Characterization and differentiation of bacteria by hyperspectral analysis

by

Rebecca Dobson Riggs

A dissertation submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy in
Biomedical Sciences

Auburn, Alabama
August 03, 2019

Keywords: Hyperspectral imaging, *Salmonella*, *Escherichia*, *Listeria*, lipopolysaccharide

Approved by

Vitaly Vodyanoy, Chair, Professor Anatomy, Physiology and Pharmacology
Iryna Sorokulova, Professor Anatomy, Physiology and Pharmacology
Sue Hudson Duran, Professor Emeritus of Clinical Sciences
Robert Judd, Professor Anatomy, Physiology, and Pharmacology
Bertram Zinner, Associate Professor of Mathematics and Sciences

Abstract

Characterization and differentiation of microorganisms by hyperspectral analysis is a promising alternative to conventional techniques in clinical, food, and environmental microbiology to be used as a rapid, non-destructive tool. In this work we used the Cytoviva hyperspectral imaging system to elicit the individual hyperspectral profiles produced by three food-borne pathogens, *Salmonella enterica* serovar Typhimurium, *Escherichia coli* O157:H7, and *Listeria monocytogenes* at 18, 21, and 24 hours of growth. Additionally, the individual hyperspectral profiles produced by four *Salmonella enterica* serovar Typhimurium lipopolysaccharide mutants at 18 hours post inoculation were also analyzed. The distinctive feature of the hyperspectral system is that it employs a super-resolution condenser that allows for the capture of high-resolution spectral data of a single microorganism. The microscopic system with hyperspectral capability produces uncomplicated single peak optical spectra of individual cells that can be analyzed by simple computerized methods. In contrast, to the conventional multi-cell hyperspectral methods that require spectral convolution and principal component analysis. Data were collected by the pushbroom method through which an image was made that has both spectral and spatial data by moving the sample pixel line by pixel line to produce distinct profiles for each cell. This super-resolution hyperspectral imaging created sharp spectral profiles based upon the unique surface property of each organism. All organisms were cultured, pelleted, washed, transferred to a poly-l-lysine coated slide, and cover-slipped to obtain the spectral profiles of live, metabolically active bacterial cells. The single-peak bacterial spectra were analyzed by the non-statistical Local Maximum method and by the non-linear fitting of spectra to a particular statistical model. Gaussian, Lorentzian, Logistic, Inverse polynomial, Gumbel, and Gram-Charlier peak functions were applied to bacterial spectra in the range of 400

– 1000 nm. A one-way ANOVA was conducted to compare fitting parameters. With this technique, we were able to characterize and differentiate *Salmonella, E. coli, Listeria*, and four *Salmonella* mutant bacteria with a high level of confidence. This novel use of hyperspectral imaging has the potential to be used as a point-of-care testing and safety scanning at all points of food production, including growing, harvesting, preparation, transportation, distribution, and storage.

Acknowledgements

Many friends and colleagues need to be acknowledged for their encouragement while I pursued my degree. To my supportive husband Michael, I thank you for your continued support day in and day out. I could not have succeeded without you by my side. To my daughter Hallie, you do not even know how inspirational you were to me to pursue my degree. I am blessed to have supportive parents, siblings and in-laws and I would like to thank each of you for your support through the years.

Dr. James Barbaree is one of the reasons I came back to Auburn to receive a graduate degree in Microbiology. I thoroughly enjoyed his class and wanted to learn all that I could from him. He agreed to take me on as a graduate student in the fall of 2013 and I am forever grateful to him. It was his idea to tackle hyperspectral imaging for use as a detection mechanism for food-borne pathogens. Without his support and faith in me as a researcher, I could not have completed the research needed to pursue my degree.

I am extremely thankful and indebted to my adviser Dr. Vodyanoy for his willingness to accept me as one of his students after the passing of Dr. Barbaree. I could not have elevated my research to the level it is and achieve my degree without his unwavering support. It was an honor to work under you this past year and I am forever grateful to you for the knowledge I received while working in your lab as well as the compassion you have for others.

Achieving my PhD was one of my dreams and I undoubtedly owe my successes to Dr. Sorokulova for directing me throughout my entire time at Auburn as a graduate student. Without her continued commitment and encouragement to achieve my dreams, I could accomplish nothing. After the loss of my advisor, she was an integral person persuading me to continue my

education and I will be forever thankful to her for all of the knowledge and encouragement she bestowed upon me over the past six years.

To my committee members, Dr. Judd, Dr. Duran and Dr. Zinner, you are all wonderful mentors who each have aided me throughout this process in different ways. Whether it was problem solving or a friendly smile in the hall, you have each helped mold me into a better student and a person.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AIC	Akaike Information Criterion
BIC	Bayesian Information Criterion
BHI	Brain Heart Infusion
CDC	Center for Disease Control
EPA	Environmental Protection Agency
GCAS	Gram-Charlier peak function for use in chromatography
HSI	Hyperspectral Imaging
LPS	Lipopolysaccharide
PCR	Polymerase Chain Reaction
PFGE	Pulse Field Gel Electrophoresis
RSS	Residual Sum of Squares
SQR	Square
WHO	World Health Organization

Chapter I. Introduction

Foodborne pathogens are among the most significant problems in maintaining the health of the population. CDC estimates that each year roughly 1 in 6 Americans (or 48 million people) are infected, 128,000 are hospitalized, and 3,000 die of foodborne diseases. *Salmonella,* Shiga toxin–producing *Escherichia coli* O157 (STEC) and *Listeria monocytogenes are* among the leading causes of foodborne illnesses in the United States. These pathogens are responsible for millions of annual cases of infectious diarrhea worldwide. Most outbreaks of these infections originate from fresh produce, sauces, and eggs and as a result affect large numbers of the population. According to USDA data, 1.4 million cases of salmonellosis, costing 2.5 billion dollars in terms of lost productivity and medical cost were recorded in the U.S. in the year 2007 alone. Infection with these pathogens results in diarrhea, acute gastroenteritis, relate to the development of long-term complications such as reactive arthritis, hemolytic uremic syndrome, acute renal failure, central nervous system abnormalities and even death. Children under 5 years of age and the elderly (≥65 years of age) are more susceptible to morbidity and mortality from foodborne-induced gastroenteritis. *Listeria*, *Salmonella*, and STEC were the most common causes of hospitalizations (82%) and deaths (82%) reported among persons in outbreaks with a single confirmed etiology. Foodborne infections have a dramatic impact on morbidity and mortality, particularly of infants and children. Rapid detection of pathogens is important for the beginning of adequate treatment and to contain the spread of the pathogens to prevent serious outbreaks. Classical identification of these pathogens is based on cultivation using selective media. More advanced methods for routine detection include immunoassays, polymerase chain reaction based methods, and biosensors. These modern methods offer many potential advantages for the rapid detection of foodborne pathogens in environmental samples and food in comparison

with traditional long-term culture methods. Despite these advantages, their applications for real time detection and environmental monitoring remain in developmental stages with significant methodological hurdles. Consequently, there is an urgent need for new, real-time detection systems for reliable recognition of foodborne pathogens. The main goal of our work is characterizing and comparing three different foodborne pathogens by hyperspectral microscopy, which is a promising tool for rapid, adequate detection of these microorganisms.

Chapter II. Literature Review

II.1 Current Methods of Identification of Foodborne Pathogens

Public health surveillance plays an important role in detecting any outbreak of disease and relies on health professionals to properly report any illness that is on the notifiable disease list. The Center for Disease Control outlines exactly what constitutes an outbreak, how outbreaks are identified and source of the outbreak. The initial, vital step is the reporting of notifiable diseases to local and state departments. This would require medical personnel to culture and have the causative agent properly identified as the cause of the patient's illness in order for the professional to report the illness properly. For an organism to be identified, the handling and testing of the culture is fundamental and is performed by trained microbiologists. Physicians usually notice a rise in the level of cases reported in a given area; and in some severe cases, suspicions will lead to immediate reporting of the illness to the local health department.

Patients suspected of incurring a foodborne illness must submit a stool sample for testing in order to identity the organism causing the illness. This sample is sent to a laboratory to be tested; different techniques are used to identify organisms including biochemical assays, genetic profiles, and molecular subtyping. Once identified pathogens, notably the organisms monitored closely by the CDC and WHO, are sent to state laboratories in order to conduct DNA fingerprinting in which a profile is extracted and entered into the CDC's PulseNet database. This database monitors local and multi-state outbreaks as well as serves as a source for physicians who may suspect similar illnesses in their area. As beneficial as this database is, it can take a few weeks from the initial doctor visit before data is entered into the PulseNet database. However, this system has helped alert the public sooner of potential outbreaks as well as discover the source of the outbreak faster than before this database was created in 1996.

Once an outbreak is suspected, it is imperative to define and discover other possible persons affected by this organism including reviewing PulseNet data, patient records and surveying people who were potentially exposed to the pathogen. Other data is collected about the patient including the timeline of the patient's illness, location of illness onset, food and drinks ingested during the potential exposure time as well as grocery lists to aid in the source of the outbreak. Often surveillance interviews are conducted days to weeks after the initial onset of symptoms and some patients have trouble remembering exactly what they ate on the days leading up to the illness. From here, epidemiological studies commence to investigate the potential exposures in an outbreak as well as if more people reported a specific exposure in order to try to identify the source of the outbreak so that others will not be victim to the same infection.

In order for a laboratory to be able to identify the causative agent in an outbreak, the specimen must be properly collected, stored in appropriate collection media and shipped or delivered to the testing laboratory so that it can be tested as soon as possible. The specimen collected must be cultured on media prior and properly isolated in order for subsequent testing to be performed. There are different types of tests that can be performed by a laboratory to properly identify causative agents including, Polymerase Chain Reaction (PCR) and automatic machines that conduct biochemical assays such as the Vitek II and further strain speciation can be performed through Pulse Field Gel Electrophoresis (PFGE). PFGE is considered the gold standard in identifying the specific strain of a type of pathogen involved in an outbreak. Through this test, an epidemiologist can ascertain whether a *Salmonella* outbreak in Georgia is the same strain causing illness in Michigan.

II.1.1 Identification through Polymerase Chain Reaction

DNA based methods are commonly used for detection purposes, most notably is Polymerase Chain Reaction (PCR). Kary Mullis first discovered this technique in 1984 using Taq polymerase, a thermostable polymerase isolated from the thermophilic bacterium *Thermus aquaticus*, capable of maintaining metabolic activity at high temperatures.[60] A drawback of using Taq polymerase is that it lacks 3 to 5' proofreading ability, often resulting in errors when replicating the target DNA sequence.[26] The use of this process hinges on using a thermostable polymerase that is capable of synthesizing DNA and has a 3→5 prime proofreading exonuclease activity to replicate DNA sequences accurately.

The overall method includes denaturing the target DNA sequence, the addition of primers to the target DNA sequence and amplification of the DNA sequence. Once the process is complete, you must visualize the results by performing gel electrophoresis and staining the gel with a dye, most commonly used is Ethidium Bromide. Through this process, a single DNA sequence can be amplified a million-fold in less than 2 hours.[36]

This breakthrough in DNA amplification has led to further advances in the field including multiplex PCR (mPCR), real-time (rt-PCR) and quantitative PCR (qPCR). mPCR allows for the simultaneous amplification of multiple target genetic sequences and therefore utilizes multiple sets of primers. This technique uses fluorescence to detect, quantify and visualize PCR products on a computer. Using mPCR, Zhou was able to detect 6 foodborne pathogens within a single sample by developing primers targeting *Salmonella enterica*, *Listeria monocytogenes*, *Staphylococcus aureus*, *Escherichia coli* O157: H7, *Shigella* spp., and *Campylobacter jejuni*.[71]

Real-time PCR is capable of monitoring nucleic acid amplification as the reaction advances by measuring the fluorescent signal produced by intercalating dyes or labeled dual probes. The intensity of the signal is directly proportional to the quantity of PCR amplicons

produced. Real-Time PCR can be simplex or multiplex dependent upon the number of primers used in the reaction. An advantage of utilizing this method is the ability to identify the presence of a pathogen and quantify the amount of pathogen present in the sample. While PCR is commonly used in the field, there are a few disadvantages to its use. Performing PCR is laborious and time consuming; its use is not ideal for high throughput use as it is challenging to automate.[32]

II.1.2 The Use of Bacteriophage for Identification

It is common knowledge that bacteria surround us in the air, soil, water, etc. and all humans risk being exposed to bacterial pathogens or viruses on a daily basis, but few realize that bacteria are also susceptible to bacteria-specific viruses known as bacteriophages. Bacteriophages are obligate intracellular parasites that rely on its host metabolism to survive as they lack ribosomes and other items needed to generate energy.[22] Bacteriophages were first discovered in 1915 by Frederick William Twort and later studied in 1917 by Felix Hubert d'Herelle who named the virus a bacteriophage. In 1998 a team of microbiologists at the University of Georgia estimated that there is over 5 million trillion trillion bacteria in the world; therefore, it is not surprising there is a vast number of phages that can be found in the natural environment too.

Bacteriophages, also called phages, can be characterized as tailed, polyhedral, and filamentous or pleomorphic yet 96.4% are tailed phages. Each type consists of a single molecule of DNA, which is often dsDNA, yet few phage groups contain ssRNA, dsRNA and ssDNA molecules. Viruses follow one of two life cycles: Lytic or Lysogenic. Lytic phages infect its host, replicates quickly within the host and induces cell lysis releasing phage particles. This phage type usually encodes host-lethal proteins that serve to disrupt replication, transcription, and

translation as well as alter the cell membrane. Lysogenic or temperate phages insert viral DNA into the hosts' genome thus replicating each time the bacteria replicates until lysis is instigated, usually by a cell disturbance. In contrast to lytic phages, temperate phages do less harm to the host and always encode a repressor protein to block transcription of other phage genes and prevent infection by other lytic phages.

Whether lytic or lysogenic, the host range can vary depending on the specific make-up of the phage. Certain phages can bind to a number of different types of bacteria while others can only infect a specific species or even a specific strain of a bacterium. Bacteriophages must be able to bind to host receptors as well as be able to overtake the restriction modification system, which is used as a defense mechanism against viral infections. The chemical makeup and location of the receptors can vary; however, the majority of receptors are located on the cell wall of the bacterium. Once infected by a phage, receptor sites are modified resulting in a resistant cell as phage tail fibers would no longer be able to attach to receptor sites on the target bacterium. The fact that bacteriophages are bacteria-specific and are not harmful toward eukaryotic hosts has led scientists to study the use of phages for detection of bacteria.

In 1987 Ulitzer and Kuhn developed a phage-based detection method for bacteria using the λ-phage, a temperate phage, with a reporter gene. Their experiment involved manipulating the λ phage to express the *lux* gene from *Vibrio fischeri,* which encodes bioluminescence, to detect *Escherichia coli* in milk. Variations of this experiment have been performed on foodborne illnesses including *E. coli* and species of *Salmonella.* One study compared using the reporter gene to detect *Listeria monocytogenes* to the traditional method of plating and culturing the bacteria. Both methods were deemed "equal in sensitivity" yet the pathogen could be detected within 24 hours using the *lux* reporter gene system where as *L. monocytogenes* requires 4 days to

grow.[34] This is an important feature as faster detection methods are constantly being explored so that physicians can properly diagnose and thus treat the infected patient. Faster detection methods would aid diagnostically while culturing would serve as a confirmation test as certain pathogens have unique growth requirements.

Advances in detection of bacteria by Ulitzer and other researchers were limited in the equipment needed to detect bioluminescence and thus were not useful in certain settings. Researchers reexamined phage-detection techniques and phage display became a new area of interest. The preferential binding of peptides to the surface of filamentous bacteriophages followed by affinity selection of the binders to a chosen target was first described and termed "phage display" in 1985 by George P. Smith.[58] When researching a phage-specific detection system toward a particular pathogen it is important to isolate the phage that binds to the bacteria of interest from a library of phage clones that exhibit "different recognition peptide specific for different targets."[41] This protocol requires skilled personnel that have meticulous techniques. Different types of phage can be used depending on the goal of the research yet the M13 phage is one of the most widely studied and used phage in many aspects of research including sequencing, mutagenesis, probes, phage display libraries, etc. M13 is an *E. coli* specific, temperate, filamentous phage with an ssDNA core surrounded by a protein coat consisting of 5 different types of proteins. Thousands of proteins known as pVIII proteins cover the majority of the phage. While only 5 copies of the pIII proteins exist on each phage, these proteins are extremely important in binding to the host target as well as transferring viral DNA into the host. Early systems focus on specific peptide sequences displayed on the pIII or pVII proteins yet later studies examine peptide sequences located on pVI.[59]

Phage display is a process that involves incubating phages of interest on a plate or on beads coated with the target bacteria. Phages that are not able to bind to the target will be washed away while those that do bind to the target can be eluted and amplified for further research. Once an individual clone has shown great potential and has undergone several rounds of binding and amplification, DNA sequencing and an ELISA must characterize it. While antibodies are sensitive and selective, "they exhibit many disadvantages for use as diagnostic bench-top sensors" and phage display has been discussed as a method to prepare phage as antibody substitute probes.[47]

Biosensors are currently being examined as an even more rapid system of detection of certain bacteria. This level of detection can only be utilized when there is a measurable signal produced when the phage probe binds to the target cell surface. Filamentous phages are often used, as it can serve as a scaffold, capable of displaying a multitude of potential antigen binding sites on its surface. Petrenko established a collection of phage clones in 1997 and which he named "landscape libraries". These landscape phages act like antibodies against certain presented antigens and receptors. In a study of *Salmonella enterica* Typhimurium, researchers focus on the detection of the pathogen using a landscape phage. The majority of earlier studies on the detection of *Salmonella* primarily focus on using antibodies as biodetectors yet this technique is costly and more time consuming. Filamentous phages are more desirable as they are durable, reusable, and can be produced at low cost while they obtain the same sensitivity and specificity of an antibody. Research had to be conducted on whether the physical adsorption of the phage to the surface of a sensor can occur and if so, the exact concentration of phage and amount of phage put on the sensor must be established. Fluorescent microscopy can be used to visualize the binding of phage to the target.

Previous research has concluded that quartz sensors serve as an analytical tool to study interactions at the solid-liquid interface.[47] Nanduri and Sorokulova explain that within this system "acoustic waves in quartz crystal microbalance (QCM) are excited by the generation of a radio frequency alternating voltage in the piezoelectric crystal. Changes in the resonance frequency are usually attributed to the effect of the added mass due to the binding at the solid–liquid interface. Acoustic wave sensors with biosensors can thus be utilized for the real-time study of the adsorption of biochemical macromolecules." [41] Data illustrates "the typical frequency changes resulting from bacterial binding as a function of increasing logarithmic concentrations of free bacteria in suspension." [43] Changes in frequency can be due to the change in mass upon binding or viscoelastic changes that occur when phage bind to bacteria at the liquid-solid interface. Petrenko and Vodynaoy demonstrated this previously in 2003 when they were able to successfully detect β- galactosidase of *E. coli* using phage display and a number of peptide binders were observed. There is sufficient evidence to suggest that phage can serve as a detection element in biosensors. However, applying bacteriophage for detection in the field can be problematic due to the potential for background interference, the length of time for detection to occur and the lack of transportability.

II.1.3 Conventional Microbiology Techniques

Conventional microbiology methods are utilized to detect foodborne pathogens, but require extensive work and extended time to properly identify the causative organism. Once samples are obtained, several steps must be followed in order to properly test the sample for possible contamination by a foodborne pathogen. The FDA and USDA have established protocols for testing food samples for specific foodborne pathogens; their recommendation for all pathogen detection includes a pre-enrichment step, growth in subsequent selective enrichment media and

growth on selective media. A bacterium that grows on the selective agar must then go through biochemical screening and when necessary, confirmed through serological testing. Following each step in the process warrants qualified personnel and time with results available within a few days to weeks. Skilled laboratory technicians following a stringent aseptic technique are needed to avoid contaminating the sample and achieving accurate results. While contamination leads to more work and time, it can also lead to the identification of false positives. This can be problematic when it comes to identifying the causative agent and such issues could cause the laboratory technician to misidentify an organism as well as inhibit them from identifying the actual causative agent. According to the United States Environmental Protection Agency (EPA) disadvantages to culturing for identification include the slow incubation times that is necessary for organisms to be identified, the risk of spreading the growing organism, the amount of labor involved as well as the necessity of trained laboratory technicians. Thus, the creation of new diagnostic techniques that are rapid and non-destructive to the sample would revolutionize the field of diagnostic microbiology.

Table II 1. Summary of different identification methods currently being utilized to identify pathogenic organisms.

Method	Advantages	Disadvantages	Sensitivity	Specificity
Conventional Microbiology	• Methods have been effective for many years • Inexpensive	• Time consuming • Requires trained personnel	Good-Excellent	Good
Polymerase Chain Reaction (PCR)	• Highly sensitive • Specific • Rapid	• Prone to errors in replication • Unknown organisms go undetected	Excellent	Excellent

Multi Locus Sequence Typing (MLST)	• High discrimination • Reproducible • Diversity within a population	• Rather Difficult to perform • Extremely expensive	Good-Excellent	Excellent
Pulse Field Gel Electrophoresis (PFGE)	• High discrimination • Capable of analyzing multiple samples at once • High reproducibility	• Time consuming (~3-4 days) • Labor intensive • Expensive	----------	----------
MALDI-TOF Mass Spectrometry (Automated Method)	• High-throughput capacity • Easy to Interpret Results • Qualitative and quantitative data	• Interference spectra due to host proteins or normal flora • Lacks ability to differentiate between closely related species	Good	Good
Enzyme-linked Immunosorbent Assay (ELISA) (Immunological Method)	• High-throughput capacity • Low cost • Quantitative and qualitative data	• Difficult to create selective antibodies • There is a detection limit for organisms with low abundance	Moderate-Good	Moderate-Good

II.2 Three Foodborne organisms used in study

Salmonella spp. are Gram-negative, rod shaped, facultative anaerobic microorganisms that possess motility through flagella; there are only two known species *Salmonella enterica* and *Salmonella bongori*. *Salmonella enterica* is further divided into 6 subspecies but within these there are over 2600 known serotypes of the 2600, 20 serotypes are commonly attributed to

clinical manifestations. [19] However, some of the uncommon serotypes can also cause disease and thus monitoring for all potential causative agents is important.[31] In 2016, *Salmonella* was reported as having the highest rate of hospitalizations as well as the highest death rates due to domestically acquired foodborne illness. It is estimated to cause as many as 1 million illnesses each year in the United States and over 80 million cases worldwide.[30] The source of infection is due to the consumption of foods or beverages that are contaminated with animal or human feces that is harboring the bacteria as this organism is usually found in the intestinal tract of both and is spread through the fecal-oral route. Salmonellosis is a zoonotic infection due to its' ability to pass between and infect humans and animals. *Salmonella* can also be spread through contaminated plants, which serve as an alternative host. Recently it was discovered that this organism is capable of surviving outside of a host as it has adapted to survive at higher temperatures and in more acidic environments than previously shown.[17] The most common food sources include eggs, contaminated water used on agricultural crops and undercooked meats, specifically poultry. In recent years there has been a rise in the number of outbreaks linked to vegetables and fruit such as cantaloupe.[4,27,7,2] An illustration of the magnitude of a single outbreak of *Salmonella* can be realized from an outbreak of *Salmonella* that occurred in 2012 that was sourced back to a cantaloupe farm who stored the fruits improperly. For this single outbreak, 261 people were sickened across 24 states; of those, 94 were hospitalized and 3 died as a result of their illness.[9] Infection begins with the ingestion of contaminated food or water so that *Salmonella* spp. reach the intestinal epithelium and trigger gastrointestinal disease. Successful infection with *Salmonella* requires different stages: attachment and adhesion to host surfaces, the production of bacterial factors, which will enable invasion and replication, and the ability to overcome or bypass host defense mechanisms. Most people with salmonellosis develop

diarrhea that may or may not contain blood, fever, vomiting and abdominal cramps 12 to 72 hours after infection. A majority of salmonellosis cases goes unreported each year because most resolve without treatment within 3-7 days and ultimately goes unreported. Rapid detection techniques would allow for earlier identification of the pathogen and in return the proper treatment could be administered to patients quickly as well as the identification of the source of the infection.

Escherichia coli is a Gram-negative, rod-shaped, motile bacterium that is associated with both disease and the normal flora in the human gastrointestinal tract. Theodor Escherich first discovered *E. coli* in 1885 from children's fecal samples. He first named this organism *Bacterium coli commune,* but it was renamed in honor of his discovery and recognized as *Escherichia coli* in 1958.[16] A majority of *E. coli* strains are harmless and play a vital role in the digestion process, but some strains are capable of inducing illness. Of the 256,000 suspected *E. coli* infections that occur in the United States each year, *E. coli* O157:H7 causes 36%. *E. coli* was first recognized as a pathogen in 1982 during an investigation into an outbreak of hemorrhagic colitis associated with consumption of hamburgers from a fast food chain restaurant. Retrospective examination of more than three thousand *E. coli* cultures obtained between 1973 and 1982 found only one isolate with serotype O157:H7 and that was a case in 1975.[33] The many strains of *E. coli* exhibit varying degrees of pathogenicity and are traditionally classified based upon the antigenic aspects of the antigens they possess. The O-antigen refers to a somatic, polysaccharide chain while the flagella antigen, referred to as H antigen, is categorized based upon the major protein component of the flagella. The production of Shiga toxins by strains of *E. coli*, including *E. coli O157:H7* can lead to hemorrhagic colitis and hemolytic uremic syndrome which can be life threatening to some patients. Cattle are common reservoirs of

E. coli O157:H7 and often are asymptomatic making it difficult for farmers to be aware of the potential of infection and subsequent spreading through contact with their crops.[43] While there are lower incident rates of *E. coli* in comparison to *Salmonella* or *Campylobacter* infections, the illnesses related to *E. coli* infections exhibit higher hospitalization and fatality rates making *E. coli* an organism of importance in relation to zoonotic foodborne disease.[33] There are specific, stringent protocols that must be followed in order to culture for enteric pathogens, including *E. coli* and they are defined by many institutions but can also be found in the journal of Clinical Infectious diseases.[66]

Listeria monocytogenes is a Gram-positive, rod-shaped, motile, intracellular pathogen that is commonly isolated from the environment (soil, water, and agriculture) because it is generally found in the fecal flora of animals and infects humans through the consumption of contaminated foods. Interestingly, this organism is capable of proliferating in environments other organisms would be inhibited in such as those with high salt concentrations, low temperatures (-2 and -42°C), and over a wide pH range.[8] This organism has unique abilities that allow for escape from host immune responses, specifically phagocytosis by a vacuole and are capable of spreading throughout a person by moving from cell to cell without being targeted by antibodies, neutrophils or antibiotics considering it is no longer located in the extracellular fluid.[28] Thirteen serotypes of *L. monocytogenes* are currently recognized while only three constitute a majority of listeriosis infections (1/2a, 1/2b, and 4b).[24] A majority of listeriosis outbreaks are associated with agricultural crops irrigated with contaminated water.[62] Of all the foodborne illnesses monitored by the CDC, *L. monocytogenes* had the 2nd highest fatality rate and the highest hospitalization rate in the year 2000. *L. monocytogenes* infections can be self-limiting; such as in cases of gastroenteritis, but can also be very dangerous in immunocompromised individuals, pregnant

women, and the elderly. Considering this organism replicates within the liver and is transferred throughout the body in the bloodstream it is thus capable of infecting multiple areas in the body including the brain, the meninges, the stomach and the intestines. *L. monocytogenes* is one of the few organisms capable of crossing the placenta and can result in infection of the fetus upon exposure that can lead to the abortion of the pregnancy, premature birth, or neonatal illness including meningitis and death. Listeriosis has a high mortality rate of 20 to 30 percent of those infected deeming *Listeria monocytogenes* an organism that needs to be closely monitored as well as needs to be rapidly identified if suspected as the source of an infection.

II.3 Bacterial Cell Wall

Bacteria are classified as Gram-positive or Gram-negative based upon the organism's type of cell wall. Hans Christian Gram in which Gram-positive organisms appear purple while Gram-negative organisms appear red under microscopic examination can visualize this in the laboratory through the Gram stain technique developed in 1884. The cell wall of bacteria is multifunctional considering it serves as a protector from outside influences as well as is capable of allowing the transport of nutrients and waste in and out of the cell.

All bacteria possess a cytoplasmic membrane that is composed of a phospholipid bilayer with multiple types of proteins and encloses the contents of the inside of the bacterial cell. Hydrophobic in nature, it acts as a barrier, preventing the leakage of the hydrophilic cytoplasmic constituents and protecting the inside of the cell from environmental insult. This bilayer separates the intracellular contents from the extracellular environment and contains many different receptors and transport systems located within the membrane that facilitate the influx of nutrients and efflux of waste within a cell. The establishment and maintenance of a stable cytoplasmic membrane is vital to cell survival.[13]

Gram-positive organisms have strong, rigid cell walls composed primarily of a thick layer of peptidoglycan, also referred to as murein. It is located outside of the plasma membrane and can make up 90% of the overall cell wall structure. Peptidoglycan is composed of two glucose derived polysaccharides known as N-acetyl glucosamine (NAG) and N-acetylmuramic acid (NAM) that are varied in length of chains. NAM molecules are interlinked through tetrapeptide bonds which provides strength to the structure. Teichoic acids are also found within the cell wall of Gram-positive bacteria and play many important roles including maintaining cell morphology, aiding in the development of proton motive force that is essential for transport across the cell wall, and protects the organism in high temperature and high salt environments. Teichoic acids can be associated with peptidoglycan or attached to the cell membrane though a lipid anchor, wherein it is referred to as lipoteichoic acid. It is important to note the porous nature of peptidoglycan allows for easy transport of most molecules through the cell wall.

Gram-negative organisms have three distinct layers: the outer membrane, a thin layer of peptidoglycan, and the inner, cytoplasmic membrane. Gram-negative organisms possess peptidoglycan within their cell envelopes however; it only constitutes 5-10% in comparison to Gram-positive bacteria. The periplasm is the aqueous compartment located between the inner and outer membranes of Gram-negative bacteria. There are significantly more proteins and oligosaccharides located within the periplasm and this produces a far more viscous compartment and is a less reduced environment than the cytoplasm.[15] Some proteins serve as periplasmic binding proteins that aid in transport of molecules within the cell. Chemoreceptors exist within this layer to aid in detection and cell signaling. It is interesting that the periplasm has been suggested to be a precursor for lysosomes as harmful agents can be isolated and degraded by detoxifying enzymes within this layer. Hydrolytic enzymes can also be found to breakdown

polymers into smaller, functional molecules that can be utilized within the cell. A small layer of peptidoglycan can be found but does not provide near the same protection as observed in Gram-positive organisms.

The outer membrane, unique to Gram-negative bacteria, is composed similarly to the inner lipid bilayer except that it possess an outer leaflet composed primarily of lipopolysaccharide (LPS).[29] The outer membrane is linked to the inner peptidoglycan layer by Braun's lipoprotein.[6] LPS benefits the cell multifold in that it protects the organism from environmental insults such as bile salts and antibiotics, provides stability to the membrane, helps keep a net negative charge on the membrane to allow for transport of molecules in and out of the cell, and is also capable of inducing inflammatory responses.[48] The LPS layer in Gram-negative bacteria consists of the outer, O antigen, or O polysaccharide, layer, a core oligosaccharide, and Lipid A which anchors LPS to the outer membrane. The genes involved in core and Lipid A biosynthesis are highly conserved while variations are observed in the genes involved in O antigen biosynthesis. The conservative nature of lipid A and core oligosaccharide biosynthesis suggests that these two processes play a role in the overall survival of Gram-negative organisms.[56]

Oligosaccharides are attached to Kdo on Lipid A through a series of glycosyl transfers by nucleotide sugar precursors and form what is referred to as the core of LPS. Enzymes responsible for catalyzing the addition of sugars are thought to be located on the cytoplasmic face of the plasma membrane. It is not surprising that *E. coli* and *S. typhimurium* share commonalities in the structure of core oligosaccharides as organisms in the same family of bacteria often share a common core oligosaccharide structure. The core is divided into two regions, the inner and outer core. The inner core is composed of two heptose oligosaccharides attached to Kdo and aids in

maintaining stability. Specifically, the transfer of phosphate groups that occurs with both heptose oligosaccharides aids in membrane stability. This group transfer provides a form of cross-linking between LPS molecules and an environment necessary for membrane interaction with proteins. Strains lacking the ability to transfer phosphate groups or attach heptose oligosaccharides to Kdo have been isolated and are classified as 'deep-rough' mutants. This affects the outer membrane in such a way that the organism becomes more vulnerable to hydrophobic compounds.[51] Considering the vital role of core oligosaccharide to an organism's ability to survive in altered environments may explain why inner core oligosaccharide genes are highly conserved in *E. coli* and *S. typhimurium*. Researchers are studying this as a potential target for future antibiotics, as 'deep-rough' mutants would have difficulty surviving in a host. The outer core region consists of a tri-hexose backbone that may or may not have attached additional side groups. This region serves as an attachment site for O antigens, a virulence factor in pathogenic organisms. The core alone is not considered a virulence factor because when isolated it has not been shown to initiate an immune response *in vivo*. While diversity of the outer core has been observed, the overall core structure is highly conserved in organisms within the same genera or family.[23] Current research focuses on antimicrobial therapies that target the formation of the core oligosaccharide, specifically the enzymes catalyzing the process.

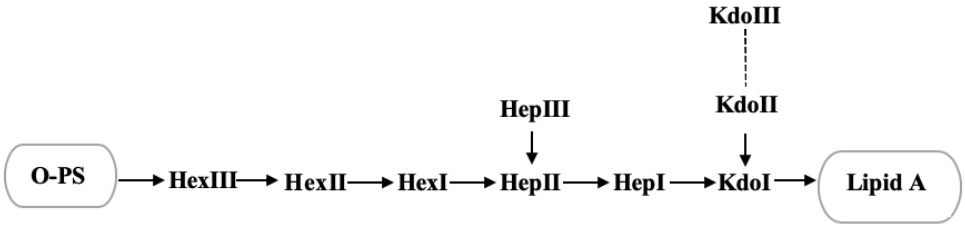


Figure II 1. Diagram of LPS layer in Gram-negative organisms.

Diversity observed in the structure of the core oligosaccharide portion of LPS is dependent upon the organism's ability to survive with the altered core.[23] Typically, alterations lead to a weakened core, which could result in cell death. While core biosynthesis genes have been thoroughly analyzed in *E. coli* and *S. typhimurium*, more so than other organisms, research has yet to prove the exact mechanisms behind this phenomenon.

In a majority of organisms, the ability to synthesize both Lipid A and 3'-deoxy-D-manno-octulosonate, referred to as Kdo, molecules are vital to the organism's overall growth as mutants unable to synthesize Lipid A have yet to be cultured.[51] The waaA operon is a bifunctional enzyme that catalyzes the attachment of two Kdo molecules to Lipid A. It is easy to confuse Kdo as a part of Lipid A because it is synthesized prior to forming a complete Lipid A molecule exhibiting it is a component of the core in LPS. Kdo is necessary to connect the core oligosaccharide to Lipid A. Lipid A is commonly referred to as endotoxin because it will activate complement within the immune system often inducing fever and diarrhea in a host. A vast majority of Gram-negative organisms produce Lipid A molecules that resemble those synthesized by *E. coli* yet the number of carbons being attached to the molecule may differ depending on the organism and the specificity of LpxA.[50] Lipid A is synthesized within the cytoplasm and once the Lipid A-core is generated it is transported to the outer membrane by an ABC transporter.

The O antigen produced by organisms can be used to help identify the genus or species, which is synthesized independently from Lipid A and requires a ligation reaction to attach the O antigen to the Lipid A-core structure and subsequently the completely formed LPS is translocated to the outer membrane.

The oligosaccharide antigen (O antigen) structure consists of a polymer of four to six molecules; the type of sugars, degree of polymerization and glycosidic linkage between sugars vary amongst different organisms. The oligosaccharide is assembled on the cytoplasmic face of the cytoplasmic membrane and must be flipped to the periplasm.[37] While in the periplasm; enzymes modify the O monomer prior to its attachment to Lipid A-core oligosaccharide. The O antigen must be attached to the core oligosaccharide to form complete LPS, classified as smooth. The O antigen is responsible for initiating an immune response toward *S. typhimurium* in a host. The structure of the O antigen is highly variable amongst organisms; even those belonging to the same genera and species can have different O antigens, yet most are heat stable and resistant to alcohol. This is especially evident in the different strains of *Salmonella* where sixty-seven different O antigens have been observed yet all show similar properties.[40] O antigens are assigned the same number when they are highly related, but some antigenic differences may be seen. Considering the high variability concerning the O antigen, it is understandable that the mechanisms of forming the O antigen are variable too. The various O antigens synthesized by a single subgroup of *S. enterica* would be too much to discuss here yet it is important to know that genetic studies are being performed to better understand why a particular strain requires one O antigen for its survival while another can survive within a host with a poorly formed O antigen. Research is also focusing on organism's ability to alter the O antigen once in a host as a possible attempt to avoid recognition by the immune system macrophages.[39] The O antigen is an important factor of LPS but is not necessarily required for survival in all Gram-negative organisms.

The ability to synthesize LPS in the Gram-negative cell envelope is important to an organism's overall ability to survive and replicate. When LPS is properly formed, the organism

is protected from an array of outside influences. While LPS cannot protect Gram-negative organisms from all possible outside influences, without the LPS layer the cell will not survive. As discussed above, modifications made to Lipid A make the organism less sensitive to molecules while the failure to make Lipid A or the use of a different lipid will result in a weaken cell envelope that leaves the organism more susceptible.[69] While core oligosaccharide synthesis is relatively conserved, a few variable reactions do occur. Understanding the exact mechanism and enzymes involved in the biosynthesis of Lipid A and core oligosaccharide synthesis, specifically the attachment of Kdo to Lipid A, have led to research teams developing antibiotics that target these key processes. While O antigens have extreme variability, they can be easily characterized in an organism while the overall role different O antigens play in the same organism is lacking. It has been shown through research that organisms possessing a strong LPS layer in their outer membranes survive better than other organisms with deficiencies in the LPS layer.

II.4. Hyperspectral imaging (HSI) and its potential for identification

There are varieties of systems that perform hyperspectral imaging (HSI) available to researchers and while the machinery can be quite different, the overall setup for each is similar in that each system has a light source, wavelength modulation system and a detector. Different light sources can be utilized and ultimately the light source will determine the range of the wavelength of light exposed to the sample. For example, tungsten halogen bulbs allow for the elicitation of hyperspectral profiles of samples from 400 to 2500 nm; however, tungsten halogen bulbs generate heat that can negatively affect the sample being analyzed.[20] The Cytoviva hyperspectral imaging system utilizes a halogen bulb known to create spectra without extraneous peaks as observed with other light sources, such as Xenon and metal halide light sources.

The second key component in the HSI system is the ability to regulate the wavelength of light exposed to the sample. This can be done through spectrographs, bandpass filter wheels, and tunable filters. The purpose of this is to detect the intensity of the reflectance of light through a sample at specific wavelengths and is coupled with spatial information.

As explained above, various systems employ different equipment, but the detector used by a specific system must be optimal for the technology being utilized. The detector used will maximize the efficacy of the wavelengths of light received by the system. A system using a Halogen bulb provides information in the Visible-Near Infrared (400-1000 nm) will likely have a camera with a charged couple device, which detects light in the 300-1000 nm range; however, the efficiency of detection declines below 400 nm and above 900 nm. Microbiological samples can be sensitive to heat and photodegradation through the exposure of the light source on the sample. In one study it was determined that microbial cells exposed to visible near infrared light between 400-420 nm become deactivated due to the production of reactive oxygen species that led to cell death.[20] An electron multiplying charged couple device can be used to elicit faster image procurement to lessen damage to microbiological samples.

HSI data can be obtained in one of three ways: “whiskbroom” which is also referred to as point mapping, “pushbroom” also termed line scanning or the staredown approach. The different ways of obtaining hyperspectral data is often determined by the type of HSI system. The whiskbroom approach takes the most time to acquire as the data is acquired for a single spatial location and through an automatic stage the sample is then moved pixel line by pixel line so that the next spectrum data can be collected.

There is increasing demand for faster, objective, and automated systems to provide analysts with accurate pathogen detection in a timely manner. Hyperspectral imaging is currently

used in a variety of research fields; one of the more recent areas of interest is its use in the detection and rapid identification of food-borne organisms, more specifically to food and agriculture. Presently, other uses of HIS include acquiring information about effects of cold storage on autolysis in salmon, detection and identification of organisms on a microscopic scale, such as being grown on artificial substrates.[61] The detection and identification of organisms grown directly on the surface of the food product can be examined.[20] Rapid detection of food borne pathogens is vital to production companies to prevent the spread of food borne illnesses. Hyperspectral analysis is being explored as an alternative to current methods.[31]

Cytoviva's hyperspectral microscopy system (Cytoviva, Inc., Auburn, Alabama), utilizes an enhanced Darkfield microscope optic, the Cytoviva Condenser (Cytoviva, Inc., Auburn, Alabama), which improves the signal to noise ratio tenfold compared to other condensers being utilized for hyperspectral imaging. The Cytoviva Hyperspectral Imaging system is capable of visualizing patterns in the Visible near Infrared (VNIR) range, 400-1000 nm, using a halogen bulb. This technology allows for high signal-to-noise optical observation as well as hyperspectral characterization and identification of samples based on their unique surface chemistries. This technology may be utilized in the biological realm as means of identifying and differentiating microbiological organisms in a variety of environments and may be applied in the areas of pathogen detection, food safety, and quality control, among others. Importantly, fluorescent markers are not necessary to confirm the presence of the object of interest using the Cytoviva Hyperspectral Imaging System; this minimalizes sample manipulation and allows for more versatility.

II.5 Need for Rapid Identification Techniques for Foodborne Pathogens

Combined, *Salmonella*, *E. coli* and *Listeria* are estimated to cause over $6.8 billion in foodborne related illness costs and are heavily monitored and studied. Current techniques used to identify pathogens include traditional microbiology, genetic assays and biochemical assays. There are benefits and concerns related to each technique currently employed yet rapid, accurate identification is of the utmost importance. Using our technique to compare the hyperspectral profile of each individual organism is novel in two ways. The first being that we use poly-l-lysine to adhere the motile organism to the slide without disrupting the cell wall; other protocols insist on drying out of the bacterial sample before rehydrating with sterile water prior to imaging. To our knowledge, we are the first to examine the hyperspectral profiles of motile, metabolically active bacterial cells. The second advantage to our approach is the Cytoviva condenser component to the microscope; the Cytoviva condenser allows for improved contrast and signal-to-noise ratio to optimize resolving power and detection capability of non-fluorescing samples through the specialized illuminator that focuses fixed-geometry, highly collimated light at oblique angels onto the sample. A resolution of 90 nm is achieved and using this device and allows for the capture of high-resolution spectral data from the organisms.[64]

Efforts are being made to expand the field of hyperspectral technology to microbiology, most specifically pathogen detection. However, there is a lot to be discovered. The various systems on the market allow for the examination of usability for particularly instruments. For example, conveyer systems are being used that have hyperspectral cameras scanning the surface of a food and when a potential pathogen is detected the lot is marked. These systems however do not allow for specific identification of the pathogen present and ultimately leads to false positives. Developing a system for identification of an organism verses the detection will be beneficial across multiple fields; most importantly, the phenomenal affect on human health.

Other research is being performed on the elicitation of hyperspectral profiles of pathogenic organisms using a microscope-based system; however, these studies were not performed on live, metabolically active organisms, which undoubtedly affected the organism's cell wall and thus the hyperspectral profile.

The ability to report and analyze the hyperspectral data collected is the most difficult aspect of this area of research. A majority of studies use common statistical approaches to report the data but as of yet, no one has analyzed the proficiency of different statistical functions when examining the data or applied these approaches to different microorganisms at different points in the stationary phase of growth. Furthermore, to this group of scientist knowledge, there has been no research into analyzing organisms through hyperspectral analysis that have undergone cell wall mutations. By doing so the research team can assess the proficiency of the system while generating new data for a growing field.

Chapter III. Objectives

While there are a variety of diagnostic methods currently implored for the identification of bacteria, each method is time consuming and testing is usually performed after an initial culture. Hyperspectral imaging for identification of microorganism would revolutionize the field of diagnostics as it has the ability, with further research, to be used to detect organisms within any environment providing rapid, nondestructive methods to identify pathogens within a sample. The ability to identify a pathogen is critical to the outcome of the patient as treatment can vary based upon the infectious organism. Previous research on hyperspectral imaging of microorganisms involved protocols which call for air drying samples in order to immobilize the bacteria and saturating the sample with distilled water prior to imaging. We aim to develop a protocol that allows for the hyperspectral imaging of live, metabolically active organisms by immobilizing the bacteria with precoated poly-l-lysine slides. In addition, using the Cytoviva system we are able to extract the spectral profile of individual, single cells of bacteria verses scanning a surface for detection of bacteria. With the Cytoviva condenser, the profiles extracted are high-resolution images thus creating high-resolution spectral data that is used for analysis. Extracting the hyperspectral profiles of three well studied foodborne pathogens at three time points within the stationary phase of growth we aim to use the data to effectively differentiate between the three organisms at the different time points by analyzing the curves produced by comparing the parameters: area, full-width half max, center and centroid of each at every time point. Next, this group will test the ability of the system to differentiate *Salmonella* LPS mutants by also analyzing the extracted high-resolution spectra in the same manner as the foodborne pathogens. Lastly, we aim to test the effectiveness of common statistical models to differentiate the three-foodborne organisms by their spectral profiles as well as the *Salmonella* LPS mutants.

Overall, there is a profound need for new, rapid diagnostics to identify pathogens and hyperspectral imaging should be further researched, as its potential to the field of detection is overwhelmingly beneficial.

Therefore, we can formulate our objective and aims as following:

Objective: Characterization and comparison of different types of bacteria by Hyperspectral Microscopy

Specific Aims:

1. Prepare bacterial samples suitable for hyperspectral imaging
2. Develop a single cell hyperspectral imaging technique
3. Differentiate three microorganisms using hyperspectral imaging
4. Differentiate bacteria after mutation by hyperspectral imaging
5. Define statistical methods to differentiate bacteria using hyperspectral results

Chapter IV. Materials and Methods

IV.1 Cultures

Salmonella enterica Typhimurium ATCC 29629, *Escherichia coli* O157:H7, and *Listeria monocytogenes* were obtained from the Auburn University (Department of Biological Sciences, Auburn, AL) culture collection.

Salmonella Lipopolysaccharide (LPS) Mutants from Salmonella Genetic Stock Center (SGSC):

Salmonella enterica Typhimurium Strain SL3770 (a) (SGSC 225)

Salmonella enterica Typhimurium SL3749 (a) (SGSC 228)

Salmonella enterica Typhimurium SL1306 (SGSC 163)

Salmonella enterica Typhimurium SL3769 (a) (SGSC 231)

Salmonella enterica Typhimurium SL1102 (SGSC 258)

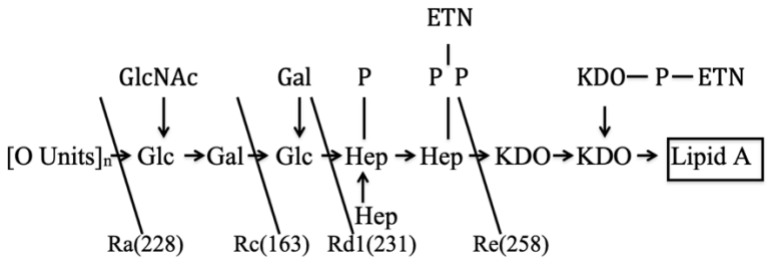


Figure IV 1. Illustration of *Salmonella* with LPS mutations and chemotype annotation.

IV.2 Media

TSA: 40g Tryptic Soy Agar (Thermo Fischer Scientific), 1000 mL

LB: 10 g Bacto-Tryptone (Becton Dickinson, BD), 5g Yeast Extract (BD), 5g NaCl into 1000 mL of Deionized (DI) Water

BHI Broth: 37g BBL Brain Heart Infusion (BD), 1000 mL of DI Water

1xPBS: 25 mL 10x PBS, 225 mL of sterile water, sterilized through filtration

80% Glycerol: 8 mL glycerol, 2 mL sterile water

IV.3 Cultivation of Bacteria

All of the organisms were initially cultured on Tryptic Soy Agar (TSA) and incubated at 37°C for 18 hours. After growth, the organisms were properly stored at 5°C for temporary storage following common microbiology procedure outlined by the Clinical and Laboratory Standards Institute. (Clinical and Laboratory Standards Institute. 2004. Approved standard M22-A3. Quality control for commercially prepared microbiological culture media, 3rd ed. CLSI, Wayne, Pa.) Stock cultures were generated for long term storage of the microorganisms after initial cultivation in Luria Bertani (LB) broth (BD, Franklin Lakes, NJ). Each organism was grown in 2 mL of LB broth for 18 hours before 375 μL of the homogenized overnight culture was transferred to a sterile 2.0 mL cryogenic tube (Corning, Corning, New York) with 625 μL of 80% Glycerol. Cryogenic tubes were stored in -80°C freezer until bacterial cells were required for experimentation.

IV. 4 Preparing bacterial suspensions for imaging

Liquid cultures were prepared by selecting a single colony from a pure specimen on respective Tryptic Soy Agar (TSA) plates and aseptically inoculated into a sterile 14 mL polypropylene tube (VWR, Radnor, Pennsylvania) prefilled with 2 mL of brain heart infusion (BHI) (BD, Franklin Lakes, NJ) broth and sealed with a snap cap lid to allow aeration. The tube was placed in a beaker in an orbital shaker-incubator at 37°C, 200 rpm, (New Brunswick Scientific, Edison, NJ) and incubated overnight. The 2 mL culture was removed and the optical density for each specimen was determined using a spectrophotometer. (Genesys 20, ThermoScientific, Waltham, MA)

Each overnight culture of bacteria were used to create fresh, standardized cultures in which each culture in 25 mL of broth in the 250 mL Erlenmeyer flasks obtained the same, final

concentration of an OD600 equal to 0.02. Each 25 mL culture was incubated in an orbital-shaker incubator at their optimal growth temperature of 37°C to insure proper aeration throughout the cells' development.

The bacterial cells were imaged at 18, 21 and 24 hours of growth due to preliminary experimentation that showed that stationary phase was the best to observe the organisms. At each selected time point, 0.5 mL of bacterial cells were aseptically transferred to a sterile microcentrifuge tube and centrifuged for 10 min at 10,000 rpm under 20°C. (Centrifuge 5430, Eppendorf, Hamburg, Germany). The supernatant was aseptically extracted and discarded while the bacterial pellet was resuspended in 0.5 mL of 1x phosphate buffer saline, pH 7.0 (0.15M NaCl, 5mM NaH_2PO_4) (1x PBS) before being pelleted over for 10 min at 10,000 rpm under 20°C. The supernatant was aseptically extracted and discarded while the bacterial cells were resuspended in 1x PBS.

All of the cultures were serially diluted in tenfold dilutions and the 10^7 dilution was used for capturing image of bacteria. 0.8 μL of each diluted sample was transferred to a sterile poly- L-lysine slide (Sigma Aldrich, St. Louis, Missouri), proven to immobilize motile cells (Colville) while not damaging the bacterial cell wall. A glass cover slip (VWR, Radnor, PA) was placed on top of the sample prior to focusing on the Dark Field Cytoviva Microscope Imaging System. (Cytoviva, Inc. Auburn, AL)

IV. 5 Microscope Design and Image Acquisition

Each of the microorganisms were transferred to sterile poly-L-lysine slides and imaged immediately under the microscope. Figure IV 1 illustrates the Cytoviva Darkfield Hyperspectral Microscope system and how the light is transmitted through the sample and subsequently captured in the spectrograph/camera portion of the system and data is stored within the software.

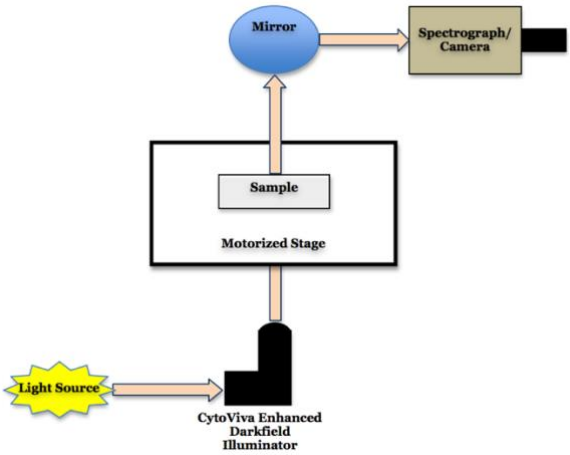


Figure IV 1. Schematic outlining the Cytoviva Microscope System

Images were acquired for each of the organisms at the designated time points in accordance with the protocol in Chapter 2 of the Cytoviva® Hyperspectral User Manual. Spectral Data was extracted following the protocol in Chapter 3 of the Cytoviva® Hyperspectral User Manual.[11]

IV. 6 Statistical analysis

Data averaging, ANOVA, curve fitting, and graph plotting were conducted using Origin 2018 (Northampton, MA) and 2010 Microsoft Excel. The comparison of means was carried out using one-way ANOVA that was followed by Tukey's multiple comparison test. The hyperspectral peaks were analyzed by the Local Maximum model and Statistical models.

Chapter V. Hyperspectral Characterization of Motile Bacteria

Hyperspectral data was collected for *Salmonella*, *E. coli*, and *Listeria* at each time point as stated above. The intensity of the reflection of light off the organism's cell wall is recorded in the range of 0-1 as light is exposed to the specimen in the visible near infrared range. For each cell, a data set is created that consists of 466 data points. Based upon accumulated data, illustrations are produced to visualize the spectra produced by the individual cell. Visually speaking, the spectra of *Salmonella*, *E. coli* and *Listeria* are quite different from each other and these differences are observed at all three time points as well.

V.1 Bacteria at 21 hours of growth

As shown below in Figure V 1, 10 *Salmonella* cells produced relatively the same spectra with peak intensities occurring within a few nm. In Figure, one can observe the average of the 10 cells spectra. This organism's spectra presents a relatively sharp rise initially creating a fairly sharp, round peak followed by a decline in intensity as the wavelength of light exposed to the organism increases.

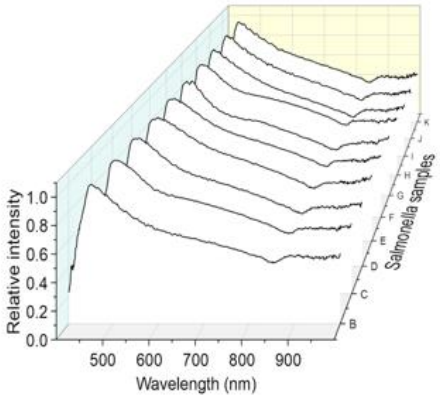


Figure V 1. Spectra of *Salmonella* cells at 21 hours of growth

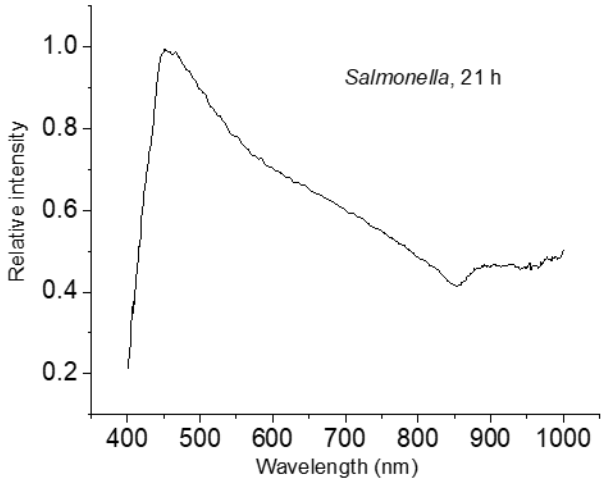


Figure V 2. Averaged spectra of *Salmonella* cells at 21 hours of growth

***E. coli* at 21 hours of growth**

The spectra extracted from the *E. coli* cells at 21 hours of growth appeared to create sharp, more pronounced peaks with a steady rate of decline in intensity as the exposure of the wavelength of light increased.

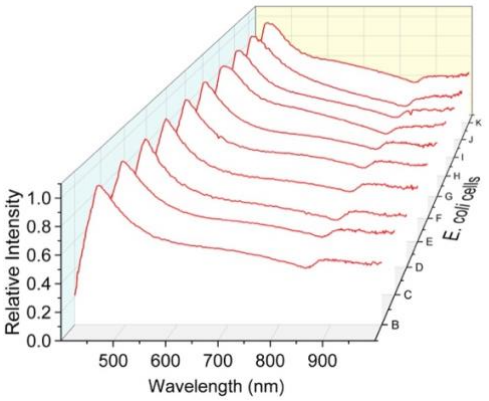


Figure V 3. Spectra of *E. coli* cells at 21 hours of growth

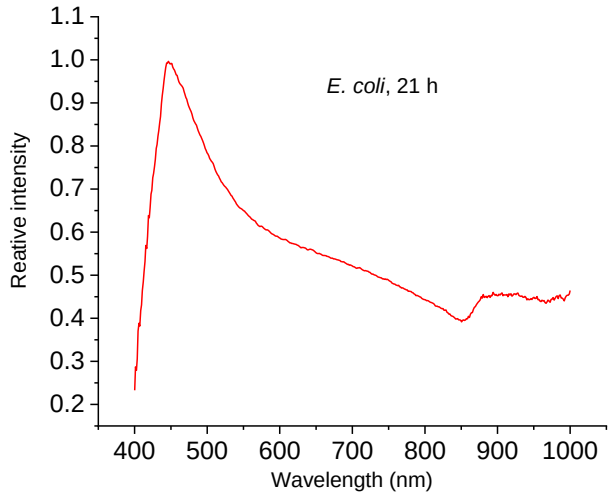


Figure V 4. Averaged spectra of *E. coli* cells at 21 hours of growth

***Listeria* at 21 hours of growth**

The spectra extracted from the 21-hour *Listeria* cells created peaks that appear to be longer and broader than previously observed. In addition, the rate of intensity declines more steadily as the wavelength of light is increased.

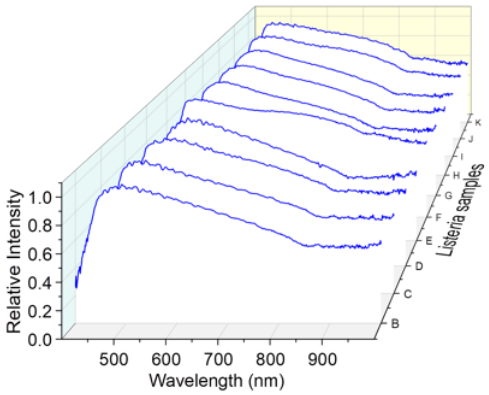


Figure V 5. Spectra of *Listeria* cells at 21 hours of growth

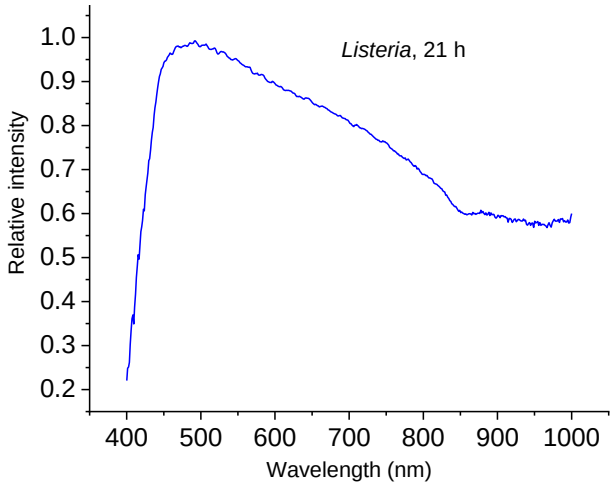


Figure V 6. Averaged spectra of *Listeria* cells at 21 hours of growth

The comparison of the averaged spectra elicited from each organism at 21 hours of growth in Figure V 7, shows how each organism produces its own, unique spectral curves and the data can be analyzed by statistical methods to confirm statistical differences as well as weight the sensitivity of the parameter being tested. Visually though, one can see distinct differences in both the peaks created by each organism as well as the rate of decline of the intensity as the wavelength of exposed light increases. While each organism reached peak intensity in the 440-500 nm range, the overall shape of each the curves are different. *Listeria* created a spectral curve with a broader peak than either *Salmonella* or *E. coli* and a less rapid rate of decline in intensity. The curves produced by *Salmonella* and *E. coli* have sharper peaks with E. coli creating the sharpest, as well as has a more rapid decline in intensity than *Salmonella.*

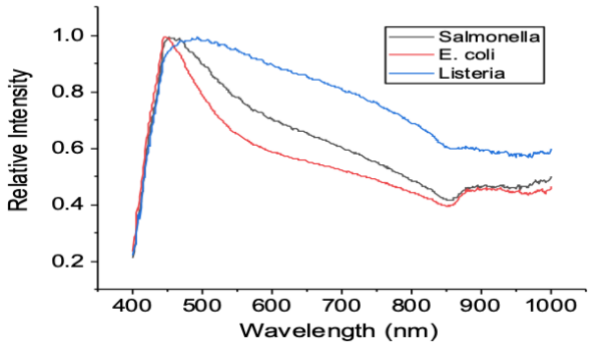


Figure V 7. Averaged spectra for each organism at 21 hours of growth

V.2 Bacteria at 18 hours of growth

The spectral curves of *Salmonella* cells at 18 hours of growth appeared to vary more in the 400-450 nm range; yet, the overall appearance is relatively similar to that observed at 21 hours as the curves produced relatively sharp peaks and a steady decline in intensity as wavelength of light increased.

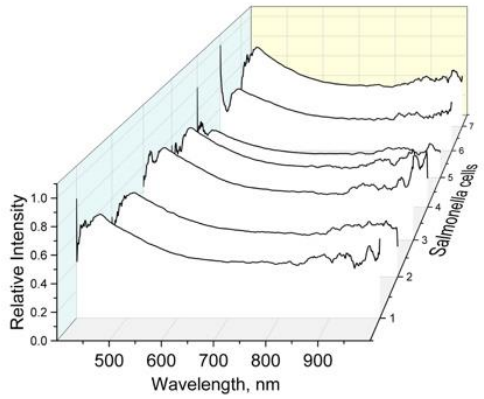


Figure V 8. Spectra of *Salmonella* cells at 18 hours of growth

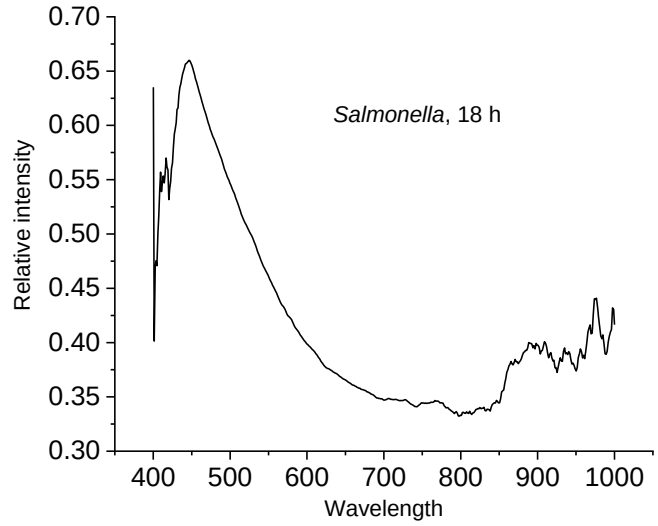


Figure V 9. Averaged spectra of *Salmonella* cells at 18 hours of growth

***E. coli* at 18 hours of growth**

The curves produced by *E. coli* at 18 hours appeared to be more static than that of *Salmonella*. The curves produced sharp peaks with a rapid decline in intensity observed as the wavelength increased.

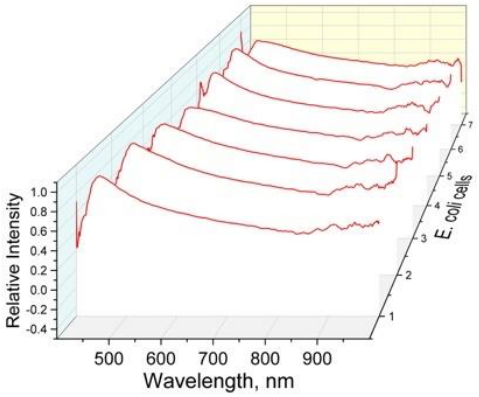


Figure V 10. Spectra of *E. coli* cells at 18 hours of growth

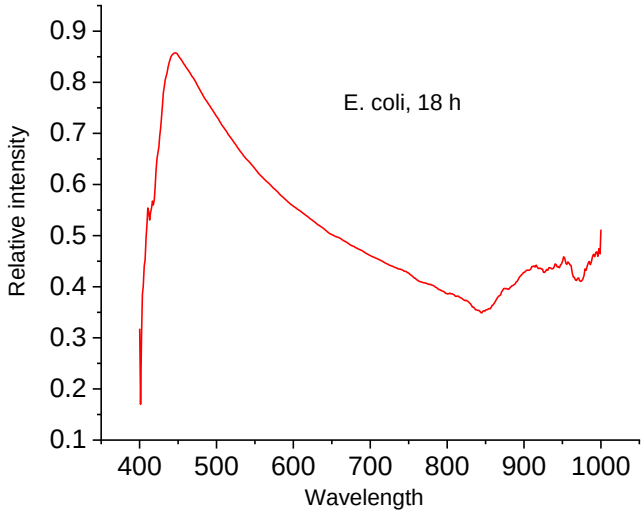


Figure V 11. Averaged spectra of *E. coli* cells at 18 hours of growth

***Listeria* at 18 hours of growth**

The curves produced by *Listeria* cells at 18 hours of growth appear to create even wider, broader peaks with a slight decline in intensity observed as the wavelength of light increased.

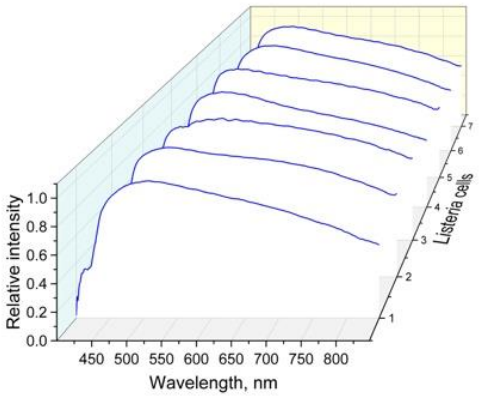


Figure V 12. Spectra of *Listeria* cells at 18 hours of growth

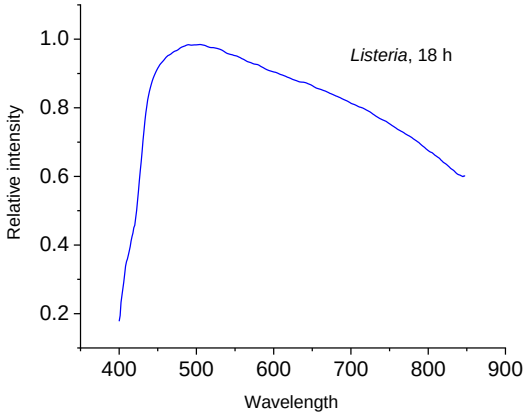


Figure V 13. Averaged spectra of *Listeria* cells at 18 hours of growth

The averaged spectral data for each organism at the 18-hour time point produced unique spectral curves for each organism analyzed. While there is more noise captured at the 18-hour time point at the early exposure of 400-450 nm of light, the curves produced are not the same. At this time point, the averaged *Salmonella* and *E. coli* spectral curves appear to be very similar while the averaged spectral curve of *Listeria* appears to be quite different prompting the necessity for adequate statistical analysis of each of these organisms.

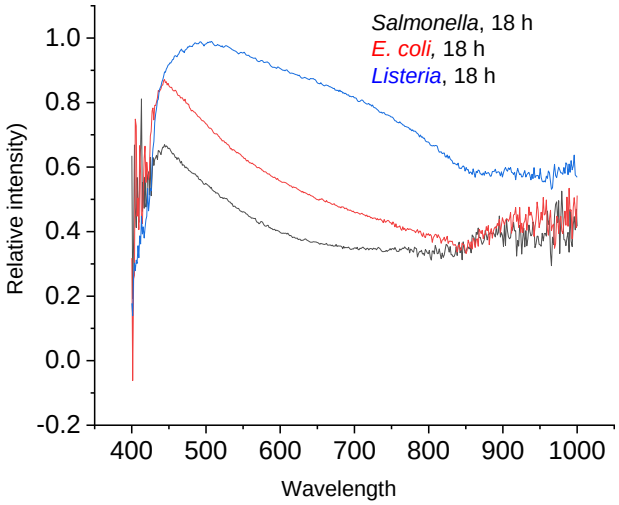


Figure V 14. Averaged spectra for each organism at 18 hours of growth

V.3 Bacteria at 24 hours of growth

The spectral curves obtained from Salmonella at 24 hours displayed sharp peaks with a rapid decline in intensity as wavelength increased. Of all of the time points, the data collected at the 24-hour time point revealed the sharpest peak created by *Salmonella* cells.

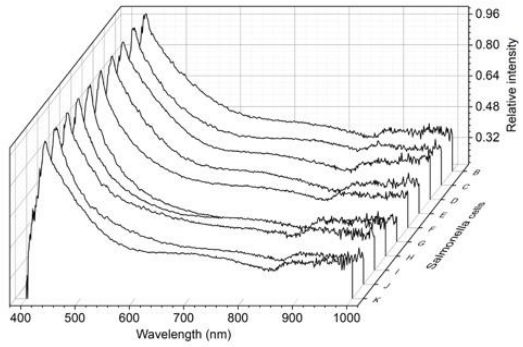


Figure V 15. Spectra of *Salmonella* cells at 24 hours of growth

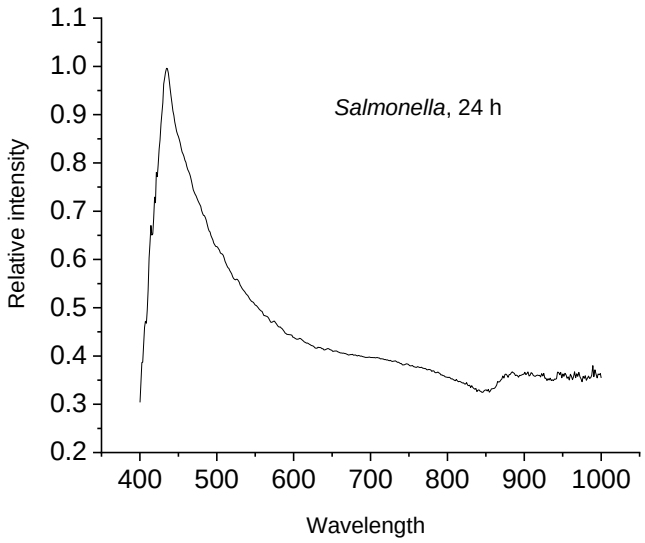


Figure V 16. Averaged spectra of *Salmonella* cells at 24 hours of growth

***E. coli* at 24 hours of growth**

Similarly, to the spectra produced by *Salmonella*, the spectra of *E. coli* created a much sharper peak. The decline in intensity as the wavelength of light increased also showed differences. An initial, rapid decline is observed followed by a steadier decline creating the sharper peak than previously observed at the other time points.

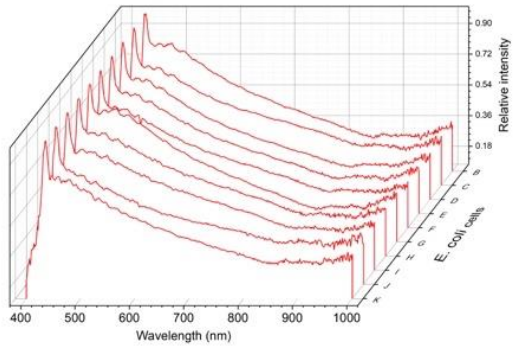


Figure V 17. Spectra of *E. coli* cells at 24 hours of growth

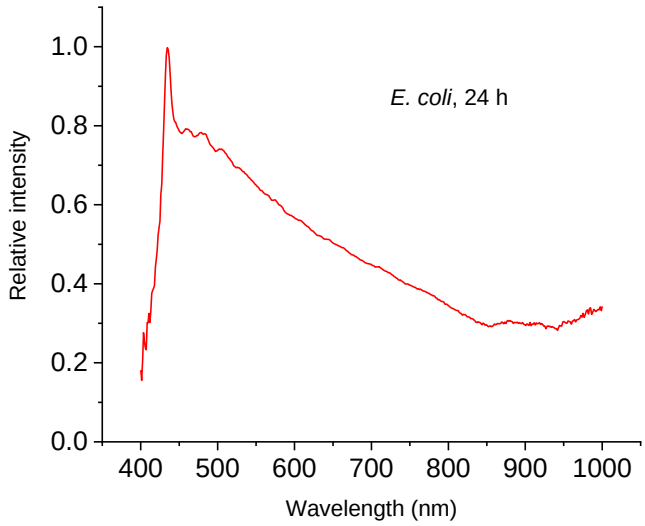


Figure V 18. Averaged spectra of *E. coli* cells at 24 hours of growth

***Listeria* at 24 hours of growth**

Listeria cells at the 24-hour time point exhibited some fluctuations in the curve, as observed with the other time points. In addition, in accordance with the spectra from the other two organisms at 24 hours, the peak of the spectral curve produced a sharper peak. This is in contrast to the broad peaks observed at 18 and 21 hours; however, of the three organisms observed, the peak of the spectral curve produced by *Listeria* at 24 hours appears broader than the other two organisms.

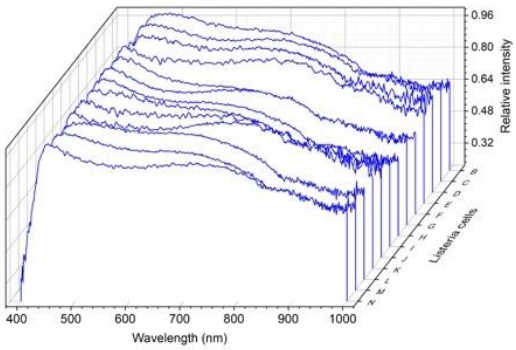


Figure V 19. Spectra of *Listeria* cells at 24 hours of growth

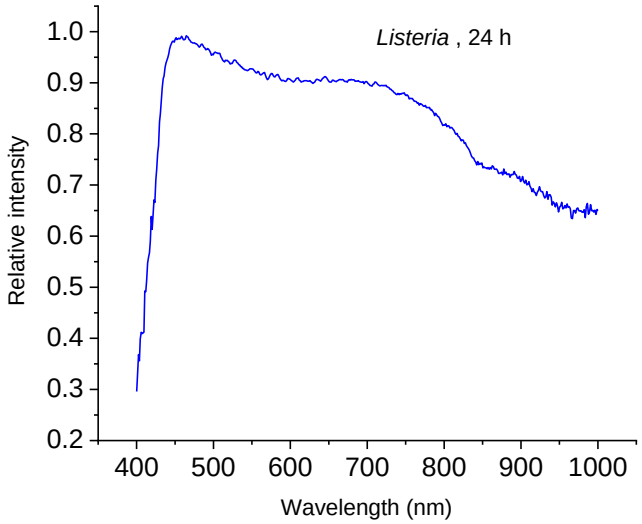


Figure V 20. Averaged spectra of *Listeria* at 24 hours of growth

The averaged spectral data for each organism at the 24-hour time point produced spectral curves for each organism analyzed. While there are more fluctuations observed within the spectral curves at the 24-hour time point, the curves produced are not visually the same. At this time point, the averaged *Salmonella* and *E. coli* spectral curves appear to create similar peaks, but differences occurred in the rate of decline in relative intensity as wavelength increased.

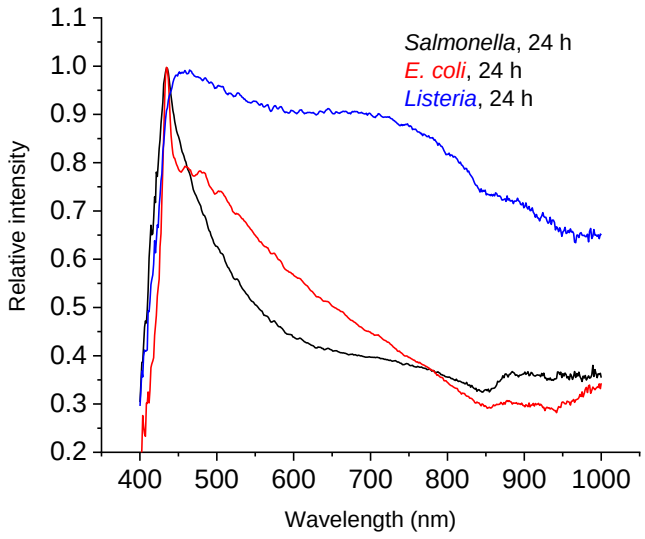


Figure V 21. Averaged spectra for each organism at 24 hours of growth

All of the data collected for the three organisms at the 18, 21, and 24-hour time points will be utilized in order to perform adequate statistical analysis of the data to determine which time point will be ideal for analyzing microorganisms by the spectral curves composed. Further analysis is necessary to gage the level of discrimination that can be applied in order to classify these organisms.

Chapter VI. Hyperspectral characterization of bacteria with mutations

The spectral curves produced by *Salmonella enterica* Typhimurium and the 4 lipopolysaccharide mutant organisms were quite similar in both the shape and location of the peak of each curve; however, the rate the intensity declined as the wavelength of light increased is quite different. Using statistical approaches, we will observe the overall shape of each of the curves and determine whether they can be properly differentiated by using statistical methods.

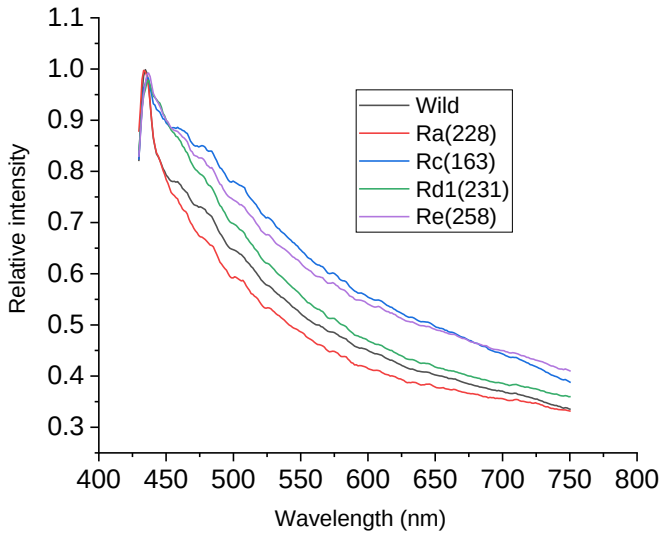


Figure VI 1. Averaged spectra for Salmonella LPS mutants and wildtype at 18 hours

Chapter VII. Statistical analysis of bacterial spectra

The aim of this chapter is to answer specific questions:

1. Which fitting model fits best?
2. Does one model fit all bacterial spectra?
3. Are fitting parameters different for different bacteria?

We will compare spectra of *Salmonella enterica Typhimurium*, *Escherichia coli* O157:H7 and *Listeria monocytogenes* obtained at 18, 21, and 24 hours of growth. Then we compare 4 *Salmonella enterica Typhimurium* lipopolysaccharide mutants with the wild type. We will use the following strategy:

1. Calculate parameters of spectra for each bacterium using the Local Maximum method and compare parameters of the spectra to determine statistically whether the bacteria can be discriminated by the calculated parameters.
2. Fit spectral peaks with statistical functions, rank models and find those that better describe the experimental spectra and compare the model parameters to examine if the statistically fitted model is able to discriminate bacteria by spectra.

VII.1 Spectra of *Salmonella, E. coli* and *Listeria* (Local Maximum model)

We will calculate parameters of *Salmonella*, *E. coli* and *Listeria* at three points of the bacteria growth: 18, 21, and 24 hours. We will start with the midpoint, 21 hours into the growth cycle.

VII.1.1 *Salmonella* at 21 hours of growth

The Local Maximum is the only not fitting statistical model in our study. It does not use any underlying function. The Local Maximum method is a maximum searching algorithm, which finds the local maximum in a moving window. A predefined number of local points determines the window size. Initially, an n-point window is placed at the start point of data stream. The maximum in this window, as well as its index, is recorded. Then the window is moved one-step

further. If the new maximum is greater than the saved maximum, update both the maximum value and index value and then move forward. By using this method, we can calculate the parameters: center, full width at half maximum (FWHM), area under peak, and centroid.

To calculate parameters of the hyperspectral peak of 21 h *Salmonella* (Figure VII 1) analysis to the representative data set of 10 cells showing the normalized relative intensity of the spectra vs the wavelength value using Origin 2019 software The set contains 466 data points for each cell in the range of 400 – 1000 nm.

Table VII 1. The sample of the initial portion of the Workbook for *Salmonella* spectra

A, nm	Cell 1	Cell 2	Cell 3	Cell 4	Cell 5	Cell 6	Cell 7	Cell 8	Cell 9	Cell 10
400.0	0.224	0.231	0.214	0.145	0.180	0.208	0.249	0.168	0.221	0.269
401.2	0.234	0.252	0.183	0.217	0.211	0.251	0.251	0.221	0.271	0.229
402.4	0.270	0.286	0.249	0.189	0.227	0.243	0.240	0.219	0.233	0.291
403.6	0.274	0.280	0.239	0.245	0.261	0.319	0.268	0.256	0.260	0.277
404.7	0.343	0.325	0.277	0.264	0.252	0.333	0.326	0.272	0.314	0.351
405.9	0.414	0.378	0.323	0.339	0.295	0.348	0.332	0.342	0.420	0.449
407.1	0.391	0.350	0.301	0.285	0.328	0.340	0.319	0.374	0.369	0.417
408.3	0.419	0.376	0.334	0.355	0.295	0.394	0.374	0.415	0.405	0.419
409.5	0.382	0.381	0.311	0.343	0.345	0.387	0.363	0.407	0.391	0.396
410.7	0.422	0.442	0.400	0.347	0.397	0.406	0.405	0.431	0.427	0.431

A – wavelength in nm, A-B relative intensities of cells from the cell 1 to 10, respectively.

In order to estimate parameters by the Local Maximum method, the mean values were calculated for each wavelength for all cells using the Origin function “Statistics on Rows” and apply a Peak Analyzer module of the software.[45] The mean value of the *Salmonella* 21 h relative spectrum intensity (Figure VII 1) is then analyzed by the Peak Analyzer using the Local Maximum function.

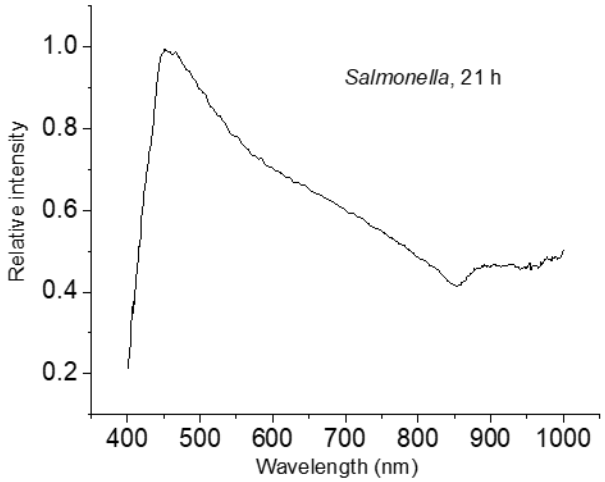


Figure VII 1. Relative intensity of the mean value of spectra for ten individual *Salmonella* cells as function of wavelength (21 h)

Starting with the Peak Analyzer Preview of the peak we find a baseline of the peak, then enter the Peak Finding Setting including number of widow points, Peak Filtering Method "By Height%", Threshold Height%. After this, choose output values of the peak: Peak Area, Peak Center, Peak Centroid, and FWHM, and Finish calculations. The output parameters will show up as shown in Table VII 2.

Table VII 2. Values of Area, WHM, Center, and Centroid of the mean relative intensity for *Salmonella* 21 h (Mean of 10 cells)

Area, nm	FWHM, nm	Center, nm	Centroid, nm
206.98	278.09	451.30	593.05

Area – area under peak, FWHM – full width at half maximum, Center – wavelength of peak maximum, Centroid – wavelength of mass center of a peak.

We will now calculate values of Area, FWHM, Center, and Centroid of the relative intensity for *Salmonella* 21 h for each of 10 individual cells from Table VII 1, using Batch Peak Analysis with Theme Parameters. The results are presented in Table VII 3.

Table VII 3. Values of Area, WHM, Center, and Centroid of the relative intensity for *Salmonella* 21 h for each of 10 individual cells.

Cells	Area, nm	FWHM, nm	Center, nm	Centroid, nm
1	196.37	250.38	450.10	592.08
2	208.70	301.78	453.71	597.63
3	221.23	291.28	465.76	596.64
4	249.80	329.64	466.96	601.05
5	230.12	307.05	465.76	599.38
6	201.43	271.01	468.17	592.98
7	206.59	305.27	453.71	599.69
8	213.47	252.48	451.30	588.54
9	190.50	230.46	451.30	584.84
10	193.62	248.95	451.30	587.93

Carrying out the statistics on columns of Table VII 3, we find mean values and standard errors for the Area, FWHM, Center, and Centroid for the group of 10 cells of *Salmonella* 21 h, shown in Table VII 4.

Table VII 4. Area, FWHM, Center, and Centroid for the group of 10 cells of *Salmonella* 21 hr.

	N total	Mean	Standard Deviation	SE of mean
FWHM	10	278.83	32.58	10.30
Area	10	211.18	18.36	5.81
Center	10	457.81	7.73	2.44
Centroid	10	594.08	5.64	1.78

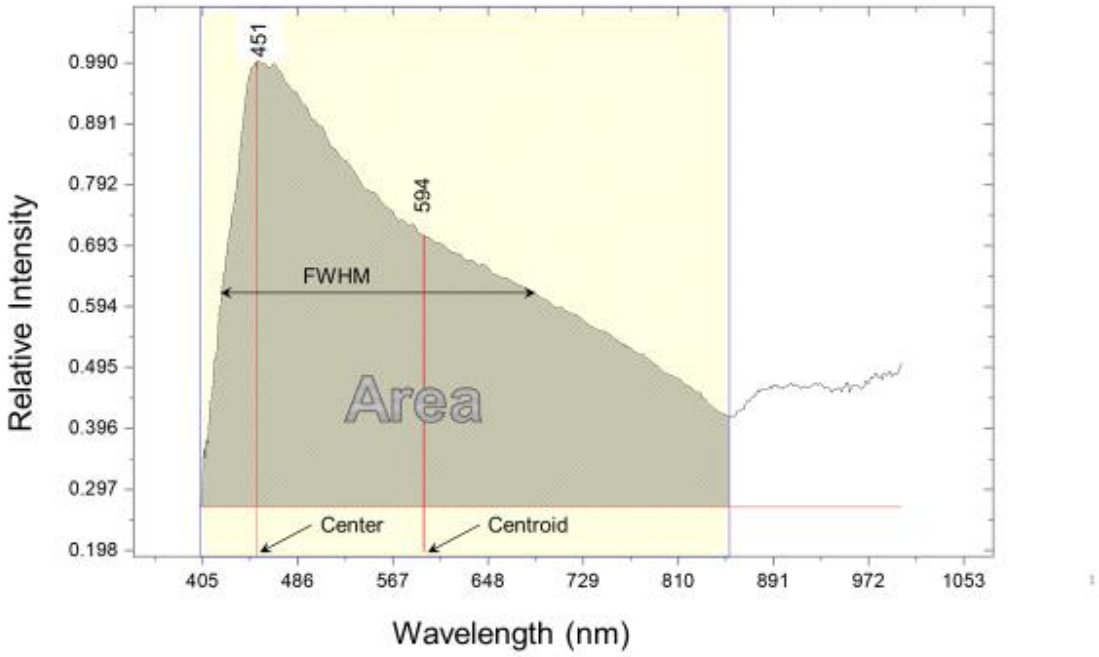


Figure VII 2. Area, FWHM, Center, and Centroid of the *Salmonella* 21 h of the mean of ten cells spectra.

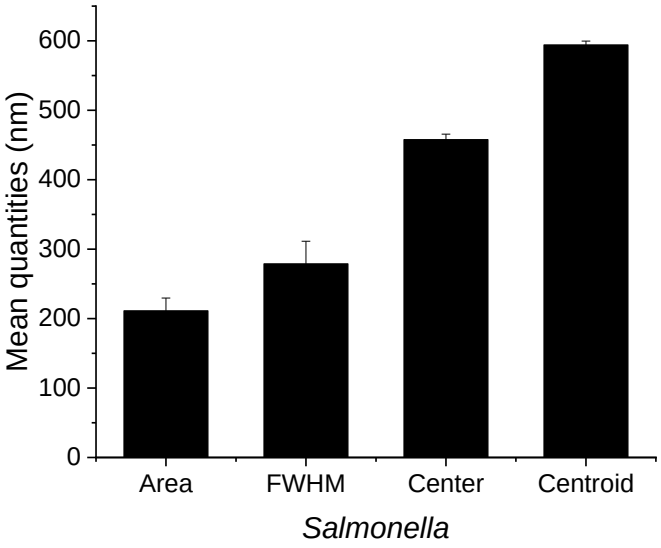


Figure VII 3. Area, FWHM, Center, and Centroid of the *Salmonella* 21 h of the mean of ten cells spectra (±SE).

The comparison of values of Area, FWHM, Center, and Centroid between the initial estimates of the parameters in Table VII 2 and Table VII 4 shows that values agree well.

VII.1.2 *E. coli* at 21 hours of growth

Using the Local maximum model, we analyze the *E. coli* 21 h spectra, similar to that of the *Salmonella* 21 h (Figure VII 4).

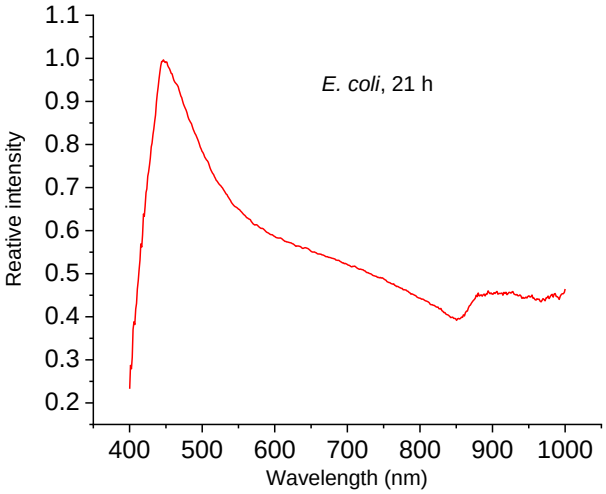


Figure VII 4. Relative intensity of the mean value of spectra for ten individual *E. coli* cells at 21 hours into growth as function of wavelength.

Using Batch Peak Analysis with Theme Parameters, we will calculate values of Area, FWHM, Center, and Centroid of the relative intensity for *E. coli* 21 h for each of 10 individual cells from the experimental spectra. The results are presented in Table VII 5.

Table VII 5. Values of Area, FWHM, Center, and Centroid of the relative intensity for *E. coli* 21 h for each of 10 individual cells.

Cells	Area, nm	FWHM, nm	Center, nm	Centroid, nm
1	177.74	161.70	446.50	588.57
2	164.88	148.27	445.30	585.70
3	150.70	120.57	444.10	576.56
4	152.10	122.49	444.10	580.25
5	169.90	145.39	444.10	587.92
6	180.46	159.05	446.50	589.58
7	183.52	221.49	451.30	586.27

8	159.72	151.61	451.30	578.69
9	145.03	151.13	446.50	559.81
10	173.86	205.95	450.10	585.35

Out of the statistics in columns on Table VII 5, we find mean values and standard errors for the Area, FWHM, Center, and Centroid for the group of 10 cells of *E. coli* 21 h, shown in Table VII 6.

Table VII 6. Area, FWHM, Center, and Centroid for the group of 10 cells of *E. coli* 21 h.

	N total	Mean	Standard Deviation	SE of mean
Area	10	165.79	13.49	4.27
FWHM	10	158.77	32.22	10.19
Center	10	446.98	2.90	0.92
Height	10	0.77	0.02	0.01
Centroid	10	581.87	8.91	2.82

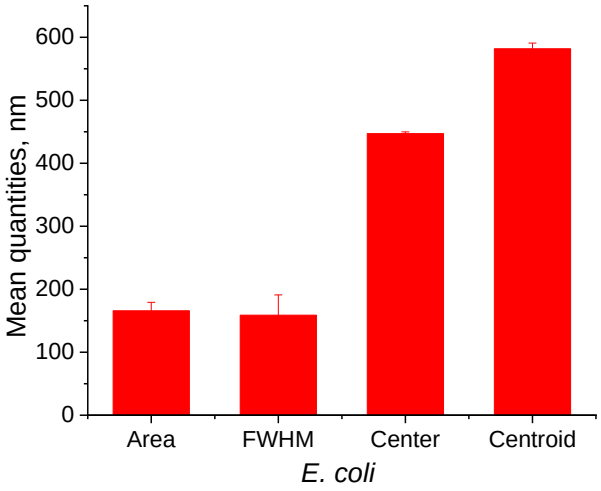


Figure VII 5. Area, FWHM, Center, and Centroid of the *E. coli* 21 h of the mean of ten cells spectra (±SE).

VII.1.3 *Listeria* at 21 hours of growth

The spectrum of *Listeria* is depicted in Figure VII 6.

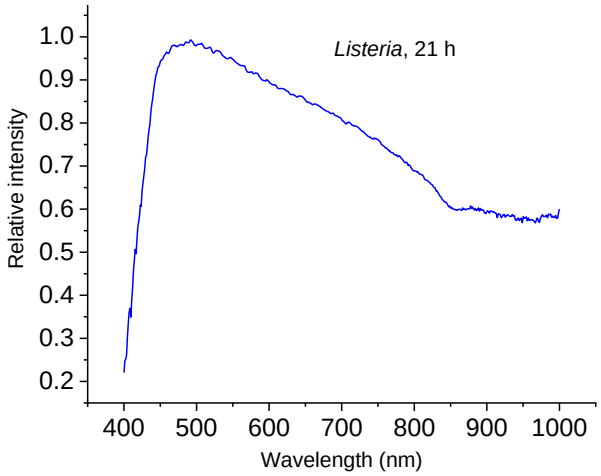


Figure VII 6. Relative intensity of the mean value of spectra for ten individual *Listeria* cells as function of wavelength.

Using Batch Peak Analysis with Theme Parameters, the software calculated the values of Area, FWHM, Center, and Centroid of the relative intensity for *Listeria* 21 h for each of 12 individual cells from the experimental spectra. The results are presented in Table VII 7.

Table VII 7. Values of Area, FWHM, Center, and Centroid of the relative intensity for *Listeria* 21 h for each of 12 individual cells

Cells	Area, nm	FWHM, nm	Center, nm	Centroid, nm
1	279.28	388.48	491.21	646.24
2	288.56	397.66	491.21	652.96
3	325.18	419.83	485.13	653.78
4	303.54	382.22	491.21	643.06
5	317.91	529.55	465.76	673.22
6	331.59	429.16	492.43	654.46
7	284.10	405.14	499.75	652.08
8	301.02	409.41	492.43	653.27
9	347.89	521.59	485.13	668.93
10	361.91	522.45	465.76	672.24
11	294.15	424.97	465.76	615.05
12	329.43	423.54	485.13	664.87

Out the statistics on columns of Table VII 7, we find mean values and standard errors for the Area, FWHM, Center, and Centroid for the group of 12 cells of *Listeria* 21 h, shown in Table VII 8.

Table VII 8. Area, FWHM, Center, and Centroid for the group of 12 cells of *Listeria* 21 h.

	N total	Mean	Standard Deviation	SE of mean
Area	12	313.72	26.20	7.56
FWHM	12	437.83	54.25	15.66
Center	12	484.24	11.86	3.42
Centroid	12	654.18	15.80	4.56

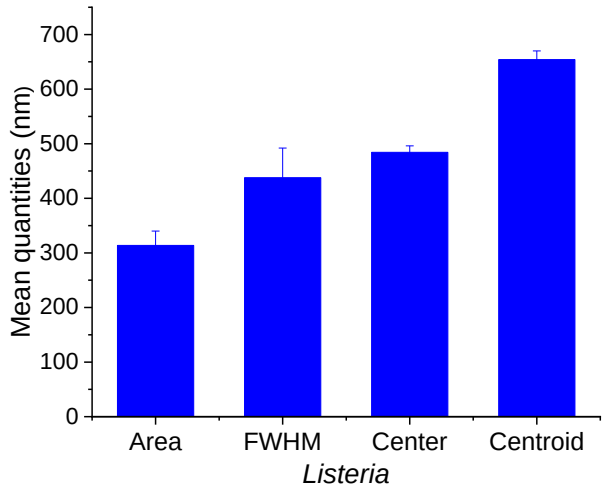


Figure VII 7. Area, FWHM, Center, and Centroid of *Listeria* at 21 hours of the mean of ten cells spectra (±SE).

VII.1.4 Comparison of *Salmonella, E. coli* and *Listeria monocytogenes* spectra at 21 hours of growth

Representative spectra of *Salmonella, E. coli* and *Listeria* at 21 h are shown in Figure VII 8.

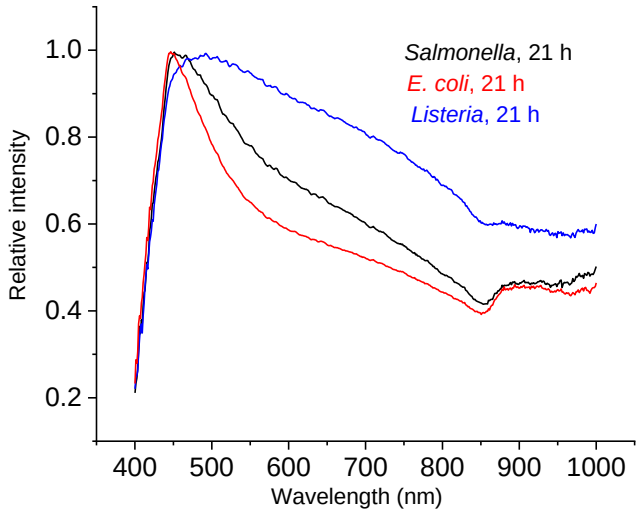


Figure VII 8. *S. enterica Typhimurium, E. coli* and *L. monocytogenes* spectra at 21 hours

The spectra of all three bacteria in a range between 400 and 1000 nm constitute broad peaks with sharp rise shoulders, single maxima around 450-500 nm, and gradual decays. *E. coli* presents the sharpest peak, and *Listeria* shows the most gradual decay. It is clear from the peak images that positions of the peak maxima (centers) cannot reliably serve to discriminate bacteria However, areas under peaks and width of the peaks seem quite different.

To compare *Salmonella, E. coli* and *Listeria* spectra at 21 h we will compare parameters of the spectra.

VII.1.4.1 Area

Taking values of Area of three bacteria from Tables VII 4-8, respectively, we can present them in Table VII 9.

Table VII 9. Area of three bacteria at 21 hours

Groups	Bacteria	Area, nm	SE
Group 1	*Salmonella*	211.18	5.81
Group 2	*E. coli*	165.79	4.27
Group 3	*Listeria*	313.72	7.56

Area of *Salmonella, E. coli* and *Listeria* spectra at 21 h is shown in Figure VII 9.

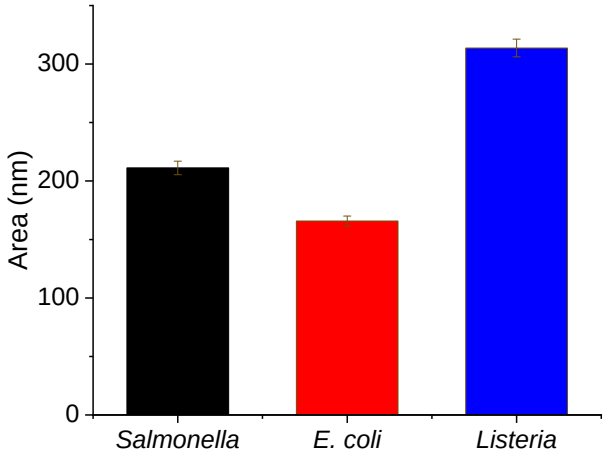


Figure VII 9. Area under spectral peaks of *Salmonella*, *E. coli* and *Listeria* 21 h (±SE).

A one-way ANOVA was conducted for comparison of areas under spectral peaks of *Salmonella*, *E. coli* and *Listeria* 21 h. The data from Table VII 9 were entered into the window of the computer program "Analysis of Variance from Summary Data" [49] and calculations were carried out as shown in Figure VII 10.

Enter your summary data here...

Group Name	N (count)	Mean	Std. Error
Group 1	10	211.18365	5.80658
Group 2	10	165.79108	4.26623
Group 3	12	313.71461	7.56248
Group 4			
Group 5			
Group 6			
Group 7			
Group 8			
Group 9			
Group 10			

Desired confidence level for post-hoc confidence intervals: 95

Compute

ANOVA Table...

Source of Variation	Sum of Squares	d.f.	Variance	F	p
Between Groups:	127916.3987	2	63958.1994	151.7611	0.0000
Within Groups:	12221.7638	29	421.4401		
Total:	140138.1625	31			

Post-hoc tests...

```
Tukey HSD Post-hoc Test...
Group 1 vs Group 2: Diff=-45.3926, 95%CI=-68.0661 to -22.7191, p=0.0001
Group 1 vs Group 3: Diff=102.5310, 95%CI=80.8228 to 124.2392, p=0.0000
Group 2 vs Group 3: Diff=147.9235, 95%CI=126.2153 to 169.6317, p=0.0000
```

Figure VII 10. Analysis of Variance from Summary Data for comparison of parameters Area under spectral peaks of *Salmonella, E. coli* and *Listeria* 21 h.

A one-way ANOVA was conducted to compare areas under hyperspectral peaks for *Salmonella, E. coli* and *Listeria*. There was significant effect of species on the area at the $p<0.05$ level for three bacteria [$F(2, 29) = 151.7$, $p=0.00001$]. Post hoc comparisons using the Tukey HSD test indicated that the mean areas under spectral peaks of *Salmonella, E. coli* and *Listeria* at 21 hours were significantly different.

VII.1.4.2 FWHM

Taking values of FWHM of three bacteria from Tables VII 4-8, respectively, we can present them in Table VII 10.

Table VII 10. FWHM of three bacteria at 21 h.

Groups	Bacteria	FWHM, nm	SE
Group 1	*Salmonella*	278.83	10.30
Group 2	*E. coli*	158.77	10.19
Group 3	*Listeria*	437.83	15.66

FWHM of *Salmonella, E. coli* and *Listeria* spectra at 21 hours is shown in Figure VII 11.

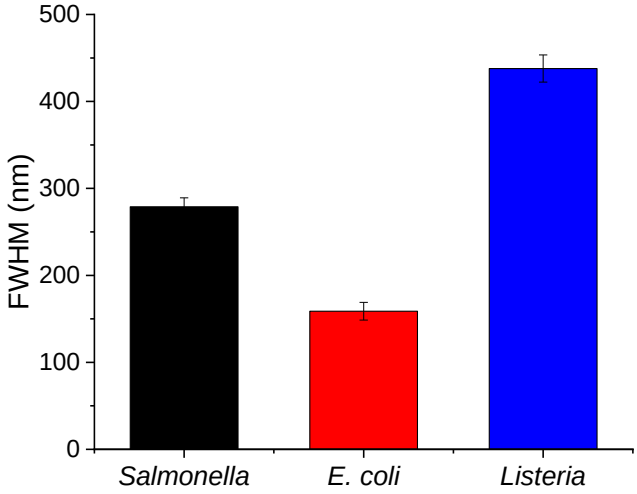


Figure VII 11. FWHM of spectral peaks of *Salmonella, E. coli* and *Listeria* 21 hours (±SE).

A one-way ANOVA was conducted for comparison of FWHM of spectral peaks of *Salmonella*, *E. coli* and *Listeria* 21 h. The data from Table VII 10 were entered into the window of the computer program "Analysis of Variance from Summary Data" [49] and calculations were carried out as shown in Figure VII 12.

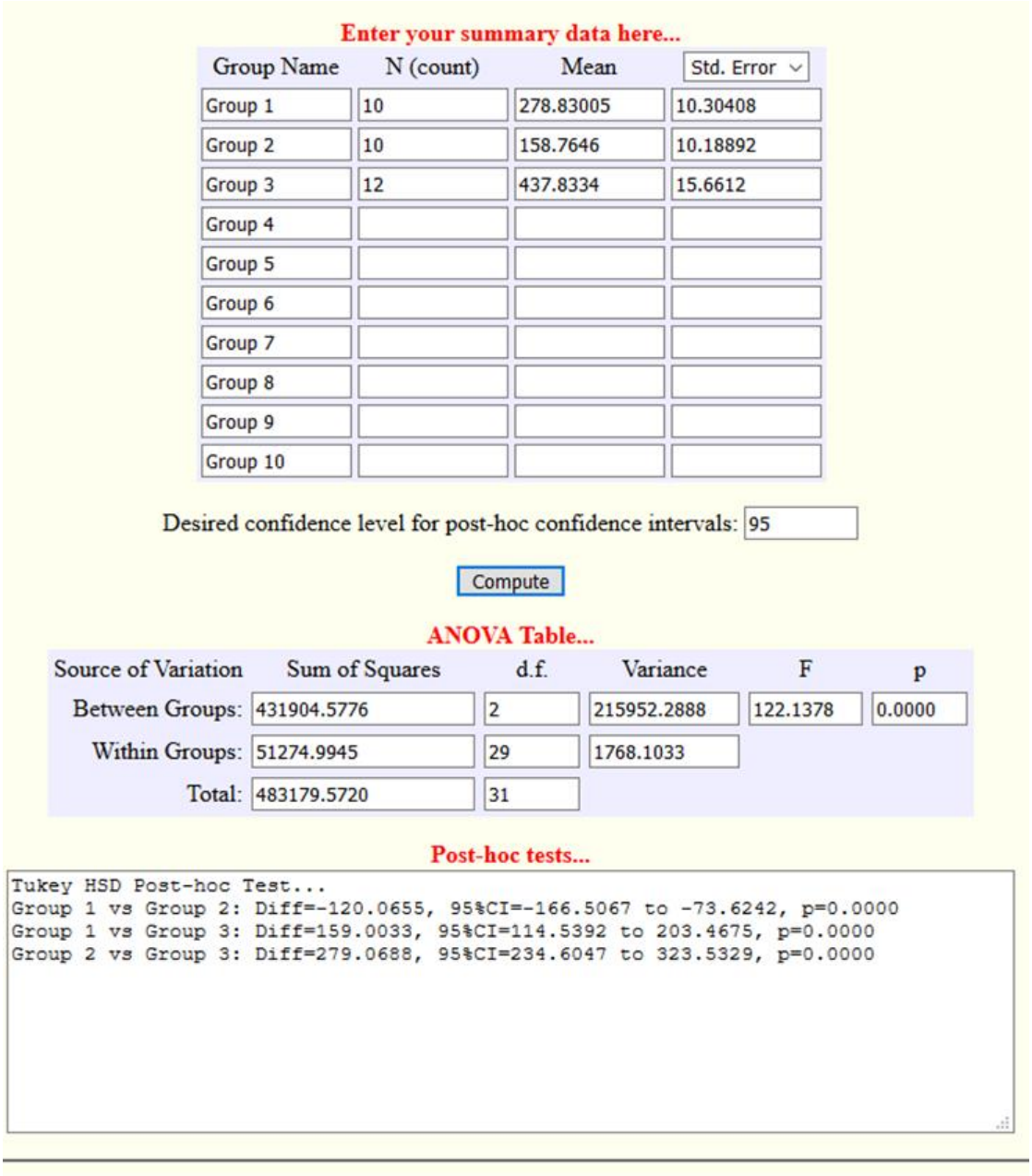


Figure VII 12. Analysis of Variance from Summary Data for comparison of parameters FWHM of peaks of *Salmonella*, *E. coli* and *Listeria* at 21 hours.

A one-way ANOVA was conducted to compare areas under hyperspectral FWHM of peaks for *Salmonella, E. coli* and *Listeria*. There was significant effect of species on the FWHM at the $p<0.05$ level for three bacteria [$F(2, 29) = 122.1$, $p=0.00001$]. Post hoc comparisons using the Tukey HSD test indicated that the mean FWHM for all three bacteria were significantly different.

VII.1.4.3 Center

Taking values of Center of three bacteria from Tables VII 4-8, respectively, we can present them in Table VII 11.

Table VII 11. Center of three bacteria at 21 h.

Groups	Bacteria	Center, nm	SE
Group 1	*Salmonella*	457.81	2.44
Group 2	*E. coli*	446.98	0.92
Group 3	*Listeria*	484.24	3.42

Centers of *Salmonella, E. coli* and *Listeria* spectra at 21 h are shown in Figure VII 13.

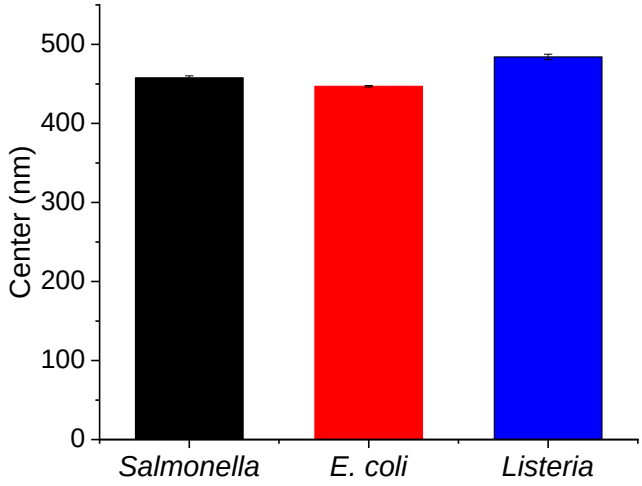
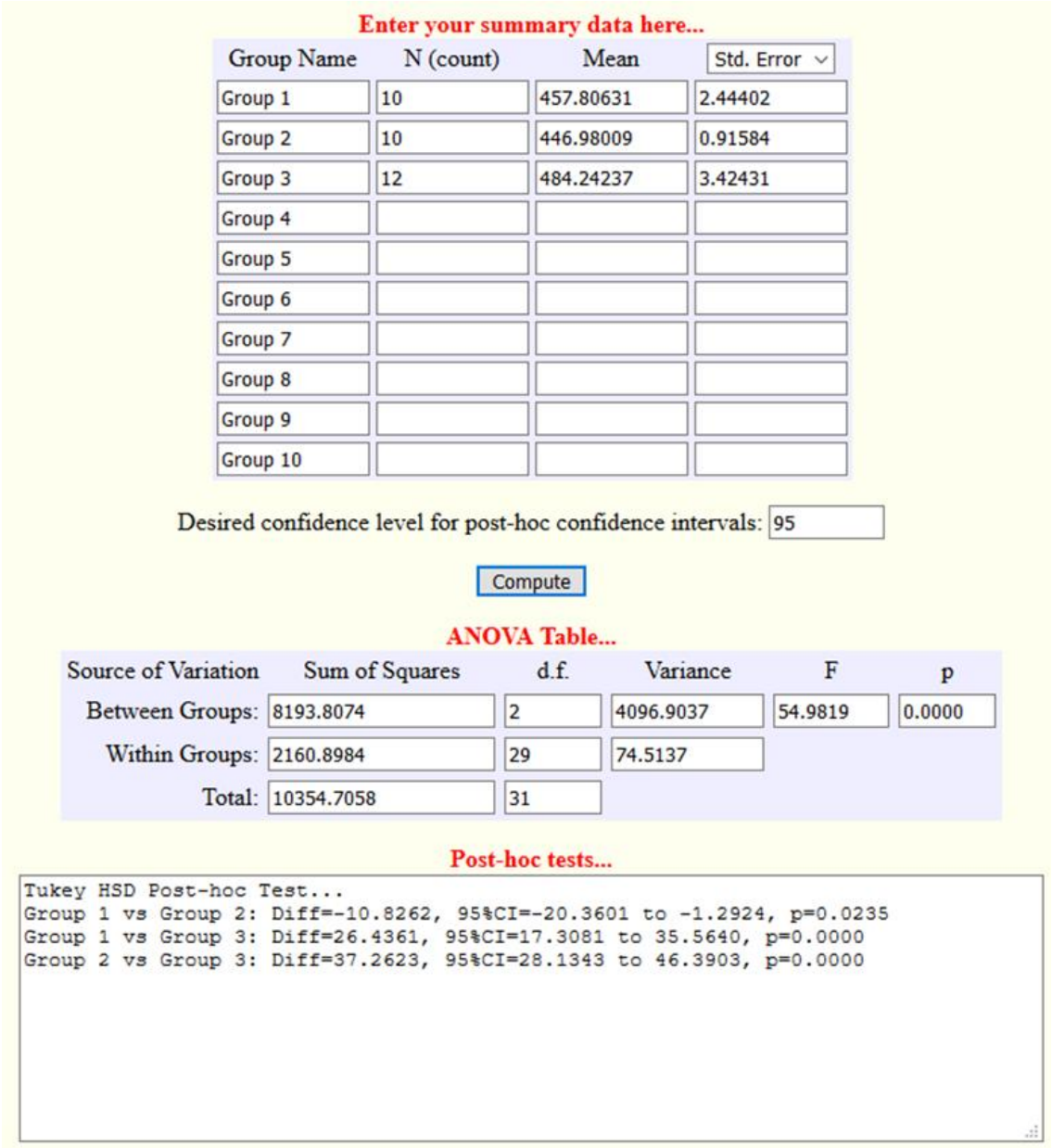


Figure VII 13. Centers of *Salmonella, E. coli* and *Listeria* spectra at 21 hours (±SE).

A one-way ANOVA was conducted for comparison of Centers of spectral peaks of *Salmonella, E. coli* and *Listeria* 21 h. The data from Table VII 11 were entered into the window of the computer program "Analysis of Variance from Summary Data" [49] and calculations were carried out as shown in Figure VII 14.



Enter your summary data here...

Group Name	N (count)	Mean	Std. Error
Group 1	10	457.80631	2.44402
Group 2	10	446.98009	0.91584
Group 3	12	484.24237	3.42431
Group 4			
Group 5			
Group 6			
Group 7			
Group 8			
Group 9			
Group 10			

Desired confidence level for post-hoc confidence intervals: 95

Compute

ANOVA Table...

Source of Variation	Sum of Squares	d.f.	Variance	F	p
Between Groups:	8193.8074	2	4096.9037	54.9819	0.0000
Within Groups:	2160.8984	29	74.5137		
Total:	10354.7058	31			

Post-hoc tests...

```
Tukey HSD Post-hoc Test...
Group 1 vs Group 2: Diff=-10.8262, 95%CI=-20.3601 to -1.2924, p=0.0235
Group 1 vs Group 3: Diff=26.4361, 95%CI=17.3081 to 35.5640, p=0.0000
Group 2 vs Group 3: Diff=37.2623, 95%CI=28.1343 to 46.3903, p=0.0000
```

Figure VII 14. Analysis of Variance from Summary Data for comparison of parameters Center of spectral peaks of *Salmonella, E. coli* and *Listeria* at 21 hours.

A one-way ANOVA was conducted to compare Centers of hyperspectral peaks for *Salmonella, E. coli* and *Listeria*. There was significant effect of species on the Centers at the $p<0.05$ level for

three bacteria [F (2, 29) =55.0, p=0.00001]. Post hoc comparisons using the Tukey HSD test indicated that the mean Centers for all three bacteria were significantly different.

VII.1.4.4 Centroid

Taking values of Centroid of three bacteria from Tables VII 4-8, respectively, we can present them in Table VII 12.

Table VII 12. Centroid of three bacteria at 21 h.

Groups	Bacteria	Centroid, nm	SE
Group 1	*Salmonella*	594.08	1.78
Group 2	*E. coli*	581.87	2.82
Group 3	*Listeria*	654.18	4.56

Centroids of *Salmonella*, *E. coli* and *Listeria* spectra at 21 h are shown in Figure VII 15.

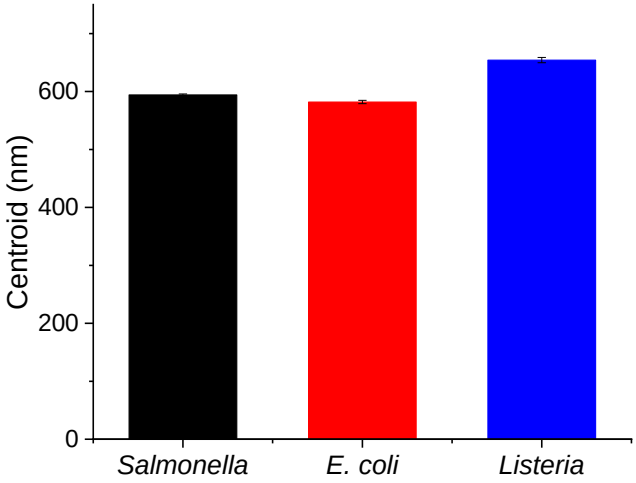
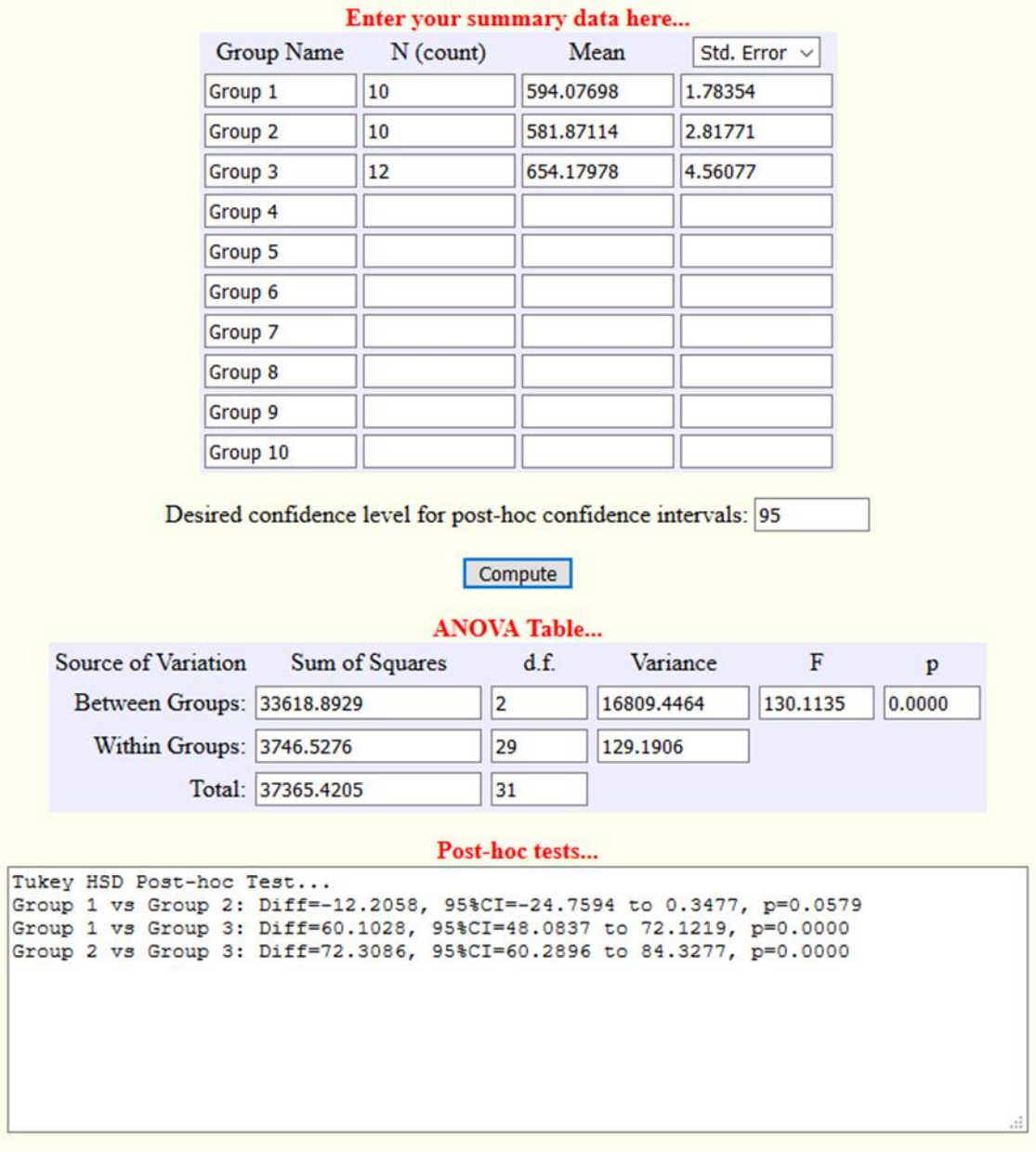


Figure VII 15. Centroids of spectral peaks of *Salmonella*, *E. coli* and *Listeria* 21 h (±SE).

A one-way ANOVA was conducted for comparison of Centroids of spectral peaks of *Salmonella*, E. coli and *Listeria* 21 h. The data from Table VII 9 were entered into the window of the computer program "Analysis of Variance from Summary Data" [49] and calculations were carried out as shown in Figure VII 16.



Enter your summary data here...

Group Name	N (count)	Mean	Std. Error
Group 1	10	594.07698	1.78354
Group 2	10	581.87114	2.81771
Group 3	12	654.17978	4.56077
Group 4			
Group 5			
Group 6			
Group 7			
Group 8			
Group 9			
Group 10			

Desired confidence level for post-hoc confidence intervals: 95

Compute

ANOVA Table...

Source of Variation	Sum of Squares	d.f.	Variance	F	p
Between Groups:	33618.8929	2	16809.4464	130.1135	0.0000
Within Groups:	3746.5276	29	129.1906		
Total:	37365.4205	31			

Post-hoc tests...

```
Tukey HSD Post-hoc Test...
Group 1 vs Group 2: Diff=-12.2058, 95%CI=-24.7594 to 0.3477, p=0.0579
Group 1 vs Group 3: Diff=60.1028, 95%CI=48.0837 to 72.1219, p=0.0000
Group 2 vs Group 3: Diff=72.3086, 95%CI=60.2896 to 84.3277, p=0.0000
```

Figure VII 16. Analysis of Variance from Summary Data for comparison of parameters of Centroids of spectral peaks of *Salmonella*, *E. coli* and *Listeria* 21 h.

A one-way ANOVA was conducted to compare Centroids of hyperspectral peaks for *Salmonella*, *E. coli* and *Listeria*. There was significant effect of species on the Centroids at the $p<0.05$ level for three bacteria [$F(2, 29) = 130.1$, $p=0.00001$]. Post hoc comparisons using the Tukey HSD test indicated that the mean Centroids for *Salmonella* vs *Listeria* and *E. coli* vs *Listeria* were

significantly different ($p<0.00001$). However, the mean Centroids for *Salmonella* vs *E. coli* were not significantly different ($p=0.058$).

VII.1.4.5 Summary of spectra comparison for three bacteria at 21 hours of growth

Results of the comparison of three bacteria at 21 h are shown in Table VII 13.

Table VII 13. Differentiation of *Salmonella, E. coli* and *Listeria* 21 h by ANOVA

Comparison	**Area**	**ANOVA**	**FWHM**	**ANOVA**	**Center**	**ANOVA**	**Centroid**	**ANOVA**
Salmonella vs *E. coli*	+	[F(2, 29)=151.7, p=0.00001]	+	[F(2, 29)=122.1, p=0.00001]	+	[F(2, 29)=55.0, p=0.00001]	-	[F(2, 29)=130.1, p=0.00001]
Salmonella vs *Listeria*	+		+		+		+	
E. coli vs *Listeria*	+		+		+		+	

+ significantly different; - not significantly different

VII.1.5 *Salmonella* at 18 hours of growth

The mean value of the *Salmonella* 18 h relative spectrum intensity as function of wavelength is shown in Figure VII 17.

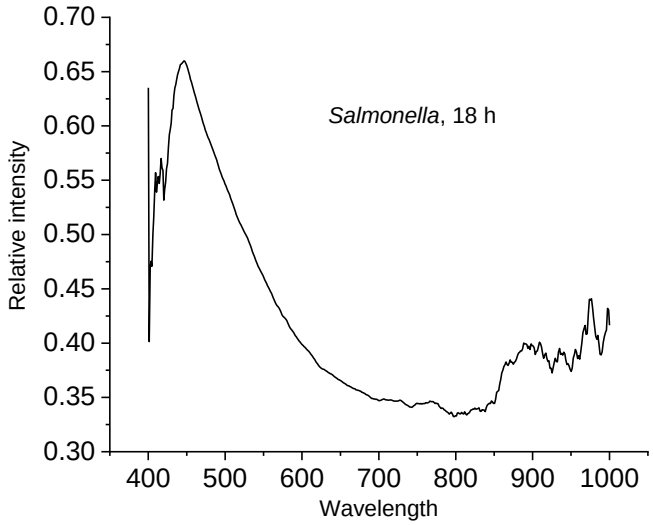


Figure VII 17. Relative intensity of the mean value of spectra for ten individual *Salmonella* cells as function of wavelength

Similar to calculation of *Salmonella* at 21 h, we calculate mean values and standard errors for the Area, FWHM, Center, and Centroid for the group of 7 cells of *Salmonella* 18 h, shown in Table VII 14.

Table VII 14. Area, FWHM, Center, and Centroid for the group of 7 cells of *Salmonella* 18 h

	N total	Mean	Standard Deviation	SE of mean
Area	7	85.60	29.23	13.07
FWHM	7	146.00	26.09	11.67
Center	7	446.26	2.74	1.22
Centroid	7	564.31	51.60	23.07

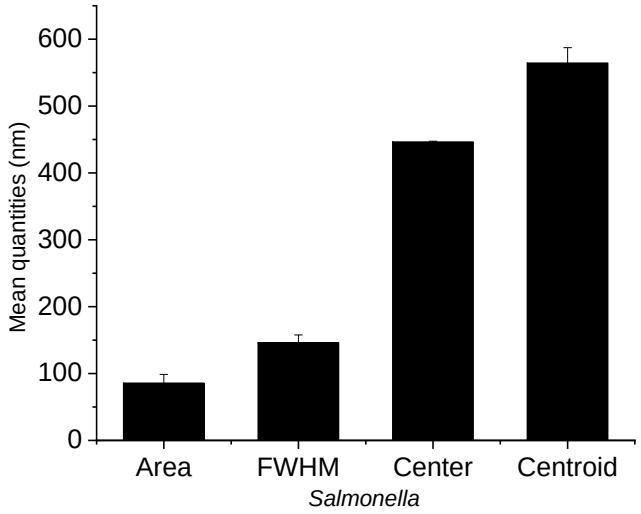


Figure VII 18. Area, FWHM, Center, and Centroid of the *Salmonella* 18 h of the mean of seven cells spectra (±SE).

VII.1.6 *E. coli* at 18 hours of growth

Using the Local maximum model, we analyze the *E. coli* 18 h spectra, similar to that of the *E. coli* 18 h (Figure VII 18).

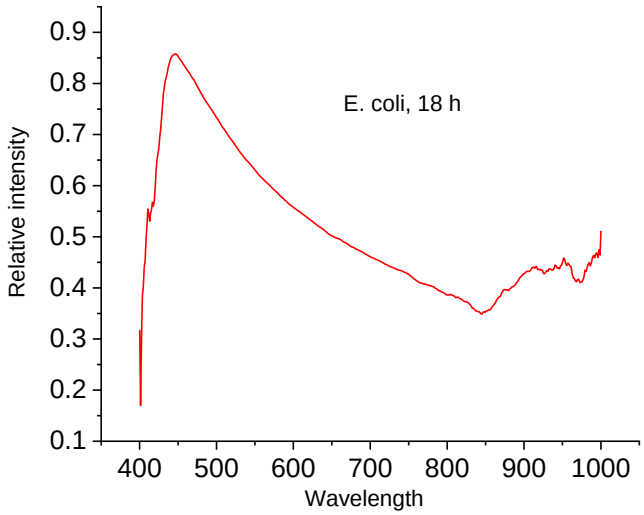


Figure VII 19. Relative intensity of the mean value of spectra for 7 individual *E. coli* cells as function of wavelength.

Similar to calculation of *E. coli* at 21 h, we calculate mean values and standard errors for the Area, FWHM, Center, and Centroid for the group of 7 cells of *E. coli* 18 h, shown in Table VII 18.

Table VII 15. Area, FWHM, Center, and Centroid for the group of 7 cells of *E. coli* at 18 h.

	N total	Mean	Standard Deviation	SE of mean
Area	7	235.75	88.55	33.47
FWHM	7	311.12	106.27	40.17
Center	7	445.81	2.28	0.86
Centroid	7	598.44	31.70	11.98

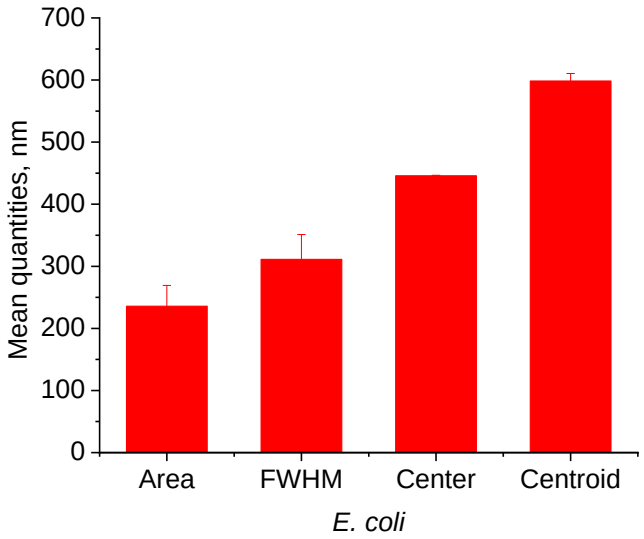


Figure VII 20. Area, FWHM, Center, and Centroid of the *E. coli* 18 h of the mean of 7 cells spectra (±SE).

VII.1.7 *Listeria at 18* hours of growth

The spectrum of *Listeria* 18 h is depicted in Figure VII 21.

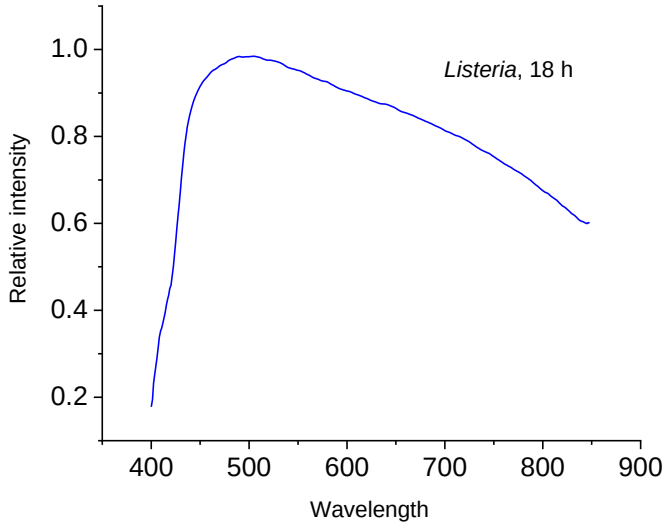


Figure VII 21. Relative intensity of the mean value of spectra for 7 individual *Listeria* cells as function of wavelength.

Using Batch Peak Analysis with Theme Parameters, we will calculate values of Area, FWHM, Center, and Centroid of the relative intensity for *Listeria* at 18 hours for each of 7 individual

cells from the experimental spectra.[21] Taking statistical averaging on the values of 7 cells, we find mean values and standard errors for the Area, FWHM, Center, and Centroid for the group of 7 cells of *Listeria* 18 h, shown in Table VII 16 and Figure VII 22.

Table VII 16. Area, FWHM, Center, and Centroid for the group of 7 cells of *Listeria* 18 h.

	N total	Mean	Standard Deviation	SE of mean
Area	7	299.97	37.45	14.16
FWHM	7	418.58	4.90	1.85
Center	7	500.66	17.83	6.74
Centroid	7	615.83	2.90	1.10

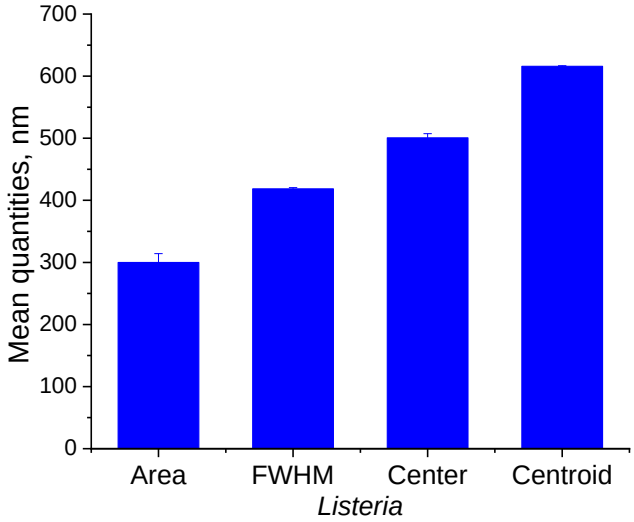


Figure VII 22. Area, FWHM, Center, and Centroid of the *Listeria* 18 h of the mean of 7 cells spectra (±SE).

VII.1.8 Comparison of *Salmonella, E. coli* and *Listeria* spectra at 18 hours of growth

Representative spectra of *Salmonella, E. coli* and *Listeria* at 18 h are shown in Figure VII 23.

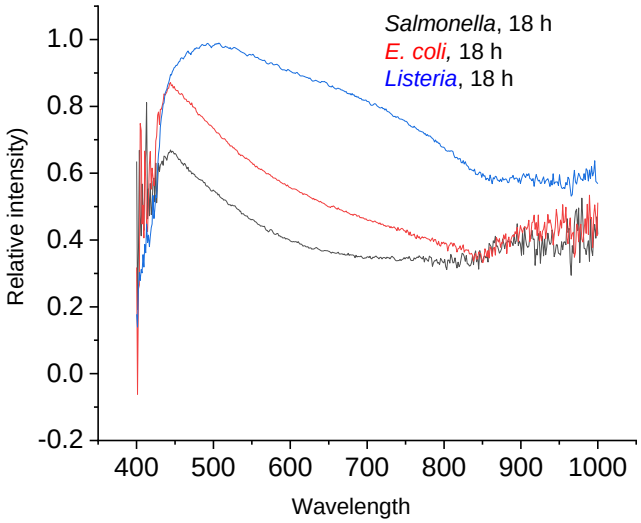


Figure VII 23. *Salmonella, E. coli* and *Listeria* spectra at 18 h

The spectra of all three bacteria in a range between 400 and 1000 nm constitute broad peaks with sharp rise shoulders, single maxima around 450-550 nm, and gradual decays. The spectra are quite noisy in the beginning and the end of curves. It is clear from the peak images that positions of the peak maxima (centers) cannot reliably serve to discriminate bacteria However, areas under peaks and width of the peaks seem quite different.

To compare *Salmonella, E. coli* and *Listeria* spectra at 18 h we will compare parameters of the spectra.

VII.1.8.1 Area

Taking values of Area of three bacteria from Tables VII 14-16, respectively, we can present them in Table VII 17.

Table VII 17. Area of three bacteria at 18 h

Groups	Bacteria	Area, nm	SE
Group 1	*Salmonella*	85.60	13.07
Group 2	*E. coli*	235.75	33.47
Group 3	*Listeria*	299.97	14.16

Area of *Salmonella, E. coli* and *Listeria* spectra at 18 h is shown in Figure VII 24.

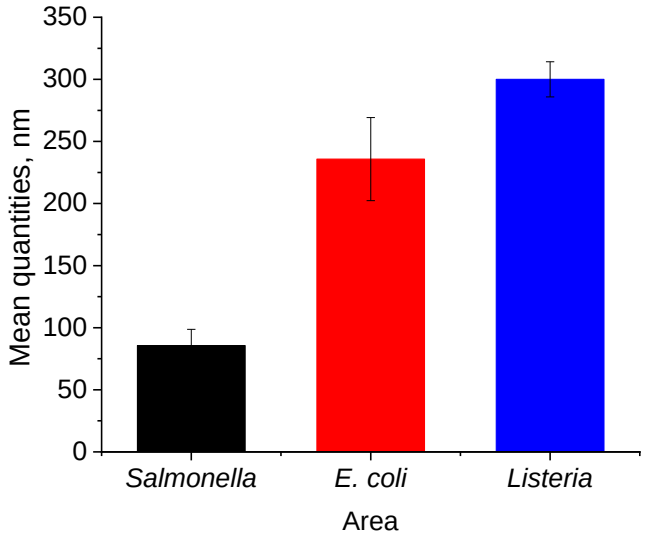


Figure VII 24. Area under spectral peaks of *Salmonella, E. coli* and *Listeria* 18 h (±SE).

A one-way ANOVA was conducted for comparison of areas under spectral peaks of *Salmonella* (Group 1), *E. coli* (Group 2), and *Listeria* (Group 3) 18 h. The data from Table VII 17 were entered into the window of the computer program "Analysis of Variance from Summary Data" [49] and calculations were carried out as shown in Table VII 18.

Table VII 18. ANOVA Area 18 h

Group Name	N (count)	Mean	Std. Error		
Group 1	7	85.60	13.07		
Group 2	7	235.75	33.47		
Group 3	7	299.97	14.16		
Source of Variation	Sum of Squares	d.f.	Variance	F	p
Between Groups:	389302.58	2	194651.29	55.94	0.00001
Within Groups:	62638.85	18	3479.94		
Total:	451941.43	20			

Tukey HSD Post-hoc Test...

Group 1 vs Group 2: Diff=150.1516, 95%CI=69.6767 to 230.6265, p=0.0004

Group 1 vs Group 3: Diff=332.9767, 95%CI=252.5017 to 413.4516, p=0.00001

Group 2 vs Group 3: Diff=182.8251, 95%CI=102.3501 to 263.3000, p=0.00001

A one-way ANOVA was conducted to compare areas under hyperspectral peaks for *Salmonella, E. coli* and *Listeria*. There was significant effect of species on the area at the $p<0.05$ level for three bacteria [$F(2, 18) = 55.93$, $p=0.0001$]. Post hoc comparisons using the Tukey HSD test indicated that the mean areas under spectral peaks of *Salmonella, E. coli* and *Listeria* 18 h were significantly different.

VII.1.8.2 FWHM

Taking values of FWHM of three bacteria from Tables VII 14-16, respectively, we can present them in Table VII 19.

Table VII 19. FWHM of three bacteria at 18 h.

Groups	Bacteria	FWHM, nm	SE
Group 1	*Salmonella*	146.00	11.67
Group 2	*E. coli*	311.12	40.17
Group 3	*Listeria*	418.58	1.85

FWHM of *Salmonella, E. coli* and *Listeria* spectra at 18 h is shown in Figure VII 25

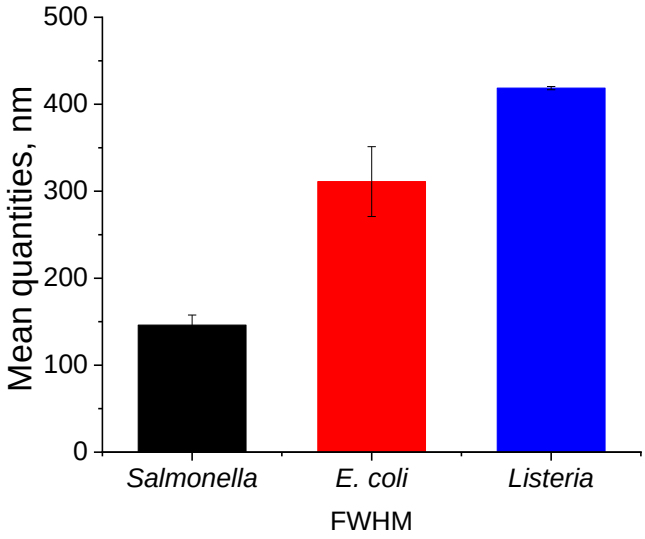


Figure VII 25. FWHM of spectral peaks of *Salmonella, E. coli* and *Listeria* 18 h (±SE).

A one-way ANOVA was conducted for comparison of FWHM of spectral peaks of *Salmonella* (Group 1), *E. coli* (Group 2), and *Listeria* (Group 3) 18 h. The data from Table VII 10 were

entered into the window of the computer program "Analysis of Variance from Summary Data" [49] and calculations were carried out as shown in Table VII 20.

Table VII 20. ANOVA FWHM 18 h

Group Name	N (count)	Mean	Std. Error
Group 1	7	146.00	11.67
Group 2	7	311.12	40.17
Group 3	7	418.58	1.85

Desired confidence level for post-hoc confidence intervals: 95%

Source of Variation	Sum of Squares	d.f.	Variance	F	p
Between Groups:	263919.16	2	131959.58	32.27	0.00001
Within Groups:	73614.23	18	4089.68		
Total:	337533.39	20			

Tukey HSD Post-hoc Test...

Group 1 vs Group 2: Diff=165.1125, 95%CI=77.8717 to 252.3533, p=0.0004 Group 1 vs Group 3: Diff=272.5760, 95%CI=185.3352 to 359.8168, p=0.00001 Group 2 vs Group 3: Diff=107.4635, 95%CI=20.2227 to 194.7042, p=0.0148

A one-way ANOVA was conducted to compare areas under hyperspectral FWHM of peaks for *Salmonella, E. coli* and *Listeria*. There was significant effect of species on the FWHM at the $p<0.05$ level for three bacteria [$F(2, 18) = 122.1$, $p=0.00001$]. Post hoc comparisons using the Tukey HSD test indicated that the mean FWHM for all three bacteria were significantly different.

VII.1.8.3 Center

Taking values of Center of three bacteria from Tables VII 14-16, respectively, we can present them in Table VII 21.

Table VII 21. Centers of three bacteria at 18 h

Groups	Bacteria	Center, nm	SE
Group 1	*Salmonella*	446.26	1.22

Group 2	*E. coli*	445.81	0.86
Group 3	*Listeria*	500.66	6.74

Centers of *Salmonella, E. coli* and *Listeria* spectra at 18 h are shown in Figure VII 26.

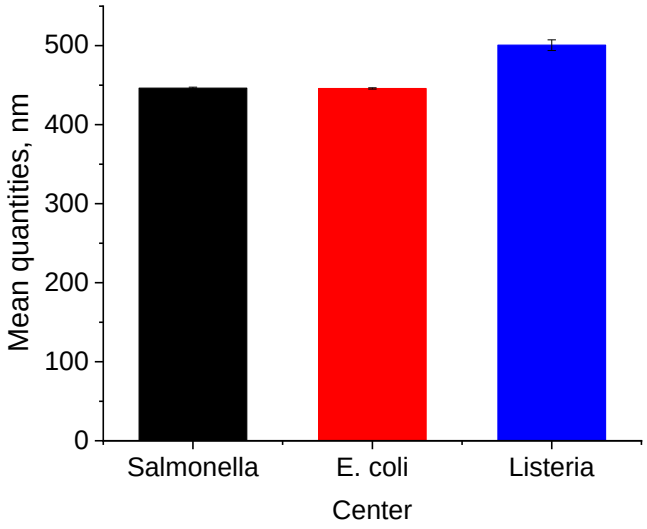


Figure VII 26. Centers of *Salmonella, E. coli* and *Listeria* spectra at 18 h (±SE).

A one-way ANOVA was conducted for comparison of Centers of spectral peaks of *Salmonella* (Group 1), *E. coli* (Group 20, and *Listeria* (Group 3), 18 h. The data from Table VII 21 were entered into the window of the computer program "Analysis of Variance from Summary Data" [49] and calculations were carried out as shown in Table VII 22.

Table VII 22. ANOVA Center 18 h.

Group Name	N (count)	Mean	Std. Error
Group 1	7	446.26	1.22
Group 2	7	445.81	0.86
Group 3	7	500.66	6.74

Desired confidence level for post-hoc confidence intervals: 95%

Source of Variation	Sum of Squares	d.f.	Variance	F	p
Between Groups:	13924.22	2	6962.11	62.61	0.00001
Within Groups:	2001.54	18	111.20		
Total:	15925.76	20			

Tukey HSD Post-hoc Test...

Group 1 vs Group 2: Diff=-0.4458, 95%CI=-14.8312 to 13.9395, p=0.9966

Group 1 vs Group 3: Diff=54.3995, 95%CI=40.0142 to 68.7849, p=0.0000

Group 2 vs Group 3: Diff=54.8454, 95%CI=40.4600 to 69.2307, p=0.0000

A one-way ANOVA was conducted to compare Centers of hyperspectral peaks for *Salmonella*, *E. coli* and *Listeria*, 18 h. There was significant effect of species on the Centers at the $p<0.05$ level for three bacteria [$F(2, 18) = 62.6$, $p=0.00001$]. Post hoc comparisons using the Tukey HSD test indicated that the mean Centers for *Salmonella* vs *E. coli* were not significantly different. Centers for *Salmonella* vs *E. coli* and *E. coli* vs *Listeria* were significantly different.

VII.1.8.4 Centroid

Taking values of Centroid of three bacteria from Tables VII 14-16, respectively, we can present them in Table VII 23.

Table VII 23. Centroid of three bacteria at 18 h.

Groups	Bacteria	Centroid, nm	SE
Group 1	*Salmonella*	564.31	23.07
Group 2	*E. coli*	598.44	11.98
Group 3	*Listeria*	615.83	1.10

Centroids of *Salmonella*, *E. coli* and *Listeria* spectra at18 h are shown in Figure VII 27.

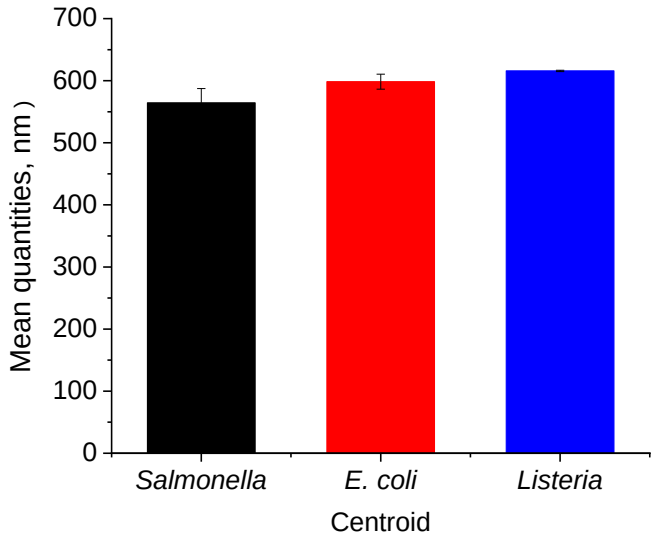


Figure VII 27. Centroids of spectral peaks of *Salmonella*, *E. coli* and *Listeria* 18 h (±SE).

A one-way ANOVA was conducted for comparison of Centroids of spectral peaks of *Salmonella* Group 1), *E. coli* (Group 2), and *Listeria* (Group 3) 18 h. The data from Table VII 23 were entered into the window of the computer program "Analysis of Variance from Summary Data" [49] and calculations were carried out as shown in Table VII 24.

Table VII 24. ANOVA Centroid 18 h

Group Name	N (count)	Mean	Std. Error
Group 1	7	564.31	23.07
Group 2	7	598.44	11.98
Group 3	7	615.83	1.10

Desired confidence level for post-hoc confidence intervals: 95%

Source of Variation	Sum of Squares	d.f.	Variance	F	p
Between Groups:	9617.47	2	4808.73	3.04	0.0727
Within Groups:	28438.56	18	1579.92		
Total:	38056.03	20			

Tukey HSD Post-hoc Test...

Group 1 vs Group 2: Diff=34.1318, 95%CI=-20.0923 to 88.3559, p=0.2685

Group 1 vs Group 3: Diff=51.5210, 95%CI=-2.7032 to 105.7451, p=0.0642

Group 2 vs Group 3: Diff=17.3892, 95%CI=-36.8349 to 71.6133, p=0.6967

A one-way ANOVA was conducted to compare Centroids of hyperspectral peaks for *Salmonella, E. coli* and *Listeria*. There was no significant effect of species on the Centroids at the $p<0.05$ level for three bacteria [$F(2, 18) = 3.04$, $p=0.0727$]. Post hoc comparisons using the Tukey HSD test indicated that the mean Centroid for all three bacteria were not significantly different ($p=0.058$).

VII.1.8.5 Summary of spectra comparison for three bacteria at 18 h

Taking together results on the comparison of three bacteria at18 h are shown in Table VII 25.

Table VII 25. Differentiation of *Salmonella*, *E. coli* and *Listeria*18 h by ANOVA

Comparison	Area	ANOVA	FWHM	ANOVA	Center	ANOVA	Centroid	ANOVA
Salmonella vs *E. coli*	+	[F(2,18)=55.93 , p=0.0001]	+	[F(2, 18)=122.1, p=0.00001]	-	[F(2, 18)=62.6, p=0.00001]	-	[F(2, 29)=130.1, p=0.00001]
Salmonella vs *Listeria*	+		+		+		-	
E. coli vs *Listeria*	+		+		+		-	

+ significantly different; - not significantly different

VII.1.9 *Salmonella* at 24 hours of growth

The mean value of the *Salmonella* 24 h relative spectrum intensity as function of wavelength is shown in Figure VII 28.

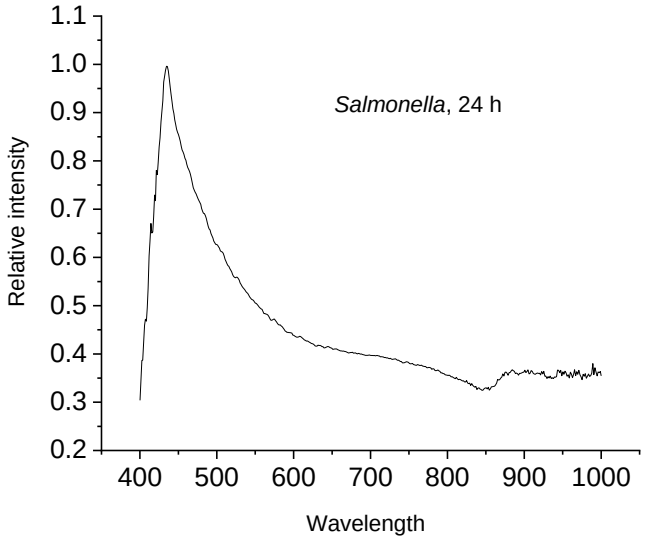


Figure VII 28. Relative intensity of the mean value of spectra for ten individual *Salmonella* cells (24 h) as function of wavelength.

Similar to calculation of *Salmonella* 21 h, we calculate mean values and standard errors for the Area, FWHM, Center, and Centroid for the group of 10 cells of *Salmonella* 24 h, shown in Table VII 26 and Figure VII 29.

Table VII 26. Area, FWHM, Center, and Centroid for the group of 10 cells of *Salmonella* 24 h

	N total	Mean	Standard Deviation	SE of mean
Area	10	106.39	26.22	8.29
FWHM	10	83.49	14.86	4.70

Center	10	434.88	0.98	0.31
Centroid	10	561.79	31.85	10.07

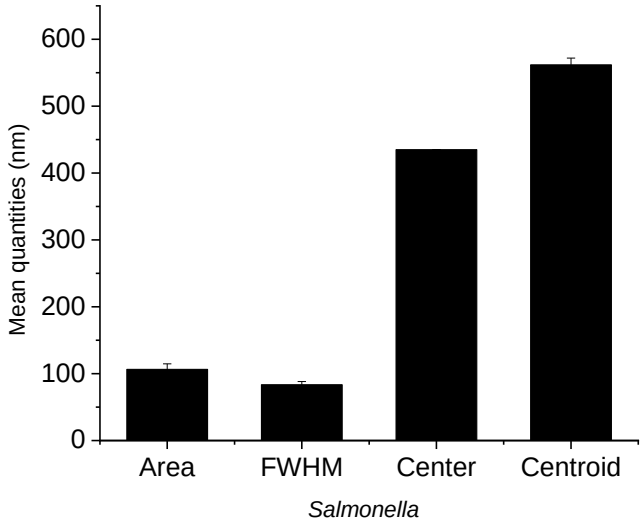


Figure VII 29. Area, FWHM, Center, and Centroid of the *Salmonella* 24 h of the mean of 10 cells spectra (±SE).

VII.1. 10 *E. coli* at 24 hours of growth

Using the Local maximum model, we analyze the *E. coli* 18 h spectra, similar to that of the *E. coli* 24 h (Figure VII 30).

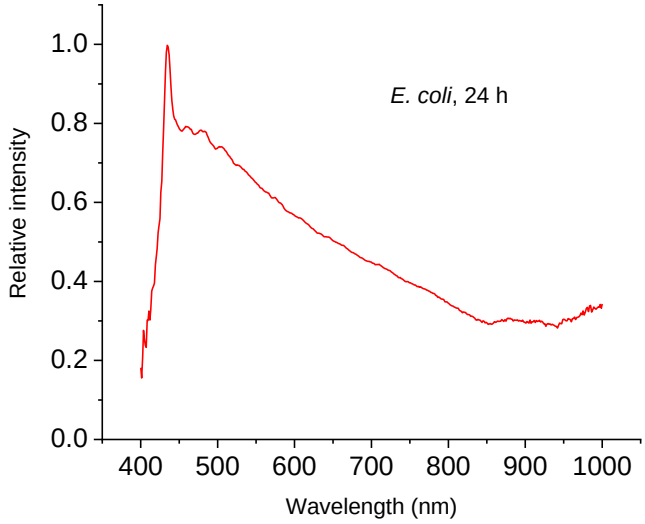


Figure VII 30. Relative intensity of the mean value of spectra for 10 individual *E. coli* cells as function of wavelength.

Similar to calculation of *E. coli* 18 h, we calculate mean values and standard errors for the Area, FWHM, Center, and Centroid for the group of 10 cells of *E. coli* 24 h, shown in Table VII 27 and Figure VII 31.

Table VII 27. Area, FWHM, Center, and Centroid for the group of 10 cells of *E. coli* 24 h.

	N total	Mean	Standard Deviation	SE of mean
Area	10	192.32	25.06	7.92
FWHM	10	182.88	21.33	6.75
Center	10	444.40	20.52	6.49
Centroid	10	613.88	9.81	3.10

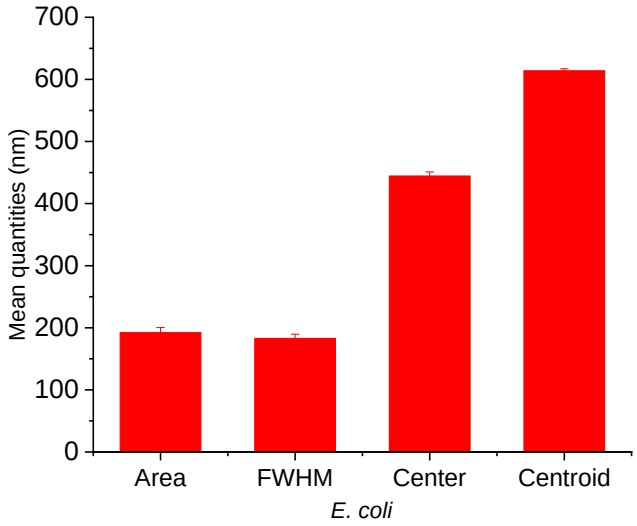


Figure VII 31. Area, FWHM, Center, and Centroid of the *E. coli* 24 h of the mean of 10 cells spectra (±SE).

VII.1.11 *Listeria* at 24 hours of growth

The spectrum of *Listeria* 24 h is depicted in Figure VII 32.

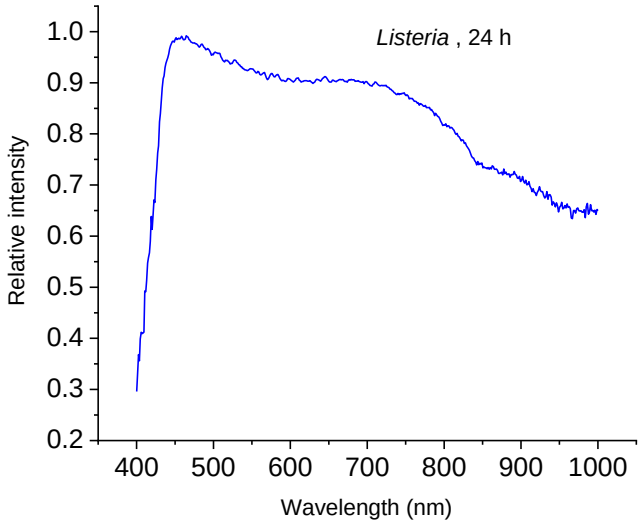


Figure VII 32. Relative intensity of the mean value of spectra for 12 individual *Listeria* cells as function of wavelength.

Using Batch Peak Analysis with Theme Parameters, we will calculate values of Area, FWHM, Center, and Centroid of the relative intensity for *Listeria* 24 h for each of 12 individual cells from the experimental spectra. Taking statistical averaging on the values of 12 cells, we find mean values and standard errors for the Area, FWHM, Center, and Centroid for the group of 12 cells of *Listeria* 18 h, shown in Table VII 28 and Figure VII 33.

Table VII 28. Area, FWHM, Center, and Centroid for the group of 12 cells of *Listeria* 24 h.

	N total	Mean	Standard Deviation	SE of mean
Area	12	195.10	53.32	15.39
FWHM	12	314.33	61.21	17.67
Center	12	455.31	5.51	1.59
Centroid	12	591.38	41.3	11.94

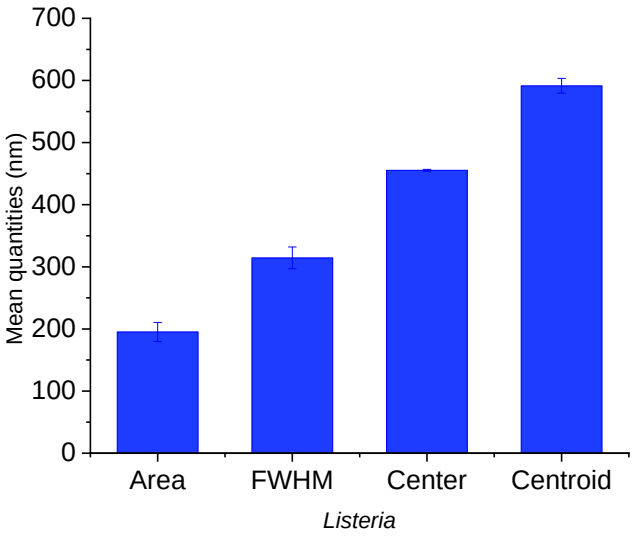


Figure VII 33. Comparison of *Salmonella, E. coli* and *Listeria* spectra at 24 hours of growth

Representative spectra of *Salmonella, E. coli* and *Listeria* at 24h are shown in Figure VII 34.

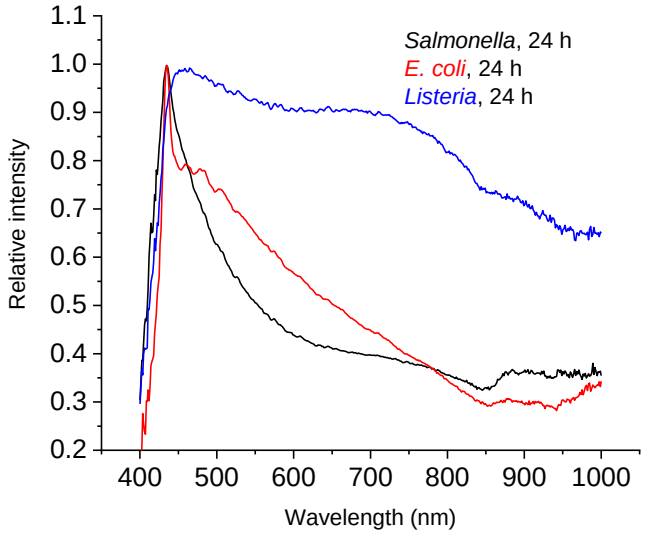


Figure VII 34. *Salmonella, E. coli* and *Listeria* spectra at 24 h.

The spectra of all three bacteria in a range between 400 and 1000 nm constitute relatively broad peaks with sharp rise shoulders, single maxima around 450-550 nm, and gradual decays. The *Listeria* spectrum is broadly bent up in the range of 650 – 850 nm. The spectra are quite noisy in the end of curves. It is clear from the peak images that positions of the peak maxima (centers)

cannot reliably serve to discriminate bacteria However, areas under peaks and width of the peaks seem quite different.

To compare *Salmonella, E. coli* and *Listeria* spectra at 24 h we will compare different parameters of the spectra.

VII.1.11.1 Area

Taking values of Area of three bacteria from Tables VII 26-28, respectively, we can present them in Table VII 29.

Table VII 29. Area of three bacteria at 24 h

Groups	N total	Bacteria	Area, nm	SE
Group 1	10	*Salmonella*	106.39	8.29
Group 2	10	*E. coli*	192.32	7.92
Group 3	12	*Listeria*	195.10	15.39

The Area of *Salmonella, E. coli* and *Listeria* spectra at 24 h is shown in Figure VII 35.

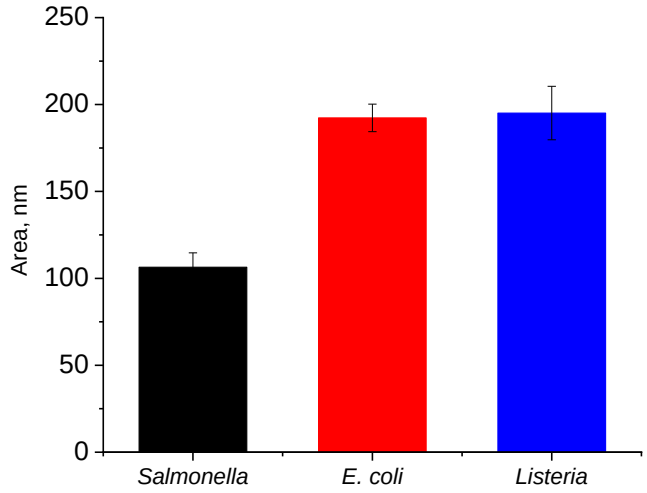


Figure VII 35. Area under spectral peaks of *Salmonella*, *E. coli* and *Listeria* 24 h (±SE).

Table VII 30. ANOVA Area 24 h

Group Name	N (count)	Mean	Std. Error
Group 1	10	106.39	8.29
Group 2	10	192.32	7.92
Group 3	12	195.10	15.39

Desired confidence level for post-hoc confidence intervals: 95%

Source of Variation	Sum of Squares	d.f.	Variance	F	p
Between Groups:	52604.78	2	26302.39	17.70	0.00001
Within Groups:	43107.74	29	1486.47		
Total:	95712.52	31			

Tukey HSD Post-hoc Test...

Group 1 vs Group 2: Diff=85.9238, 95%CI=43.3415 to 128.5061, p=0.0001

Group 1 vs Group 3: Diff=88.7007, 95%CI=47.9313 to 129.4701, p=0.00001

Group 2 vs Group 3: Diff=2.7769, 95%CI=-37.9925 to 43.5463, p=0.9845

A one-way ANOVA was conducted to compare areas under hyperspectral peaks for *Salmonella, E. coli* and *Listeria*. There was significant effect of species on the area at the $p<0.05$ level for three bacteria [F (2, 29) =17.6, p=0.00001]. Post hoc comparisons using the Tukey HSD test indicated that the mean areas under spectral peaks of *Salmonella* vs *E. coli* and *Salmonella* vs *Listeria* 24 h were significantly different. However, when comparing *E. coli* vs *Listeria,* no significant difference was found.

VII.1.11.2 FWHM

Taking values of FWHM of three bacteria from Tables VII 26-28, respectively, we can present them in Table VII 31.

Table VII 31. FWHM of three bacteria at 24 h.

Groups	N total	Bacteria	FWHM, nm	SE
Group 1	10	*Salmonella*	83.49	4.70
Group 2	10	*E. coli*	182.88	6.75
Group 3	12	*Listeria*	314.33	17.67

FWHM of *Salmonella, E. coli* and *Listeria* spectra at 24 h is shown in Figure VII 36.

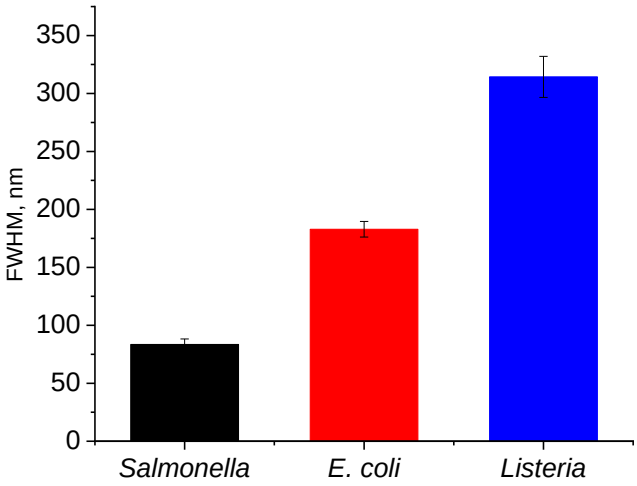


Figure VII 36. FWHM of spectral peaks of *Salmonella*, *E. coli* and *Listeria* 24 h (±SE).

A one-way ANOVA was conducted for comparison of FWHM of spectral peaks of *Salmonella* (Group 1), *E. coli* (Group 2), and *Listeria* (Group 3) 24 h. The data from Table VII 31 were entered into the window of the computer program "Analysis of Variance from Summary Data" [49] and calculations were carried out as shown in Table VII 32.

Table VII 32. ANOVA FWHM 24 h

Group Name	N (count)	Mean	Std. Error
Group 1	10	83.49	4.70
Group 2	10	182.88	6.75
Group 3	12	314.33	17.67

Desired confidence level for post-hoc confidence intervals: 95

Source of Variation	Sum of Squares	d.f.	Variance	F	p
Between Groups:	295513.14	2	147756.57	90.60	0.00001
Within Groups:	47294.36	29	1630.84		
Total:	342807.50	31			

Tukey HSD Post-hoc Test...

Group 1 vs Group 2: Diff=99.3956, 95%CI=54.7934 to 143.9978, p=0.00001

Group 1 vs Group 3: Diff=230.8481, 95%CI=188.1448 to 273.5514, p=0.00001

Group 2 vs Group 3: Diff=131.4525, 95%CI=88.7492 to 174.1558, p=0.00001

A one-way ANOVA was conducted to compare areas under hyperspectral FWHM of peaks for *Salmonella, E. coli* and *Listeria*. There was significant effect of species on the FWHM at the $p<0.05$ level for three bacteria [$F(2, 29) = 90.6$, $p=0.00001$]. Post hoc comparisons using the Tukey HSD test indicated that the mean FWHM for all three bacteria were significantly different.

VII.1.11.3 Center

Taking values of Center of three bacteria from Tables VII 26-28, respectively, we can present them in Table VII 33 and Figure VII 37.

Table VII 33. Centers for three bacteria at 24 h

Groups	Bacteria	Center, nm	SE
Group 1	*Salmonella*	434.88	0.31
Group 2	*E. coli*	444.40	6.49
Group 3	*Listeria*	455.31	1.59

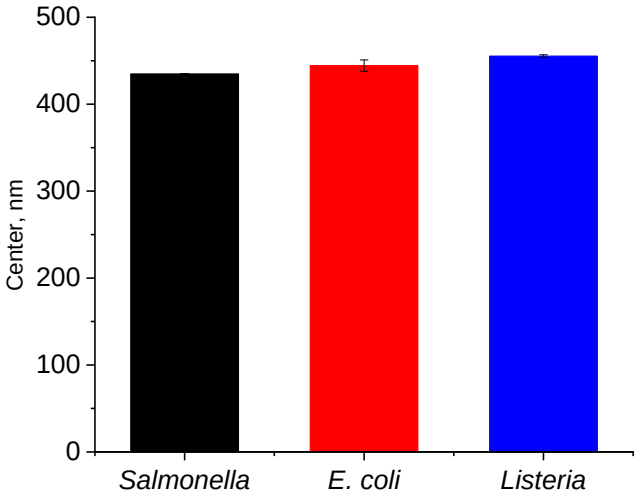


Figure VII 37. Centers of *Salmonella, E. coli* and *Listeria* spectra at 24 h (±SE).

Table VII 34. ANOVA Center 24 h.

Group Name	N (count)	Mean	Std. Error
Group 1	10	434.88	0.31
Group 2	10	444.40	6.49
Group 3	12	455.31	1.59

Desired confidence level for post-hoc confidence intervals: 95%

Source of Variation	Sum of Squares	d.f.	Variance	F	p
Between Groups:	2295.05	2	1147.52	8.05	0.0017
Within Groups:	4132.74	29	142.51		
Total:	6427.79	31			

Tukey HSD Post-hoc Test.

Group 1 vs Group 2: Diff=9.5181, 95%CI=-3.6666 to 22.7028, p=0.1931

Group 1 vs Group 3: Diff=20.4310, 95%CI=7.8076 to 33.0544, p=0.0011

Group 2 vs Group 3: Diff=10.9129, 95%CI=-1.7105 to 23.5363, p=0.1003

A one-way ANOVA was conducted to compare Centers of hyperspectral peaks for *Salmonella, E. coli* and *Listeria*, 24 h. There was significant effect of species on the Centers at the $p<0.05$ level for three bacteria [F (2, 29) =8.05, p=0.001]. Post hoc comparisons using the Tukey HSD test indicated that the mean Centers for *Salmonella* vs *E. coli* and *Salmonella* vs *Listeria* were not significantly different. Centers for *E. coli* vs *Listeria* were significantly different.

VII.1.11.4 *Centroid*

Taking values of Centroid of three bacteria from Tables VII 26-28, respectively, we can present them in Table VII 35 and Figure VII 38.

Table VII 35. Centroid of three bacteria at 24 h.

Groups	Bacteria	Centroid, nm	SE
Group 1	*Salmonella*	561.79	31.85
Group 2	*E. coli*	613.88	3.10
Group 3	*Listeria*	591.38	11.94

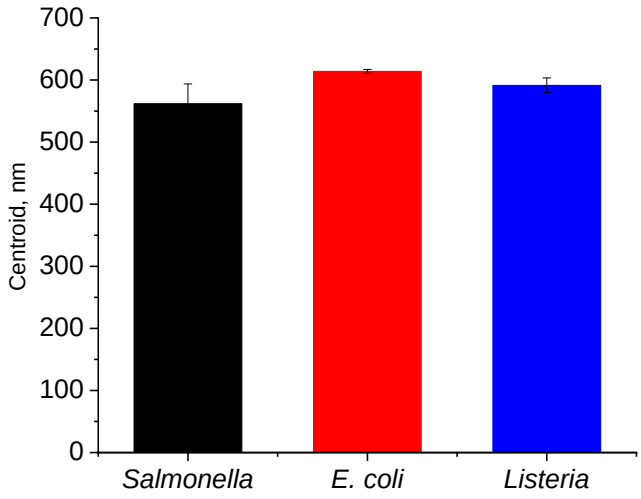


Figure VII 38. Centroids of spectral peaks of *Salmonella*, *E. coli* and *Listeria* 24 h (±SE).

A one-way ANOVA was conducted for comparison of Centroids of spectral peaks of *Salmonella* Group 1), *E. coli* (Group 2), and *Listeria* (Group 3) 24 h. The data from Table VII 35 were entered into the window of the computer program "Analysis of Variance from Summary Data" [49] and calculations were carried out as shown in Table VII 36.

Table VII 36. ANOVA Centroid 24 h.

Group Name	N (count)	Mean	Std. Error
Group 1	10	561.79	31.85
Group 2	10	613.88	3.10
Group 3	12	591.38	11.94

Desired confidence level for post-hoc confidence intervals: 95%

Source of Variation	Sum of Squares	d.f.	Variance	F	p
Between Groups:	13660.94	2	6830.47	1.79	0.1857
Within Groups:	110970.79	29	3826.58		
Total:	124631.72	31			

Tukey HSD Post-hoc Test...

Group 1 vs Group 2: Diff=52.0894, 95%CI=-16.2318 to 120.4106, p=0.1617

Group 1 vs Group 3: Diff=29.5926, 95%CI=-35.8200 to 95.0052, p=0.5114

Group 2 vs Group 3: Diff=-22.4968, 95%CI=-87.9094 to 42.9158, p=0.6759

A one-way ANOVA was conducted to compare Centroids of hyperspectral peaks for *Salmonella, E. coli* and *Listeria*. There was no significant effect of species on the Centroids at the $p<0.05$ level for three bacteria [$F(2, 29) = 1.78$, $p=0.185$]. Post hoc comparisons using the Tukey HSD test indicated that the mean Centers for all three bacteria were not significantly different ($p=0.058$).

VII.1.11.5 *Summary of spectra comparison for three bacteria at 24 h (Local Maximum model)*

Results on the comparison of three bacteria at 24 h are shown below in Table VII 37.

Table VII 37. Differentiation of *Salmonella*, *E. coli* and *Listeria* 24 h by ANOVA

Comparison	Area	ANOVA	FWHM	ANOVA	Center	ANOVA	Centroid	ANOVA
Salmonella vs *E. coli*	+	[F (2, 29) =17.6 p=0.00001].	+	[F (2, 29) =90.6, p=0.00001]	-	[F (2, 29) =8.05 p=0.001]	-	[F (2, 29) =1.78, p=0.185]
Salmonella vs *Listeria*	+		+		+		-	
E. coli vs *Listeria*	-		+		-		-	

+ significantly different; - not significantly different

VII.2 Spectra of *Salmonella, E. coli* and *Listeria* (Statistical models)

We will use non-linear fitting of our peak spectral data to particular statistical model to find which model fits best to experimental data. We can describe the process of nonlinear curve fitting as below.[45]

1. Generate an initial function curve from the initial values.
2. Iterate to adjust parameter values to make data points closer to the curve.
3. Stop when minimum distance reaches the stopping criteria to get the best fit

Origin software (Origin) provides options of different algorithm, which have different iterative procedure and statistics to define minimum distance. To describe the hyperspectral peaks, we

tried to fit 24 explicit functions. We have chosen six functions that converged with our peak experimental data. The list of functions that we used in our work is given in Table VII 38.

Table VII 38. Functions for fitting experimental hyperspectral peaks.

Name	Function/Parameters	Brief description	Ref
Extreme. Gumbel distribution	$y = y_0 + Aexp[-\exp\left[-\left(\frac{x-x_c}{w}\right)\right] - \left(\frac{x-x_c}{w}\right) + 1]$ Parameters: Number: 4; Names: y0, xc, w, A; Meanings: y0 = offset, xc = center, w = width, A = amplitude, Lower Bounds: w > 0.0, Upper Bounds: none	Gumbel distribution	[21]
GaussAmp.	$y = y_0 + Ae^{-\frac{(x-xc)^2}{2w^2}}$ Number: 4; Names: y0, xc, w, A Meanings: y0 = offset, xc = center, w = width, A = amplitude; Lower Bounds: w > 0.0; Upper Bounds: none	Amplitude version of Gaussian peak function	[68]
GCAS	$\boldsymbol{f(z) = y_0 + \frac{A}{w\sqrt{2\pi}} e^{\frac{-z^2}{2}}(1 + / \sum_{i-3}^{4} \frac{a_i}{i!} H_i(z)/)}$; $z = \frac{x - x_c}{w}$ $H_3 = z^3 - 3z$ $H_4 = z^4 - 6z^3 + 3$ Parameters: Number: 6; Names: y0, xc, A, w, a3, a4; Meanings: y0 = offset, xc = center, A=area, w=half width, a3 = unknown, a4 = unknown; Lower Bounds: w > 0.0; Upper Bounds: none	Gram-Charlier peak function for use in chromatography.	[14]
Lorentz (Cauchy distribution).	$y = y_0 + \frac{2A}{\pi} \frac{w}{4(x - x_c)^2 + w^2}$ Parameters: Number: 4; Names: y0, xc, w, A; Meanings: y0 = offset, xc = center, w = FWHM, A = area; Lower Bounds: w > 0.0; Upper Bounds: none	Lorentzian peak function with bell shape and much wider tails than Gaussian function	[68]
InvsPoly. Inverse polynomial peak function with center	$\boldsymbol{y = y_0 + \frac{A}{1 + A_1(2\frac{x-x_c}{w})^2 + A_2(2\frac{x-x_c}{w})^4 + A_3(2\frac{x-x_c}{w})^6}}$ Parameters: Number: 7; Names: y0, xc, w, A, A1, A2, A3; Meanings: y0 = offset, xc = center, w = width, A = amplitude, A1 = coefficient, A2 = coefficient, A3 = coefficient; Lower Bounds: w > 0.0; Upper Bounds: none	Inverse polynomial peak function with center	[68]
Logistpk,	$y = y_0 + \frac{4Ae^{-\frac{x-xc}{w}}}{\left(1 + e^{-\frac{x-xc}{w}}\right)^2}$	Logistic peak function, also called Hubbert function	[68]

	Parameters: Number: 4; Names: y0, xc, w, A; Meanings: y0 = offset, xc = center, w = width, A = amplitude; Lower Bounds: w > 0.0; Upper Bounds: none		

VII.2 *Salmonella* at 21 hours of growth

VII.2.1 Salmonella 21 h Nonlinear Curve Fit by GCAS

Table VII 39. Parameters of the *Salmonella* 21 h Nonlinear Curve Fit by GCAS

	Value	SE	t-Value	Pro>|t|	Dependency
y0	0.468	0.00343	136.4231	0	0.61996
xc	505.898	0.92996	544.0013	0	0.56495
A	100.3784	1.81735	55.23347	0	0.73995
w	72.66418	0.81076	89.62503	0	0.58995
a3	1.18782	0.0467	25.4365	0	0.40089
a4	-1.4025	0.04348	-32.2595	0	0.4734

SE: standard error; t-value: the test statistic for t-test; Prob>|t|: The p-value for t-test.

t-Value: the test statistic for t-test

t-Value = Fitted value/Standard Error, for example the t-Value for y0 is 0.468/0.00342 = 136.4231.

For this statistical t-value, it usually compares with a critical t-value of a given confident level alpha (5%). If the t-value is larger than the critical t-value ($|t|>t_{\alpha/2}$ t, it can be said that there is a significant difference. However, the Prob>|t| is easier to interpret, and it is recommended that we ignore t-Value and judge by Prob>|t|.

Prob>|t|: The p-value for t-test

If Prob>|t| < alpha (5%), that means we find enough evidence to reject null hypothesis of t-test. The smaller Prob>|t|, the more unlikely the parameter is equal to zero.

Dependency: The dependency value, which is computed from the variance-covariance matrix, typically indicates the significance of the parameter in your model. For example, if some dependency values are close to 1, this could mean that there is mutual dependency between those

parameters. In other words, the function is over-parameterized, and the parameter may be redundant.

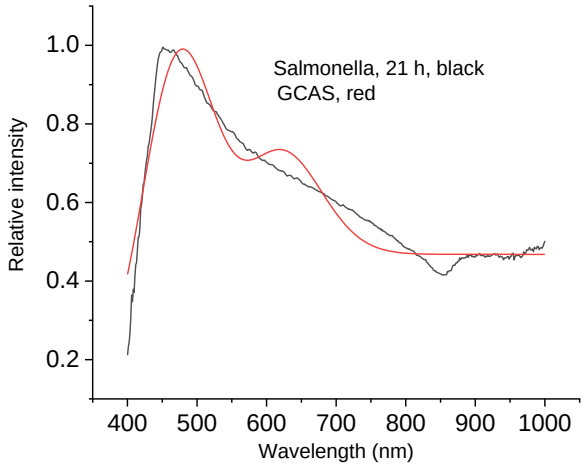


Figure VII 39. Fitting 21 h *Salmonella* with GCAS

Table VII 40. Statistics of the *Salmonella* 21 h Nonlinear Curve Fit by GCAS

	Mean
Number of Points	466
Degrees of Freedom	460
Reduced Chi-Sqr	0.00208
Residual Sum of Squares	0.95869
R-Square (COD)	0.93223
Adj. R-Square	0.93149
Fit Status	Succeeded (100)

Residual Sum of Squares

Residual Sum of Squares is usually abbreviated to RSS. It is actually the sum of the square of the vertical deviations from each data point to the fitting regression line. It can be inferred that your data is perfect fit if the value of RSS is equal to zero. This statistic can help to decide if the fitted regression line is a good fit for your data, the smaller the residual sum of squares, the better your model fits your data.

Scale Error with sqrt (Reduced Chi-Sqr)

The Reduced Chi-square value, which is also called Scale Error with square, is equal to the residual sum of square (RSS) divided by the degree of freedom. Typically, a Reduced Chi-square

value close to 1 indicates a good fit result, and it implies that the difference between observed data and fitted data is consistent with the error variance. If the error variance is over-estimated, the Reduced Chi-square value will be much less than 1. For under-estimated error variance, it will be much greater than 1. Note that it needs to select the correct variance in fitting procedure for Reduced Chi-square. For example, if the y data is multiplied by a scaling factor simply, the Reduced Chi-square will be scaled as well. Only if you also scale the error variance by a correct factor, the value of Reduced Chi-square will turn back into a normal value.

Pearson's r

Pearson's correlation coefficient, denoted by Pearson's r, can help to measure the strength of linear relationship between paired data. The value of Pearson's r is constrained between -1 to 1. In linear regression, a positive value of Pearson's r indicates that there is positive linear correlation between predictor variable(x) and response variable(y), while a negative value of Pearson's r indicates that there is negative linear correlation between predictor variable(x) and response variable(y). The value of zero indicates that there is no linear correlation between data. What's more, the closer the value is to -1 or 1, the stronger linear correlation is.

R-Square (COD)

R-square, which is also known as the coefficient of determination (COD), is a statistical measure to qualify the linear regression. It is a percentage of the response variable variation that explained by the fitted regression line, for example the R-square suggests that the model explains approximately more than 89% of the variability in the response variable. Hence, R-square is always between 0 and 1. If R-square is 0, it indicates that fitted line explains none of the variability of the response data around its mean; while if R-square is 1, it indicates that, the fitted

line explains all the variability of the response data around its mean. In general, the larger the R-square, the better the fitted line fits your data.

Adj. R-Square

R-square can be used to quantify how well a model fits the data, and R-square will always increase when a new predictor is added. It is a misunderstanding that a model with more predictors has a better fit. The Adj. R-square is a modified version of R-square, which is adjusted for the number of predictors in the fitted line. Thus, it can be used to compare with the fitted lines with different numbers of predictors. If the number of predictors is greater than 1, Adj.-square is always smaller than R-square.

Table VII 41. Summary of the *Salmonella* 21 h Nonlinear Curve Fit by GCAS

y0	SE	xc	SE	A	SE	w	SE
0.468	0.00343	505.898	0.92996	100.3784	1.81735	72.66418	0.81076

a3	SE	a4	SE	Chi-Sqr	R-Sqr
1.18782	0.0467	-1.4	0.04348	0.00208	0.93149

VII.2.2 Salmonella 21 h Nonlinear Curve Fit Lorentz

Table VII 42. Parameters of the *Salmonella* 21 h Nonlinear Curve Fit by Lorentz

	Value	Standard Error	t-Value	Prob>|t|	Dependency
y0	0.45315	0.00773	58.58805	5.41E-216	0.76021
xc	501.9773	2.36556	212.2027	0	0.06284
w	191.691	10.35343	18.51474	1.18E-57	0.82185
A	140.58	7.66991	18.32877	8.56E-57	0.90861
H	0.46688	0.01178			

Table VII 43. Statistics of the *Salmonella* 21 h Nonlinear Curve Fit by Lorentz

	Mean
Number of Points	466
Degrees of Freedom	462
Reduced Chi-Sqr	0.00668
Residual Sum of Squares	3.08834
R-Square (COD)	0.78168

Adj. R-Square	0.78027
Fit Status	Succeeded (100)

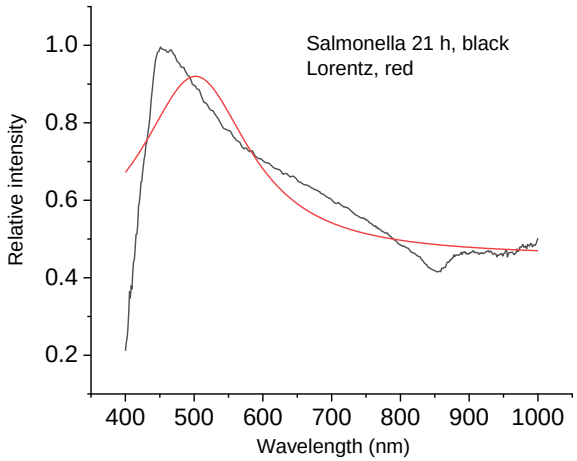


Figure VII 40. Fitting *Salmonella* 21 h with Lorentz

Table VII 44. Summary of the *Salmonella* 21 h Nonlinear Curve Fit Lorentz

y0	SE	xc	SE	w	SE	A	SE
0.45315	0.00773	501.9773	2.36556	191.691	10.35343	140.58	7.66991

H	SE	Chi-Sqr	R-Sqr
0.46688	0.01178	0.00668	0.78027

VII.2.3 Salmonella 21 h Nonlinear Curve Fit by Extreme

Table VII 45. Parameters of the *Salmonella* 21 h Nonlinear Curve fit by Extreme

Value	SE	t-Value	Prob>|t|		Dependency
y0	0.47985	0.0053	90.46629	0	0.63933
xc	493.5513	1.48427	332.5205	0	0.06992
w	66.41225	1.94797	34.09311	0	0.59982
A	0.46459	0.00948	48.99905	0	0.47369

Table VII 46. Statistics of the *Salmonella* 21 h Nonlinear Curve Fit by Extreme

	Mean
Number of Points	466
Degrees of Freedom	462
Reduced Chi-Sqr	0.00473
Residual Sum of Squares	2.18462
R-Square (COD)	0.84557
Adj. R-Square	0.84457

Fit Status	Succeeded (100)

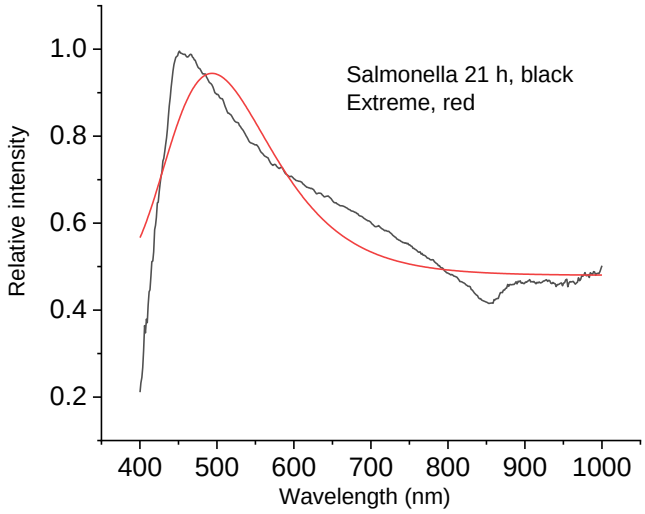


Figure VII 41. Fitting *Salmonella* 21 h with Extreme

Table VII 47. Summary of the *Salmonella* 21 h Nonlinear Curve Fit by Extreme

y0	SE	xc	SE	w	SE	A	SE	Chi-Sqr	R-Sqr
0.4798	0.0053	493.551	1.48427	66.4122	1.94797	0.4645	0.00948	0.00473	0.84457

VII.2.4 Salmonella 21 h Nonlinear Curve Fit by InvsPoly

Table VII 48. Parameters of the *Salmonella* 21 h Nonlinear Curve Fit by InvsPoly

	Value	SE	t-Value	Prob>|t|	Dependency
y0	0.52318	0.00522	100.285	0	0.35469
xc	500.3035	1.36185	367.3706	0	0.99166
w	206.0305	1721371	1.20E-04	0.9999	1
A	0.33537	0.0145	23.13127	4.84E-79	0.70952
A1	-2.76446	46193.84	-5.98E-05	0.99995	1
A2	10.88345	363722	2.99E-05	0.99998	1
A3	-2.46334	123486.5	-1.99E-05	0.99998	1

Table VII 49. Statistics of the *Salmonella* 21 h Nonlinear Curve Fit by InvsPoly

	Mean
Number of Points	466
Degrees of Freedom	459
Reduced Chi-Sqr	0.00818
Residual Sum of Squares	3.7566
R-Square (COD)	0.73445

Adj. R-Square	0.73097
Fit Status	Succeeded (100)

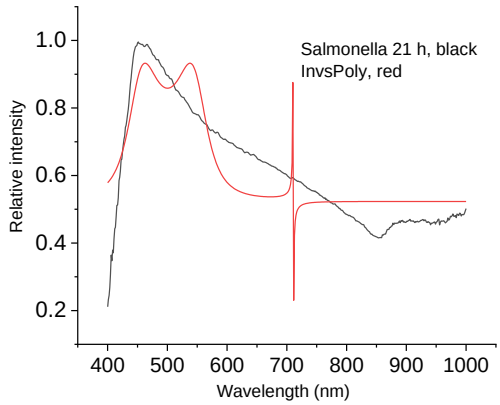


Figure VII 42. Fitting 21 h *Salmonella* with GCAS with InvsPoly.

Table VII 50. Summary of the *Salmonella* 21 h Nonlinear Curve Fit by InvsPoly

y0	SE	xc	SE	w	SE	A	SE
0.52318	0.00522	500.3035	1.36185	206.0305	1721371	0.33537	0.0145

A1	SE	A2	SE	A3	SE	Chi-Sqr	R-Sqr
-2.76446	46193.84	10.88345	363722	-2.46334	123486.5	0.00818	0.73097

VII.2.5 Salmonella 21 h Nonlinear Curve Fit by GaussAmp

Table VII 51. Parameters of the *Salmonella* 21 h Nonlinear Curve Fit by GaussAmp

	Value	SE	t-Value	Prob>|t|	Dependency
y0	0.48176	0.00635	75.85215	0	0.60373
xc	513.4477	2.88059	178.2438	0	0.116
w	91.07984	3.70175	24.60454	0	0.55352
A	0.39824	0.01087	36.65001	0	0.50121
FWHM	214.48	8.71696			
Area	90.92	3.91622			

Table VII 52. Statistics of the *Salmonella* 21 h Nonlinear Curve Fit by GaussAmp

	Mean
Number of Points	466
Degrees of Freedom	462
Reduced Chi-Sqr	0.00745
Residual Sum of Squares	3.44146
R-Square (COD)	0.75672
Adj. R-Square	0.75514

Fit Status	Succeeded (100)

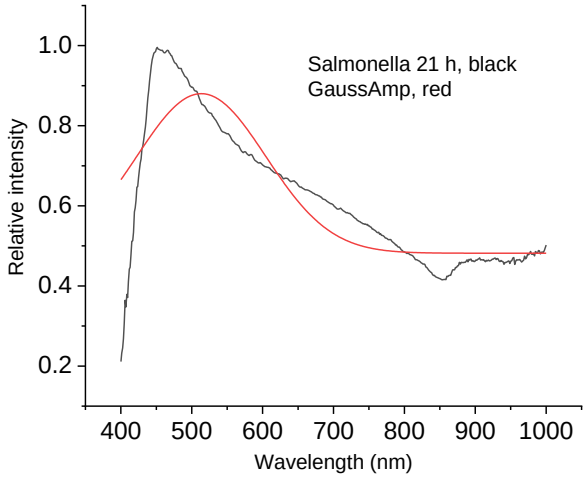


Figure VII 43. Fitting 21 h *Salmonella* with GCAS with GaussAmp

Table VII 53. Summary of the *Salmonella* 21 h Nonlinear Curve Fit by GaussAmp

y0	SE	xc	SE	w	SE	A	SE
0.48176	0.00635	513.4477	2.88059	91.07984	3.70175	0.39824	0.01087

FWHM	SE	Area	SE	Chi-Sqr	R-Sqr
214.48	8.72	90.92	3.92	0.00745	0.75514

VII.2.6 Salmonella 21 Nonlinear Curve Fit Logistpk

Table VII 54. Parameters of the *Salmonella* 21 Nonlinear Curve Fit Logistpk

	Value	SE	t-Value	Prob>|t|	Dependency
y0	0.47806	0.00658	72.66224	0	0.64789
xc	508.9701	2.68537	189.5344	0	0.09881
w	57.74138	2.56477	22.51327	0	0.61285
A	0.41596	0.01101	37.77036	0	0.51086
FWHM	203.5669	9.04209			

Table VII 55. Statistics of the *Salmonella* 21 h Nonlinear Curve Fit Logistpk

	Mean
Number of Points	466
Degrees of Freedom	462
Reduced Chi-Sqr	0.0071
Residual Sum of Squares	3.28141
R-Square (COD)	0.76804

Adj. R-Square	0.76653
Fit Status	Succeeded(100)

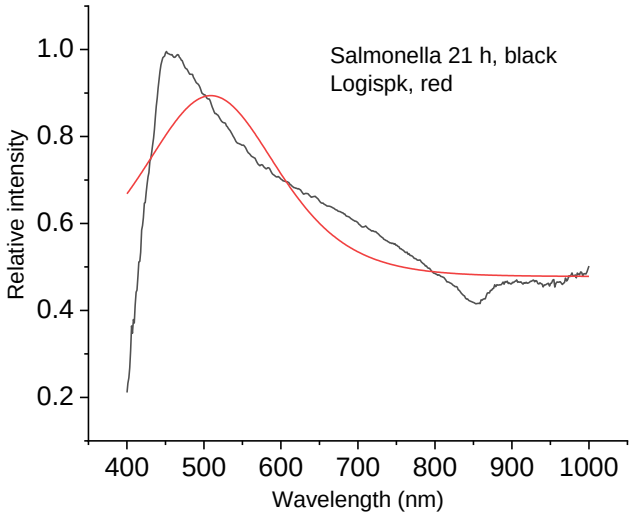


Figure VII 44. Fitting *Salmonella* 21 h with Logistpk

Table VII 56. Summary of the *Salmonella* 21 h Nonlinear Curve Fit Logistpk.

y0	SE	xc	SE	w	SE	A	SE	FWHM	SE	Chi-Sqr	R-sqr
0.478	0.006	508.9	2.68	57.74	2.56	0.41	0.011	203.56	9.04	0.007	0.76

VII.2.7 Rank of non-linear fitting of Salmonella 21 h by statistical functions.

Table VII 57. Ranks of fitting of the *Salmonella* 21 h mean spectrum by GCAS, Extreme, Lorentz, Logistpk, GaussAmp, and InvsPoly.

Functions	Status	AIC	BIC	Adj. R-Sqr	Residual Sum of Sqrs	Chi-Sqr	Outcome
GCAS	Succeeded	-2868.85	-2839.84	0.9314	0.95869	0.00208	Fit converged
Extreme	Succeeded	-2489.04	-2468.32	0.8445	2.18462	0.00473	Fit converged
Lorentz	Succeeded	-2327.71	-2306.99	0.7802	3.08834	0.00668	Fit converged
Logistpk	Succeeded	-2299.45	-2278.73	0.7665	3.28141	0.0071	Fit converged
GaussAmp	Succeeded	-2277.26	-2256.54	0.7551	3.44146	0.00745	Fit converged
InvsPoly	Succeeded	-2230.43	-2197.28	0.7309	3.7566	0.00818	Fit converged

Chi-Sqr tolerance value of 1E-9 was reached.

Goodness of Fit:[38]

- Akaike information criterion (AIC)[38]

Recommended criteria for nonlinear curve fitting. The model with smaller AIC value is more likely to be correct.

- Bayesian information criterion (BIC)[38]

Recommended criteria for nonlinear curve fitting. The model with smaller BIC value is more likely to be correct.

- Adj R-Square

The model with smaller Adj R-Square value is more likely to be correct.

- Residual Sum of Squares

If there are two independent variables in the regression model, the model with smaller residual sum of squares is more likely to be correct.

- Reduced Chi-Sqr

The reduced chi-squares obtained by dividing the residual sum of squares (RSS) by the degrees of freedom (DOF)

According to Table VII 60, the function GCAS is the best fitting function for *Salmonella* 21 h spectral peak.

VII.3 *E. coli* at 21 hours of growth

Rank of non-linear fitting of E. coli 21 h by statistical functions.

Table VII 58. Ranks of fitting of the *E. coli* 21 h mean spectrum by GCAS, Extreme, Lorentz, Logistpk, GaussAmp, and InvsPoly.

Functions	Status	AIC	BIC	Adj. R-Square	Residual Sum of Sqrs	Chi-Sqr
GCAS	Succeeded	-2937.03	-2908.02	0.9275	0.8282	0.0018
Extreme	Succeeded	-2696.96	-2676.24	0.8782	1.3982	0.00303
Lorentz	Succeeded	-2536.51	-2515.79	0.8282	1.9730	0.00427
Logistpk	Succeeded	-2465.74	-2445.02	0.8000	2.2966	0.00497
GaussAmp	Succeeded	-2439.44	-2418.72	0.7884	2.4299	0.00526
InvsPoly	Succeeded	-2321.67	-2288.52	0.7293	3.0886	0.00673

All fits converged; Chi-Sqr tolerance value of 1E-9 was reached.

According to Table VII 58, the function GCAS is the best fitting function for *E. coli* 21 h spectral peak.

VII.4 *Listeria* at 21 hours of growth

Rank of non-linear fitting of Listeria 21 h by statistical functions.

Table VII 59. Ranks of fitting of the *Listeria* 21 h mean spectrum by GCAS, Extreme, Lorentz, Logistpk, GaussAmp, and InvsPoly.

Functions	Status	AIC	BIC	Adj. R-Square	Residual Sum of Sqrs	Chi-Sqr
GCAS	Succeeded	1514.54	1543.55	0.94953	11663.61	25.35566
Extreme	Succeeded	1955.59	1976.31	0.86941	30310.84	65.60787
InvsPoly	Succeeded	2057.12	2090.28	0.83865	37208.01	81.0632
Lorentz	Succeeded	2093.70	2114.42	0.82436	40767.71	88.2418
Logistpk	Succeeded	2108.25	2128.98	0.81879	42061.02	91.04117
GaussAmp	Succeeded	2123.09	2143.81	0.81293	43421.29	93.98548

All fits converged; Chi-Sqr tolerance value of 1E-9 was reached.

According to Table VII 59, the function GCAS is the best fitting function for *Listeria* 21 h spectral peak.

VII.5 Comparison of *Salmonella, E. coli* and *Listeria* spectra at 21 hours of growth

For all 3 bacteria, GCAS appeared to give the best fit. Therefore, we can compare fitting data of *Salmonella*, *E. coli* and *Listeria* spectra at 21 h using Compare Databases Module of Origin 2019. [45]

VII.5.1 Salmonella vs E. coli 21 h.

Results of comparison for *Salmonella* and *E. coli* are given in Tables VII 57-58.

Table VII 60. Comparisons of *Salmonella* 21 h and *E. coli* 21 h spectra

Input Data	Mean relative intensity of *Salmonella* 21 h spectrum as function wavelength	Mean relative intensity of *E. coli* 21 h spectrum as function wavelength
Function	GCAS	GCAS
Number of Points	466	466
Number of Parameters	6	6

Table VII 61. F-test, comparison of the *Salmonella* and *E. coli* 21 h spectra fitted by GCAS.

F	Numerator DF	Denominator. DF	Prob > F
123.04	6	920	0

At the 0.05 significance level, the *Salmonella* and *E. coli* 21 h spectra are statistically different.

Prob>F: The p-value for F-test. If Prob>F< lower than level of significance (usually be 5%), that means we find enough evidence to reject 0-hypothesis of F-test. The smaller Prob>F, the more unlikely the parameter is equal to zero.

VII.5.2 Salmonella vs Listeria, 21 h.

Similarly, to the comparison of *Salmonella* vs *E. coli* 21 h spectra, we compared the 21-hour spectra of *Salmonella* and *Listeria*, the F-test result is shown in Table VII 62.

Table VII 62. F-test, comparison of the *Salmonella* and *Listeria* 21 h spectra fitted by GCAS.

F	Numerator DF	Denominator DF	Prob > F
322.18	6	920	0

At the 0.05 significance level, *Salmonella* and *Listeria* 21 h spectra are statistically different.

VII.5.3 E. coli vs Listeria, 21 h

Likewise, the comparison of *Salmonella* vs *E. coli* 21 h spectra, we have for comparison of *E. coli* vs *Listeria*, 21 h F-test result shown in Table VII 63.

Table VII 63. F-test, comparison of the *E. coli* and *Listeria* 21 h spectra fitted by GCAS.

F	Numerator DF	Denominator DF	Prob > F
795.18	6	920	0

At the 0.05 significance level, the *E. coli* and *Listeria* 21 h spectra are statistically different

VII.5.4 *Summary of spectra comparison for three bacteria at 21 h (Statistical model)*

Taking together results on the comparison of three bacteria at 21 h are shown in Table VII 64.

Table VII 64. Differentiation of *Salmonella, E. coli* and *Listeria* 21 h fitted by GCAS by F-test

Comparison	F	Significance
Salmonella vs *E. coli*	F(6, 920)=123.04, 0	+
Salmonella vs *Listeria*	F(6, 920)=322.18, 0	+
E. coli vs *Listeria*	F(6, 920)=795.17, 0	+

At the 0.05 significance level, the *Salmonella, E. coli* and *Listeria* 21 spectra are statistically different.

Thus, comparing 6 parameters and fitting statistics of GCAS, the *Salmonella, E. coli* and *Listeria* 21 h spectra are significantly different at the 0.05 level. Because the degrees of freedom are the same for all comparisons and Prob>F=0, the higher F value indicates a higher degree of discrimination between bacteria. Thus, the discrimination level for *E. coli* vs *Listeria* > *Salmonella* vs *Listeria*> *Salmonella* vs *E. coli*.

VII.6 *Salmonella* at 18 hours of growth

Rank of non-linear fitting of Salmonella 18 h by statistical functions

Table VII 65. Ranks of fitting of the *Salmonella* 18 h mean spectrum by GCAS, Extreme, Lorentz, Logistpk, GaussAmp, and InvsPoly.

Functions	Status	AIC	BIC	Adj. R-Square	Residual Sum of Sqrs	Chi-Sqr
GCAS	Succeeded	-3527.69	-3498.68	0.94097	0.23317	5.07E-04
InvsPoly	Succeeded	-3424.46	-3391.31	0.92649	0.28974	6.31E-04
Extreme	Succeeded	-3300.5	-3279.78	0.90347	0.38293	8.29E-04
GaussAmp	Succeeded	-3189.85	-3169.13	0.8776	0.48556	0.00105

Logistpk	Succeeded	-3186.18	-3165.46	0.87664	0.4894	0.00106
Lorentz	Succeeded	-3183.43	-3162.71	0.87591	0.49229	0.00107

Fits were converged.

According to Table VII 65, the function GCAS is the best fitting function for *Salmonella* 18 h spectral peak.

VII.7 *E. coli* at 18 hours of growth

Rank of non-linear fitting of E. coli 18 h by statistical functions.

Table VII 66. Ranks of fitting of the *E. coli* 18 h mean spectrum by GCAS, Extreme, Lorentz, Logistpk, GaussAmp, and InvsPoly.

Functions	Status	AIC	BIC	Adj. R-Square	Residual Sum of Sqrs	Chi-Sqr
GCAS	Succeeded	-3102.79	-3073.78	0.93747	0.5803	0.00126
Extreme	Succeeded	-2886.31	-2865.59	0.90008	0.93139	0.00202
Lorentz	Succeeded	-2735.58	-2714.86	0.86192	1.28709	0.00279
Logistpk	Succeeded	-2700.32	-2679.60	0.85106	1.38826	0.003
GaussAmp	Succeeded	-2666.26	-2645.54	0.83977	1.49351	0.00323
InvsPoly	Succeeded	-2338.40	-2305.25	0.67824	2.97969	0.00649

All Fits converged. Chi-Sqr tolerance value of 1E-9 was reached

According to Table VII 66, the function GCAS is the best fitting function for *E. coli* 18 h spectral peak.

VII.8 *Listeria* at 18 hours of growth

Rank of non-linear fitting of Listeria 18 h by GCAS and Extreme functions.

Table VII 67. Comparison of fits of the *Listeria* 18 h mean spectrum by GCAS and Extreme functions.

	RSS	N	Parameters	AIC	Akaike Weight
GCAS	0.6495	355	6	-2223.80	1
Extreme	2.9313	355	4	-1692.82	5.00E-116

The GCAS model has lower Residual Sum of Squares (RSS) and Akaike information criterion (AIC) values and so more likely to be correct.[38, 67] Additionally, value of Akaike Weights indicate that the GCAS model is many times more likely to be correct.

VII.9 Comparison of *Salmonella, E. coli* and *Listeria* spectra at 18 hours of growth

For all 3 bacteria, GCAS appeared to give the best fit. Therefore, we can compare fitting data of *Salmonella* , *E. coli* and *Listeria* spectra at 18 h using Compare Databases Module of Origin 2019. [45]

VII.9.1 Salmonella vs E. coli 18 h.

Results of comparison are given in Table VII 68.

Table VII 68. F-test, comparison of the *Salmonella* and *E. coli* 18 h spectra fitted by GCAS.

F	Numerator DF	Denominator DF	Prob > F
865.02	6	809	0

At the 0.05 significance level, the *Salmonella* and *E. coli* 18 h spectra are statistically different.

VII.9.2 Salmonella vs Listeria, 18 h.

Similarly, the comparison of *Salmonella* vs *E. coli* 21 h spectra, we have for comparison of *Salmonella* vs *Listeria,* 18 h F-test result shown in Table VII 69.

Table VII 69. F-test, comparison of the *Salmonella* and *Listeria* 18 h spectra fitted by GCAS.

F	Numerator DF	Denominator DF	Prob > F
4730.53	6	809	0

At the 0.05 significance level, the *E. coli* and *Listeria* 18 h spectra are statistically different.

VII.9.3 E. coli vs Listeria, 18 h

Likewise, the comparison of *Salmonella* vs *E. coli* 21 h spectra, we have for comparison of *E. coli* vs *Listeria,* 18 h F-test result shown in Table VII 69.

Table VII 70. F-test, comparison of the *E. coli* and *Listeria* 18 h spectra fitted by GCAS.

F	Numerator DF	Denominator DF	Prob > F

1737.26	6	809	0

At the 0.05 significance level, the *E. coli* and *Listeria* 18 h spectra are statistically different.

VII.9.4 Summary of spectra comparison for three bacteria at 18 h (Statistical model)

Taking together results on the comparison of three bacteria at 18 h are shown in Table VII 70.

Table VII 71. Differentiation of *Salmonella, E. coli* and *Listeria* 18 h fitted with GCAS by F-test.

Comparison	F	Significance
Salmonella vs *E. coli*	F(6, 809)= 865.0225, 0	+
Salmonella vs *Listeria*	F(6, 809)= 4730.528, 0	+
E. coli vs *Listeria*	F(6, 809)= 1737.262, 0	+

At the 0.05 significance level, the *Salmonella, E. coli* and *Listeria* 18 h spectra are statistically different.

Thus, comparing 6 parameters and fitting statistics of GCAS, the *Salmonella, E. coli* and *Listeria* 18 h spectra are significantly different at the 0.05 level. Because, degrees of freedom are the same for all comparisons and Prob>F=0, the higher F value indicates higher discrimination between bacteria. Thus, the discrimination level for *Salmonella* vs *Listeria*> *E. coli* vs *Listeria*> *Salmonella* vs *E. coli* O157:H7.

VII.10 *Salmonella* at 24 hours of growth

Rank of non-linear fitting of Salmonella 24 h by statistical functions

Table VII 72. Ranks of fitting of the *Salmonella* 24 h mean spectrum by GCAS, Extreme, Lorentz, Logistpk, GaussAmp, and InvsPoly.

Functions	Status	AIC	BIC	Adj. R-Square	Residual Sum of Sqrs	Chi-Sqr
GCAS	Succeeded	-3170.8	-3141.79	0.95556	0.5015	0.00109
Extreme	Succeeded	-2952.19	-2931.47	0.92866	0.8086	0.00175
Lorentz	Succeeded	-2837.43	-2816.71	0.90873	1.0344	0.00224
Logistpk	Succeeded	-2716.78	-2696.06	0.88176	1.34007	0.0029
GaussAmp	Succeeded	-2664.36	-2643.64	0.86769	1.49962	0.00325

InvsPoly	Failed	Fit did not converge.

Chi-Sqr tolerance value of 1E-9 was reached for all converged fits.

According to Table VII 72, the function GCAS is the best fitting function for *Salmonella* 24 h spectral peak.

VII.11 *E. coli* at 24 hours of growth

Rank of non-linear fitting of E. coli 24 h by statistical functions.

Table VII 73. Ranks of fitting of the *E. coli* 24 h mean spectrum by GCAS, Extreme, Lorentz, Logistpk, GaussAmp, and InvsPoly.

Functions	Status	AIC	BIC	Adj. R-Square	Residual Sum of Sqrs	Chi-Sqr
GCAS	Succeeded	-2639.93	-2610.92	0.89336	1.56684	0.00341
Extreme	Succeeded	-2468.54	-2447.82	0.8453	2.28287	0.00494
Lorentz	Succeeded	-2365.89	-2345.17	0.80718	2.84543	0.00616
Logistpk	Succeeded	-2344.37	-2323.64	0.79806	2.97993	0.00645
GaussAmp	Succeeded	-2323.77	-2303.05	0.78894	3.11456	0.00674
InvsPoly	Succeeded	-1982.71	-1949.55	0.56398	6.39242	0.01393

All fits converged; Chi-Sqr tolerance value of 1E-9 was reached.

According to Table VII 73, the function GCAS is the best fitting function for *E. coli* 24 h spectral peak.

VII.12 *Listeria* at 24 hours of growth

Rank of non-linear fitting of Listeria 24 h by statistical functions.

Table VII 74. Ranks of fitting of the *E. coli* 24 h mean spectrum by GCAS, Extreme, Lorentz, Logistpk, GaussAmp, and InvsPoly.

Functions	Status	AIC	BIC	Adj. R-Square	Residual Sum of Sqrs	Chi-Sqr
GCAS	Succeeded	-2748.26	-2719.26	0.83557	1.24182	0.0027
Extreme	Succeeded	-2387.96	-2367.24	0.64222	2.71377	0.00587
GaussAmp	Succeeded	-2323.53	-2302.81	0.58916	3.11621	0.00675
Logistpk	Succeeded	-2321.07	-2300.35	0.58699	3.13267	0.00678
Lorentz	Succeeded	-2318.95	-2298.22	0.58510	3.14700	0.00681

InvsPoly	Failed	Fit did not converge

Chi-Sqr tolerance value of 1E-9 was reached for all converged fits.

According to Table VII 74, the function GCAS is the best fitting function for *Listeria* 24 h spectral peak.

VII.13 Comparison of *Salmonella, E. coli* and *Listeria* spectra at 24 hours of growth

For all 3 bacteria at 24 hours of growth, GCAS appeared to give the best fit. Therefore, we can compare fitting data of *Salmonella, E. coli* and *Listeria* spectra at 24 h using Compare Databases Module of Origin 2019.[45]

VII.13.1 Salmonella vs E. coli 24 h.

Results of comparison are given in Table VII 75.

Table VII 75. F-test, comparison of the *Salmonella* and *E. coli* 24 h spectra fitted by GCAS.

F	Numerator DF	Denominator DF	Prob > F
176.28	6	920	0

At the 0.05 significance level, the *Salmonella* and *E. coli* 24 h spectra are statistically different.

VII.13.2 Salmonella vs Listeria, 24 h.

Likewise, the comparison of *Salmonella* vs *E. coli* 24 h spectra, we have for comparison of *Salmonella* vs *Listeria,* 24 h F-test result shown in Table VII 76.

Table VII 76. F-test, comparison of the *E. coli* and *Listeria* 24 h spectra fitted by GCAS.

F	Numerator DF	Denominator DF	Prob > F
3284.42	6	920	0

At the 0.05 significance level, the *E. coli* and *Listeria* 24 h spectra are statistically different.

VII.13.3 E. coli vs Listeria, 24 h

Likewise, the comparison of *Salmonella* vs *E. coli* 24 h spectra, we have for comparison of *E. coli* vs *Listeria,* 24 h F-test result shown in Table VII 77.

Table VII 77. F-test, comparison of the *E. coli* and *Listeria* 24 h spectra fitted by GCAS.

F	Numerator DF	Denominator DF	Prob > F
1704.58	6	920	0

At the 0.05 significance level, the *E. coli* and *Listeria* 24 h spectra are statistically different.

VII.13.4 Summary of spectra comparison for three bacteria at 24 h (Statistical model)

Combining results on the comparison of three bacteria at 24 h are shown in Table VII 78.

Table VII 78. Differentiation of *Salmonella, E. coli* and *Listeria* 24 h fitted by GCAS by F-test

Comparison	F	Significance
Salmonella vs *E. coli*	F(6, 920)= 176.28, 0	+
Salmonella vs *Listeria*	F(6, 920)= 3284.41, 0	+
E. coli vs *Listeria*	F(6, 920)= 1704.57, 0	+

At the 0.05 significance level, the *Salmonella, E. coli* and *Listeria* 24 h spectra are statistically different. Because, degrees of freedom are the same for all comparisons and Prob>F=0, the higher F value indicates higher discrimination between bacteria. Thus, the discrimination level for *Salmonella* vs *Listeria*> *E. coli* vs *Listeria*> *Salmonella* vs *E. coli.*

VII.14 Comparison of *Salmonella* at 18, 21, 24 hours of growth.

Figure VII 45 shows mean values of relative intensity of *Salmonella* spectra at 18, 21, and 24 h of growth.

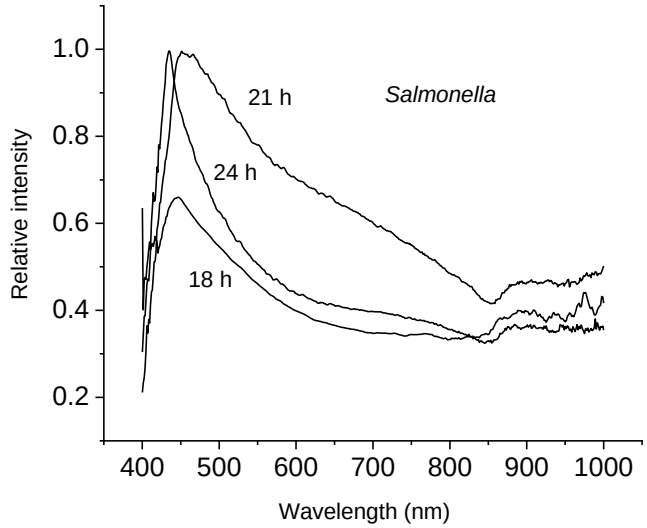


Figure VII 45. Spectra of *Salmonella* at 18, 21, and 24 h of growth.

It is obvious from Figure VII 45; the spectra look different at the different times of growth. To examine this change statistically we will compare these spectra using Compare Databases Module of Origin [11].

VII.14.1 Salmonella 18 h vs Salmonella 21 h.

Results of comparison are given in Table VII 79 (F-test).

Table VII 79. F-test, comparison of the *Salmonella* 18 and *Salmonella* 21 h spectra fitted by GCAS.

F	Numerator DF	Denominator DF	Prob > F
3372.75	6	920	0

At the 0.05 significance level, the *Salmonella* 18 and *Salmonella* 21 h spectra are statistically different.

VII.14.2 Salmonella 18 h vs Salmonella 24 h.

Similarly, the comparison of *Salmonella* 18 vs *Salmonella* 24 h spectra, F-test result shown in Table VII 80.

Table VII 80. F-test, comparison of the *Salmonella* 18 and *Salmonella* 24 h spectra fitted by GCAS.

F	Numerator DF	Denominator DF	Prob > F
298.37	6	920	0

At the 0.05 significance level, the *Salmonella* 18 and *Salmonella* 24 h spectra are statistically different.

VII.14.3 Salmonella 21 h vs Salmonella 24 h.

Likewise, the comparison of *Salmonella* 21 vs *Salmonella* 24 h spectra, F-test result shown in Table VII 81.

Table VII 81. F-test, comparison of the *Salmonella* 21 and *Salmonella* 24 h spectra fitted by GCAS.

F	Numerator DF	Denominator DF	Prob > F
932.67	6	920	0

At the 0.05 significance level, the *Salmonella* 21 and *Salmonella* 24 h spectra are statistically different.

Summary of spectra comparison for Salmonella at 18, 21, and 24 h

Table VII 82. Differentiation of *Salmonella,* at 18, 21, and 24 h fitted by GCAS by F-test

Comparison	F	Significance
Salmonella 18 h vs *Salmonella* 21 h	F(6, 920)= 3372.748, 0	+
Salmonella 18 h vs *Salmonella* 24 h	F(6, 920)= 298.3711, 0	+
Salmonella 21 h vs *Salmonella* 24 h	F(6, 920)= 932.6696, 0	+

At the 0.05 significance level, the *Salmonella* at 18, 21, and 24 h spectra are statistically different. Because, degrees of freedom are the same for all comparisons and Prob>F=0, the higher F value indicates higher discrimination between bacteria. Thus, the discrimination level for *Salmonella* 18 h vs *Salmonella* 21 h> *Salmonella* 21 h vs *Salmonella* 24 h> *Salmonella* 18 h vs *Salmonella* 24 h.

VII.15 Comparison of *E. coli* at 18, 21, 24 hours of growth.

Figure VII 46 shows mean values of relative intensity of *E. coli* spectra at 18, 21, and 24 hours of growth.

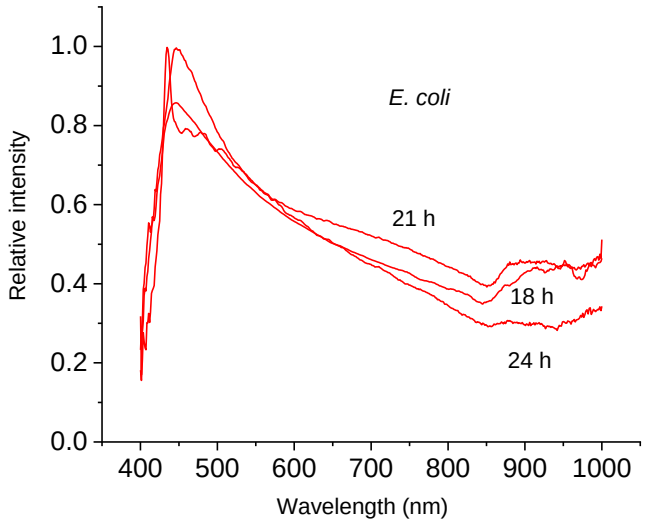


Figure VII 46. Spectra of *E. coli* at 18, 21, and 24 h of growth.

It is clear from Figure VII 46; the spectra look somewhat similar at the different times of growth. To examine this statistically we will compare these spectra using Compare Databases Module of Origin.[45]

VII.15.1 E. coli 18 h vs E. coli 21 h.

Results of comparison are given in Table VII 82 (F-test).

Table VII 83. F-test, comparison of the *E. coli* 18 and *E. coli* 21 h spectra fitted by GCAS
.

F	**Numerator DF**	**Denominator DF**	**Prob > F**
21.49	6	920	0

At the 0.05 significance level, the *E. coli* 18 and *E. coli* 21 h spectra are statistically different.

VII.15.2 E. coli 18 h vs E. coli 24 h.

Similarly, the comparison of *E. coli* 18 vs *E. coli* 24 h spectra, F-test result shown in Table VII 84.

Table VII 84. F-test, comparison of the *E. coli* 18 and *E. coli* 24 h spectra fitted by GCAS.

F	Numerator DF	Denominator DF	Prob > F
298.37	6	920	0

At the 0.05 significance level, the *E. coli* 18 and *E. coli* 24 h spectra are statistically different.

VII.15.3 E. coli 21 h vs E. coli 24 h.

Similarly, the comparison of *E. coli* 21 vs *E. coli* 24 h spectra, F-test result shown in Table VII 85.

Table VII 85. F-test, comparison of the *E. coli* 21and *E. coli* 24 h spectra fitted by GCAS.

F	Numerator DF	Denominator DF	Prob > F
932.67	6	920	0

At the 0.05 significance level, the *E. coli* 21 and *E. coli* 24 h spectra are statistically different.

Summary of spectra comparison for E. coli at 18, 21, and 24 h

Table VII 86. Differentiation of *E. coli* at 18, 21, and 24 h fitted by GCAS by F-test

Comparison	F	Significance
E. coli 18 h vs *E. coli* 21 h	F(6, 920)= 3372.748, 0	+
E. coli 18 h vs *E. coli* 24 h	F(6, 920)= 298.3711, 0	+
E. coli 21 h vs *E. coli* 24 h	F(6, 920)= 932.6696, 0	+

At the 0.05 significance level, the *E. coli* at 18, 21 and 24 h spectra are statistically different.

Because, degrees of freedom are the same for all comparisons and Prob>F=0, the higher F value indicates higher discrimination between bacteria. Thus, the discrimination level for *E. coli* 18 h vs *E. coli* 21 h> *E. coli* 21 h vs *E. coli* 24 h>*E. coli* 18 h vs *E. coli* 24 h

Comparison for Listeria spectra at 18, 21, and 24 h.

Figure VII 47 shows mean values of relative intensity of *Listeria* spectra at 18, 21, and 24 h of growth.

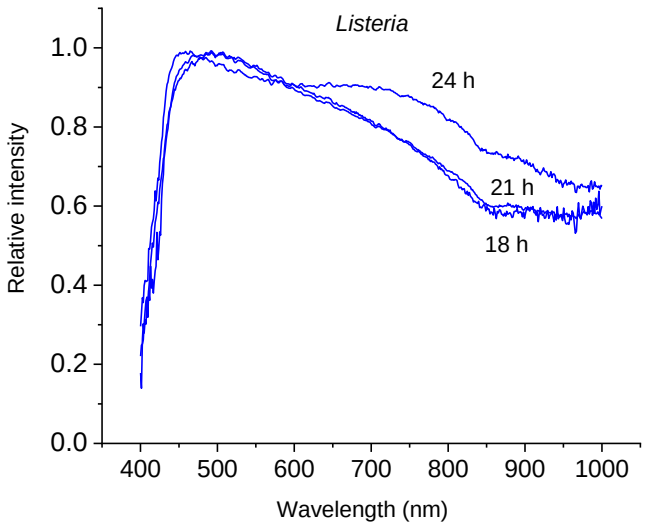


Figure VII 47. Spectra of *Listeria* at 18, 21, and 24 hours of growth.

Figure VII 47 shows that *Listeria* spectra are similar at 18 and 21 h of growth, and somewhat different from that of 24 h of development. We have examined these spectra statistically using Compare Databases Module of Origin.[45]

VII.16 Comparison of *Listeria* at 18, 21, and 24 h.

VII.16.1 Listeria 18 h vs Listeria 21 h.

Results of comparison are given in Table VII 86 (F-test).

Table VII 87. F-test, comparison of the *Listeria* 18 and *Listeria* 21 h spectra fitted by Extreme.

F	Numerator DF	Denominator DF	Prob > F
1.70	4	924	0.14775

At the 0.05 significance level, the *Listeria* 18 and *Listeria* 21 h spectra are not statistically different.

VII.16.2 Listeria 18 h vs Listeria 24 h.

Similarly, the comparison of *Listeria* 18 vs *Listeria* 24 h spectra, F-test result shown in Table VII 88.

Table VII 88. F-test, comparison of the *Listeria* 18 and *Listeria* 24 h spectra fitted by Extreme.

F	Numerator DF	Denominator DF	Prob > F

84.73	4	924	0

At the 0.05 significance level, the *Listeria* 18 and *Listeria 24* h spectra are statistically different.

VII.16.3 Listeria 21 h vs Listeria 24 h.

Similarly, the comparison of *Listeria* 21 vs *Listeria* 24 h spectra, F-test result shown in Table VII 89.

Table VII 89. F-test, comparison of the *Listeria* 21 and *Listeria* 24 h spectra fitted by Extreme.

F	Numerator DF	Denominator .DF	Prob > F
78.10	4	924	0

At the 0.05 significance level, the *Listeria* 21 and *Listeria* 24 h spectra are statistically different.

Summary of spectra comparison for Listeria at 18, 21, and 24 h

Table VII 90. Differentiation of *Listeria*, at 18, 21, and 24 h fitted with Extreme by F-test.

Comparison	F	Significance
Listeria 18 h vs *Listeria* 21 h	F(6, 920)= 1.70, p=0.14775	-
Listeria 18 h vs *Listeria* 24 h	F(6, 920)= 84.73253, p=0	+
Listeria 21 h vs *Listeria* 24 h	F(6, 920)= 78.09705, p=0	+

At the 0.05 significance level, the *Listeria* 18 and *Listeria* 21 h spectra are not statistically different. At the 0.05 significance level, the *Listeria* 18 and *Listeria* 24 h spectra and *Listeria* 21 h vs *Listeria* 24 h are statistically different.

VII.3 Spectra of *Salmonella* lipopolysaccharide mutants and the wild type

Figure VII 48 shows relative intensity of the wild type *Salmonella* and four lipopolysaccharide (LPS) mutants as function of wavelength.

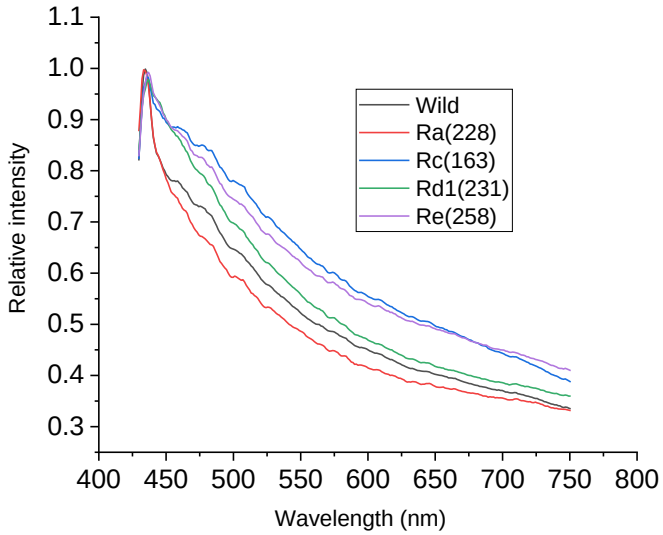


Figure VII 48. Relative intensity of the wild type *Salmonella* and four mutants as function of wavelength. Wild type, and LPS mutants: Ra(228), Rc(163), Rd1(231), and Re(258).

VII.3.1 *Salmonella* 18 h and four LPS mutants.

Rank of non-linear fitting of Salmonella 18 h and 4 mutants by statistical functions.

Table VII 91. Ranks of fitting of *Salmonella* 18 h and 4 mutants by statistical functions

Functions	Status	AIC	BIC	Adj. R-Square	Residual Sum of Sqrs	Chi-Sqr
Lorentz	Succeeded	-2122.01	-2104.27	0.99032	0.06414	2.54E-04
Extreme	Succeeded	-1989.77	-1972.02	0.98381	0.1073	4.24E-04
GaussAmp	Failed					
GCAS	Failed					
InvsPoly	Failed					
Logistpk	Failed					

For converged fits, Chi-Sqr tolerance value of $1E^{-9}$ was reached.

According to Table VII 91, the function Lorentz is the best fitting function for the fitting of *Salmonella* 18 h and 4 mutants.

VII.3.1.1 Wild Salmonella 18 h vs Ra (228) Salmonella 18 h mutant.

A comparison of Wild *Salmonella* 18 h and Ra (228) *Salmonella* 18 h mutant by F-test is shown in table VII 92 (Lorentz model).

Table VII 92. F-test comparison of Wild *Salmonella* 18 h and Ra (228) *Salmonella* 18 h mutant.

F	Numerator DF	Denominator DF	Prob > F
165.46	4	506	0

At the 0.05 significance level, Wild *Salmonella* 18 h and Ra(228) *Salmonella* 18 h mutant are statistically different.

VII.3.1.2 Wild Salmonella 18 h vs Rc (163) Salmonella 18 h mutant.

A comparison of Wild *Salmonella* 18 h and Rc (163) *Salmonella* 18 h mutant by F-test is shown in table VII 93(Lorentz model).

Table VII 93. F-test comparison of Wild *Salmonella* 18 h and Rc(163) *Salmonella* 18 h mutant.

F	Numerator DF	Denominator DF	Prob > F
1640.27	4	506	0

At the 0.05 significance level, Wild *Salmonella* 18 h and Rc(163) *Salmonella* 18 h mutant are statistically different.

VII.3.1.3 Wild Salmonella 18 h vs Rd1(231) Salmonella 18 h mutant.

A comparison of Wild *Salmonella* 18 h and Rd1(231) *Salmonella* 18 h mutant by F-test is shown in table VII 94 (Lorentz model).

Table VII 94. F-test comparison of Wild *Salmonella* 18 h and Rd1(231) *Salmonella* 18 h mutant.

F	Numerator DF	Denominator DF	Prob > F
279.82	4	506	0

At the 0.05 significance level, Wild *Salmonella* 18 h and Rd1(231) *Salmonella* 18 h mutant are statistically different.

VII.3.1.4 Wild Salmonella 18 h vs Re(258) Salmonella 18 h mutant.

A comparison of Wild *Salmonella* 18 h and Re(258) *Salmonella* 18 h mutant by F-test is shown in table VII 95(Lorentz model).

Table VII 95. F-test comparison of Wild *Salmonella* 18 h and Re(258) *Salmonella* 18 h mutant.

F	Numerator DF	Denominator DF	Prob > F
1223.75	4	506	0

At the 0.05 significance level, Wild *Salmonella* 18 h and Re(258) *Salmonella* 18 h mutant are statistically different.

VII.3.1.5 Summary of spectra comparison for Salmonella 18 h and four mutants.

Table VII 96. Comparison of Salmonella 18 h and 4 mutants.

Comparison	F	Significance
Salmonella vs Ra(228)	F(4, 506)= 165.4638, p=0	+
Salmonella vs Rc(163)	F(4, 506)= 1640.266, p=0	+
Salmonella vs Rd1(231)	F(4, 506)= 279.8166, p=0	+
Salmonella vs Re(258)	F(4, 506)= 1223.752, p=0	+

At the 0.05 significance level, Wild *Salmonella* 18 h and 4 mutants are statistically different. Because the degrees of freedom are the same for all comparisons and Prob>F=0, the higher F value indicates higher discrimination between the wildtype *Salmonella* 18 h and mutants. Thus, level of discrimination for *Salmonella* vs Rc(163) > *Salmonella* vs Re(258) > *Salmonella* vs Rd1(231) > *Salmonella* vs Ra(228).

Chapter VIII. Discussion

Unique spectral profiles from *Salmonella, E. coli* and *Listeria* were elicited for each organism as well as distinct changes in the spectral curve were notable when comparing the 18, 21 and 24 h curves generated for each organism. It was hypothesized that the area under the curves and width of the peaks, if analyzed properly, could be used to differentiate and with further research be used to identify each of these organisms. Many methods are currently used to analyze hyperspectral data however no study has been conducted on the veracity of the application used by comparing the method to other available applications. In our study, we used the local maximum method to decrease our dataset while not altering the datasets. By using this method, we were able to determine the center, FWHM, area under the peak as well as the centroid, or mass of gravity, for each of the curves analyzed. These four parameters can then be further analyzed to determine statistical differences between the three organisms at a time point.

21 Hour Data on Salmonella, E. coli and Listeria

Spectral data collected on the three organisms at 21 h over the 400-1000 nm range show broad peaks with sharp rise shoulders, a single maximum between 450-500 nm and varying rates of gradual decay. The sharpest peak was created by *E. coli* while *Listeria* displayed the most gradual decay. Considering all of the organisms produced similar peak maxima, this parameter cannot be reliably used to discriminate between the three organisms. The area under the peaks as well as the width of the curves produced do appear different and can be used to compare the spectra.

Area

The first parameter examined was the area under each of the curves at 21 h-the midpoint value in our study. The area under the curve for *Salmonella* was 211.2±5.8 nm, *E. coli* O157: H7 was

165.8±4.3 nm and 313.7±7.6 nm for *Listeria*. Using these values, a one-way ANOVA was conducted and showed that there was a significant effect of species on the area with a p-value <0.05. The post hoc Tukey HSD test was performed and proved the mean areas under the spectral curves for each of the organisms at 21 hours were significantly different.

FWHM

The full width at half mass value was calculated for each organism at 21 hours: *Salmonella* 278.8 nm±10.3 nm, *E. coli* 158.8nm±10.2 nm, and *Listeria* 437.8nm±15.7 nm. The one-way ANOVA was performed on the datasets and showed there was a significant effect of species on the FWHM with a p-value <0.05. Further, the Tukey HSD test showed the mean FWHM for all three organisms were significantly different from each other.

Center

The center or peak intensity for each organism was analyzed next. Little variation is seen between the values: *Salmonella* 457.8±2.4 nm, *E. coli* 447.0±0.9 nm, and *Listeria* 484.2±3.4 nm. Interestingly, at this parameter there is also a much smaller deviation from the averaged value and shows the importance of evaluating this parameter. The one-way ANOVA test proved there was a significant effect of species in the center for each organism analyzed. The Tukey post hoc HSD test also indicated the averaged center for each organism were significantly different from one another with a p-value <0.05 .

Centroid

Finally, the centroid values were evaluated. The averaged values for *Salmonella* 594.1±1.8 nm, *E. coli* 581.9±2.8 nm, and *Listeria* 654.2±4.6 nm were entered into a one-way ANOVA test to compare the centroids. It was found that there was significant effect on species on Centroids with a p-value <0.05. The post-hoc Tukey HSD test indicated that the Centroid parameter was

sufficient in differentiating *Salmonella* vs *Listeria* and *E. coli* vs *Listeria*; however, it was found that no statistical difference existed between the centroid of *Salmonella* and *E. coli* (p=0.058).

18 Hour Data on Salmonella, E. coli and Listeria

Spectral data collected from the 18 h time point showed noise at the beginning and end of the spectral curves and a peak maximum for each of the three organisms, but this parameter cannot solely be used to discriminate bacteria. The area under the peaks and the width of the peaks are different for each organism and can be used to analyze other parameters that may be useful for discrimination purposes.

Area

The area under the spectral curve was found for each organism and averaged to calculate the mean area under the curve. These values are: *Salmonella* 85.6±13.1 nm, *E. coli* 235.8±33.5 nm and *Listeria* 300.0±14.2 nm. A one-way ANOVA was conducted and determined that species had a significant effect on the Area under the spectral peaks with a p-value <0.05. Tukey HSD post hoc test was then performed and determined that the Area under the curve was significantly different for each organism when compared to one another.

FWHM

The averaged full width at half mass was calculated for each organism and the results were as follows: *Salmonella* 146.0±11.7 nm, *E. coli* 311.1±40.2 nm, and *Listeria* 418.6±1.9 nm. A one-way ANOVA was performed on the mean FWHM hyperspectral data and it was found that species did play a significant effect on the FWHM parameter with a p-value <0.05. The Tukey HSD test was used for post hoc comparisons and determined the mean FWHM was significantly different for each of the organisms.

Center

The average center or peak of each curve was determined for each of the three organisms and the results were as follows: *Salmonella* 446.3±1.2 nm, *E. coli* 445.8±0.9 nm, and *Listeria* 500.7±6.7 nm. A one-way ANOVA was performed for the comparison of Centers of hyperspectral peaks and it was discovered that species had a significant effect on this parameter. Post hoc comparisons were performed using the Tukey HSD test and it was found that the mean Centers for *Salmonella* vs *E. coli* were not significantly different; however, the mean Centers for *Salmonella* vs *Listeria* and *E. coli* vs *Listeria* were significantly different.

Centroid

The mean Centroid for each organism was calculated and the results were as follows: *Salmonella* 564.3±23.1nm, *E. coli* 598.4±12.0 nm, and *Listeria* 615.8±1.1 nm. A one-way ANOVA was performed for the comparison of the Centroids of spectral peaks for each organism and there was no significant effect of species on the Centroid value. Post hoc comparisons were performed using the Tukey HSD test and indicated the mean Centroid for all three organisms were not significantly different (p=0.058)

24 Hour Data on Salmonella, E. coli and Listeria

Spectral data collected at the 24-hour time point display broad peaks with sharp rise shoulders, a single maxima around 450-500nm and show gradual decay with noise at the end of each of the curves. Visually it is difficult to discriminate the three organisms based upon a single maxima but the area under the peaks and width of the peaks vary by organism.

Area

The area under the spectral curve was found for each organism and averaged to calculate the mean area under the curve. These values are: *Salmonella* 106.4.6±8.3 nm, *E. coli* 192.3±7.9 nm and *Listeria* 195.1±15.4 nm. A one-way ANOVA was conducted and determined that species

had a significant effect on the Area under spectral peaks with a p-value <0.05. Tukey HSD post hoc test was then performed and determined that the Area under the curve was significantly different for *Salmonella* vs *E. coli* 0157:H7 and *Salmonella* vs *Listeria* but was not significantly different when comparing the areas under the spectral peaks *E. coli* 0157:H7 vs *Listeria*.

FWHM

The averaged full width at half mass was calculated for each organism and the results were as follows: *Salmonella* 83.5±4.7 nm, *E. coli* 182.9±6.7 nm, and *Listeria* 314.3±17.7 nm. A one-way ANOVA was performed to compare the mean area under hyperspectral FWHM of peaks and it was found that species did play a significant effect on the FWHM parameter with a p-value <0.05. The Tukey HSD test was used for post hoc comparisons and determined the mean FWHM was significantly different for each of the organisms.

Center

The average center or peak of each curve was determined for each of the three organisms and the results were as follows: *Salmonella* 434.8±0.3 nm, *E. coli* 444.4±6.5 nm, and *Listeria* 455.3±1.6 nm. A one-way ANOVA was performed for the comparison of Centers of hyperspectral peaks and it was discovered that species had a significant effect on this parameter. Post hoc comparisons were performed using the Tukey HSD test and it was found that the mean Centers for *Salmonella* vs *E. coli* and *Salmonella* vs *Listeria* were not significantly different; however, the mean Centers for *E. coli* vs *Listeria* were significantly different.

Centroid

The mean Centroid for each organism was calculated and the results were as follows: *Salmonella* 561.8±31.8 nm, *E. coli* 613.9±3.1 nm, and *Listeria* 591.4±11.9 nm. A one-way ANOVA was performed for the comparison of the Centroids of spectral peaks for each organism and there was

no significant effect of species on the Centroid value (p=0.185). Post hoc comparisons were performed using the Tukey HSD test and indicated the mean Centroid for all three organisms were not significantly different (p=0.058)

Spectra of Salmonella, E. coli and Listeria Testing Statistical Methods

Using the Origin Software, 24 different explicit functions were applied to the hyperspectral peak data. Of the 24, six functions converged with our data and are explained in detail within Table VII 37.

Salmonella 21 h

Six statistical models were applied to the *Salmonella* 21 h spectral data and through our analysis it was determined that all models were successful in being applied to the data. The model that served as the best fit would prove to have the lowest AIC and BIC, the highest adjusted R-Sqr, and the lowest Chi-Sqr value. The best model to apply to the *Salmonella* 21 h data was determined to be GCAS followed by Extreme, Lorentz, Logistpk, GaussAmp, and finally InvsPoly as shown in Table VII 56.

***E. coli* 21 h**

The same six statistical models were applied to the *E. coli* 21 h spectral data and through our analysis it was determined that all models were successful in being applied to the data. The model that served as the best fit would prove to have the lowest AIC and BIC, the highest adjusted R-Sqr, and the lowest Chi-Sqr value. The best model to apply to the *E. coli* 21 h data was determined to be GCAS followed by Extreme, Lorentz, Logistpk, GaussAmp, and finally InvsPoly as shown in Table VII 57.

***Listeria* 21 h**

The same six statistical models were applied to the *Listeria* 21 h spectral data and through our analysis it was determined that all models were successful in being applied to the data. The model that served as the best fit would prove to have the lowest AIC and BIC, the highest adjusted R-Sqr, and the lowest Chi-Sqr value. The best model to apply to the *Listeria* 21 h data was determined to be GCAS followed by Extreme, InvsPoly, Lorentz, Logistpk, and finally GaussAmp as shown in Table VII 58.

Comparison of Salmonella, E. coli and Listeria 21 h Statistical Models

GCAS proved to be the best statistical model for all three organisms at this time point. Considering the same model is best for all three we were able to further compare the organisms. This study showed that by using GCAS and a subsequent F-test, the spectral curves produced by *Salmonella* vs *E. coli* and *Listeria* are statistically different with a 0.5 significance level. The degrees of freedom are the same for all three organisms and therefore the higher the F value the higher the degree of discrimination. For the 21 h data sets it was determined that the discrimination level is: *E. coli* vs *Listeria* > *Salmonella* vs *Listeria* > *Salmonella* vs *E. coli* as shown in Table VII 63.

Salmonella 18 h

The best model to apply to the *Salmonella* 18 h data was determined to be GCAS followed by InvsPoly, Extreme, GaussAmp, Logistpk, and finally Lorentz as shown in Table VII 64.

***E. coli* 18 h**

The best model to apply to the *E. coli* 18 h data was determined to be GCAS followed by Extreme, Lorentz, Logistpk, GaussAmp, and finally InvsPoly as shown in Table VII 65.

***Listeria* 18 h**

The same six statistical models were applied to the *Listeria* 18 h spectral data and through our analysis it was determined that only two models were a successful fit for being applied to the data. The best model to apply to the *Listeria* 18 h data was determined to be GCAS followed by Extreme as shown in Table VII 66.

Comparison of Salmonella, E. coli and Listeria 18 h Statistical Models

GCAS proved to be the best statistical model for all three organisms at this time point. Considering the same model is best for all three and we were able to further compare the organisms using Compare Databases Module of Origin 2019. This study showed that by using GCAS and a subsequent F-test, the spectral curves produced by *Salmonella*, *E. coli* and *Listeria* at 18 h are statistically different with a 0.5 significance level. The degrees of freedom are the same for all three organisms and therefore the higher the F value the higher the degree of discrimination. For the 18 h data sets it was determined that the discrimination level is: *Salmonella* vs *Listeria* > *E. coli* vs *Listeria* > *Salmonella* vs *E. coli.*

Salmonella 18 h

The best model to apply to the *Salmonella* 18 h data was determined to be GCAS followed by InvsPoly, Extreme, GaussAmp, Logistpk, and finally Lorentz as shown in Table VII 64.

***E. coli* 18 h**

The best model to apply to the *E. coli* 18 h data was determined to be GCAS followed by Extreme, Lorentz, Logistpk, GaussAmp, and finally InvsPoly as shown in Table VII 65.

***Listeria* 18 h**

The same six statistical models were applied to the *Listeria* 18 h spectral data and through our analysis it was determined that only two models were a successful fit for being applied to the

data. The best model to apply to the *Listeria* 18 h data was determined to be GCAS followed by Extreme as shown in Table VII 66.

Comparison of Salmonella, E. coli and Listeria 18 h Statistical Models

GCAS proved to be the best statistical model for all three organisms at this time point. Considering the same model is best for all three and we were able to further compare the organisms using Compare Databases Module of Origin 2019. This study showed that by using GCAS and a subsequent F-test, the spectral curves produced by *Salmonella*, *E. coli* and *Listeria* at 18 h are statistically different with a 0.5 significance level. The degrees of freedom are the same for all three organisms and therefore the higher the F value the higher the degree of discrimination. For the 18 h data sets it was determined that the discrimination level is: *Salmonella* vs *Listeria* > *E. coli* vs *Listeria* > *Salmonella* vs *E. coli* as shown in Table VII 70 .

Salmonella 24 h

The best model to apply to the *Salmonella* 24 h data was determined to be GCAS followed by, Extreme, Lorentz, Logistpk, GaussAmp, and finally InvsPoly as shown in Table VII 71.

***E. coli* 24 h**

The best model to apply to the *E. coli* 24 h data was determined to be GCAS followed by Extreme, Lorentz, Logistpk, GaussAmp, and finally InvsPoly as shown in Table VII 72.

***Listeria* 24 h**

The best model to apply to the *Listeria* 24 h data was determined to be GCAS followed by Extreme, GaussAmp, Logistpk, and finally Lorentz. Using the InvsPoly model the fit did not converge as shown in Table VII 58.

Comparison of Salmonella, E. coli and Listeria 24 h Statistical Models

GCAS proved to be the best statistical model for all three organisms at this time point. Considering the same model is best for all three and we were able to further compare the organisms using Compare Databases Module of Origin 2019. This study showed that by using GCAS and a subsequent F-test, the spectral curves produced by *Salmonella*, *E. coli* and *Listeria* at 24 h are statistically different with a 0.5 significance level. The degrees of freedom are the same for all three organisms and therefore the higher the F value the higher the degree of discrimination. For the 24 h data sets it was determined that the discrimination level is: *Salmonella* vs *Listeria* > *E. coli* vs *Listeria* > *Salmonella* vs *E. coli* as shown in Table VII 77.

Comparison of Salmonella at 18, 21, and 24 hours of growth

The spectral curves produced by *Salmonella* vary depending on the time point and the changes in the spectra can be examined statistically using Compare Databases Module of Origin. The spectra produced at each time point was shown to be statistically different with a significance level of 0.05. The degrees of freedom are the same for all three *Salmonella* timepoints and therefore the higher the F value the higher the degree of discrimination. For the *Salmonella* data sets it was determined that the discrimination level is: *Salmonella* 18 h vs *Salmonella* 21 h > *Salmonella* 21 h vs *Salmonella* 24 h > *Salmonella* 18 h vs *Salmonella 21* h

Comparison of E. coli at 18, 21, 24 hours of growth

The spectral curves produced by *E. coli* are quite similar at the different time points and was examined statistically using Compare Databases Module of Origin. The spectra produced at each time point are statistically different with a significance level of 0.05. The degrees of freedom are the same for all three *E. coli* timepoints and therefore the higher the F value the higher the degree of discrimination. For the *E. coli* data sets it was determined that the discrimination level is: *E. coli* 18 h vs *E. coli* 21 h > *E. coli* 21 h vs *E. coli* 24 h > *E. coli* 18 h vs *E. coli 21* h

Comparison of Listeria at 18, 21, and 24 hours of growth

The spectral curves produced by *Listeria* are quite similar at the 18 and 21 time points and appear slightly different than the spectra at the 24 h timepoint. The data was examined statistically using Compare Databases Module of Origin. The spectra produced by *Listeria* at 18 and 21 h are not statistically different from each other; however, when comparing the spectral data of *Listeria* at 18 and 24 h as well as *Listeria* at 21 and 24 h statistical differences were found.

Spectra of Salmonella lipopolysaccharide mutants and the wild type at 18 h

Visually, the hyperspectral peaks generated by the lipopolysaccharide mutants and the wild type appear quite similar in the formation of a shark initial peak but the decay for each chemotype varies. Previously explained statistical models were applied to the data and only the Lorentz and Extreme models were effectively applied with Lorentz serving as the best fit as outlined in Table VII 90. The four mutants and wild type spectra were all found to be statistically different with a significance level of 0.05. The degrees of freedom are the same for all chemotypes and therefore the higher the F value the higher the degree of discrimination. For the 18 h *Salmonella* lipopolysaccharide mutant and wild type data sets it was determined that the discrimination level is: *Salmonella* vs Rc(163)> *Salmonella* vs Re(258)> *Salmonella* vs Rd1(231)> *Salmonella* vs Ra (228) as outlined in Table VII 95.

Chapter IX. Summary and Conclusion

In order to properly characterize the bacteria based upon the spectral profiles of each organism, it was necessary to run a controlled experiment to eliminate alterations in the organism's profiles that could be attributed to the media type, buffer solutions, etc. Using different media to grow the organisms and different buffer solutions to wash the cells would likely result in different hyperspectral patterns, as the nutrient sources and the background of the sample would differ. No variations in culturing procedures occurred due to the three organism's ability to propagate under the same conditions. *S. enterica* Typhimurium, *E. coli* and *L. monocytogenes* were cultured, sampled, and imaged under identical methods to ensure uniformity across experiments and to strengthen the validity of the research thus preparing samples suitable for hyperspectral imaging. Further, we developed a technique that can be used to image live, metabolically active, single cell organisms using poly-l-lysine slides to immobilize motile organisms without compromising the cell wall integrity.

It was clear from the spectral data collected that visually the spectra for each organism had their own unique properties and could be analyzed statistically based upon the different areas under the peaks and overall widths of the peaks created. For the 21 h time point, the spectra produced by the three organisms show that species plays an effect on the averaged areas under the curve, FWHM, Center and the Centroid parameter. Tukey post hoc HSD tests determined the averaged area under the curve, FWHM and Centers for each organisms' spectra were significantly different from one another. The Centroid parameter was found to be sufficient in differentiating both *Salmonella* and *E. coli* from *Listeria* but there was no statistical difference between the centroid of *Salmonella* and *E. coli* at the 21 h timepoint. For the 18 h timepoint, it was discovered that only the parameters of area under the curve and FWHM could be used to differentiate the three

organisms, as they were all statistically different. Species did have an effect on the Center parameter for all 3 organisms but the mean centers for *Salmonella* and *E. coli* were not statistically different; however, *Salmonella* and *E. coli* could be differentiated from *Listeria* using this parameter. Species played no significant effect on the Centroid parameter and it was found that the mean Centroid for all three organisms were not statistically different. For the 24 h timepoint, species had a significant effect on the Area, FWHM, and Center parameters for all 3 organisms while for the Centroid parameter species had no significant effect. For the Area under the curves, *Salmonella* vs *E. coli* 0157:H7 and *Salmonella* vs *Listeria* proved to be significantly different from one another but not for *E. coli* 0157:H7 vs *Listeria*. The average FWHM for each organism was determined to be statistically different for all three organisms. It was determined that the mean Centers for *Salmonella* vs *E. coli* and *Salmonella* vs *Listeria* were not significantly different; however, the mean Centers for *E. coli* vs *Listeria* were significantly different. The compilation of our research has proved that our statistical method can successfully be applied to differentiate the three bacteria used in our experiment.

Further, this technology was tested for its ability to display unique spectral curves for cell wall mutants using *Salmonella* lipopolysaccharide mutants from the Salmonella Genetic Stock Center as well as being able to apply statistical models suitable for differentiating between the mutants and the wild type. While visually these curves appeared to be quite similar, using the Lorentz model the scientists were able to differentiate the mutants from the wildtype as well as observe the discrimination level for each comparison.

Chapter X. Future Work

Hyperspectral imaging is an ever-growing field of science that will continue to be researched as this technology provides numerous applications in various industries. Future research in this area should focus on the overall ability to assess the data at hand. Proper statistical analysis is crucial in this area and must be studied further. It is most intriguing that hyperspectral libraries can be created to store the spectral data of an organism, tissue, or any sample and saved for later comparison. Research should include building these libraries so that there is a database established that can be used to identify unknown bacterial samples. We envision that hyperspectral analysis of bacteria will be developed further and techniques will be adapted to clinical isolates. With future development, such a system could be utilized in hospital settings to detect the presence as well as identify the pathogen type. This technology could be an invaluable diagnostic machine if there were ample libraries created for all known pathogens at all stages of growth as well as libraries created for each pathogen in different environments such as blood, saliva, soil, etc. This would undoubtedly take a long time to produce but would be an invaluable tool for the medical profession as rapid detection and subsequent identification is essential to meeting patients' needs and will ultimately save lives.

Ideally, for the area of hyperspectral technology to improve, new statistical models should be developed and especially computerized programs that would allow for even more rapid analysis of sample. The ability to share libraries would be essential for this to happen.

References

1. Alexandrakis, D., Downey, G., and Scannell, A.G. (2012). Rapid non-destructive detection of spoilage of intact chicken breast muscle using near-infrared and Fourier transform mid-infrared spectroscopy and multivariate statistics. *Food Bioprocess Technology 5*(1): 338– 47.
2. Amrutha B, Sundar, K., Shetty, P.H. (2017). Study on E. coli and Salmonella biofilms from fresh fruits and vegetables. *Food Science and Technology 54*(5):1091-1097.
3. Bottini, G., Losito, F., De Ascenti, A., Priolisi, F.R., Mari, A., Antonini, G. (2011). Validation of the micro biological survey method for total viable count and *E. coli* in food samples. *American Journal of Food Technology 6*: 951-962.
4. Braden, C.R. (2006). Salmonella enterica serotype Enteritidis and eggs: a national epidemic in the United States. *Clinical Infectious Diseases 43*(4):512-517.
5. Brahmbhatt, H.N., Wyk, P., Quigley, N.B. and Reeves, P.R. (1988). Complete PhysicalMap oftherfbGene ClusterEncoding Biosynthetic Enzymes for the 0 Antigen of Salmonella typhimuriumLT2. *Journal of Bacteriology* 170(1): 98-102. http://jb.asm.org/content/170/1/98.full.pdf+html
6. Braun, V. (1975) Covalent lipoprotein from the outer membrane of Escherichia coli. *Biochimica et Biophysica Acta (BBA) - Reviews on Biomembranes* 415 (3): 335-377. https://doi.org/10.1016/0304-4157(75)90013-1.
7. Bryan, F. L. and Doyle, M.P. (1995). Health risks and consequences of Salmonella and Campylobacter jejuni in raw poultry. *Journal of Food Protection 58*(3):326-344.
8. Bucur, F. I., Grigore-Gurgu, L., Crauwels, P., Riedel, C. U., Nicolau, A. I. (2018). Resistance of Listeria monocytogenes to Stress Conditions Encountered in Food and

Food Processing Environments. *Frontiers in Microbiology 9*:2700. doi:10.3389/fmicb.2018.02700

9. Centers for Disease Control and Prevention. CDC outbreak database. October (2012).http://www.outbreakdatabase.com/reports/Salmonella_Newport_Linked_to_Cantaloupe,_CDC_Final_Update.pdf
10. Croxen, M. A., Law, R. J., Scholz, R., Keeney, K. M., Wlodarska, M., & Finlay, B. B. (2013). Recent advances in understanding enteric pathogenic Escherichia coli. *Clinical Microbiology Reviews 26*(4): 822–880. doi:10.1128/CMR.00022-13
11. Cytoviva, Inc. User Manual. https://cytoviva.com/wp-content/documents/hyperspectral/products/HSI-User-Manual-8-11-11.pdf
12. Dalebroux, Z.D., Matamouros, S., Whittington, D., Bishop, R.E. and Miller, S.I. (2014). PhoPQ regulates acidic glycerophospholipid content of the *Salmonella* Typhimurium outer membrane. *PNAS 111*(5): 1963-1968. http://www.pnas.org/content/111/5/1963.short
13. Depelteau, J. S., Brenzinger, S., and Briegel, A. (2019). Bacterial and Archaeal Cell Structure. Reference Module in Life Sciences. *Elsevier* https://doi.org/10.1016/B978-0-12-809633-8.20679-1
14. Dondi, F., Betti, A., Blo, G., and Bighi, C. (1981). Statistical analysis of gas chromatographic peaks by the Gram-Charlier series of type A and the Edgeworth-Cramer series. *Analytical Chemistry 53*(3):496-504.
15. Duong, F., Eichler, J., Price, A., Leonard, M. R., and Wickner, W. (1997). Biogenesis of the Gram-Negative bacterial envelope. *Cell 91*(5): 567-573. http://www.sciencedirect.com/science/article/pii/S0092867400804444?np=y

16. (2015). Etymologia: Escherichia coli. *Emerging Infectious Diseases 21*(8), 1310. https://dx.doi.org/10.3201/eid2108.et2108.

17. Fàbrega, A., & Vila, J. (2013). Salmonella enterica serovar Typhimurium skills to succeed in the host: virulence and regulation. *Clinical Microbiology Reviews 26*(2), 308–341. doi:10.1128/CMR.00066-12

18. Fakruddin, M.D., Sultana, M., Ahmed, M.M., Chowdhury, A., and Choudhury, N. (2013). Multiplex PCR (polymerase chain reaction) assay for detection of E. coli, Salmonella sp., Vibrio cholerae Vibrio parahaemolyticus in spiked shrimps (Penaeus monodon). *Pakistan Journal of Biological Sciences 16*: 267-274.

19. Gal-Mor, O., Boyle, E. C., & Grassl, G. A. (2014). Same species, different diseases: how and why typhoidal and non-typhoidal Salmonella enterica serovars differ. *Frontiers in Microbiology 5*:391. doi:10.3389/fmicb.2014.00391

20. Gowen, A.A., Feng, Y., Gaston, E., and Valdramidis, V. (2015). Recent applications of hyperspectral imaging in microbiology. *Talanta 137*:43-54. https://doi.org/10.1016/j.talanta.2015.01.012.

21. Gumbel E.J. (1941). The Return Period of Flood Flows. *The Annals of Mathematical Statistics 12*(2):163-190.

22. Hagens, S. and Loessner, M.J. (2007). Application of bacteriophages for detection and control of foodborne pathogens. *Applied Microbiology Biotechnology 76*: 513-519.

23. Heinrichs, D.E., Yethon, J.A., and Whitfield, C. (1998). Molecular basis for structural diversity in the core regions of lipopolysaccharides of *Escherichia coli* and *Salmonella* enterica. *Molecular Microbiology 30*(2): 221-232. http://onlinelibrary.wiley.com/doi/10.1046/j.1365-2958.1998.01063.x/epdf

24. Hernandez-Milian, A., & Payeras-Cifre, A. (2014). What is new in listeriosis? *BioMed Research International* Article ID 358051. https://doi.org/10.1155/2014/358051

25. Huang, H., Liu, L., Ngadi, M.O. (2014). Recent developments in hyperspectral imaging for assessment of food quality and safety. *Sensors 14*(4): 7248– 76.

26. Ishino, S., & Ishino, Y. (2014). DNA polymerases as useful reagents for biotechnology - the history of developmental research in the field. *Frontiers in Microbiology 5*:465. doi:10.3389/fmicb.2014.00465

27. Islam, M., Morgan, J., Doyle, M.P., Phatak, S.C., Millner, P., and Jiang, X. (2004). Persistence of Salmonella enterica Serovar Typhimurium on lettuce and parsley and in soils on which they were grown in fields treated with contaminated manure composts or irrigation water. *Foodborne Pathogens and Disease 1*(1): 27-35. https://doi.org/10.1089/153531404772914437

28. Janakiraman V. (2008). Listeriosis in pregnancy: diagnosis, treatment, and prevention. *Reviews in Obstetrics & Gynecology*, *1*(4): 179–185.

29. Kamio, Y. and Nikaido, H. (1976). Outer membrane of Salmonella typhimurium: accessibility of phospholipid head groups to phospholipase C and cyanogen bromide activated dextran in the external medium. *Biochemistry 15*(12): 2561-2570 doi: 10.1021/bi00657a012

30. Keerthirathne, T. P., Ross, K., Fallowfield, H., & Whiley, H. (2016). A Review of Temperature, pH, and Other Factors that Influence the Survival of Salmonella in Mayonnaise and Other Raw Egg Products. *Pathogens 5*(4): 63. doi:10.3390/pathogens5040063

31. Kirkland, L., Herr, K., Keim, E., Adams, P., Salisbury, J., Hackwell, J., Treiman, A. (2002). First use of an airborne thermal infrared hyperspectral scanner for compositional mapping. *Remote Sensing of Environment 80*(3): 447-459. https://doi.org/10.1016/S0034-4257(01)00323-6

32. Law, J. W.-F., Ab Mutalib, N.-S., Chan, K.-G., & Lee, L.-H. (2014). Rapid methods for the detection of foodborne bacterial pathogens: principles, applications, advantages and limitations. *Frontiers in Microbiology 5*:770. http://doi.org/10.3389/fmicb.2014.00770

33. Lim, J. Y., Yoon, J., & Hovde, C. J. (2010). A brief overview of Escherichia coli O157:H7 and its plasmid O157. *Journal of Microbiology and Biotechnology 20*(1), 5–14.

34. Loessner, M. J., Rudolf, M. and Scherer, S. (1997). Evaluation of luciferase reporter bacteriophage A511:luxAB for detection of Listeria monocytogenes in contaminated foods. *Applied Environmental Microbiology 63*(8): 2961-2965.

35. Losito, F., Bottini, G., De Ascentis, A., Priolisi, F.R., Mari, A., Tarsitani, G. and Antonini, G. (2012). Qualitative and quantitative validation of the micro biological survey method for Listeria spp., Salmonella spp., enterobacteriaceae and Staphylococcus aureus in food samples. *American Journal of Food Technology 7*:340-351.

36. Mandal, P.K., Biswas, A.K., Choi, K. and Pal, U.K. (2011). Methods for Rapid Detection of Foodborne Pathogens: An Overview. *American Journal of Food Technology, 6*: 87-102. doi: 10.3923/ajft.2011.87.102

37. McGrath, B.C. and Osborn, M.J. (1991). Localization of the terminal steps of O-Antigen Synthesis in Salmonella typhimurium. *Journal of Bacteriology 173*(2): 649-654. http://jb.asm.org/content/173/2/649.full.pdf+html

38. Mohammed E.A., Naugler, C., Far, B.H. (2015). Chapter 32 - Emerging Business Intelligence Framework for a Clinical Laboratory Through Big Data Analytics. In: *Emerging Trends in Computational Biology, Bioinformatics, and Systems Biology.* Edited by Tran Q.N., Arabnia H. Boston: Morgan Kaufmann: 577-602.

39. Murray, G.L., Attridge, S.R. and Morona, R. (2006). Altering the length of lipopolysaccharide O antigen has an impact on the interaction of Salmonella enterica Serovar Typhimurium with macrophages and complement. *Journal of Bacteriology 188*(7): 2735-2739. http://jb.asm.org/content/188/7/2735.full.pdf+html

40. Murray, G.L., Attridge, S.R. and Morona, R. (2003). Regulation of Salmonella typhimurium lipopolysaccharide O antigen chain length is required for virulence; identification of FepE as second Wzz. *Molecular Microbiology 47*(5): 1395-1406.

41. Nanduri, V., Sorokulova, I.B., Samoylov, A.M., Simonian, A.L., Petrenko, V.A., and Vodyanoy, V.J. (2007). Phage as a molecular recognition element in biosensors immobilized by physical adsorption. *Biosensors and Bioelectronics 22*(6): 986-992. http://www.sciencedirect.com/science/article/pii/S0956566306002193

42. Nugen, SR, Baeumner, A. (2008). Trends and opportunities in food pathogen detection. *Analytical and Bioanalytical Chemistry 391*(2): 451– 54.

43. Olsen, E.V., Sorokulova, I.B., Petrenko, V.A., Chen, I., Barbaree, J.M., and Vodyanoy, V.J. (2006). Affinity-selected filamentous bacteriophage as a probe for acoustic wave biodetectors of Salmonella typhimurium. *Biosensors and Bioelectronics 21*(8): 1434-1442. http://www.sciencedirect.com/science/article/pii/S0956566305001831#

44. Oporto B. E., Esteban J. I., Aduriz, G., Juste, R.A., and Hurtado, A. (2008). Escherichia coli O157:H7 and Non-O157 Shiga Toxin-producing E. coli in Healthy Cattle, Sheep and Swine Herds in Northern Spain. *Zoonoses Public Health 55*:73-81.

45. OriginLab_Corporation: Origin(Pro). (2019). Version 2019b, Northampton, MA, USA. Inc.

46. Osborn, M.J., Gander, J.E., Parisi, E., and Carson, J. (1972). Mechanism of assembly of the outer membrane of Salmonella typhimurium. *Journal of Biological Chemistry 247*(12): 3962-3972

47. Petrenko, V.A. and Vodyanoy, V.J. (2003). Phage display for detection of biological threat agents. *Journal of Microbiological Methods 53*: 253-262. http://ac.els-cdn.com/S0167701203000290/1-s2.0-S0167701203000290-main.pdf?_tid=3be2a6dc-172a-11e4-b89a00000aacb360&acdnat=1406643268_e53632aaba2af2be3819bb8d00407528

48. Pfalzgraff, A., Wilmar, C., Heinbockel, L., Schromm, A.B., Lübow, C., Gisch, N., Martinez-de-Tejada, G., Brandenburg, K., Weindl,G. (2019). LPS-neutralizing peptides reduce outer membrane vesicle-induced inflammatory responses. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids* https://doi.org/10.1016/j.bbalip.2019.05.018.

49. Pezzullo JC, Merser S, Miller B, Weiner K. (2016). Analysis of Variance from Summary Data. http://statpages.info/anova1sm.html.

50. Qureshi, N., Takayama, K., Ribi, E. (1982). Purification and structural determination of non-toxic lipid A obtained from lipopolysaccharide of *Salmonella typhimurium*. *Journal of Biological Chemistry 257*: 11808-11815.

51. Raetz, C.R.H., Guan, Z., Ingram, B.O., Six, D.A., Song, F., Wang, X. and Zhao, J. (2009). Discovery of new biosynthetic pathways: the Lipid A Story. *Journal of Lipid Research 50*(Suppl): S103-S108. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2674688/

52. Raetz, C.R.H., Reynolds, C.M., Trent, M.S., and Bishop, R.E. (2008). Lipid A Modification Systems in Gram-Negative Bacteria. *Annual Review of Biochemistry 76*: 295-329. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2569861/

53. Raetz. C.R.H., and Whitfield, C. (2002). Lipopolysaccharide endotoxins. *Annual Review of Biochemistry 71*: 635-700. http://cell.ccrc.uga.edu/~rcarlson/Raetz.pdf

54. Rich, P.D., Fung, L.W., Ho, C., Osborn, M.J. (1977). Lipid A mutants of Salmonella typhimurium. Purification and characterization of a lipid A precursor produced by a mutant in 3-deoxy-n-mannooctulosonate-8-phosphate synthetase. *Journal of Biological Chemistry 252*: 4904-4912.

55. Rietschel, E.T., Horst-Werner, W., Zahringer, U., and Luderitz, O. (1982). Lipid A, the lipid component of bacterial lipopolysaccharides: Relation of chemical structure to biological activity. *Journal of Molecular Medicine 60*(14): 705-709.

56. Scallan, E., Hoekstra, R. M., Angulo, F. J., Tauxe, R. V., Widdowson, M. A., Roy, S. L., Griffin, P. M. (2011). Foodborne illness acquired in the United States--major pathogens. *Emerging Infectious Diseases 17*(1): 7–15. doi:10.3201/eid1701.P11101

57. Silhavy, T. J., Kahne, D. and Walker, S. (2010). The Bacterial cell envelope. *Cold Spring Harbor Perspectives of Biology* http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2857177/

58. Smith, G.P. (1985). Filamentous fusion phage: Novel expression vectors that display cloned antigens on the virion surface. *Science 228*: 1315-1317.

59. Smith, G.P., Petrenko, V.A. (1997). Phage display (Review). *Chemical Reviews 97*: 391-410

60. Solanki, R., Vanjari, L., Subramanian, S., B. A., E. N., & Lakshmi, V. (2014). Comparative Evaluation of Multiplex PCR and Routine Laboratory Phenotypic Methods for Detection of Carbapenemases among Gram Negative Bacilli. *Journal of Clinical and Diagnostic Research 8*(12): DC23–DC26. doi:10.7860/JCDR/2014/10794.5322

61. Sone, I., Olsen, R.L., Dahl, R., and Heia, K. (2011). Visible/near-infrared spectroscopy detects autolytic changes during storage of Atlantic salmon (Salmo salar L.). *Journal of Food Science 76*(3): S203– 09.

62. Steele, M. and Odumeru, J. (2004). Irrigation water as a source of foodborne pathogens on fruit and vegetables. *Journal of Food Protection 67* (12): 2839-2849. https://doi.org/10.4315/0362-028X-67.12.2839

63. Swaminathan, B., Barrett, T. J., Hunter, S. B., Tauxe, R. V., & CDC PulseNet Task Force. (2001). PulseNet: the molecular subtyping network for foodborne bacterial disease surveillance, United States. *Emerging Infectious Diseases 7*(3): 382–389. doi:10.3201/eid0703.010303

64. U.S. Environmental Protection Agency. (2012). *Microbial Risk Assessment Guideline:Pathogenic Microorganisms with Focus on Food and Water;* U.S. Environmental Protection Agency, Interagency Microbial Risk Assessment Guideline Workgroup. EPA/100/J-12/001.

65. Vainrub, A. Pustovyy, O. and Vodyanoy, V. (2006). Resolution of 90 nm (λ/5) in an optical transmission microscope with an annular condenser. Opt. Lett. 31: 2855-2857

66. Voetsch, A. C., Frederick J., Angulo, F. J., Rabatsky-Ehr, T., Shallow, S., Cassidy, M., Thomas, S. M., Swanson, E., Zansky, S. M., Hawkins, M. A., Jones, T. F., Shillam, P. J., Van Gilder, T. J., Wells, J. G., Griffin, P.M, for the Emerging Infections Program FoodNet Working Group. (2004). Laboratory Practices for Stool-Specimen Culture for Bacterial Pathogens, Including *Escherichia coli* O157:H7, in the FoodNet Sites, 1995–2000. *Clinical Infectious Diseases 38*(3,15): S190–S197. https://doi.org/10.1086/381586

67. Wagenmakers E-J, Farrell S. (2004). AIC model selection using Akaike weights. *Psychonomic Bulletin & Review 11*(1):192-196.

68. Walck C. (2007).Hand-book on statistical distributions for experimentalists. Stockholm: University.

69. Yoshii, M., Okamoto, A. and Ohta, M. (2013). Lipopolysaccharide-deficient mutants of Salmonella enterica have increased sensitivity to catechins. *Advances in Microbiology 3*(5): 399-402. http://file.scirp.org/Html/1-2270197_35906.htm

70. Zhang, G., Meredith, T. C., & Kahne, D. 2013. On the essentiality of lipopolysaccharide to Gram-negative bacteria. *Current Opinion in Microbiology 16*(6), 779–785. doi:10.1016/j.mib.2013.09.007

71. Zhou B., Xiao J., Liu S., Yang J., Wang Y., Nie F., et al. (2013). Simultaneous detection of six food-borne pathogens by multiplex PCR with GEXP analyzer. *Food Control 32*: 198–204. doi: 10.1016/j.foodcont.2012.11.044