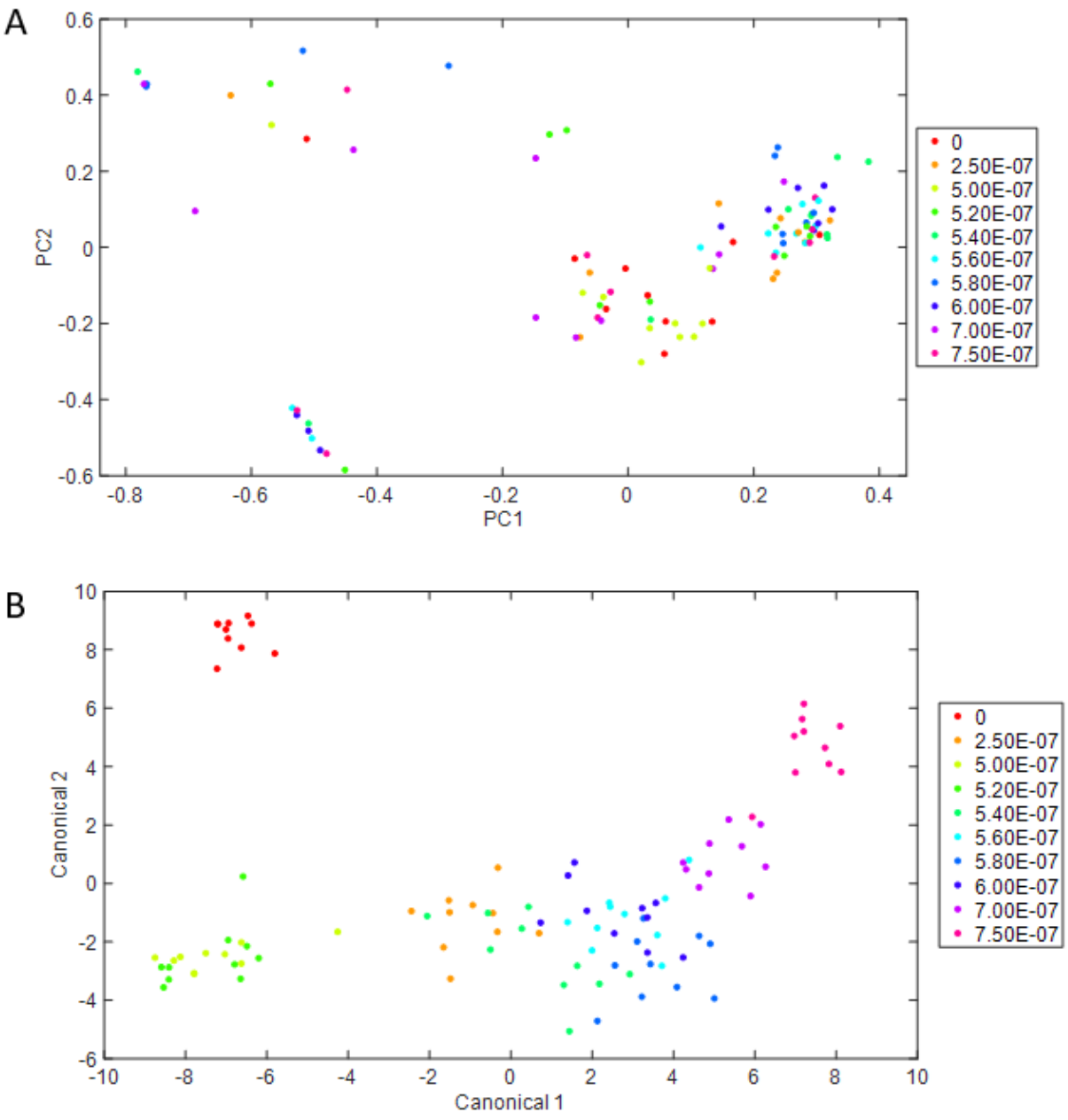


**Figure 3-11: PCA and DAPC of AgNP to 2.5E-7 wt %**

**A.** PCA of AgNP at a concentration range from 2.5E-7 to 7.5E-7 wt %. **B.** DAPC of the same AgNP data, utilizing 50 PCs to account for variability within the system.

**Table 3-8: AgNP P-Values**

P-values from the Dunnett's test showing similarities in the phenotypes of AgNP concentrations from 2.5E-7 to 2.5E-5 wt %.

| Concentration (wt %) | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 0 | 2.5E-7 | 5.0E-7 | 5.2E-7 | 5.4E-7 | 5.6E-7 | 5.8E-7 | 6.0E-7 | 7.0E-7 | 7.5E-7 |
| P-Value | N/A | 0.0000 | 0.8484 | 0.7254 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |

# 3.9: Multicomponent Test- Kanamycin, $CoCl_2$, Acetic Acid

## 3.9.1: $CoCl_2$, Kanamycin, Acetic Acid

For this multicomponent test, all concentrations were mixed using the maximum value of AA for added components (3.4 mL LB and a mixture of AA, $CoCl_2$, and kanamycin solutions as well as water to make 0.6 mL). It was initially run against the initial control for AA; however, the results were better shown if no control was used and only plotted with the data for the single-component tests. Both the PCA (figure 3-12A) and DAPC (figure 3-12B) showed that the mixed value was significantly different from all three single-component samples. The DAPC utilized 20 PCs to account for 99.19 % of variability in the data. This value was chosen due to 50 PCs being half of the total number of samples in previous experiments. These differences were confirmed with ANOVA and a Dunnett's comparison test over canonical 1 (Table 3-9), showing that all single-component samples were significantly different ($p < 0.05$) from the multicomponent test. It was noted that along all canonicals that the mixture behaved most similarly to $CoCl_2$, and further testing was needed to confirm this.

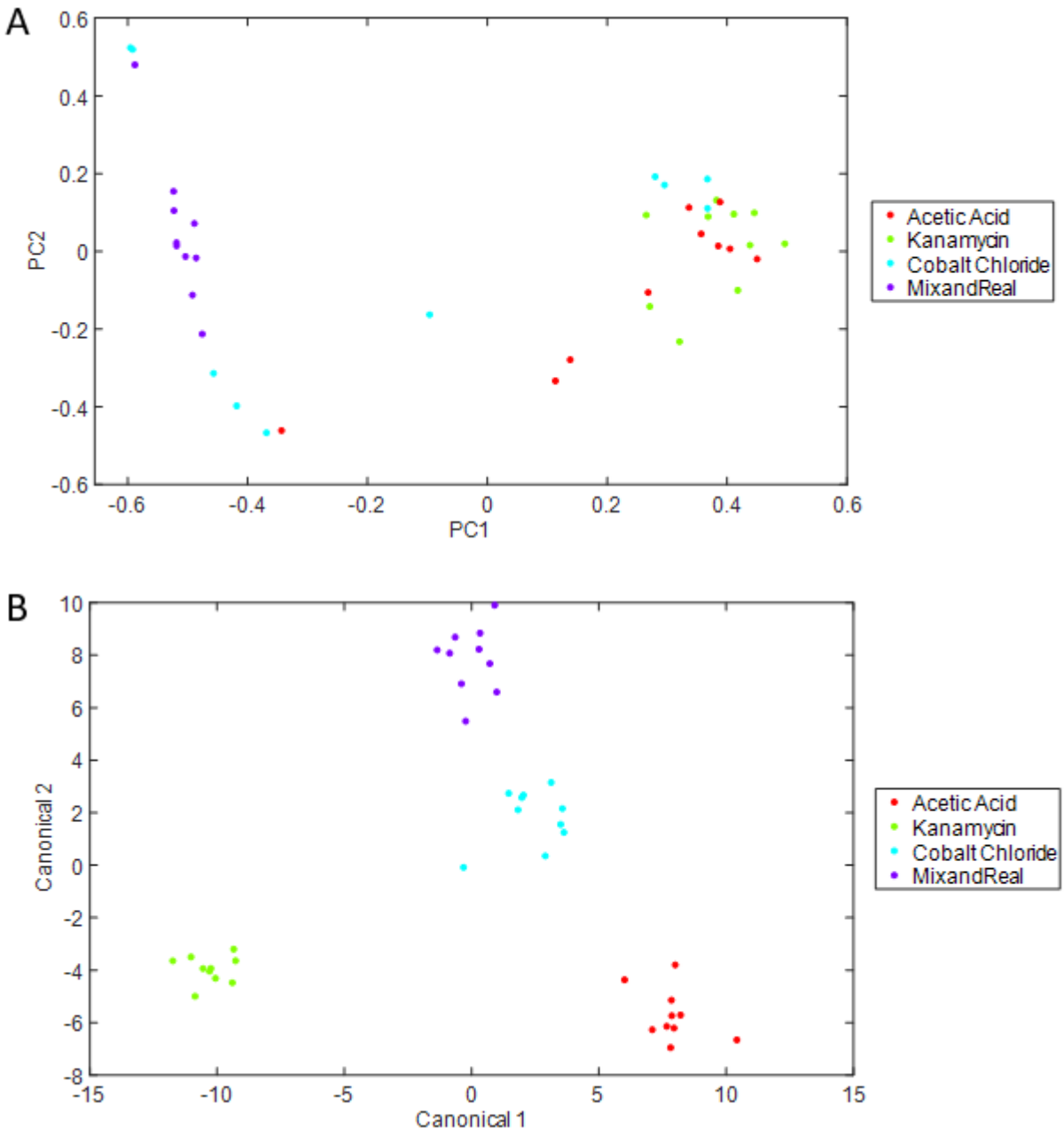


**Figure 3-12: PCA and DAPC of Multicomponent Test**

**A.** PCA of the tertiary mixture of AA, kanamycin, and $CoCl_2$. **B.** DAPC of the same tertiary mixture data, with 20 PCs utilized to account for variability within the system.

**Table 3-9: Multicomponent P-Values**

P-values from the Dunnett's test showing similarities in the phenotypes of the multicomponent test and corresponding single test concentrations.

| Chemical | P-Value |
|---|---|
| Multicomponent | N/A |
| AA | 0.2613E-5 |
| Kanamycin | 0.2613E-5 |
| $CoCl_2$ | 0.2613E-5 |

To confirm the suspicions that $CoCl_2$ was having the most effect on the phenotype, the average canonical values and the TCD, as shown in equation 3-1, with x representing canonical 1, y representing canonical 2, and z representing canonical 3[79].

**Equation 3-1: Distance Equation**

$$TCD = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2}$$

The TCD between each component and the mixture was found (shown below in Table 3-10), and the average values along each canonical show that $CoCl_2$ is the closest to the mixture, as the TCD value is the shortest of the three single components. Kanamycin seems to have the next-highest effect on the phenotype, with AA having the least effect, based on TCD values.

**Table 3-10: Multicomponent TCD**

The average canonical values of each test along the canonicals and the TCD between each component in AU.

| Component | Canonical 1 (AU) | Canonical 2 (AU) | Canonical 3 (AU) | TCD (AU) |
|---|---|---|---|---|
| AA | 7.90 | -5.71 | -1.72 | 15.72 |
| Kanamycin | -10.27 | -3.98 | -0.29 | 11.01 |
| $CoCl_2$ | 2.38 | 1.84 | -4.44 | 5.36 |
| Mixture | -0.014 | 7.85 | -2.43 | N/A |

## 3.9.2: $CoCl_2$, $ZnCl_2$, $MgCl_2$

To further strengthen the results from the previous multicomponent test, two mixtures of $CoCl_2$, $ZnCl_2$, and $MgCl_2$ were created. The first mixture kept all three components at 5.0E-6 wt %, and the second had 5E-5 wt % $CoCl_2$ and 5E-6 wt % $ZnCl_2$ and $MgCl_2$. The basis for the sample volumes was based on $CoCl_2$'s maximum value (3.98 mL LB and 0.02 mL of contaminants and water) to keep the amount of nutriment consistent. As before, no water control was used for these samples in order to clearly see the relationships between each group.

### 3.9.2.1: Consistent Concentrations

For the test with consistent concentrations of contaminants, it can be seen from the PCA (figure 13-13A) and DAPC (figure 13-13B) utilizing 20 PCs to account for 98.84 % of the variability that the mixture can be differentiated from each individual component. However,

when run through a Dunnett's test (table 3-11), $ZnCl_2$ was statistically the same ($p > 0.05$) as the mixture along canonical 1. It was also noted that it could be difficult to find which single group the mixture behaved the most like, though $MgCl_2$ seemed closest.

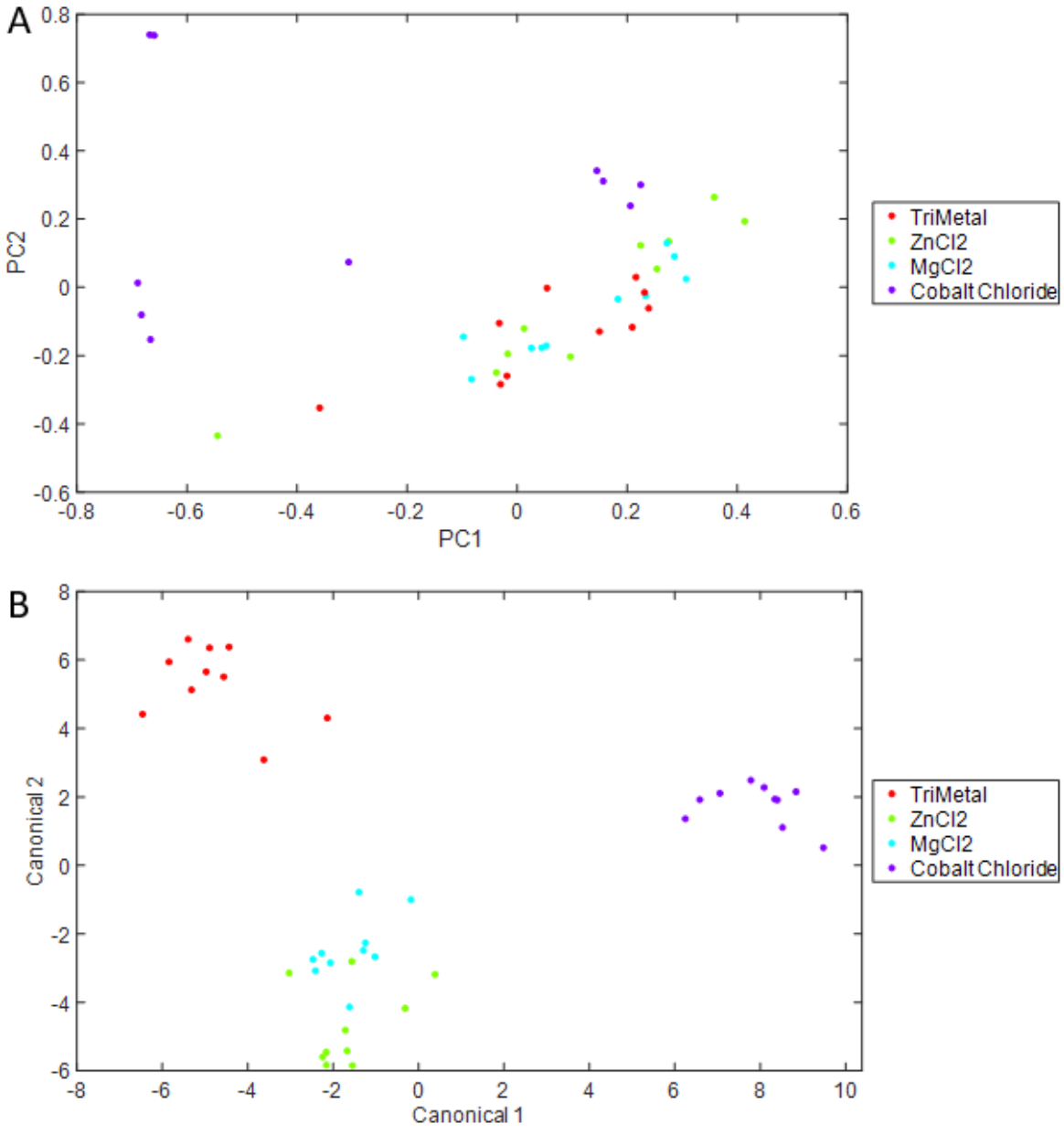


**Figure 3-13: PCA and DAPC of Metal Mixture Test**

**A.** PCA of the tertiary mixture of $ZnCl_2$, $MgCl_2$, and $CoCl_2$. **B.** DAPC of the same tertiary mixture data, with 20 PCs utilized to account for variability within the system.

**Table 3-11: Metal Mixture P-Values**

P-values from the Dunnett's test showing similarities in the phenotypes of the multicomponent metal test and corresponding single test concentrations.

| Chemical | P-Value |
|---|---|
| Multicomponent | N/A |
| $ZnCl_2$ | 1.0000 |
| $MgCl_2$ | 0.0000 |
| $CoCl_2$ | 0.0000 |

As with the previous mixture, the similarity of groups and their effects were found through TCD. The distances shown in table 3-12 implied that $MgCl_2$ is closest, as could be seen in the DAPC, followed by $ZnCl_2$ and then very closely by $CoCl_2$. This was interesting, as this shows that $CoCl_2$ and $ZnCl_2$ are very similar in their overall phenotypic effect, though their clusters are appearing in separate locations on the plots.

**Table 3-12: Metal Mixture TCD**

The average canonical values of each test along the canonicals and the TCD between each component in AU.

| Component | Canonical 1 (AU) | Canonical 2 (AU) | Canonical 3 (AU) | TCD (AU) |
|---|---|---|---|---|
| $CoCl_2$ | 7.94 | 1.77 | -0.19 | 10.97 |
| $ZnCl_2$ | -1.59 | -4.63 | -2.45 | 10.62 |
| $MgCl_2$ | -1.59 | -2.46 | 3.21 | 6.06 |
| Mixture | -4.76 | 5.33 | -0.58 | N/A |

### 3.9.2.2 Higher $CoCl_2$ Concentrations

After the interesting results when $ZnCl_2$, $MgCl_2$, and $CoCl_2$ have consistent concentrations, an experiment was conducted where $CoCl_2$ was run with a higher concentration of 5E-5 wt %. As shown in the PCA (figure 3-14A) and DAPC (3-14B) with 20 PCs accounting for 98.59 % of the variability, the mixture is separate from each individual component. Our Dunnett's test (table 3-13) backed this up, showing that all groups are significantly different ($p < 0.05$). As before, all three canonicals seemed to show that $MgCl_2$ is the most similar to the group.

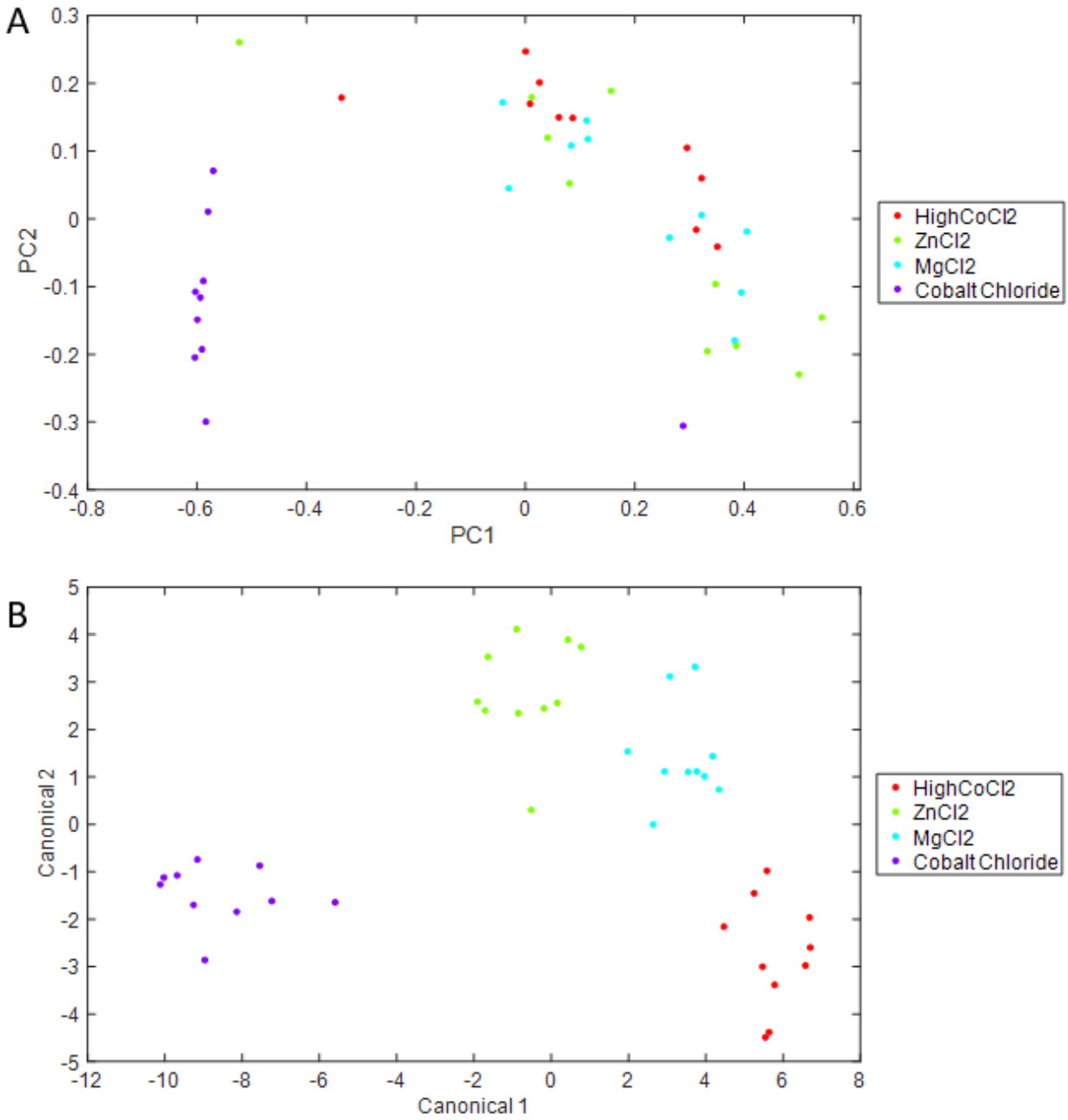


**Figure 3-14: PCA and DAPC of High-$CoCl_2$ Test**

**A.** PCA of the tertiary mixture of $ZnCl_2$, $MgCl_2$, and higher amounts of $CoCl_2$. **B.** DAPC of the same tertiary mixture data, with 20 PCs utilized to account for variability within the system.

**Table 3-13: High-$CoCl_2$ Mix P-Values**

P-values from the Dunnett's test showing similarities in the phenotypes of the multicomponent test and corresponding single test concentrations.

| Chemical | P-Value |
|---|---|
| Multicomponent | N/A |
| $ZnCl_2$ | 0.2614E-5 |
| $MgCl_2$ | 0.2613E-5 |
| $CoCl_2$ | 0.2613E-5 |

TCD for each single component was once again found, with $MgCl_2$ being closest to the mixture, shown in table 3-14). Like before, $ZnCl_2$ was the next closest, followed by $CoCl_2$ showing the lowest effect on the phenotype changes. Unlike before, the differences between $CoCl_2$ and $ZnCl_2$ can be clearly seen through both distance and the plots, rather than only the plots.

**Table 3-14: High-$CoCl_2$ Mix TCD**

The average canonical values of each test along the canonicals and the TCD between each component in AU.

| Component | Canonical 1 (AU) | Canonical 2 (AU) | Canonical 3 (AU) | TCD (AU) |
|---|---|---|---|---|
| $CoCl_2$ | -8.57 | -1.48 | -0.54 | 12.54 |
| $ZnCl_2$ | -0.63 | 2.78 | 2.23 | 8.54 |
| $MgCl_2$ | 3.42 | 1.44 | -2.78 | 6.58 |
| Mixture | -4.76 | 5.33 | -0.58 | N/A |

# 3.10: Impure Water Test

As a proof-of-concept of the distinguishing power of this Raman method with Rametrix™, tests with media made from non-DI water were completed. These tests included LB media made with water from Virginia Tech's Duck Pond, Stroubles Creek, and water from a standard tap in the laboratory. The control for this experiment was LB made with DI water. Tests were otherwise run in the same manner as the single-component and multicomponent analyses. For data analysis, the spectra were truncated from 400-1800 $cm^{-1}$ with only baseline corrections and no vector normalization. This was because concentrations of cellular components as well as components within the water should be taken into account to show the results.

As seen in figure 3-13, the PCA (3-13A) and DAPC (3-13B) showed that all three samples were different from the control, as expected. 20 PCs were utilized for the same reason as the multicomponent test, which accounted for 99.84 % of variability in the data. This was confirmed

with ANOVA and a Dunnett's comparison test along canonical 1 (Table 3-11), showing that there was significant difference ($p < 0.05$) between all samples, with tap water appearing closest to the control along both canonical 1 and canonical 2.

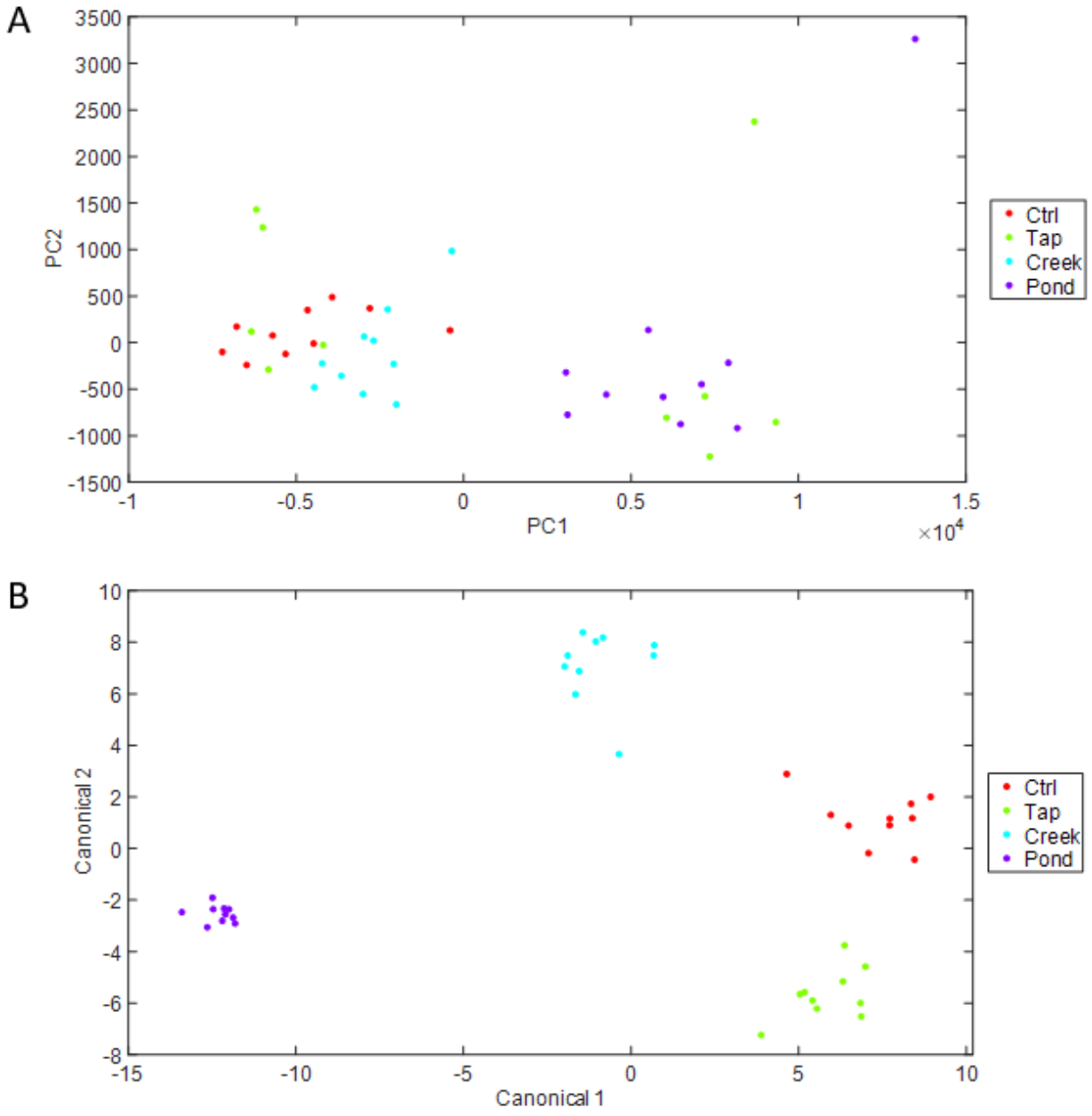


**Figure 3-15: Impure Water PCA and DAPC**

**A.** PCA of the water samples taken from a stagnant, running, and tap source. **B**. DAPC of that same data, incorporating 50 PCs to account for variability.

**Table 3-15: Impure Water P-Values**

P-values from the Dunnett's test showing similarities in the phenotypes of the impure water test and and the DI control.

| Chemical | P-value |
|---|---|
| Control | N/A |
| Tap | 0.0044 |
| River | 0.0000 |
| Pond | 0.0000 |

As with the multicomponent test, the TCD covering all three canonicals was found utilizing equation 3-1. As expected from the DAPC data and the Dunnett's comparison test, the closest group to the control is tap water, followed by the creek water and finally the pond, with the data shown in table 3-12.

**Table 3-16: Impure Water TCD**

The average canonical values of each concentration along the canonicals and the total distances between each component in AU.

| Component | Canonical 1 (AU) | Canonical 2 (AU) | Canonical 3 (AU) | TCD (AU) |
|---|---|---|---|---|
| Control | 7.38 | 1.13 | 2.53 | N/A |
| Tap | 5.85 | -5.67 | -1.78 | 8.19 |
| Creek | -0.93 | 7.09 | -1.54 | 14.46 |
| Pond | -12.30 | -2.55 | 0.79 | 15.09 |

# Chapter 4: Discussion

Raman technology, specifically SERS, is seeing increasing use in multiple fields of science; from chemistry to environmental engineering. The method's advantages include its portability, minimal sample preparation, and its low sensitivity to the spectrum of water, among others. Its uses include, but are not limited to; determining the chemical compositions of samples, as a near-real time screening technique, monitoring option for contaminants in water, a biosensor, and a technology to quickly and easily find new phenotypes of cells for specialized purposes.

While Raman has been used to monitor for single components in water samples as well as monitoring cell phenotypes, the combination of these two methods to look at all contaminants within water systems has not been well-studied. This research looked at the viability of a simple SERS method to test water for contaminants. The tests involved finding the limits of single components in water, a tertiary test involving $CoCl_2$, AA, and kanamycin to see how it compares to the respective single tests, and tests of water from a relatively stagnant source, a running source, and from a standard tap to show their behavior. From these results, the viability of this method and the directions in which it should go was determined.

## 4.1: Single Component Tests

The single-component tests showed that using SERS with the phenotypes of *E. coli* was able to detect the presence of contaminants from multiple different classes down to levels comparable to or below standard tests. Differences in *E. coli* phenotypes from ionic strength were

able to be detected to within the realm of $10^{-9}$ wt % in NaCl. A potential reason for this is that *E. coli* is has low halotolerance[91], and small amounts of extra salt could affect the phenotype and cause changes to combat toxicity. $CoCl_2$ was able to be detected to a concentration in the realm of $10^{-12}$ wt % with reliability, though a "hard" limit was unable to be found. $CoCl_2$ itself is toxic to bacteria and humans in fairly low concentrations, but trace amounts of it are required in many lifeforms to create essential nutrients and biological components such as the cofactor vitamin B12. With this knowledge, it can be assumed that low concentrations altered the phenotype to allow for higher production of cobalt-related products, while higher concentrations induced changes to adapt to a toxic, but sublethal environment.

Organic acids like CA and AA and solvents such as EtOH tend to have a higher limit than the others, with both acids only showing differences in the realm of $10^{-4}$ wt % and EtOH going to the realm of $10^{-7}$ wt %. For the acids, CA is an essential part of almost all organisms, being an intermediate within the Krebs cycle while AA is able to be used as a carbon source for bacteria in low concentrations. This implies that once the concentrations of CA and AA stop having a noticeable effect on the pH of the media, they become feedstock for the cells similar to the glucose and yeast extract within the LB, giving the cells no reason to adapt and their phenotype does not change. While bacteria can adapt and gain tolerance to EtOH, the EtOH most likely is having negative effects down to that concentration, and at $10^{-7}$ wt % evaporation may be changing it into the gaseous state and causing a minimal effect on bacterial phenotypes.

Changes in phenotype due to ampicillin and kanamycin were detectable to relatively low concentrations, with detection limits in the realms of $10^{-10}$ and $10^{-11}$ wt %, respectably. In both cases, it can be assumed that the cells were adapting to the increased antibiotic content and, in

the case of kanamycin, resistance may be conveyed by the increased but sub-lethal dose, causing a change in the phenotype based on new cellular machinery. Ampicillin being detectable to such low concentrations was surprising, since the *E. coli* utilized incorporated a pUC19 plasmid to convey resistance to it. A potential reason is that the concentration was far enough over the resistance "limit" to add additional stress to the cell or create a reason to gain additional resistance mechanisms.

The final component tested, AgNP, also had a somewhat low detection limit that fell into the realm of $10^{-7}$ wt %. This was not surprising given silver's antibacterial properties, as well as how nanoparticles have the ability to permeate membranes and potentially spread the silver further than the ionic version. Tests with other nanoparticles will need to be completed in order to glean more insight into this relatively new environmental contaminant.

# 4.2: Multicomponent Test

## 4.2.1: $CoCl_2$, Kanamycin, Acetic Acid

For the first multicomponent test, a mixture of kanamycin, AA, and $CoCl_2$ was added in concentrations that were around the center of each range tested. The DAPC (figure 3-12 B) shows that along canonical 1 and canonical 2, which show the most change and have the most effect on the outcome, that it is close to being an even spread of all three effects. Slight differences in location and behavior may be due to synergistic effects of each contaminant working at once, or some cellular machinery being used for multiple purposes. What can be seen is that along those canonicals, it appears that $CoCl_2$ affects the phenotype the most, causing it to behave in a manner

more suited to an environment rich in it. This was confirmed by looking at the TCD between the mixture of contaminants and their single counterparts. Since $CoCl_2$ could be detected to a lower concentration than any of the other contaminants, the data begins to make sense. At this concentration, its toxicity mechanisms are holding the most effect, followed by kanamycin and finally AA. A potential insight from this trial is that when looking at a mixed or real system, it is helpful to show the single components in order to get an idea of the contaminant which has the most effect, signifying what may be the largest contributor. More insight into the behavior can be gained through further tests on mixed components, with this specific combination and a combination of the other contaminants.

### 4.2.2: $CoCl_2$, $ZnCl_2$, $MgCl_2$

To further characterize behavior from multiple components and differentiate between their phenotype changes, tests were performed on similar salts: $CoCl_2$, $MgCl_2$, and $ZnCl_2$. This was to prove that the new Raman scanning methods could be utilized for similar components and can differentiate between similar concentrations and different concentrations.

#### 4.2.2.1: Consistent Concentrations

While the differences were able to be seen in PCA and DAPC (Figure 3-13), it was noted that the Dunnett's test shows that $ZnCl_2$ appears statistically similar to the mixture. This could be due to $ZnCl_2$ being more toxic at this concentration, or the phenotype being similar to the mixture due to smaller changes induced by the other two contaminants, as well as the fact that only the differences along canonical 1 are being observed. Along with this, TCD showed that

while the DAPC shows clear differences in the locations of each phenotype cluster, the overall distances of $ZnCl_2$ and $CoCl_2$ were extremely similar, at 10.62 and 10.97 AU, respectively. This could be due to distance differences adding up to be similar, which could cause a false positive when looking for results.

#### 4.2.2.2: Higher $CoCl_2$ Concentrations

The second of these metallic tests involved a higher concentration of $CoCl_2$ while keeping the concentrations to $ZnCl_2$ and $MgCl_2$ consistent. All of the tests run showed that each group was significantly different from each other. It was also seen from the TCD that $CoCl_2$ and $ZnCl_2$ were distinct from each other, most likely due to the increased concentration of $CoCl_2$. It is interesting that $CoCl_2$ appears to have the least effect on the overall phenotype. This could potentially be due to it being near the lower bounds of a toxic effect or minimal beneficial effect compared to the effects of the other two contaminants. More insight into the behavior can be gained through further tests on single components and other mixtures of similar chemicals and proteins, be it in structure or their toxic effects on the cell.

## 4.3: Impure Water Test

The impure water test showed some interesting, but expected results. From what can be gleaned, all three impure samples were fairly different from the DI water, due to the higher concentration of dissolved solids and other pollutants which were not filtered out or evaporated when autoclaved. Along canonical 1, it was observed that tap water is the closest to the control, as it has been processed and made potable before being sent to the private and public sectors.

Unlike the deionization process, processing tap water does not remove all dissolved solids, but focuses on sterilizing it and removing any toxic components that are present in the source used.

The next closest was the creek water. While the its placement along canonical 1 and the TCD (-0.93 and 14.46 AU, respectively) were much further than with tap water, it showed more similarity to it than the pond water did. Factors which play into this may involve the stream is running and there was little rainfall in the days prior to sampling, so the sample was most likely more dilute than more stagnant sources, making it "purer" than a source such as the Duck Pond. This minimized the number of extraneous metabolites and dissolved solids that may make their way into the samples taken.

The final and furthest component from the DI control was the relatively stagnant water from the Duck Pond. Unlike the creek water, any particulates in the water will either be suspended or sink to the bottom, increasing the overall concentration and "pollution" of the source. Along with this, the lack of movement leaves the top layers of the pond oxygen-rich, while the lower layers tend to be deficient in oxygen. This would lead to the buildup of metabolites from anaerobes near the bottom, which will dissolve and diffuse throughout the pond, with a general concentration buildup over time. The presence of these metabolites will most likely affect the phenotype of *E. coli* compared to the lack of these metabolites from the creek water. Future work could involve finding the types of contaminants associated with these "real" water samples and determining which ones are in the highest concentration based on where the phenotype shows, as well as determining if combinations of unrelated contaminants may have synergistic effects and amplify a certain type of phenotype.

# Chapter 5

## 5.1: Conclusions

SERS is a technology that has a multitude of uses in multiple fields, biological and otherwise. The methods outlined show that a SERS method involving changes in *E. coli* phenotypes due to environmental contaminants is able to reliably detect said contaminants in single-component tests to levels comparable to portable and lab-based standard tests, and it has been shown that it can detect differences within tertiary system that falls within the boundaries of the three corresponding single-component tests, and shows that while all metabolites have an effect on the phenotype displayed, one contaminant can pull the phenotype towards it due to being more toxic or having a higher concentration than the other ones, allowing for researchers to pinpoint the problem more easily, though contaminants which are similar may be more difficult to differentiate with the current method. In impure water samples, it was shown that the phenotypes change and are displayed according to their relative purity, with tap water being the closest to DI while relatively stagnant pond water is the furthest away. Though outside of the scope of this study, a long-term goal of this study is to compare real samples to libraries of single-component and multicomponent phenotypes to learn what pollutants are in water samples and determine the best courses of action to remedy the problems.

## 5.2: Future Directions

In the future, tests with more single and mixed components will be useful to further catalog effects on cell phenotype and provide references for any tests of water in the field or tests of unknown aqueous solutions within a laboratory setting. These references will help to narrow down what is within the water and if the contaminants and/or their concentrations should be of concern to the ones near the water source. Along with this, determination of possible synergistic effects unrelated components have on phenotype may broaden the range and lower the chance of false positives when looking at real water samples.

While the method for inoculation seems to work fairly well, there is plenty of room for improvement in that as well. Experiments with inoculating all tubes at once or with a smaller number of samples could be a good starting point. Other potential options may include keeping everything at 37 $^{o}$C to prevent cooling and aeration to make sure the cells get a constant amount of oxygen.

Along with other goals listed, development of a functional and portable field test is required. As stated prior and shown in figure 2-3, the Raman spectrometer is a small device; with a weight of under 8 pounds. For the system to truly be portable, the ability to make the spectrometer battery-powered would be ideal, though the ability to plug into a generator would serve as a good alternative. The cells could potentially be kept frozen in glycerol stock and brought to the field when testing is ready, and changes in phenotype could be measured in real-time as growth occurs. The sampling time(s) would need to be determined through further experimentation, though many types of schemes could work.

# References


1. Athamneh, A. I. M., Alajlouni, R. A., Wallace, R. S., Seleem, M. N. & Senger, R. S. Phenotypic Profiling of Antibiotic Response Signatures in Escherichia coli Using Raman Spectroscopy. *Antimicrobial Agents and Chemotherapy* **58**, 1302–1314 (2014).
2. Athamneh, A. I. M. & Senger, R. S. Peptide-Guided Surface-Enhanced Raman Scattering Probes for Localized Cell Composition Analysis. *Applied and Environmental Microbiology* **78**, 7805–7808 (2012).
3. Zu, T., Athamneh, A. & Senger, R. Characterizing the Phenotypic Responses of Escherichia coli to Multiple 4-Carbon Alcohols with Raman Spectroscopy. *Fermentation* **2**, 3 (2016).
4. Zu, T. N. K., Athamneh, A. I. M., Wallace, R. S., Collakova, E. & Senger, R. S. Near-Real-Time Analysis of the Phenotypic Responses of Escherichia coli to 1-Butanol Exposure Using Raman Spectroscopy. *Journal of Bacteriology* **196**, 3983–3991 (2014).
5. Godsey, W. Fresh, Brackish, or Saline Water for Hydraulic Fracs: What are the Options?
6. Sofos, J. N. ANTIMICROBIAL EFFECTS OF SODIUM AND OTHER IONS IN FOODS: A REVIEW. *Journal of Food Safety* **6**, 45–78 (1984).
7. 2009_03_13_estuaries_monitor_chap14.pdf.
8. *Strategies to reduce sodium intake in the United States*. (National Academies Press, 2010).
9. Hajmeer, M., Ceylan, E., Marsden, J. L. & Fung, D. Y. C. Impact of sodium chloride on Escherichia coli O157:H7 and Staphylococcus aureus analysed using transmission electron microscopy. *Food Microbiology* **23**, 446–452 (2006).

10. Doyle, M. P. & Roman, D. J. Response of Campylobacter jejuni to sodium chloride. *Appl. Environ. Microbiol.* **43**, 561–565 (1982).

11. Wu, X., Altman, R., Eiteman, M. A. & Altman, E. Adaptation of Escherichia coli to Elevated Sodium Concentrations Increases Cation Tolerance and Enables Greater Lactic Acid Production. *Applied and Environmental Microbiology* **80**, 2880–2888 (2014).

12. Chemical Disinfectants | Disinfection & Sterilization Guidelines | Guidelines Library | Infection Control | CDC. Available at: https://www.cdc.gov/infectioncontrol/guidelines/disinfection/disinfection-methods/chemical.html. (Accessed: 3rd April 2019)

13. Fletcher, M. The Effects Of Methanol, Ethanol, Propanol and Butanol On Bacterial Attachment To Surfaces. *Microbiology* **129**, 633–641 (1983).

14. Mazzola, P. G., Jozala, A. F., Novaes, L. C. de L., Moriel, P. & Penna, T. C. V. Minimal inhibitory concentration (MIC) determination of disinfectant and/or sterilizing agents. *Brazilian Journal of Pharmaceutical Sciences* **45**, 241–248 (2009).

15. Haft, R. J. F. *et al.* Correcting direct effects of ethanol on translation and transcription machinery confers ethanol tolerance in bacteria. *Proceedings of the National Academy of Sciences* **111**, E2576–E2585 (2014).

16. Dombek, K. M. & Ingram, L. O. Effects of ethanol on the Escherichia coli plasma membrane. *J. Bacteriol.* **157**, 233–239 (1984).

17. Dias, F. S., da Silva Ávila, C. L. & Schwan, R. F. In situ Inhibition of Escherichia coli Isolated from Fresh Pork Sausage by Organic Acids. *Journal of Food Science* **76**, M605–M610 (2011).

18. Breidt, F., Hayes, J. S. & McFeeters, R. F. Independent effects of acetic acid and pH on survival of Escherichia coli in simulated acidified pickle products. *J. Food Prot.* **67**, 12–18 (2004).

19. Yuk, H.-G. & Marshall, D. L. Influence of acetic, citric, and lactic acids on Escherichia coli O157:H7 membrane lipid composition, verotoxin secretion, and acid resistance in simulated gastric fluid. *J. Food Prot.* **68**, 673–679 (2005).

20. Dey, T. & Naughton, D. Cheap non-toxic non-corrosive method of glass cleaning evaluated by contact angle, AFM, and SEM-EDX measurements. *Environmental Science and Pollution Research* **24**, 13373–13383 (2017).

21. Abd-Allah, R. Chemical cleaning of soiled deposits and encrustations on archaeological glass: A diagnostic and practical study. *Journal of Cultural Heritage* **14**, 97–108 (2013).

22. Hong, S. & Elimelech, M. Chemical and physical aspects of natural organic matter (NOM) fouling of nanofiltration membranes. *Journal of Membrane Science* **132**, 159–181 (1997).

23. Menconi, A. *et al.* Effect of different concentrations of acetic, citric, and propionic acid dipping solutions on bacterial contamination of raw chicken skin. *Poultry Science* **92**, 2216–2220 (2013).

24. Almeida, B. *et al.* Yeast protein expression profile during acetic acid-induced apoptosis indicates causal involvement of the TOR pathway. *PROTEOMICS* **9**, 720–732 (2009).

25. Umezawa, H. *et al.* Production and isolation of a new antibiotic: kanamycin. *J. Antibiot.* **10**, 181–188 (1957).

26. Zhang, L. *et al.* Determination of kanamycin using a molecularly imprinted SPR sensor. *Food Chemistry* **266**, 170–174 (2018).

27. Zengin, A., Tamer, U. & Caykara, T. Extremely sensitive sandwich assay of kanamycin using surface-enhanced Raman scattering of 2-mercaptobenzothiazole labeled gold@silver nanoparticles. *Analytica Chimica Acta* **817**, 33–41 (2014).

28. Leung, K.-H. *et al.* An oligonucleotide-based switch-on luminescent probe for the detection of kanamycin in aqueous solution. *Sensors and Actuators B: Chemical* **177**, 487–492 (2013).

29. Wirmer, J. & Westhof, E. Molecular Contacts Between Antibiotics and the 30S Ribosomal Particle. in *Methods in Enzymology* **415**, 180–202 (Elsevier, 2006).

30. Zaman, S. B. *et al.* A Review on Antibiotic Resistance: Alarm Bells are Ringing. *Cureus* (2017). doi:10.7759/cureus.1403

31. CDC. About Antimicrobial Resistance. *CDC*

32. Burmeister, A. R. Horizontal Gene Transfer: Figure 1. *Evolution, Medicine, and Public Health* **2015**, 193–194 (2015).

33. Acred, P., Brown, D. M., Turner, D. H. & Wilson, M. J. Pharmacology and chemotherapy of ampicillin--a new broad-spectrum penicillin. *Br J Pharmacol Chemother* **18**, 356–369 (1962).

34. Merola, G., Martini, E., Tomassetti, M. & Campanella, L. New immunosensor for β-lactam antibiotics determination in river waste waters. *Sensors and Actuators B: Chemical* **199**, 301–313 (2014).

35. Cha, J. M., Yang, S. & Carlson, K. H. Trace determination of β-lactam antibiotics in surface water and urban wastewater using liquid chromatography combined with electrospray tandem mass spectrometry. *Journal of Chromatography A* **1115**, 46–57 (2006).

36. Nestorovich, E. M., Danelon, C., Winterhalter, M. & Bezrukov, S. M. Designed to penetrate: Time-resolved interaction of single antibiotic molecules with bacterial pores. *Proceedings of the National Academy of Sciences* **99**, 9789–9794 (2002).

37. Kobayashi, M. & Shimizu, S. Cobalt proteins. *European Journal of Biochemistry* **261**, 1–9 (1999).

38. Ranquet, C., Ollagnier-de-Choudens, S., Loiseau, L., Barras, F. & Fontecave, M. Cobalt Stress in *Escherichia coli*: *THE EFFECT ON THE IRON-SULFUR PROTEINS*. *Journal of Biological Chemistry* **282**, 30442–30451 (2007).

39. Thorgersen, M. P. & Downs, D. M. Cobalt Targets Multiple Metabolic Processes in Salmonella enterica. *Journal of Bacteriology* **189**, 7774–7781 (2007).

40. Majtan, T., Frerman, F. E. & Kraus, J. P. Effect of cobalt on Escherichia coli metabolism and metalloporphyrin formation. *BioMetals* **24**, 335–347 (2011).

41. Morby, A. P. & Brocklehurst, K. R. Metal-ion tolerance in Escherichia coli: analysis of transcriptional profiles by gene-array technology. *Microbiology* **146**, 2277–2282 (2000).

42. Nie, S. Probing Single Molecules and Single Nanoparticles by Surface-Enhanced Raman Scattering. *Science* **275**, 1102–1106 (1997).

43. Campion, A. & Kambhampati, P. Surface-enhanced Raman scattering. *Chemical Society Reviews* **27**, 241 (1998).

44. Drescher, D. & Kneipp, J. Nanomaterials in complex biological systems: insights from Raman spectroscopy. *Chemical Society Reviews* **41**, 5780 (2012).

45. Sajid, M. *et al.* Impact of nanoparticles on human and environment: review of toxicity factors, exposures, control strategies, and future prospects. *Environmental Science and Pollution Research* **22**, 4122–4143 (2015).

46. Taghavi, S. M., Momenpour, M. & Azarian, M. Effects of Nanoparticles on the Environment and Outdoor Workplaces. *Electronic physician* 706–712 (2013). doi:10.14661/2013.706-712

47. Nowack, B. & Bucheli, T. D. Occurrence, behavior and effects of nanoparticles in the environment. *Environmental Pollution* **150**, 5–22 (2007).

48. Sharma, V. K., Siskova, K. M., Zboril, R. & Gardea-Torresdey, J. L. Organic-coated silver nanoparticles in biological and environmental conditions: Fate, stability and toxicity. *Advances in Colloid and Interface Science* **204**, 15–34 (2014).

49. Clement, J. L. & Jarrett, P. S. Antibacterial Silver. *Metal-Based Drugs* **1**, 467–482 (1994).

50. Maillard, J.-Y. & Hartemann, P. Silver as an antimicrobial: facts and gaps in knowledge. *Critical Reviews in Microbiology* **39**, 373–383 (2013).

51. D19 Committee. *Test Method for Lithium, Potassium, and Sodium Ions in Brackish Water, Seawater, and Brines by Atomic Absorption Spectrophotometry*. (ASTM International). doi:10.1520/D3561-16

52. D19 Committee. *Test Methods for Chloride Ion In Water*. (ASTM International). doi:10.1520/D0512-12

53. D19 Committee. *Test Methods for Cobalt in Water*. (ASTM International). doi:10.1520/D3558-15

54. D19 Committee. *Test Methods for Electrical Conductivity and Resistivity of Water*. (ASTM International). doi:10.1520/D1125-14

55. D19 Committee. *Test Method for Sodium in Water by Atomic Absorption Spectrophotometry*. (ASTM International). doi:10.1520/D4191-15

56. Yasokawa, D. *et al.* Toxicity of Methanol and Formaldehyde Towards Saccharomyces cerevisiae as Assessed by DNA Microarray Analysis. *Applied Biochemistry and Biotechnology* **160**, 1685–1698 (2010).

57. D19 Committee. *Test Method for Volatile Alcohols in Water by Direct Aqueous-Injection Gas Chromatography*. (ASTM International). doi:10.1520/D3695-95R13

58. Tomlins, K. I., Baker, D. M. & McDowell, I. J. HPLC method for the analysis of organic acids, sugars, and alcohol in extracts of fermenting cocoa beans. *Chromatographia* **29**, 557–561 (1990).

59. Yarita, T. *et al.* Determination of ethanol in alcoholic beverages by high-performance liquid chromatography–flame ionization detection using pure water as mobile phase. *Journal of Chromatography A* **976**, 387–391 (2002).

60. D19 Committee. *Test Methods for Acidity or Alkalinity of Water*. (ASTM International). doi:10.1520/D1067-16

61. Metheny, N. A. *et al.* Effect of pH Test-Strip Characteristics on Accuracy of Readings. *Critical Care Nurse* **37**, 50–58 (2017).

62. Mueller, H. F., Larson, T. E. & Lennarz, W. J. Chromatographic Identification and Determination of Organic Acids in Water. *Analytical Chemistry* **30**, 41–44 (1958).

63. Ibáñez, A. B. & Bauer, S. Analytical method for the determination of organic acids in dilute acid pretreated biomass hydrolysate by liquid chromatography-time-of-flight mass spectrometry. *Biotechnology for Biofuels* **7**, 145 (2014).

64. Barrett, J. R. Agriculture: Tracking Antibiotics in Groundwater. *Environmental Health Perspectives* **112**, (2004).

65. Bailón-Pérez, M. I., García-Campaña, A. M., del Olmo-Iruela, M., Gámiz-Gracia, L. & Cruces-Blanco, C. Trace determination of 10 β-lactam antibiotics in environmental and food samples by capillary liquid chromatography. *Journal of Chromatography A* **1216**, 8355–8361 (2009).

66. Serrano, J. M. & Silva, M. Rapid and sensitive determination of aminoglycoside antibiotics in water samples using a strong cation-exchange chromatography non-derivatisation method with chemiluminescence detection. *Journal of Chromatography A* **1117**, 176–183 (2006).

67. Manyanga, V., Dhulipalla, R. L., Hoogmartens, J. & Adams, E. Improved liquid chromatographic method with pulsed electrochemical detection for the analysis of kanamycin. *Journal of Chromatography A* **1217**, 3748–3753 (2010).

68. Liu, Q. *et al.* Simultaneous determination of aminoglycoside antibiotics in feeds using high performance liquid chromatography with evaporative light scattering detection. *RSC Advances* **7**, 1251–1259 (2017).

69. Verdier, M.-C. *et al.* Simultaneous Determination of 12 β-Lactam Antibiotics in Human Plasma by High-Performance Liquid Chromatography with UV Detection: Application to Therapeutic Drug Monitoring. *Antimicrobial Agents and Chemotherapy* **55**, 4873–4879 (2011).

70. Antibiotic Tests. *Charm Sciences*

71. Sierra-Rodero, M., Fernández-Romero, J. M. & Gómez-Hens, A. Determination of aminoglycoside antibiotics using an on-chip microfluidic device with chemiluminescence detection. *Microchimica Acta* **179**, 185–192 (2012).

72. Laborda, F. *et al.* Detection, characterization and quantification of inorganic engineered nanomaterials: A review of techniques and methodological approaches for the analysis of complex samples. *Analytica Chimica Acta* **904**, 10–32 (2016).

73. Clark, L. C. & Lyons, C. Electrode systems for continuous monitoring in cardiovascular surgery. *Ann. N. Y. Acad. Sci.* **102**, 29–45 (1962).

74. Mohanty, S. P. & Kougianos, E. Biosensors: a tutorial review. *IEEE Potentials* **25**, 35–40 (2006).

75. Malhotra, B. D., Singhal, R., Chaubey, A., Sharma, S. K. & Kumar, A. Recent trends in biosensors. *Current Applied Physics* **5**, 92–97 (2005).

76. Bilal, M. & Iqbal, H. M. N. Microbial-derived biosensors for monitoring environmental contaminants: Recent advances and future outlook. *Process Safety and Environmental Protection* **124**, 8–17 (2019).

77. Raman, C. V. & Krishnan, K. S. A New Type of Secondary Radiation. *Nature* **121**, 501–502 (1928).

78. Matousek, P. Deep non-invasive Raman spectroscopy of living tissue and powders. *Chemical Society Reviews* **36**, 1292 (2007).

79. Fisher, A. K. *et al.* The Rametrix™ LITE Toolbox v1.0 for MATLAB ®. *Journal of Raman Spectroscopy* **49**, 885–896 (2018).

80. Zu, T. N. K. *et al.* Assessment of *ex vivo* perfused liver health by Raman spectroscopy: Assessment of *ex vivo* perfused liver health by Raman spectroscopy. *Journal of Raman Spectroscopy* **46**, 551–558 (2015).

81. Olson, M. L. *et al.* Characterization of an evolved carotenoids hyper-producer of Saccharomyces cerevisiae through bioreactor parameter optimization and Raman spectroscopy. *Journal of Industrial Microbiology & Biotechnology* **43**, 1355–1363 (2016).

82. Freedman, B. G., Zu, T. N. K., Wallace, R. S. & Senger, R. S. Raman spectroscopy detects phenotypic differences among *Escherichia coli* enriched for 1-butanol tolerance using a metagenomic DNA library. *Biotechnology Journal* **11**, 877–889 (2016).

83. Chrimes, A. F., Khoshmanesh, K., Stoddart, P. R., Mitchell, A. & Kalantar-zadeh, K. Microfluidics and Raman microscopy: current applications and future challenges. *Chemical Society Reviews* **42**, 5880 (2013).

84. Fleischmann, M., Hendra, P. J. & McQuillan, A. J. Raman spectra of pyridine adsorbed at a silver electrode. *Chemical Physics Letters* **26**, 163–166 (1974).

85. American Society of Nephrology | Kidney Week - Abstract Details. Available at: https://www.asn-online.org/education/kidneyweek/2017/program-abstract.aspx?controlId=2784319. (Accessed: 9th April 2019)

86. Singh, M. *et al.* 3D printed conformal microfluidics for isolation and profiling of biomarkers from whole organs. *Lab on a Chip* **17**, 2561–2571 (2017).

87. Moreira, L. P. *et al.* Raman spectroscopy applied to identify metabolites in urine of physically active subjects. *Journal of Photochemistry and Photobiology B: Biology* **176**, 92–99 (2017).

88. Lee, S. & Lee, D. K. What is the proper way to apply the multiple comparison test? *Korean Journal of Anesthesiology* **71**, 353–360 (2018).

89. Dunnett, C. W. Pairwise Multiple Comparisons in the Unequal Variance Case. *Journal of the American Statistical Association* **75**, 796–800 (1980).

90. Pokala, N. *dunnett.m*. (2012).

91. *Handbook of nutrition and food*. (CRC Press, Taylor & Francis Group, 2014).

92. How, J. A. *et al.* Adaptation of *Escherichia coli* ATCC 8739 to 11% NaCl. *Dataset Papers in Biology* **2013**, 1–7 (2013).

# Appendices

## Appendix A: Code for ANOVA and Comparison Tests

Full_Clustsigscript.m

```
%Parts of anovamatrix found in keystone.RDAModel.DAPCStats.canon(x:y,1), y being 1:10,
11:20, further increments of 10. Can extend this to as many percentages as you'd like, simply
add more columns to anovamatrix and more anova tests later.
anovamatrix = zeros(100,1); %Holds every Principal Component used in separate columns
%anovamatrix(1:100,1)=keystone.RDAModel.DAPCStats.canon(1:100,1)';
anovamatrix(1:10,1)=keystone.RDAModel.DAPCStats.canon(1:10,1)';
anovamatrix(11:20,1)=keystone.RDAModel.DAPCStats.canon(11:20,1)';
anovamatrix(21:30,1)=keystone.RDAModel.DAPCStats.canon(21:30,1)';
anovamatrix(31:40,1)=keystone.RDAModel.DAPCStats.canon(31:40,1)';
anovamatrix(41:50,1)=keystone.RDAModel.DAPCStats.canon(41:50,1)';
anovamatrix(51:60,1)=keystone.RDAModel.DAPCStats.canon(51:60,1)';
anovamatrix(61:70,1)=keystone.RDAModel.DAPCStats.canon(61:70,1)';
anovamatrix(71:80,1)=keystone.RDAModel.DAPCStats.canon(71:80,1)';
anovamatrix(81:90,1)=keystone.RDAModel.DAPCStats.canon(81:90,1)';
anovamatrix(91:100,1)=keystone.RDAModel.DAPCStats.canon(91:100,1)';
group={}; %labelling array
for n=1:length(anovamatrix)
   if n<=10
      group{1,n}='Control';
   end
   if (10<n) && (n<=20)
      group{1,n}='Low';
   end
   if (20<n) && (n<=30)
      group{1,n}='LowLowMid';
   end
   if (30<n) && (n<=40)
      group{1,n}='LowMid';
   end
   if (40<n) && (n<=50)
      group{1,n}='HighLowMid';
   end
   if (50<n) && (n<=60)
      group{1,n}='Mid';
```

```
    end
    if (60<n) && (n<=70)
        group{1,n}='LowHighmid';
    end
    if (70<n) && (n<=80)
        group{1,n}='Highmid';
    end
    if (80<n) && (n<=90)
        group{1,n}='HighHighmid';
    end
    if (90<n) && (n<=100)
        group{1,n}='High';
    end
end
[p,tbl,stats]=anova1(anovamatrix,group); %Overall ANOVA
p=dunnett(stats) %Showing if any of your percentages differ from the control. Function found
at bottom of script.
Fall=tbl{2,5}; %Saves the F Statistic
pall=tbl{2,6}; %Saves the p-value
%After this, these are individual comparisons
group1={};
group1=group(1:20);
[p,tbl,stats]=anova1(anovamatrix(1:20),group1);
[c,~,~,gnames]=multcompare(stats); %Tukey Comparison Test
F1=tbl{2,5};
p1=tbl{2,6};
group2={};
group2=group([1:10 21:30]);
[p,tbl,stats]=anova1(anovamatrix([1:10 21:30]),group2);
[c,~,~,gnames]=multcompare(stats);
F2=tbl{2,5};
p2=tbl{2,6};
group3={};
group3=group([1:10 31:40]);
[p,tbl,stats]=anova1(anovamatrix([1:10,31:40]),group3);
[c,~,~,gnames]=multcompare(stats);
F3=tbl{2,5};
p3=tbl{2,6};
group4={};
group4=group([1:10 41:50]);
[p,tbl,stats]=anova1(anovamatrix([1:10,41:50]),group4);
[c,~,~,gnames]=multcompare(stats);
F4=tbl{2,5};
p4=tbl{2,6};
```

```
group5={};
group5=group([1:10 51:60]);
[p,tbl,stats]=anova1(anovamatrix([1:10,51:60]),group5);
[c,~,~,gnames]=multcompare(stats);
F5=tbl{2,5};
p5=tbl{2,6};
group6={};
group6=group([1:10 61:70]);
[p,tbl,stats]=anova1(anovamatrix([1:10,61:70]),group6);
[c,~,~,gnames]=multcompare(stats);
F6=tbl{2,5};
p6=tbl{2,6};
group7={};
group7=group([1:10 71:80]);
[p,tbl,stats]=anova1(anovamatrix([1:10,71:80]),group7);
[c,~,~,gnames]=multcompare(stats);
F7=tbl{2,5};
p7=tbl{2,6};
group8={};
group8=group([1:10 81:90]);
[p,tbl,stats]=anova1(anovamatrix([1:10,81:90]),group8);
[c,~,~,gnames]=multcompare(stats);
F8=tbl{2,5};
p8=tbl{2,6};
group9={};
group9=group([1:10 91:100]);
[p,tbl,stats]=anova1(anovamatrix([1:10,91:100]),group9);
[c,~,~,gnames]=multcompare(stats);
F9=tbl{2,5};
p9=tbl{2,6};

function p = dunnett(stats, expt_idx, ctrl_idx)
% p = dunnett(stats, expt_idx, ctrl_idx)
% p is vector of p-values for comparing experimental value(s) (expt_idx) w/ a single
% control value (ctrl_idx), corrected by the Dunnett multiple comparison test
% stats is output from anova1
% idx are indicies of means of interest within stats datastructure
% p = dunnett(stats), then ctrl_idx=1, expt_idx=2:length(stats.means)
% p = dunnett(stats, expt_idx), then ctrl_idx=1
% p = dunnett(stats, [], ctrl_idx), then expt_idx=1:length(stats.means), but NOT ctrl_idx
% p = Dunnett's probability for non-zero difference between ctrl_idx and expt_idx means
% Based on Behavior Research Methods & Instrumentation (1981), vol. 13 (3), 363-366
% Dunlap, Marx, and Agamy Fortran IV source code adapted to Matlab
% The output from this function are consistent w/ the Dunnett test implemented in Prism 5.0a
```

```
%
% % Example
% % generate random data
% % groups ctrl and one are zero centered
% % groups two, three, and four are 2,3,4 centered respectively
%
% groupnames = {'ctrl','one','two','three','four'};
% datavector = [];
% k=1;
% for(i=1:length(groupnames))
%    len = rand*20;
%    while(len<10)
%       len = rand*20;
%    end
%    if(i>2)
%       datavector = [datavector i*rand(1,len)];
%    else
%       datavector = [datavector rand(1,len)];
%    end
%    for(j=1:len)
%       group{k} = groupnames{i};
%       k=k+1;
%    end
% end
% [p,t,stats] = anova1(datavector,group); % perform one-way ANOVA
% p = dunnett(stats)
%
% p =
%
%     NaN    0.9238    0.1639    0.0106    0.0002
%
%  p(1) = 'ctrl' vs 'ctrl' = NaN p-value
%  p(4) means 'three' is different from 'ctrl' w/ p-value 0.0106
%  p(5) means 'four' is different from 'ctrl' w/ p-value 0.0002
%
% Navin Pokala 2012


if(nargin<1)
   disp('p = dunnett(stats, [expt_idx], [ctrl_idx])')
   return
end

if(nargin<2)
```

```
   ctrl_idx=1;
   p=[NaN];
   for(expt_idx=2:length(stats.means))
      p = [p dunnett(stats, expt_idx, ctrl_idx)];
   end
   return;
end

if(nargin<3)
   ctrl_idx=1;
end

if(isempty(expt_idx))
   p=[];
   for(expt_idx=1:length(stats.means))
      if(expt_idx~=ctrl_idx)
         p = [p dunnett(stats, expt_idx, ctrl_idx)];
      else
         p = [p NaN];
      end
   end
   return;
end

if(length(expt_idx)>1)
   p=[];
   for(i=1:length(expt_idx))
      if(expt_idx~=ctrl_idx)
         p = [p dunnett(stats, expt_idx(i), ctrl_idx)];
      else
         p = [p NaN];
      end
   end
   return;
end


DF = stats.df;
n_expt_groups = length(stats.means)-1;

mean_ctrl = stats.means(ctrl_idx);
n_ctrl = stats.n(ctrl_idx);

mean_expt = stats.means(expt_idx);
```

```
n_expt = stats.n(expt_idx);

% both are practically zero
if(abs(mean_ctrl)<=eps('single') && abs(mean_expt)<=eps('single'))
   p = 1;
   return;
end

T = (mean_ctrl - mean_expt)/(stats.s*sqrt(1/n_ctrl + 1/n_expt));
Q=abs(T);

R = n_expt/(n_expt+n_ctrl);

BT=sqrt(1-R);
TP=sqrt(R);

if(DF > 2000)
   p = 1 - DUN(Q, n_expt_groups);
   return;
end

% outer integral 0->inf
A1=0;
S=0.14/sqrt(DF);
X0=1;
F0=DUN(Q*X0,n_expt_groups)*SD(X0);

ctr=0;
SUB = 1;
while(A1/SUB < 1e7 || ctr==0)
   Xl=X0+S;
   F1=DUN(Q*Xl,n_expt_groups)*SD(Xl);
   X2=Xl+S;
   F2=DUN(Q*X2,n_expt_groups)*SD(X2);
   SUB=S/3*(F0+4*F1+F2);
   A1=A1+SUB;
   X0=X2;
   F0=F2;
   S=S*1.05;
   ctr=ctr+1;
end

% OUTER INTEGRAL FROM 1 TO 0
A2 = 0;
```

```
S = -0.14/sqrt(DF);
XINC = 1.05;
if(DF <= 12)
   S = -0.03125;
   XINC = 1;
end
X0=1;
F0=DUN(Q*X0,n_expt_groups)*SD(X0);

for(KK=1:16)
   X1=X0+S;
   F1=DUN(Q*X1,n_expt_groups)*SD(X1);
   X2=X1+S;
   F2=DUN(Q*X2,n_expt_groups)*SD(X2);
   SUB = -S/3*(F0+4*F1+F2);
   A2 = A2+SUB;
   if(A2/SUB > 1e7)
      break;
   end
   X0 = X2;
   F0 = F2;
   S = S*XINC;
end
p = 1 - A1 - A2;
if(p < 0)
   p=0;
end

   function dn = DUN(Q,n_expt_groups)
      SP = sqrt(1/(2*pi));
      AREA = 0;
      if(Q < 0)
         dn = 0;
         return;
      end
      sig = 0.07;
      x0 = 0;
      f0 = SP*exp(-x0*x0/2)*(ZPRB((TP*x0+Q)/BT)-ZPRB((TP*x0-Q)/BT))^n_expt_groups;

      ct=0;
      sub=1;
      while(AREA/sub < 1e7 || ct==0)
         x1=x0+sig;
         f1=SP*exp(-x1*x1/2)*(ZPRB((TP*x1+Q)/BT)-ZPRB((TP*x1-Q)/BT))^n_expt_groups;
```

```
    x2=x1+sig;
    f2=SP*exp(-x2*x2/2)*(ZPRB((TP*x2+Q)/BT)-ZPRB((TP*x2-Q)/BT))^n_expt_groups;
    sub=sig/3*(f0+4*f1+f2);
    AREA=AREA+sub;
    x0=x2;
    f0=f2;
    sig=sig*1.05;
    ct=ct+1;
  end

  dn=2*AREA;
end

function g = SD(S)
  g = ((DF^(DF/2))*(S^(DF-1))*exp(-DF*S*S/2))/(gamma(DF/2)*2^(DF/2-1));
end

function zprb = ZPRB(Z)
  %  COMPUTES THE CUMULATIVE PROBABILITY OF NORMAL DEVIATE Z
  % (INTEGRAL OF THE NORMAL DISTRIBUTION FROM -INFINITY TO Z)

  x=abs(Z);

  zprb=0;

  if(x > 12)
    if (Z > 0)
      zprb=1-zprb;
    end
    return
  end

  q=sqrt(1/(2*pi))*exp(-x*x/2);

  if(x > 3.7)
    zprb=q*(sqrt(4+x*x)-x)/2;
    if (Z > 0)
      zprb=1-zprb;
    end
    return;
  end

  t=1/(1.+0.2316419*x);
  P=0.31938153*t;
```

```
      P=P-0.356563782*t^2;
      P=P+1.78147937*t^3;
      P=P-1.821255978*t^4;
      P=P+1.330274429*t^5;
      zprb=q*P;

      if (Z > 0)
         zprb=1-zprb;
      end

   end

return
end
```

# Appendix B: Spreadsheet of Barcodes and Conditions in Samples

**Table B: Example of the Barcode Layout Used**

This shows the number of the scan used, the barcode that the Rametrix™ LITE Toolbox uses, and the concentration of each scan, in particular.

| Index | Barcode | Project | Instrument | Scan Date | Scanned By | Chemical | Concentration | Biological Replicate |
|---|---|---|---|---|---|---|---|---|
| 775 | ENV775 | Environmental Monitoring | Alpha | 2019403 | Hunter Flick | Acetic Acid | 0.00E+00 | 1 |
| 776 | ENV776 | Environmental Monitoring | Alpha | 2019403 | Hunter Flick | Acetic Acid | 0.00E+00 | 2 |
| 777 | ENV777 | Environmental Monitoring | Alpha | 2019403 | Hunter Flick | Acetic Acid | 1.00E-04 | 1 |
| 778 | ENV778 | Environmental Monitoring | Alpha | 2019403 | Hunter Flick | Acetic Acid | 1.00E-04 | 2 |
| 779 | ENV779 | Environmental Monitoring | Alpha | 2019403 | Hunter Flick | Acetic Acid | 2.50E-04 | 1 |
| 780 | ENV780 | Environmental Monitoring | Alpha | 2019403 | Hunter Flick | Acetic Acid | 2.50E-04 | 2 |
| 781 | ENV781 | Environmental Monitoring | Alpha | 2019403 | Hunter Flick | Acetic Acid | 2.60E-04 | 1 |
| 782 | ENV782 | Environmental Monitoring | Alpha | 2019403 | Hunter Flick | Acetic Acid | 2.60E-04 | 2 |
| 783 | ENV783 | Environmental Monitoring | Alpha | 2019403 | Hunter Flick | Acetic Acid | 2.80E-04 | 1 |
| 784 | ENV784 | Environmental Monitoring | Alpha | 2019403 | Hunter Flick | Acetic Acid | 2.80E-04 | 2 |
| 785 | ENV785 | Environmental Monitoring | Alpha | 2019403 | Hunter Flick | Acetic Acid | 3.00E-04 | 1 |
| 786 | ENV786 | Environmental Monitoring | Alpha | 2019403 | Hunter Flick | Acetic Acid | 3.00E-04 | 2 |
| 787 | ENV787 | Environmental Monitoring | Alpha | 2019403 | Hunter Flick | Acetic Acid | 4.00E-04 | 1 |
| 788 | ENV788 | Environmental Monitoring | Alpha | 2019403 | Hunter Flick | Acetic Acid | 4.00E-04 | 2 |
| 789 | ENV789 | Environmental Monitoring | Alpha | 2019403 | Hunter Flick | Acetic Acid | 5.00E-04 | 1 |
| 790 | ENV790 | Environmental Monitoring | Alpha | 2019403 | Hunter Flick | Acetic Acid | 5.00E-04 | 2 |

# Appendix C: Visualization of Data

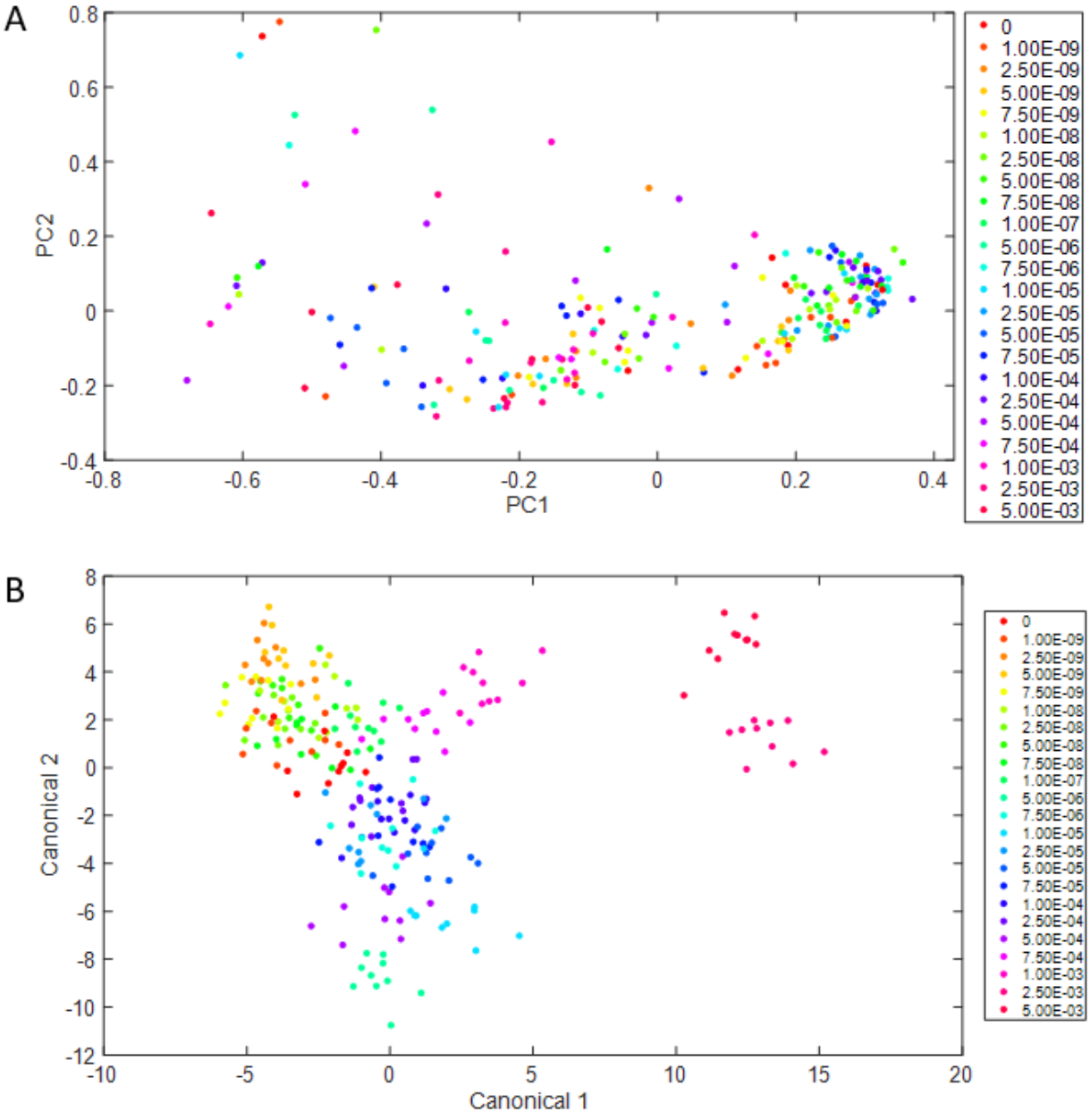


**Figure C-1: PCA and DAPC of All NaCl Concentrations**

**A**: PCA ranging from 5.0E-3 to 1.0E-9 wt %. **B**: DAPC with the same range. There were 230 PCs total leading to 115 used to explain the variability.

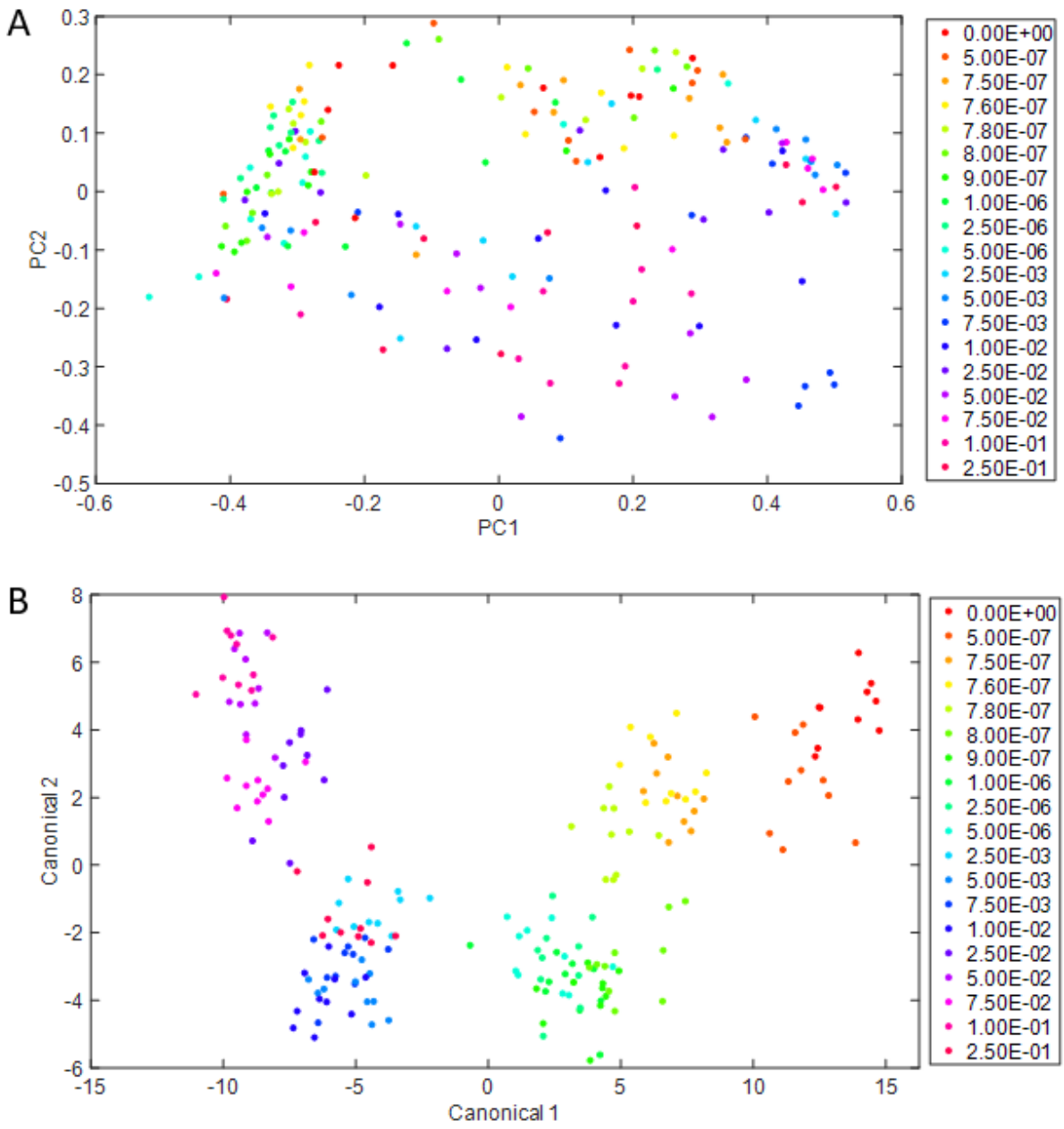


**Figure C-2: PCA and DAPC of All EtOH Concentrations**

**A**: PCA ranging from 0.25 to 5.0E-7 wt %. **B**: DAPC with the same range. There were 180 PCs total leading to 95 used to explain the variability.

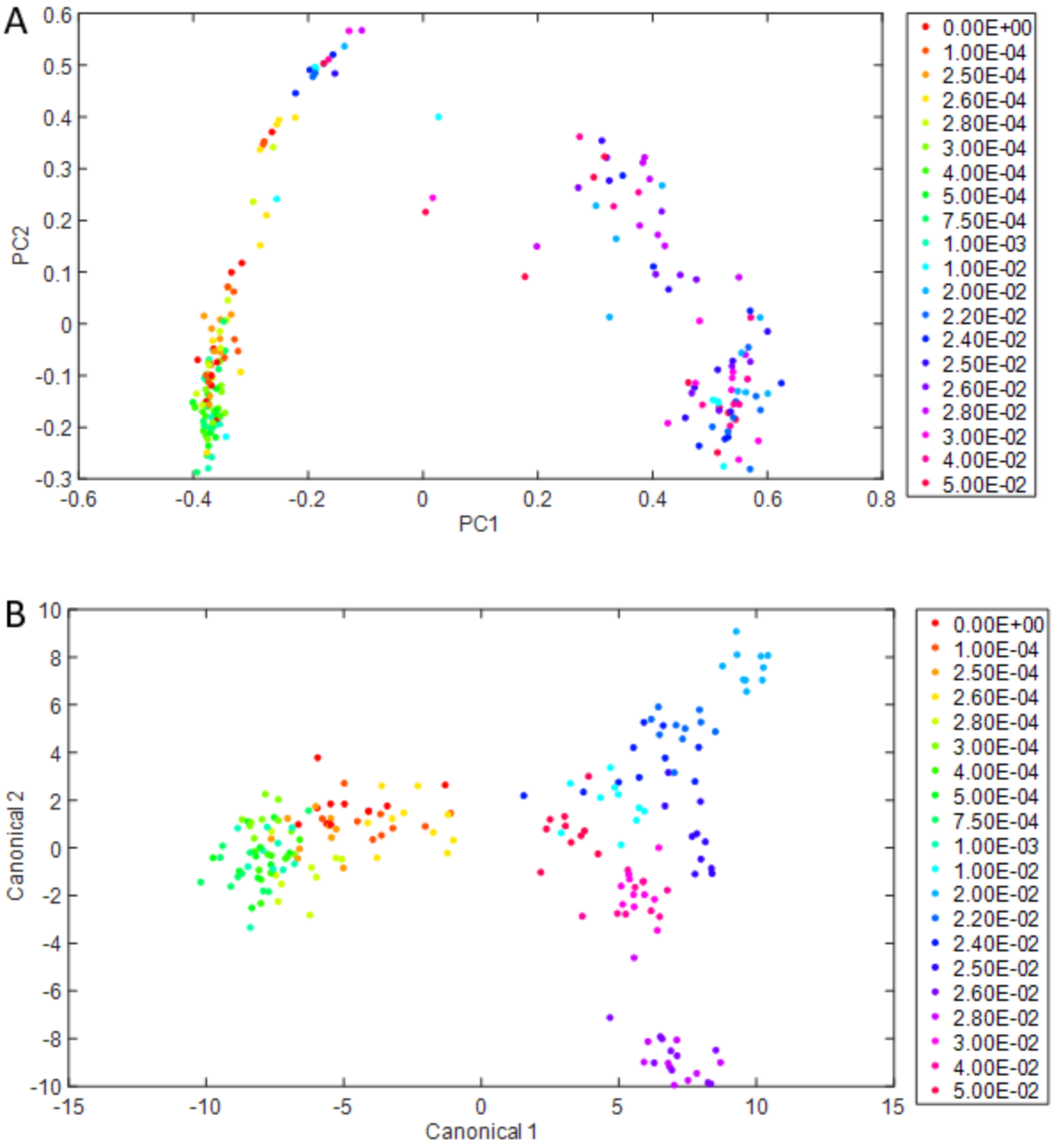


**Figure C-3: PCA and DAPC of All AA Concentrations**

**A**: PCA ranging from 0.05 to 1.0E-4 wt %. **B**: DAPC with the same range. There were 200 PCs total leading to 100 used to explain the variability.

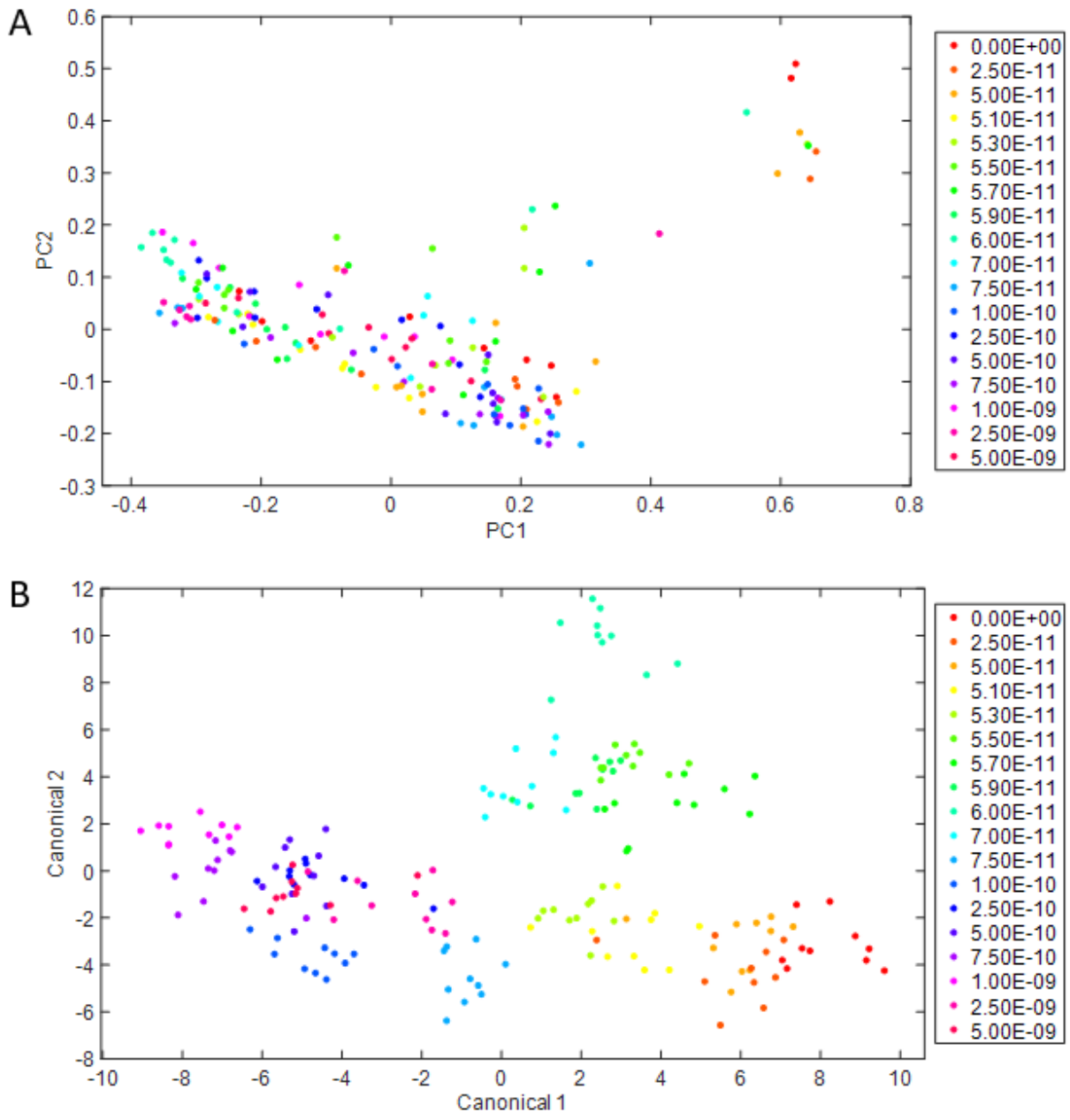


**Figure C-4: PCA and DAPC of All Kanamycin Concentrations**

**A**: PCA ranging from 5.0E-9 to 2.5E-11 wt %. **B**: DAPC with the same range. There were 180 PCs total leading to 95 used to explain the variability.

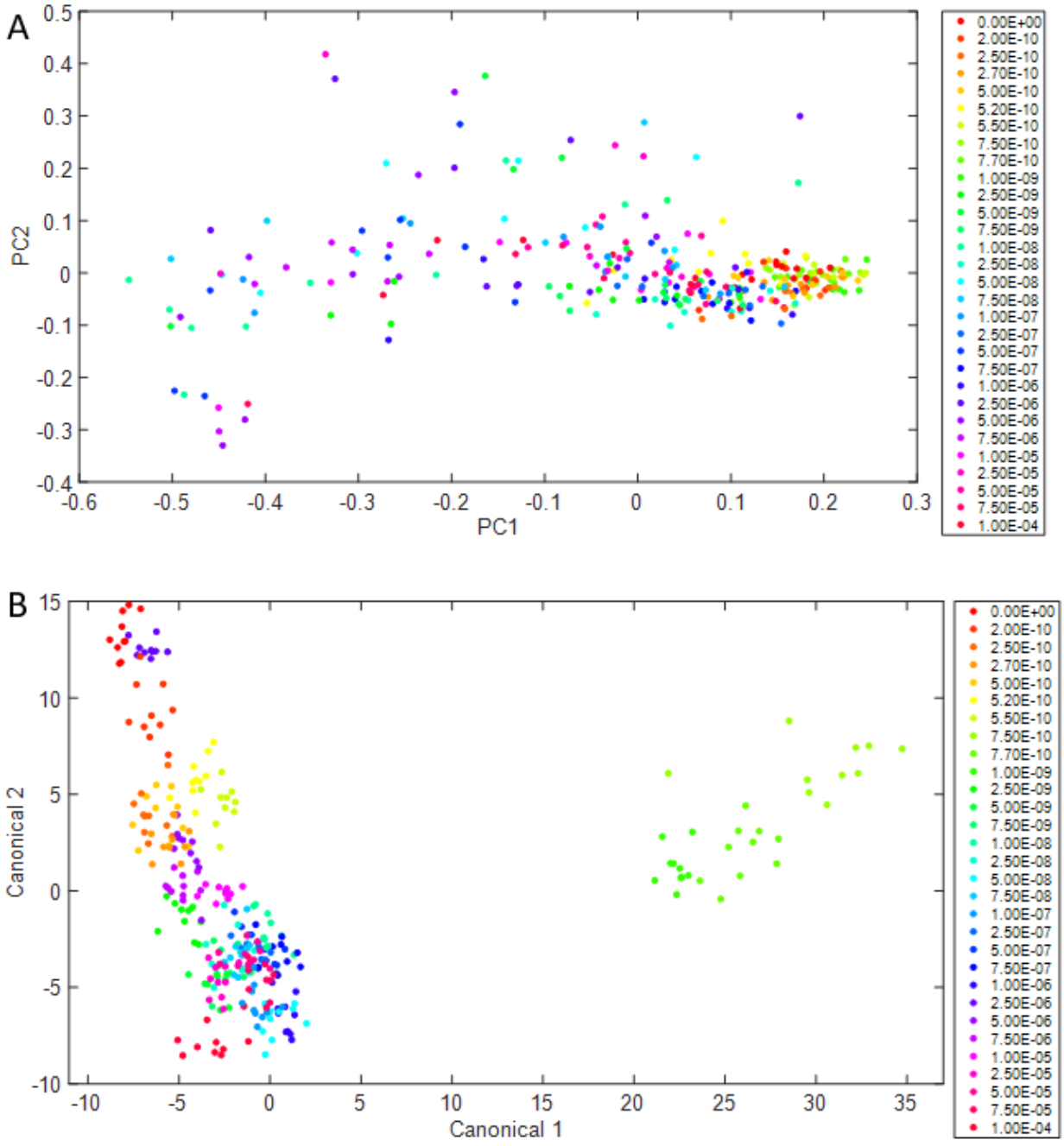


**Figure C-5: PCA and DAPC of All Ampicillin Concentrations**

**A**: PCA ranging from 1.0E-4 to 2.0E-10 wt %. **B**: DAPC with the same range. There were 300 PCs total leading to 150 used to explain the variability.

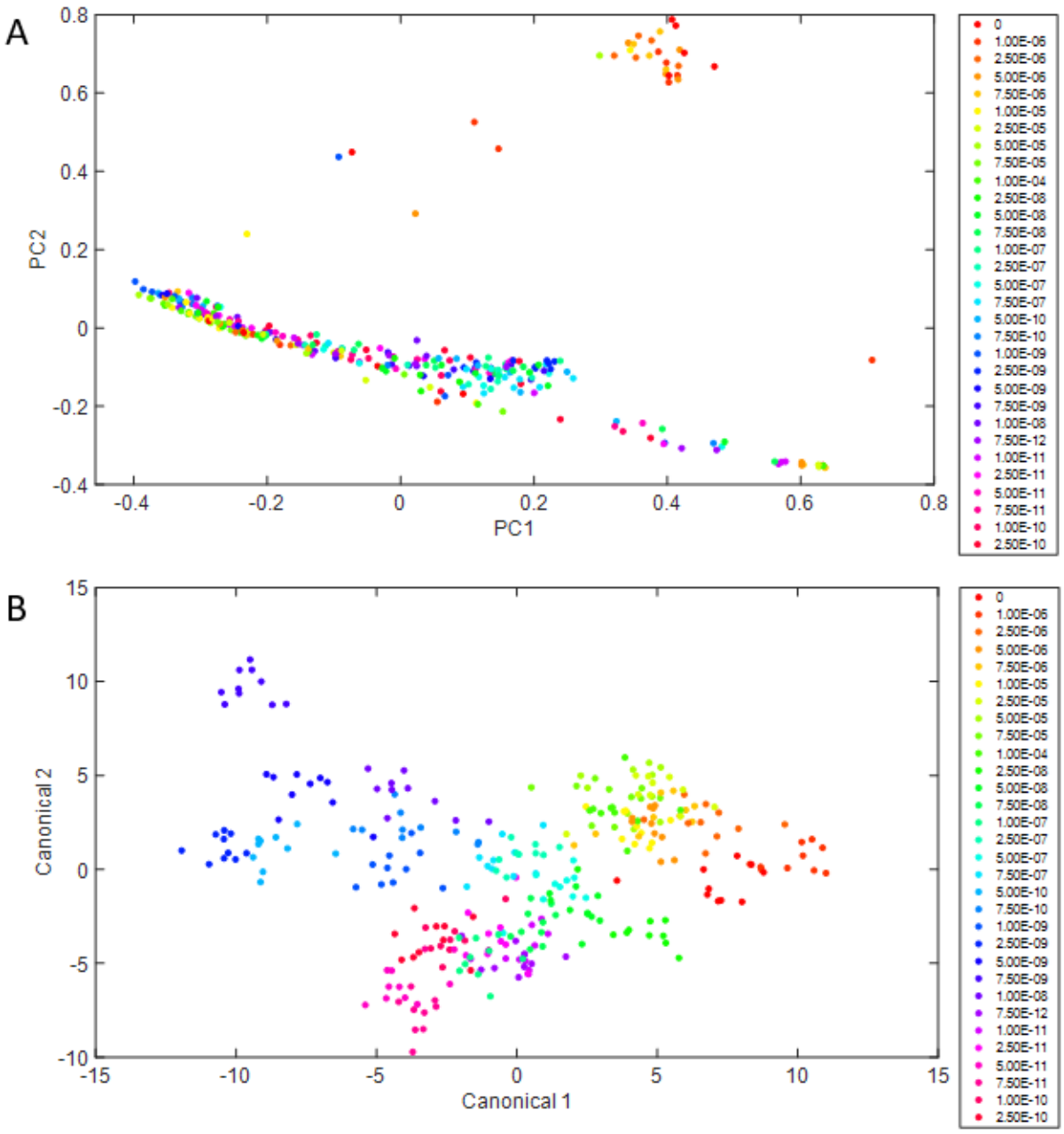


**Figure C-5: PCA and DAPC of All $CoCl_2$ Concentrations**

**A**: PCA ranging from 1.0E-4 to 5.0E-11 wt %. **B**: DAPC with the same range. There were 310 PCs total leading to 155 used to explain the variability.

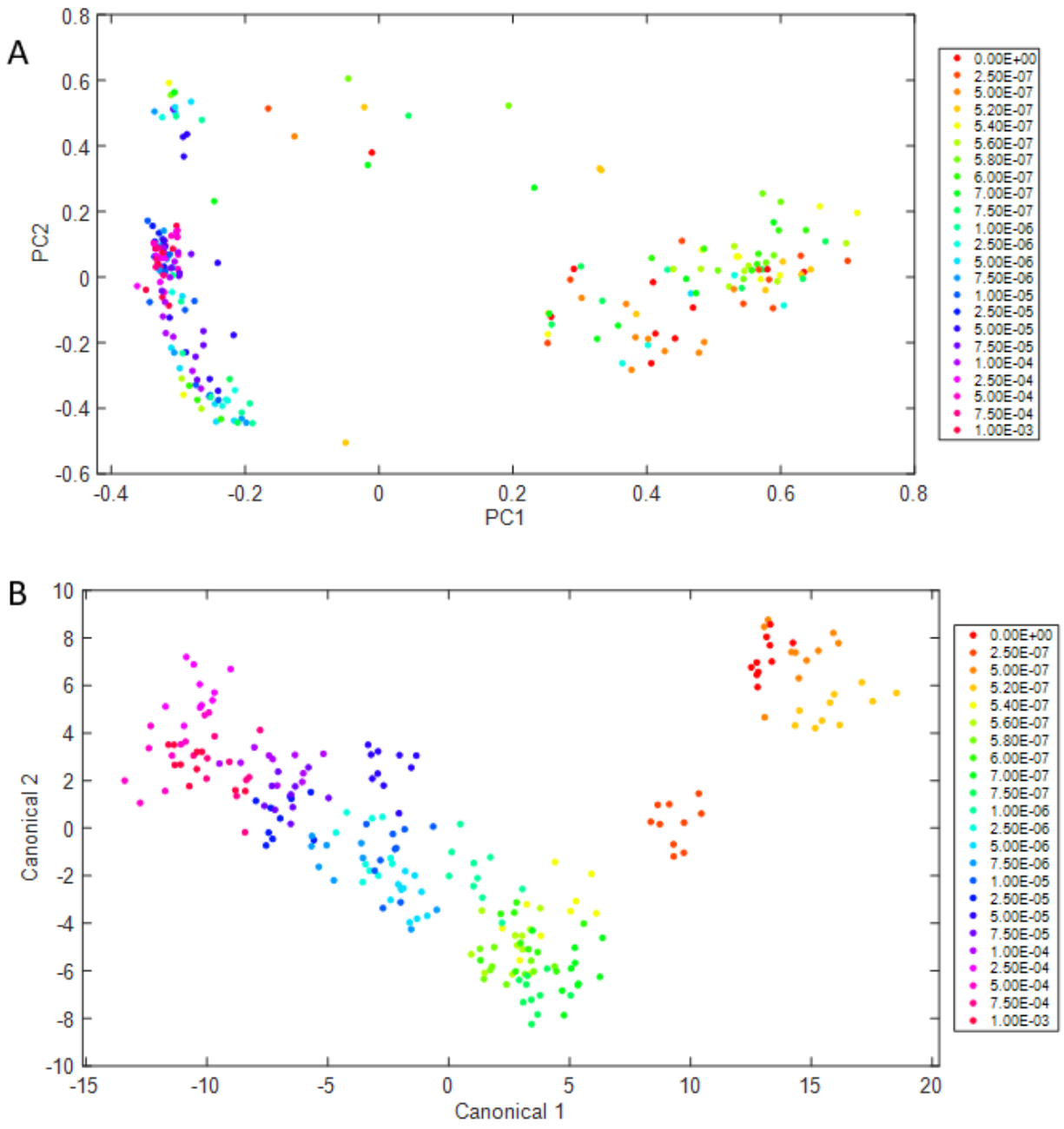


**Figure C-5: PCA and DAPC of All AgNP Concentrations**

**A**: PCA ranging from 1.0E-3 to 2.5E-7 wt %. **B**: DAPC with the same range. There were 230 PCs total leading to 115 used to explain the variability.