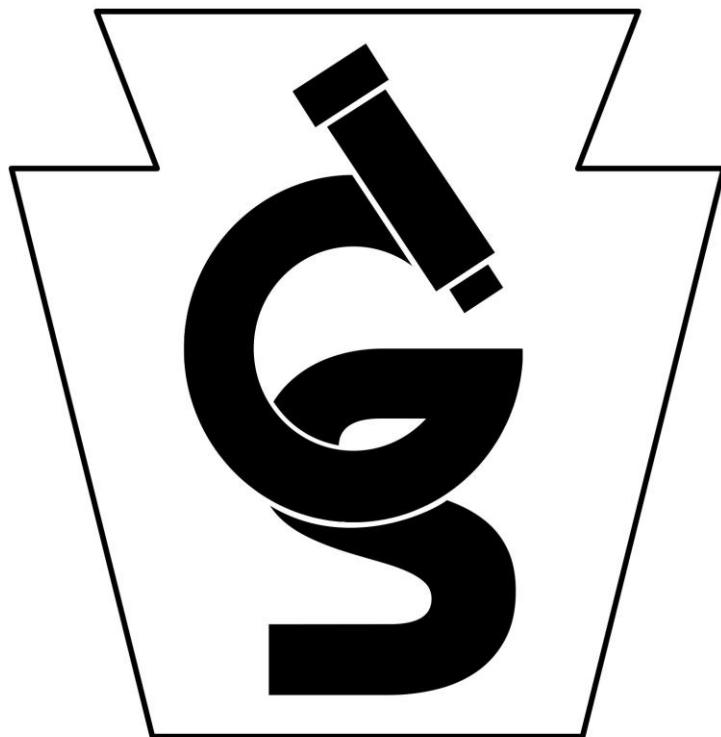


**Journal of the  
Pennsylvania Governor's School for the Sciences**

Class of 2022



Volume 37, 2022

Mellon College of Science  
Carnegie Mellon University, Pittsburgh, PA 15213



Journal  
of the  
Pennsylvania Governor's School for the Sciences

*Class of 2022*

**Volume 37, 2022**

**Copyright (c) 2022, by The Pennsylvania Governor's School for the Sciences  
and Carnegie Mellon University**

**Permission** is granted to quote from this journal with the customary acknowledgement of the source. To reprint a figure, table or other excerpt, requires, in addition, the consent of one of the original authors and notification of the PGSS.



**The Pennsylvania Governor's School for the Sciences**

**In cooperation with**



**Major funding provided by**



**(Alumni, Parents and Friends of PGSS)**

**Corporate Partners:**



**UPMC HEALTH PLAN**



### **Educational Partners:**

Pittsburgh Public Schools  
Lancaster-Lebanon Intermediate Unit (IU 13)  
Selinsgrove Area School District  
Central Susquehanna Intermediate Unit (IU 16)  
Loyalsock Township School District  
Chester County Intermediate Unit (IU 24)  
The School District of Philadelphia  
CCRES, Inc.

The PGSS program and its director  
gratefully acknowledge the support of  
Carnegie Mellon University and the  
**Mellon College of Science.**

# **Carnegie Mellon University**

## Table of Contents

Preface .....	1
Acknowledgements .....	2

### Biology Team Projects

<b>Modeling Effects of SPF on Ultraviolet Radiation Induced DNA Damage Using <i>Saccharomyces cerevisiae</i></b>	
Fennick, Fick, Ko, Krimin, Levin, Lyn-Sue, Palle, Tu, A. Wang .....	5
<b>Host-Microbiota Interactions in <i>C. elegans</i> Disease Models</b>	
Chinnasamy, Devlin, Goldsmith, Lin, McBride, Olson, Pandula, Riddell, C. Smith .....	47

### Biology / Computational Biology Team Projects

<b>Who's That Bacteria? A Pipeline for Computational Analysis of DNA Fingerprints</b>	
Barlow, Mauriello, Strickland, Zapotocky .....	71
<b>Identifying Lactic Acid Bacteria Through Comparison of 16S Genetic Sequence</b>	
Green, Matich, Reber, Tang, D. Wang, K. Zhang.....	101

### Chemistry Team Project

<b>HPLC Analysis of Active Ingredient Mass in Expired and Unexpired OTC Analgesics</b>	
Guo, Jain, Khan, León, Lynch, Magdeburg, Niebauer, Noon, Qu, G. Tout-Puissant, T. Tout-Puissant, C. Zhou .....	119

### Computer Science Team Projects

<b>Reversing Reversi: Using Recursion and Minimax Algorithms to Master Turn-Based Game Strategy</b>	
Hemminger, Hundiwala, Rastogi, Ricafort.....	143
<b>Artificial Intelligence Architecture for Turn-Based Strategy Games: Nine Men's Morris</b>	
Cao, Cui, S. Wang, D. Zhang .....	167

### Mathematics Team Project

<b>Applying Stochastic Processes to Decision Making</b>	
Bryant, Gordon, Hasker, Quinodoz, Sparling, Strocko, Tandon .....	191

**Physics Team Projects**

<b>Raman Spectroscopy: The Use of Inelastic Light Scattering to Compare Models in Classical Physics to Quantum Mechanics</b>	
Blocher, Gombos,.....	<b>229</b>
<b>Determining Particle Diameter With Elastic Light Scattering and Lasers</b>	
Gong, Kenning, Meigs.....	<b>249</b>
<b>Optical Tweezers</b>	
DeWilde, Mejia-Toro, Weller .....	<b>261</b>
<b>The Behavior of Magnetic Moments at Low Temperature</b>	
Krakauskas, Olsen, Solai, Vizza, Wehbe, Wehler .....	<b>277</b>
<b>Behind the Beauty of the Carina Nebula: Star Formation, Ionization Fronts and Interstellar Reddening</b>	
Rajesh, Wilson, E. Zhang.....	<b>307</b>

**Appendix**

The Students, Faculty, and Teaching Assistants/Counselors of the Pennsylvania Governor's School for the Sciences for 2022 .....	<b>345</b>
--	------------

## Preface

The Pennsylvania Governor's School for the Sciences (PGSS) is a five-week summer residential program for talented Pennsylvania high school students. The PGSS was initiated by the Pennsylvania Department of Education (PDE) in 1982 and is hosted by the Mellon College of Science on the Pittsburgh campus of Carnegie Mellon University.

The PGSS class of 2022 consisted of 72 students – 36 male and 36 female – chosen from a pool of 317 applicants from all over the Commonwealth of Pennsylvania. The 37th session of PGSS was conducted in 2022 from June 26 through July 30.

The PGSS academic program consists of five key components:

- Core Courses in biology, chemistry, computer science, mathematics and physics provide a common educational experience for all students.
- Elective Courses, which vary from year to year, may be chosen according to the interest of each individual student.
- Formal Laboratory Courses are offered in biology, chemistry, computer science, mathematics, and physics. Each student chooses one formal laboratory course.
- Panel discussions on topics of current interest are required of all students.
- Team Research Projects are offered on a wide variety of topics. Each student chooses one project and works with a team of students under the supervision of a faculty member.

The PGSS Team Research Projects involve the investigation of an original problem or the solution of a problem by techniques original to the student investigators. While a faculty member is available for initial direction and advice, most of the accomplishments come from the students' own initiative.

The Journal of the Pennsylvania Governor's School for the Sciences reports the results of the students' efforts and is the official record of these team research projects. Each team investigated their chosen problem using resources and techniques appropriate for that topic. The following reports were written by the student team members. The faculty advisor for each project reviewed the final report and made necessary minor corrections.

This journal was edited by Melissa Lessure, Assistant to the Director of PGSS, who assumed the responsibility of converting all the reports to a similar format, reproducing figures, and dealing with other publication issues. Otherwise, the original character of the student-authored papers was maintained to the maximum feasible extent.

## **Acknowledgements**

The Pennsylvania Governor's School for the Sciences was supported this year by grants from the PGSS Campaign, Inc. (Alumni, Parents and Friends of PGSS) and the Pennsylvania Department of Education. We are particularly grateful this year for donations from our corporate sponsors: Air Products, EQT Foundation, UPMC Health Plan, and Weis Markets. The PGSS gratefully acknowledges the continuing support of the Mellon College of Science at Carnegie Mellon University and generous contributions from the parents and guardians of the PGSS Class of 2022.

Barry B. Luokkala, Ph.D.  
Program Director of the Pennsylvania Governor's School for the Sciences  
October 2022

**BIOLOGY  
TEAM PROJECTS**



# **Modeling Effects of SPF on Ultraviolet Radiation Induced DNA Damage Using *Saccharomyces cerevisiae***

Brea R. Fennick, Owen T. Fick, Peter S. Ko, Olivia M. Krimin, Daniel S. Levin,  
Rachel E. Lyn-Sue, Abhinav S. Palle, Daniel R. Tu, Amy Wang

## **Abstract**

Skin cancer is a prevalent concern in the entire world. Skin cancer typically arises as a result of repeated overexposure to ultraviolet radiation (UVR) without proper skin protection. There are three main types of UVR: UVC, UVB, and UVA. UVB is an important causal factor of skin cancers and sunburns. Since sunburns can cause mutations in skin cell genomes, sunscreen is recommended to protect against the risks that UVR poses. This project tests the efficacy of various SPF-containing products in protecting against the damaging effects of UVR using the model organism *Saccharomyces cerevisiae*. Cell viability assays show that *S. cerevisiae* exposed to UVR without protection died within three minutes. An organic or inorganic chemical makeup of the sunscreen appears to offer the same level of protection, higher SPF values tend to provide more defense against UVR, and clear sunscreen has higher efficacy than its opaque counterpart. Nucleosome fragmentation assays reveal the extent of apoptosis triggered by cell stress. DNA shearing increases with UV exposure, indicating a heightened activation of programmed cell death pathways in *S. cerevisiae*. Indirect immunofluorescence microscopy suggests that an SPF coating prevents UV-induced mutations. These results confirm that UVR damages DNA and sunscreen protects cells from UVR.

## **I. Introduction**

### **A. Ultraviolet Radiation**

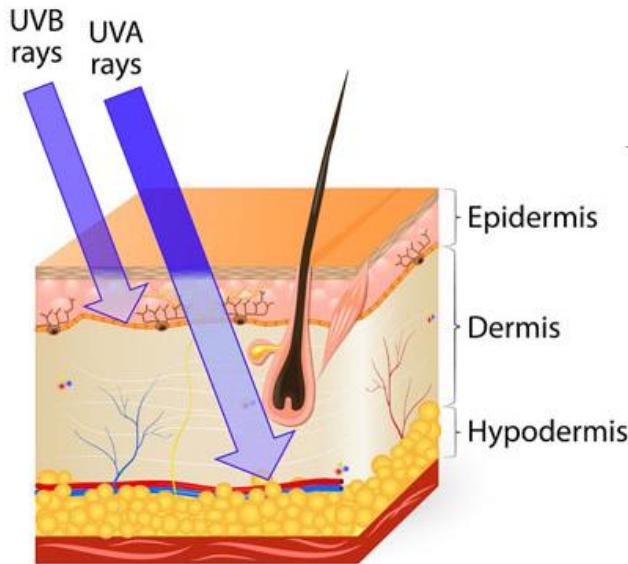
#### **1. Sources**

Ultraviolet radiation (UVR) is emitted by the sun, tanning beds, welding torches, certain types of lasers, mercury vapor lighting (which can be found in stadiums and school gyms), and halogen and fluorescent lighting.<sup>1,2</sup> The concentration of UVR exposure varies based on time of day, geographical location, and the environment.<sup>1</sup> For example, UV radiation is strongest between 10am and 4pm, when the sun's rays are most concentrated on the earth.<sup>1,3</sup> UV rays become stronger closer to the equator and during spring and summer.<sup>1</sup> Additionally, higher elevations increase the concentration of UV exposure.<sup>1,3</sup> Furthermore, since clouds do not block UVR entirely, UVR can still penetrate one's skin when it's cloudy or overcast.<sup>1,3</sup> UVR can be scattered by incomplete cloud cover, and the radiation that reaches the earth can still reflect off of other surfaces.<sup>1,3</sup> Similarly, reflection off of other objects, such as water or sand, can also result in sun exposure.<sup>1,3</sup> Consequently, individuals in the shade can still be harmed by UV rays as a result of reflected surfaces.<sup>1,3</sup>

#### **2. Wavelengths, Types, and Interactions of UVR**

Radiation is when a source gives off energy in the form of a wave that is invisible to the human eye.<sup>1</sup> The energy spectrum ranges from sources that give off low energy waves, like radio waves, to high energy waves, like x-rays.<sup>1</sup> Low energy waves are typically longer in their wavelength, with their peaks spaced farther apart.<sup>1</sup> High energy waves are the opposite, with their wavelength peaks spaced closer together.

UV radiation falls in the middle of the spectrum, with more energy than visible light, which ranges from 380 to 700 nm, but less than x-rays, which ranges from 0.01 to 10 nanometers.<sup>1,4,5</sup>



**Figure 1: UVB and UVA Skin Penetration<sup>6</sup>**

Figure 1 illustrates the layers of the skin and which type of UVR each layer absorbs.

There are three categories of UV radiation: UVA, UVB, and UVC.<sup>1,2</sup> UVA has the longest wavelength (315 - 400 nm), carries the least amount of energy, and is not absorbed by the earth's ozone layer; therefore, all of its energy reaches the earth's surface.<sup>1,2</sup> This wavelength's longer, lower energy grants it the ability to penetrate into the dermis, the second layer of the skin (Figure 1).<sup>7</sup> Consequently, UVA can cause various problems in the dermis, such as premature skin wrinkling.<sup>1</sup> UVB has energy levels between UVA and UVC (280 - 314 nm). Only some of its energy reaches the surface of the earth because the majority of it is absorbed by the ozone layer.<sup>1,2</sup> Very little of this wavelength penetrates the dermis, as most of it is absorbed by the epidermis (Figure 1).<sup>7</sup> UVB is considered the most damaging of the UVRs that reach the earth's surface due to its higher energy.<sup>8</sup> It's also typically responsible for the development of sunburns as well as a large portion of skin cancers.<sup>8</sup> UVC rays are made up of the shortest wavelength (100 - 279 nm) of the three UVRs, and thus carry the most energy.<sup>1,2,6</sup> Ozone gas in the ozone layer absorbs UVC entirely. This interaction splits the ozone molecules apart into an oxygen molecule ( $O_2$ ) and an oxygen atom, effectively absorbing the energy from the UVC rays.<sup>9</sup> As a result, none of UVC's energy reaches the earth's surface. Thus, humans are not exposed to UVC ray's energy.<sup>7</sup>

## B. Impact of UVR on Skin

When considering the impact of UV radiation on humans, the primary bodily organ in consideration is the skin. There are three main layers to the skin, the epidermis, dermis, and hypodermis.<sup>7</sup> The epidermis is the thinnest and outermost layer, and the body's first line of defense against the outside world.<sup>10</sup> New skin cells are produced here as well as melanin, which gives skin its color.<sup>7</sup> The dermis is the second layer of the skin and the thickest of the three layers.<sup>11</sup> The dermis contains oil glands, which help control body temperature, and sebaceous glands, which protect the skin from drying out. It also contains hair follicles, the areas in the

skin from which hair grows.<sup>11</sup> The hypodermis is the innermost layer of the skin, which attaches the dermis to other muscles and tissues. It functions to maintain body temperature, and contains fat tissue, which act as insulators.<sup>12</sup>

Previous papers have detailed the numerous effects of UVR on human skin.<sup>13,14, 15</sup> These include sunburn, premature aging of the skin, the development of skin cancer, and various other environmentally-influenced skin disorders.<sup>14</sup> However, at the same time, UV radiation has been shown to benefit human health through the mediation of the synthesis of vitamin D and endorphins in the skin.<sup>16</sup> Nevertheless, despite the beneficial effects of UV on human health, excessive exposure to UV is associated with serious risks.

Humans react to UV exposure in various ways across different populations. Because the human epidermis serves as the medium of contact through which the body is exposed to its environment, the phenotypic makeup of its biological characteristics serve a crucial role in the body's response to external stressors such as UV radiation. The most abundant type of cell in the epidermis is keratinocytes. The primary purpose of keratinocytes is to form a physical barrier within the epidermis, which is highly effective in serving its role in protecting the body from most stressors. Additionally, keratinocytes accrue melanin pigments as they mature, which functions to block UV penetration into skin.<sup>17</sup> When UV light hits melanin, instead of becoming reactive it releases the extra energy as heat.<sup>81</sup> Thus, melanin in keratinocytes is thought to function as a type of natural sunscreen. Because of this interaction, the higher the melanin concentration in the skin, the more UV light is shielded from the rest of the epidermis and the cells underneath. Genetic factors influence the amount and concentration of melanin produced in melanocytes.<sup>82,83</sup> Thus, the extent to which UV radiation penetrates skin is largely determined by genetic factors.

With this in consideration, it is natural to assume that genetic factors also play an important role in the development of UV-mediated skin diseases. Fair-skinned people are more likely to be UV sensitive in comparison to darker-skinned individuals, due to the relative lack of melanin pigment in the epidermis. The primary determinant of skin complexion is the concentration of the specific melanin type, eumelanin, in the epidermis. In comparison to other types of melanin found in the body, eumelanin is the most effective at absorbing energy from UV radiation.<sup>84</sup> Thus, there is a strong correlation between skin complexion and protection against UV radiation and susceptibility to UV-mediated skin diseases. Accordingly, a previous study investigated "Minimal erythematous dose" (MED), which is a quantitative method to determine the minimum amount of UV required to induce sunburn in human skin.<sup>18</sup> They demonstrated that the fairer the skin, the lower the MED, while the darker the skin, the more UV radiation was required to cause sunburn.<sup>18</sup>

The process of melanogenesis—pigment generation—and tanning occurs in the epidermis with the second most abundant cell type, melanocytes. When melanocytes are exposed to UVB radiation, melanocytes produce melanin pigment, which travels through the epidermis and is absorbed by other skin cells.<sup>19</sup> When this melanin is exposed to UVA light, it oxidizes, causing it to darken.<sup>19</sup> As previously discussed, this darkening process is the epidermis's primary way of protecting itself against excessive UV exposure.

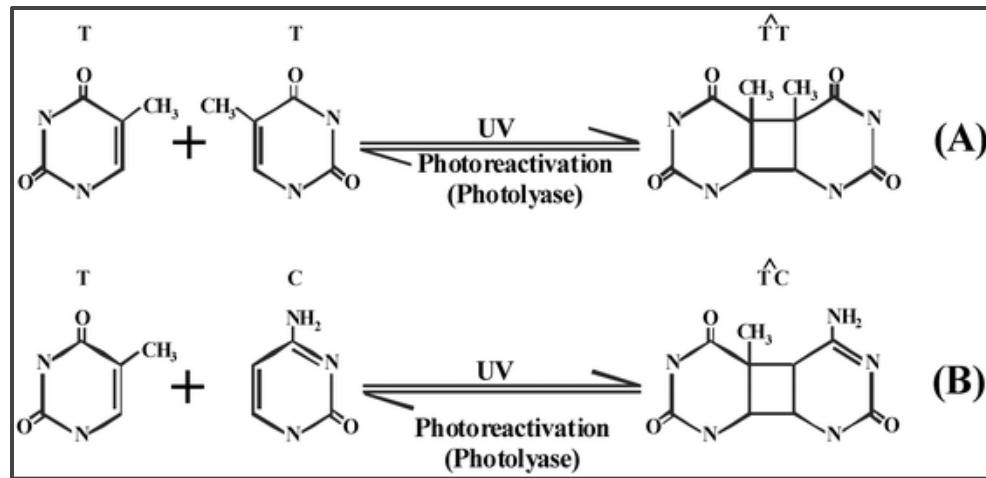
The interactions between the human epidermis and UV light are also responsible for the production of vitamin D. Humans need vitamin D for natural growth and development. Thus, lack thereof can contribute to conditions such as osteoporosis, as vitamin D is essential to help the intestines absorb calcium. Vitamin D is also important for boosting the immune system and protecting against a variety of diseases. When a protein called 7-DHC is exposed to UVB rays, it converts into D3, which in turn isomerizes into vitamin D3, the active form of vitamin D required by the body.<sup>20</sup> Consequently, a lack of UV exposure can lead to disturbance of vitamin D concentration and the disruption of bodily function.

### C. DNA Damage Caused by UVR

The UVR that is absorbed by human skin cells (namely UVB radiation) reaches the DNA contained in the nucleus of skin cells. DNA is double stranded, meaning that two long DNA molecules are bonded across each strand between the nucleotide bases through a process called hydrogen bonding. These base pairs are important during protein synthesis, since the bases are read by replication enzymes.<sup>21</sup> DNA is transcribed into RNA which is translated into proteins. The proteins then express certain genes which allows many organic systems within the body to function.

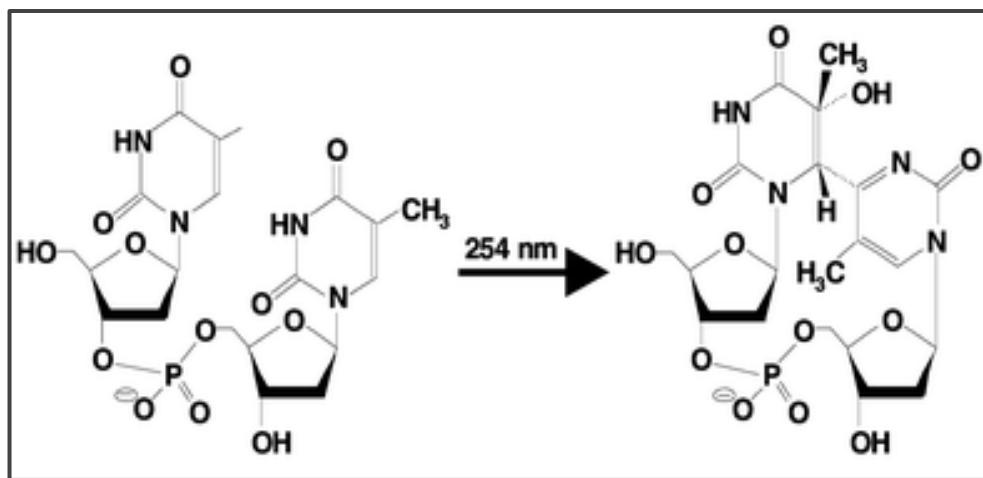
DNA molecules are composed of nucleotides, the monomers (base units) of DNA, which form the entire human genome. Nucleotides contain a deoxyribose sugar ring, a phosphate group, and a base. The bases pair with other molecules from the opposite strand of DNA via hydrogen bonding. DNA has 4 bases: Adenine (A), Thymine (T), Cytosine (C), Guanine (G). Pairing of the bases across the strands is complimentary, meaning that A pairs with T and C pairs with G. This sequencing continues down the strand of DNA for millions of base pairs.

When UVB Radiation comes into contact with DNA, the chemical structure of the nucleotides change. T and C are classified as pyrimidines, bases that have a single ringed structure. A and G are classified as purines, bases that have a double ringed structure. When UVB radiation comes into contact with DNA, the atoms that comprise the pyrimidine can become excited, making the structure unstable. To correct this, pyrimidines that are adjacent to each other form two single bonds between each other instead of forming hydrogen bonds between strands. This creates a “bubble” on one side of a DNA strand, due to the lack of base pairing, referred to as a lesion. These lesions form between two thymine residues, two cytosine residues, and between adjacent thymine and cytosine residues. The specific lesions regarding pyrimidines that occur in human skin cells are called cyclobutane pyrimidine dimers (CPDs - Figure 2) and pyrimidine-(6-4)-pyrimidine photoproducts (6-4PPs - Figure 2).<sup>22</sup> CPDs have been found to be more cytotoxic (harmful to other cells) than 6-4PPs, but 6-4PPs may have more lethal impacts on human health.



**Figure 2: Synthesis of Cyclobutane Pyrimidine Dimers<sup>22</sup>**

Figure 2 shows the process of two pyrimidine nucleotides (T with T and T with C) interacting with UV radiation to form a CPD. The affected nucleotides no longer bond across DNA strands and bond side to side with a cyclobutane ring.

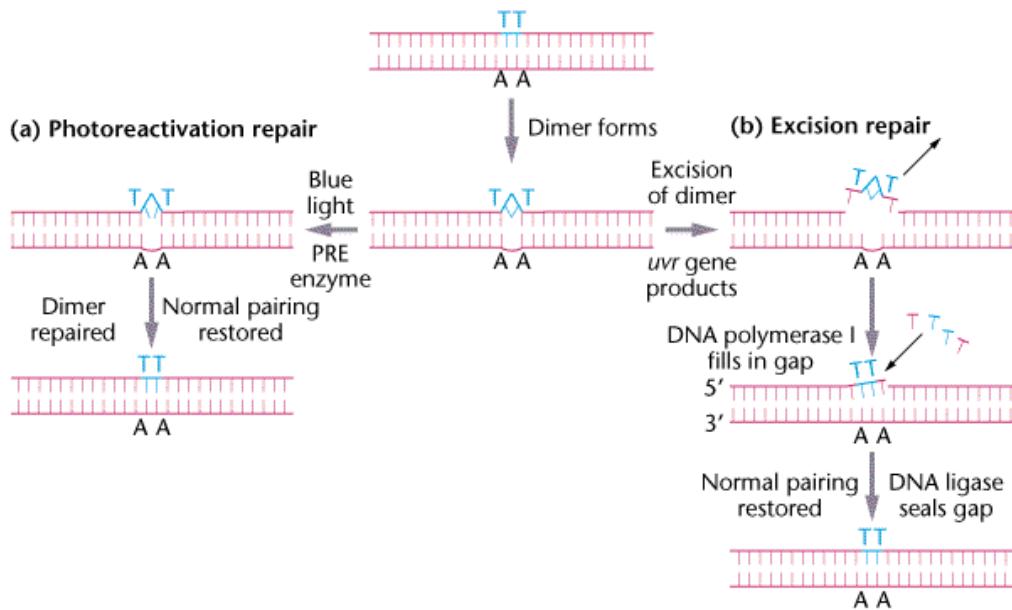


**Figure 3: Formation of Pyrimidine-(6-4)-pyrimidine Photoproducts<sup>22</sup>**

Figure 3 shows the process of two pyrimidine nucleotides interacting with UV radiation to form a 6-4 PP. The affected nucleotides no longer bond across DNA strands and a bond forms between nucleotides, creating a large lesion on the DNA molecule.

Both CPDs and 6-4PPs can prove deleterious to DNA replication and human health.<sup>23</sup> When DNA is being replicated, an enzyme called DNA polymerase scans each DNA strand, synthesizing a copy which is used to complete a new DNA molecule. When lesions are present on a DNA strand, the DNA polymerase cannot move along the strand or read the unexpected DNA structure. By inhibiting DNA polymerase, UV-induced lesions would prevent cells from being correctly reproduced, potentially leading to problems including developmental issues, melanoma, and other forms of skin cancer. Not only would the function and ability of DNA polymerase to replicate DNA be negatively impacted, but so would the function of RNA polymerase, another enzyme which helps synthesize proteins. RNA polymerase II can become stuck at these lesions when transcribing DNA which can inhibit protein formation. Without the production of correct proteins, these molecules could function incorrectly, which could again lead to cancer and many other diseases. A lack of proteins could also cause disease and health problems.

First, UV rays will hit the cells in the epidermis. Upon impact, a mutation may occur, forming a thymine-thymine bond that will create a bump in the DNA strand. This bump is called a thymine dimer, and can increase the risk of skin cancer in an individual. However, the body has a specific excision repair to get rid of this thymine dimer before it causes too much harm. While multiple repair processes occur, eukaryotes generally use excision repair to fix the lesion (Figure 4). First, UVR gene products will encode proteins that will locate the dimer, and an endonuclease will cleave the thymine-thymine dimer out of the DNA strand. Then, DNA polymerase I will build up new nucleotides in the cleaved region, and DNA ligase will mesh the new segment back into the strand.

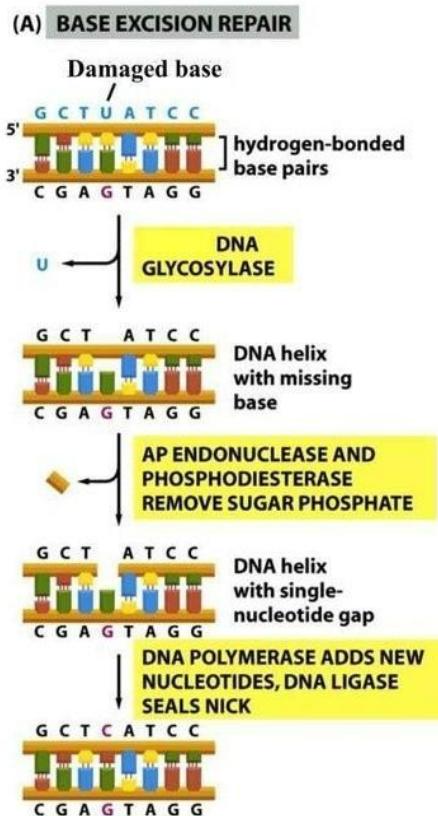


**Figure 4: Thymine Dimer Repair<sup>66</sup>**

Figure 4 shows the process in which a thymine dimer is created and excised from a DNA strand. The thymine dimer is a small mutation in the DNA strand and can increase the risk of skin cancer in an individual.

#### D. DNA Repair Mechanisms

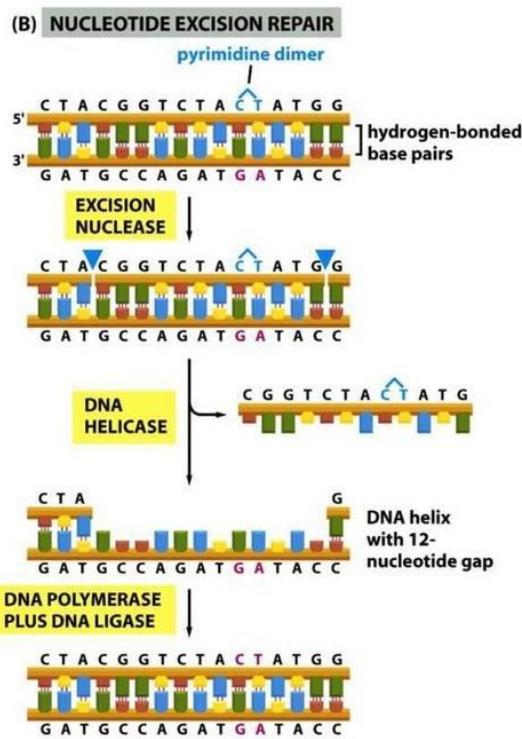
To combat UV induced DNA damage, the cell must repair the DNA. In eukaryotes, such as humans and yeast, base excision repair (BER) and nucleotide excision repair (NER) are two ways this is accomplished.<sup>24</sup> BER focuses on repairing the effects of abnormal bases that do not disrupt the DNA structure (Figure 5). BER uses DNA glycosylases, enzymes that remove a damaged base from DNA by breaking the N-glycosidic bond between the DNA base and its backbone, creating an apurinic or apyrimidinic (AP) site without a base.<sup>25,26</sup> Different DNA glycosylases are used for different types of DNA damage.<sup>25</sup> Once the damaged base is removed, one of several DNA repair pathways is implemented, where an AP endonuclease or AP lyase cuts the DNA strand 3' or 5' to the AP site.<sup>25</sup> Then, an AP endonuclease and a phosphodiesterase removes the remaining DNA backbone at the site.<sup>25</sup> Finally, the DNA damage is repaired when a DNA polymerase replaces the missing base and a DNA ligase reseals the strand.<sup>25</sup>



**Figure 5: Base Excision Repair<sup>27</sup>**

Figure 5 shows the BER process and specifies the proteins involved in each step.

The second method that cells use to repair UV induced damage is NER (Figure 6). This repair mechanism is found in many organisms and is principally responsible for repairing large DNA lesions, such as CPDs, which include thymine-thymine dimers and 6-4PPs, that can disrupt DNA transcription and replication.<sup>25</sup> Additionally, there are two types of NER: global genome NER (GG-NER) and transcription-coupled NER (TC-NER).<sup>28</sup> GG-NER repairs abnormalities anywhere in the genome while TC-NER repairs abnormalities specifically in transcribed DNA. While the process isn't completely determined, the general process has been identified. One proposed NER mechanism states that in GG-NER, the XPC-hHR23B protein identifies DNA damage and binds to that location, and in TC-NER, DNA damage is identified when RNA polymerase II encounters the damage during transcription.<sup>25,28</sup> Then, both types of NER trigger the transcription factor TFIIH to attach to the site, verifying the presence of DNA damage. Next, DNA helicases contained in TFIIH begin to unravel the damaged DNA.<sup>25,28</sup> Simultaneously, the excision nuclease XPA is recruited to the site, working with the single-stranded DNA binding protein RPA to correctly position the XPG and XPF-ERCC1 endonucleases to cut the DNA 3' and 5' to the lesion, respectively.<sup>25,28,31</sup> Then, DNA polymerase fills in the removed section with the correct bases, using the complementary strand as a guide, followed by DNA ligase sealing the DNA.<sup>25,28</sup> People with defective NER function are up to 1000 times more likely to acquire skin cancer as their cells are unable to repair DNA lesions caused by UV rays.<sup>25</sup>



**Figure 6: Nucleotide Excision Repair<sup>27</sup>**

Figure 6 shows the NER process and specifies the proteins involved in each step.

Beyond BER and NER, eukaryotic cells can also repair damaged DNA through homologous recombination (HR) and non-homologous end joining (NHEJ). When UV damage causes a double-stranded DNA break, which can have extreme and fatal impacts on a cell, either HR or NHEJ can be implemented.<sup>25,28</sup> In HR, the cell uses the homologous DNA strand from a sister chromatid to replace the damaged DNA strand.<sup>25,29</sup> However, if HR cannot be performed, as occurs when the ends of two incongruous DNA strands are present, the less accurate NHEJ is used. In NHEJ, there is no homologous strand that can be used, so, in mammals, a Ku complex protein is bound to each end. Then, the catalyzing portion of DNA protein kinase (DNA-PKcs) recruits various other repair proteins to the site and aids in bridging the gap between each end by interacting with the Ku complexes and other proteins. In yeast, a similar process occurs with proteins homologous to Ku complexes and DNA-PKcs.<sup>24</sup>

Finally, apoptosis, also known as programmed cell death, is used to prevent carcinogenic mutations from being replicated.<sup>30</sup> When DNA is damaged by UV radiation, apoptosis can be induced through either extrinsic or intrinsic pathways.<sup>30</sup> Extrinsicly, there are several different receptor types that receive extracellular signals in the presence of UV radiation.<sup>30</sup> Intrinsically, when UV radiation damages cell DNA, the p53 gene is activated, and if the mitochondria or mitochondrial DNA is damaged, cytochrome c is released into the cytoplasm of the cell.<sup>30</sup> Both the extrinsic and intrinsic pathways trigger the executioner pathway, leading the cell to die in order to prevent duplication of damaged DNA.<sup>32</sup> For example, once exposed to UV radiation, “sunburn” keratinocytes are eliminated through both the extrinsic and intrinsic pathways in order to prevent their mutations from being passed down to daughter cells.<sup>33</sup>

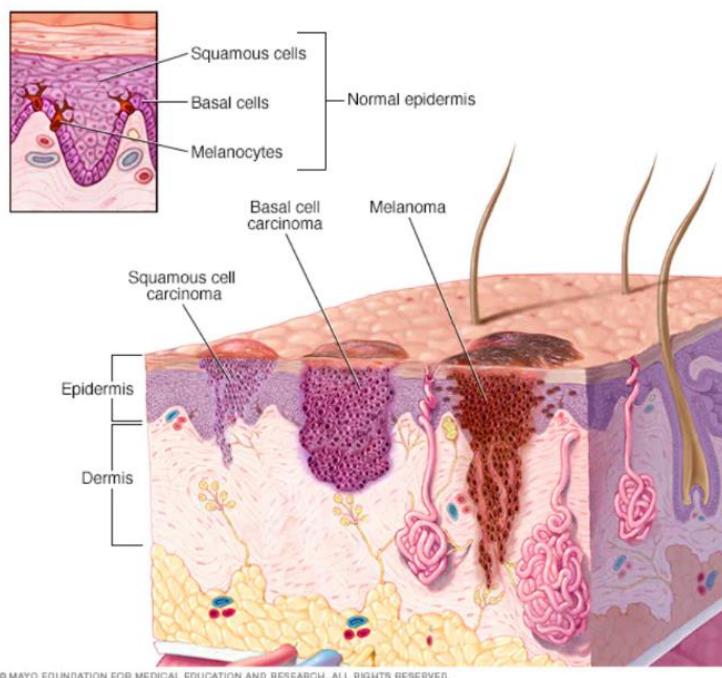
## E. UVR and Skin Cancer

By age 70, approximately 1 in 5 Americans will develop skin cancer. UVR is the primary cause of skin cancer and is associated with 90% of nonmelanoma skin cancer cases.<sup>34,35</sup> More than 5 million people are diagnosed with skin cancer yearly in the U.S., contributing to over \$8.1 billion in medical expenses.<sup>36</sup>

Approximately 3.6 million people in the U.S. are diagnosed with basal cell carcinoma (BCC) yearly, which makes up around 80% of all non-melanoma skin cancer diagnoses.<sup>37</sup> Since basal cells are located in the outermost layer of the epidermis (Figure 7), they are exposed to some of the highest levels of UV radiation the body encounters, putting them at an increased risk for DNA damage. UV-induced mutations in the PTCH1 gene can turn on the sonic hedgehog pathway and trigger continuous signaling, which can cause basal cells to grow abnormally and form carcinomas.<sup>38</sup> The sonic hedgehog signaling pathway is active during embryonic maturation and plays a key role in tissue and organ formation. After birth the pathway is normally deactivated. When treated early, basal cell carcinoma has a 5-year survival rate of 100%.<sup>30</sup> If treatment is inadequate or delayed, the tissue undergoes significant damaged in the form of open sores and discolored lesions.<sup>39</sup>

Over 1.8 million people are diagnosed with squamous cell carcinoma (SCC) yearly, representing around 20% of all skin cancer diagnoses.<sup>37</sup> The EGFR gene helps to control squamous cell proliferation. UV radiation can mutate the EGFR gene which can thrust squamous cells into a hyperproliferative state and cause a tumor to form in the epidermis.<sup>40</sup> Squamous cell carcinoma is more likely to metastasize than basal cell carcinoma. There are limited options for treatment and a poor prognosis for squamous cell carcinomas, resulting in more than 15,000 deaths each year from the disease.<sup>41</sup>

Though melanoma is less common than BCC and SCC (207,930 diagnoses in the U.S. 2021), it is the most invasive and deadly form of skin cancer.<sup>42</sup> In its metastatic form, the 5-year survival rate is 5-19%.<sup>43</sup> A single blistering sunburn during childhood can double a person's risk of developing melanoma.<sup>44</sup> There has been a 320% increase in melanoma from 1975 to 2018 (in part due to the growth in popularity of tanning beds that expose people to concentrated UV radiation).<sup>76</sup> Melanocytes protect the skin by absorbing harmful UV radiation. This radiation puts the cells at an increased risk of mutation in the BRAF gene that regulates their growth.<sup>45</sup> Mutations in BRAF alone may not lead to malignancy. However, many UV-induced driver mutations such as NF1, PTEN, CDKN2A, TP53, and RAC1, can synergize with mutated BRAF and cause melanoma to develop.<sup>46</sup> Early stages of melanoma in the epidermis can be removed surgically. When cancerous melanocytes push down through their normal position in the epidermis and into the dermis and subcutaneous tissue (Figure 7), they can invade blood vessels and lymph nodes. These act as highways that allow melanoma to migrate to other bodily tissues well beyond their place of origin.



**Figure 7: Types of Skin Cancer<sup>39</sup>**

Figure 7 illustrates the normal arrangement of basal cells, squamous cells, and melanocytes, and the layout of cutaneous and subcutaneous tissue with these cells in a malignant form.

UVR can suppress the immune system by compromising its function in the skin, which reduces its ability to treat malignancies. Langerhans cells are the primary antigen presenting cells found in the epidermis and constantly survey the region for abnormalities. When an antigen is recognized, Langerhans cells play a central role in triggering an immune response by activating T-cells via the major histocompatibility complex 2 (MHC-II) and intracellular adhesion molecule 1 (ICAM-1). Exposure to UVR can form reactive oxygen species, such as hydrogen peroxide, that compromise the expression of MHC-II and ICAM-1, making it difficult for Langerhans cells to interact with CD4<sup>+</sup> T cells and initiate an immune response.<sup>47,48</sup> An increased production of the immunosuppressive cytokine IL-10 by keratinocytes and regulatory T cells has been observed after unhealthy levels of UV radiation.<sup>49</sup> The abundant IL-10 in the affected area of the epidermis can then enter the bloodstream and cause systemic immunosuppression. The mutagenic effects of UVR coupled with its immunosuppressive capabilities make it difficult for the immune system to mediate damage, thus increasing susceptibility to photocarcinogenesis and UV-induced skin tumors.<sup>50</sup>

## F. Sunburn

A sunburn is the outcome of overexposure to UVR, resulting in damage to the DNA.<sup>51</sup> The epidermis becomes inflamed as a result of the blood flow being directed to the burnt area to provide it with nutrients to repair the damage.<sup>52</sup> The inflammation gives the sunburn its characteristic red appearance and feel.<sup>52</sup> However, this is only the beginning of the repair process as there are four main stages an individual goes through when they become sunburnt. The first stage, known as exposure, is the time range when a person gets exposed to UVR resulting in a sunburn. The time required for sunburn varies as it is very person specific; this can be a matter of minutes for one person or a few hours for another.<sup>53</sup> One factor that determines how long a person can be exposed to UV rays before getting a sunburn development is the

amount of melanin present in their skin. As mentioned previously, melanin is a substance in the body that is responsible for the skin's pigmentation.<sup>52</sup> People with fair skin are more likely to get a sunburn than those who have a darker skin tone.<sup>51</sup> According to a CDC study, 66% of white women and 65% of white men received a sunburn in the past year. On the other hand, 38% of Hispanic women and 32% of Hispanic men experienced a sunburn. Additionally, 13% of black women and 9% of black men experienced a sunburn. Another variable that can affect one's risk is skin tone. The melanin in the skin acts as a natural barrier and can absorb some of the UV rays. However, if an individual has lighter skin, the amount of melanin in the skin is considerably less, meaning that it is more likely that they will burn. On the other hand, if an individual has darker skin, there is more melanin in their skin, making it less likely that they will burn.<sup>53</sup> The next stage, the burn stage, occurs when the epidermis becomes more inflamed, leading to swelling and a painful burning feeling. This causes an erythema, which is most commonly recognized as the bright red color of a sunburn.<sup>52</sup> This stage can be felt as early as two hours after the initial exposure.<sup>54</sup> The third stage is when peeling starts to occur. As a way to protect the damaged skin, the upper portion of the dermis will create a barrier with the keratinocytes cells to protect the skin from further damage. Although keratinocytes are also skin cells, they are a special variety since instead of flaking off, they stick together, which results in the peeling.<sup>52</sup> The fourth stage in the sunburn journey is to understand what causes sunburns and how to avoid them at the individual level since every skin type is different. By knowing simple techniques, people can learn that sunscreen should be reapplied every two hours, especially after swimming or sweating. They can also learn other important information that is suited for their specific skin type, since not everyone's skin is the same.

Another factor to take into account when determining the type and amount of sun protection that may be necessary, one can look at the UV index. The UV index is a measure of the strength of the UV rays; this is converted to a number to help people identify the level of caution they should exercise. The scale ranges from 1-11, with values one and two being considered the least dangerous values on the UV index.<sup>55</sup> However, it is still recommended to wear sunglasses during this time even though the level of risk is low, as there is still a risk of exposure especially if on a beach or mountaintop.<sup>55</sup> A UV index of 3-5 is considered to be generally moderate for exposure. Wearing and reapplying sunscreen during this time is generally recommended especially after swimming or sweating.<sup>56</sup> A UV index of 6-7 is considered high and requires sun protection in order to prevent burning.<sup>55</sup> The time it takes someone to burn varies from person to person, but the average time during this range is usually around 15 minutes. For any value higher than 8, it is recommended to wear sunscreen at all times and to avoid going outside during the hours of 10 AM and 4 PM when the sun is at its highest peak.<sup>56</sup> Getting a sunburn during this time is very common, as the time frame for exposure is very short.

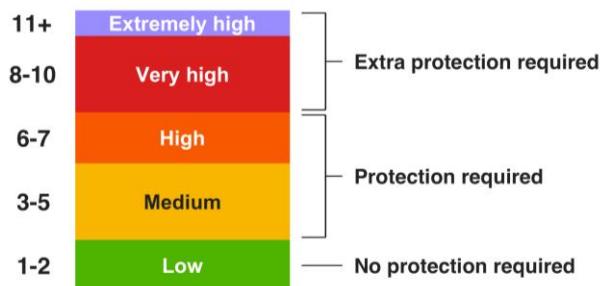
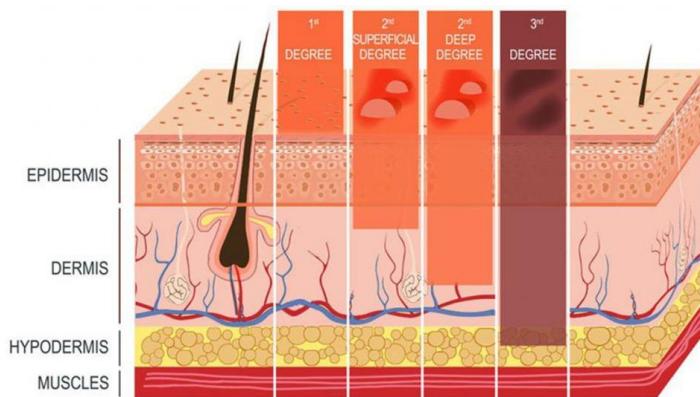


Figure 8: Measure of the UV Index<sup>57</sup>

Figure 8 shows the UV index, and it is a measure of the UV rays at a certain time or place and they show how much protection a person needs. As the value on the UV index increases, so does the level of protection needed to protect oneself.

After getting a sunburn, the degree of burn is categorized. The most common, a first-degree burn, is when the epidermis becomes damaged.<sup>58</sup> This form of a sunburn will usually go away after a few days. Second degree burns are more serious and dangerous since the effects can reach down to the dermis layer. Blistering becomes very common during this stage as a protection mechanism from any outside factors that may irritate the burn more. Third degree burns are the most rare and serious forms of a sunburn as the damage destroys the dermis and epidermis areas and, in some cases, reaches to the fat part of the skin.<sup>59</sup> There is almost no sensation since all of the nerves are destroyed and it requires immediate medical attention. In general, to treat a first-degree sunburn, one should cleanse it with cool water and soap and apply cream such as aloe vera to hydrate the skin. The sunburned area should be kept out of the sun and if venturing outside, it should be protected with sunscreen.<sup>58</sup>



**Figure 9: Damage Caused By Different Degrees of Sunburn<sup>60</sup>**

Figure 9 shows the damages caused by UV rays for the different types of sunburns. As the degree for the sunburn increases, so does the damage caused by the rays.

Although it seems like UV rays are harmful, there are some benefits. When UVB rays reach our skin, a protein in our skin known as 7-DHC can synthesize the rays into vitamin D3, an active form of Vitamin D.<sup>61</sup> Although there are many myths out there such that wearing sunscreen can cause vitamin D deficiency, many clinical studies have never found evidence that links sunscreen to vitamin D deficiency.<sup>62</sup> It does not take much time outdoors for your body to produce the amount of vitamin D it needs; usually it only requires around 10-15 minutes of exposure on the arms, legs, and face a few times a week. Unfortunately, the same waves that produce vitamin B are the same ones that can cause skin cancer, which is why it is important to exercise caution and wear sun protection when outside.<sup>63</sup>

## G. Sunscreen

Sunscreen protects humans by providing a defensive barrier on the skin against harmful UV rays from the sun. However, sunscreen itself has varying factors that may influence its ability to protect the skin against such harmful rays. Factors include organic vs. inorganic ingredients, varying SPF levels, sport vs. clear, etc. Inorganic sunscreen contains titanium dioxide and zinc oxide, semiconducting materials that are able to use their shielding abilities to reflect a portion of the UV rays away from the skin. Conversely, organic sunscreen contains molecules, including oxybenzone and avobenzone, that absorb the light. Inorganic sunscreens are known as “physical” filters, whereas organic sunscreens are referred to as “chemical” filters.<sup>64</sup> These names derive from the fact that inorganic sunscreens act as physical blockers that absorb UV rays, while organic sunscreens have chemical backbones that absorb them.

Many scientists have discussed whether or not sunscreens work, and if they are ever harmful. While the general consensus is that sunscreens are effective, they also include some ingredients that have been proven to be harmful in high amounts towards humans. For instance, oxybenzone, an ingredient commonly found in many sunscreens, is known to be a hormone disruptor, commonly causing bloating, fatigue, infertility, and more. Another active ingredient in some sunscreens include retinyl palmitate, a form of vitamin D, which has been found to increase the risk of skin cancer. Furthermore, higher SPF sunscreens contain more chemicals than sunscreens with a lower SPF. These harmful chemicals may cause a variety of health risks, such as tissue damage, hormone disruption, and allergic reactions to the ingredients in the sunscreen.<sup>65</sup> However, these side effects still do not deter from our main point: using sunscreen is better than not using sunscreen at all.

Sun protection factor, most commonly known as SPF, indicates how many UV rays it takes to reach the skin's surface. For example, if someone applies SPF 30, that means 1/30 rays will reach your skin; alternatively, it will take 30 times longer to burn compared to no sunscreen application.<sup>67</sup> SPF with a higher number will protect the skin more because less rays can reach and potentially damage your cells. Additionally, the ability of the sunscreen to block UV rays must be considered. Broad spectrum sunscreens block both UVA and UVB rays, while narrow spectrum sunscreens only block one type (Figure 10).

There are more significant factors of sunscreen compared to buying a sunscreen based on its price or brand. The ingredients in the sunscreen are what determine the efficacy of the sunscreen. Every sunscreen and its SPF number must be approved by the FDA. Therefore, name brand or generic products and prices may not make any difference because they should all have the same efficacy.<sup>68</sup>

Sunscreens have characteristics that make them differ from each other. Regular sunscreens are not technically waterproof. There are two different ranges of waterproof sunscreens, water resistant and very water resistant. Water resistant sunscreen can last for 40 minutes in the water and very water-resistant sunscreen can be worn in the water for up to 80 minutes.<sup>67</sup> Traditional non-water resistant sunscreens will not protect your skin once it has come in contact with water. Sport sunscreens are waterproof sunscreens because they were made for people to sweat or have water on them.<sup>68</sup> Clear sunscreens were made so the face or body doesn't have a white or chalky film (which is an effect of normal or tinted sunscreens). The different forms of sunscreen, including spray, stick, or lotion, should not make a difference on shielding someone from the sun's rays because the products were made for people to select based on their preference. When using spray sunscreens, users must be careful not to inhale the sunscreen's propellant and to be cautious that they are putting enough on their body.<sup>69</sup>

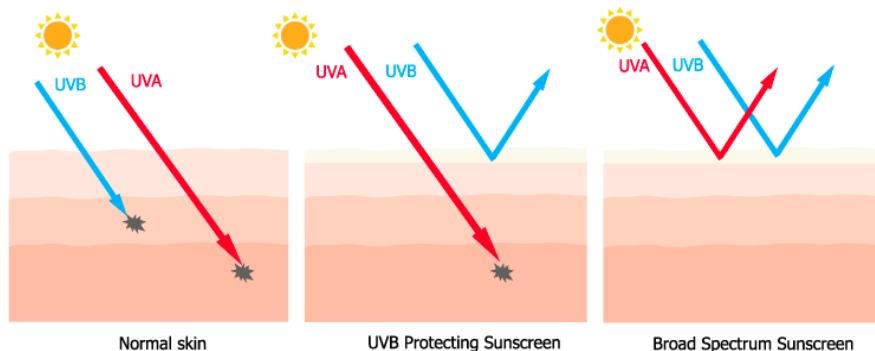


Figure 10: Protection of Broad vs Narrow Sunscreen<sup>70</sup>

Figure 10 describes that normal unprotected skin will absorb both UVB and UVA rays. Narrow spectrum sunscreens will only block UVB rays. Even though UVA is much less harmful compared to UVB, the UVA can still have effects since it reaches the dermis. Therefore, it is ideal to buy a broad-spectrum sunscreen to protect your skin from both UVB and UVA rays.

## H. *Saccharomyces cerevisiae*

*Saccharomyces cerevisiae*, also known as baker's or budding yeast, is a useful model organism that effectively illustrates the effects of UV light on cellular health. The simple requirements for growth, the similarities to other eukaryotic organisms, and the extensive knowledge about the yeast genome clearly show why baker's yeast was a good model organism.

Yeast is a single-celled organism that grows as a single cell, and multiple yeast cells live together in multicellular colonies. The yeast cells reproduce through budding; this occurs when the daughter yeast cell grows a protrusion (a "bud") that continually gets bigger until it is the same size as the mother yeast cell.<sup>71</sup> The bud then cleaves to become a separate single-celled organism. This process is very fast, as budding yeast can double every 45-90 minutes under ideal laboratory conditions, which includes 32°C incubation temperature and the ideal growth media, YEPD. YEPD is a media containing yeast extract, peptone, and dextrose, and with this, scientists are able to give the yeast a food source containing proteins and sugar. In comparison, a mammalian cell typically only divides every 24 hours.<sup>71</sup> The fast reproduction cycle and relatively simple and inexpensive growth also makes yeast resource conscious, which is beneficial for sustainability.

Yeast have been studied extensively, and this depth of knowledge contributes greatly to being able to determine the genes or activity that is responsible for a certain cellular response, such as cell checkpoints or DNA damage repair mechanisms. The complete genome of *S. cerevisiae* has been published; it was determined that the genome is 12,157,105 base pairs in length and contains 6,692 genes.<sup>72</sup> Since this organism is considerably more simple than many other organisms, scientists are able to work out the connections between genes and proteins using homologies found within the organisms.

Evidently, there are many similarities between both yeast and humans. Both are eukaryotic, which means they have a nucleus, though yeast is considered as one of the simplest eukaryotes. In addition, the yeast genome is homologous to at least 20 percent of human genes known to have a role in disease.<sup>72</sup> It is known that diseases result from the disruption of very basic cellular processes, such as in DNA repair and cell division, and yeast conserve these same basic cellular processes. Importantly, it has been found that much of the yeast genome that regulates cell division in yeast have homologs that control cell division in higher organisms, including humans.<sup>73</sup> Since cancer is an unregulated cell division, yeast can be used as a model to determine factors that contribute to unregulated cell growth.

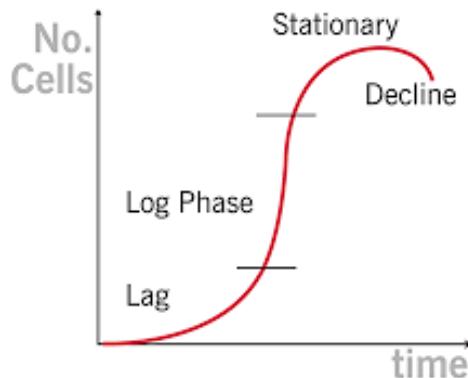
The response of yeast cells to DNA damage is closely related to the human DNA repair process. Even in the relatively simple genetic framework, checkpoints function in coordinating cell cycle progression and DNA repair within the yeast cell.<sup>74</sup> Several sensor proteins function to receive the signal of damaged DNA and transfer the damage signal to effector kinase, and after the activation of the checkpoint, a series of cellular processes is catalyzed.<sup>74</sup> This includes DNA damage repair pathways such as homologous recombination, non-homologous end joining, and nucleotide excision repair, all of which closely model the DNA repair methods in human cells.<sup>74</sup> Since these DNA repair pathways closely resemble those found in humans, tests conducted in yeast can be correlated to how human DNA repair pathways will behave after DNA damage, specifically in excision repair.

## II. Methods

### A. General Microbiology Techniques

#### 1. Yeast Propagation

When yeast is cultured in a controlled medium, they often follow a well-established pattern for microbial growth consisting of three phases: lag phase, log phase, and stationary phase.<sup>86</sup> The lag phase characterizes the beginning of yeast growth. The log phase describes the exponential increase of yeast growth. The stationary phase begins when the yeast have reached their carrying capacity and the cell population stabilizes and remains constant. For the purpose of this study *S. cerevisiae* were irradiated in the log phase because they were actively growing and dividing. During this period of the growth cycle, yeast cells check DNA for damage and actively replicate. The cells were incubated at 32°C for 48 hours in YEPD (yeast extract peptone dextrose) growth media.



**Figure 11: Stages of Yeast Growth<sup>75</sup>**

As shown in Figure 11, the growing stages of yeast are described as lag phase, log phase, and stationary phase. Log phase has an exponential increase until the stationary phase is reached. In this phase, the yeast have reached carrying capacity.

#### 2. Serial Dilutions

Serial dilutions were used to reduce the concentration of *S. cerevisiae* in liquid media by consecutively resuspending using a diluent. Each tube of dilution should have a 1:9 ratio of *S. cerevisiae* and sterile water, meaning 1 mL of yeast solution was transferred into 9mL of water for every dilution. Before every sample was extracted, the tubes were vortexed in order to mix the yeast. After vortexing, the tube of solution was passed over the Bunsen burner to maintain sterilization. We repeated this process until we had completed dilutions to  $10^{-5}$ .

## B. Sunscreen Exposure

### 1. Sunscreen Application

In order to test the efficacy of SPF in protecting cells from UV damage, it was necessary to test the application of different types of sunscreen on yeast cells in a controlled environment. To avoid applying sunscreen directly to the yeast samples, yeast were spread plated onto YEPD agar plates. Then 0.5 grams of sunscreen was evenly spread onto plastic wrap. The plastic wrap was used to cover the agar plate. For the control groups, agar plates were covered with plastic wrap without sunscreen.

### 2. UV Irradiation

For controlled exposure to UV radiation, the plates were placed into UV hoods for differing periods of time, which are specified in each experiment. At specific time points, the plates were taken out and lids were placed on to be stored and incubated.

Additionally, in the nucleosome fragmentation and immunofluorescence, it is specified whether or not the yeast was irradiated under the protection of a sunscreen. Products with a positive (+) SPF had sunscreen spread over the plastic wrap cover on the agar plate. On the other hand, products with negative (-) SPF had only a clear plastic wrap covering the plate.

In order to test the efficacy of the various types of sunscreen products available, several different products were tested, specified in each individual experiment.

### 3. Viable Count and % Survivor Calculations

In order to accurately compare the amount of yeast that survived UV irradiation, viable count, and percent survivor was calculated. Viable count indicates the number of cells that survived UV irradiation and were capable of reproduction. Viable count was calculated by dividing the number of colonies on the plate (CFU) by the volume spread plated and multiplying by the dilution factor (one divided by the dilution concentration). In order to accurately compare data between varying samples spread plated by different groups, a percent survivor calculation was necessary. Percent survival was calculated by dividing the viable count of the irradiated yeast sample by the viable count of the control, the unirradiated yeast, and multiplying by 100%. A graph of Time vs. Percent Survivor was then created with the data to demonstrate any correlations between the length of time the yeast were irradiated and percent of colonies that survived the irradiation.

## C. Nucleosome Fragmentation Assay

### 1. Genomic DNA Extraction

DNA extraction is required to assess DNA damage using agarose gel electrophoresis. The genomic DNA (gDNA) of *S. cerevisiae* was extracted using the Yeast Genomic Extraction kit from Thermo Fisher Scientific according to the manufacturer's instructions. In brief, yeast cells were irradiated and then frozen at -20 degrees C. Once thawed, the cells were resuspended in Y-PER reagent and incubated at 65°C for ten minutes. The cells were then centrifuged, and the supernatant was discarded. The cell pellet was resuspended with DNA Releasing Agent A and DNA Releasing Agent B via inversion, and subsequently incubated at 65°C for 10 minutes. Protein Removal Reagent was added to the tubes before being mixed by inversion a second time. The suspension was centrifuged, and the supernatant was transferred to sterile

conical tubes. DNA was concentrated using ethanol precipitation. Isopropanol was added to the tubes and mixed in by inversion. The tubes were centrifuged, and the supernatant was removed. Ethanol was then added and mixed in by inversion. The tubes were centrifuged, the supernatant was discarded, and the pellets were left to dry overnight. The DNA pellets were resuspended in TE buffer by gentle rocking.

## 2. Agarose Gel Electrophoresis

After extraction, genomic DNA was loaded into 0.8% ethidium bromide containing agarose gels. A current was applied such that a negative anode is situated at the gel wells and a positive cathode is situated on the other side of the gel. The negatively charged DNA fragments will migrate through the gel towards the positive charge. Smaller fragments would move farther towards the positive charge than relatively larger fragments. Because of this, the sizes of the DNA fragments could be determined.

Once the genomic DNA was extracted from the irradiated yeast strains, it was prepared for electrophoresis. Each extracted sample contained eight microliters of extracted gDNA. 10X dye was micropipetted into the samples so that the movement of the gDNA across the gel was visible. The samples were flicked and then centrifuged three times to ensure that the dye and the gDNA were completely mixed. The samples were then loaded along with a 1KB DNA ladder and a positive control. If a sample of DNA was irradiated and damaged, then the sample lane will look similar to the positive control. After an hour, the current was stopped and the gel was imaged for analysis.

## 3. Spectrophotometer Calculations

A spectrophotometer was used to measure the purity, concentration, and yield of the DNA obtained for nuclear fragmentation. The spectrophotometer gave readings for the DNA's absorbance at wavelengths of 230 nm, 260 nm, and 280 nm. The aromatic rings of DNA and RNA absorb at 260 nm while 230 nm and 280 nm indicate the presence of protein contamination. Calculating the purity of the sample included dividing the reading at 260 nm by the reading at 280 nm. If the resulting ratio was between 1.7 and 1.9, the sample was pure. To calculate the concentration, the absorbance at 260 nm was multiplied by the constant 50 ug/mL along with the dilution factor. To calculate yield, the concentration was multiplied by the volume of the final extracted sample in mL (50  $\mu$ L converted to 0.05 mL).

## D. Indirect Immunofluorescence Microscopy

### 1. Preparation of Yeast Cells

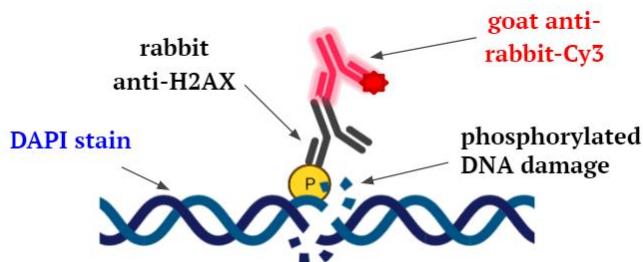
A 5 mL yeast culture was grown overnight and diluted the next day to grow until the log phase. Afterward, the cells were transferred into a petri dish and irradiated in a UV hood for the chosen time points. In addition to the BY4741 strain, a 2DENQ strain (a kind gift from Dr. Anthony Schwacha, University of Pittsburgh) was also irradiated with and without an SPF coating (only for the longest time point). Then, 1 mL of cells were transferred into an Eppendorf tube and centrifuged for 1 minute at 13,000 rpm to separate the yeast cells media. The supernatant was removed, and the 1 mL of yeast cells was harvested twice. The Eppendorf tube was centrifuged at the same speed and was resuspended in phosphate buffer (2.74 mL 1M KH<sub>2</sub>PO<sub>4</sub>, 1.26 mL 1M K<sub>2</sub>HPO<sub>4</sub>, 1.5% BSA, 0.5% Tween20, 1M MgCl<sub>2</sub> + 0.1% Triton X-100 in 50 mL PBS) containing 4% paraformaldehyde to fix the cells and were incubated overnight at 4°C.

To prepare the slides for immunofluorescence, 10  $\mu$ L of poly L-lysine was applied onto each of the wells and incubated at room temperature for 10 minutes. The wells were rinsed 5 times with 20  $\mu$ L of distilled

water. All liquid was removed before adding the next wash and any excess liquid was removed after the final wash. The cells were washed twice using 500  $\mu$ L of a phosphate buffer and to eliminate any unnecessary proteins and RNA that remained. Afterwards, it was resuspended in 100  $\mu$ L of sorbitol buffer (2.2 g sorbitol, 0.274mL 1M KH<sub>2</sub>PO<sub>4</sub>, 0.126mL 1M K<sub>2</sub>HPO<sub>4</sub>, 5 $\mu$ L 1M MgCl<sub>2</sub> in 10 mL water). This was followed by addition of 5  $\mu$ L of lyticase. The lyticase was added to break down the cell wall and therefore release the DNA from the cells. The cells are incubated for 15 minutes, shaking at 32°C. After the cells incubate, centrifuge for 1 minute at 5000 rpm. Wash the cell pellets twice in 100  $\mu$ L of the sorbitol buffer and then resuspend in 100  $\mu$ L of the sorbitol buffer. 20  $\mu$ L of cells were applied onto each well of the slide. Incubate at room temperature for 10 minutes. Remove any of the excess buffer after the incubation period and then transfer to the methanol chamber for 6 minutes followed by the acetone chamber for 30 seconds. As it dries, make a humid chamber to prevent the slide from drying out. The humid chamber is an empty 1000 micropipette tip container with Wattman paper that has been soaked by water. Pieces of parafilm are put on top of the Wattman paper. The slides go into the humid chamber after a 1:1000 dilution of rabbit anti- $\gamma$ -H2AX in PBS is added to the wells. Wrap the humid chamber in foil to prevent light from going in and let it incubate overnight in the fridge.

## 2. Addition of Antibodies

The cells were then incubated for two days in the fridge. The cells were washed 5 times with 1X PBS to remove excess antibody and were subsequently incubated with PBS for 10 min at room temperature. The cells were then exposed to a 1:500 dilution of goat anti-rabbit-Cy3 in 1X PBS and incubated for 1hr at room temperature (Figure 12). Incubation was conducted in the dark because Cy3 can undergo photobleaching that reduces its fluorescent capabilities. The cells were then washed 5 times and incubated with 1X PBS for 10 min at room temperature to remove excess antibody. A 50:50 DAPI (4',6-diamidino-2-phenylindole) / fluoromountG stain was applied for 10 min. A cover slip was then added to the slide and bubbles were removed with slight pressure for 10 min. Clear nail polish was used to seal the cover slide to the glass slide.



**Figure 12: Antibody Complex Attached to Damaged DNA**

Figure 12 illustrates the stacking of antibodies on phosphorylated DNA damage for indirect immunofluorescence microscopy.

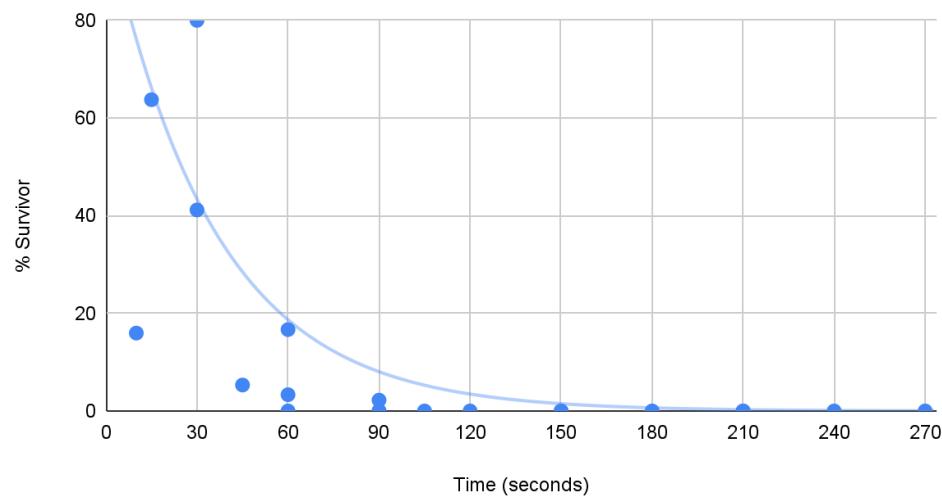
## 3. Imaging

An epifluorescent microscope was used to image the cells. Appropriate filters for DAPI and Cy3 were used, and brightfield images were taken. The DAPI images were false colored blue and Cy3 images were false colored red using ImageJ software. The blue and red images were merged using ImageJ to colocalize the fluorescent signals.

### III. Results

#### A. Killing Curves

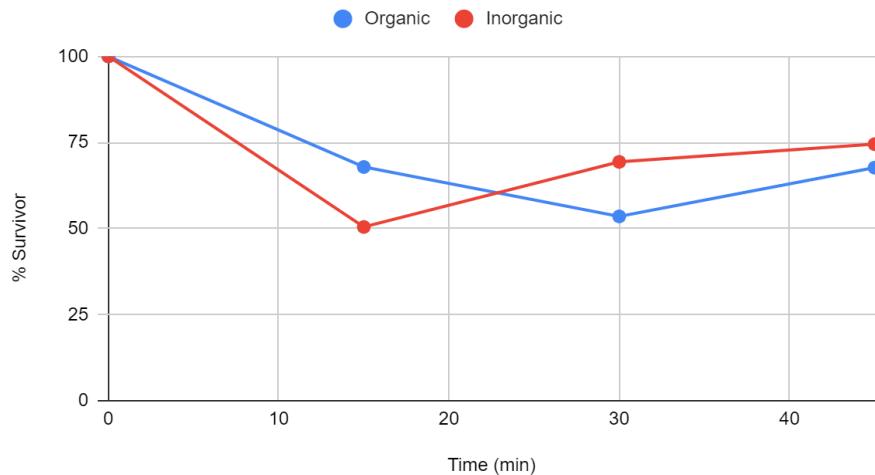
% Survivor of Yeast Colonies vs. Time



**Figure 13: The Percent Survivor of Cells Irradiated Without Sunscreen Protection**

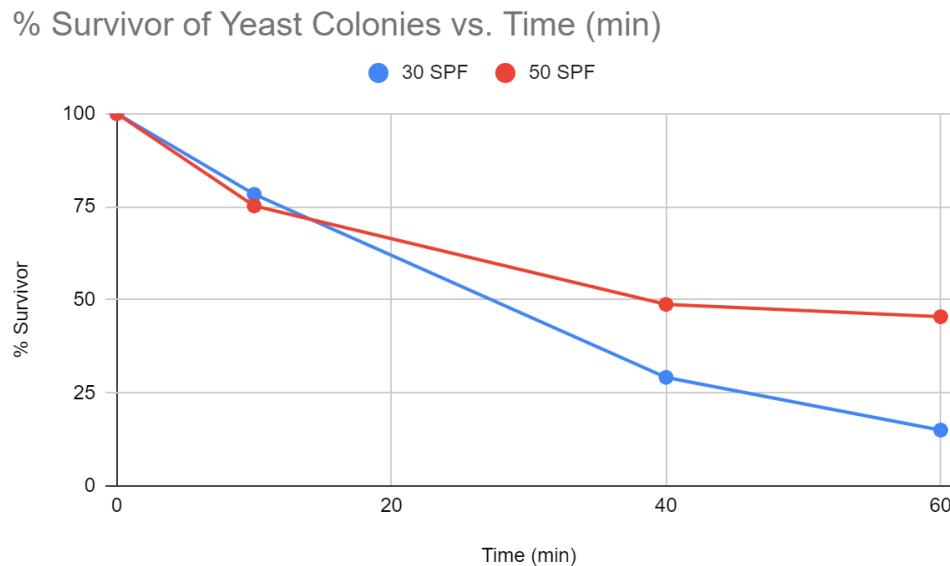
Figure 13 shows the killing curve for the yeast cells as they were irradiated in the UV hood over a time frame of 5 minutes without the addition of any SPF containing product.

% Survivor of Yeast Colonies vs. Time (min)



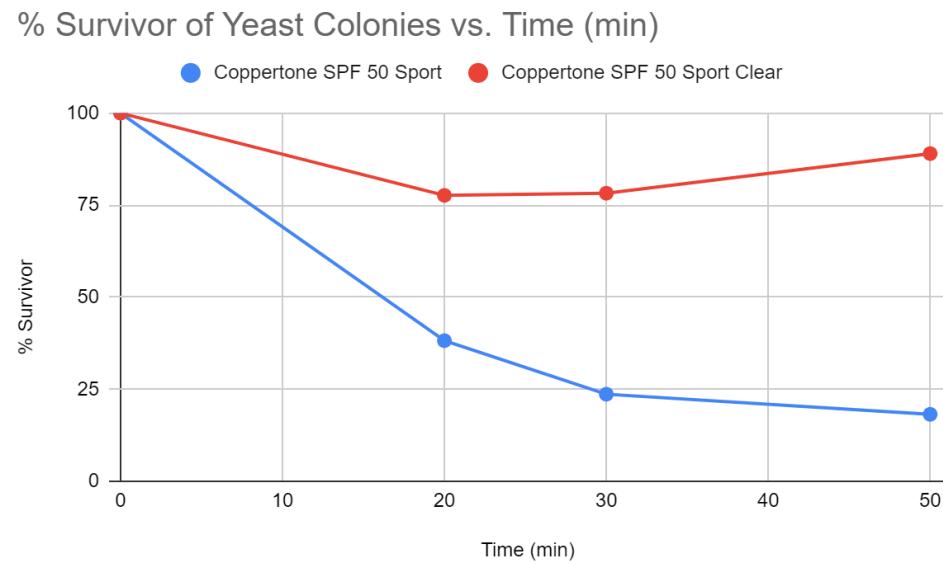
**Figure 14: Comparison of SPF 30 Organic and Inorganic Sunscreen**

Figure 14 shows the effect of both organic (blue line) and inorganic sunscreen (red line) applied to the yeast. The organic sunscreen had an active ingredient of avobenzone, while the inorganic sunscreen had an active ingredient of zinc oxide. Yeast samples were irradiated for a total of 45 minutes.



**Figure 15: Comparison of SPF 30 and SPF 50**

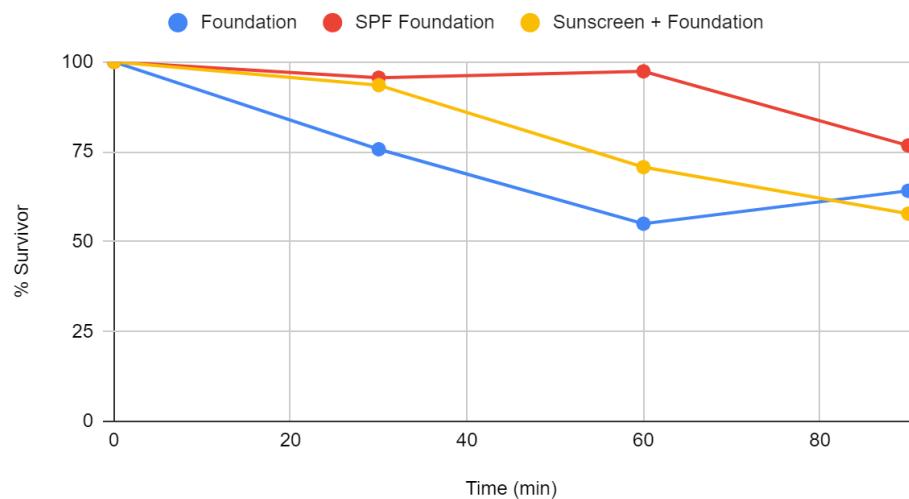
Figure 15 shows the effects of SPF 30 (blue line) and SPF 50 (red line) on the yeast cells. Yeast samples were irradiated for a total of 60 minutes.



**Figure 16: Comparison of SPF 50 Coppertone Sport and Clear**

Figure 16 shows the Coppertone SPF 50 sport clear (red line), and Coppertone SPF 50 sport (blue line) are shown. Yeast samples were irradiated for a total of 50 minutes.

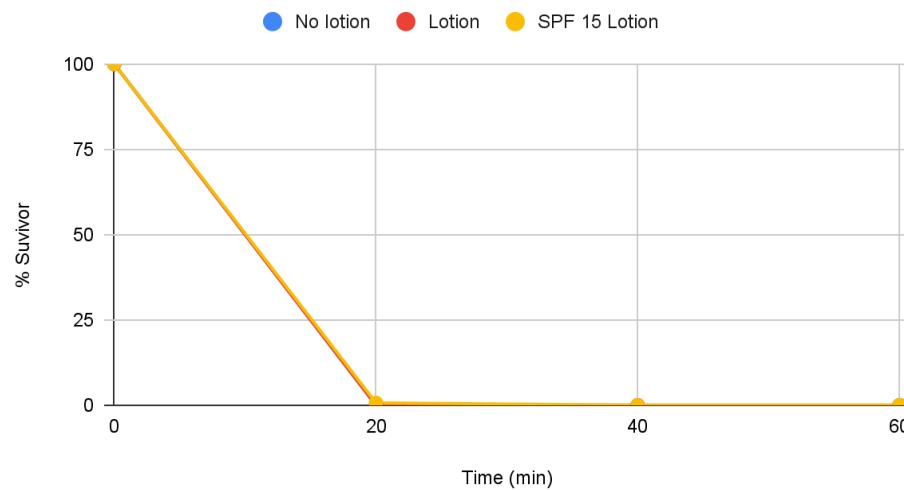
### Foundation, SPF Foundation and Sunscreen + Foundation



**Figure 17: Comparison of Foundation, Foundation with SPF, and Foundation with Sunscreen**

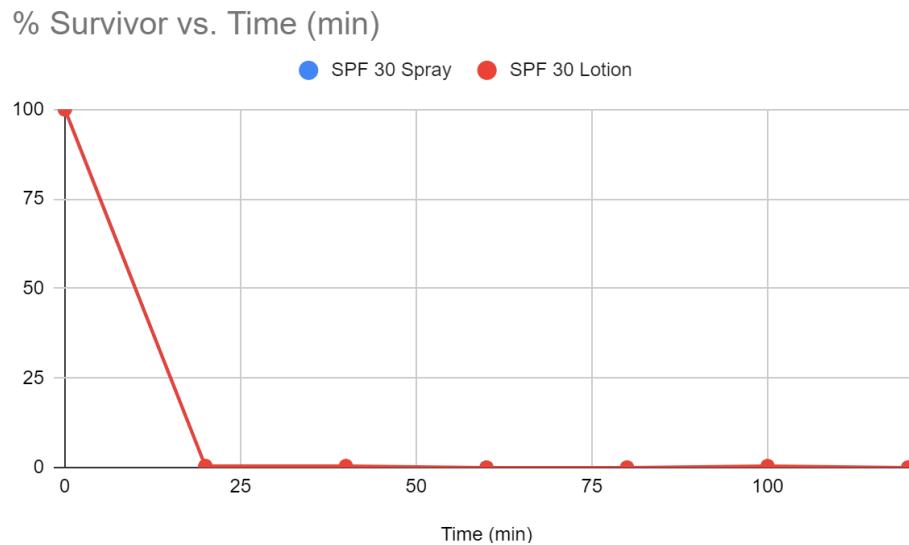
Figure 17 shows the comparison of three different variables that compare regular foundation (blue line), foundation containing SPF 50 (red line), and sunscreen SPF50 followed by foundation (yellow line). Yeast samples were irradiated for a total of 90 minutes.

### % Survivor vs Time (min)



**Figure 18: Comparison of No Lotion, Lotion, and SPF 15 Lotion**

Figure 18 shows killing curves of the yeast with no lotion (blue line), lotion (red line), and lotion containing SPF 15 (yellow line). Yeast samples were irradiated for a total of 60 minutes.



**Figure 19: Comparison of SPF 30 Spray and SPF 30 Lotion Simulating Water Exposure**

Figure 19 shows a killing curve protected by the SPF 30 spray (blue line) and the SPF 30 lotion (red line) simulating water exposure after sunscreen application. Plates were covered in sunscreen, followed by spraying of water. Each sample was gently wiped to remove the water layer. The yeast was irradiated for a total of 120 minutes.

## B. Nucleosome Fragmentation Assay

### 1. gDNA Extraction

Tables 1-4 below show the calculated purities, concentrations, and yields of the DNA samples extracted from yeast irradiated for different amounts of time with different types of sunscreen protection. +SPF means that the sunscreen was used to protect that sample during irradiation, and -SPF means that no sunscreen was used to protect that sample during irradiation.

**Table 1: Coppertone SPF 50 Spray**

Group 1	Purity	Concentration ( $\mu\text{g/mL}$ )	Yield ( $\mu\text{g}$ )
+SPF t=0	2.7931	8.10	0.405
+SPF t=5	1.8	9.45	0.4725
+SPF t=25	2.2837	16.10	0.805
+SPF t=60	1.5574	36.60	1.830

-SPF t=0	1.7633	18.25	9.125
-SPF t=5	2.1350	24.35	1.2175
-SPF t=25	1.9293	53.25	2.6625
-SPF t=60	1.6980	21.65	1.0825

**Table 2: Coppertone SPF 50 Sport Clear**

Group 2	Purity	Concentration ( $\mu\text{g/mL}$ )	Yield ( $\mu\text{g}$ )
+SPF t=0	1.81	15.15	0.7575
+SPF t=10	1.94	6.5	0.325
+SPF t=20	1.36	70.35	3.5175
+SPF t=50	1.34	18.85	0.9425
-SPF t=0	1.71	8.5	0.425
-SPF t=10	1.29	84.05	4.2025
-SPF t=20	1.91	169.5	8.475
-SPF t=50	1.33	75.45	3.7725

**Table 3: Coppertone SPF 50 Sport Spray**

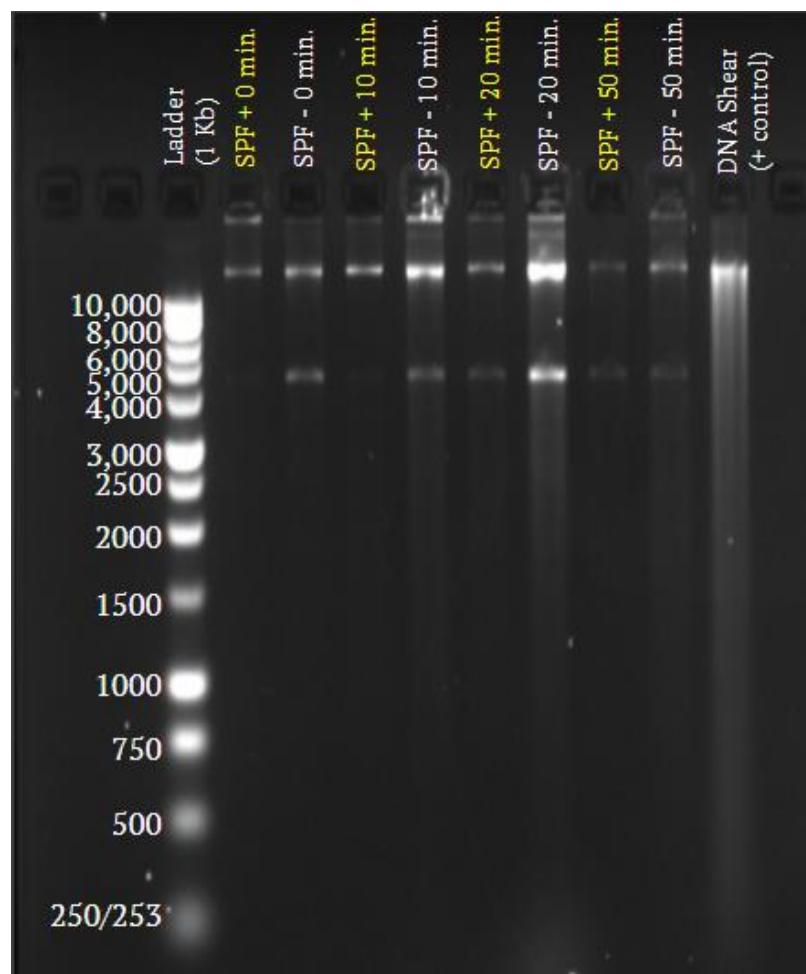
Group 3	Purity	Concentration ( $\mu\text{g/mL}$ )	Yield ( $\mu\text{g}$ )
+SPF t=0	1.746565	18.457	1.16742
+SPF t=20	1.485	14.867	0.74335
+SPF t=40	1.5096153	7.539	0.37695
+SPF t=60	1.625	7.132	0.3566
-SPF t=0	2.7281553	-12.038	-0.6019
-SPF t=20	1.8131868	-16.489	-0.82445
-SPF t=40	20.166667	-6.032	-0.3016
-SPF t=60	2.30555	-4.16	-0.208

**Table 4: Banana Boat Sport SPF 30**

Group 4	Purity	Concentration ( $\mu\text{g/mL}$ )	Yield ( $\mu\text{g}$ )
+SPF t=0	1.30	10.202	0.51
+SPF t=20	1.62	36.221	1.81
+SPF t=35	1.66	51.405	2.57
+SPF t=50	1.33	-35.001	-1.75
-SPF t=0	1.29	56.456	2.82
-SPF t=20	1.38	16.309	0.82

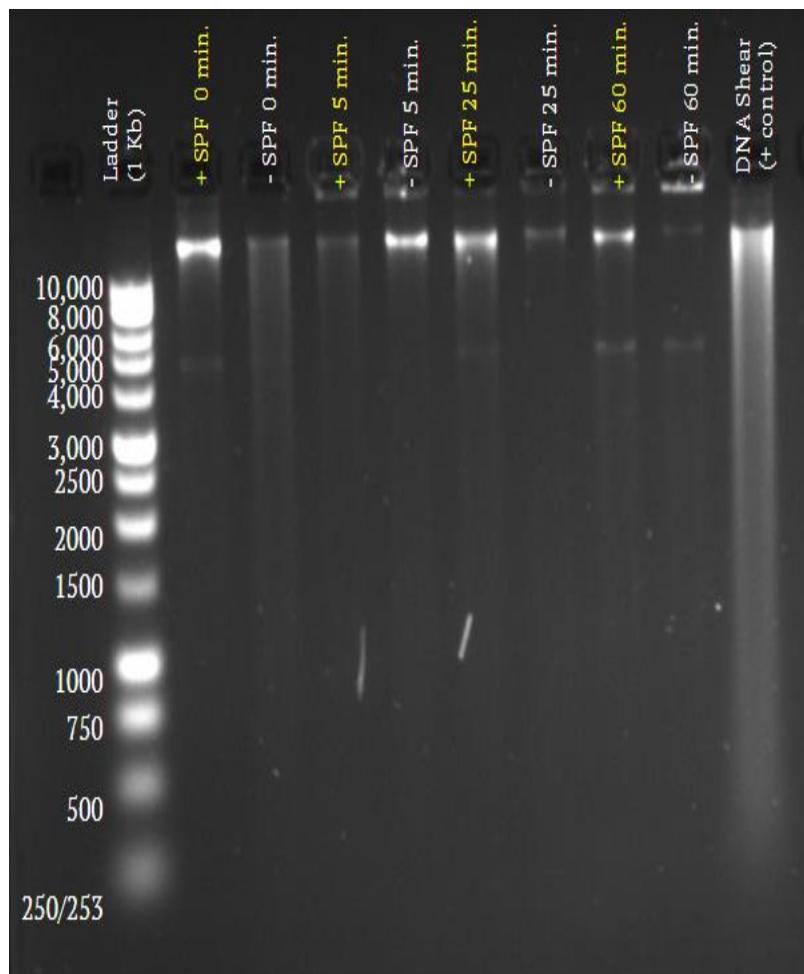
-SPF t=35	1.64	-6.675	-0.334
-SPF t=50	1.52	9.318	0.97

## 2. Gel Electrophoresis



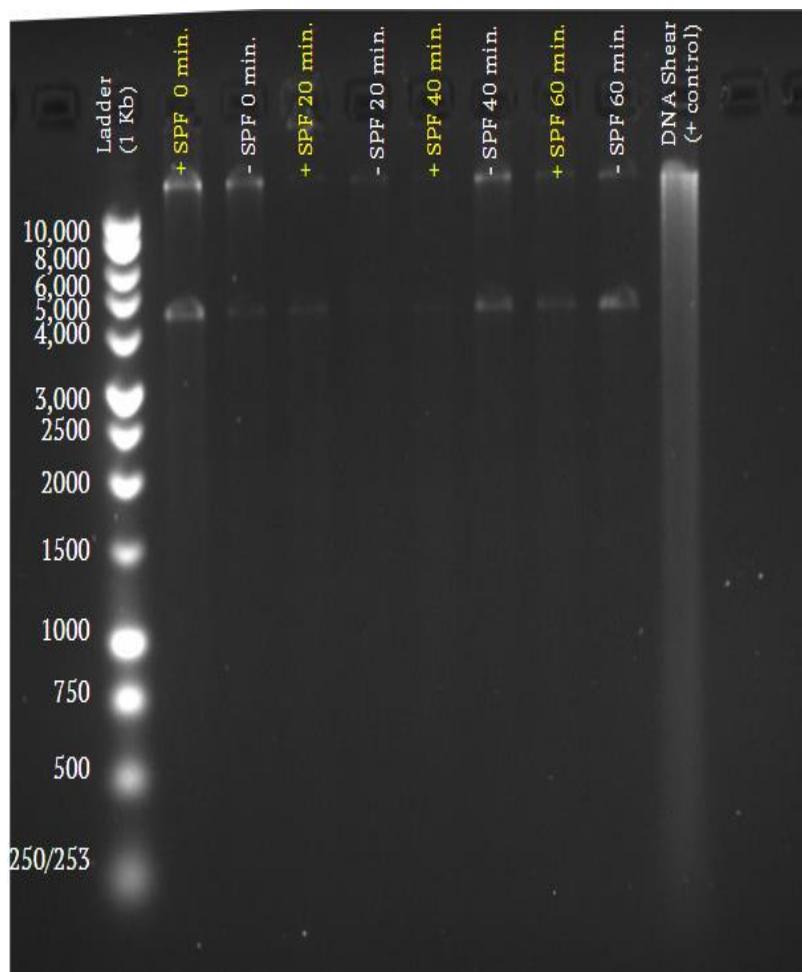
**Figure 20: Agarose Gel Electrophoresis of Genomic DNA Comparing Coppertone 50 SPF Sport Clear Sunscreen**

In Figure 20, the first lane is the 1 KB DNA ladder. The next eight lanes are samples of gDNA that were irradiated at different timepoints either with or without sunscreen. Lanes denoted as +SPF contained samples of gDNA that were covered by sunscreen during UV irradiation. Lanes denoted as -SPF contained samples of gDNA that were not covered by sunscreen during UV irradiation. The first two of these eight lanes are samples of gDNA that were not exposed to UV radiation. The next two were exposed to UV radiation for 10 minutes. The next two were exposed to UV radiation for 20 minutes. The next two were exposed to UV radiation for 50 minutes. The last lane on the gel is a DNA positive control which contains sheared DNA.



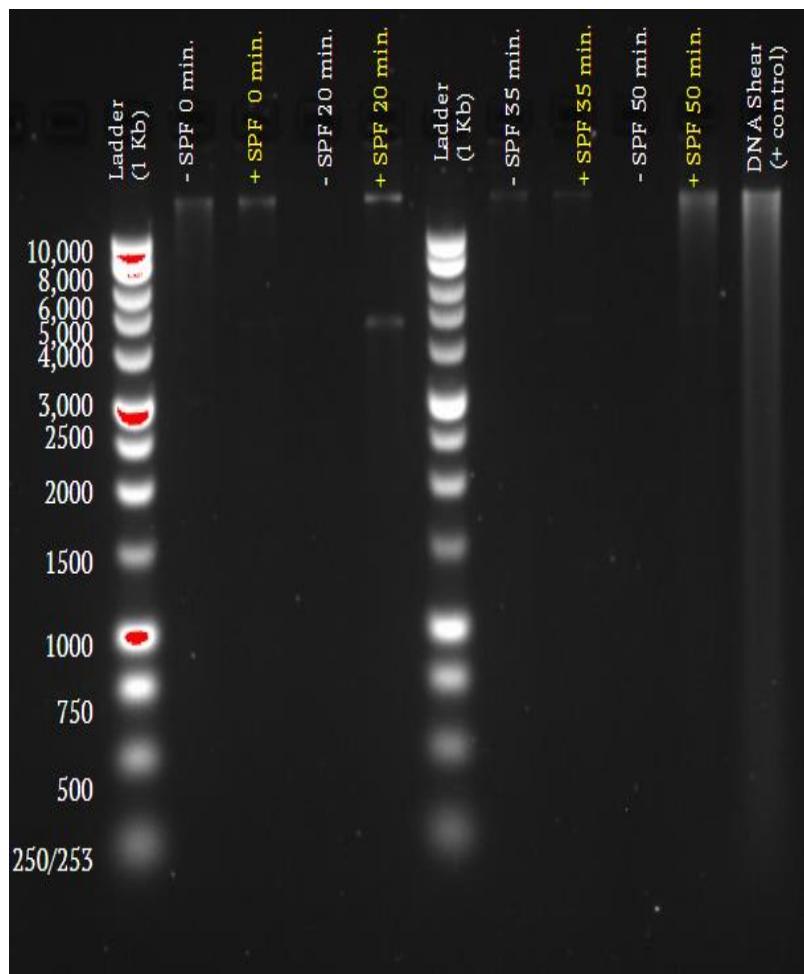
**Figure 21: Agarose Gel Electrophoresis of Genomic DNA Comparing Coppertone SPF 50 Spray**

In Figure 21, the first lane is the 1 KB DNA ladder. The next eight lanes are samples of gDNA that were irradiated at different timepoints either with or without sunscreen. Lanes denoted as +SPF contained samples of gDNA that were covered by sunscreen during UV irradiation. Lanes denoted as -SPF contained samples of gDNA that were not covered by sunscreen during UV irradiation. The first two of these eight lanes are samples of gDNA that were not exposed to UV radiation. The next two were exposed to UV radiation for 5 minutes. The next two were exposed to UV radiation for 25 minutes. The next two were exposed to UV radiation for 60 minutes. The last lane on the gel is a DNA positive control which contains sheared DNA.



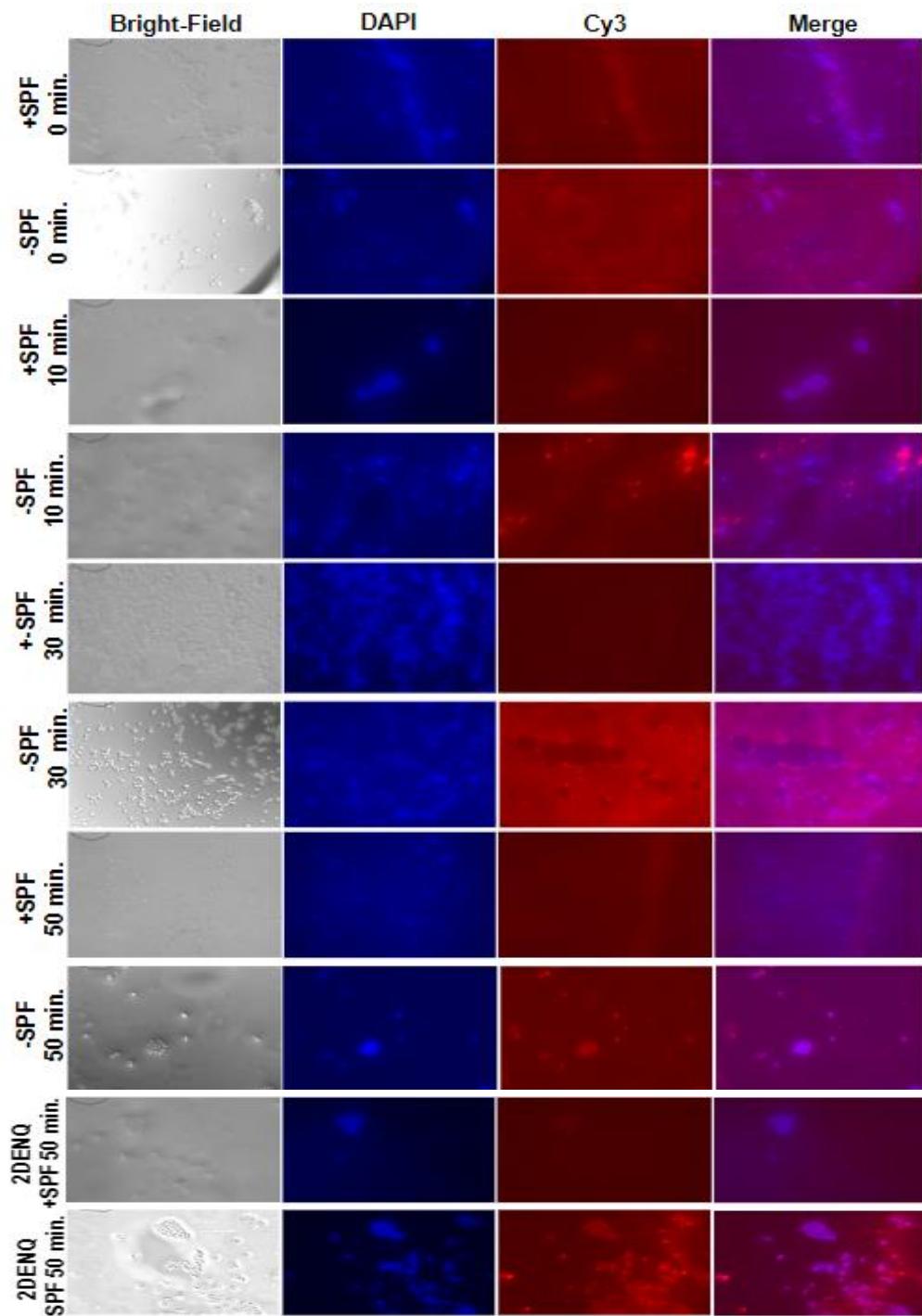
**Figure 22: Agarose Gel Electrophoresis of Genomic DNA Comparing Coppertone SPF 50 Sport Spray**

In Figure 22, the first lane is the 1 KB DNA ladder. The next eight lanes are samples of gDNA that were irradiated at different timepoints either with or without sunscreen. Lanes denoted as +SPF contained samples of gDNA that were covered by sunscreen during UV irradiation. Lanes denoted as -SPF contained samples of gDNA that were not covered by sunscreen during UV irradiation. The first two of these eight lanes are samples of gDNA that were not exposed to UV radiation. The next two were exposed to UV radiation for 20 minutes. The next two were exposed to UV radiation for 40 minutes. The next two were exposed to UV radiation for 60 minutes. The last lane on the gel is a DNA positive control which contains sheared DNA.

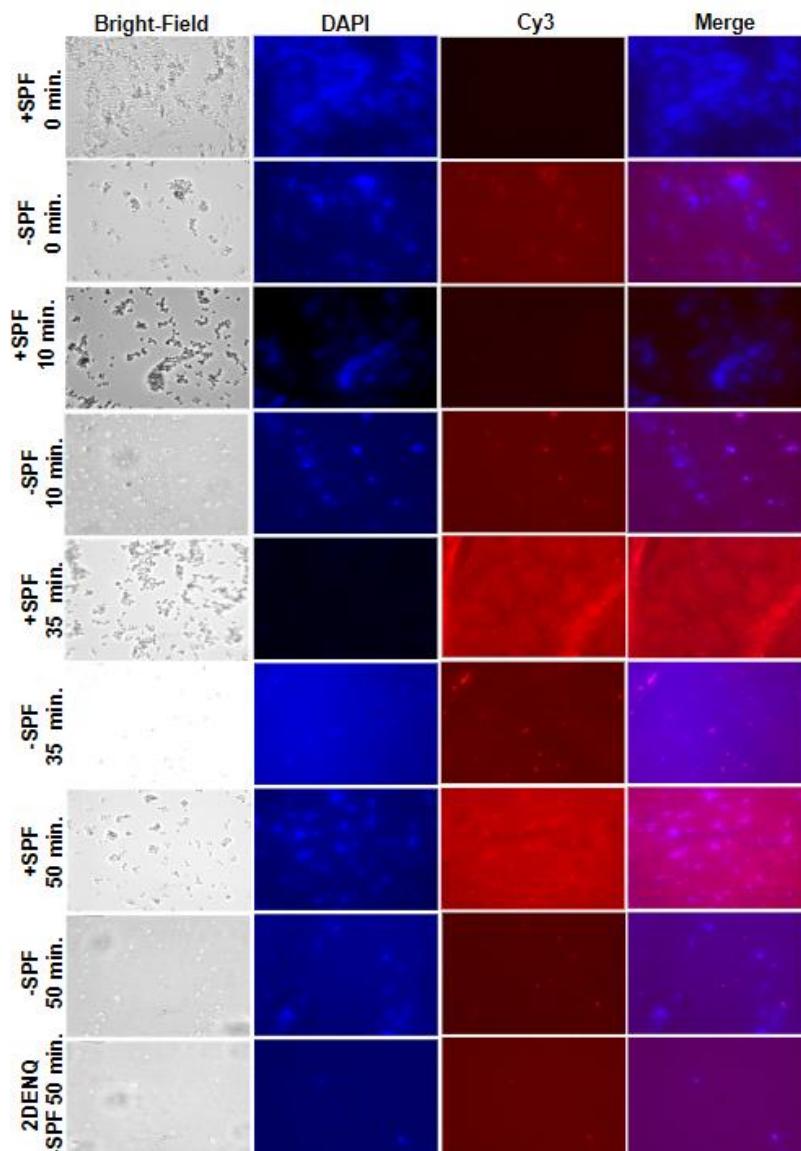


**Figure 23: Agarose Gel Electrophoresis of Genomic DNA Comparing Banana Boat SPF 30 Sport Spray**

In Figure 23, the first lane is the 1 KB DNA ladder. The next eight lanes are samples of gDNA that were irradiated at different timepoints either with or without sunscreen. Lanes denoted as +SPF contained samples of gDNA that were covered by sunscreen during UV irradiation. Lanes denoted as -SPF contained samples of gDNA that were not covered by sunscreen during UV irradiation. The first two of these eight lanes are samples of gDNA that were not exposed to UV radiation. The next two were exposed to UV radiation for 20 minutes. The next two were exposed to UV radiation for 35 minutes. The next two were exposed to UV radiation for 50 minutes. The last lane on the gel is a DNA positive control which contains sheared DNA.

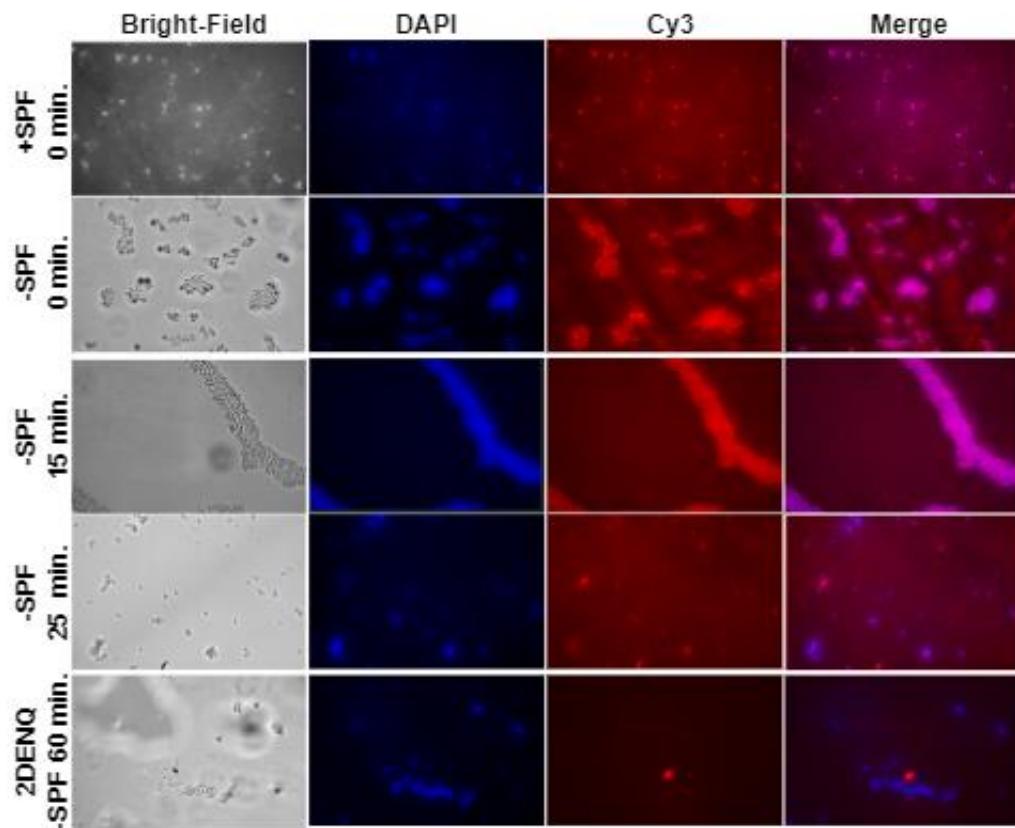
**C. Immunofluorescence****Figure 24: Immunofluorescence Imaging with SPF 70 Sunscreen**

In Figure 24. Illustrates yeast cells irradiated with and without Neutrogena SPF 70 sunscreen over various time points. Time points consisted of a 0-minute control, 10 minutes, 30 minutes, and 50 minutes.



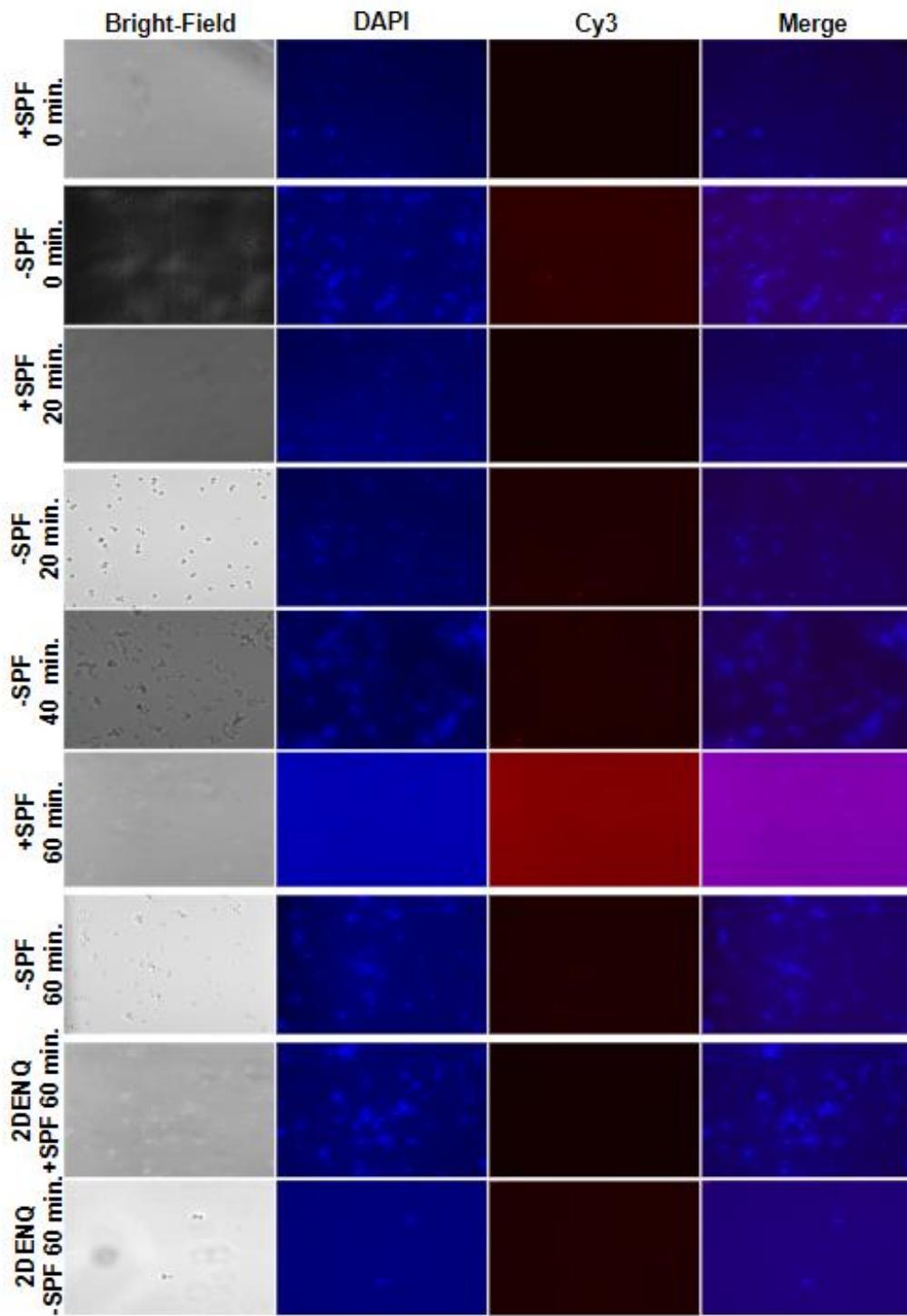
**Figure 25: Immunofluorescence Imaging with SPF 45 Banana Boat Sunscreen Lip Balm**

Figure 25 displays the indirect immunofluorescence imaging of yeast cells with and without SPF45 Banana Boat lip balm sunscreen after irradiation for various time points. Yeast were irradiated for a total of 50 minutes. Due to technical errors, images for +SPF 2DENQ were not captured, and the CY3 image for +SPF at 0 min is unusually dark.



**Figure 26: Immunofluorescence Imaging with SPF 50 Sunscreen**

Figure 26 displays the results of indirect immunofluorescence imaging of yeast cells with and without SPF50 sunscreen after irradiation for various time points. Yeast were irradiated for a total of 60 minutes. Slide 7, specifically wells 2-5, did not have significant amounts of yeast to be able to be observed under the light microscope. In addition, yeast cells were unable to be found in the 4th well in the 8th slide.



**Figure 27: Immunofluorescence Imaging with SPF 30 Sunscreen**

Figure 27 displays the results of direct immunofluorescence DAPI, Cy3, and a Merged column for yeast irradiated for a total of 60 minutes. Cells are identifiable in all conditions in the Bright-Field column. There is a relatively small red fluorescent signal in the CY3 and Merged columns.

## IV. Conclusions

### A. Killing Curves

The killing curves were constructed in order to examine the effectiveness of SPF containing products on yeast viability. When yeast were exposed to UV radiation exposure for 5 minutes without any protection, the majority of the yeast cells died after 90 seconds (Figure 13). Additionally, all yeast cells were dead after 5 minutes of exposure. It can thus be concluded that, in unprotected cells, UV radiation causes relatively high levels of cell death in a short period of time. There appears to be no significant difference in the percent survivors between organic and inorganic sunscreens of SPF 30 (Figure 14). As a result, it appears that organic and inorganic sunscreens of the same SPF provide similar levels of protection; based on the data, yeast protected by both sunscreens maintained greater than 50% survivors at every time point from 0 to 45 minutes. When comparing different levels of SPF in the same sunscreen brand, the percent survivors present in yeast protected by the SPF 50 sunscreen appeared to be greater than those protected by the SPF 30 sunscreen (Figure 15). At 60 minutes, the percent survivors of the SPF 50 plate were 45.45%, while those of the SPF 30 plate were 15%. Therefore, it appears that sunscreen of a higher SPF offers greater protection than sunscreen of a lower SPF.

When comparing different applications of the same brand of sunscreen, the percent survivors of yeast protected by the Coppertone SPF 50 Sport Clear appears to be significantly greater than in the yeast protected by the Coppertone SPF 50 Sport (Figure 16). While both sunscreens still offer UV protection, the clear sunscreen is more expensive and may be preferential for its consistency and lack of white color. Interestingly, it appears that the SPF 50 foundation offered the most UV protection as the percent survivors of the yeast were generally the highest out of those compared (Figure 17), at 95.61% after 30 minutes, 97.37% at 60 minutes, and 76.76% at 90 minutes. Furthermore, it appears that the foundation without SPF offered the least UV protection with % survivors at 75.70% after 30 minutes, 54.98% at 60 minutes, and 64.14% at 90 minutes. The combination of the foundation without SPF and the SPF 50 sunscreen offered the middle amount of protection, with % survivors at 93.54% at 30 minutes, 70.72% at 60 minutes, and 57.79% at 90 minutes. When examining the efficacy of non-SPF containing lotion and SPF 15 lotion, as well as yeast with no protection (Figure 18) nearly all of the yeast cells appeared to be dead after being irradiated for 20 minutes in all three cases. This indicates that lotion lacking SPF and lotion containing SPF 15 do not offer significant protection from UV radiation as the data was similar to the cells irradiated without protection. Furthermore, the data on the unprotected yeast cells supports the rapid and dramatic death of cells irradiated without protection. When comparing the efficacy of SPF 30 spray and the SPF 30 lotion when exposed to water (Figure 19), the majority of the cells died after 20 minutes of irradiation. As a result, it appears that sunscreens' effectiveness diminishes greatly after exposure to water, regardless of its nature as a spray or a lotion. However, more research will need to be conducted on this topic as we completed only one trial.

Overall, it can be concluded that without sunscreen protection, yeast cells die rapidly with less than 5 minutes of UV exposure. Based on the data, when sunscreen protection is used, yeast cells survive longer. Additionally, consumers can choose between organic and inorganic sunscreens at the SPF of their choice as they offer similar protection, and sunscreen with a higher SPF offers greater protection than sunscreen with a lower SPF. It also appears that clear sunscreen offers more protection than white sunscreen, and a foundation that includes SPF is ideal for UV protection while wearing makeup. Furthermore, lotion that does not contain SPF and SPF 15 lotion do not offer substantial protection, and when sunscreen is exposed to water, protection is also minimal.

## B. Nucleosome Fragmentation

The nucleosome fragmentation assay allows for examination of apoptosis induced DNA damage resulting from UV exposure. The 1KB DNA ladder in the first lane of each image marked known molecular weights of DNA fragments to compare the results of the samples to the ladder. The positive control in the far right lane of each image contains previously sheared DNA. As illustrated in Figure 20, there are very small amounts of DNA shearing in the + SPF samples at 10, 20, and 50 minutes of irradiation. In comparison, the - SPF samples at 10, 20, and 50 minutes of irradiation had greater amounts of DNA shearing. This shows a greater amount of DNA damage was induced by UV radiation when there was no sunscreen covering the yeast DNA samples. Also, the amount of DNA shearing increases between 10 and 20 minutes for the - SPF samples. The shearing in the - SPF 20-minute sample has a greater intensity and brightness than the shearing that occurred in the - SPF 10 minutes. However, this trend doesn't continue to the - SPF 50-minute sample, as the DNA shearing that occurs is much dimmer than the DNA shearing at the SPF 20 minute sample. There are two explanations for this unexpected data. During the extraction of genomic DNA, the amount of pure genomic DNA that was collected varied for each sample. The SPF 50-minute samples had a lower concentration and yield of genomic DNA than that of the other irradiated samples. As seen in Table 2, concentrations and yields for the 50 minute samples were less than both the 10 minute and 20 minute samples. Since the same volume was loaded into the gel, there would be less DNA present in the lane compared to the other samples. Another possible explanation is that many yeast cells could have died during the prolonged exposure to UV irradiation of the time period of 50 minutes. This would decrease the amount of viable genomic DNA within a sample to observe, which could also account for the dimmer shearing. Nevertheless, sunscreen was still effective in protecting against genomic DNA damage from UV radiation in the assay.

Figures 21 through 23 are images of other agarose gel electrophoresis assays that were taken during the nucleosome fragmentation experiment. Although the assays were conducted similarly to the data shown in Figure 20, there are no observable trends. This is because there is a lack of gDNA in many lanes per gel. Either not enough gDNA was loaded into the agarose gel for electrophoresis or not enough viable gDNA was able to be collected after extended periods of UV irradiation.

## C. Immunofluorescence

Indirect immunofluorescence microscopy was performed to visualize *S. cerevisiae* with DNA damage. Bright-Field images reveal the presence of *S. cerevisiae* under the slide. DAPI stained the nuclei blue. All cells should have a nucleus, whether damaged by UV radiation or not and should thus fluoresce with this stain. Cy3 is expected to reveal cells with damaged DNA, as the antibody complex was able to bind to the phosphorylated DNA (Figure 12). This phosphate group was initially attached to the disfigured DNA and flagged the site for nucleotide excision repair. DAPI and Cy3 were combined in the Merge image to distinguish between *S. cerevisiae* with DNA damage in the nucleus (which presents as red on blue) from debris and artifacts that can be picked up in the Bright-Field, DAPI, or Cy3 images individually. 2DENQ is genetically modified to have a higher rate of DNA damage, and thus acted as a positive control for indirect immunofluorescence microscopy.<sup>85</sup>

In Figure 24, which compares SPF 70 application, it's expected that there will be little to no DNA damage at T=0 for both SPF samples. This is supported by the corresponding Cy3 images, as they contain minimal bright red areas. Furthermore, there is minimal purple coloration in the Merge images for T=0, which is an indication of little to no sustained DNA damage in the yeast. Other expectations include increased purple coloration in the samples as time increases as well as increased purple in the -SPF rows compared with

the +SPF ones. For T=10 in the +SPF row, the results are very similar to those of T=0, with there being minimal DNA damage present in Cy3 and very little overlap of red and blue in the Merge image. However, in the -SPF row for T=10, there are bright red areas in the Cy3 image and purple coloration in the Merge image. This indicates DNA damage occurred during the ten-minute exposure time for the -SPF sample. The pattern continues for T=30. There is no observable DNA damage in the yeast for +SPF condition and consequently no purple coloration in the Merge image. Furthermore, the -SPF sample row contains damaged DNA from the Cy3 image and displays a few purple areas in the Merge image. For T=50, there is also very little bright red in the +SPF Cy3 image with minimal purple in the Merge image. By comparison, -SPF shows a larger amount of DNA damage with regards to the other -SPF timepoints. 2DENQ, being a positive control, indicates DNA damage in yeast and acts as a comparison mechanism for the other samples. This gives validity that our results are expected and we can detect DNA damage. Overall, it can be conferred that sunscreen is effective at protecting yeast cell DNA from UVR induced damage.

When comparing irradiated, unprotected yeast cells and cells protected with SPF 45 Banana Boat Lip Balm (Figure 25), very little conclusive evidence can be drawn about the effectiveness of sunscreen in protecting yeast cells from DNA damage. Due to technical errors, the Cy3 image for +SPF at 0 minutes and 35 minutes is unusually dark and unusable. As a result, the merge images for these data points cannot be compared with the -SPF images at those time points. Additionally, the +SPF sample of 2DENQ was lost, which provided no positive control for the +SPF category. Also, the remaining 2DENQ positive control sample appeared to have very few yeast cells to observe, perhaps due to the length of time the cells were irradiated without protection, thus making it difficult to analyze the -SPF 2DENQ as well. The only data point that appears to be unaffected by these errors is the 10-minute time point, where we can see evidence of a greater amount of DNA damage in the -SPF merged image than in the +SPF merged image, shown by the increased purple coloration on the -SPF image.

In Figure 26, Coppertone Spray SPF 50 was placed on half of the plates to observe the relationship between DNA damage and SPF protection. In the merge column, there is evidence that UV damage occurred in the unprotected yeast cells, though it was difficult to compare due to the lack of +SPF yeast in the microscope. The +SPF slide in times 10-40 minutes were unable to be recorded, as an insufficient number of colonies were observed after extended searching. However, there was a direct relationship between the time and DNA damage in the plates that were imaged, as there were a greater number of purple cells (DNA damaged cells) found when time increased. In time 0 minutes for both the +SPF and -SPF, there were many colonies that appeared purple, but after comparison between the images, much less actually sustained DNA damage from the UV radiation. Due to the coloring software of ImageJ, many more cells appeared purple in the composite image than actually overlapped between the DAPI and CY3. At 15 minutes, there was a significant increase in DNA damaged cells, which is likely the direct result of increased UV ray exposure. While observing the 25-minute sample, it was noted that there were significantly less yeast cells nuclei, perhaps due to the increased exposure to UV rays that caused increased cell death. However, almost all of the nuclei found from the DAPI were also found to be damaged in Cy3. The 2DENQ should have had many more damaged DNA, but due to a mistake in the data collection process, there was little to no DNA damage displayed. In addition, the well in time 50 in the -SPF slide was unable to be recorded as the number of colonies observed was insufficient. This may be the result of too many of the cells being killed before we measured them. In general, though this trial could be used to conclude that increased UV radiation causes increased DNA damage, the lack of proper +SPF comparisons causes any determination to be inconclusive based on just this trial.

DNA damage was expected to increase with UV exposure. In the experiment conducted with SPF 30 shown in Figure 27, a few cells in the -SPF 0, -SPF 20, and -SPF 40 seem to have undergone DNA damage. However, a relatively small Cy3 signal is generated for all of the conditions without respect to UV exposure

time. The lack of a strong Cy3 signal in the 2DENQ positive control indicates that the dampened signal in all conditions may not be a result of less DNA damage. Errors in the application of the antibody complex may have left damaged DNA unmarked, which may be why it is not seen on most of the images.

The cells that underwent UV treatment for 60 minutes with SPF 30 in Figure 27 are difficult to identify in the Bright-Field image and do not have a strong signal in the DAPI, Cy3, or Merge images. A possible explanation for this is an insufficient number of cells or inadequately stained cells on the slide. The lack of intensified fluorescent signals may cause the ImageJ software to false color the entire image with a homogeneous shade. Cells irradiated for 40 minutes were not able to be located on the epifluorescent microscope, so no fluorescent imaging was conducted on this population.

## V. Error Analysis

There are areas of the experiments we conducted where errors may have occurred that could have affected the results. In our serial dilutions, it is possible that the tubes of sterilized water were not properly sterilized throughout the experiment. This could have been done by not replacing micropipette tips after each use of the micropipette and by not heating the tip of the sterile water tube before and after mixing in the yeast culture. Improperly sterilizing throughout this experiment risks bacteria forming on the agar plates along with the yeast cells which could influence the results of the experiment. Also, the yeast culture may not have been mixed thoroughly with the vortex mixer during the serial dilution. This could have led to inaccurate dilutions of the yeast culture in sterile water and smaller amounts of yeast would be present in the solution.

One potential error that may have impacted the data is the difference in lab technique between groups. As there were many different trials being conducted by different members of this study, the difference in lab technique may have shifted the data depending on the group. While we tried to be as similar as possible, there were still minute differences in technique. Specifically, different groups used different techniques to spread the sunscreen over the plates, such as placing the plastic wrap over the plates before/after spreading. This may have resulted in either an increase in concentrated sunscreen that covered the yeast cells or a decrease in sunscreen concentration, as it may have been applied to more than just on top of the yeast cells. We could remedy this by having an individual carry out a certain experimental procedure or by continuing to have thorough discussions of experimental procedures before beginning experimentation.

Another error we made during the nucleosome fragmentation experiment was assuming that each sample of gDNA had the same amount of gDNA extracted. Small differences in extraction methods could lead to different concentrations and yields among the gDNA samples. A variety of gDNA concentrations and yields would allow for inconsistencies in data analysis on the gel electrophoresis assays. We could fix this issue by varying the amounts of gDNA that are loaded into the gel. A constant ng amount of DNA could be calculated from the concentrations determined. We could dilute the different amounts of gDNA with sterile water in order to load the same amounts of solution into each well.

## VI. Future Research

Given the conclusions of the study, there are several additional avenues that could have been discussed and explored.

## 1. Physical Barriers

In addition to testing the efficacy of sunscreen in protecting cells from UV radiation damage, other physical barriers could be tested. A relevant and interesting example is testing the impact of different types of fabrics in protecting cells from damage, simulating the effects of wearing clothing as a protective barrier on skin. The CDC recommends wearing long sleeve protective clothing when going outside to mitigate the harmful effects of UV rays. Thus, clothing fabric could act as an additional barrier in protecting the epidermis DNA damage. Experiments could be conducted to test the extent different types of fabric are effective in this regard. In a similar fashion to the application of sunscreen, the different fabrics could be placed over the agar plates before being irradiated and then tested at different time points to generate killing curves for analysis. This would help provide more insight into the effects of fabric in absorbing or reflecting harmful UV radiation.

## 2. Reflected UV light

Experiments could also be conducted to test the differences in the effects of reflected UV radiation. UV rays can be reflected off of a variety of sources, such as open water, sand, snow, grass, and concrete. Surfaces that are soft and rugged are less likely to reflect UV radiation in comparison to hard and smooth surfaces. For example, according to the WHO, snow can reflect as much as 80% UV radiation, while dry sand can reflect around 15%.<sup>77</sup> Future studies could conduct experiments with similar methods, utilizing different materials to mediate the reflection of UV radiation to irradiate the agar plates.

## 3. DNA Sequencing

Future research could investigate the DNA damage on a molecular level. The genomes of the irradiated yeast could be isolated and sequenced to provide more insight into the specific locations of the mutations. The results from these DNA sequences could then be compared with the human genome, which could help indicate conclusions about UV-induced DNA damage in the human genetics. Perhaps some regions are more susceptible to UVR than others. Determining if there is a relationship could help identify future methods to protect against UVR exposure.

## 4. Cell Staining for Free Radicals

Reactive oxygen species (ROS) can form in cells as a byproduct of UV exposure. ROS can mutate DNA, induce oxidative stress, and kill cells. To expand an understanding of the conditions that generate these molecules, the fluorescent probe DHE (dihydroethidium) could be used to mark super hydroxide molecules, and DCFDA (5-carboxy-2',7'-dichlorodihydrofluorescein diacetate) could be used to stain for free radicals including peroxy and hydroxyl radicals, and hydrogen peroxide.<sup>78,79</sup> Furthermore, highly destructive protein oxidation that occurs as a result of ROS exposure could be assessed using a protein carbonyl (PC) assay.<sup>80</sup> The levels of ROS found in each cell could be compared to the results of a comet assay (carried out via electrophoresis) to determine a relationship between the amount of free radicals and the amount of single and double stranded DNA breaks after UVR exposure. This would provide more insight on specific types of UV induced DNA damages and how different skin cancers can form.

## VII. References

1. The American Cancer Society medical and editorial content team. (2019, July 10). *Ultraviolet (UV) Radiation*. American Cancer Society. Retrieved July 16, 20233, from <https://www.cancer.org/healthy/cancer-causes/radiation-exposure/uv-radiation.html>
2. Centers for Disease Control and Prevention. (2022, July 5). *UV Radiation*. Centers for Disease Control and Prevention. Retrieved July 16, 2022, from <https://www.cdc.gov/nceh/features/uv-radiation-safety/index.html>
3. World Health Organization. (2016, March 9). *Radiation: Ultraviolet (UV) radiation*. World Health Organization. Retrieved July 28, 2022, from [https://www.who.int/news-room/questions-and-answers/item/radiation-ultraviolet-\(uv\)#:~:text=UV%20levels%20vary%20mainly%20with,most%20direct%20path%20to%20earth](https://www.who.int/news-room/questions-and-answers/item/radiation-ultraviolet-(uv)#:~:text=UV%20levels%20vary%20mainly%20with,most%20direct%20path%20to%20earth).
4. National Aeronautics and Space Administration. (2022, July 26). *Visible Light*. NASA. Retrieved July 26, 2022, from [https://science.nasa.gov/ems/09\\_visiblelight#:~:text=More%20simply%2C%20this%20range%20of,from%20380%20to%20700%20 nanometers](https://science.nasa.gov/ems/09_visiblelight#:~:text=More%20simply%2C%20this%20range%20of,from%20380%20to%20700%20 nanometers)
5. Swinburne University of Technology. (n.d.). *X-Rays*. COSMOS - The SAO Encyclopedia of Astronomy. Retrieved July 26, 2022, from <https://astronomy.swin.edu.au/cosmos/X/X-rays>
6. AIM. (2022, January 27). What is Ultraviolet Radiation? AIM at Melanoma Foundation. Retrieved July 25, 2022, from <https://www.aimatmelanoma.org/melanoma-101/prevention/what-is-ultraviolet-uv-radiation/>
7. D'Orazio, J., Jarrett, S., Amaro-Ortiz, A., & Scott, T. (2013, June 7). *UV radiation and the skin*. International journal of molecular sciences. Retrieved July 16, 2022, from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3709783/>
8. University of Iowa Hospitals & Clinics. (2018, April 4). What is the difference between UVA and UVB rays? University of Iowa Hospitals & Clinics. Retrieved July 16, 2022, from <https://uihc.org/health-topics/what-difference-between-uva-and-uvb-rays>
9. UN environment programme. (n.d.). *All about ozone and the ozone layer*. Ozone and You | Ozone Secretariat. Retrieved July 26, 2022, from <https://ozone.unep.org/ozone-and-you#:~:text=Ozone%20is%20being%20produced%20and,oxygen%20molecules%20into%20oxygen%20atoms>.
10. a Cleveland Clinic medical professional; (2021, October 19). *Epidermis*. Cleveland Clinic. Retrieved July 16, 2022, from <https://my.clevelandclinic.org/health/body/21901-epidermis>
11. a Cleveland Clinic medical professional. (2022, February 7). *Dermis*. Cleveland Clinic. Retrieved July 16, 2022, from <https://my.clevelandclinic.org/health/body/22357-dermis#:~:text=from%20Cleveland%20Clinic-,Your%20dermis%20is%20the%20middle%20layer%20of%20skin%20in%20your,take%20care%20of%20your%20dermis>.
12. a Cleveland Clinic medical professional. (2021, October 19). *Hypodermis (Subcutaneous Tissue)*. Cleveland Clinic. Retrieved July 16, 2022, from <https://my.clevelandclinic.org/health/body/21902-hypodermis-subcutaneous-tissue#:~:text=The%20hypodermis%20is%20the%20bottom%20layer%20of%20skin%20in%20your,your%20skin%20starts%20to%20sag>.
13. Ultraviolet (Uv) radiation. (n.d.). Retrieved July 28, 2022, from <https://www.cancer.org/healthy/cancer-causes/radiation-exposure/uv-radiation.html>
14. D'Orazio, J., Jarrett, S., Amaro-Ortiz, A., & Scott, T. (2013). UV radiation and the skin. International journal of molecular sciences, 14(6), 12222–12248. <https://doi.org/10.3390/ijms140612222>

15. Elwood, J. M., & Jopson, J. (1997). Melanoma and sun exposure: an overview of published studies. *International journal of cancer*, 73(2), 198–203. [https://doi.org/10.1002/\(sici\)1097-0215\(19971009\)73:2<198::aid-ijc6>3.0.co;2-r](https://doi.org/10.1002/(sici)1097-0215(19971009)73:2<198::aid-ijc6>3.0.co;2-r)
16. Nair, R., & Maseeh, A. (2012). Vitamin D: The "sunshine" vitamin. *Journal of pharmacology & pharmacotherapeutics*, 3(2), 118–126. <https://doi.org/10.4103/0976-500X.95506>
17. Bowman, S. L., & Marks, M. S. (2018). Shining a Light on Black Holes in Keratinocytes. *The Journal of investigative dermatology*, 138(3), 486–489. <https://doi.org/10.1016/j.jid.2017.11.002>
18. Kawada A. (2000). Risk and preventive factors for skin phototype. *Journal of dermatological science*, 23 Suppl 1, S27–S29. [https://doi.org/10.1016/s0923-1811\(99\)00074-2](https://doi.org/10.1016/s0923-1811(99)00074-2)
19. Ortonne J. P. (1990). The effects of ultraviolet exposure on skin melanin pigmentation. *The Journal of international medical research*, 18 Suppl 3, 8C–17C.
20. Wacker, M., & Holick, M. F. (2013). Sunlight and Vitamin D: A global perspective for health. *Dermato-endocrinology*, 5(1), 51–108. <https://doi.org/10.4161/derm.24494>
21. Cooper, G. M., Hausman, R. E., & Hausman, R. E. (2007). *The cell: a molecular approach* (Vol. 4, pp. 649-656). Washington, DC: ASM press. <https://library.wur.nl/WebQuery/titel/1837475>
22. Sinha, R. P., & Häder, D. P. (2002). UV-induced DNA damage and repair: a review. *Photochemical & Photobiological Sciences*, 1(4), 225-236. <https://doi.org/10.1039/B201230H>
23. de Lima-Bessa, K. M., Armelini, M. G., Chiganças, V., Jacysyn, J. F., Amarante-Mendes, G. P., Sarasin, A., & Menck, C. F. M. (2008). CPDs and 6-4PPs play different roles in UV-induced cell death in normal and NER-deficient human cells. *DNA repair*, 7(2), 303-312. <https://doi.org/10.1016/j.dnarep.2007.11.003>
24. Rastogi, R. P., Richa, Kumar, A., Tyagi, M. B., & Sinha, R. P. (2010). Molecular mechanisms of ultraviolet radiation-induced DNA damage and repair. *Journal of nucleic acids*, 2010, 592980. <https://doi.org/10.4061/2010/592980>
25. Sinha, R. P., & Häder, D. P. (2002). UV-induced DNA damage and repair: a review. *Photochemical & photobiological sciences: Official journal of the European Photochemistry Association and the European Society for Photobiology*, 1(4), 225–236. <https://doi.org/10.1039/b201230h>
26. Libretexts. (2020, August 11). 3.11: 3.34 N-glycosidic bonds. Chemistry LibreTexts. Retrieved July 27, 2022, from [https://chem.libretexts.org/Courses/Middle\\_Georgia\\_State\\_University/MGA\\_CHEM\\_1152K\\_Survey\\_of\\_Chemistry\\_II/Chapters/03%3A\\_More\\_Reactions\\_and\\_Properties/3.11%3A\\_3.34\\_N-glycosidic\\_bonds](https://chem.libretexts.org/Courses/Middle_Georgia_State_University/MGA_CHEM_1152K_Survey_of_Chemistry_II/Chapters/03%3A_More_Reactions_and_Properties/3.11%3A_3.34_N-glycosidic_bonds)
27. Pellegrino, S. (n.d.). *Example of ber (left) and NER (right) pathways. A) in ber a single ...* Retrieved July 29, 2022, from [https://www.researchgate.net/figure/Example-of-BER-left-and-NER-right-pathways-A-In-BER-a-single-damaged-base-is\\_fig4\\_280046956](https://www.researchgate.net/figure/Example-of-BER-left-and-NER-right-pathways-A-In-BER-a-single-damaged-base-is_fig4_280046956)
28. Schärer O. D. (2013). Nucleotide excision repair in eukaryotes. *Cold Spring Harbor perspectives in biology*, 5(10), a012609. <https://doi.org/10.1101/cshperspect.a012609>
29. Sidransky, E. (n.d.). *Homologous recombination*. Genome.gov. Retrieved July 27, 2022, from <https://www.genome.gov/genetics-glossary/homologous-recombination>
30. Lee, C. H., Wu, S. B., Hong, C. H., Yu, H. S., & Wei, Y. H. (2013). Molecular Mechanisms of UV-Induced Apoptosis and Its Effects on Skin Residential Cells: The Implication in UV-Based Phototherapy. *International journal of molecular sciences*, 14(3), 6414–6435. <https://doi.org/10.3390/ijms14036414>
31. Fadda, E. (2015, December 8). *Role of the XPA protein in the NER pathway: A perspective on the function of structural disorder in Macromolecular Assembly*. Computational and structural biotechnology journal. Retrieved July 27, 2022, from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4710682/#:~:text=Xeroderma%20pigmentosum%>

- 20complementation%20group%20A%20(XPA)%20is%20an%20essential%20protein,both%20eu karyotic%20and%20prokaryotic%20cells.
- 32. Elmore S. (2007). Apoptosis: a review of programmed cell death. *Toxicologic pathology*, 35(4), 495–516. <https://doi.org/10.1080/01926230701320337>
  - 33. Kulms, D., & Schwarz, T. (2000). Molecular mechanisms of UV-induced apoptosis. *Photodermatology, photoimmunology & photomedicine*, 16(5), 195–201. <https://doi.org/10.1034/j.1600-0781.2000.160501.x>
  - 34. Stern, RS. Prevalence of a history of skin cancer in 2007: results of an incidence-based model. *Arch Dermatol* 2010; 146(3):279-282
  - 35. Koh HK, Geller AC, Miller DR, et al. Prevention and early detection strategies for melanoma and skin cancer: Current status. *Arch Dermatol* 1996; 132(4):436-442.
  - 36. Chen JT, Kempton SJ, Rao VK. The Economics of Skin Cancer; AN Analysis of Medicare Payment Data. *Plastic and Reconstructive Surgery—Global Open*. 2016; 4(9):868
  - 37. Rogers HW, Weinstock MA, Feldman SR, Coldiron BM. Incidence estimate of nonmelanoma skin cancer (keratinocyte carcinomas) in the US population, 2012. *JAMA Dermatol* 2015; 151(10):1081-1086.
  - 38. Daya-Grosjean L, Couve-Privat S. Sonic hedgehog signaling in basal cell carcinomas. *Cancer Lettres* 2005; 225(2):181-192
  - 39. Mayo Clinic Staff. Basal cell carcinoma. Mayo Foundation for Medical Education and Research. 2021
  - 40. Maner BS, Dupuis L, Su A, Jueng JJ, Harding TP, Meisenheimer J, Siddiqui FS, Hardack MR, Aneja S, Solomon JA. Overview of genetic signaling pathway interactions within cutaneous malignancies. *J Cancer Metastasis Treat* 2020;6:37
  - 41. Mansouri B, Housewright C. The treatment of actinic keratoses—the rule rather than the exception. *J Am Acad Dermatol* 2017; 153(11):1200.
  - 42. Cancer Facts & Figures 2022, American Cancer Society (ACS), Atlanta, Georgia, 2022. <https://www.cancer.org/facts-statistics/>
  - 43. Sandru A, Voinea S, Panaiteescu E, Blidaru A. Survival rates of patients with metastatic melanoma. *J Med Life*. 2014;7(4):572-576
  - 44. Lew RA, Sober AJ, Cook N, et al. Sun exposure habits in patients with cutaneous melanoma: a case study. *J Dermatol Surg Onc* 1983; 12:981-6.
  - 45. Bauer J, Büttner P, Murali R, et al. BRAF mutations in cutaneous melanoma are independently associated with age, anatomic site of the primary tumor, and the degree of solar elastosis at the primary tumor site. *Pigment Cell Melanoma Res*. 2011;24(2):345–351.
  - 46. Viros A, Sanchez-Laorden B, Pedersen M, et al. Ultraviolet radiation accelerates BRAF-driven melanomagenesis by targeting TP53. *Nature*. 2014;511(7510):478–482.
  - 47. Caceres-Dittmar G, Ariizumi K, Xu S, Tapia FJ, Bergstresser PR, Takashima A. Hydrogen peroxide mediates UV-induced impairment of antigen presentation in a murine epidermal-derived dendritic cell line. *Photochem. Photobiol.* 1995;62:176–183.
  - 48. Aberer W, Schuler G, Stingl G, Honigmann H, Wolff K. Ultraviolet light depletes surface markers of Langerhans cells. *J. Invest. Dermatol.* 1981;76:202–210.
  - 49. Toews GB, Bergstresser PR, Streilein JW. Epidermal Langerhans cell density determines whether contact hypersensitivity or unresponsiveness follows skin painting with DNFB. *J. Immunol.* 1980;124:445–453.
  - 50. Ullrich, SE. Mechanisms underlying UV-induced immune suppression. *Mutation Research*. 2005;571(1-2):185-205
  - 51. *Sunburn: Causes, symptoms and treatment*. Cleveland Clinic. (n.d.). Retrieved July 27, 2022, from <https://my.clevelandclinic.org/health/diseases/21858-sunburn>

52. *Science of skincare: The stages of a Sunburn.* Hale Cosmeceuticals Inc - Your Skin HEALTH Company. (2021, March 5). Retrieved July 27, 2022, from <https://www.halecosmeceuticals.com/blog/stages-of-a-sunburn/>
53. Sharkey, L. (2019, July 30). *Can dark-skinned people get sunburn? what to know.* Healthline. Retrieved July 27, 2022, from <https://www.healthline.com/health/can-black-people-get-sunburn#sunburn-risk>
54. *Infographic: The 4 stages of Sunburn.* Treated.com. (n.d.). Retrieved July 27, 2022, from <https://www.treated.com/blog/infographic-the-4-stages-of-sunburn>
55. Premier Dermatology. (2018, March 22). *UV index: The Sun Safety Scale.* Premier Dermatology. Retrieved July 27, 2022, from <https://pdskin.com/blogs/uv-index-the-sun-safety-scale/>
56. "UV Index Scale." EPA, Environmental Protection Agency, 28 Dec. 2016, [https://19january2017snapshot.epa.gov/sunsafety/uv-index-scale-1\\_.html](https://19january2017snapshot.epa.gov/sunsafety/uv-index-scale-1_.html).
57. "UK Weather: What Is the UV Index and Why Could It Break UK Records?" BBC News, BBC, 2 June 2020, <https://www.bbc.com/news/explainers-53178988>.
58. *Sunburn: Causes, symptoms and treatment.* Cleveland Clinic. (n.d.). Retrieved July 27, 2022, from <https://my.clevelandclinic.org/health/diseases/21858-sunburn>
59. *Four different types of sunburn (and how to avoid them).* Doctors in Italy. (2022, July 21) Retrieved July 27, 2022, from <https://www.doctorsinitaly.com/b/4-types-sunburn-and-how-to-avoid-them/>
60. Admin. "Identifying a Severe Burn." Elite Medical Center, 27 Aug. 2020, <https://elitelv.com/identifying-a-severe-burn/>.
61. *Vitamin D and UV.* Science Learning Hub. (n.d.). Retrieved July 27, 2022, from <https://www.sciencelearn.org.nz/resources/1313-vitamin-d-and-uv>
62. Kannan, S., & Lim, H. W. (2013, December 9). *Photoprotection and vitamin D: A Review - Wiley Online Library.* Wiley Online Library. Retrieved July 27, 2022, from <https://onlinelibrary.wiley.com/doi/10.1111/phpp.12096>
63. Engelsen, O. (2010, May 4). *The relationship between ultraviolet radiation exposure and vitamin D status.* MDPI. Retrieved July 27, 2022, from <https://www.mdpi.com/2072-6643/2/5/482>
64. What is Organic & Inorganic Sunscreen? Cosmetics Business. (n.d.). Retrieved July 27, 2022, from [https://www.cosmeticsbusiness.com/news/article\\_page/What\\_is\\_organic\\_and\\_inorganic\\_sunscreen/153042](https://www.cosmeticsbusiness.com/news/article_page/What_is_organic_and_inorganic_sunscreen/153042)
65. Fiza Pirani, A. J.-C. (2019, May 17). Study finds 73% of sunscreens don't even work - how to find one that does. Boston 25 News. Retrieved July 27, 2022, from <https://www.boston25news.com/news/trending-now/study-finds-73-of-sunscreens-dont-even-work-how-to-find-one-that-does/950232314/>
66. [https://www.mun.ca/biology/scarr/Thymine\\_Dimer\\_Repair.html](https://www.mun.ca/biology/scarr/Thymine_Dimer_Repair.html)
67. *All About Sunscreen.* The Skin Cancer Foundation. (2022, January 28). Retrieved July 15, 2022, from <https://www.skincancer.org/skin-cancer-prevention/sun-protection/sunscreen/>
68. Byrne, A. (2020, April 29). *Is expensive sunscreen worth it?* dealnews. Retrieved July 15, 2022, from <https://www.dealnews.com/features/Does-Price-Matter-When-Buying-Sunscreen/1992845.html>
69. *Which is best for optimal sun protection - sprays or lotions?* Harvard Health. (2017, June 1). Retrieved July 15, 2022, from <https://www.health.harvard.edu/skin-cancer/which-is-best-for-optimal-sun-protection-sprays-or-lotions>
70. Wing, Janet. "What SPF Should You Use? Here's What Dermatologists Recommend." Florida Medical Clinic, 9 Aug. 2021, <https://www.floridamedicalclinic.com/blog/what-spf-should-you-use/>.
71. Hsu, Charlotte. (2020). *Department of Biological Sciences.* University at Buffalo. (2022, July 20). Retrieved July 27, 2022, from <https://arts-sciences.buffalo.edu/biological-sciences.html>

72. Why use yeast in research? (n.d.). Yourgenome. Retrieved July 27, 2022, from <https://www.yourgenome.org/facts/why-use-yeast-in-research/>
73. Botstein, D., Chervitz, S. A., & Cherry, M. (1997). Yeast as a model organism. *Science*, 277(5330), 1259-1260.
74. Yao, S., Feng, Y., Zhang, Y., & Feng, J. (2021). DNA damage checkpoint and repair: From the budding yeast *Saccharomyces cerevisiae* to the pathogenic fungus *Candida albicans*. *Computational and Structural Biotechnology Journal*, 19, 6343-6354.
75. "The OD600 Basics: Best OD600 Tool to Generate Microbial Growth Curves." *Implen NanoPhotometers. The Best in Spectrophotometers.*, 5 Nov. 2021, <https://www.implen.de/od600-diluphotometer/od600/>.
76. Saginala K, Barouk A, Aluru JS, Rawla P, Barsouk B. Epidemiology of Melanoma. *Med Sci*. 2021;9(4):63
77. Radiation: Ultraviolet (UV) radiation. (2016). World Health Organization. [https://www.who.int/news-room/questions-and-answers/item/radiation-ultraviolet-\(uv\)](https://www.who.int/news-room/questions-and-answers/item/radiation-ultraviolet-(uv))
78. Ubezio P., Civoli F. Flow cytometric detection of hydrogen peroxide production induced by doxorubicin in cancer cells. *Free Radical Biology & Medicine*. 1994;16(4):509–516. doi: 10.1016/0891-5849(94)90129-5 AND
79. Benov L., Sztejnberg L., Fridovich I. Critical evaluation of the use of hydroethidine as a measure of superoxide anion radical. *Free Radical Biology & Medicine*. 1998;25(7):826–831. doi: 10.1016/S0891-5849(98)00163-4
80. Berlett B. S., Stadtman E. R. Protein oxidation in aging, disease, and oxidative stress. *Journal of Biological Chemistry*. 1997;272(33):20313–20316. doi: 10.1074/jbc.272.33.20313
81. Hopkins, R. (n.d.). How ultraviolet light reacts in cells. Scitable; Nature. [https://www.nature.com/scitable/blog/scibytes/how\\_ultraviolet\\_light\\_reacts\\_in/#:~:text=Melanin%20is%20different.,like%20DNA%2C%20from%20UV%20exposure.](https://www.nature.com/scitable/blog/scibytes/how_ultraviolet_light_reacts_in/#:~:text=Melanin%20is%20different.,like%20DNA%2C%20from%20UV%20exposure.)
82. Ancans, J., Tobin, D. J., Hoogduijn, M. J., Smit, N. P., Wakamatsu, K., & Thody, A. J. (2001). Melanosomal pH controls rate of melanogenesis, eumelanin/pheomelanin ratio and melanosome maturation in melanocytes and melanoma cells. *Experimental cell research*, 268(1), 26–35. <https://doi.org/10.1006/excr.2001.5251>
83. Naik, P. P., & Farrukh, S. N. (2022). Influence of Ethnicities and Skin Color Variations in Different Populations: A Review. *Skin pharmacology and physiology*, 35(2), 65–76. <https://doi.org/10.1159/000518826>
84. Nasti, T. H., & Timares, L. (2015). MC1R, eumelanin and pheomelanin: their role in determining the susceptibility to skin cancer. *Photochemistry and photobiology*, 91(1), 188–200. <https://doi.org/10.1111/php.12335>
85. Vijayraghavan, S., Tsai, F-L., Schwacha, A. (2016). A Checkpoint-Related Function of the MCM Replicative Helicase Is Required to Avert Accumulation of RNA:DNA Hybrids during Sphase and Ensuing DSBs during G2/M. *PLoS Genetics*, 12(8). <http://doi.org/10.1371/journal.pgen.1006277>
86. O'Connor, C. (2018, December 28). 4.3: Yeast growth phases. Biology LibreTexts. [https://bio.libretexts.org/Bookshelves/Cell\\_and\\_Molecular\\_Biology/Book%3A\\_Investigations\\_in\\_Molecular\\_Cell\\_Biology\\_\(O'Connor\)/04%3A\\_Working\\_with\\_Yeast/4.03%3A\\_Yeast\\_growth\\_phases](https://bio.libretexts.org/Bookshelves/Cell_and_Molecular_Biology/Book%3A_Investigations_in_Molecular_Cell_Biology_(O'Connor)/04%3A_Working_with_Yeast/4.03%3A_Yeast_growth_phases)

# **Host-Microbiota Interactions in Disease Model *C. Elegans***

Krishna Chinnasamy, Charlotte Devlin, Korina Goldsmith, Xi Lin,  
Lindsay McBride, Lanee Olson, Neel Pandula, Emma Riddell, Catie Smith

## **Abstract**

While the microbiota has been increasingly correlated with human health and wellness, the role of individual bacteria in disease expression and prevention remains unclear. The nematode *Caenorhabditis elegans* serves as a relevant model for a variety of human diseases and conditions ranging from neurodegenerative disorders like Parkinson's disease to the aging process due to its well understood genome and simple nervous system. Our project focused on the relationship between the bacterial diets fed to disease model *C. elegans* and behavior patterns observed through both qualitative analysis and thrashing assays. We studied motor-disease models of *C. elegans* for Amyotrophic Lateral Sclerosis, Duchenne Muscular Dystrophy, and Parkinson's Disease and used known *E. coli* bacterial strains as well as locally-sourced samples. 16S rRNA sequencing identified three out of seven strains collected; results indicated that these strains affected the worms in varying degrees of compatibility, with one isolated from the Mellon Institute sterile room light switch inhibiting thrashing. Known bacteria differentially altered behavior as well; subgroups observed the highest relative initial thrashing rates in both disease models and the wild-type strain given the HT115 bacterial diet. Disease models were more affected than wild type by different bacteria diets. Generally, bacterial diet influences behavioral responses in *C. elegans*, which has future implications for personalized medicine that may involve the microbiota.

## **I. Introduction**

### **A. Relevance of the Microbiota**

A "hidden organ" of the human body, the microbiota is home to a diverse community of microorganisms that interact with each other and their human host. In the case of a host, the microbiota interacts with it in a complex, symbiotic manner (2). Recent advancements in bioanalytical technology have ushered in an explosion of studies aiming to better understand this symbiotic relationship as well as the composition and function of the human microbiota (1). These studies indicate that host-microbiota interactions can be beneficial to the host organism, as is the case of the human gastrointestinal tract, where native microorganisms aid in digestion and protect against foreign pathogens. However, they can also be harmful (3). In the case of humans, disrupting the host-microbiota relationship has been correlated with conditions like cancer, obesity, and autism (4). This appears like a drastic leap, but recent studies have begun to elucidate the role of microbes in disease. The pathogenesis of Alzheimer's Disease and Parkinson's Disease both begin with an imbalance in the microbiota, and possibly from microbes themselves — it is theorized that infection with the herpes simplex virus type 1 may disable protective proteins implicated in Alzheimer's (5). Research is ongoing, but broadly, improving understanding of such mechanisms and effects of host-microbiome interactions has the potential to lead to improved treatment approaches and options for a broad spectrum of diseases.

### **B. Use of *C. elegans* as a Model Organism**

*Caenorhabditis elegans* has been identified as a useful model organism to study host-microbiota interaction because of its short life cycle, well-understood physiology, and ease of cultivation (5). A nematode worm found primarily in damp vegetative environments, *C. elegans* was the first multicellular

organism to have its genome sequenced. With only 306 neurons in comparison to humans' 86 billion, the simplicity of *C. elegans* combined with its well-understood genetic similarity to humans makes it a useful model for studying human disease. Every cell fate has been mapped in *C. elegans*. For this reason, many novel genomic techniques are tested in the species. *C. elegans* research is applicable to human diseases like cancer, neurodegeneration (e.g. Alzheimer's or ALS), neurodevelopmental conditions, muscular dystrophies, and aging (not a disease itself, but a contributing factor) (6).

*C. elegans* is known to feed on a variety of bacteria; nearly all observed behaviors of *C. elegans* are affected by their diet (7). In a laboratory setting, *C. elegans* are most commonly fed the *E. coli* strain OP50 due to its availability and clear visibility on an agar plate (8). While the natural diet of *C. elegans* is significantly more varied than the lab diet, the OP50 bacterial strain has been shown to have no pathogenic effect on the nematodes. Additionally, there is limited research in regards to ecologically-relevant bacteria. (8).

### **C. Studying Human Disease in *C. elegans***

The biological simplicity of *C. elegans* allows for effective modeling of various diseases through mutant strains of the organism (9). *C. elegans*' short life span combined with a manageable-sized genome allows experts to induce genetic mutations that mimic human disorders, which can then be manipulated and studied. Each strain can vary in terms of size, movement patterns, reproduction, color, and other physical and behavioral characteristics (10). The most commonly recognized natural strain of *C. elegans* is known as the N2 Bristol wild type strain, which is generally used as a control variable in disease research (11).

When studying diseases like Parkinson's disease, amyotrophic lateral sclerosis (ALS), or Duchenne muscular dystrophy, specific model strains of *C. elegans* such as NL5901, a model of Parkinson's disease, AM725 and CB1091, models of ALS, and LS587, a model of Duchenne muscular dystrophy, may all be treated with various bacteria and observed for changes in symptoms or behavior (12). By studying the behavior of mutant strains of *C. elegans* alongside N2 wild type strains, potential conclusions can be drawn about certain diseases through facile *in vivo* experimentation. In this study, we primarily tested the impact of bacterial strains on *C. elegans* by studying their movement patterns using thrashing assays (12).

#### **1. Parkinson's Disease**

Parkinson's disease is a progressive neurodegenerative disorder that can be characterized by both motor and non-motor-related symptoms. Resulting from a loss or degeneration of dopaminergic neurons in the substantia nigra of the midbrain, Parkinson's disease is associated with risk factors including aging, family history, pesticide exposure, and environmental chemicals (13). Affected individuals classically present with tremors, rigidity, bradykinesia, and stooping posture. Biomarkers are required to detect the disease in the early stages when physical symptoms are almost non-existent (14), so Parkinson's is most often diagnosed in its later stages after neuron degeneration. Various biomarkers include those of imaging, cerebrospinal fluid, oxidative stress, neuroprotection, and inflammation (14). There is no definitive neuroprotective treatment for Parkinson's, but possible symptom suppressing routes are available like pharmaceuticals, behavioral therapies, and deep brain stimulation.

*C. elegans* is a useful model organism to elucidate the disease mechanisms associated with Parkinson's. The transgenic strain of worms we used were engineered to express  $\alpha$ -synuclein, which causes the exhibition of phenotypic deficits, including the loss of dopamine neurons, disruption of dopamine-dependent behaviors, increased sensitivity to stress, age-dependent aggregation, and deficits in movement (15). As a result, these phenotypes can be used as outcome measures to gain insight into disease pathogenesis and to identify disease modifiers (15).

## 2. Duchenne Muscular Dystrophy

Duchenne Muscular Dystrophy, also known as DMD, is an X-linked recessive genetic disease found primarily in males (16). Considered the most severe form of the muscular dystrophies (17), it results in a complete lack of dystrophin, a protein complex that protects and strengthens muscle fibers. This causes frail, easily damaged muscle cells. The disease has an early onset in childhood often between two and three years with a high rate of mortality. It affects muscles closest to the core of the body before causing progressive deterioration towards the extremities (17). In addition to the gradual loss of motor functions such as walking, children with DMD may also display calf muscle hypertrophy (abnormal enlargement of the calves), learning disabilities, and may later develop cardiac and respiratory inadequacies (16).

The dystrophin gene in humans and the *dys-1* gene in *C. elegans* are strikingly similar, and the induction of mutation to the gene in *C. elegans* produces clinically significant phenotypes; for example, *C. elegans* with mutated *dys-1* genes display the motor and strength decline, gait abnormalities, and decreased lifespan symptoms comparable to human sufferers of DMD (18). While there are a few *C. elegans* strains for DMD, strain LS587 *dys-1(cx18;hlh-1)* was used exclusively in this project.

## 3. Amyotrophic Lateral Sclerosis

Amyotrophic lateral sclerosis (ALS) is a progressively fatal neurodegenerative disease characterized by the degeneration of both the upper and lower motor neurons (19). Although the pathogenesis that drives the degradation is still uncertain, scientists have identified some genetic factors: dysfunction in RNA metabolism and protein homeostasis, defects in nucleocytoplasmic trafficking, impaired dynamics of ribonucleotide bodies (RNA granules, liquid-liquid separation) and the induction of stress in the endoplasmic reticulum (20). First identified in 1995, TAR DNA protein 43 (TDP43) is a RNA/DNA protein whose hyperphosphorylated and ubiquitinated deposits are inclusion bodies in the brain and spinal cord of patients with ALS. While most ALS cases (90-95%) are sporadic rather than familial, the majority of patients contain the TDP43 neuronal inclusions, which suggests that the TDP43 could have a pivotal role in ALS pathology. In 2006, TDP43 was identified as a key factor in the brains of ALS patients; pathological hallmarks include protein truncation of toxic C-terminal TDP43 fragments, protein aggregation, among others. However, the etiology of the deposits are being debated as to whether TDP43 deposited cells have amyloid-like aggregate features or filament-like structures that don't stain with amyloid-specific dyes (20).

The only drugs approved by the FDA are riluzole<sup>79</sup> and edaravone, which only provide a limited improvement to patients; the former suppresses excessive motor neuron firing and the latter suppresses oxidative stress. Currently, ALS patients can only rely on managing symptoms via interventions such as nasogastric feeding and cough-assist devices (19). *C. elegans* are an analogous model for the disease pathology of ALS in humans, thus they are a fantastic aid in research for better and more-effective cures. Compared to the N2 wildtype strain of *C. elegans*, the ALS model is equivalent in size and early movement patterns, but as the disease model progresses, the strain will begin to express behavioral phenotypes that are correlated to human symptoms of the disease. An ALS model *C. elegans* exhibits a slower and more-coiled swimming behavior in M9 buffer solution that becomes increasingly inhibited as the phenotypes increase in severity.

## 4. Application of Disease Models to This Study

In the course of this study, disease model strains of *C. elegans*, as well as the wildtype N2 strain, were fed different strains of bacteria to assess the effects of the bacterial diet (that is, microbiota) on behavior. Based

on the results of both behavioral assays and qualitative analysis of these variables, inferences were made about the potential effects of the interaction between *C. elegans* and its microbiota.

## II. Methods

### A. *C. elegans* Strains

*C. elegans* were raised on nematode growth media (NGM) plates seeded with a thin layer of bacterial growth for food. The growth of *C. elegans* is highly temperature dependent. For the purpose of this experiment, worms in the final larval stage (L4) were studied. Strains used in this study include N2 Bristol (wild type), AM725 (Amyotrophic Lateral Sclerosis model), NL5901 (Parkinson's Disease model), and LS587 (Duchenne Muscular Dystrophy model). Strains were provided by the Caenorhabditis Genetics Center (CGC), which is funded by the National Institute of Health (NIH) Office of Research Infrastructure Programs (21).

#### 1. Behavioral Patterns



Figure 1: N2 wild type *C. elegans* strain



Figure 2: LS587 *C. elegans* DMD model

N2 *C. elegans* move actively in consistent sinusoidal waves (Fig. 1) on agar plates; movement rates change when roaming/searching for food, dwelling, or recovering from eating.



**Figure 3: AM725 *C. elegans* ALS model**



**Figure 4: NL5901 *C. elegans* Parkinson's disease model**

## B. Bacterial Strains

*C. elegans* were raised on 6 cm NGM plates seeded with OP50. The bacterial strain OP50 is the most common food source for *C. elegans* raised in the lab, but there are a multitude of bacterial strains that these organisms are known to eat. The worms raised on OP50 plates were collected and seeded with various bacterial diets. These additional diets included both HT115 and HB101, which are *E. coli* strains. Although all worms were acquired from OP50 plates, the OP50 bacterial strain was still incorporated into additional experimentation. These strains of *E. coli* were used as control strains due to the fact that they are commonly used in the lab. As previously mentioned, they are safe for the worms to ingest. The relationship between the worms and their microbiomes show evident effects on functions such as development, aging, metabolism, and response to infection.

## C. Seeding Nematode Growth Media (NGM) Plates

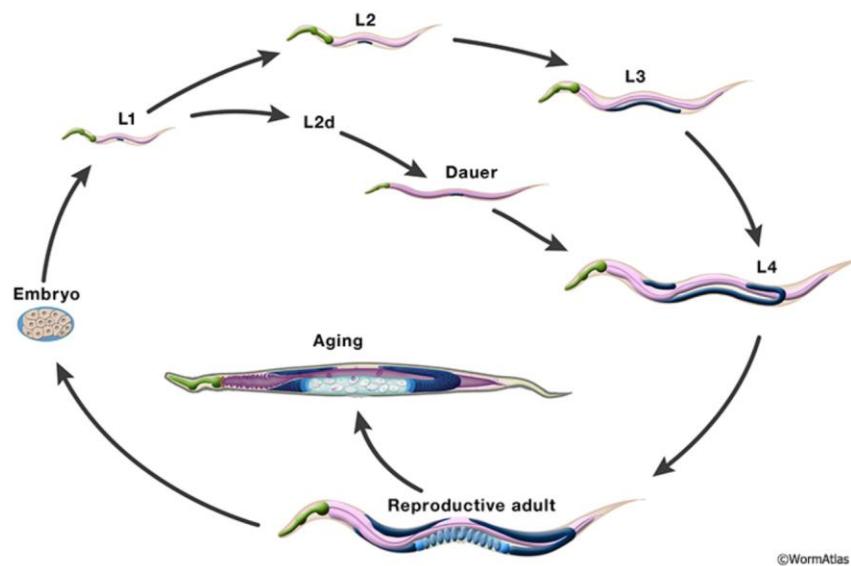
A saturated bacterial culture was measured out into an increment of 50  $\mu\text{L}$ . A nematode growth media (NGM) agar plate is prepared, and the bacterial solution was micropipetted into the middle of the plate. The bacterial lawn was created by spreading the bacterial fluid to the edges of the plate using a cell spreader, leaving a small rim of space for the future planted worms to be placed. The plates were incubated overnight at 20°C. After incubation, the plates were stored at 4°C until needed.

## D. Bacterial Plate Streaking

The effects of bacteria on *C. elegans* disease models were tested using externally sourced, known bacterial strains, as well as manually collected bacterial colonies. In order to obtain species of bacteria isolated from the environment, each team member used a sterile cotton swab to collect microorganisms from a location of interest within the Mellon Institute. Each swab was streaked onto an LB agar plate, then incubated overnight at 37° C. Single colonies were then inoculated into LB media in sterile tubes that were placed overnight in a shaking water bath to promote growth. Once the bacterial colonies had grown in the tubes, portions of the culture were exported for 16S rRNA sequencing to the company Genewiz (Azenta). The remaining cultures were used to seed NGM plates.

## E. Small-Scale Staging of Worms

Staging is the process of preparing a plate of age-synchronous worms. Synchronized larval stages are critical to *C. elegans* research because life cycle dictates behavior — it would be inaccurate to compare thrashing data taken from an adult worm with a larval stage 1 worm to suggest that a variable like bacteria is the reason for the variance in behavior. If data is collected from worms in different stages or of different sizes, this could introduce extraneous variables. For this reason, synchronizing cultures was a vital part of *C. elegans* maintenance and allowed us to obtain reproducible data. Figure 5 depicts the life cycle of *C. elegans*, a 3 day process characterized by multiple larval stages.



**Figure 5: *C. elegans* life cycle<sup>22</sup>**

Because the L4 stage is most commonly used in published behavior studies utilizing *C. elegans*, we chose to synchronize our cultures at this stage for consistency; worm behavior can differ by stage. Adult worms are sensitive to bleach, while embryos possess an egg shell that protects from cell death. If bleached long enough, the embryos will not survive; monitoring the duration of exposure ensures that *C. elegans* embryos successfully remain on the plate.

To carry out this staging, 60 µL worm bleach (per 5 mL: 1 mL NaClO, 1.25 mL 1M NaOH, and 2.75 mL dH<sub>2</sub>O) were placed on the edge of an NGM plate. At least ten gravid hermaphrodites were placed in the bleach one at a time and observed under the microscope until dissolved. More worm bleach was added if needed. The process was repeated for every plate. Eggs were incubated at 15°C or 20°C for between 16 and 48 hours depending on the *C. elegans* strain until eggs hatched and age-synchronized populations developed. Some plates were incubated for up to three days to obtain L4 stage worms (23).

## F. Large-Scale Staging of Worms

Similar to the small-scale staging, large-scale staging harvests gravid hermaphrodite worms from NGM plates seeded with OP50. M9 buffer solution (per 1 L: 3g KH<sub>2</sub>PO<sub>4</sub>, 6g Na<sub>2</sub>HPO<sub>4</sub>, 5g NaCl, 1mL MgSO<sub>4</sub>, H<sub>2</sub>O) was added to the OP50 plate(s) and a probe was used to gently disrupt the contents of the plate before collecting them in a plastic tube. This solution was then added to a conical tube and centrifuged in order to create a pellet of worms. Excess supernatant was removed using a micropipette. The pellet was

then exposed to an addition of bleach solution for five minutes (including the 1 minute centrifuge). At this point, only eggs in their protective coating remained. The eggs were processed through a series of washes to remove excess bleach and then transferred onto new plates and incubated at either 15 °C or 20 °C for at least 24 hrs and up to 72 hrs to obtain synchronized larval populations.



**Figure 6: *C. elegans* embryos following worm bleaching**

## G. Worm Picking

Worm picking was used to safely transfer individual worms onto M9 buffer which acts as a medium for worms to swim in for the thrashing assay. *C. elegans* were transferred using a worm pick that was made out of glass wrapped around a 0.2 mm diameter platinum iridium wire (Figure 7). The worm picks are malleable and heat and cool very quickly. An ethanol burner was used to sterilize the pick before moving the worms. When the edge of the instrument turned bright orange, the pick was removed from heat as any potential contaminant on it had been sterilized. The worms were picked under the microscope. They were scooped from the surface of the plate using the flat part of the pick. The worms were then transferred to the M9 buffer that had been placed on a microscope slide for thrashing analysis. The worm picks were then re-sterilized.

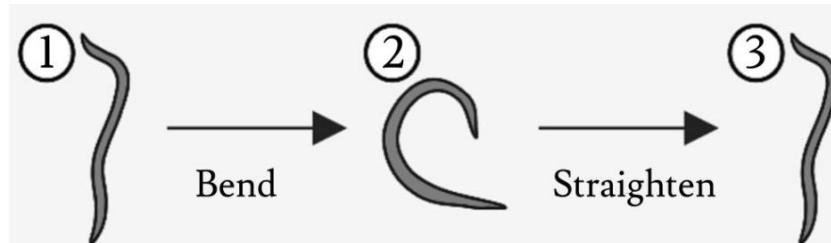


**Figure 7: Platinum worm pick**

## H. Quantifying *C. elegans*' Movement via Thrashing Assay

Thrashing assays are designed to quantify the movement and behavior of *Caenorhabditis elegans*. *C. elegans* navigate through agar plates via sinusoidal, wavelike movement, characterized by their normal behavior patterns (24). The worms display movements called thrashes when swimming in liquid, that are exaggerated compared to movement on solid media. Phenotypes that are not readily apparent on agar can be observed in liquid (in this study, M9 buffer). Quantifying locomotion as thrashes per minute is a commonly used way to assess locomotion in a way (25). This is a simple yet powerful way of identifying the mutant's affected behavior and observing neuronal function (26.). Thrashing rates can be used as an index for observing the effects of drugs, chemicals, and bacteria on *C. elegans* (25). Generally, *C. elegans* move by creating waves of dorsal-ventral and their movements can be classified as in a C-shape or in an S-shape (27). The *C. elegans* move forward by propagating from head to tail and backwards by propagating from tail to head, (24). A wildtype strain of *C. elegans* (N2) has an average thrashing rate of 100 thrashes per minute (28). Although behavioral assays are powerful tools, *C. elegans*' behavior is sensitive to external variables. Thrashing rate can easily be influenced by small changes such as variation in culture or assay condition (26).

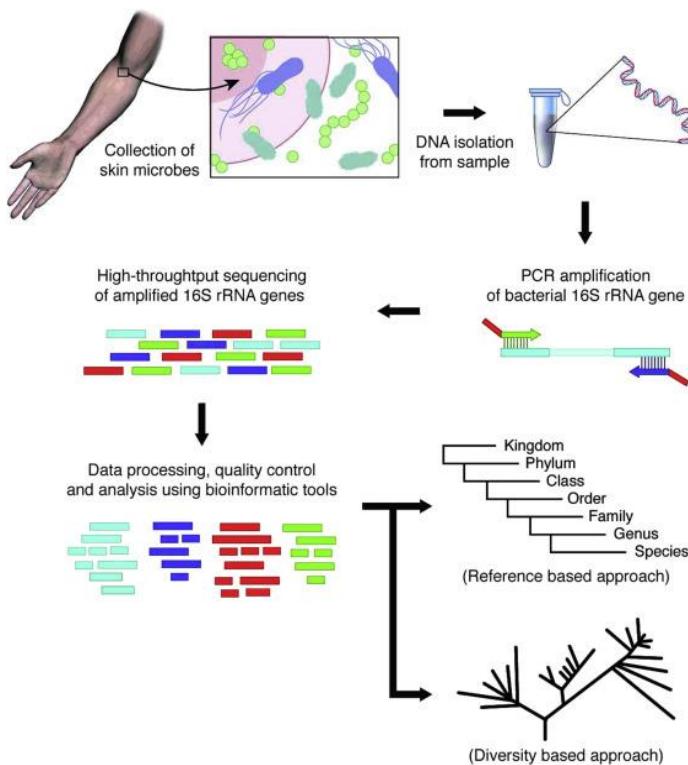
L4 worms were picked into 40  $\mu$ L of M9 buffer on a microscope slide. After a minute, a 30 second video was recorded via microscope imaging. The light on the microscope was deactivated and 5 minutes elapsed; after the 5 minutes, another 30 second video was recorded. These videos were then quantified by judging by *C. elegans*' bodily position with hand tally counters (Fig. 8). The manual approach is limited by frame rate speed, and is fully constrained to scoring one animal at a time. This may delay quantification, but was not a major concern given our sample size.



**Figure 8: Thrashing patterns in *C. elegans***

## I. 16S rRNA sequencing

Portions of the bacterial cultures were exported for 16S rRNA sequencing, a sequencing method widely used to identify bacterial species based on highly variable regions within the 1500bp 16S ribosomal RNA subunit gene, which is found in virtually all bacteria (29). Bacterial genomes are first amplified using PCR. Next, Sanger sequencing is used to identify the exact (or near exact) sequence of nucleotides within this gene. Sanger sequencing involves mixing the amplified bacterial DNA isolate with deoxynucleotides and fluorescently-labeled chain-terminating dideoxynucleotides, resulting in the formation of DNA fragments of varied length, each with a fluorescent dideoxynucleotide marker attached. These fragments are then separated using gel electrophoresis and scanned with an excitatory laser as they pass through the gel, which causes their dideoxynucleotide markers to fluoresce. A machine known as a basecaller detects these fluoresces and the order in which they occur and is able to translate them into a readable nucleotide sequence (30). This sequence can then be cross-referenced with known 16S sequences from existing databases, like the National Center for Biotechnology Information Basic Local Alignment Search Tool (NCBI BLAST), to identify the bacterial species from which the DNA was isolated.



**Figure 9: Process of 16S rRNA sequencing<sup>31</sup>**

16S rRNA sequencing enabled us to identify, or at least narrow down, the species of bacteria contained in a particular sample. The NIH BLAST tool compared our samples against its database of nucleotide sequences. An important consideration is the choice of dataset, as BLAST offers a wide array of programs; a database including representatively sampled sequences will return more accurate matches and contain fewer sequencing errors or vector contamination. We used the 16S ribosomal RNA sequences (bacteria and archaea) search set.

## J. Microscopy

All images and videos in this study were acquired using Motic Images Plus 2.0 attached to Olympus Stereo Zoom Dissecting Microscopes. The transparent nature of *C. elegans*' body is conducive to microscopic imaging. Each species was observed qualitatively and then recorded via Motic Images for behavioral assay.

## III. Results and Discussion

### A. Bacterial Sequencing

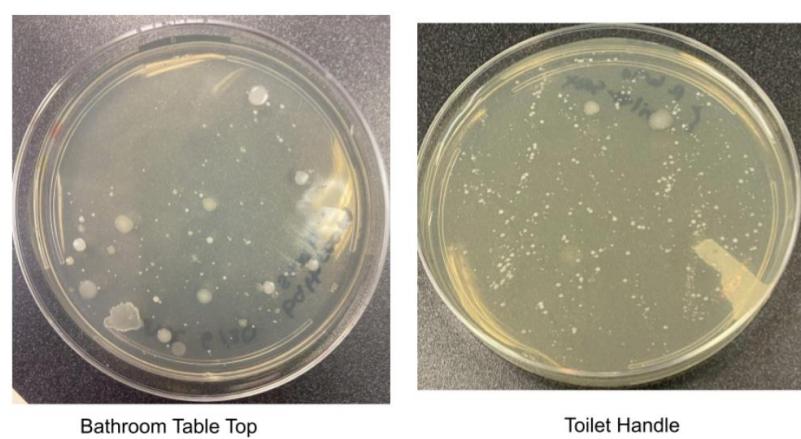
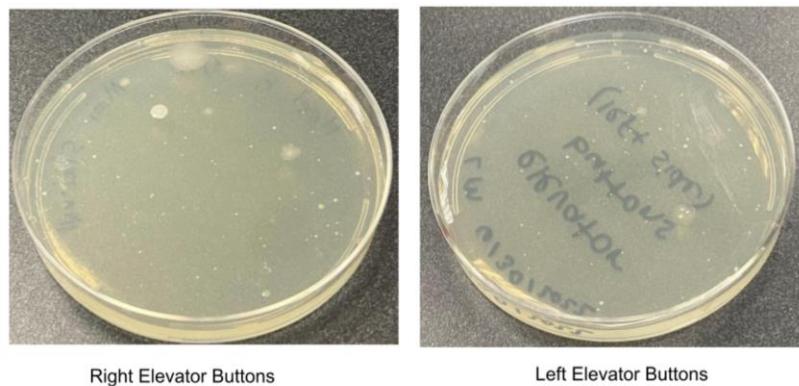
**Table 1: 16S rRNA sequencing results of locally-collected bacteria**

<i>Initials of Experimenter*</i>	<i>Location</i>	<i>Color</i>	<i>Result</i>	<i>Percent Match</i>
XL	Bathroom table top	Light yellow	<i>Bacillus toyonensis</i>	94.2%
<b>NP</b>	Stairs	White	Not enough sequence	N/A
KG	Sterile room table	Opaque milky, yellow-white	<i>Staphylococcus hominis</i>	99%
<b>LO</b>	Bathroom table top	Dark yellow	Not enough sequence	N/A
CD	Right elevator buttons	Opaque milky yellow	<i>Staphylococcus epidermidis</i>	98.06%
<b>CS</b>	Sterile room light switch	Opaque bright yellow	Too many repetitive nucleotides, could not narrow down the bacteria, nothing conclusive	N/A
<b>ER</b>	Toilet handle	White	Too many bacteria strains, nothing conclusive	N/A

\* Initials are not relevant to samples that were identified via 16S rRNA sequencing, as those will be referred to with their species name. Bold text indicates the initials will be relevant later.

After the bacteria were swabbed and cultivated on the agar plates (as shown in Figure 9), seven samples were sent for sequencing. Once bacteria sequencing results came back in the form of nucleotide strings, sequences were input into the Basic Local Alignment Search Tool (BLAST). BLAST is an online database through the NIH that allows us to match the unknown collected bacteria sequences to the known bacterial sequences in the database. We were able to use BLAST to identify three samples with high probability of accuracy. The light yellow bathroom table top sample had a likely positive match with *Bacillus toyonensis* (94.2% sequence match). The opaque milky, yellow-white bacteria from the sterile room table was identified as *Staphylococcus hominis* (99% sequence match). The opaque milky yellow sample collected from the right elevator buttons of the Mellon Institute matched with *Staphylococcus epidermidis* (98.06% sequence match).

We were unable to identify the remaining four bacterial samples. This could be due to a number of factors, including cross-contamination within the samples and the potential for not isolating enough of the initial sample before sending it for sequencing. The white bacterial sample from the Mellon Institute stairs and the dark yellow bacteria from the bathroom table top did not provide enough sequence for BLAST to identify them accurately. The opaque bright yellow bacteria from the sterile room light switch had many redundant nucleotides in its sequence, and BLAST could not narrow down the bacteria to a conclusive result. The same is true of the white bacteria from the toilet handle; there were too many bacterial options and we were unable to accurately identify the strain.

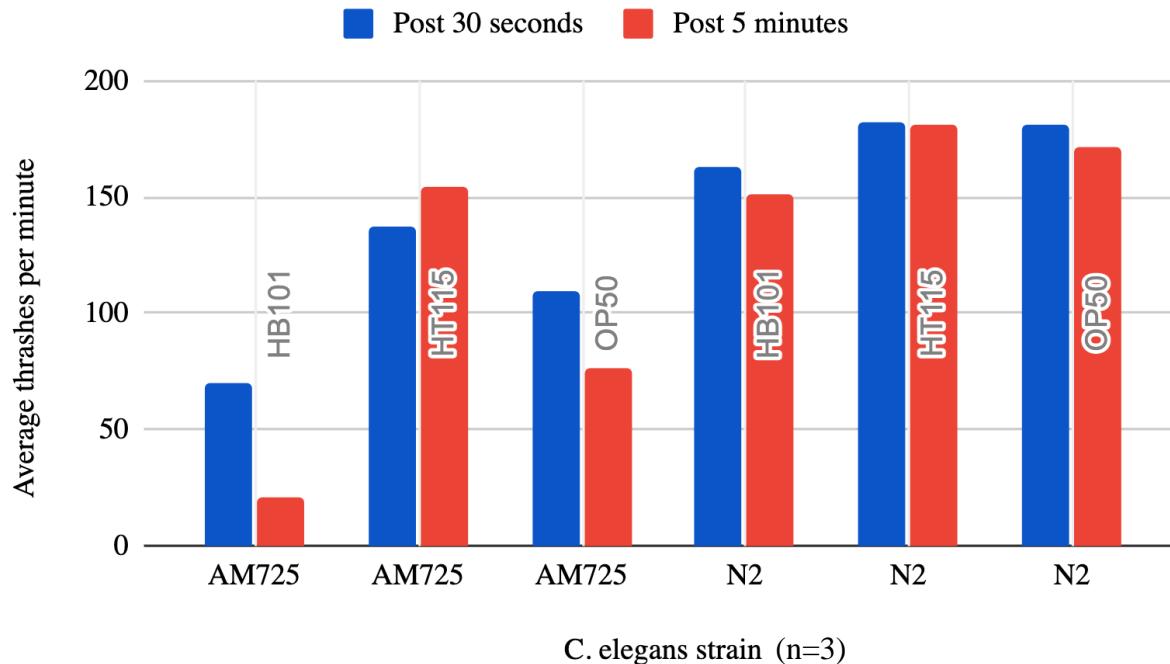


**Figure 10: Growth of local bacterial samples on LB plates**

### B. Subgroup 1: Comparing the Effects of Pre-Obtained & Locally-Sourced Bacteria on AM725 and N2 Strains of *C. elegans*

**1. Phase I:** We compared the effect of three different bacterial strains, HT115, HB101, and OP50, on ALS model AM725 worms' and control N2 Bristol worms' thrashing behavior.

## AM725 vs. N2 Thrashing on Varying Bacteria

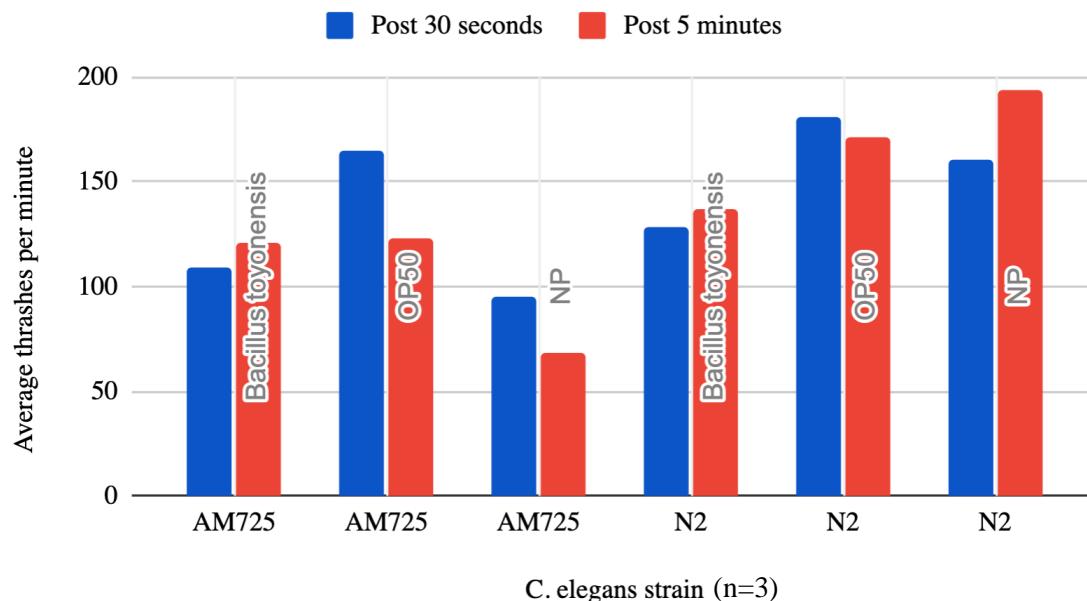


**Figure 11: Comparison of thrashing rates in different known bacterial strains of *C. elegans***

All three bacterial strains appeared to produce similar thrashing behavior in N2 Bristol worms, whereas the three bacterial strains appeared to produce dissimilar thrashing behavior in AM725 worms. AM725 worms fed HB101 appeared to thrash significantly less when compared to AM725 worms fed OP50, which in turn appeared to thrash significantly less when compared to AM725 worms fed HT115. Additionally, N2 Bristol thrashing activity appeared to decrease post five minutes regardless of the bacterial strain which the worms were fed. The AM725 worms displayed a similar decrease in thrashing behavior, except for the AM725 worms fed HT115, which displayed an increase in thrashing behavior after five minutes. AM725 worms also exhibited significantly less thrashing overall when compared to N2 Bristol worms fed the same bacteria.

**2. Phase II:** We compared the effect of three different bacterial strains: OP50, *Bacillus toyonensis* (isolated from a Mellon Institute bathroom table top), and an unidentified species (“NP”) from a Mellon Institute staircase, on ALS model AM725 worms’ and control N2 Bristol worms’ thrashing behavior.

## AM725 vs. N2 on Various Locally Sourced Bacteria



**Figure 12: Comparison of thrashing rates of AM725 (ALS) strain and N2 (wild type)**

The unidentified “NP” bacterial strain and OP50 appeared to produce similar thrashing behavior in N2 Bristol worms. *Bacillus toyonensis*, however, appeared to produce significantly less thrashing behavior in N2 Bristol worms. The three bacterial strains appeared to produce dissimilar thrashing behavior in AM725 worms. AM725 worms fed unidentified “NP” bacteria appeared to thrash significantly less when compared to AM725 worms fed *Bacillus toyonensis*, which in turn appeared to thrash significantly less when compared to AM725 worms fed OP50. Additionally, N2 Bristol worms fed OP50 exhibited a decrease in thrashing post five minutes. N2 Bristol worms fed unidentified “NP” bacteria and *Bacillus toyonensis* exhibited an increase in thrashing behavior after five minutes. The AM725 worms displayed similar decreases and increases in thrashing behavior, except for the AM725 worms fed unidentified “NP” bacteria, which displayed a decrease in thrashing behavior post five minutes instead of an increase.

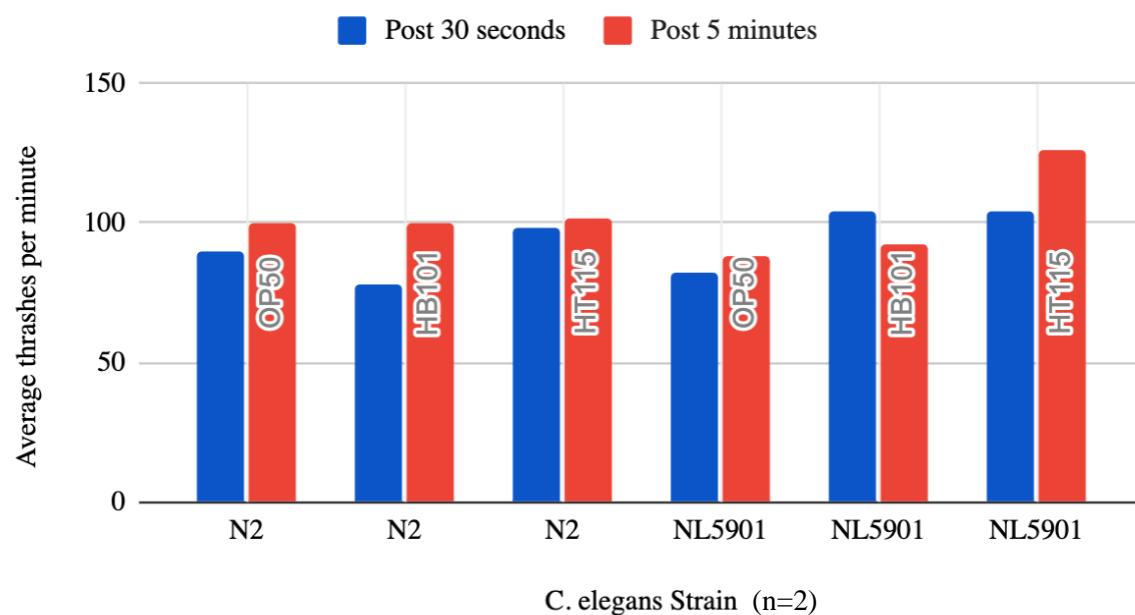
### C. Subgroup 2: Comparing the Effects of Pre-Obtained & Locally-Sourced Bacteria on NL5901 and N2 Strains of *C. elegans*

In Phase I, the behavior of the NL5901 Parkinson’s model as well as the N2 wild type strain was monitored in response to three known *E. coli* strain diets. The goal of this phase was to identify whether or not bacteria has a larger impact on thrashing rates of *C. elegans* disease models compared to the wild type.

**1. Phase I:** Wild type (N2) *C. elegans* worms were treated with three control *E. coli* bacterial diets: HB101, OP50, and HT115. The thrashing numbers were fairly similar amongst the three tested wild types. HT115 had slightly higher numbers than OP50 and HB101, which both ended with the same numbers after five minutes. HB101 had the lowest initial thrashing rate, but otherwise the results were mostly consistent for the first half of Phase I. The numbers followed similar data in other studies we found.

The second half of Phase I tested how the NL5901 (Parkinson's) strain of worms reacted to the same three control bacterial diets that were given to the wild type. For this portion of the experiment, we found higher initial numbers in the HB101 and HT115 in comparison to the OP50, but after five minutes the HB101 dropped and the OP50 increased slightly. There was a noticeable difference in the HT115 after the time interval, with a large increase in thrashing numbers after five minutes. The NL5901 strain showed greater data variation in comparison to the wild type, despite being fed the same bacterial diets. We found that the three control bacteria strains have a much higher effect in thrashing number on genetically modified worm strains, such as the NL5901. This is supported by data collected in the phase that indicates a higher thrashing rate difference in the disease model strains as compared to the relatively constant thrashing rates seen in the wild type.

## NL5901 vs. N2 Thrashing on Varying Bacteria



**Figure 13: Comparison of thrashing rates of Parkinson's model NL5901 to N2 wildtype**

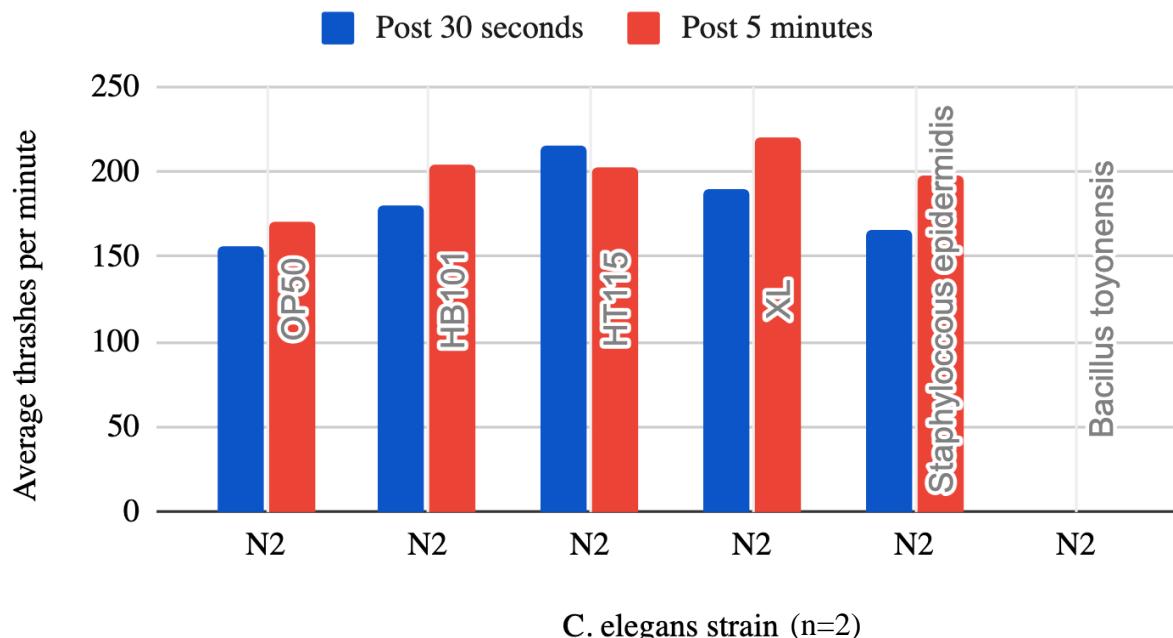
The three *E. coli* diets given to N2 wildtype worms resulted in relatively consistent numbers. This data allows for assumptions to be made about results for the Parkinson's NL5901 model. The NL5901 strain treated with HT115 resulted in the largest increase in thrashing rate overall. The data within Figure 12 suggests that the NL5901 *C. elegans* strains treated with control bacterial diets presented with a greater change in thrashing rate compared to the N2 wild type.

In Phase II, data from Phase I was used as a foundation to expand research of not only the impact of the three known *E. coli* strains, but also three locally sourced bacterial strains. Phase II consisted of studying the effects of the OP50, HB101, and HT115 controls alongside *Staphylococcus epidermidis*, *Bacillus toyonensis*, and an unidentifiable bacteria referred to as "XL", on the given *C. elegans* strains.

**2. Phase II:** To test how different bacteria would affect the thrashing rates of the N2 wildtype, the worm population was given six different bacterial diets, including three *E. coli* bacteria strains and three bacteria strains we swabbed: HB101, OP50, HT115, *Bacillus toyonensis*, *Staphylococcus epidermidis*, and one

unknown strain (XL). Unfortunately, the *Bacillus toyonensis* plate had mold contamination and, as a result, no data could be retrieved.

## N2 on Various Locally Sourced Bacteria



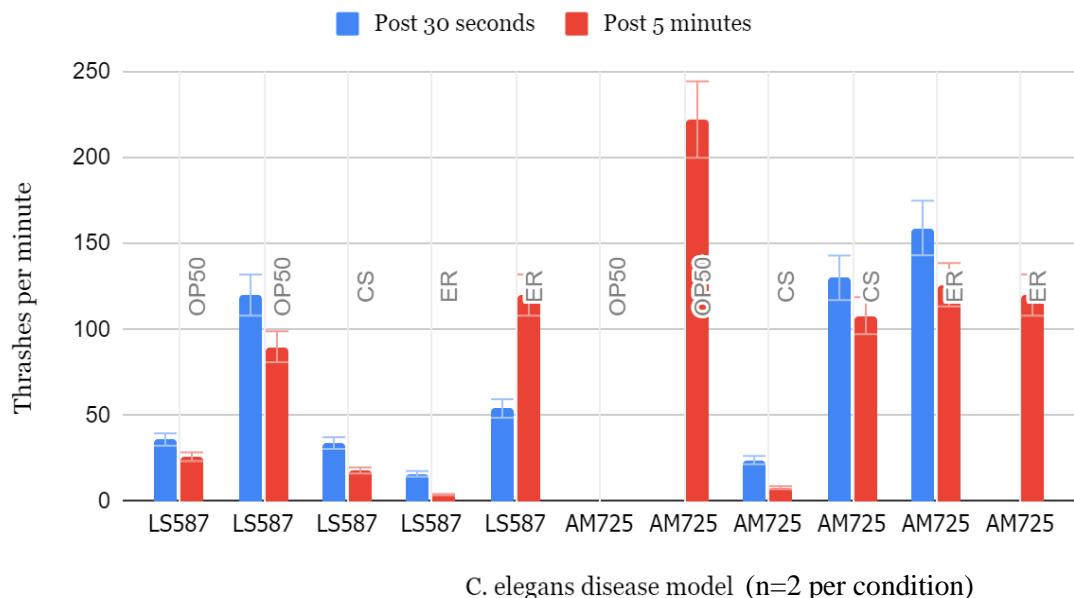
**Figure 14: Thrashing rates of N2 wildtype for control *E. coli* and locally sourced bacteria**

A general trend may be derived from Figure 12; a consistent increase in thrashing rates after 5 minutes. OP50, HB101, and the staph sample all had an increase in thrashing rates post 5 minutes. The XL sample had the highest thrashing rate post 5 minutes out of all the other samples. The only exception to the trend was the N2 grown on the HT115 bacterial strain, which saw a decrease in thrashing rate after 5 minutes. The initial thrashing rate of the *C. elegans* grown on HT115 was the highest of all the *C. elegans* with different diets. We concluded that different bacterial diets play a significant role in thrashing rates of the *C. elegans*.

## D. Subgroup 3: Comparing the Effects of Pre-Obtained & Locally-Sourced Bacteria on AM725 and LS587 Strains of *C. elegans*

**1. Phase I:** Subgroup 3 compared the effect of different known and unidentified bacterial strains on the ALS model (AM725) and DMD model (LS587) via a thrashing assay and qualitative analysis. The OP50 bacteria was used as the control group, while the experimental group included two locally-sourced but unidentified bacteria labeled "CS" and "ER", which show a large variation in the thrashing behaviors for each worm ( $n=2$ ) from the thrashing assay at both the initial and post-5 minute time intervals. While the data collected in Phase I does not support any clear behavioral relationship between the different disease strains and the bacteria they were tested on, a notable difference was found between the thrash rate of the LS587 and AM725 worms. An average number of thrashes per minute of all of the AM725 and LS587 worms across all the bacterial strains used was calculated for both the initial and post-five minute time intervals. It was found that the AM725 strain had an experimentally significant increase (approx. 50%) in average number of thrashes per minute at either time interval when compared to the LS587 worms from the same phase.

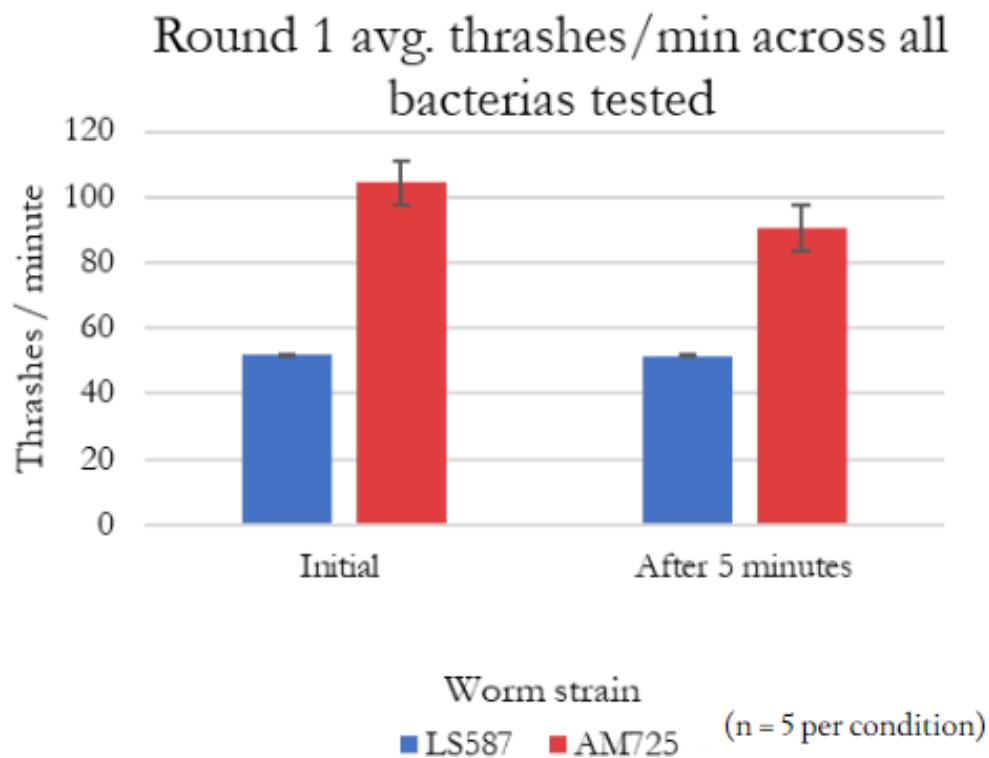
### LS587 v. AM725 thrashing assays on varying bacteria



**Figure 15: Comparison of thrashing rates of individual LS587 (DMD) strains to that of AM725 (ALS) strains**

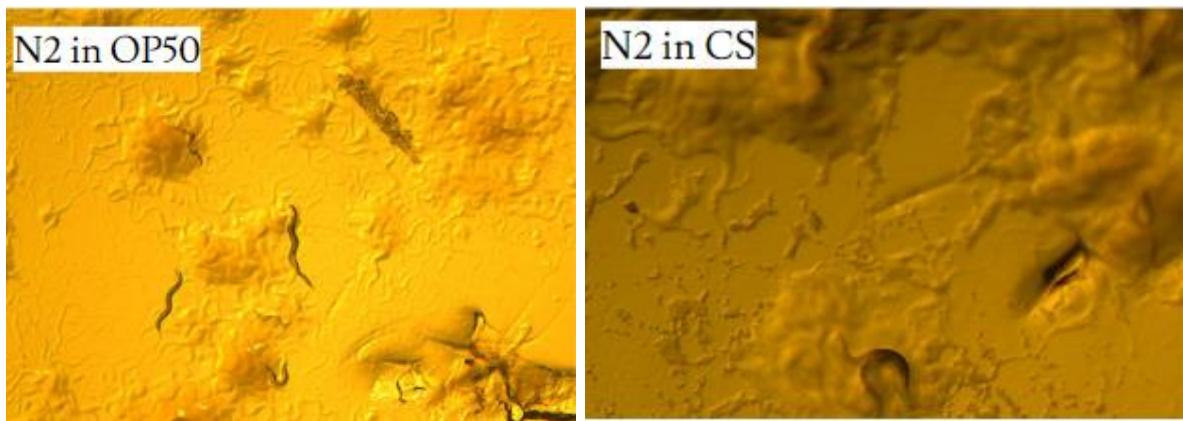
As shown in Figure 15, the thrashing assays of phase 1 resulted in a wide variety of detected behaviors. Contrary to previously displayed graphs, which showed the average thrashing rate of each *C. elegans* phenotype per bacterial strain, this graph shows each individual *C. elegans* sample to highlight the variety in the data. Another abnormality in the displayed graph is the lack of data points for several AM725 samples, which is due to a malfunction in the recording software in which the frame rate was not high enough to record quantifiable data. Though there is a large variety in thrashing behaviors, one observable trend is the consistently higher thrashing rate of the AM725 phenotype compared to that of the LS587. When the thrashing rates of all AM725 samples are averaged and compared to that of the LS587 (represented in Figure 15) there is nearly a 100% increase from the LS587 thrashing rate to the thrashing rate of the AM725.

**2. Phase II:** Based on the results collected in Phase I, we wanted to take a closer look at the effects of the unidentified “CS” strain on the ALS (AM725) and DMD (LS587) disease models as well as the N2 wildtype strain. Bacterial strain OP50 was used as a control for comparison. The results gathered from this phase support the conclusions drawn by subgroups 1 and 2, with regards to the ALS strain. The key takeaways from this analysis were the toxicity and overall incompatibility of the “CS” bacterial strain with *C. elegans*. Plates seeded with OP50 grew small uniform colonies, and the bacterial lawn did not obscure *C. elegans*. Compared to the OP50 control strain, “CS” results in visibly thicker, bright yellow colonies (see Fig 16) that nearly conceal the worms when viewed under the microscope for visual analysis. When worms grown on “CS” were placed into the thrashing assay, they died quickly (often <1 min.) after being placed in the buffer solution. They experienced a failure to thrash and additionally appeared to be weaker and smaller than other worms of the same strain grown on the control bacteria.



**Figure 16: Average number of thrashes per minute of all AM725 and LS587 worms across all bacterias tested in this case at both the initial and post-5 minute time intervals**

It is worth noting that in both phases of this subgroup, the phenotypic expressions of the disease models AM725 and LS587 were present. The AM725 strain displayed a slower and more coiled thrashing behavior than the N2 wildtype strain. Comparatively, LS587 exhibited further impacted behaviors. While the worm itself was visibly more “kinky” and frail looking than either the ALS or especially the wildtype strain, it also exhibited an inhibited gait with thrashing patterns originating inconsistently from the abdomen of the worm as compared to the top pharynx region of the specimen seen in the wildtype strains. These gait abnormalities are likely a contributing factor to the significant decrease in thrashes per minute seen in the DMD model in phase I compared to the ALS model.



**Figure 17: N2 (wild type) *C. elegans* with OP50 bacteria (left) and N2 (wild type) *C. elegans* with “CS” bacteria (right)**

As described above, the CS bacteria created a different bacterial lawn for *C. elegans* to navigate through; these characteristics are directly visible on the agar (Fig. 17). Although we were unable to identify the bacteria, its drastic effect on *C. elegans* underscores the importance of bacterial diet in health. As gut bacteria reflect the composition of the microbiota, we observe the relationship between microorganisms and animal behavior. The CS bacteria virtually eliminated thrashing in strains of *C. elegans* and physically decreased the sizes of the worms.

## IV. Conclusion

### A. General Trends

*C. elegans* is a useful research tool for modeling complex human diseases difficult to study *in vitro* or in unicellular models. Additionally, their well understood genome and intestinal tract makes them ideal candidates for studying the microbiota and effects of the relationship between a host organism and its microbiota. In the course of our study, we found that, as a general trend, the bacterial diet fed to strains of *C. elegans* does appear to play a role in movement patterns of the worm. While the extent of this role varied from bacterial strain to bacterial strain or among the different strains of *C. elegans* we tested, the majority of the data collected in the subgroups of this study supports the hypothesis that different bacterial diets affect the overall health of *C. elegans*. However, due to the variety of the variables tested within the subgroups, we were not able to determine any other trends across all the subgroups. The trends and extrapolations from the results of the sub-group studies discussed earlier are therefore applicable only within the context of each subgroup due to the variance in disease models and bacterial strains used.

### B. Future Applications

In the future, further research could be conducted to ensure more consistent analysis across the variable groups tested. This could be done through repetition of the same assays completed in this study or through the use of larger sample sizes to eliminate the discrepancies in data found in our experiment. Additionally, future experimentation could elucidate the molecular mechanisms behind the relationship between host and microbiota. The results of this study illuminated the possible “what” of this relationship, but the question of “why” is just as valid. It would be worthwhile to conduct research into the molecular interactions between

the various bacterial diets and their digestion in *C. elegans*. Genetic similarity between *C. elegans* and humans permits comparison of molecular mechanisms, and will inform future therapeutic & biotechnology developments. Through the lens of *C. elegans*, we can search for a cure.

## V. Acknowledgements

We extend our utmost appreciation to Carnegie Mellon University, the Department of Biology, the PGSS Alumni Campaign, Dr. Barry Luokkala, Ms. Melissa Lessure and Dr. Carrie Doonan for providing the resources that enabled this research, and to Dr. Emily Drill and Cameron Blackwood-Short for their endless support. We would not be here without them.

## VI. References

1. Marchesi, J. R., & Ravel, J. (2015, July 30). The vocabulary of microbiome research: A proposal - microbiome. SpringerLink. Retrieved July 25, 2022, from <https://link.springer.com/article/10.1186/s40168-015-0094-5>
2. Malard, F., Dore, J., Gaugler, B., & Mohty, M. (2020, December 9). Introduction to host microbiome symbiosis in health and disease. *Nature News*. Retrieved July 25, 2022, from <https://www.nature.com/articles/s41385-020-00365-4>
3. Kayama, H., Okumura, R., & Takeda, K. (2020, April). Interaction between the microbiota, epithelia, and immune cells in the intestine. Retrieved July 25, 2022, from <https://www.annualreviews.org/doi/abs/10.1146/annurev-immunol-070119-115104>
4. Schwabe, R. F., & Jobin, C. (2013, October 17). The microbiome and cancer. *Nature News*. Retrieved July 25, 2022, from <https://www.nature.com/articles/nrc3610>
5. Abbott, A. (2020, Nov. 4). Are infections seeding some cases of Alzheimer's disease? *Nature News*. Retrieved July 28, 2022, from <https://www.frontiersin.org/articles/10.3389/fcimb.2021.625913/full>
- 5b. Riddle D.L., Blumenthal T., Meyer B.J., et al. (1997, December 31). The biological model - *C. elegans* II - NCBI bookshelf. (n.d.). Retrieved July 25, 2022, from <https://www.ncbi.nlm.nih.gov/books/NBK20086/>
6. Lundquist, E.A. (n.d.) What is *C. elegans* and why do scientists use it to study human development and disease? Why study *C. elegans*? (n.d.). Retrieved July 25, 2022, from [http://www.people.ku.edu/~erikl/Lundquist\\_Lab/Why\\_study\\_C.\\_elegans.html](http://www.people.ku.edu/~erikl/Lundquist_Lab/Why_study_C._elegans.html)
7. Borisovich Shtonda, B., & Avery, L. (2006, January). Dietary choice behavior in *Caenorhabditis elegans*. *The Journal of Experimental Biology*. Retrieved July 25, 2022, from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1352325/>
8. Samuel, B. S., Rowedder, H., Braendle, C., Félix, M.-A., & Ruvkun, G. (2016, June 17). *Caenorhabditis elegans responses to bacteria from its natural habitats*. *PNAS*. Retrieved July 25, 2022, from <https://www.pnas.org/doi/pdf/10.1073/pnas.2202856119>

9. Kaletta, T., & Hengartner, M. O. (2006, April 21). *Finding function in novel targets: C. elegans as a model organism*. Nature News. Retrieved July 13, 2022, from <https://www.nature.com/articles/nrd2031>
10. Corsi, A. K. (2006, December 1). *A Biochemist's Guide to Caenorhabditis elegans*. Analytical Biochemistry. Retrieved July 13, 2022, from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1855192/>
11. Urban, N. D., Cavataio, J. P., Berry, Y., Vang, B., Maddali, A., Sukpraphrute, R. J., Schnell, S., & Truttmann, M. C. (2021, November 15). *Explaining inter-lab variance in C. elegans N2 lifespan: Making a case for standardized reporting to enhance reproducibility*. Experimental Gerontology. Retrieved July 13, 2022, from <https://www.sciencedirect.com/science/article/pii/S0531556521004046>
12. CGC-registered strains. Strain List - Caenorhabditis Genetics Center (CGC) - College of Biological Sciences. (n.d.). Retrieved July 13, 2022, from <https://cgc.umn.edu/strains?vt=all>
13. Beitz, J. M. (2014). *Parkinson's disease: A Review*. Frontiers in Bioscience (Scholar edition). Retrieved July 25, 2022, from <https://pubmed.ncbi.nlm.nih.gov/24389262/>
14. Matsos, S. (2022, July 21). *Neurology and neurosurgery*. Johns Hopkins Medicine. Retrieved July 25, 2022, from [https://www.hopkinsmedicine.org/neurology\\_neurosurgery/](https://www.hopkinsmedicine.org/neurology_neurosurgery/)
15. Cooper, J. F., & Van Raamsdonk , J. M. (2018). *Modeling Parkinson's disease in C. elegans*. Journal of Parkinson's disease. Retrieved July 25, 2022, from <https://pubmed.ncbi.nlm.nih.gov/29480229/>
16. U.S. Department of Health and Human Services. (n.d.). *Duchenne muscular dystrophy - about the disease*. Genetic and Rare Diseases Information Center. Retrieved July 26, 2022, from <https://rarediseases.info.nih.gov/diseases/6291/duchenne-muscular-dystrophy/>
17. Duchenne Muscular Dystrophy. (n.d.a) Johns Hopkins Medicine. Retrieved July 25, 2022, from <https://www.hopkinsmedicine.org/health/conditions-and-diseases/duchenne-muscular-dystrophy>
18. Ellwood, R. A., Piasecki, M., & Szewczyk, N. J. (2021). *Caenorhabditis elegans as a Model System for Duchenne Muscular Dystrophy*. International Journal of Molecular Sciences, **22**(9), 4891. Retrieved July 25, 2022, from <https://doi.org/10.3390/ijms22094891>.
19. Mayo Foundation for Medical Education and Research. (2022, February 22). *Amyotrophic lateral sclerosis (ALS)*. Mayo Clinic. Retrieved July 25, 2022, from <https://www.mayoclinic.org/diseases-conditions/amyotrophic-lateral-sclerosis/symptoms-causes/syc-20354022>
20. Ustyantseva, E. I., Medvedev, S. P., & Zakian, S. M. (2020). Studying ALS: Current Approaches, Effect on Potential Treatment Strategy. *Advances in experimental medicine and biology*, **1241**, 195–217. Retrieved July 25, 2022, [https://doi.org/10.1007/978-3-030-41283-8\\_11](https://doi.org/10.1007/978-3-030-41283-8_11)
21. Strain information. (n.d.). Caenorhabditis Genetics Center (CGC). Retrieved July 25, 2022, from <https://cgc.umn.edu/strain/OP50-GFP>
22. Altun, Z. F., & Hall, D. H. (2009). *Introduction to C. elegans Anatomy*. Handbook - introduction. Retrieved July 28, 2022, from <https://www.wormatlas.org/hermaphrodite/introduction/mainframe.htm>
23. Lemons, M.L. 2016. An Inquiry-Based Approach to Study the Synapse: Student-Driven Experiments Using C. elegans. *J Undergraduate Neuroscience Education* **15**(1): A44-A55.

24. Zhao, B., Khare, P., Feldman, L., & Dent, J. A. (2003). Reversal Frequency in *Caenorhabditis elegans* Represents an Integrated Response to the State of the Animal and Its Environment. *The Journal of Neuroscience*, **23**(12), 5319–5328. Retrieved July 25, 2022, from <https://doi.org/10.1523/jneurosci.23-12-05319.2003>
25. Buckingham, S. D., & Sattelle, D. B. (2009). Fast, automated measurement of nematode swimming (thrashing) without morphometry. *BMC Neuroscience*, **10**(1). Retrieved July 25, 2022, from <https://doi.org/10.1186/1471-2202-10-84>
26. Hart, A. (n.d.). *Behavior*. [Www.wormbook.org](http://www.wormbook.org/chapters/www_behavior/behavior.html). Retrieved July 25, 2022, from [http://www.wormbook.org/chapters/www\\_behavior/behavior.html](http://www.wormbook.org/chapters/www_behavior/behavior.html)
27. Pierce-Shimomura, J. T., Chen, B. L., Mun, J. J., Ho, R., Sarkis, R., & McIntire, S. L. (2008). Genetic analysis of crawling and swimming locomotory patterns in *C. elegans*. *Proceedings of the National Academy of Sciences of the United States of America*, **105**(52), 20982–20987. Retrieved July 25, 2022, from <https://doi.org/10.1073/pnas.0810359105>
28. Zhang, H., Gao, S., & Chen, W. (2022). Automated recognition and analysis of head thrashes behavior in *C. elegans*. *BMC Bioinformatics*, **23**(1). Retrieved July 25, 2022, from <https://doi.org/10.1186/s12859-022-04622-0>
29. Janda, M. J., & Abbott, S. L. (2006, September 1). *16S rRNA Gene Sequencing for Bacterial Identification in the Diagnostic Laboratory: Pluses, Perils, and Pitfalls*. American Society for Microbiology. Retrieved July 26, 2022, from <https://journals.asm.org/doi/10.1128/JCM.01228-07?cookieSet=1>
30. Valencia, C.A., Pervaiz, M.A., Husami, A., Qian, Y., Zhang, K. (2013). *Sanger Sequencing Principles, History, and Landmarks*. In: Next Generation Sequencing Technologies in Medical Genetics. SpringerBriefs in Genetics. Springer, New York, NY. [https://doi.org/10.1007/978-1-4614-9032-6\\_1](https://doi.org/10.1007/978-1-4614-9032-6_1)
31. Jo, J.-H., Kennedy, E. A., & Kong, H. H. (2016, February 19). *Research techniques made simple: Bacterial 16S ribosomal RNA gene sequencing in cutaneous research*. Journal of Investigative Dermatology. Retrieved July 28, 2022, from <https://www.sciencedirect.com/science/article/pii/S0022202X16003481>



**BIOLOGY/  
COMPUTATIONAL  
BIOLOGY  
TEAM PROJECTS**



# **Who's That Bacteria?: A Pipeline for Computational Analysis of DNA Fingerprints**

Abigail Barlow, Drew Mauriello, Emma Strickland, Toby Zapotocky

## **Abstract**

Probiotics have been used by humans for millennia in order to enrich our diet and improve gastrointestinal health. The relationship between the bacteria in probiotics and our microbiota is an incredibly important and precise documentation of that relationship for scientific and personal health purposes is becoming more needed, but present methods are mostly inefficient. The purpose of this project is to develop a robust, inexpensive, and adaptable pipeline of simple biological and computational techniques that will permit the identification of bacteria present within probiotic products. The biological techniques include, but are not limited to, isolation of genomic DNA from bacteria, ultraviolet (UV) scans of genomic DNA, agarose gel electrophoresis, polymerase chain reaction (PCR), PCR purification, DNA sequencing, and restriction enzyme digests. The final results showed that for each probiotic product, there appears to be a common bacterium that was present in all samples extracted from that product. Potential future directions for this project include refinement of this pipeline by increasing the efficiency of individual steps/techniques or expanding this methodology to be used to identify different types of bacteria or to identify bacteria for other purposes.

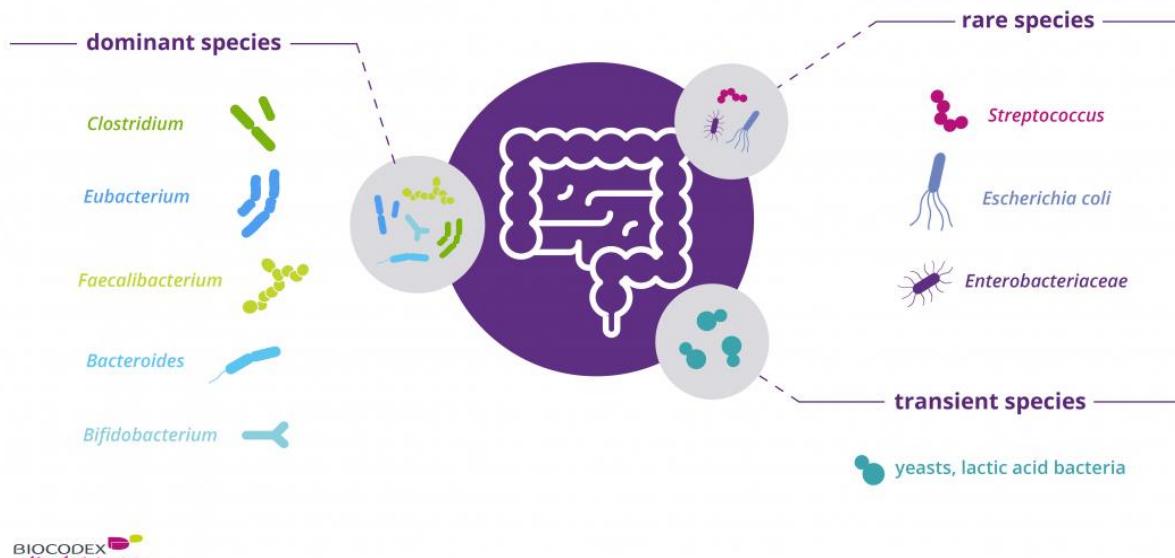
## **I. Introduction**

### **A. The Human Microbiota**

The human gastrointestinal (GI) tract is one of the largest organs in the human body that interacts with outside factors along with the human microbiota. When food passes through your GI tract, microorganisms play a role in breaking down and absorbing nutrients. These microorganisms can include funguses, viruses, parasites, and bacteria. The community of the bacteria living in the GI tract is termed the gut microbiota (Figure 1). Some species of bacteria will thrive in the GI tract due to being more suitable for their pH and temperature, and others will not because the conditions are not right for their survival<sup>1</sup>.

## COMPOSITION OF THE GUT MICROBIOTA

over 100,000 billion microorganisms live in our intestines!



**Figure 1: Many different species of bacteria are found in the human gut. There are dominant species (left) and rare species (right)<sup>2</sup>**

Everyone's microbiota is different from one another and not a lot is known as to why this is. Some researchers speculate that the differences result from environmental factors such as diet and antibiotic use. A recent study found that one's birth delivery method (vaginal vs cesarean section) can affect their microbiota. Those that are born vaginally have 72% resemblance to their mother's microbiota, while someone delivered by C-section only resembles 41% of their mother's microbiota<sup>1</sup>. At earlier stages of development, the diversity of your microbiota is very low and is mostly made up of two phyla: Actinobacteria and Proteobacteria<sup>3</sup>. In the first year of life, your microbiota grows more diverse, differentiating from other's microbiota due to environmental factors<sup>4</sup>. Around 2.5 years of age your microbiota will be fully matured and will resemble an adult-like microbiota. Your microbiota remains stable from this young age through adulthood. As you become elderly, the diversity of the microbiota declines.

To protect itself and maintain homeostasis, the GI tract limits the interactions between its microbiota and the immune system. They can do this because they have an intestinal barrier that is made up of physical, biochemical, and immunological factors (e.g., mucus layers, enzymes, antimicrobial proteins, IgA antibodies, and epithelia-associated immune cells)<sup>5</sup>. The bacteria that live in your GI tract will only continue to inhabit it if they contribute to your digestive process. If a certain species of bacteria is not contributing, it might be purged by species that contribute more.

The microbiota's role in the GI tract does not stop at digesting food; it also helps with resistance to certain infections, in the development of the immune system, and affects neurodevelopment<sup>3</sup>. One of the most important roles that the microbiota plays is assisting in maintaining the integrity of the mucosal barrier<sup>6</sup>. This helps provide nutrients to the host and to protect against pathogens. The interaction between the

microbiota and the mucosal barrier is also crucial for the immune system. Some bacteria express an enzyme that allows them to ferment complex carbohydrates and generate metabolites. These metabolites are absorbed by epithelial cells and regulate cell processes such as gene expression, differentiation, and proliferation<sup>9</sup>.

## B. Probiotics and Antibiotics

The word “probiotic” is derived from the Latin *pro* and the Greek *βιος* and describes a substance in which bacteria are present that is vital for the development of gut microbiota and can survive the digestive process to colonize the gut<sup>7</sup>. Thus, probiotics can become part of a host’s microbiota. Since ancient times, probiotics have been used by humans. Starting around 10,000 BCE, coinciding with a shift from nomadic familial groups to agricultural communities, some of the first fermented foods were produced. Fermented foods became a staple of these communities. Early writings include descriptions of these foods which were said to improve health and cure sickness. For example, ancient Hindu scriptures include details of a Hindu community consuming probiotic products like yogurt and sustaining themselves for long periods<sup>8</sup>. Although evidence was beginning to emerge of the usefulness of probiotics, their true potential would not be discovered until the 19th century.

In the modern day, probiotic food products are found all over the world in a variety of forms. Some of the most popular of these products include milk and its derivatives (butter, cheese, etc.) as well as sourdough bread, miso soup, kimchi, kombucha, and fermented foods like pickles and sauerkraut. These products are the third most popular dietary supplement in the world, just behind vitamins and minerals<sup>9</sup>.

In contrast to probiotics, antibiotics are compounds that kill or inhibit the growth of bacteria<sup>10</sup>. While antibiotics have saved millions of lives and played an important part in global health through the treatment of various diseases, these benefits do not come without drawbacks.

Due to their mechanism of action, antibiotics can have negative impacts on the microbiome of an individual. When an antibiotic is introduced into the body, it has no way of differentiating between pathogenic bacteria and beneficial bacteria leading to side effects such as nausea and diarrhea<sup>10</sup>. In more serious cases, antibiotic use can ultimately lead to what is known as dysbiosis, or the loss of healthy bacteria as a result of pathogenic bacteria overgrowth<sup>11</sup>. For many patients in critical care, where antibiotics are frequently administered, dysbiosis can lead to hospital-acquired infection, sepsis, and organ failure<sup>11</sup>. In an effort to combat the negative effects of antibiotics, some researchers are investigating the use of probiotic therapy in addition to antibiotic treatment. For instance, in cases of antibiotic-associated diarrhea, several studies show that approximately 40% of patients notice a decrease in symptoms after using probiotics<sup>11</sup>.

In addition to treating mild symptoms and side effects of antibiotics, probiotics can also play a role in solving a much larger health concern: antibiotic resistance<sup>1</sup>. Antibiotic resistance occurs whenever pathogenic bacteria adapt to survive in conditions where antibiotics are present<sup>12</sup>. In critical care patients, dysbiosis due to antibiotic use results in a microbiota that contains large amounts of pathogenic, multi-drug resistant bacteria<sup>11</sup>. Specifically, drug-resistant infection by the bacteria *C. difficile* has become increasingly common<sup>1</sup>. Patients with *C. difficile* experience severe, potentially fatal, gastrointestinal symptoms that do not respond to antibiotic treatment<sup>11</sup>. Promising research shows that probiotics can cure *C. difficile* infection in 64% of patients who are also taking antibiotics<sup>11</sup>. Similarly, fecal transplants, which are used to replenish the microbiota, can cure *C. difficile* colitis infection in up to 90% of cases<sup>11</sup>. These findings demonstrate the importance of probiotics and their effect on the microbiota.

Probiotics are also beneficial to gut health and foster a hearty microbiota. They are claimed to alleviate acute and chronic conditions like diarrhea, constipation, Inflammatory Bowel Syndrome, and irritable bowel syndrome. Having a fully functional and healthy microbiota has been shown to enhance resistance to certain transmissible and foodborne infections, contribute to a well-balanced immune system, as well as facilitate the digestion of food and the absorption of nutrients by the gastrointestinal tract<sup>13</sup>.

### C. Global Context

In many parts of the world, particularly in rural communities, people ferment their own food quite often. These communities rely on dishes such as doubanjiang, an essential southern Chinese ingredient that is composed of fermented beans, or Lutefisk, a common Scandinavian dish composed of codfish that has been dried and rehydrated using lye, to sustain themselves. Other more ubiquitous products, like yogurts, are essential to diets in a multitude of communities from around the world<sup>14</sup>. In most cases, these products are homemade and are distributed to customers or community members without labels. These products have the potential to house new and potentially revolutionary strains of bacteria within them that are not well-studied. Establishing a method that would allow us to test these products and identify the bacteria they contain without the need for expensive/heavy equipment or long wait times would have immense potential. This experimental method would enable rapid and accurate identification of potentially undiscovered strains of bacteria within these fermented food products, any number of which could house a multitude of potential health benefits. As evidenced above, bacteria in probiotic food products have been proven to have the potential to improve gastrointestinal health and strengthen our internal environment. Thus, the identification of new bacteria could provide any number of new health benefits to our microbiota.

### D. Finding A New Method of Analysis

Although it is true that processes for identifying bacteria in food products already exist, they are not easily accessible. They require immobile and expensive lab equipment in order to be performed, to which the vast majority of biology laboratories do not have access. Currently, smaller labs must send their samples to larger labs nearby so they can have their samples sequenced for a price. The process is ultimately time-consuming and not applicable to situations in which quick identification is needed.

Methods like DNA sequencing are certainly invaluable but often prove to be impractical. DNA sequencing, for the vast majority of biology labs, requires one's samples to be sent off to a larger laboratory or company for the sequencing to be performed there, due to the smaller labs' lack of proper equipment. Even sending the samples can be risky—unreliable mail services can damage the samples or otherwise render them useless. Actually, sequencing the DNA is the most expensive part—one has to pay \$7 per run per sample. This, in addition to transportation costs, can add up very quickly considering the fact that even in this project, we sent two runs for 21 samples at a total cost of \$294.

Other methods like TRFLP (Terminal Restriction Fragment Length Polymorphism) use a process that consists of amplifying a certain region of bacteria through PCR and tagging the ends of the primers with fluorescent molecules. These PCR products are then digested by a restriction enzyme and the ends of the new segments are located and the distance between them is measured using a very expensive capillary electrophoresis machine that runs them through thin polymer capillaries<sup>15</sup>.

Our objective is to assemble a sequence of simple, affordable, and common techniques that nearly any biology laboratory in any environment can perform to produce a reliable result. These basic molecular biology techniques include genomic DNA isolation, polymerase chain reaction (PCR), restriction enzyme digests, and agarose gel electrophoresis. The final product that we generate is a “fingerprint” of the 16S gene of the microbe of interest. This 16S subunit is a highly conserved region, meaning that it exhibits many of the same characteristics and is structurally similar between different bacteria. Although this is true, the 16S region contains a few species-specific sequences within hypervariable regions that allow it to be identified along with its bacteria and compared to other slightly different 16S regions and their bacteria. Also, mutations are not commonly found in the 16S region, indicating that there is very little opportunity for the identification process to go wrong<sup>16</sup>.

To analyze the 16S fingerprint data that we collect, the Computation Biology Team Project developed a new computational analysis method<sup>17</sup>. Our combined efforts have resulted in a pipeline that can accurately identify the genus if not the species of lactic acid bacteria using standard molecular biology techniques and basic computer hardware.

## II. Methods

### A. Isolation of Microbes from Dairy Products

The isolation of microbes was completed prior to the start of the project. However, so that readers may gain a more comprehensive understanding of the complete process of our project, we have included the steps that took place prior to the beginning of our time in the lab.

Samples were collected from a variety of dairy products, including blueberry-flavored Lifeway brand kefir and Schneider’s buttermilk with live cultures. These samples were collected using aseptic technique. Once these samples were collected, they were diluted because the number of living bacteria cells within the food product is too high to immediately be able to visualize properly on an agar plate. Without this step, the colonies would have been too large and too close together to be able to properly quantify the number of them that were present on the plate.

#### 1. Serial Dilutions

Serial dilutions were used to generate single well-isolated colonies from the dairy samples on agar plates. First, the mass of 1 mL of dairy product was determined. Then, that 1mL sample was mixed with 9 mL of sterile peptone water by vortexing, creating a  $10^{-1}$  dilution of our original sample. This step was repeated with the  $10^{-1}$  dilution and each successive, less-concentrated dilution to create a  $10^{-6}$  dilution of the original sample. The portions of each dilution that were not used to create the next dilution were kept for later use.

#### 2. Spread Plating

At this point, the cells were at a concentration that would be easily visible and distinguishable on an agar plate (a few hundred cells per 0.1mL of solution). To ensure that the cells were evenly distributed throughout the solution in which they were present, the cells were resuspended by vortexing. 0.1 mL each of the  $10^{-4}$ ,  $10^{-5}$ , and  $10^{-6}$  dilutions were then spread on MRS + Chx plates following standard

microbiology techniques. In short, spreaders were sterilized using ethanol and flame. Then the 0.1 mL of diluted sample was spread around the full surface of the agar plate and allowed to dry for 2 minutes before being incubated agar-side-up at 37C for 48 - 72 hours until visible colonies formed.

### 3. Streak Plating

Once colonies had formed on the dilution plates, streak plating was used to purify the colonies and ensure each was a clonal population of a single microbe. A single, well-isolated colony was picked using a sterile inoculating loop. The colony was then streaked on MRS+Chx media for single colonies. These plates were incubated at 37C for 48-72 hours until visible colonies formed.

## B. Genomic DNA Extraction

All Genomic DNA protocol follows the typical steps for Gram-positive bacteria, which was the only type of bacteria involved in this experiment. We carried out the gDNA using the QiaAMP DNA Mini Kit according to the manufacturer's instructions. In short, we pelleted the isolated bacteria, suspended the pellet in an enzyme solution, incubated, added an enzyme, incubated again, and added the product to a QiaAMP Mini Spin Column. Then, we repeated the sequence of adding a solution and then washing it through the column by centrifuging a few times. After these steps, we were left with isolated genomic DNA.

We isolated DNA in a total of 16 candidates (B1 - B8 and K1 - K8). A volume of 10mL of each candidate was cultured overnight at 37°C. During the genomic DNA extraction process, we spun down 3mL of each sample.

First, we centrifuged the media and the bacteria within it for a total duration of 10 minutes at 7500 rpm in order to remove all bacteria from the liquid medium and concentrate it in a small volume. This pellet was resuspended in a total volume of 180µL of 20 mg/mL lysozyme in 20 mM Tris-HCl, pH 8.0; 2mM EDTA; 1.2% Triton. We then incubated them for 30 minutes rotating at 37°C.

After incubation, we added 20µL of proteinase K, which serves to digest proteins, and 200µL of Buffer AL, which promotes the lysis of the cell membrane and encourages the rapid denaturation of proteins and other macromolecules that may act as contaminants. The solution was mixed by vortexing and incubated at 56°C for 30 minutes and then 15 minutes at 95°C, which served to denature proteinase K. We then added 200µL of 95% ethanol to the sample and mixed by pulse-vortexing for 15 seconds.

We added the resulting mixture (including the precipitate) to a QIAamp Mini spin column (which was itself situated in a 2mL collection tube). The spin column and collection tubes were then centrifuged at 8,000 rpm for 1 minute.

Next, we added 500µL of Buffer AW1 to the spin column and centrifuged it at 8,000 rpm for 1 minute. The spin column was then put into a clean 2mL collection tube. This step was repeated with buffer AW2, except that the centrifuge was set instead to 14,000 rpm for a total time of 3 minutes. Both buffer AW1 and buffer AW2 removed any unwanted substances from the DNA contained within the spin column.

To elute the gDNA, we placed the spin column into a new 1.5mL microcentrifuge tube. We then added 200µL of Buffer AE, incubated at room temperature for 1 minute, and centrifuged the column and filtrate tube for 1 minute at 8,000 rpm. Buffer AE allowed the DNA attached to the spin column to flow to the bottom of the filtrate tube and become part of the solution. This step was repeated using the same 200µL volume to increase yield.

### C. Ultraviolet (UV) Spectroscopy

The purpose of using UV spectroscopy is to quantify the yield of genomic DNA. DNA and RNA have a peak absorbance at 260nm, and the aromatic amino acids tyrosine, phenylalanine, and tryptophan absorb at 280nm. The peptide bonds within protein chains absorb at 230nm<sup>18</sup>.

We used NanoDrop (ThermoFisher) UV spectrophotometers to complete this step. After each scan, the NanoDrop creates a graph of the sample's absorbance at each wavelength of light.  $\text{Abs}_{230}$ ,  $\text{Abs}_{260}$ , and  $\text{Abs}_{280}$  values are given by the scanner. Using these values, we calculated the purity ratio of the DNA using this formula:

$$\text{Purity} = \text{Abs}_{260} / \text{Abs}_{280}$$

where the value given is the purity of the sample with respect to protein contaminants. Ideally, each sample should have a purity ratio between 1.7 and 1.9. Having a purity below 1.7 suggests that protein contamination may have occurred, and a ratio above 1.9 could suggest RNA contamination<sup>18</sup>.

We used another formula to calculate the concentration of DNA within the sample:

$$1\text{Abs}_{260} = 50\mu\text{g} / 1\text{mL DNA}$$

Double-stranded DNA is present at a concentration of 50 $\mu$ g/mL at 260nm if it has an absorbance of 1<sup>19</sup>.

We also performed a UV scan with the NanoDrop on each sample after PCR purification (see below).

### D. Agarose Gel Electrophoresis

Agarose gel electrophoresis is a technique used to separate charged particles based on mass. The process works by running an electrical current through a gel, which pulls negatively charged DNA fragments to the positively charged end of the gel. Since the larger DNA fragments travel slower through the gel compared to smaller fragments, the fragments are separated by size (measured in base pairs (bp)).

0.7% (w/v) agarose gels were used to analyze the concentration and purity of our gDNA preps. Mini gels were run at 100V until the necessary resolution was achieved to observe differences between the gDNA and any RNA contamination. 5 $\mu$ L of sample was mixed with loading dye and loaded on the gel.

0.7% (w/v) agarose midi gels were used to confirm PCR product (PCR method described below). Gels were run at 100V with a 1kb ladder until the ladder had resolved enough to clearly see each standard. 5 $\mu$ L of each PCR product was checked on the gel.

1.2% (w/v) agarose midi gels were used to generate the genomic fingerprint of each candidate sample after restriction enzyme digests. 30 $\mu$ L of each digest was loaded on the gel and run at 100V until the 100bp ladder had fully resolved.

## E. Polymerase Chain Reaction (PCR)

Polymerase Chain Reaction (PCR) is a technique used to isolate and then amplify a very specific sequence of DNA by replicating it many times in succession. In this experiment, we amplified the 16S region of our genomic DNA. This region was chosen because it is highly conserved, is not prone to mutations, and varies slightly between bacteria. The small differences between each bacterium's 16S region help to identify those bacteria.

The PCR reaction that we used in this project has five components: a buffer solution to provide an optimal environment for reaction, the genomic DNA that serves as a template for replication, dNTPs (new nucleotides added in abundance to be used to replicate the DNA), DNA polymerase (the enzyme that drives the formation of the new DNA strand), and forward and reverse primers to tell the DNA polymerase where to begin the replication process.

In this experiment, we performed three reactions for each sample of template DNA using ThermoFisher Phusion Flash high fidelity PCR master mix (ThermoFisher). Each reaction contained a total of 1 $\mu$ L of template DNA (25-250ng of DNA), a final concentration of 0.5 $\mu$ M of each forward and reverse primer (Table 1), 2X master mix solution, and water to a final volume of 50 $\mu$ L.

**Table 1: PCR primers and their respective sequences and melting temperatures**

Primer	Sequence (5' → 3')	Melting Temperature (°C)
WLBA1F	TCG CTT TAC GCC CAA TAA ATC CGG A	71
NLAB2R	GGC GGC GTG CCT AAT ACA TGC AAG T	74

The initial denaturation step took place at 98°C for 10s. Then, there were 30 cycles of 1s at 98°C, annealing at 69.4°C for 5 seconds, and extension at 72°C for 45 seconds. There was a final extension at 72°C for 5 minutes. After this, the thermocycler stopped the reaction by cooling the reaction environment to 4°C, at which the products were held until we removed them and used them for purification the next day.

We carried out PCR for each candidate in triplicate and combined the replicates into one clean product during PCR purification.

After the reaction stopped, we were left with an extremely high concentration of PCR product, which we then ran on an agarose gel to properly view its molecular weight using a 1kb ladder as a reference.

### 1. PCR Purification

After the PCR reaction took place, we were left with a solution that still contained the components other than the template DNA that was added to facilitate the reaction, such as dNTPs, primers, and DNA polymerase. Although they were no longer functioning due to the temperature drop that took place at the end of the PCR reaction, they still needed to be separated from the replicated DNA so that the DNA could be used without interference.

We used GeneJet PCR purification kits (ThermoFisher) following the manufacturer's instructions. First, we added 45 $\mu$ L of binding buffer to our PCR product and mixed thoroughly. We then combined each replicate with product in a GeneJET purification column and centrifuged for 1 minute at 13,000 rpm and discarded the flow-through. Then, we added 700 $\mu$ L of wash buffer to the column, centrifuged for 1 minute, and discarded the flow-through. To be cautious, we centrifuged the empty column for another minute to ensure that all of the residual wash buffers had exited the column.

The clean product was transferred to a 1.5mL microcentrifuge tube. After the transfer, 50 $\mu$ L of elution buffer was added to the GeneJET column, which was then centrifuged for 1 minute and discarded once the elution containing the purified DNA had flown completely through the column. This purified DNA elution was stored in the 1.5mL microcentrifuge tube at 4C until needed.

## F. Restriction Enzyme Digests

Restriction enzymes are proteins found in bacteria whose objective is to sever foreign DNA at very specific points within a sequence called recognition sequences<sup>20</sup>. These enzymes need several components to function properly. First, all restriction enzymes require Mg<sup>2+</sup>, which permits interaction between the enzyme and the DNA to take place, to function properly. Enzymes also require unique buffer solutions that serve to replicate the most optimal conditions in which their respective enzyme can reach their maximum activity. Many enzymes also require a specific temperature to work properly<sup>20</sup>. When restriction enzymes are not in use, they must be kept at cold temperatures to inhibit any unwanted activity. This unwanted activity is called star activity, which (in general) is a scenario in which an enzyme is in a suboptimal environment and cuts at non-specific locations.

In this experiment, we performed two digests per sample of PCR product. One with the enzyme MseI, while the other was carried out with Hpy188I (NEB). MseI cuts the sequence TTAA in between the second T and the first A (TT/AA) and Hpy188I cuts the sequence TCNGA in between the N and the G (TCN/GA). 1 $\mu$ g of each PCR product was digested following the manufacturer's instructions for a 50 $\mu$ L digest:

10X rCutSmart buffer	1X
PCR Product	1 $\mu$ g
Restriction Enzyme	1L
Water	Q.S. 50 $\mu$ L

We incubated the samples at 37°C for 1 hour to ensure a complete digest while reducing the risk of star activity.

After the incubation, we put the digests on ice and added DNA dye containing EDTA to stop the reaction. We prepared the undigested samples 1:1 (5 $\mu$ L uncut sample + 5 $\mu$ L sterile water) to create a 10 $\mu$ L dilution. We loaded 5  $\mu$ L of each uncut sample and 30 $\mu$ L of our digested samples on a 1.2% agarose gel as well as 10 $\mu$ L of 100bp ladder as a reference. We ran the gel at 100V until the ladder was fully resolved.

## G. Analysis of PCR and Restriction Enzyme Gels

To analyze the data from PCR and restriction enzyme digests and calculate the products' exact molecular weights, we created a graph of distance traveled (in cm) vs. the log of molecular weight (in base pairs) for the ladder used in each gel. The molecular weight of each band on the ladder is a predetermined value and we used a ruler to measure the distance from the well (starting point) to the far end of the band

(endpoint) to get the total distance that that band of DNA migrated along the gel. After graphing these values as a standard, we then found the line of best fit (and its associated equation) to represent the data points and used that equation to calculate the exact molecular weights of the PCR or RE digest products using the distance migrated as the x value. Our equations were as follows:

#### **Initial PCR Product**

Samples B1, B2, B3, B4:  $\log(\text{bp}) = -0.208(\text{distance migrated}) + 4.61$  ( $R^2 = 0.990$ )  
Samples B5, B6, B7, B8:  $\log(\text{bp}) = -0.262(\text{distance migrated}) + 5.47$  ( $R^2 = 0.994$ )  
Samples B3, B4, B5, B6:  $\log(\text{bp}) = -0.231(\text{distance migrated}) + 4.96$  ( $R^2 = 0.997$ )  
Samples K1, K2, K3, K4:  $\log(\text{bp}) = -0.224(\text{distance migrated}) + 4.94$  ( $R^2 = 0.996$ )  
Samples K3, K4, K5, K6:  $\log(\text{bp}) = -0.404(\text{distance migrated}) + 5.41$  ( $R^2 = 0.995$ )

#### **Restriction Enzyme Digests**

Samples B1, B2, B3, B4:  $\log(\text{bp}) = -0.181(\text{distance migrated}) + 4.62$  ( $R^2 = 0.965$ )  
Samples B5, B6, B7, B8:  $\log(\text{bp}) = -0.181(\text{distance migrated}) + 4.61$  ( $R^2 = 0.982$ )  
Samples B3, B4, B5, B6:  $\log(\text{bp}) = -0.179(\text{distance migrated}) + 4.69$  ( $R^2 = 0.982$ )  
Samples K1, K2, K3, K4:  $\log(\text{bp}) = -0.189(\text{distance migrated}) + 4.61$  ( $R^2 = 0.969$ )  
Samples K3, K4, K5, K6:  $\log(\text{bp}) = -0.233(\text{distance migrated}) + 5.20$  ( $R^2 = 0.985$ )

To calculate the molecular weight (bp), we had to take the log of each of their equations. Example for PCR product samples B1 through B4:

$$\text{Molecular weight} = 10^{-0.208(\text{distance migrated}) + 4.61}$$

## **H. DNA Sequencing**

We preserved a portion of each of our purified PCR products in order to send them for DNA sequencing. To prepare these samples for sequencing, we diluted each 1:100, producing new samples with 1 $\mu\text{L}$  of product and 99 $\mu\text{L}$  of sterile water. A portion of each of these new dilutions was then added to however much water would make up 20 $\mu\text{L}$  when added to the dilution. For example, if we needed 9 $\mu\text{L}$  of a sample, we would add 11 $\mu\text{L}$  of water to create a total volume of 20 $\mu\text{L}$ . We found the amount of dilution that we needed to add by dividing the concentration of DNA in each sample (found from the UV scans we performed after PCR) by 100 and then dividing 20 by that quotient.

We requested that Azenta Life Sciences use Sanger sequencing, also known as the chain termination method, which we chose because it remains the gold standard for accuracy when it comes to DNA sequencing.

## **III. Results & Discussion**

### **A. Results**

#### **1. UV spectroscopy of gDNA**

**Table 2: UV Absorbance Data of gDNA Samples**

Sample	Concentration (ng/µL)	Abs. at 230nm	Abs. at 260nm	Abs. at 280nm	Purity Ratio
B1	198.451	1.601	3.969	1.887	2.1
B2	201.715	1.614	4.034	1.934	2.09
B3	134.768	1.106	2.695	1.289	2.09
B4	94.11	0.796	1.882	0.917	2.05
B5	84.323	0.73	1.686	0.834	2.02
B6	89.881	0.75	1.798	0.876	2.05
B7	130.161	1.021	2.603	1.256	2.07
B8	98.758	0.861	1.975	0.974	2.03
K1	65.519	0.558	1.31	0.639	2.05
K2	82.698	1.052	1.654	0.82	2.02
K3	27.365	0.35	0.547	0.288	1.9
K4	21.285	0.377	0.426	0.212	2.01
K5	32.568	0.646	0.651	0.418	1.56
K6	31.15	1.075	0.623	0.451	1.38
K7	20.823	0.537	0.416	0.287	1.45
K8	15.587	0.417	0.312	0.227	1.37

This is the data that we obtained after performing spectroscopy on our samples. The concentration levels for our samples are all over the place which suggests that some of our gDNA isolation techniques might have not gathered enough gDNA then we would have liked. Our absorption values are good for samples B1-B8 and samples K1-K4. Our absorption values for samples K5-K8 are less than we would have liked. When looking at our purity of our samples we ideally want our purity to fall between 1.7-1.9. Only one of our samples (K3) falls into this category. This can suggest that our samples might have been contaminated with RNA or another possibility is that our samples could be contaminated by proteins. If the absorption value is above the range that suggests that RNA contamination is present, if the purity is lower then it suggests protein contamination.

## 2. Genomic DNA Preps

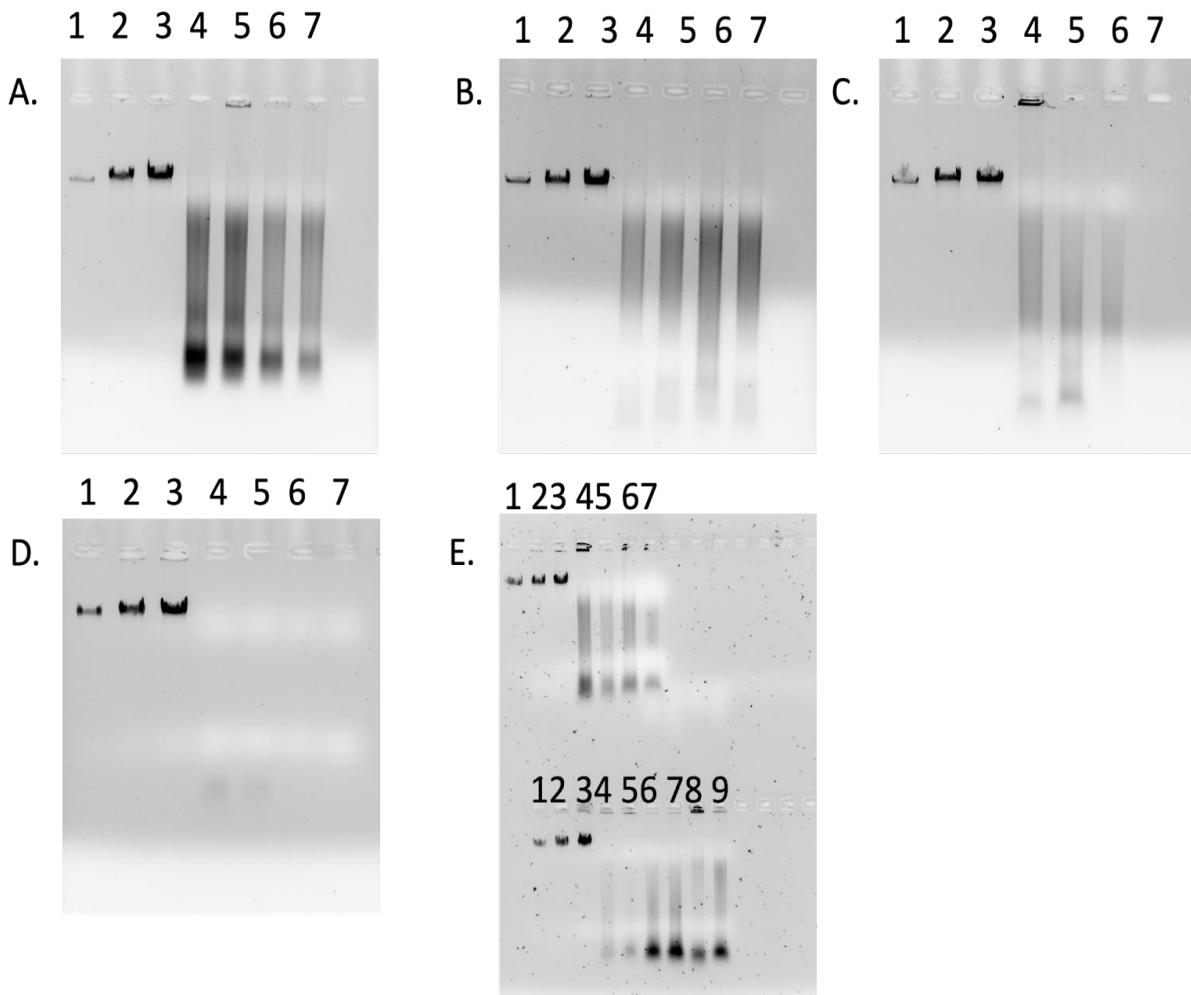


Fig. 2 | Gel electrophoresis of samples after DNA isolation

- A. Samples B1, B2, B3, and B4. Lane 4 is B1, lane 5 is B2, lane 6 is B3, lane 7 is B4
- B. Samples B5, B6, B7, and B8. Lane 4 is B5, lane 5 is B6, lane 6 is B7, lane 7 is B8
- C. Samples K1, K2, K3, and K4. Lane 4 is K1, lane 5 is K2, lane 6 is K3, lane 7 is K4
- D. Samples K5, K6, K7, and K8. Lane 4 is K5, lane 5 is K6, lane 6 is K7, lane 7 is K8
- E. Samples B3, B4, B5, and B6 (top). Lane 4 is B3, lane 5 is B4, lane 6 is B5, lane 7 is B6; samples K3, K4, K5, and K6 (bottom). Lane 4 is K3, lane 5 is K4, lane 6 is K5, lane 7 is K6

After isolating the gDNA from the bacteria samples we want to check to see if we had any bacterial gDNA to work with. The bands in the gels signify the presence of DNA. The darker, more concentrated patches towards the bottoms of the lanes indicate the presence of RNA. Protein itself does not appear on agarose gels, but DNA that it is attached to will be seen at the top of the gel near the wells. The first three lanes on all the gels are mass markers which lets us see the concentration of the proteins, the darker the color the more concentration. From lanes 1 to 3, the mass markers became darker the more of the mass marker we put into the wells, which is the result that we wanted. gDNA will be located about a quarter after the

wells and will continue until about the middle of the gel. The ideal situation when looking over these gels would be to have a solid line that starts about a quarter down the gel and fades towards the end.

In gels 2B, and 2E (top) you can see this ideal result. The dark spots that are still located in the wells show that gDNA is trapped in the well, likely due to being stuck in protein contamination. The solid band shows that there is gDNA within the samples. Since the line fades towards it suggests that there is little to no RNA contamination within the sample.

In gels 2A and 2E (bottom) we can see that both samples have a dark spot still present within the wells indicating possible protein contamination as discussed above. The darker spots of low molecular weight indicate RNA contamination. Taking this into account we considered revisiting these samples and doing an RNA cleanup. To test if cleanup was necessary, we did a test run of PCR with a few of the gDNA preps (data not shown). Each showed good amplification, so no RNA cleanup was necessary.

In gel 2C you can see that gDNA is only present in 3 of the 4 lanes. That means that the sample in lane 4 was not loaded into the gel or that it had little to no gDNA within it. After seeing this we considered isolating another sample, but after a small test round of PCR, we found PCR products meaning there was no need for the reisolating.

In gel 2D there is no gDNA. We suspect that during the gDNA prep there was no DNA isolated. However, the test PCR (data not shown) and the UV scans (Figure X) suggested that DNA was present in the sample, so we decided to use these samples for the experimental PCR run.

### 3. Polymerase Chain Reaction

**Table 3: UV Spectroscopy of PCR Product**

Sample	Concentration (ng/ $\mu$ L)	Abs. at 260nm	Abs. at 230nm	Abs. at 280nm	Purity Ratio
B1	223.713	4.472	3.21	2.382	1.88
B2	225.132	4.503	2.316	2.4	1.88
B3	203.81	4.076	2.441	2.179	1.87
	181.874	3.637	3.532	1.936	1.88
B4	190.445	3.809	3.843	2.026	1.88
	176.63	3.533	2.815	1.893	1.87
B5	221.051	4.421	2.495	2.338	1.89
	129.965	2.599	5.689	1.499	1.73
B6	324.499	6.49	6.662	3.454	1.88
	196.772	3.935	4.217	2.229	1.77
B7	229.256	4.585	5.137	2.487	1.84
B8	178.98	3.58	2.857	1.937	1.85
K1	207.47	4.149	3.94	2.288	1.81
K2	229.524	4.59	4.545	2.704	1.7
K3	191.803	3.836	3.514	2.08	1.84
K4	136.351	2.727	3.376	1.477	1.85

The leftmost column represents the concentration in nanograms per microliter of the PCR product samples. All of the samples have at least relatively high concentrations, leading to the conclusion that they are usable in future steps of our project. This conclusion is further supported by the purity ratios that can be seen in the rightmost column, which are all within the ideal range of 1.7 - 1.9. This means that the samples are not contaminated by anything that would affect the subsequent procedures or data collection. The high absorbance values at 260nm signify the extremely high DNA concentration that should be present after PCR takes place. Overall, the values in this table tell us that these PCR product samples are worth using for restriction enzyme digests and data collection.

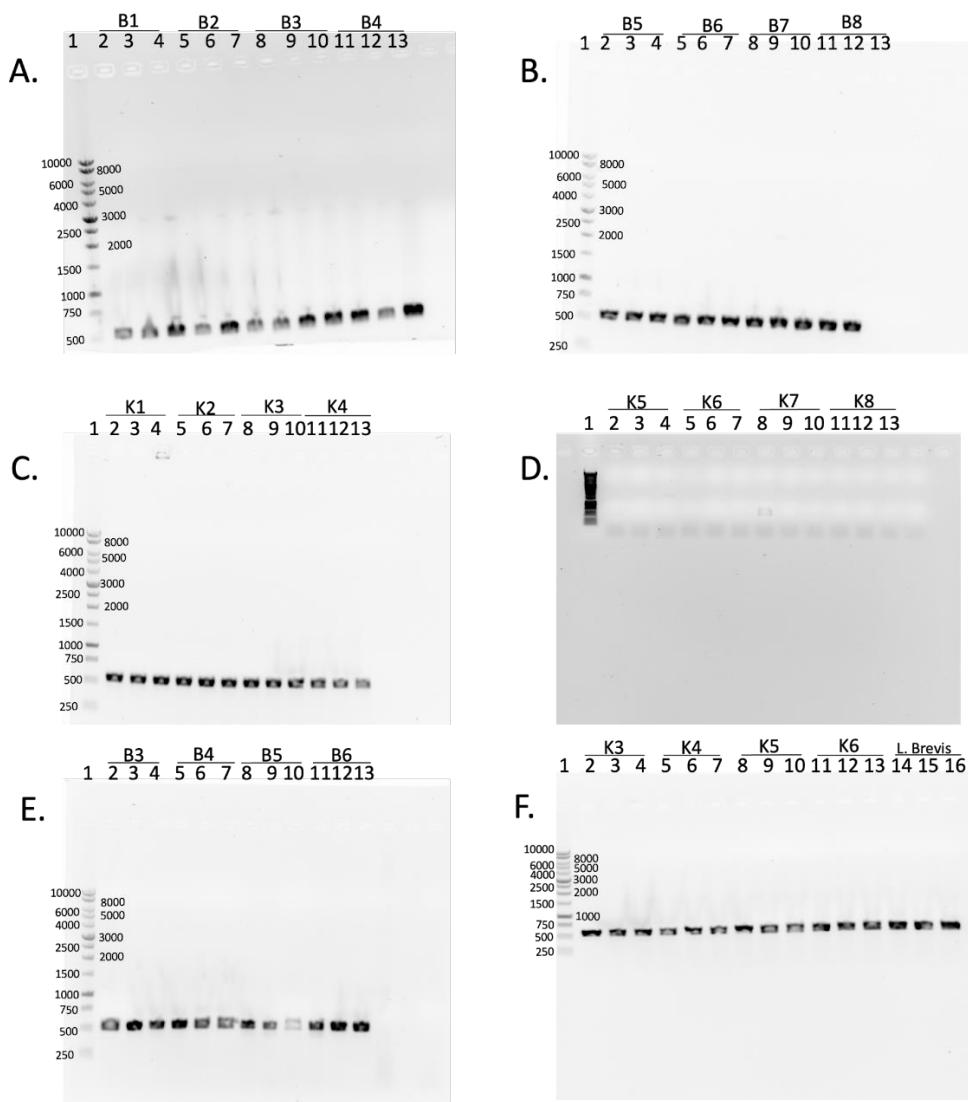


Fig. 3 | Gel electrophoresis after PCR

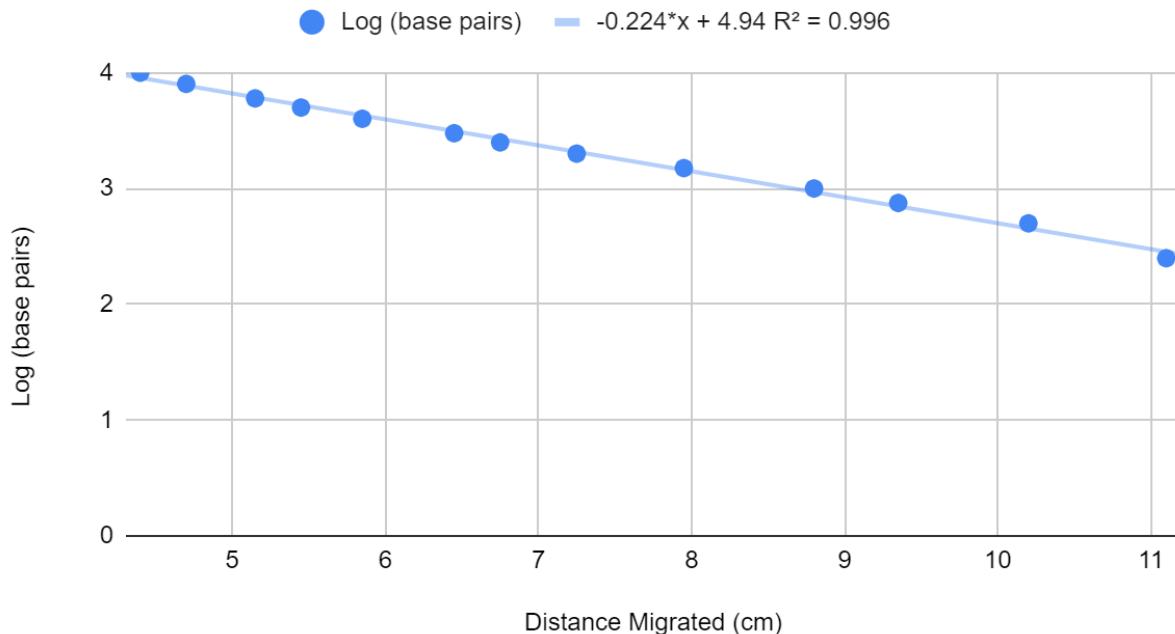
- A. Samples B1, B2, B3, B4. Lanes 2-4 are B1, lanes 5-7 are B2, lanes 8-10 are B3, lanes 11-13 are B4
- B. Samples B5, B6, B7, B8. Lanes 2-4 are B5, lanes 5-7 are B6, lanes 8-10 are B7, lanes 11-13 are B8
- C. Samples K1, K2, K3, K4. Lanes 2-4 are K1, lanes 5-7 are K2, lanes 8-10 are K3, lanes 11-13 are K4
- D. Samples K5, K6, K7, K8. Lanes 2-4 are K5, lanes 5-7 are K6, lanes 8-10 are K7, lanes 11-13 are K8
- E. Samples B3, B4, B5, B6. Lanes 2-4 are B3, lanes 5-7 are B4, lanes 8-10 are B5, lanes 11-13 are B6
- F. Samples K3, K4, K5, K6, L. Brevis. Lanes 2-4 are K3, lanes 5-7 are K4, lanes 8-10 are K5, lanes 11-13 are K6 lanes 14-16 are L. Brevis

Following the PCR reaction, we ran a small portion of each product on an agarose gel (Figure 3) to ensure there was product before continuing with the experiment. One ladder and 4 samples (in triplicate) were run on each gel. The 10kb ladder that was used ranged from 10,000 base pairs to 250 base pairs. To confirm that there was PCR product, we looked for a single, clear band in each lane that was loaded with the PCR samples. Our control sample, *L. del*, was supposed to have a PCR product of 500 base pairs, which is seen in figure 3F.

In figures 3A., 3C., 3E., and 3F, there is PCR product in each of the three lanes for every one of the 4 samples that were included in the test. In figure 3B., there is no PCR product for the third replicate of sample B8 (see lane 13). The gel in figure 3D. had no PCR product. This was likely due to an error during

the gDNA prep that prevented any DNA from being gathered. These samples (K5, K6, K7, K8) were not used in subsequent analysis.

### PCR Products of K1, K2, K3, and K4



**Figure 4: Graphs relating distance traveled of the ladder and base pairs of the ladder (see the rest in appendix)**

The graphs were made by measuring the distance that the ladder traveled down the gel. We related that to the log<sub>10</sub> of the molecular weight or the Kilobase pairs. After making a graph out of them we find the line of best fit and use the equation against our product lengths. This results in a close estimation of the size of each PCR product.

**Table 4: Average Molecular Weights of PCR Products**

<b>Sample</b>	<b>Average Molecular Weight (base pairs)</b>	
	Replicate 1	Replicate 2
<b>B1</b>	489.3	n/a
<b>B2</b>	497.0	n/a
<b>B3</b>	497.0	479.6
<b>B4</b>	505.2	505.7
<b>B5</b>	248.9	519.4
<b>B6</b>	220.6	492.5
<b>B7</b>	207.7	n/a
<b>B8</b>	187.8	n/a
<b>K1</b>	440.7	n/a
<b>K2</b>	411.3	n/a
<b>K3</b>	404.3	514.9
<b>K4</b>	397.4	514.9
<b>K5</b>	n/a	514.9
<b>K6</b>	n/a	514.9

After calculating the line of best fit, we calculated the estimated size of the PCR products in base pairs. Table 5 shows these estimations, and we can see that all samples have PCR products. The right-most column is the cross-over of the same products that were tested twice.

#### 4. Restriction Enzyme Digests

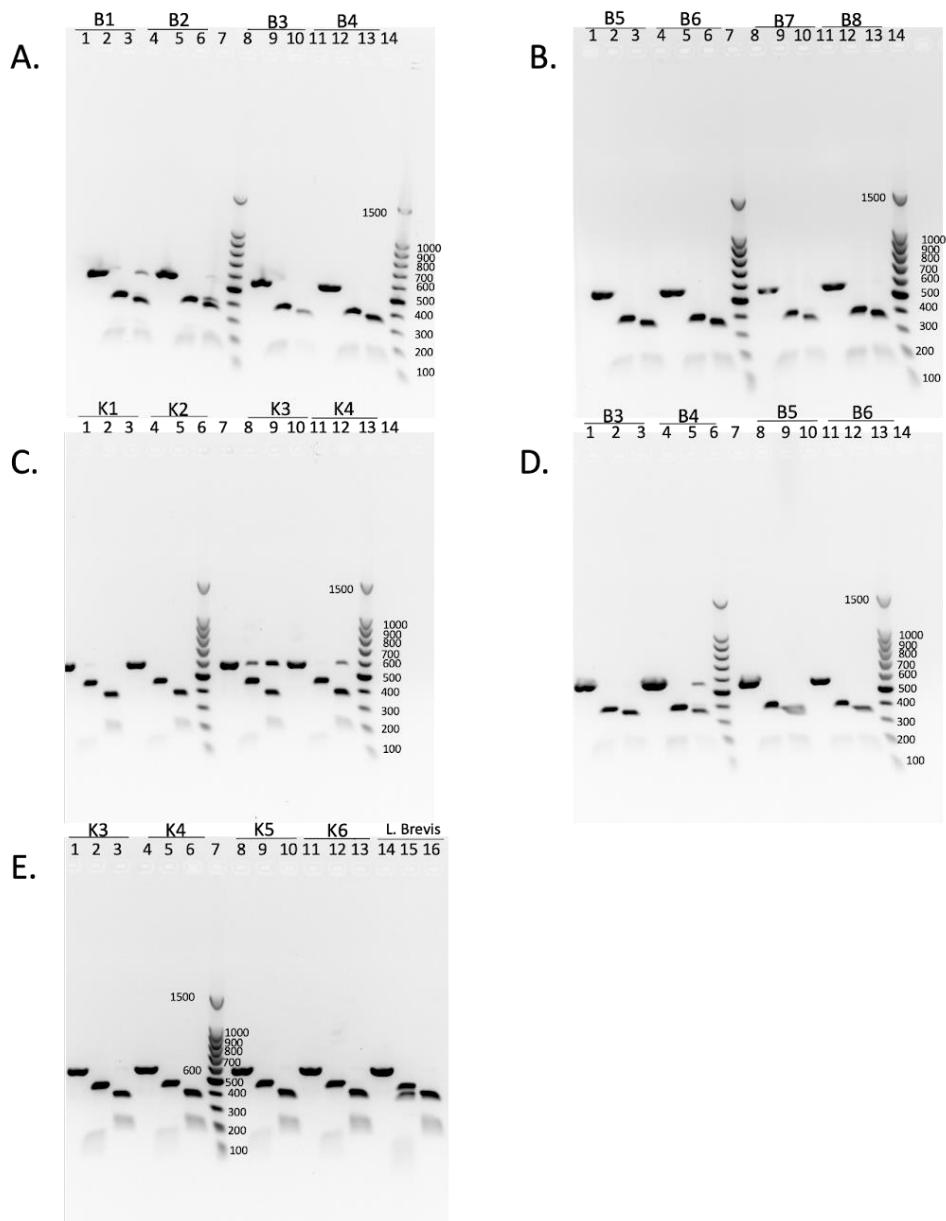


Fig. 6 | Gels that were done once we had finished the final digest

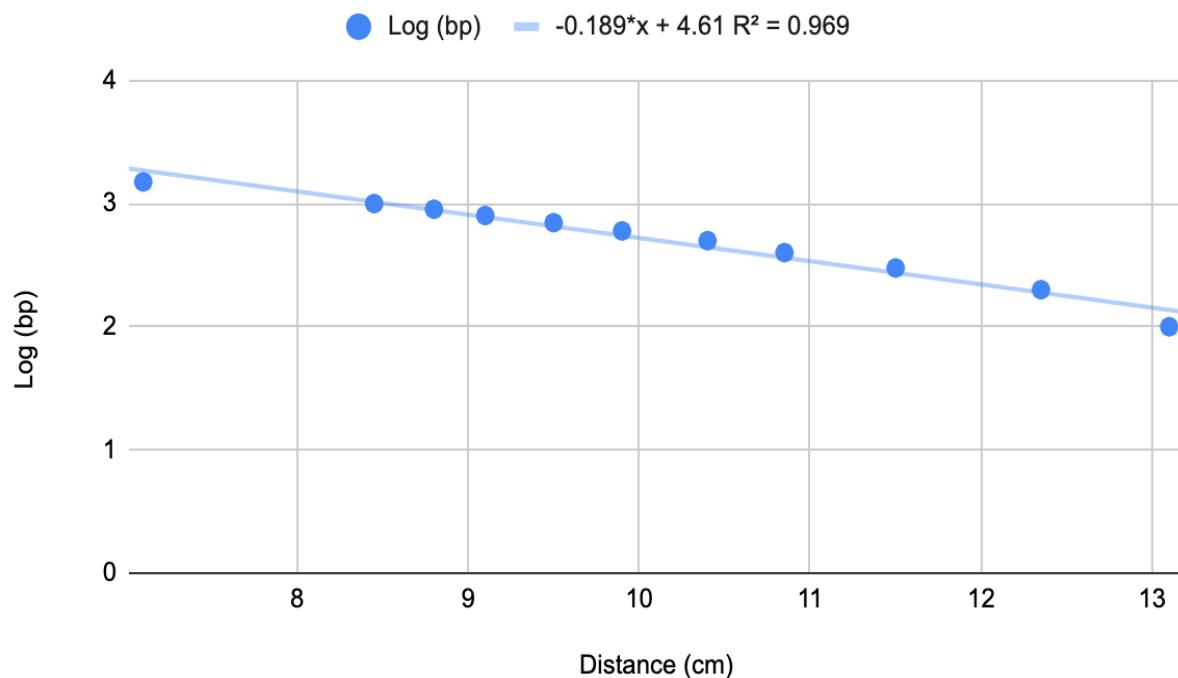
- A. Samples B1, B2, B3, B4. Lanes 1-3 are B1, lanes 4-6 are B2, lanes 8-10 are B3, lanes 11-13 are B4
- B. Samples B5, B6, B7, B8. Lanes 1-3 are B5, lanes 4-6 are B6, lanes 8-10 are B7, lanes 11-13 are B8
- C. Samples K1, K2, K3, K4. Lanes 1-3 are K1, lanes 4-6 are K2, lanes 8-10 are K3, lanes 11-13 are K4
- D. Samples B3, B4, B5, B6. Lanes 1-3 are B3, lanes 4-6 are B4, lanes 8-10 are B5, lanes 11-13 are B6
- E. Samples K3, K4, K5, K6, L. Brevis. Lanes 1-3 are K3, lanes 8-10 are K5, lanes 11-13 are K6, lanes 14-16 are L. Brevis

Loading order for each sample is uncut, MseI digest, Hpy188I digest

After confirmation that the PCR produced products and was cleaned up, we performed restriction digests to generate a unique DNA fingerprint for each sample. When digesting our DNA, we used two different restriction enzymes MseI and Hyp188I. Each enzyme will make cuts at different regions along the gDNA

sequencing. Agarose gel electrophoresis was used to determine the size of each digest fragment (Figure 6). We used the 100bp ladder which measures 1500bp to 100bp combined with some computational analysis to determine this. When looking for our results we wanted bands around the 500bp ladder. This will allow us to see how many segments the enzymes will cut our pieces of DNA into and also the size of each segment. Using the 100bp ladder we created graphs (Fig. 7) that allow us to estimate the sizes of the kilobases of the digest products.

## Digest Product of K1, K2, K3, K4



**Figure 7: Graphs of digest products relating the distance the ladder traveled (x-axis) and the log of the molecular weight of the ladder (y-axis) (see the rest in appendix)**

The graphs were made by measuring the distance traveled by each band on the 100 bp ladder of the digested gel and then relating it to the log<sub>10</sub> value of the molecular weight (base pairs) of each band. The equations generated by these graphs were used to calculate the estimated sizes of the bands on the PCR digest gels. After collecting the data, we can determine the bacteria genus of each of our samples using computer software.

**Table 5: Estimated Molecular Weight of Restriction Enzyme Digest Products**

	Uncut (bp)	Msel (bp)	Hyp188I (bp)	Uncut (bp)	Msel (bp)	Hyp188I (bp)
<b>B1</b>	524.2	375.6	345.5			
		345.5	184.9			
<b>B2</b>	524.2	375.6	345.5			
		177.4	170.1			
<b>B3</b>	482.3	360.2	345.5	484.4	341.2	320.8
		163.2	177.4		187.7	191.6
<b>B4</b>	482.3	360.2	317.9	504.8	341.2	320.8
		170.1	163.2		183.9	176.5
<b>B5</b>	482.3	345.5	331.4	515.3	370.6	341.2
		163.2	177.4		180.1	191.6
<b>B6</b>	482.3	345.5	331.4	537.0	370.6	355.6
		163.2	170.1		195.6	195.6
<b>B7</b>	502.8	360.2	317.9			
		170.1	170.1			
<b>B8</b>	502.8	360.2	331.4			
		177.4	163.2			
<b>K1</b>	422.2	431.5	347.1			
		142.2	197.1			
<b>K2</b>	536.4	431.5	347.1			
<b>K3</b>	536.4	422.2	354.7	566.9	433.5	369.1
		151.8	210.4		152.3	210.1
<b>K4</b>	536.4	422.2	354.7	582.3	457.4	379.1
		158.6	205.9		221.7	215.8
<b>K5</b>	582.3	274.8	389.4			
		169.5	215.8			
<b>K6</b>	598.1	482.6	400.0			
		221.7	221.7			
<b>L. delbruekii</b>	598.1	469.8	389.4			
		140.5	215.8			

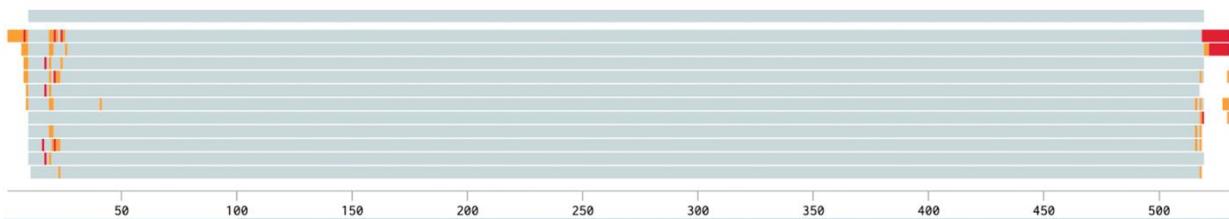
The table illustrates the two restriction enzymes that were used to digest our samples. The results we got were how many segments there were once digested. It will also tell us what the estimated sizes of each of the bands from the gel are. We used this data and plug it into the computer science team's algorithm, which able to give a list of possible bacteria that our samples may be.

## 5. DNA Sequencing

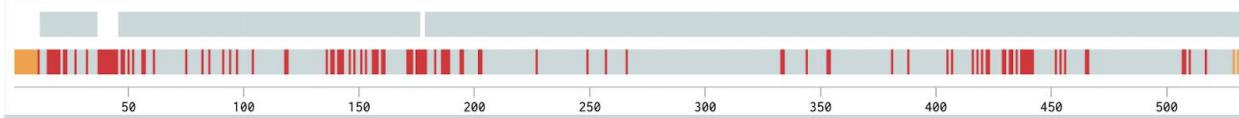
### A. K-F Alignments



### B. B-F Alignments



### C. K1 F - B1 F Alignment



**Figure 8: DNA sequencing of samples**

- A. K samples using the forward primer
- B. B samples using the forward primer
- C. B1 and K1 compared to each other

Sequencing was completed by Azenta Life Sciences. The data that was sent back to us included the DNA sequences of the samples that we sent to them. We wanted to first compare our buttermilk samples against each other. We do this by aligning the sequence data that was previously mentioned and see where different base pairs are located. The red lines mark a spot where the samples have a different base pair when compared to the first line. The gray areas show where the samples have the same base pair compared to the first line. We also know that there will be differences towards the beginning and ending of our sample just from the Sanger Method not because the samples are different. From further inspection of the beginning and ending of our data we found that most of these mismatches are a result of the sequencing machine not recognizing the nucleotide which is a limitation of the Sanger Method of

sequencing. The alignment suggests that all the B samples are the same bacteria, and all the K bacteria are the same bacteria since there are no red markers within the middle of the sequencing data. It also suggests that the K samples are different bacteria from the B samples. This means that we have collected two different bacteria species within the 20 samples that we tested.

There are several possible situations that may have occurred that could explain why we only obtained two bacteria species out of 20 candidates. The first possibility is that, when we initially gathered our unknown microbe samples during the streak plating portion of our method, we could have gathered from the same bacterial colonies. Another possibility is that multiple species of bacteria were present in the buttermilk and kefir, but that only two were able to survive the experimental protocols. That means that we may have collected more than one species of bacteria, however, after completing all of the protocols, some of the bacterial gDNA was lost due to non-ideal conditions for the bacteria's survival.

**Table 6: Results of possible bacteria matches after running them through the BLAST program**

	Description	Scientific name	Query Cover	E Value	Percent Identical
B1	Lactococcus lactis strain 1881 16S ribosomal RNA gene, partial sequence	Lactococcus lactis	99%	0	100%
K1	Lacticaseibacillus paracasei strain HBUAS68383 16S ribosomal RNA gene, partial sequence	Lacticaseibacillus casei	99%	0	100%
L. Del	Lactobacillus brevis strain Sh49 16S ribosomal RNA gene, partial sequence	Levilactobacillus brevis	100%	0	100%

After the alignment, we used NCBI Nucleotide Blast to identify the species of each unique sample (Table 6). From the alignment analysis, we knew there was only one unique species isolated from Kefir and from buttermilk and we chose to only do blast searches on one candidate from each dairy product. The program determined the possible bacteria species that we cultured by checking the gDNA sequences that were obtained and comparing them against known bacterial gDNA. From the BLAST program we can compare our gDNA to the known gDNA (Query Cover), the likelihood of the match being from random chance (E-value), and how much of our gDNA sequence matched the query (Percent Identical).

Sample B1 was identified as *Lactococcus lactis*. Since we know that the DNA sequence is the same for all B samples, we know that all B samples are *Lactococcus lactis*. We can also see that there is a 100% identical match, meaning that the exact sequence of gDNA that was sequenced matched *Lactococcus lactis* 100%.

Sample K1 was identified as *Lactobacillus casei*. Since we know that the DNA sequence is the same for all K samples, we know that all K samples are *Lactobacillus casei*. We can also see that there is a 100% identical match meaning that the exact sequence of gDNA that we got matched *Lactobacillus casei* 100%. We also ran a control, *L. delbruekii*, sample through the same protocols that our other samples went through. However, when running the sequencing DNA that we obtained, it matched for *Levilactobacillus*

*brevis* instead. There are several errors that could have occurred to cause this result. One is that the company that we purchased the sample from did not send the proper sample. Another thing that could have happened is that the sample could have been contaminated with *Levilactobacillus brevis*, which in turn caused us to culture it instead. We also could have cultured the wrong sample during the initial steps of the project.

## IV. Conclusion

Each step of this project taught us something different about the process and brought us closer to our final result.

The first step, extracting genomic DNA from the bacteria samples that were isolated for us, gave us the template that we would use for the rest of the project from PCR to digests. We familiarized ourselves with some of the more common techniques that we would end up using later on in the process, such as elutions and working with spin columns. This extraction step was successful for twelve of our 16 samples. Although we would not know this until later, some aspect of our gDNA prep did not produce the desired result (a product with genomic DNA) for samples K5, K6, K7, and K8. The first indication of this came when we ran K5 - K8 on an agarose gel and saw no product. We became more suspicious when we performed UV scans of the supposed gDNA of samples K5 - K8 and saw that the DNA concentrations and purity ratios were well below the normal range. Because there was evidence of some type of error before PCR took place, we believe that the gDNA prep went wrong in some capacity for samples K5, K6, K7, and K8. The most likely explanation for this was a procedural error in the preparation of the genomic DNA for those four samples.

Despite this shortcoming, the agarose gels and UV scans were clear and within the normal range for samples B1 - B8 as well as for samples K1 - K4.

After the first round of UV scans, we carried out PCR. PCR served as the center point of the preparation process and allows us to study the 16S region specifically in the latter portion of the project. We learned about the process of PCR, the components of the reaction, as well as how those components combine to allow PCR to take place as it does. The PCR product was invaluable in our studies and allowed us to carry out the execution portion of our project. After PCR, we ran another agarose gel and UV scan in order to detect any contamination within our PCR product, which we then purified in order to eliminate any aforementioned contamination. Our suspicions of some form of error with samples K5 - K8 were confirmed at this point in the project when we performed PCR on those samples and found that there was no PCR product present after the reaction.

To note, at this point, samples K5, K6, K7, and K8 were no longer included in our dataset. Instead, we carried out the rest of the steps beginning with PCR purification with duplicates of samples B3, B4, B5, and B6. That is why there appears to be two sets of samples B3, B4, B5, and B6 in many of our tables.

Perhaps the most important step in the process that we carried out was the restriction enzyme digests. These digests cut the 16S region into pieces of specific sizes—sizes that would form the backbone of the values that we handed off to the computer science team in order to continue the project as planned. Performing these digests in the manner that we did was possible due to the favorable properties of the 16S region that we mentioned throughout this paper. As evidenced by the expected DNA bands on each

post-digest gel being present in their expected locations, this step went well and produced the desired results. This step went exactly as we predicted it would.

The purpose of analyzing the PCR and enzyme digest gels was to obtain data that could be given to the Computational Biology Team. By using the ladder as a reference and creating a standard curve, we were able to use the resulting equation to determine the approximate sizes of our digest products. These molecular weight estimates were then run through the computer program to generate a list of possible lactic acid bacteria "identity" matches. During the process of creating the enzyme digest graphs, we found that some bands on the gel migrated much less than others, resulting in a molecular weight value that was higher than 500 base pairs. These fragments were likely undigested pieces of DNA that were only visible when we looked at high contrast images of the gels. Because these pieces of data could not be used due to the fact that they represented undigested DNA, they were excluded from final tables and data sets. This analysis, since it relied on the success of the restriction enzyme digests, was successful as well. Although we had a few unexpected bits of product that we had to exclude because, in all likelihood, they were just undigested, everything else worked out as evidenced by the fact that each enzyme cut our PCR product into pieces whose sizes, when added together, are equivalent to the approximate size of the undigested PCR product.

The Genewiz DNA sequencing provided us with a way to confirm the validity of the results that were generated by the computational analysis program that was created by the Computational Biology Team. DNA sequencing results indicated that a single species of bacteria was found in the B samples and that a single, different species was present in all of the K samples. It also allowed us to see that our control bacteria were different from what we had originally believed it to be.

Finally, the BLAST analysis was used to determine the specific names of the bacteria species in the B samples, K samples, and control samples. The results were achieved by comparing the DNA sequences provided by the sequencing company to a database of known bacterial DNA sequences. We can rely on these results being accurate because of the favorable query cover, percent identity, and e-values that correspond to the BLAST results.

## A. Future Directions

The next step for this project would be to refine the pipeline. Although the general sequence of the project will not change, there are some smaller things that we could do differently given more time. For example, we could have tried using more than two restriction enzymes to digest the PCR product to gain additional fingerprinting data and make a more refined computational analysis. We also could streamline our technique throughout the project. For example, many of our agarose gels came out a little bit blurry and overblown when we ran our PCR product on them. If we had the opportunity to refine our project, we would load less of the DNA samples or a less concentrated version of the samples in order to get more precise results from the gels. This would allow a more precise calculation of band sizes.

Another direction that we could take with this project would be to attempt to apply it to other scenarios such as water testing for companies like the Eastern Pennsylvania Coalition for Abandoned Mine Reclamation (EPCAMR) that test river water and need to know its specific bacterial content. Although identifying bacteria is an integral part of how we consider and study probiotics, processes like water testing use much of the same concepts and have many of the same problems that bacteria identification has.

Testing can be expensive, uncommon, and inefficient, but with a few tweaks, we could tailor our pipeline to fit very well as an alternative to the current system.

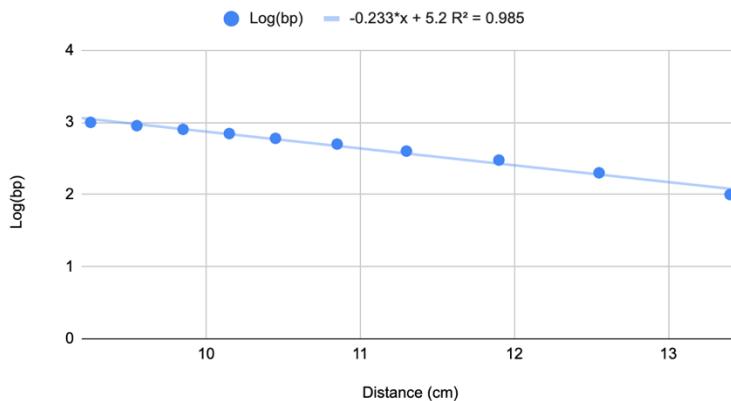
## V. Acknowledgments

We would like to thank Dr. Barry Luokkala, Dr. Natalie McGuier, Mr. Michael Sobol, Melissa Lessure, and the PGSS Campaign, Inc. for their tremendous help and support throughout the course of this project.

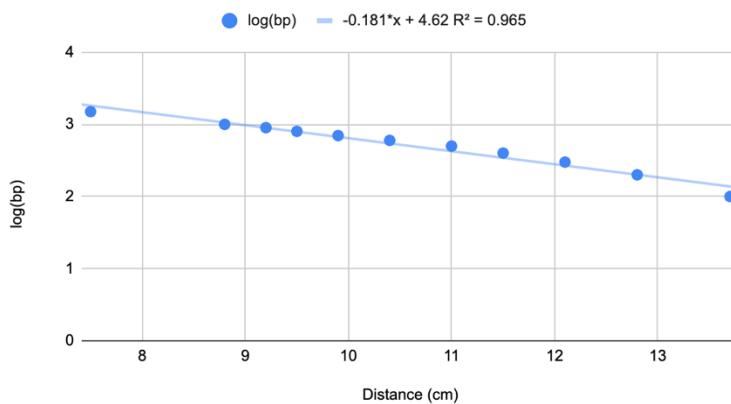
We would also like to extend special thanks to our parents, guardians, and other family members for the support they have given us during our time at PGSS. We could not have done it without them.

## Appendix A: Restriction Enzyme Digest Product Graphs

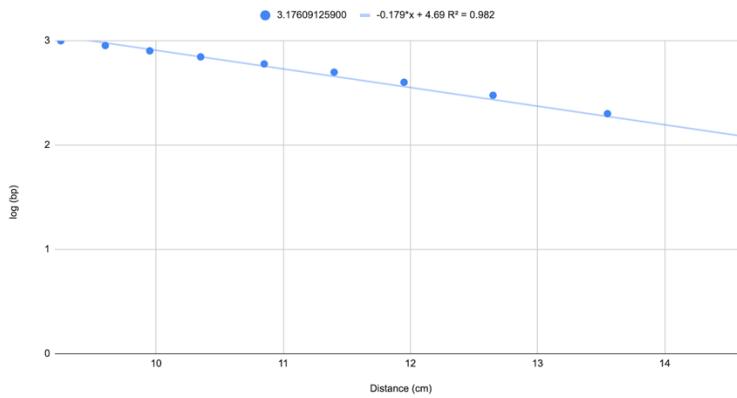
Digest Product of K3, K4, K5, K6



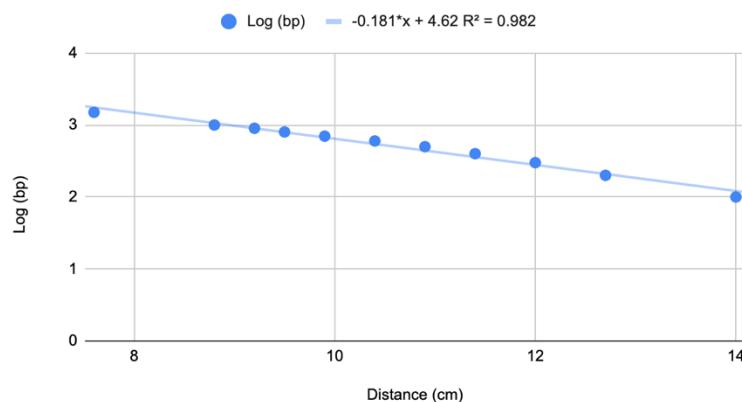
Digest Products of B1, B2, B3, B4



Digest Product of B3, B4, B5, B6



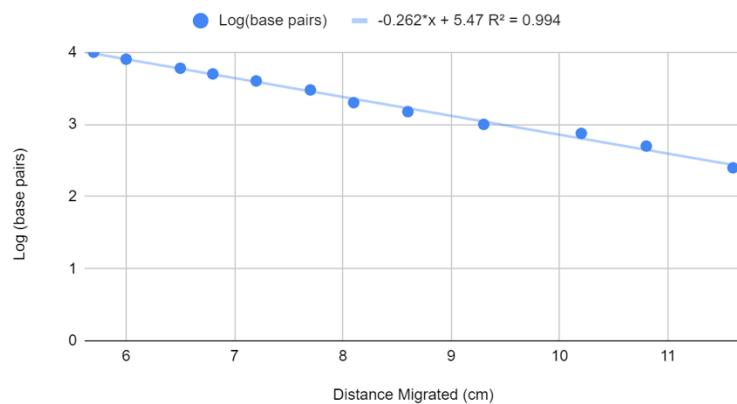
### Digest Product of B5, B6, B7, B8



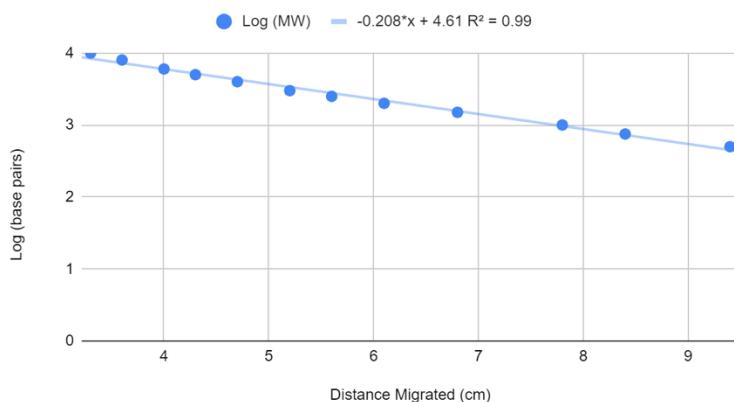
The graphs were made by measuring the distance traveled by each band on the 100 bp ladder of the digested gel and then relating it to the log<sub>10</sub> value of the molecular weight (base pairs) of each band. The equations generated by these graphs were used to calculate the estimated sizes of the bands on the PCR digest gels. After collecting the data we can determine the bacteria genus of each of our samples using computer science.

## Appendix B: PCR Product Graphs

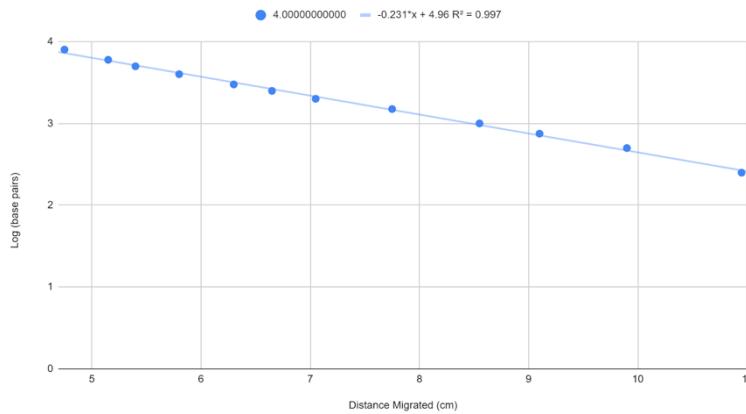
### PCR Products of B5, B6, B7, and B8



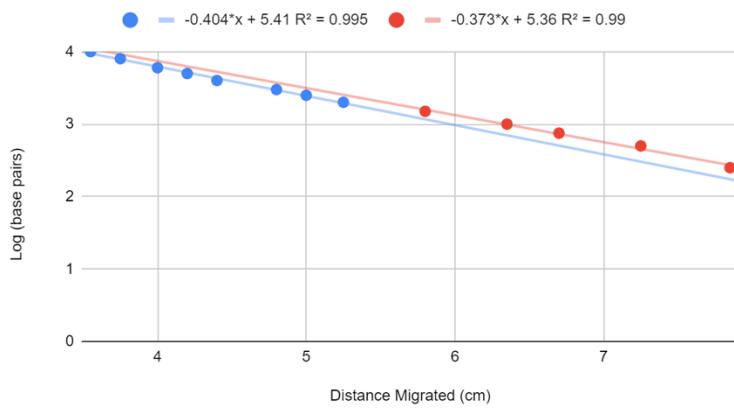
## PCR Products of B1, B2, B3, and B4



## PCR Products of B3, B4, B5, and B6



## PCR Products of K3, K4, K5, K6



The graphs were made by measuring the distance that the ladder traveled down the gel. We related that to the log<sub>10</sub> of the molecular weight or the Kilobase pairs. After making a graph out of them we find the line of best fit and use the equation against our product lengths. This results in a close estimation of the size of each PCR product.

## VII. References

<sup>1</sup>McCartney, B. (2021, July). *Microbiota* [Google Slides].

<https://canvas.cmu.edu/courses/29562/pages/e5-you-and-your-microbiota>.

<sup>2</sup>The gut microbiota. (2021, August 2). Biocodex Microbiota Institute. Retrieved July 21, 2022, from <https://www.biocodexmicrobiotainstitute.com/en/gut-microbiota>

<sup>3</sup>Thursby, E., & Juge, N. (2017, May 16). *Introduction to the human gut microbiota - PMC*. NCBI.

Retrieved July 21, 2022, from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5433529/>

<sup>4</sup>Klaenhammer, T. R. (2010, July 28). *Colloquium Paper: Succession of microbial consortia in the developing infant gut microbiome*. NCBI. Retrieved July 21, 2022, from

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3063592/>

<sup>5</sup>Musso, G. (2010, October). *Obesity, Diabetes, and Gut Microbiota - PMC*. NCBI. Retrieved July 21, 2022, from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2945175/>

<sup>6</sup>Neish, A. S. (2008, November 19). *Microbes in gastrointestinal health and disease*. PubMed. Retrieved July 26, 2022, from <https://pubmed.ncbi.nlm.nih.gov/19026645/>

<sup>7</sup>Gasbarrini, G. (2016, December). *Probiotics History*. Journal of Clinical Gastroenterology. Retrieved June 19, 2022, from [https://journals.lww.com/jcge/fulltext/2016/11001/probiotics\\_history.3.aspx](https://journals.lww.com/jcge/fulltext/2016/11001/probiotics_history.3.aspx)

<sup>8</sup>Gogineni, V. K. (n.d.). Probiotics: History and Evolution. *Journal of Infectious Diseases and Preventative Medicine*. 10.4172/2329/8731.100107

<sup>9</sup>NCCIH. (2019, August). *Probiotics: What You Need to Know* | NCCIH. National Center for Complementary and Integrative Health. Retrieved July 28, 2022, from <https://www.nccih.nih.gov/health/probiotics-what-you-need-to-know>

<sup>10</sup>CDC. (2021, October 6). *Antibiotic Use Questions and Answers* | Antibiotic Use. CDC. Retrieved July 25, 2022, from <https://www.cdc.gov/antibiotic-use/q-a.html>

<sup>11</sup>Wischmeyer, P. E., McDonald, D., & Knight, R. (2016, October 13). *Role of the microbiome, probiotics, and 'dysbiosis therapy' in critical illness*. NCBI. Retrieved July 25, 2022, from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5065053/>

<sup>12</sup>CDC. (2021, August 23). *Antibiotic Resistance Questions and Answers* | Antibiotic Use. CDC. Retrieved July 25, 2022, from <https://www.cdc.gov/antibiotic-use/antibiotic-resistance.html>

<sup>13</sup>Cleveland Clinic. (2020, March 9). *Probiotics: What is it, Benefits, Side Effects, Food & Types*.

Cleveland Clinic. Retrieved July 28, 2022, from <https://my.clevelandclinic.org/health/articles/14598-probiotics>

<sup>14</sup>Sherman, E. (2020, June 30). *Traditional Fermented Foods from Around the World*. Matador Network.

Retrieved July 25, 2022, from <https://matadornetwork.com/read/traditional-fermented-foods/>

<sup>15</sup>Walker, L. M., & Silliker, M. E. (2022). Terminal Restriction Fragment Length Polymorphism.

*ScienceDirect*.

<sup>16</sup>16S rRNA, One of the Most Important rRNAs. (2018, October 17). CD Genomics. Retrieved July 26, 2022, from <https://www.cd-genomics.com/blog/16s-rna-one-of-the-most-important-rrnas/>

<sup>17</sup>Green, A., Matich, M., Reber, M., Tang, S., Wang, D., & Zhang, K. (n.d.). Computational Biology: Identifying Lactic Acid Bacteria Through Comparison of 16S Genetic Sequence. *The PGSS Journal*.

<sup>18</sup>Ultraviolet Spectrophotometry. (2019). *ScienceDirect*.

<https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/ultraviolet-spectrophotometry>

<sup>19</sup>Ganske, F. (2014, 12). *UV absorbance DNA quantitation | BMG LABTECH*. BMG Labtech. Retrieved July 25, 2022, from <https://www.bmglabtech.com/en/application-notes/uv-absorbance-dna-quantitation/>

<sup>20</sup>restriction enzyme | Definition, Function, & Types | Britannica. (n.d.). Encyclopedia Britannica.

Retrieved July 26, 2022, from <https://www.britannica.com/science/restriction-enzyme>

# **Computational Biology: Identifying Lactic Acid Bacteria Through Comparison of 16S Genetic Sequence**

Aidan Green, Mason Matich, Malina Reber, Siming Tang, Daniel Wang, Kurtis Zhang

## **Abstract**

Lactic acid bacteria (LAB) microorganisms are widely found in various fermented foods and plants, possessing the ability to act as probiotics that are highly beneficial to the human body. Thus, recent research of LAB microorganisms yields new discoveries, though many limitations to identification still exist. Our project focused on developing an algorithm using computational tools that identifies LAB based on the 16S rRNA subregion of bacterial genes. Open source packages like Biopython, Fasta-reader and EMBOSS and the open access data source Ribosomal Data Project (RDP) allowed our team to build a comparison pipeline that could differentiate between bacterial strains. Sorting algorithms and statistical analysis of specific DNA fragment cuts helped identify possible matches between input experimental LAB data and the RDP database data. Our results indicate that our methodology was largely successful in filtering the database to a smaller list that contains the correct bacteria and is open to future application for other bacteria.

## **I. Introduction**

### **A. Background**

Bacteria are microscopic organisms with complex and far-reaching impacts on Earth's ecological systems. Thriving in atmospheric, aquatic, subterranean, and organic environments (i.e. inside an organism's body), microbes can be found in locations ranging from the human gut and plants to hydrothermal vents and cirrus clouds. Their influences range accordingly, from intestinal processes to seemingly abiotic occurrences in nature such as the cyclic path of nitrogen<sup>4</sup>. Although microbes generally are simpler structures compared to multicellular organisms, their influence is imperative to biological and ecological functions, as they pervade most aspects of larger organisms and the environments in which they depend on, and are largely responsible for essential processes of life cycles such as decomposition and disease causation and prevention.

Applications of microbial abilities are extremely beneficial for human life in particular, as their decaying ability can help balance ecosystems, and they can be used in the production of probiotics, insulin, and fermented nutritional products<sup>2</sup>. Some examples of their wide-spread application include their ability to consume and filter oil spills, function in modern biotech drugs to mitigate disease, fuel nutritional campaigns, balance ecosystems, and consume plastic. Microorganisms are so proliferative in human life, in fact, that it is estimated that there are more bacteria in the human body than human cells<sup>4</sup>. The identification and understanding of microbes and their functions is essential for utilizing their full potential to enhance the environmental, nutritional, and preventative aspects of human health. Our project is mainly focused on the identification of a microorganism most people consume daily: lactic acid bacteria, or LAB.

Lactic Acid Bacteria (LAB) are of the *Clostridium* branch of the Gram positive bacteria, meaning that they primarily grow in clusters or chains<sup>26</sup>. They have many applications, including the fermentation of food and the development of probiotics which are used to improve disease resistance via gut health<sup>16</sup>.

## 1. Fermented Food

The oldest and most wide-spread use of LAB is the fermentation process. Some examples of its benefits to the food industry include its ability to improve yogurt textures through its consumption of carbs and release of exopolysaccharides, as well as its ability to accelerate the maturation of cheese through the use of adjunct cultures that make up for the loss of biodiversity in the pasteurization process<sup>2</sup>. Besides convenience and dietary preference, fermentation is also used to package food because the fermentation process makes food more nutritious, less perishable, and safer to consume<sup>16</sup>. Some probiotic LAB are added purely for nutritional benefit, for their release of antifungal elements and bacteriocins, peptide toxins that inhibit the growth of certain bacteria by damaging cell membranes<sup>7</sup>. LAB is also able to produce hydrogen peroxide through oxidation of reduced NADH which reacts with gaseous oxygen, creating an antibiotic effect for organisms that may cause spoilage and food-borne illness. LAB can also lower the pH of an organism's systems to that of an uninhabitable environment for other competing organisms that could include harmful bacteria<sup>16</sup>.

Today, LAB are utilized in the manufacturing process as a starter culture for fermented foods such as cheese, yogurt, fermented meats, pickled vegetables, etc. As a process that has low demands in terms of cost and energy, LAB is a desirable method of food preservation for locations in which canned and frozen foods are unavailable or too expensive. By identifying specific LAB in fermented products, we can differentiate between product labels that include the word "probiotic" as a buzzword, and products that sufficiently provide nutritional benefits through the use of particular lactobacillus strains.

## 2. Probiotics

Within the past two decades, the use of probiotics has recently been recognized as a potential supplement to human health. LAB are one of the most commonly used bacteria in probiotic foods because they are able to protect a host (human in application) against harmful microorganisms, allowing the host to maintain a well-balanced microbial environment within their digestive system. Antibiotics are found increasingly ineffective against certain strains of bacteria, and so the bacteriocins of lactic acid bacteria can serve as a countermeasure. Bioengineered with low toxicity to human health, they are potent against some antibiotic-resistant bacteria, and they can have wide ranges of activity levels which gives scientists options for their application. In addition, recent research studies have found promising results regarding the positive correlation between probiotic immunity and mental health as well as reduced metabolic disorders<sup>32</sup>. This has led to an increased interest in the possibilities for probiotic enrichment of the human microbiome. It is expected that these health benefits will lead to greater growth in the probiotic industry. The global market in 2021 was around 58.2 Billion USD, with the US market growing from 8.8 Billion USD to 9.3 Billion USD from 2020 to 2021<sup>21</sup>.

## B. Purpose

The goal of our algorithm is to increase the efficiency and accessibility of bacteria identification processes by computationally augmenting ARDRA to achieve the performance of the more expensive TRFLIP. Faster identification will allow individuals outside of large organizations who cannot purchase identification equipment to verify food and drug ingredients. In order to design an algorithm that analyzes experimental data from gel electrophoresis and offers a range of possible bacteria identities, we decided to narrow our focus to the identification of a bacteria that affects our microbiome, the lactobacillus found in cultured dairy products such as buttermilk and kefir. Probiotic lactic bacteria are often listed in vague detail on product food labels as a nutritional buzzword either abbreviating the bacteria's binomial nomenclature or excluding

a description of the lactobacteria altogether. Our lactic acid bacteria identification system aims to be an approachable and accessible method for identifying microbes in daily life, such as the bacteria in our yogurt.

## C. Methods of Bacterial Identification

Current methods of identification are expensive processes involving complex machinery costly in time, effort, and material resources. Almost all methods first involve the culturing, or growing, of bacteria in order to gather an adequate sample for testing and identification.

### 1. FISH

One such method is Fluorescence *in Situ* Hybridization. This process locates specific genes within chromosomes through the use of a probe which localizes a DNA or RNA sequence in a tissue sample. The probe is then denatured using chemicals or heat, washing off excess RNA content. The probe is also tagged with fluorescent dye that can be used to identify bacteria and tissue under a microscope. Fluorescence *In situ* Hybridization is greatly influenced by the ability or limits of fluorescence microscopy and has its limitations such as the inherently short-lived nature of the fluorescent signal<sup>13</sup>.

### 2. ARDRA

Amplified Ribosomal DNA Restriction Analysis is a modified version of PCR ribotyping. DNA that is taken from a sample is subjected to PCR using primers at the end of the conserved regions of the 16S rDNA gene. The amplified regions are cut by tetra cutter restriction enzymes, producing restricted fragments which are then put on agarose or gels for analysis. The restriction fragment patterns are able to be used to further identify diversity and similarities of bacteria based on the proportion of fragments. Though useful for identification of bacteria in pure cultures or in microbial communities, ARDRA is mainly only useful for providing genus and species differentiation in bacteria because the first DNA fragment is unknown<sup>13</sup>. The computational model described here aims to augment this deficiency by comparing the ARDRA results to a preprocessed database.

### 3. T-RFLP

Terminal Restriction Fragment Length Polymorphism is another PCR based genetic fingerprinting technique for studying microbial communities in the 16S rRNA region. T-RFLP is useful for estimating the abundance of dominant bacterial species. Primer pairs in the 16S region amplify the DNA coding for the region through PCR. The PCR product is then digested by restriction enzymes and terminal restriction fragments (TRF) are separated with a DNA analyzer and labeled with fluorescent dye. The species composition of the bacteria can then be determined by measuring the level of intensity of fluorescence emission. Diversity of bacterial communities can also be estimated by analyzing the pattern in TRF. However, sample variation in DNA during extraction that is caused by partial cell lysis, PCR bias and uneven distribution of marker genes during amplification, and poor phylogenetic resolution make T-RFLP a potential pitfall<sup>20</sup>.

### 4. MALDI-TOF MS

Matrix-assisted laser desorption/ionization-time of flight Mass Spectrometry takes a microbial sample for analysis by coating and mixing the sample with an organic compound called a matrix. The matrix then crystallizes, causing the sample within the matrix to co-crystallize, preparing it to be ionized by a laser beam.

The laser beam creates protonated ions that can be analyzed: specifically, the m/z ratio of an ion is measured by determining the time required for the ion to travel the length of the flight tube is measured, and based on time of flight analyzers (TOF), a peptide mass fingerprint (PMF) can be generated. The PMF is then used for comparison of the unknown bacteria or microorganism against PMFs stored in a known proteome database. Additionally, the MS spectrum of microbial isolates can be compared, virtually allowing the genus and species of a microorganism to be established. However, limitations including reliability on a constant power supply and heavy dependence on a vast database that contains a sufficient number of organisms are important in order for MALDI-TOF MS to be successful<sup>24</sup>.

## 5. Real Time PCR

The technique of real-time PCR allows for data collection of the PCR process as it occurs, hence the name "real-time". This combines the steps of amplification and detection<sup>30</sup>. The data is able to be collected in real time by looking at the fluorescence as the process occurs; the intensity of the fluorescence represents the amount of DNA amplicons present<sup>15</sup>. By using the change of fluorescence over time, it can be used to calculate the amount of amplicons created in every cycle. This technique allows for the quantitative analysis of precise changes in gene expression. It can also be used for detection of genetically modified organisms and cancer phenotyping.

## 6. DNA Microarrays

DNA microarrays are a tool that can be used to find whether or not there is a mutation in a gene. In DNA microarrays, an array of thousands to millions of known nucleic acid fragments are bound to a solid surface known as a "chip". This chip is then used to see if these fragments are able to bind to a "normal" complementary sequence or a mutated complementary sequence. Whichever sequence these fragments bind to will show whether or not there was a mutation in the DNA. The application of this technique is for large-scale studies on changes in genes. It can be used to look at the extent to which certain genes are turned on and off in cells<sup>8</sup>.

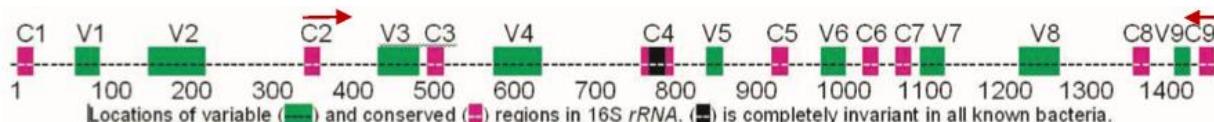
## 7. RAPD fingerprinting

Randomly Amplified Polymorphic DNA fingerprinting is a modified version of PCR that uses a single arbitrary primer to amplify a number of discrete fragments from a given DNA template. Using gel electrophoresis will create a unique "fingerprint". This can help show alterations in DNA sequences as small as one base. This technique allows for the potential in identifying mutations in bacterial or human cells. Though, RAPD fingerprinting does have its limitations as operations must be conducted under controlled conditions, and the reproductive accuracy of RAPD patterns oftentimes varies<sup>30</sup>.

## D. 16S Subregion Analysis

The 16S as shown in figure 1 is a highly conserved subregion of the 30S subunit and is used for the gel electrophoresis process. What makes it attractive for identification is that each species will have a slightly different 16S region, acting as a metaphorical fingerprint for each bacteria. Using primers, this region is selectively amplified using PCR, the process of generating DNA copies through heating and cooling cycles in which the polymerase chain reaction takes place<sup>8</sup>. The copies are then cut using the Hpy188I and MseI restriction enzymes, which have been experimentally shown to be the optimal enzymes for identifying bacteria in beverages<sup>33</sup>. This process yields the overall 16S length of the unknown bacteria as well as the

lengths of the corresponding DNA fragments. These unique sets of fragments are then analyzed to determine the identity of the bacteria.



1 AAATTG**AAGAGTTTGATCATGGCTCAGATTGAAC**GCTGGCGGCAGGCCTAACACATGCAAGTCGAAC**GG**  
 70 **TAACAGGAAGAACGTTGCTCTTGCTGACG**AGTGGCGGACGGGTGAGTAATGTCTGGGAAACTGCC**TGA**  
 139 **TGGAGGGGGATAACTACTGGAAACGGTAGCTAATACCGCATAACGTCGCAAGACCAAAGAGGGGGACCT**  
 208 **TCGGGCCTCTTGCCATCG**GATGTGCCAGATGGGATTAGCTAGTAGGTGGGTAACGGCTCACCTAGGC  
 277 GACGATCCCTAGCTGGTCTGAGAGGATGACCAGCCACACTGGAACGTGAGACACGGTCCAGAC**TCCTACG**  
 346 **GGAGGCAGCAGT**GGGAAATTGCACAATGGCGCAAGCCTGATGCAGCCATGCCGCGTGTATGAAGAA  
 415 GGCCTTCGGGTTGTAAAGTACTTT**CAGCGGGAGGAAGGGAGTAAAGTTAATACCTTGCTCATTGACG**  
 484 **TTACCCGCAGAA**GAAGCACCGGCTAACTCC**GTGCCAGCAGCCGCGGTAA**...901nts...**GTAGGTAG**  
 1442 **CTTAACCTTCGGAGGGCGCTTA**CCACTTGTGATTGACTGGGTG**AAGTCGTAACAAGGTAACCG**  
 1511 **TAGGGGAACCTGCGGTTGGATCACCTCCTTA**

Figure 1: 16S subregion

## E. Gel Electrophoresis

Gel electrophoresis is a method of visualizing the length of DNA sequences by separating them from each other based on weight, which directly corresponds to length. A gel made of agarose and buffer is used for separation as its structure creates small pores which the DNA can flow through. The gel has holes, or wells, at one end which samples can be loaded into. This gel is placed in a gel box filled with a buffer containing salt, so it is electrically conductive. The gel box is connected to a power source with positive at one end and negative at the other. To run a gel, samples stained with a DNA binding dye are loaded into the wells along with at least one DNA ladder. DNA ladders are used to measure the length of the samples after the gel is completed. Since DNA is negatively charged, when a voltage is applied across the gel the DNA will move towards the positive pole. Shorter pieces of DNA will travel faster through the gel than longer strands, over time causing separation. DNA sequences of the same length form bands which show up as bright lines. The brighter the band, the higher the concentration of DNA.

Gel electrophoresis was chosen for this project because it is accessible to wider populations including intermediate schools and field labs where equipment is more basic. Our team combined a more accessible biology process, gel electrophoresis, with our own computational algorithm that identifies lactic bacteria based on the length of its fragments within the 16S sequence of its DNA.

## II. Methodology

### A. Open Source Software Used

#### 1. Biopython

Biopython<sup>5</sup> is a package offered by Python that includes a set of libraries and modules used for bioinformatic data analysis<sup>3</sup>. Specifically, we used the Bio.Seq and the Bio.Restriction modules to find the lengths of sequence fragments.

#### 2. Fasta-reader

A FASTA file is a text-based file format for representing DNA or peptide sequences with single-letter codes denoting the nucleobases of DNA: adenine, cytosine, thymine, and guanine into A, C, T, and G respectively. The fasta-reader package provides tools to read FASTA files as lists; we used this feature to read the DNA sequences imported from the RDP.

#### 3. EMBOSS

EMBOSS is a set of open source software for computational analysis of DNA. The primersearch program was used for the extraction of lactic acid bacteria from 16S regions from the raw database to cut down on unnecessary processing.

### B. Open Access Databases

#### 1. Ribosomal Database Project (RDP)

The Ribosomal Database Project (RDP)<sup>6</sup> is a database of over 3 million Bacterial and Archaeal 16S rRNA sequences and 125 thousand Fungal 28S rRNA sequences. 189 thousand 16S rRNA sequences were taken from the database in the format of a FASTA file as sample data to be used in comparison to user input data<sup>16</sup>.

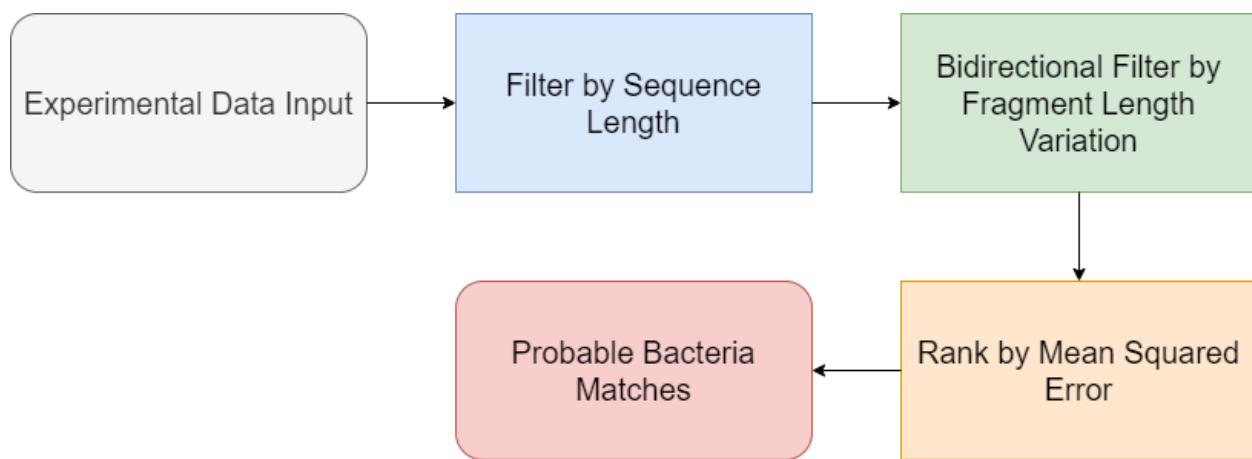
### C. Sorting and Formatting of the Database

The RDP contains approximately 3 million species of bacteria, but only bacteria in the order Lactobacillales are relevant to this project. To cut down the size of the database, all non-lactobacillales were removed. Any experimental 16S samples will be generated by PCR using primers, meaning the entire 16S region will not always be included. To account for this, EMBOSS primersearch was employed to perform virtual PCR on the Lactobacillales bacteria. Finally, a virtual restriction digest using Msel and Hpy188I was performed on the virtual PCR product to bring the database fully in line with the experimental procedure. The formatted database was then exported into a CSV file, with each bacteria having its own line as shown in figure 2 sorted in the order of: definition, sequence length, Msel cut lengths, and finally the Hpy188I cut lengths.

```
{
'defline': 'S003447426 uncultured bacterium;
LIB090_013_1165\tLineage=Root;rootrank;Bacteria;domain;Firmicutes;phylum;Bacilli;class;Lactobacillales;order;Aerococcaceae;family;Abiotrophia;genus',
'seq_length': 201,
'fragments_MseI': [201],
'fragments_Hpy188I': [124, 77]
}
```

**Figure 2: Example of finalized database entry**

#### D. Filtering the Data and Finding Matches



**Figure 3: Flowchart of Analysis Pipeline**

Our analysis pipeline as shown in figure 3, referred to as the Computational Modeling Method (CMM), uses the transformed database csv and then further formats the data for efficient analysis, and filters bacteria for possible matches to the experimental data. Filters are set to compare available bacteria data and user input data on the total Sequence Length of the 16S region and the resulting fragment lengths after the application of the enzymes Msel and Hpy1881.

The final, filtered list is then stored in a .txt file for further filter iterations.

##### 1. Inputting the Sample Data for Method Refinement

Lists in the CMM eliminate sample sequence data that correspond to incorrect matches to the input data. Sequence fragments greater than the total length of a sample bacteria sequence are removed because that represents an undigested DNA sample, not a DNA fragment.

Our dictionary of bacteria fragments is also iterated with a function that ignores fragments which are less than or equal to 100 base pairs(bp) when comparing database values to experimental values. These fragments less than 100 bp in length are ignored because gel electrophoresis results of this size cannot

be easily read and reported by program users, and this discrepancy can lead to false identification or missed potential matches.

## 2. Filtering via Sequence Length Comparison

The sequence filtering process as shown in figure 4 is used to compare the lengths of the experimental (input) and database (RDP) sequences. If the absolute value of the difference between the two sequences is less than the preset buffer value, in our case 100bp<sup>34</sup> as that is the resolution possible with gel electrophoresis<sup>1</sup>, then the database bacteria is added to a new list. If the absolute value is greater than the buffer value, then the bacteria is not added. Once the function is finished running, it will return the new filtered list of the remaining candidate bacteria. This method ensures that the experimental sequence and any potential database matches have roughly the same undigested sequence length.

### Example List 1

Total Exp Sequence Length: 561

Total DB Sequence Length: 710

$$| 561 - 710 | > 100 \longrightarrow \boxed{\text{Not a Match}}$$

### Example List 2

Total Exp Sequence Length: 561

Total DB Sequence Length: 600

$$| 561 - 600 | <= 100 \longrightarrow \boxed{\text{Possible Match}}$$

Figure 4: Sequence Length Filtering

## 3. Filtering via Comparing Fragment Lengths

In this step, the team essentially needed to filter through the DNA fragments to narrow down the search for possible matches to the experimental samples. The pipeline performed a bi directional analysis of the fragment length variations and found corresponding matches from the database to the experimental samples, and vice versa. Additionally, the model was able to test for 2 things about possible bacteria matches: The length of each DNA fragment and if the relative size of the experimental and database

fragments were in range of each other. So, by using a 100 base pair range, the team first compared the database DNA fragment values to the experimental DNA fragment values, and if they were in the 100 bp range of each other, the test moved onto the next DNA fragment in the database and so on. Furthermore, any fragments in the database that were less than 100 bp were considered too small and not considered in deciding a match. This is because many unaccounted-for factors could have affected the gel's ability to record down smaller fragment sizes, even though they could be matched to the database, so they should not have been ruled out.

Then, in order to account for the experimental DNA fragment sizes, the test was run backwards, comparing the experimental fragment values to the database, allowing the team to account for any unmatched fragment in either direction. Therefore, the test was able to account for all the different fragments within range of each other, not omitting any possible matches. An example, as seen in Figure 5: DNA Fragment Filtering (Numerical), indicates that since the 500 base pair DNA fragment is not within 100 bp range of the other database fragments, the program could conclude that that database sample was not a viable match for the team's experimental sample. Thus, running a bi directional test was crucial; If the test was not run backwards, the 500 bp DNA fragment in Test 1 would not have been considered in determining a match. However, Test 2 helped to check Test 1 by concluding that the database sample was not a good match to the experimental sample. Therefore, the bi directional fragment comparing functions were able to be run through all the database samples to narrow down matches with experimental samples received from the biology team.

The fragment comparison procedure as shown in Figure 5 and 6 will do a bi-directional analysis of the data.

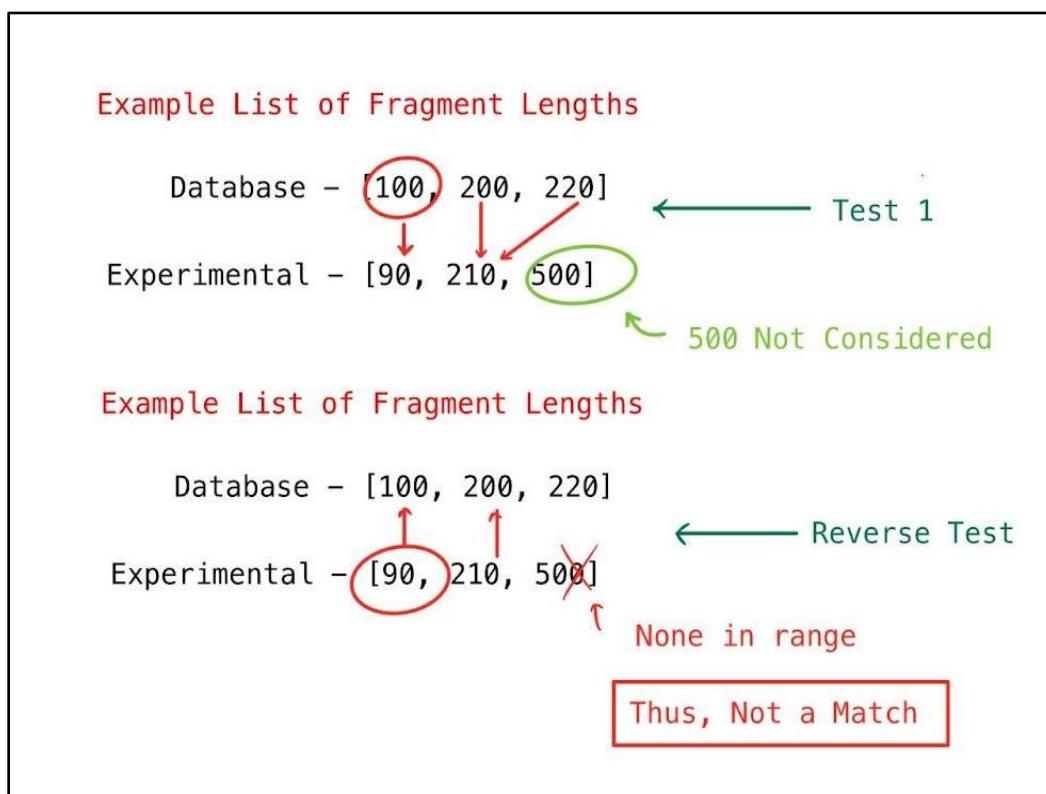


Figure 5: DNA Fragment Filtering (Numerical)

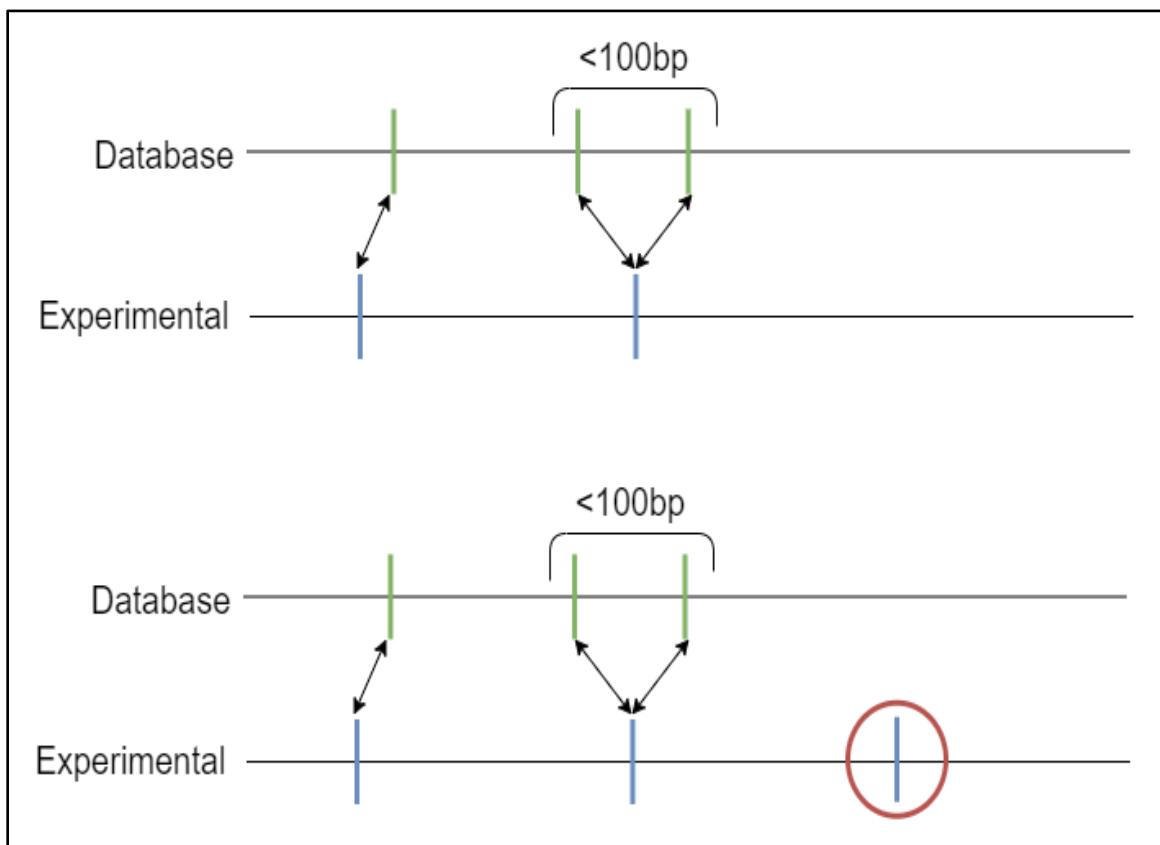


Figure 6: DNA Fragment Filtering (Visual)

#### 4. Ordering the Matches through Ranking

Once the list of potential matches is completely filtered, as described above, The data are then sorted by the distance between experimental cut values and the database cut values and vice versa. To accomplish this, each experimental cut length is compared to each database bacterial cut length of the same enzyme in order to find the smallest absolute difference, which is then used to calculate Mean Squared Error (MSE), a commonly used calculation to determine error value. Once the MSE has been calculated for each enzyme, it is then repeated in the reverse direction, comparing the database bacterial cut lengths to the experimental cut lengths. These values are then summed up and averaged, and that value is then appended to each of the bacteria as an additional data point. Finally, the bacteria in the filtered list are then ranked according to their MSE which is shown in figure 7, from least to greatest. The function then returns the database with a MSE value appending to each entry.

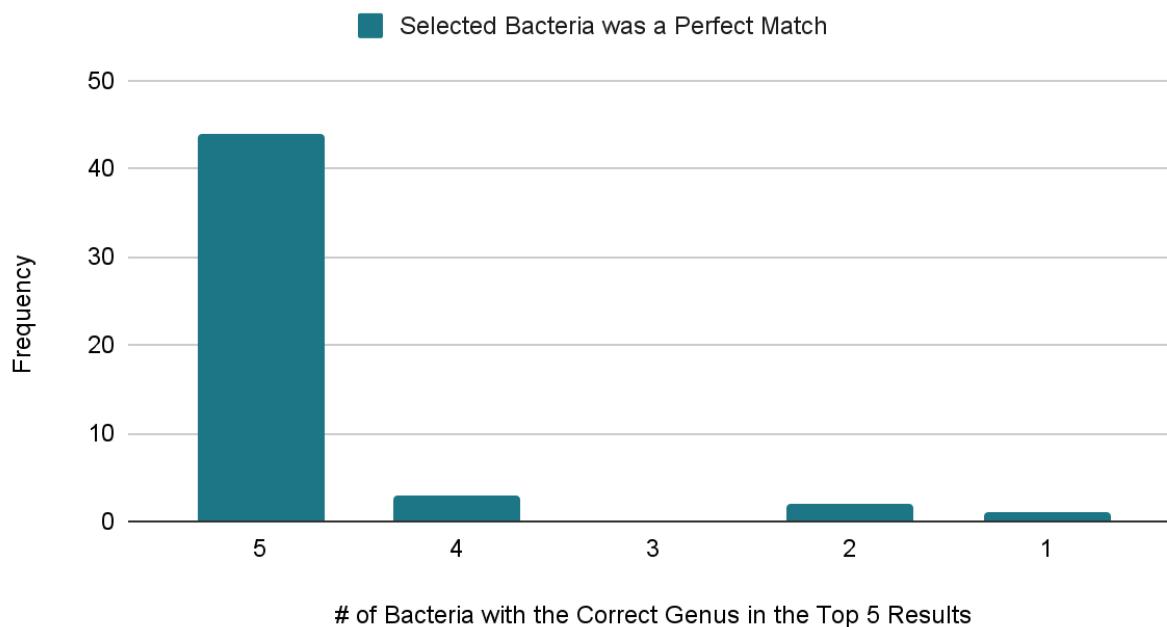
$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$$

Figure 7: MSE formula

### III. Results

#### A. Validation of Our Approach

##### Ideal Sample Data



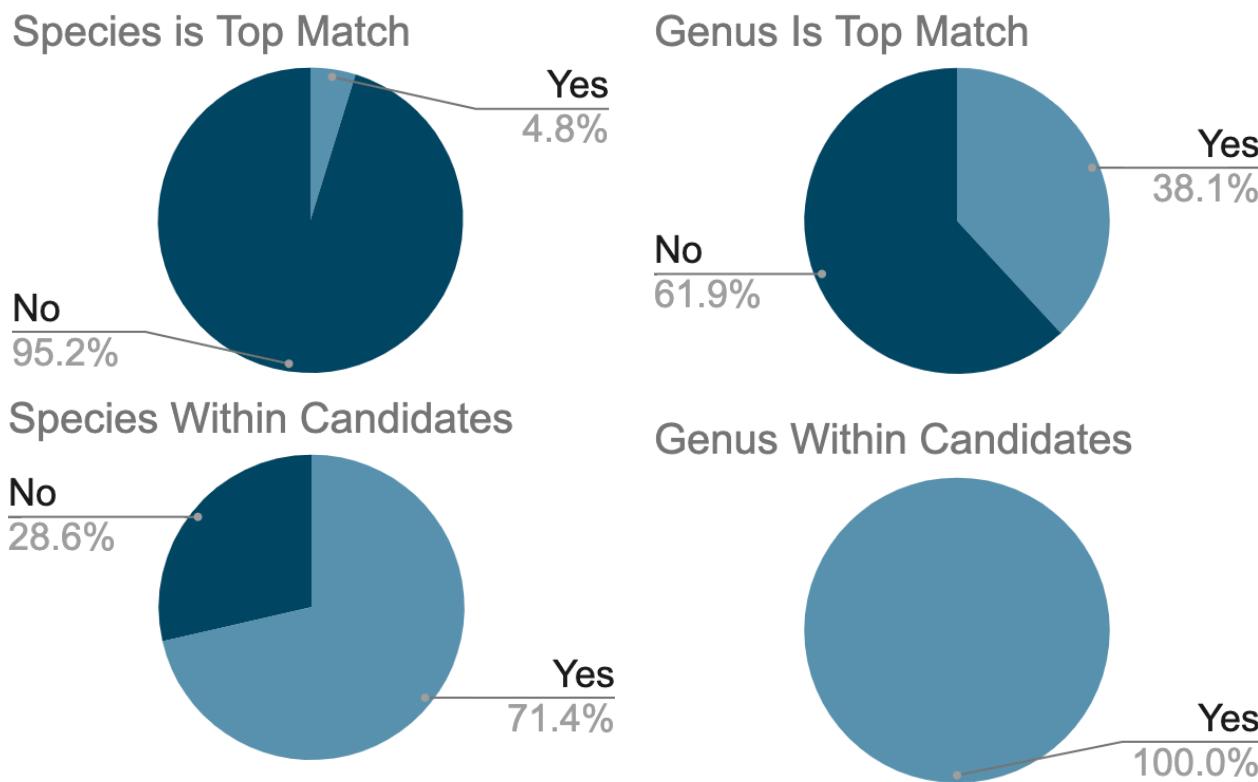
**Figure 8: Results Using Sample Database Data** The Figure represents the frequency of the correct bacteria as the top result of the algorithm, categorized by the number of bacteria with the correct genus in the top 5 results.

In order to validate the matching and ranking algorithm, we randomly sampled 50 bacteria from our locally stored, pre-digested database as shown in figure 8. These bacterial samples were then entered into the algorithm and the final sorted list for each sample was placed into separate .TXT files. In every iteration of this validation, the selected bacteria was correctly identified as a perfect match. In 44 out of the 50 tested samples, the following 4 top results were of the correct genus. In 3 of the tested samples, 4 of the top 5 results were of the correct genus. In 2 of the tested samples, 2 of the top 5 results were of the same genus. In only 1 of the tested samples, only the actual bacteria was the only one with the correct genus in the top 5 results. This information indicates that, when given ideal data, the algorithm is capable of not only identifying the correct bacteria, but also ranking similar bacteria highly.

## B. Confirmation with Experimental Sample Data

**Table 1: Results of Statistical Model on Experimental Data**

Cultures	Correct Species in Candidates	Correct Genus As Top Choice	Correct Genus in Candidates	Candidates in Filtered Set
<b>B1</b>	No	No	Yes	14
<b>B2</b>	Yes	No	Yes	166
<b>B3a</b>	Yes	No	Yes	70
<b>B3b</b>	Yes	No	Yes	84
<b>B4a</b>	Yes	No	Yes	87
<b>B4b</b>	Yes	No	Yes	142
<b>B5a</b>	Yes	No	Yes	70
<b>B5b</b>	Yes	No	Yes	143
<b>B6a</b>	Yes	No	Yes	82
<b>B6b</b>	Yes	No	Yes	141
<b>B7</b>	Yes	No	Yes	234
<b>B8</b>	Yes	No	Yes	138
<b>K1</b>	No	No	Yes	69
<b>K2</b>	No	Yes	Yes	43
<b>K3a</b>	Yes	Yes	Yes	141
<b>K3b</b>	Yes	Yes	Yes	243
<b>K4a</b>	Yes	Yes	Yes	195
<b>K4b</b>	No	Yes	Yes	45
<b>K5</b>	No	Yes	Yes	95
<b>K6</b>	No	Yes	Yes	37
<b>Control</b>	Yes	Yes	Yes	132



**Figure 9: Results Using Experimentally Obtained Data**

The pie charts represent the percentage of the experimental results where the correct species was within the filtered candidates, the genus was identified as the top match, and the correct genus was within the filtered candidates.

The statistical model was run on the experimental data derived from Barlow et al. as shown in Figure 9. The model's accuracy was tested on its ability to identify the sample's species in the genus as the top choice or within the filtered set. As seen in Table 1, the bacteria candidates were narrowed from the original 102,000 to an average of 112.9 bacteria.

#### IV. Discussion

Given the match results of the CMM algorithm when inputted with 50 randomly sampled RDP database bacteria, as described in Table 1, it can be concluded that when ideal data is provided, the algorithm performs as expected. That is to say, when the sample data provided to the algorithm is of the very highest quality, the algorithm is able to provide results with exceptional accuracy.

It is more difficult to reach a simple conclusion based on the algorithm's results. In the majority of cases, when experimental data is input into the algorithm, the species is in the highly filtered, final list of potential matches. As mentioned in Figure 9, that highly filtered list is substantially refined as compared to the original, non-filtered data set.

For future development of this algorithm there are several adjustments that could be implemented to our methodology. Firstly, additional enzymes could be used in the gel electrophoresis and the virtual PCR of our reference data set in addition to Msel and Hpy188I. Secondly, the ranking procedure can be optimized to improve the ranking of the actual species, though improvements would likely be small due to the inherent limitation of the resolution of a gel. This can be achieved by refining our algorithm through increased testing and consideration of a larger data set. We used these enzymes to filter our matches based on a study solely relating to lactic acid bacteria in beverages<sup>33</sup>. Therefore, it is possible that more enzymes may be required to filter matches in dairy products. If we were to implement a third enzyme, it could further filter our potential matches by cross checking more 16S fragments.

## V. Acknowledgements

We would like to express our gratitude and thanks to team project advisors Dr. Natalie McGuier and Mr. Andrew McGuier for their helpful critiques and guidance throughout the research process and project presentation process. We would also like to express our thanks to teaching assistants Michael Sobol and Shiyang Tian for their encouragement and support. Lastly, a special thanks to Dr. Barry Luokkala, Melissa Lessure, and the PGSS Faculty and Staff for organizing and running the PGSS experience. This opportunity would not be possible without the efforts of the PGSS Corporate Sponsors and the PGSS Campaign, Inc. to keep the program provided for, so that students can have access to research opportunities and professional guidance. Finally, we are grateful to Carnegie Mellon University and the Mellon College of Science for hosting PGSS and providing access to their facilities. Completion of this project was made possible through PGSS's efforts to foster a thought-provoking learning environment for aspiring scientists. We cannot thank you enough.

## VI. References

- <sup>1</sup>Baron, S. (Ed.). (1996). Medical Microbiology. (4th ed.). University of Texas Medical Branch at Galveston.
- <sup>2</sup>Bintsis, T. (2018). Lactic acid bacteria: Their applications in foods. *Bacteriology & Mycology: Open Access*, 6(2).
- <sup>3</sup>Bio Rad. (2022). *What is real-Time PCR (qPCR)?* [Advertisement]. Retrieved July 28, 2022, from <https://www.bio-rad.com/en-us/applications-technologies>
- <sup>4</sup>Blaser, M. J. (2014). Missing microbes: How the overuse of antibiotics is fueling our modern plagues. First edition. New York: Henry Holt and Company.
- <sup>5</sup>Cock PA, Antao T, Chang JT, Chapman BA, Cox CJ, Dalke A, Friedberg I, Hamelryck T, Kauff F, Wilczynski B and de Hoon MJL (2009). Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics*, 25, 1422-1423
- <sup>6</sup>Cole, J. R., Q. Wang, J. A. Fish, B. Chai, D. M. McGarrell, Y. Sun, C. T. Brown, A. Porras-Alfaro, C. R. Kuske, and J. M. Tiedje. (2014). Ribosomal database project: Data and tools for high throughput rRNA analysis. *Nucl. Acids Res.* 42(Database issue):D633-D642; doi: 10.1093/nar/gkt1244 [PMID: 24288368]
- <sup>7</sup>Cotter, P. D., Ross, R. P., & Hill, C. (2012). Bacteriocins — a viable alternative to antibiotics? *Nature Reviews Microbiology*, 11(2), 95-105. <https://doi.org/10.1038/nrmicro2937>
- <sup>8</sup>DNA microarray technology fact sheet (About Genomics). (2020, August 15).
- <sup>9</sup>Ercolini, D., & Cocolin, L. (2014). Identification methods | culture-independent techniques. In C. A. Batt & M. L. Tortoello (Eds.), *Encyclopedia of food microbiology* (2nd ed., pp. 259-266). Academic Press. <https://doi.org/10.1016/B978-0-12-384730-0.00438-9>
- <sup>10</sup>Feucherolles, M., Poppert, S., Utzinger, J., & Becker, S. L. (2019). MALDI-TOF mass spectrometry as a diagnostic tool in human and veterinary helminthology: A systematic review. *Parasites & Vectors*. <https://doi.org/10.1186/s13071-019-3493-9>

- <sup>11</sup>Florou-Paneri, P., Christaki, E., & Bonos, E. (2013). Lactic Acid Bacteria as Source of Functional Ingredients. In (Ed.), Lactic Acid Bacteria - R & D for Food, Health and Livestock Purposes. IntechOpen. <https://doi.org/10.5772/47766>
- <sup>12</sup>Govindarajan, R., Duraiyan, J., Kaliyappan, K., & Palanisamy, M. (2012). Microarray and its applications. *Journal of pharmacy & bioallied sciences*, 4(Suppl 2), S310–S312. <https://doi.org/10.4103/0975-7406.100283>
- <sup>13</sup>Karki, G. (n.d.). Fluorescent in-situ hybridization: Advantages, limitations and preparation of probe. In *Online Biology Notes: Genetics*. <https://www.onlinebiologynotes.com> (Original work published 2020)
- <sup>14</sup>Kashyap, S. K., Maherchandani, S., & Kumar, N. (2014). Chapter 18 - ribotyping: A tool for molecular taxonomy. In A. S. Verma & A. Singh (Eds.), *Animal biotechnology: Models in discovery and translation* (pp. 327-344). Academic Press. <https://doi.org/10.1016/B978-0-12-416002-6.00018-3>
- <sup>15</sup>Kralik, P., & Ricchi, M. (2017). A basic guide to real time PCR in microbial diagnostics: Definitions, parameters, and everything. *Frontiers in Microbiology*.
- <sup>16</sup>Mohania, D., Nagpal, R., Kumar, M., Bhardwaj, A., Yadav, M., Jain, S., Marotta, F., Singh, V., Parkash, O., & Yadav, H. (2008). Molecular approaches for identification and characterization of lactic acid bacteria. *Journal of Digestive Diseases*, 9(4), 190-198. <https://doi.org/10.1111/j.1751-2980.2008.00345.x>
- <sup>17</sup>Narvhus, J. A., & Axelsson, L. (2003). Lactic Acid Bacteria [Introduction]. In B. Caballero (Ed.), *Encyclopedia of food sciences and nutrition* (2nd ed.). Elsevier.
- <sup>18</sup>Neelja, S., Manish, K., K. K. P., & S. V. J. (2015). MALDI-TOF mass spectrometry: An emerging technology for microbial identification and diagnosis. *Frontiers in Microbiology*, 6. <https://doi.org/10.3389/fmicb.2015.00791>
- <sup>19</sup>Pelczar, M. J., & Pelczar, R. M. (2022, May 1). *Types of microorganisms*. Encyclopædia Britannica. Retrieved July 28, 2022, from <https://www.britannica.com/science/microbiology/Types-of-microorganisms>
- <sup>20</sup>Prakash, O., Pandey, P. K., Kulkarni, G. J., Mahale, K. N., & Shouche, Y. S. (2014). Technicalities and Glitches of Terminal Restriction Fragment Length Polymorphism (T-RFLP). *Indian journal of microbiology*, 54(3), 255–261. <https://doi.org/10.1007/s12088-014-0461-0>
- <sup>21</sup>Probiotics market size (Nutraceuticals & Functional Foods). (2017).
- <sup>22</sup>Python Software Foundation. (2022, July 25). CSV - csv file reading and writing¶. csv - CSV File Reading and Writing - Python 3.10.5 documentation. Retrieved July 25, 2022, from <https://docs.python.org/3/library/csv.html>
- <sup>23</sup>qPCR probes (Oligonucleotides). (n.d.). Bio-Synthesis. <https://www.biosyn.com/qpcrprobes.aspx>
- <sup>24</sup>Rychert, J. (2019). Benefits and limitations of MALDI-TOF mass spectrometry for the identification of microorganisms. *Journal of Infectiology and Epidemiology*. <https://doi.org/10.29245/2689-9981/2019/4.1142>
- <sup>25</sup>Sandle, T. (Ed.). (2016). *Pharmaceutical microbiology*. Woodhead publishing. <https://doi.org/10.1016/B978-0-08-100022-9.00009-8>
- <sup>26</sup>Sizar, O., & Unakal, C. G. (2022). *Gram positive bacteria*.
- <sup>27</sup>Solomon, B. (2022, July 28). *In situ hybridization* [Audio recording]. National Human Genome Research Institute. <https://www.genome.gov/genetics-glossary/in-situ-hybridization>
- <sup>28</sup>Van Rossum, G., & Drake, F. L. (2009). Python 3 Reference Manual. Scotts Valley, CA: CreateSpace.
- <sup>29</sup>What is FASTA format? (n.d.). University of Michigan.
- <sup>30</sup>Wong, M. L., & Medrano, J. F. (2018). Real-time PCR for mRNA quantitation. *BioTechniques*, 39(1).
- <sup>31</sup>Zielińska, D., & Kolożyn-Krajewska, D. (2018). Food-Origin lactic acid bacteria may exhibit probiotic properties: Review. *Biomed Research International*, 2018(Food Microbiology). <https://doi.org/10.1155/2738>
- <sup>32</sup>Laranjo, L., Dunn, A. G., Tong, H. L., Kocaballi, A. B., Chen, J., Bashir, R., Surian, D., Gallego, B., Magrabi, F., Lau, A., & Coiera, E. (2018). Conversational agents in healthcare: a systematic review. *Journal of the American Medical Informatics Association : JAMIA*, 25(9), 1248–1258. <https://doi.org/10.1093/jamia/ocy072>
- <sup>33</sup>Bokulich, Nicholas & Mills, David. (2012). Next-generation approaches to the microbial ecology of food fermentations. *BMB reports*. 45. 377-89. 10.5483/BMBRep.2012.45.7.148.

<sup>34</sup>Lee PY, Costumbrado J, Hsu CY, Kim YH. Agarose gel electrophoresis for the separation of DNA fragments. *J Vis Exp.* 2012 Apr 20;(62):3923. doi: 10.3791/3923. PMID: 22546956; PMCID: PMC4846332.

<sup>35</sup>Maidak, Gary J. Olsen, Niels Larsen, Ross Overbeek, Michael J. McCaughey, Carl R. Woese, The Ribosomal Database Project (RDP), *Nucleic Acids Research*, Volume 24, Issue 1, 1 January 1996, Pages 82–85, <https://doi.org/10.1093/nar/24.1.82>.

**CHEMISTRY  
TEAM PROJECT**



# HPLC Analysis of Active Ingredient Mass in Expired and Unexpired OTC Analgesics

Quiana Guo, Sahil Jain, Harris Khan, Graciela León, Miranda Lynch, Sarah Magdeburg, Annaliese Niebauer, Della Noon, Amanda Qu, Gianna Tout-Puissant, Tianna Tout-Puissant, and Christopher Zhou

## Abstract

Over-the-counter (OTC) pain medications and analgesics are highly common in US households. However, there exists a level of uncertainty regarding whether or not to consume expired OTCs. Three analgesics – aspirin, acetaminophen, and ibuprofen – were quantitatively analyzed using high-performance liquid chromatography (HPLC) to determine if there was a significant change in mass of the active pharmaceutical ingredient (API) after a medication's expiration date. The expiration date is the time at which the drug starts to become unstable from a variety of environmental factors. Aspirin and acetaminophen showed no statistically significant difference in active ingredient mass between unexpired and up to six years expired tablets; ibuprofen that was twelve years past its expiration date was the only analgesic in which any statistically significant difference was observed in the API mass between expired and non-expired tablets. Aspirin and acetaminophen tablets showed little decomposition in periods less than a decade while ibuprofen seemed to reveal some degradation, likely due to its older expiration date.

## I. Introduction

Over-the-counter (OTC) analgesics, commonly known as pain relievers, are ubiquitous in U.S. households.<sup>1</sup> They are often kept past their expiration date, prompting consumers to question the drug's safety and efficacy past that date. The Food and Drug Administration's definition of a drug's expiration date is the time period when the drug is known to be stable. In other words, its strength, quality, and purity are all intact when stored in the conditions labeled on the bottle.<sup>2</sup> The expiration date is determined through stability testing, which ensures that a medicine's potency and integrity remains optimal over a period of time.<sup>3</sup> Factors like moisture, heat, and light may affect the quantity of the active ingredient and the quality of an analgesic prior to its expiration date. From a broader perspective, there are not only potency but safety issues that can arise from the use of expired medications, such as antibacterial growth and antibiotic resistance.<sup>4</sup>

Analgesic consumption in the US has been common for decades; a 1988-1994 National Health and Nutrition Examination Survey revealed that 147 million adults in the US reported monthly analgesic use.<sup>5</sup> As of 2019, nearly nine out of ten Americans—more than 260 million individuals—use over-the-counter pain medications daily.<sup>6</sup> Common non-prescription analgesics include acetaminophen (paracetamol), aspirin, and ibuprofen; their chemical structures are shown in Figure 1. In this study, the aspirin, acetaminophen, and ibuprofen content in expired/unexpired OTC medications will be quantified to determine the relationship between expiration date and active ingredient mass.

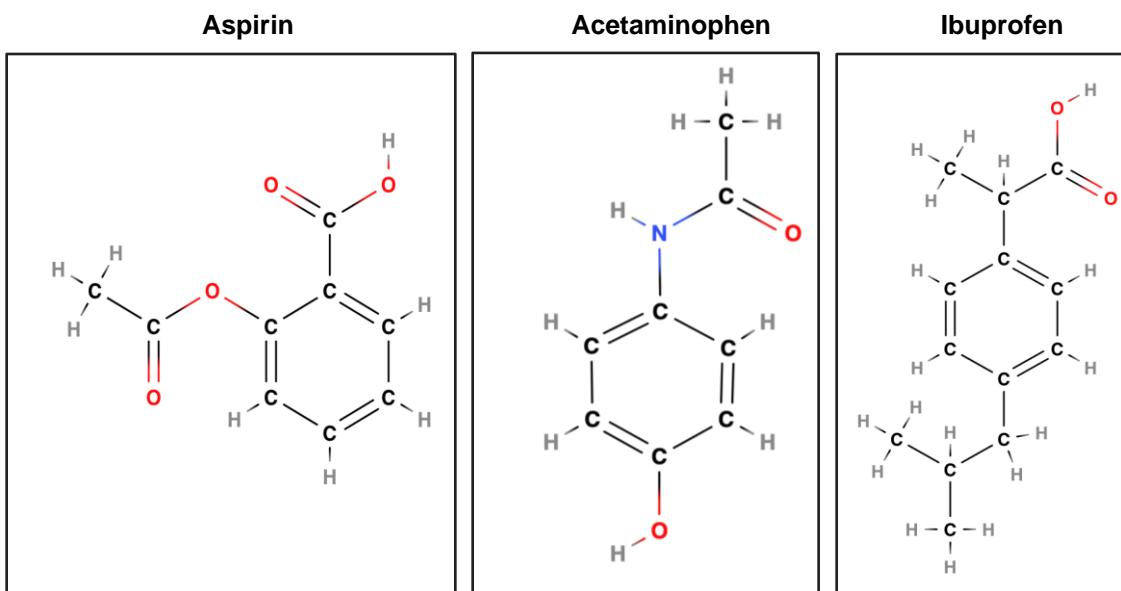


Figure 1: Chemical structures of aspirin, acetaminophen, and ibuprofen

## A. Aspirin

### 1. History

The history of aspirin dates back to roughly 3,500 years ago when willow tree bark was utilized as a painkiller and antipyretic (i.e. fever-reducer) first by Sumerians and Egyptians and later by Romans and the Greeks.<sup>7</sup> Aspirin's precursor, salicylate, was discovered by Reverend Stone in 1763, who analyzed its antipyretic effects. Italian chemist Raffaele Piria was able to extract salicylic acid from salicylate in 1838. In 1897, Bayer chemist Felix Hoffmann successfully synthesized acetylsalicylic acid in a stable form, creating aspirin. Bayer is typically associated with aspirin due to Hoffmann's success in synthesizing what would soon be known as "the miracle drug" because of acetylsalicylic acid's versatility. Common brands of aspirin in the U.S. besides Bayer, include Ascriptin, Aspergum, Aspirtab, Easprin, Ecotrin, Ecpirin, Genacote, Halprin, Ninoprin, and Norwich Aspirin. Bayer Extra Strength Aspirin for Migraines was approved by the FDA in 2001 and quickly became popular among consumers. Later, the FDA approved the capsule form of 325 mg aspirin in 2013. Aspirin was originally used as an antipyretic and anti-inflammatory drug; however, it later became known for its antiplatelet properties. This has made it useful in the prevention of cardiovascular and cerebrovascular diseases, as well as in trials testing its chemopreventive effect against colorectal and various other forms of cancer. Hence, aspirin continues to live up to its nickname as the miracle drug.<sup>7</sup>

### 2. Application

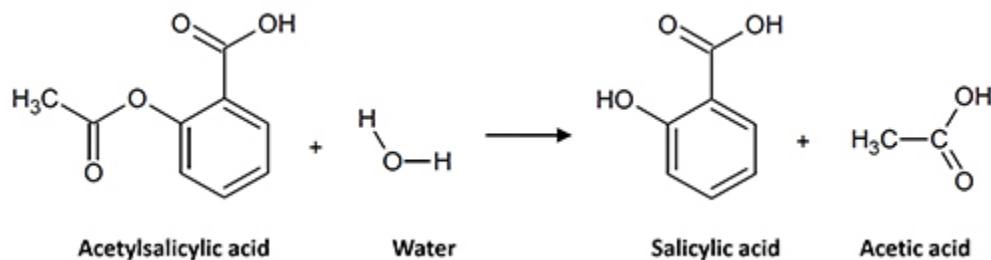
Aspirin is one of the most common pain relievers in the United States. John Robert Vane discovered the mechanism through which aspirin acts to relieve pain. He proved that NSAIDs inhibit an enzyme, cyclooxygenase, or COX. COX leads to the production of prostaglandins (PGs) that produce swelling and fever.<sup>8</sup> Studies from 2002 from the Antithrombotic Trialists (ATT) on patients with cardiovascular diseases reveal the implications and use of aspirin<sup>9</sup>; in secondary prevention - reducing the impact of a disease early on - there was a 25% decrease in the risk of vascular conditions across 75 mg to 1500 mg daily doses.

Meanwhile, doses below 75 mg yielded a 13% reduction rate, albeit the latter was based on a mere three randomized trials. In patients with existing acute myocardial infections (MI's), there was a 49% reduction rate in non-fatal MI, 46% decrease in non-fatal stroke, 23% reduction in vascular death - all with a 162.5 mg daily dose of aspirin after 5 weeks. For both primary prevention - preventing the onset of a disease - and secondary prevention, a 75 mg dose was sufficient to inhibit thromboxane B<sub>2</sub>, a substance known to cause blood clots. For patients with acute MI conditions, though, a dose of 162.5 mg to 325 mg was needed for an antithrombotic effect. Furthermore, higher doses of 500-1500 mg were not needed for such outcomes. Overall, the ATT study evidently suggests that aspirin has a statistically significant benefit for preventing vascular events.<sup>9</sup>

### 3. Expiration

The United States Pharmacopeia Aspirin guidelines call for no less than 90% and no more than 110% of the labeled active ingredient in an aspirin tablet. In regards to expiration dates, a study by the Food and Drug Administration reveals that about 90% of both prescription and over-the-counter drugs were generally safe and effective even 15 years past their respective expiration date.<sup>10</sup>

Aspirin, like many other pharmaceuticals, decomposes in the solid state. This process is first order with respect to concentration. Vapor pressure and temperature are two prominent factors known to drive this reaction, although over a period of 50 days, aspirin shows little degradation at 80 °F or below.<sup>11</sup> From a molecular perspective, the decomposition begins with the hydrolysis of the ester group in acetylsalicylic acid - resulting in salicylic acid and acetic acid. Through moisture, bonds in the active ingredient are broken due to the addition of water. This step is known to be responsible for the vinegar-esque odor that expired aspirin produces. Through heat, temperature increases the rate constant of the decomposition reaction; bonds are broken as energy is absorbed and the substance degrades faster. Either way, the process is endergonic and undergoes the same mechanism.<sup>11</sup>



**Figure 2: Decomposition reaction of acetylsalicylic acid through hydrolysis of the ester group<sup>12</sup>**

## B. Acetaminophen

### 1. History

Though acetaminophen was first synthesized in 1878, it wasn't until 1950 that it became ubiquitous for medicinal use. Prior to the discovery of acetaminophen, aspirin and quinine, two compounds found in the bark of willow and cinchona trees, respectively, had long been used as pain relievers. But as the cinchona tree became scarce, people began searching for alternatives.<sup>13</sup> In 1899, acetaminophen was found to be a metabolite of acetanilide, a fever-reducing drug and alternative to aspirin since 1886.<sup>14</sup> However, these results were largely overlooked until another study found that acetaminophen catalyzed an analgesic effect

when acetanilide was consumed. Acetanilide was also found to cause a non-lethal blood condition, which encouraged further research on acetaminophen, since the latter did not exhibit the toxic side effects that acetanilide did. Following clinical studies and further research in the 1950s, acetaminophen was sold generically or under the brand name Tylenol in the US.<sup>13</sup>

## 2. Application

Unlike aspirin and ibuprofen, acetaminophen is not classified as a non-steroidal anti-inflammatory (NSAID) medication, meaning it cannot reduce swelling. However, it does serve as an analgesic, relieving mild to moderate pain for a variety of symptoms. Acetaminophen is commonly used to alleviate headaches, muscle aches, and joint aches.<sup>15</sup> Because acetaminophen is not addictive like its prescription counterparts such as opioids, it is often used as the first line of treatment for chronic pain. In addition to being classified as an analgesic, acetaminophen is also considered an antipyretic, a fever reducer, making it an active ingredient in many cold and flu medications.<sup>14</sup>

Furthermore, acetaminophen is considered a safe and effective analgesic for people with heart conditions. Since it does not raise blood pressure or cause fluid retention, acetaminophen does not raise the risk for heart attack or worsen congestive heart failure symptoms like ibuprofen does. Acetaminophen is also more gentle on the stomach than other analgesics, and it is not directly associated with gastrointestinal irritation, ulcers, or bleeding. While it is a safe analgesic for people with heart conditions, acetaminophen is not ideal for people with liver conditions. Those with preexisting liver conditions or those at risk should avoid acetaminophen altogether due to its risk of worsening symptoms.<sup>16</sup>

The mechanism by which acetaminophen works for pain and fever relief is still unknown.<sup>17</sup> Existing theories point towards acetaminophen potentially inhibiting the production of prostaglandins by limiting the COX pathway. The COX pathway is a major enzyme system that is involved in the synthesis of prostaglandins, lipids involved in the biological processes related to inflammation and pain.<sup>18</sup> However, unlike NSAIDS which have been found to specifically target COX to reduce fever inflammation and pain, acetaminophen shows only weak inhibition of COX and consequently does not act as an anti-inflammatory.<sup>19</sup>

## 3. Expiration

Listed expiration dates of acetaminophen are generally within four to five years of its manufacture.<sup>20</sup> When acetaminophen degrades, it commonly yields impurities including: 4-aminophenol, 4-nitrophenol, *p*-benzoquinone, hydroquinone, 4'-chloroacetanilide.<sup>21</sup> 4-aminophenol is hazardous with teratogenic and nephrotoxic effects, and is the most frequent impurity that has been established to emerge during pharmaceutical preparation and storage. Degrading acetaminophen is transformed by hydrolysis and oxidation reactions which can be driven by presence of moisture, high temperatures, microorganisms, and other contaminants.<sup>22</sup>

Research done by the Shelf Life Extension Program (SLEP) in conjunction with the FDA has supported that the actual shelf life of acetaminophen exceeds its listed expiry dates.<sup>23</sup> Using HPLC analysis, the SLEP found that shelf life of combined acetaminophen-pseudoephedrine capsules could be extended by an average of two years.<sup>24</sup> In an analysis of long expired prescription medications, medications and their drug compounds were analyzed for degradation. Hycomine, which had expired 28-40 years prior, was found to contain 100% of the 250 mg of labeled active ingredient acetaminophen. The declared amount of acetaminophen present in the hycomine was 250 mg compared to the measured amount after expiry which was 249 mg with a standard deviation of 38.3 mg.<sup>25</sup>

## C. Ibuprofen

### 1. History

Ibuprofen was initially synthesized in 1961 by Dr. Stewart Adams, Dr. John Nicholson, and technician Colin Burrows while working for Boots Pure Drug Company Ltd. Dr. Adams was searching for a new treatment for rheumatoid arthritis that would not cause the symptoms of typical corticosteroid treatments. He began to look for a compound that would be similar to the NSAID aspirin. The result of ten years of research was a new NSAID: 2-(4isobutylphenyl) propionic acid, also known as ibuprofen. Ibuprofen was patented in 1962. However, since the new drug was required to undergo extensive clinical trials, it wasn't until 1969 that ibuprofen was marketed to the public in the United Kingdom and approved for over-the-counter sales in 1983.<sup>26</sup> It was later introduced to the United States in 1974 and sold over-the-counter in 1984.<sup>27</sup> The drug ibuprofen has since grown in popularity, commonly recognized across the world by brand names such as Advil, Motrin, Nuprin, and IBU.<sup>28</sup>

### 2. Application

Ibuprofen is primarily used as a pain reliever for many different types of conditions, including, but not limited to: headaches, menstrual pain, muscle aches, and arthritis.<sup>29</sup> It has also been shown that, in higher dosages, ibuprofen helps slow the progression of cystic fibrosis and mild lung disease.<sup>30</sup> Ibuprofen functions by inhibiting the body's ability to produce prostaglandins—chemicals that are linked to inflammation and several other immune responses.<sup>31</sup> Over-the-counter ibuprofen can be found and used as an oral medicine. Usually taken every 4-6 hours by mouth with a glass of water, the dosage is most commonly 200 mg in a tablet form, but can vary depending on the condition and prescription.<sup>32</sup> The full benefit of the drug may not occur until after two weeks of consistent usage.

### 3. Expiration

According to previous studies, unexpired ibuprofen maintained over 90% efficacy within the first three months of analysis after production. However, after six to seven months, the concentration of ibuprofen in each tablet began to significantly deviate from the initial concentration. After 12 months of analysis, the ibuprofen tablets were, on average, only approximately 73% effective. As ibuprofen tablets became less effective over time, the toxic degradation product, 4-isobutylacetophenone, was produced. This study was carried out on samples stored in a normal bathroom cabinet in which spikes of temperature and humidities both occurred corresponding to the outside temperature. This study also concluded that vapor pressure and temperature play a prominent role in the degradation of ibuprofen.<sup>33</sup> Studies have also shown that common excipients, such as polyethylene glycol and polysorbate 80 accelerate the degradation of ibuprofen.<sup>34</sup> Thus, the rate of degradation of ibuprofen varies depending on its surroundings.

## D. HPLC

High performance liquid chromatography (HPLC) has been previously used to determine the mass of the active ingredient in aspirin, acetaminophen, and ibuprofen.<sup>35</sup> HPLC is a laboratory technique that can be used to separate, identify, and quantify chemical components in a liquid mixture. In this experiment, reversed phase HPLC was used, i.e. the mobile phase was more polar than the stationary phase. The sample was dissolved in a polar mobile phase of methanol and injected to the HPLC system. It was delivered by a controlled flow of mobile phase to the stationary phase, a C18 resin column which exhibits very low polarity. Different compounds were separated in the column due to their different degrees of interaction with the stationary phase. Non-polar molecules interact strongly with the non-polar stationary

phase; therefore, they remain in the column for a longer time and elute slowly. Vice versa, the polar molecules interact stronger with the polar mobile phase than the stationary phase so they stay in the column shortly before elution. As such, compounds were eluted at different times to reach the detector that measures the absorption of a specified wavelength of light of each compound. The workstation picked up the signal from the detector to generate peaks with integrated peak areas proportional to the concentration of the compound.<sup>36</sup> Figure 3 illustrates the path of the mobile fluid and sample through the system for HPLC analysis.

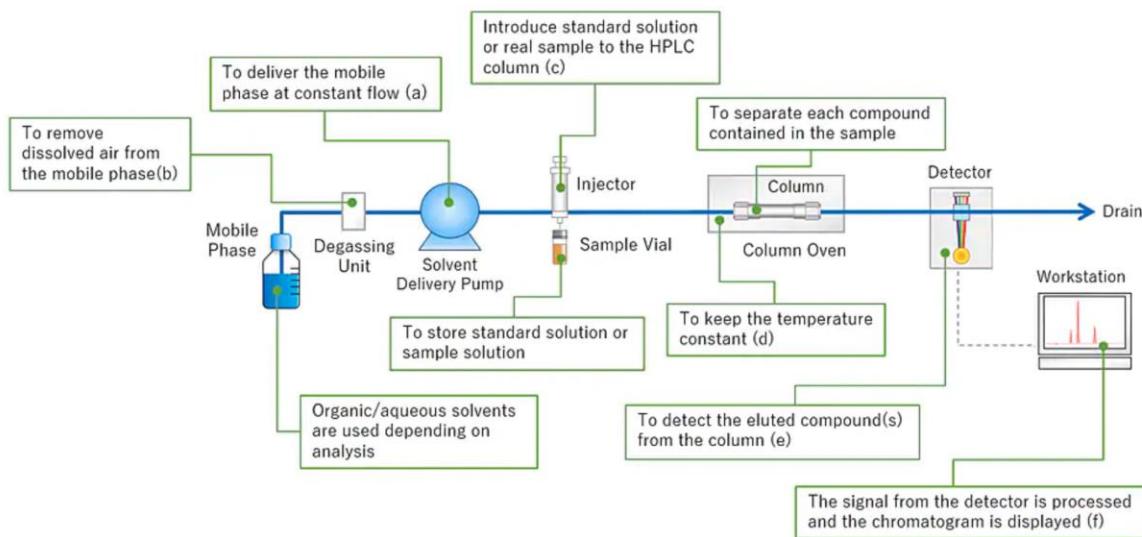


Figure 3: HPLC flow chart<sup>35</sup>

## II. Experimental Methods

### A. Chemicals

HPLC-grade water (99% purity) was obtained from the Fisher Scientific biotechnology company. Acetic acid (99.7% purity), acetonitrile ( $\geq$  99.9% purity), HPLC-grade methanol (99.9% purity), and triethylamine ( $\geq$  99% purity) were purchased from Sigma Aldrich. Acetaminophen (98% purity) and acetysalicylic acid (99% purity) were sourced from Acros Organics. Ibuprofen (99% purity) was purchased from Spectrum. The following unexpired medications were purchased from a local retailer: Advil Ibuprofen (200 mg tablets), Equate Ibuprofen (200 mg tablets), Top Care Acetaminophen (500 mg tablets), Tylenol (500 mg tablets), Bayer Aspirin (325 mg tablets). Expired OTC analgesics were crowdsourced from the local community.

### B. Equipment

#### 1. UV-vis Spectrometer

UV spectra were collected with a Cary 8454 UV-vis spectrometer operated by UV-Vis ChemStation software version B.05.04, which were purchased from Agilent Technologies.

## 2. HPLC

The HPLC Detector SPD-20A, HPLC Pump LC-20AT, C-R8A Chromatopac Data Processor manual start integrator, and Column Microsorb-MV 100-5 C18 150 \* 4.6 mm were sourced from Shimadzu Manufacturing Inc and used for the aspirin, acetaminophen, and ibuprofen analysis apparatus. The syringe used was obtained from Hamilton.

## C. Preparation of Standards

### 1. Aspirin Standards

A 3128 mg/L stock solution of acetylsalicylic acid was diluted to produce a series of six standards with concentrations ranging from 50.56 mg/L to 1251.2 mg/L. All solutions were prepared using a solvent consisting of water: acetonitrile: triethylamine: acetic acid (94.1:5.5:0.2:0.2 v/v/v/v).

### 2. Acetaminophen Standards

A 125.2 mg/L stock solution of acetaminophen was diluted to produce a series of seven standards within the concentration range of 10.0 mg/L to 125.2 mg/L. All solutions were prepared using a solvent consisting of water : acetonitrile : triethylamine : acetic acid (94.1 : 5.5 : 0.2 : 0.2 v/v/v/v).

### 3. Ibuprofen Standards

A 2035 mg/L stock solution of ibuprofen already in solution with methanol was diluted using methanol to produce five standards within the concentration range of 41 mg/L to 407 mg/L.

## D. Preparation of Samples

The acetaminophen and aspirin samples were prepared using a solvent of 94.1:5.5:0.2:0.2 (v/v/v/v) water: acetonitrile: triethylamine: acetic acid; ibuprofen samples, on the other hand, were prepared using methanol.

### 1. Aspirin Samples

Three tablets from each of the three bottles of Aspirin were obtained. Each of these nine pills was individually dissolved in 75 mL of solvent by mechanical stirring and then diluted to 100 mL. Five milliliters of each solution were diluted with different amounts of solvent to 25 mL for analysis.

## 2. Acetaminophen Samples

Three tablets for each sample were individually dissolved in 50 mL of solvent through mechanically stirring, which were then diluted to 100 mL using the solvent. Two milliliters were taken from each 100 mL solution and diluted to 50 mL. Four milliliters were taken from each 50 mL solution and diluted to 10 mL for analysis.

## 3. Ibuprofen Samples

Three tablets from each of the five bottles of ibuprofen were selected and individually dissolved in 50 mL methanol through mechanical stirring. Once dissolved, the samples were diluted to 100 mL with methanol. Prior to HPLC analysis, 1 mL of each sample solution was further diluted to 10 mL with methanol. The 10 mL sample solutions were filtered via 0.22  $\mu\text{m}$  syringe filters and then analyzed through HPLC.

# E. Analysis

## 1. UV-vis Spectroscopy

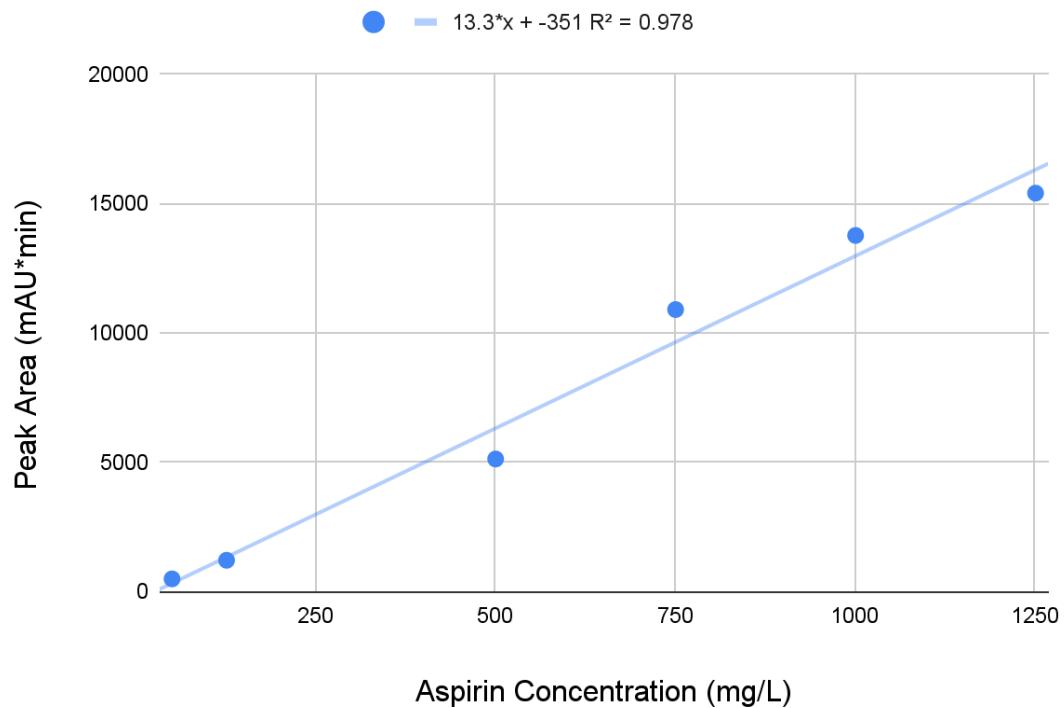
A UV spectrum for the highest concentration standard for each analgesic was collected using a 1-cm glass cuvette to identify the optimal wavelength for detection during HPLC analysis.

## 2. High Performance Liquid Chromatography

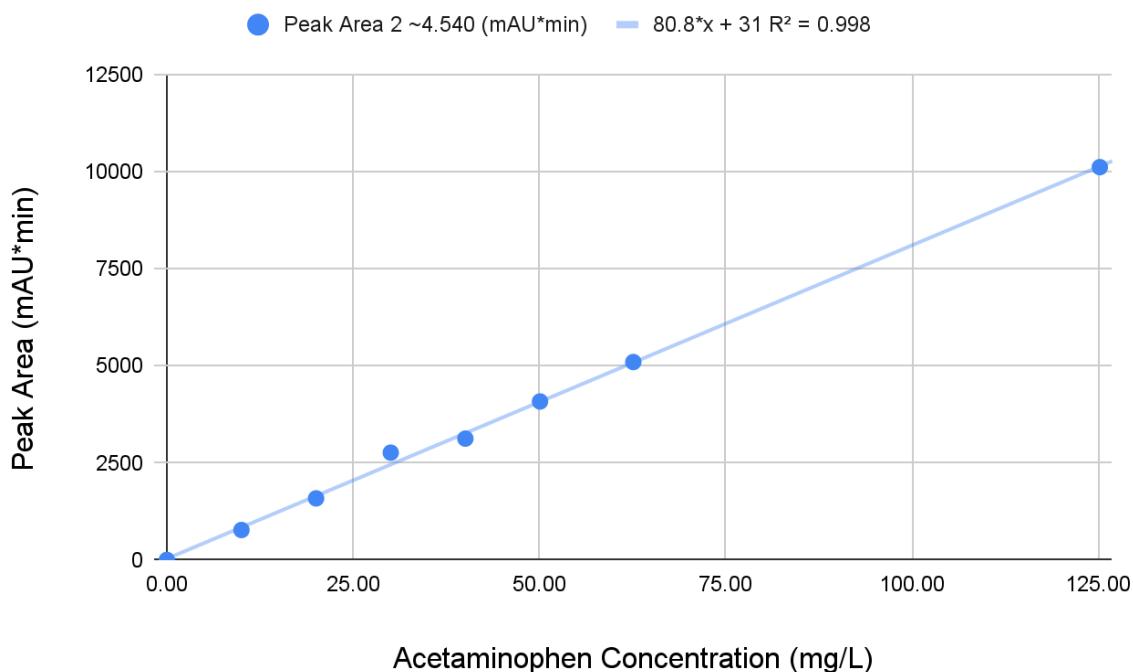
Aspirin and acetaminophen standards and samples used a mobile phase comprised of 94.1:5.5:0.2:0.2 (v/v/v) water: acetonitrile: triethylamine: acetic acid. The flow rate was 1.5 mL/min and the wavelength of detection was 270 nm. The injected samples contained about 40  $\mu\text{L}$  of volume. Acetaminophen eluted at approximately 2.2 and 4.5 min. Aspirin eluted at approximately 8.9 min. Ibuprofen standards and samples were analyzed in HPLC using the following parameters: a mobile phase of HPLC grade methanol, a flow rate of 1 mL/min, a wavelength of detection of 221 nm, and a volume of 20  $\mu\text{L}$  for injected samples. Ibuprofen eluted at approximately 1.7 and 2.4 min.

# III. Results

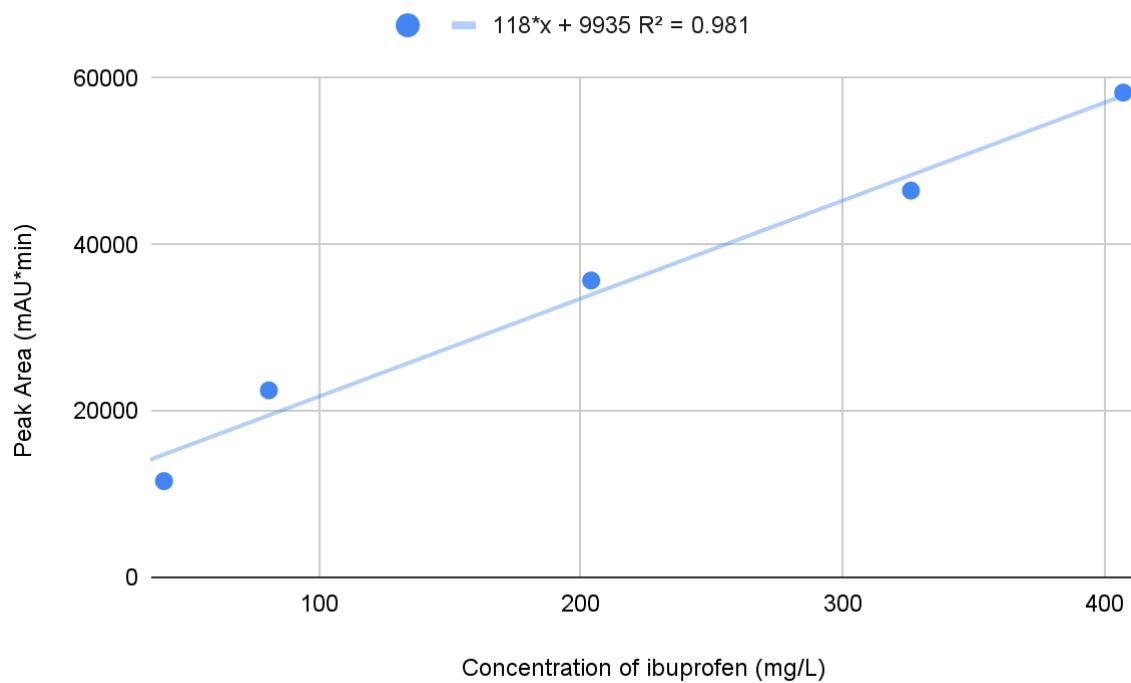
Calibration curves for each analgesic are shown in Figures 4-6. The standard curves were created by running a set of standards diluted from a stock solution for each medication through HPLC and finding a line of best fit with the data points. The  $R^2$  value indicates the linearity of the curve and its overall accuracy. The closer  $R^2$  is to one, the stronger correlation between the concentration and peak areas. The line of best fit of the standard curves was used to calculate the concentration of the active ingredient for each respective drug. Using the dilution factors, the mass of active ingredient for each sample was found and the mass of active ingredients in expired and unexpired medication were compared (Table 1).



**Figure 4: HPLC peak area as a function of aspirin concentration**



**Figure 5: Acetaminophen calibration curve of standards**



**Figure 6: Ibuprofen calibration curve of standards**

**Table 1: Mass of active ingredients in OTC analgesics**

Active Ingredient	Brand	Dosage (mg)	Expiration Date	Mass of Active Ingredient (mg)	Average Mass of the Active Ingredient (mg) ± margin of error		
Aspirin	Bayer	325	March 2016	282	292 ± 43		
				301			
				294			
	Bayer		Nov. 2024	280	265 ± 24		
				293			
				221			
	CVS		June 2017	199	187 ± 11		
				199			
				163			
Acetaminophen	Top Care	500	May 2019	506	518 ± 17		
				515			
				534			
	Sept. 2023			495	489 ± 11		
				478			
				493			
	Tylenol		March 2022	499	507 ± 11		
				503			
				518			
	March 2026			493	524 ± 26		
				557			
				522			
Ibuprofen	Advil	200	July 2022	277	237 ± 39		
				220			
				214			
	Equate		Sept. 2023	206	216 ± 11		
				224			
				219			
	Equate		Oct. 2019	149	176 ± 27		
				192			
				189			
	Kirkland		Oct. 2009	175	153 ± 67		
				198			
				86			

Active Ingredient	Brand	Dosage (mg)	Expiration Date	Mass of Active Ingredient (mg)	Average Mass of the Active Ingredient (mg) ± margin of error	
Aspirin	Bayer	325	March 2016	282	$292 \pm 43$	
				301		
				294		
	Bayer		Nov. 2024	280	$265 \pm 24$	
				293		
				221		
	CVS		June 2017	199	$187 \pm 11$	
				199		
				163		
	Top Care		March 2022	259	$214 \pm 49$	
				173		
				211		

## A. Aspirin

Figure 4 depicts the linear regression ( $R^2=0.978$ ) obtained from the six standards relating the calculated acetylsalicylic acid concentration (mg/L) and peak area (mAU\*min). In simpler terms, given a specific concentration of aspirin in solution, the theoretical peak area could be calculated, and vice versa. The aspirin row in Table 1 reveals the different brands and their respective dosage, expiration dates, active ingredient mass, and average active ingredient mass for each of three types of non-expired Bayern, expired CVS, and expired Bayer aspirin.

The mass of the active ingredient was calculated using the linear regression:

$$y = 13.3x - 351, \text{ where } y \text{ represents the peak area and } x \text{ represents the acetylsalicylic acid concentration}$$

For all samples, their exhibited peak area was substituted to solve for the experimental concentration, which represented the value after dilution. The latter was then used to find an experimental, original concentration of the aspirin solution before it was diluted through a dilution equation with an original volume of 5 mL and final volume of 25 mL. Statistical analyses using two-sample t-tests were used to test for significance at a significance level of  $\alpha = 0.05$  between the Bayer expired and unexpired. If the calculated p-value falls below  $\alpha$ , the data is significant and likely not caused by chance, while if the calculated p-value falls above, the data is not significant and can be attributed to the random error. The p-value was calculated to be 0.346, therefore, the null hypothesis was not rejected and there was no statistical significance to any comparisons between the expired and unexpired Bayer aspirins.

## B. Acetaminophen

Fig. 5 displays the standard curve, created with eight standards, used to determine the mass of active ingredients in the Tylenol and Topcare samples. The linear regression was very strong with an  $R^2$  value of 0.998. Table 1 presents the expired and unexpired varieties of medication and the masses of active ingredients they contain.

To obtain mass from peak area and linear regression:

$$\text{Linear regression: Peak area mAU * min} = 80.8(\text{acetaminophen concentration}) + 31$$

$$\begin{aligned} \text{Mass (mg)} &= \frac{(\text{Peak area mAU * min} - 31)}{80.8} \times \frac{\text{Dilution 2 V2 (final volume, mL)}}{\text{Dilution 2 V1 (initial volume, mL)}} \\ &\quad \times \frac{\text{Dilution 1 V2 (final volume, mL)}}{\text{Dilution 1 V1 (initial volume, mL)}} \times 0.1 \end{aligned}$$

**Table 2: Two-sample t-test p-values**

	Top Care Expired $\neq$ Unexpired	Tylenol Expired $\neq$ Unexpired
P-value	0.051	0.336

Within both brands of acetaminophen used, no significant decline of the mass of the active ingredient was observed in the expired medicine when compared to that of the unexpired. Because the average mass of the unexpired Top Care was consistently less than that of the expired, an unusual result, a two-sample t-test was performed to determine whether the difference between the averages was at all statistically significant. A p-value of 0.051 was found, so at the significance level of  $\alpha = 0.05$ , the difference between the averages of the unexpired and expired acetaminophen masses is not statistically significant. The Tylenol unexpired acetaminophen mass was consistently greater than the acetaminophen mass of the expired acetaminophen, and a two-sample t-test was used to determine whether the difference between the masses of acetaminophen in the expired and unexpired were significant. With a p-value of 0.336 at a significance level of  $\alpha = 0.05$ , there was also no significant difference between the acetaminophen mass in the unexpired and expired Tylenol.

### C. Ibuprofen

Three samples derived from three expired and two unexpired ibuprofen tablets were analyzed through HPLC following the development of the standard curve. A linear regression was used to determine the relationship between peak areas and the concentration of ibuprofen. Using peak areas and ibuprofen concentrations as measured by HPLC analysis of the standard solutions, a standard curve was created with a concentration range of 41  $\mu\text{g/L}$  to 407  $\mu\text{g/L}$ . The retention rate was observed to be 2.4 minutes on average for the standards and 2.5 minutes for the samples. A direct, positive correlation was found between the concentration of ibuprofen and peak area as shown by Figure 6. Table 1 displays the average mass of the active ingredient in ibuprofen tablets from each of the ibuprofen brands. At a significance value of 0.05, the Grubbs' test shows no statistically significant outlier in the collected data, even in the Kirkland ibuprofen data. Notably, a clear trend of the expired samples having a lower average mass of ibuprofen than the unexpired samples was observed.

To determine whether the resulting amounts of ibuprofen found in unexpired versus expired samples had a statistically significant difference, two two-sample t-tests were performed. Testing at a significance level of  $\alpha = 0.05$ , the first two-sample t-test tested for a statistically significant difference between the average masses of ibuprofen in unexpired and expired Equate, comparing the same brand. The t-test resulted in a p-value of 0.0582, providing convincing evidence that there is not a statistically significant difference between the average masses of ibuprofen in the samples of unexpired and expired Equate.

A second two-sample t-test was performed to test for a statistically significant difference between the average mass of ibuprofen in long-expired Kirkland tablets from 2009 and the average mass of ibuprofen in unexpired Equate and Advil tablets. Also testing at a significance level of  $\alpha = .05$ , the t-test resulted in a p-value of 0.0390, providing convincing evidence that there is a statistically significant difference between the average masses of ibuprofen in the samples of expired Kirkland and the unexpired Equate and Advil.

Given how generally similar all of the values for the average masses of ibuprofen are in samples of Advil, Top Care, and Equate, it can be concluded that there is no significant loss of ibuprofen in short-term expired medications. However, the significance found in the second two-sample t-test provides a case for a significant loss of ibuprofen for medications that have been expired for longer durations of time, such as over a decade.

## IV. Discussion

### A. Aspirin

#### 1. Data Analysis

Based on data, Bayer aspirin's active ingredient mass was contrary to expectations: the expired samples were found to have higher amounts of acetylsalicylic acid than the non-expired ones. For expired CVS aspirin, though, the active ingredient mass was considerably lower at 188 mg compared to the other two values of 265 mg and 292 mg for non-expired and expired Bayer aspirin tablets, respectively—albeit there was not an unexpired CVS aspirin for direct comparison of this group. Ultimately, it seems plausible to conclude that the time elapsed after expiration did not have a significant impact on at least Bayer's acetylsalicylic acid concentration. Simply put, it seems that the six years after Bayer's expiration date had relatively little to no impact on the concentration and mass of the active ingredient in aspirin tablets.

#### 2. Sources of Error

Possible errors in the experiment could have occurred during the various dilutions, potentially contributing to the varied results for unexpired Bayer aspirin. In particular, the standard calibration curve is the tool used to determine experimental active ingredient mass. In other words, an error in preparing the standards for the regression would lead to errors in analysis of the experimental aspirin concentrations before and after the dilutions of the sample. To illustrate, an over-dilution of a standard would lead to a lower concentration of aspirin than the theoretical value calculated through the standards' dilution set-up. This would result in a lower peak area and the slope would subsequently be smaller and more shallow. In turn, this would explain the relatively low mass of active ingredient in the non-expired Bayer aspirin as concentration and absorbance are directly proportional. Therefore, a decrease in absorbance (peak area) would correspond with a decrease in concentration and finally a lower experimental mass of acetylsalicylic acid.

#### 3. Areas for Improvement

In future experiments, an unexpired CVS brand tablet should be used in order to directly compare with the expired CVS aspirin results obtained from this study. That way, there is an actual comparison group for both Bayer and CVS brands; in simpler terms, there were expired and non-expired Bayer aspirin tablets, but there were only expired CVS tablets. In regards to lowering experimental error and inaccuracy, another process for diluting the tablets could include crushing numerous tablets and measuring out the mass of one tablet to obtain higher accuracy. This would increase an experimental yield since only one tablet was crushed for each sample. More importantly, though, more focus should be placed on the preparation of standards as these are used for the regression model that forms the basis of the results.

### B. Acetaminophen

#### 1. Data Analysis

The results of HPLC analysis of the unexpired and expired acetaminophen medications revealed no significant decline in active ingredient mass. This finding is consistent with previous findings in research on the stability of medications. The expired Top Care was past its expiration date by 38 months while the expired Tylenol was expired by five months. In previous research, acetaminophen was found to remain at

high potency following years of its expiry date<sup>37</sup>. The findings of this research support that acetaminophen does not significantly degrade and lose potency in the immediate months and even years after listed expiration.

In the course of HPLC analysis, the primary acetaminophen sample and standard peaks were preceded by smaller peaks. When running the stock solution to create the standard curve, two peaks were graphed. The first peak had a retention time at approximately two minutes and the second peak had a retention time at approximately four and a half minutes. The smaller peak was proportional to the larger peak in all of the standards, and the impurity remained at the same retention time in the standards of varying concentrations. It is hypothesized that the impurity was either a byproduct of acetaminophen synthesis or an excess reactant. The first and third samples of the unexpired Tylenol did not present the first peak, meaning that the two tablets used were very pure. No other impurities were found, a finding consistent with the absence of acetaminophen degradation.

## **2. Sources of Error**

Several sources of error were present throughout the experiment, contributing to inconsistencies in the calculated active ingredient mass. In the expired Top Care and both Tylenol samples, the average mass of acetaminophen exceeded the 500 mg dosage on the label. There are multiple explanations to account for this discrepancy. Most likely, the increased acetaminophen mass resulted from human error, stemming from errors in the sample preparation. The samples in the volumetric flasks may not have been diluted to the intended mark, or the samples may not have been mixed thoroughly enough before being run through the HPLC. Releasing more of the higher concentrated solution than necessary when diluting as a result of using a graduated pipette causes a higher concentration than calculated and thus a higher mass than expected. In contrast, releasing less of the higher concentration solution results in a lower mass than expected. Each sample was created by a different group member, which could also explain the inconsistencies in results. Another possible rationale for the elevated acetaminophen mass could be a manufacturing error. Medications are not always produced to include exactly specified amounts of active ingredients. The FDA specifies that during the listed shelf life of medications, they can include up to 105% of active ingredient concentrations<sup>38</sup>.

## **3. Areas for Improvement**

If this experiment were to be repeated, multiple changes could be made to reduce error and draw additional conclusions. Either one person could make all the samples, or each group member could perform one step in each sample, so samples are consistently made. Consistency in creating samples, such as taking greater care to pay attention to the entire dilution process, decreases the chance of human error. More specifically, using a volumetric pipette instead of a graduated pipette would reduce measurement errors. In the future, testing a greater number of trials would reduce the amount of random error. More samples of varying expiration dates, especially older medicines, could be used in the future in order to investigate the same question over a longer period of time.

## C. Ibuprofen

### 1. Data Analysis

The five different samples were chosen due to two reasons. The Equate that expired as of Oct. 2019 and the Equate that expired in Sept. 2023 were chosen because they were both under the brand Equate, but they had different expiration dates. The Equate that expired in Oct. 2019 had been expired for three years while the Equate that expired in Sept. 2023 was unexpired. Thus, these were used for a direct comparison between expired and unexpired ibuprofen. The other three samples of Advil, Kirkland Signature, and Top Care were used for a larger sample of different expiration dates. The length of expiration of these three varied from one month to over 10 years.

As illustrated by the data table, the Kirkland ibuprofen tablets that expired over 10 years ago contained significantly less ibuprofen than all other ibuprofen tablets. Conversely, the Advil ibuprofen tablets that expired within the last month contain significantly higher amounts of ibuprofen than other tablets. The Top Care ibuprofen tablets that expired in March of 2022 contained similar levels of ibuprofen when compared to the Advil ibuprofen tablets. The trend in the results were consistent with past studies, but since the condition of storage of each sample is unknown, no definite conclusions can be drawn.

### 2. Sources of Error

During testing, several cases of ibuprofen showed residual solids being left over after mechanical stirring. However, this residual is not confirmed to be the actual ibuprofen, as it may also be other parts of the tablets. Each sample was mechanically stirred for 30 minutes or more, allowing for the assumption that the undissolved particles may not have been ibuprofen. Also, the appearance of the solution was inconsistent in color and purity. Another source for possible errors and inconsistencies may come from the method in which the samples were prepared. Instead of one individual conducting all of the samples, the results came from multiple individuals conducting the testing of samples. This may lead to inconsistencies in the data and lab research process.

### 3. Areas for Improvement

In order to obtain more accurate results if the experiment were to be performed again, some procedures could be done differently. To control consistency, only one individual could prepare all of the standards and samples or one person could prepare each group of samples. This would prevent errors with methods due to variations in experimental performance. In addition, it would likely create more consistency if comparisons were drawn from only one brand that had several different expiration dates, seeing as different brands could produce medications of different qualities. Expiration dates could have also been spread out more evenly to demonstrate exactly how long after the expiration date the active ingredient begins to degrade. Finally, a small amount of sodium hydroxide could be used to dissolve the ibuprofen tablets better and extract the full amount of active ingredient from the medication as a potential alternative extraction method.

## V. Conclusion

### A. Aspirin

The purpose of the investigation was to determine if the decomposition of the active ingredient was revealed through a change in its mass. Simply put, the HPLC analysis showed that there was no significant notable statistical difference between the expired and non-expired Bayer aspirin tablets such that the former showed any considerable degradation in comparison to the latter. In simpler terms, the half-decade elapsed after the expiration date did not necessarily reveal heavy degradation, considering the mass of active ingredients in all samples.

Although time may not be a considerable aspect in the breakdown of an active ingredient, there are certainly other factors that can be involved. Moisture and temperature, in particular, are two conditions which heavily influence the endothermic reaction of acid degradation. Since energy is absorbed to break bonds, an increase in temperature or a decrease in concentration of aspirin from water would be followed by an equilibrium shift towards the product side to favour the formation of salicylic acid and acetic acid after the ester group detaches. To this end, increasing levels of moisture and temperature in a non-expired aspirin tablet may have allowed the reaction to reach equilibrium faster and would perhaps then have a considerable effect on the acetylsalicylic acid mass compared to a tablet that would be kept in lower-temperature storage conditions.

Nevertheless, provided that all solutions and standards are prepared accurately, HPLC is a relatively reliable and efficient method of determining a solution's concentration from its absorbance based on molecules' intermolecular forces and polarities. The former can have many implications, and this experiment demonstrates its usage in determining acid concentration in analgesics and over-the-counter pain medications.

### B. Acetaminophen

Using HPLC, the components of acetaminophen were quantified, allowing the mass of active acetaminophen to be determined. The differences between the unexpired and expired medicines for both Tylenol and Top Care were statistically insignificant, showing that acetaminophen may not degrade heavily over short periods of time, so a belief that recently expired acetaminophen may not be as effective may not actually be true, as approximately the same concentration of active ingredient was found to still be present. This finding is in accordance with previous research demonstrating the stability of acetaminophen for extended periods following its expiration. Several error sources such as inconsistent sample preparation and manufacturing error could have resulted in biased data. Additionally, simulated conditions could model environmental conditions of acetaminophen contaminants in the environment<sup>37</sup>. As acetaminophen is a common pollutant, research analyzing its degradation products in various environments is relevant to understanding its potential impacts and dangers.

### C. Ibuprofen

This experiment aimed to determine if the mass of the active ingredient in over-the-counter analgesics, such as ibuprofen, would change after the expiration date. Ultimately, the samples of the unexpired ibuprofen tablets (Sample Y) had a similar concentration of ibuprofen to the concentration of the ibuprofen in the tablets that had expired within the past three years (Samples U, T, and X). However, there is evidence of a decrease in the active ingredient's mass in the tablets that expired in 2009 (Sample F). While the data

shows little to no degradation of the active ingredient in the recently expired ibuprofen tablets, there is a significant degradation in the ibuprofen tablets that expired several years before testing.

Variables, such as the consistency of experimenters, the solvent used to dissolve the tablets, the brands of tablets used, and the distribution of expiration dates, could be changed if the experiment were performed again. These changes would likely help the experimenter achieve more accurate and consistent results. Future testing with these corrections could confirm or disprove the findings of the experiment.

In conclusion, data collected from HPLC analysis suggests that the mass of the active ingredient in over-the-counter ibuprofen tablets does not decrease immediately after expiration, but it does decrease several years after the expiration date.

#### D. General

The results of all three analgesics reveal that the actual mass of active ingredients did not decrease within months or even a couple of years after the labeled expiration date provided by the manufacturer. When the expiration date has passed for more than a decade, however, there was an observed reduction in the amount of detected ibuprofen. A similar result might be obtained from the other two drugs if long-expired samples could be obtained. The results are consistent with previous studies which state that these drugs do not easily degrade unless under extreme spikes of high moisture level and high temperatures.

#### E. Future Considerations

In future studies, rather than testing naturally degrading samples, environmental conditions could be adjusted to provoke more significant degradation. These would model the extreme conditions common improperly disposed of analgesics would degrade in. Samples could be placed in extreme heat or elevated moisture environments to expedite the degradation process. Because drugs like acetaminophen and aspirin do not degrade significantly over the span of a few years, this procedure could allow researchers to yield more measurable differences in active ingredient concentration between expired and unexpired analgesics. Moreover, it would allow for researchers to gain insight into how these medications degrade in the environment and what hazards might arise from this process.

Furthermore, a greater variety of analgesics could be sourced from different locations, which would enable more concrete conclusions to be drawn. Testing different forms of medications could also be considered. In this study, only pill forms of OTCs were used; future experiments can test liquid, suspension, or gel forms of medications.

Beyond the lab, other implications regarding expired analgesics should also be considered. Namely, in light of recent environmental concerns, disposal methods for OTCs must be addressed. Many people are not aware of proper disposal methods, and which can result in adverse environmental effects. One of the fastest ways to safely dispose of expired medications is to bring them to drug take back programs, which are convenient and safe disposal sites funded on both the local and federal level<sup>59</sup>.

## V. Acknowledgements

First and foremost, we would like to thank the Pennsylvania Governor's School for the Sciences alumni and donors for providing us with the opportunity to conduct this research. We are appreciative of Carnegie Mellon University and the Mellon College of Science for allowing us to use their facilities and resources during this program. Thank you to our team project advisor, Dr. Diane Nutbrown, for her guidance through the project. We are also incredibly grateful for the help from our three teaching assistants, Clara Dou, Mars Hong, and Jennifer Zhou, for the assistance and support they provided throughout our research. Finally, we would like to thank all PGSS faculty and staff, especially the director, Dr. Barry Luokkala, for making PGSS an incredible experience.

## VI. References

1. Paulose-Ram, R.; Hirsch, R.; Dillon, C.; Losonczy, K.; Cooper, M.; Ostchega, Y. Prescription and Non-Prescription Analgesic Use among the US Adult Population: Results from the Third National Health and Nutrition Examination Survey (NHANES III). *Pharmacoepidemiology and Drug Safety* 2003, 12 (4), 315–326. <https://doi.org/10.1002/pds.755>.
2. *Expiration Dates- Questions and Answers*. <https://www.fda.gov/drugs/pharmaceutical-quality-resources/expiration-dates-questions-and-answers>
3. Drug Expiration Dates: How Accurate Are They? | Winchester Hospital <https://www.winchesterhospital.org/health-library/article?id=73446#:~:text=Pharmaceutical%20manufacturers%20determine%20a%20drug> (accessed 2022 -07 -24).
4. Don't Be Tempted to Use Expired Medicines <https://www.fda.gov/drugs/special-features/dont-be-tempted-use-expired-medicines#:~:text=Expired%20medical%20products%20can%20be> (accessed 2022 -07 -24).
5. Trygstad, T. Americans Love OTC Medications <https://www.pharmacytimes.com/view/americans-love-otc-medications> (accessed 2022 -07 -25).
6. Montinari MR, Minelli S, De Caterina R. The first 3500 years of aspirin history from its roots - A concise summary. *Vascul Pharmacol*. 2019 Feb;113:1-8. doi: 10.1016/j.vph.2018.10.008. Epub 2018 Nov 2. PMID: 30391545.
7. Vane JR, Botting RM. The mechanism of action of aspirin. *Thromb Res*. 2003 Jun 15;110(5-6):255-8. doi: 10.1016/s0049-3848(03)00379-7. PMID: 14592543.
8. Hennekens, C. H.; Sechenova, O.; Hollar, D.; Serebruany, V. L. Dose of Aspirin in the Treatment and Prevention of Cardiovascular Disease: Current and Future Directions. *Journal of Cardiovascular Pharmacology and Therapeutics* 2006, 11 (3), 170–176. <https://doi.org/10.1177/1074248406292263>.
9. Acetaminophen, Aspirin, and Caffeine Tablets [https://doi.usp.org/USPNF/USPNF\\_M240\\_01\\_01.html](https://doi.usp.org/USPNF/USPNF_M240_01_01.html) (accessed 2022 -07 -26).
10. Leeson, L. J.; Mattocks, A. M. Decomposition of Aspirin in the Solid State\*\*College of Pharmacy, University of Michigan. Ann Arbor. *Journal of the American Pharmaceutical Association (Scientific ed.)* 1958, 47 (5), 329–333. <https://doi.org/10.1002/jps.3030470509>.

11. Kelly, C. A. Determination of the Decomposition of Aspirin. *Journal of Pharmaceutical Sciences* 1970, 59 (8), 1053–1079. <https://doi.org/10.1002/jps.2600590802.1>
12. FORMULATION DEVELOPMENT - A Quick Approach for Evaluation of Drug-Excipient Compatibility: Case of Acetylsalicylic Acid. (2021, August 26). Retrieved from Drug Development and Delivery website: <https://drug-dev.com/131805-2/>
13. Acetaminophen. <https://www.newworldencyclopedia.org/entry/acetaminophen> (accessed Jul 22, 2022)
14. acetaminophen | Classification, Uses, & Side Effects <https://www.britannica.com/science/acetaminophen>.
15. Acetaminophen: Medlineplus Drug Information. [https://medlineplus.gov/druginfo/meds/a681004.html#:~:text=Acetaminophen%20is%20used%20to%20relieve,\)%2C%20and%20to%20reduce%20fever.](https://medlineplus.gov/druginfo/meds/a681004.html#:~:text=Acetaminophen%20is%20used%20to%20relieve,)%2C%20and%20to%20reduce%20fever.) (accessed Jul 22, 2022).
16. Avitzur, O.; M.D. Acetaminophen Is the Best Pain Reliever for Heart Patients <https://www.consumerreports.org/drugs/acetaminophen-best-pain-reliever-heart-patients/>.
17. Acetaminophen - American Chemical Society <https://www.acs.org/content/acs/en/molecule-of-the-week/archive/a/acetaminophen.html>.
18. Ohashi, N.; Kohno, T. Analgesic Effect of Acetaminophen: A Review of Known and Novel Mechanisms of Action. *Frontiers in Pharmacology* 2020, 11. <https://doi.org/10.3389/fphar.2020.580289>.
19. Kirkpatrick, P. New Clues in the Acetaminophen Mystery. *Nature Reviews Drug Discovery* 2005, 4 (11), 883–883. <https://doi.org/10.1038/nrd1887>.
20. Kar, A. K.; Kar, B. In-Vitro Comparative Dissolution Study of Commercially Available Paracetamol Tablet. *Journal of Drug Delivery and Therapeutics* 2020, 10 (1), 18–23. <https://doi.org/10.22270/jddt.v10i1.3817>.
21. Calinescu, O.; Badea, I. A.; Vladescu, L.; Meltzer, V.; Pincu, E. HPLC Separation of Acetaminophen and Its Impurities Using a Mixed-Mode Reversed-Phase/Cation Exchange Stationary Phase. *Journal of Chromatographic Science* 2012, 50 (4), 335–342. <https://doi.org/10.1093/chromsci/bmr043>.
22. Li, J.; Ye, Q.; Gan, J. Degradation and Transformation Products of Acetaminophen in Soil. *Water Research* 2014, 49, 44–52. <https://doi.org/10.1016/j.watres.2013.11.008>.
23. Commissioner, O. of the. Expiration Dating Extension. *FDA* 2022.
24. Lyon, R. C.; Taylor, J. S.; Porter, D. A.; Prasanna, H. R.; Hussain, A. S. Stability Profiles of Drug Products Extended beyond Labeled Expiration Dates. *Journal of Pharmaceutical Sciences* 2006, 95 (7), 1549–1560. <https://doi.org/10.1002/jps.20636>.
25. Cantrell, L.; Suchard, J. R.; Wu, A.; Gerona, R. R. Stability of Active Ingredients in Long-Expired Prescription Medications. *Archives of Internal Medicine* 2012, 172 (21), 1685–1687. <https://doi.org/10.1001/archinternmed.2012.4501>.
26. The Hangover That Led to the Discovery of Ibuprofen. *BBC News*. November 15, 2015.

27. Nonprescription use of ibuprofen and the risk of gastrointestinal toxicity  
[https://web.archive.org/web/20130815040113/https://www.fda.gov/ohrms/dockets/ac/02/briefing/3882b2\\_06\\_international%20ibuprofen%20foundation.htm](https://web.archive.org/web/20130815040113/https://www.fda.gov/ohrms/dockets/ac/02/briefing/3882b2_06_international%20ibuprofen%20foundation.htm) (accessed 2022 -07 -26).
28. Ogbru, O.; Marks, J. W. ibuprofen (Advil, Motrin) Side Effects (Alcohol), Uses, Dosage & Pregnancy Safety  
[https://www.medicinenet.com/ibuprofen/article.htm#what\\_else\\_should\\_i\\_know\\_about\\_ibuprofen](https://www.medicinenet.com/ibuprofen/article.htm#what_else_should_i_know_about_ibuprofen).
29. Drugs & Medications <https://www.webmd.com/drugs/2/drug-5166-9368/ibuprofen-oral/ibuprofen-oral/details#:~:text=Ibuprofen%20is%20used%20to%20relieve>.
30. Konstan, M. W.; Byard, P. J.; Hoppel, C. L.; Davis, P. B. Effect of High-Dose Ibuprofen in Patients with Cystic Fibrosis. *New England Journal of Medicine* **1995**, 332 (13), 848–854.  
<https://doi.org/10.1056/nejm19950330321303>.
31. Ricciotti, E.; FitzGerald, G. A. Prostaglandins and Inflammation. *Arteriosclerosis, Thrombosis, and Vascular Biology* **2011**, 31 (5), 986–1000. <https://doi.org/10.1161/atvaha.110.207449>.
32. Bushra, R.; Aslam, N. An Overview of Clinical Pharmacology of Ibuprofen. *Oman medical journal* **2010**, 25 (3), 155–1661. <https://doi.org/10.5001/omj.2010.49>.
33. Archibald, T.; Brown, S. Monitoring Commercial Ibuprofen Potency Changes over 1 Year When Stored in a Household Setting. *Journal of Pharmacy Technology* **2019**, 36 (1), 16–21.  
<https://doi.org/10.1177/8755122519877808>.
34. Cory WC, Harris C, Martinez S. Accelerated degradation of ibuprofen in tablets. *Pharm Dev Technol.* **2010** Dec;15(6):636-43. <https://doi.org/10.3109/10837450903426518>.
35. Ferguson, G.K. Quantitative HPLC Analysis of an Analgesic/Caffeine Formulation: Determination of Caffeine. *J. Chem. Educ.* 1998, 75(4), 467-469.
36. What is HPLC (High Performance Liquid Chromatography) ?  
[https://www.shimadzu.com/an/service-support/technical-support/analysis-basics/basic/what\\_is\\_hplc.html#2](https://www.shimadzu.com/an/service-support/technical-support/analysis-basics/basic/what_is_hplc.html#2).
37. Zilker, M.; Sörgel, F.; Holzgrabe, U. A Systematic Review of the Stability of Finished Pharmaceutical Products and Drug Substances beyond Their Labeled Expiry Dates. *Journal of Pharmaceutical and Biomedical Analysis* 2019, 166, 222–235.  
<https://doi.org/10.1016/j.jpba.2019.01.016>.
38. Aguilar, C. A.; Montalvo, C.; Ceron, J. G.; Moctezuma, E. Photocatalytic Degradation of Acetaminophen. *International Journal of Environmental Research* 2011, 5 (4), 1071–1078.  
<https://doi.org/10.22059/ijer.2011.465>.

**COMPUTER SCIENCE  
TEAM PROJECTS**



# Reversing Reversi: Using Recursion and Minimax Algorithms to Master Turn-Based Game Strategy

Alena Hemminger, Diya Hundiwala, Jai Rastogi, Bea Angela Ricafort

## Abstract

This report discusses the implementation of recursion and minimax algorithms into a smart computer program for the turn-based game Reversi, or Othello. The Python-based project explores both the benefits and challenges of assigning values to moves through the use of a board evaluation function in order to create a program that can successfully beat a lower-intelligence computer most, if not all, of the time.

## I. Introduction

### A. What is Reversi?

Reversi, also known as Othello, is a two-player game played on an 8 by 8 board with two-sided (usually black and white) disks. Figure 1 shows the beginning layout, which uses the letters A-H to organize its rows and the numbers 1-8 to label its columns. It contains two white disks in D4 and E5 and two black disks in D5 and E4. The game does not include any randomization factors, such as dice roll or physical ability, so it only involves pure strategy. Moreover, Reversi ends when either player runs out of legal moves, regardless of the current color in play. The winner is the player with the most disks of their color on the board.

	1	2	3	4	5	6	7	8
A								
B								
C								
D				W	B			
E				B	W			
F								
G								
H								

Figure 1: Starting Gameboard

### 1. A Brief History

Although the two titles are used interchangeably, Reversi and Othello *do* have differences. Patented in 1888, Reversi predates Othello, but the former was commercially advertised as the latter in the 1970s<sup>1,2</sup>. These two games mostly have the same gameplay, but Reversi originally lacked the “Othello starting configuration” – the four beginning pieces could be arranged in any way at the center of the board by the

participants in earlier versions of Reversi. (Reversi later adopted this configuration, as displayed in Figure 1.) Othello, however, terminates when *both* players have no more legal moves and simply skips the turn of the player that can no longer play, while a game of Reversi ends as soon as one player no longer has any legal moves<sup>3</sup>, and this rule is retained within this program.

## 2. How to Play

In Reversi, the player with black disks opens the game. They can make a move by placing a disk of their color in an unoccupied spot that is next to at least one disk of the opposite color and across from a disk of their own color in the same direction. (A player must “trap” at least one opposite color disk between two of their own disks. They can also trap multiple disks, as long as they form an uninterrupted horizontal, vertical, or diagonal line.) For instance, Figure 1 shows that C4 is unoccupied and presents a space above a white disk (D4). The disk at D4 is above a black disk at E4. Similarly, E6 is both unoccupied and to the right of a white disk at E5, and *this* disk is to the right of a black disk at E4. C4 and E6 are clearly in the set of legal moves for the black player at the start of the game. Moreover, the black player has two other moves in this set: F5 and D3. All of these combinations add one black disk to the board and swap one white disk, leading to four black disks and one white disk on the board.

All possible moves for the white player afterwards causes the board to again have an equal number of black and white disks. (The grid would contain three black and three white disks, but each pair of moves does not always balance the number of each disk color.) After a couple of turns, as shown in Figure 2, the score of the board can be very unbalanced. The board placements in Figure 2, however, still provides black with multiple available moves that swap multiple white disks, such as B3 and B4, meaning that the black player can still easily gain its footing in the game. Reversi ends when one or both players have no more remaining legal moves. This scenario can occur when a given player has no pieces left or when the grid is completely filled with disks.

	1	2	3	4	5	6	7	8
A								
B								
C			W	W	W			
D			W	W	W			
E			W	B	W			
F			B					
G								
H								

Figure 2: A “Lopsided” Board

## 3. Methods/Strategy

With a maximum of 60 disks to place on the board and a finite 8 by 8 board, some moves provide a much easier path to victory than other moves. Acquiring a corner is strategically the most important move because

no combination of disks can flip it. (Gaining a corner would cement a point for the successful player!) Similarly, the remaining 12 positions around the corners are harder to flip than any other generic piece, so corner-adjacent pieces would be the second-most advantageous positions in the game. A player can then prevent the opposing player from claiming an edge position first by using an early-game strategy – which would occur when only 16 disks are on the board – that features the appropriately named “danger zone.” (This region is later defined on the zone map, or Figure 18.) As the name implies, if a player has a disk in this zone, they would be in danger of losing a major strategic advantage because that piece could bridge the opponent to the much-desired edges. These positions are the worst locations for a player to place their disks because they can bridge the opponent to the corner. A player should then try to avoid these areas and allow the opposing player to claim part of the danger zone so that the player can later claim the edge or corner<sup>4</sup>. Granted, venturing into the danger zone is sometimes unavoidable before the game board has 16 or more disks placed, but the game must continue, regardless of faulty early-game positions!

## B. Turn-Based Games

### 1. What is a Turn-Based Game?

When people refer to strategy games, they usually talk about games that fall into two categories: real-time strategy (RTS) or turn-based strategy (TBS). Reversi is the latter because the players take turns, while real-time strategy games are played simultaneously. Turn-based strategy games may also give each player a certain amount of time to execute their turn<sup>5</sup>. Additionally, this category of strategy games does not include any element of randomization.

### 2. Examples of Turn-Based Games

Board games are often based on turn-based strategy. Examples of turn-based strategy games, aside from Reversi, include Chess, Checkers, Connect 4, and Nine Men’s Morris.

## II. Game Functionality

### A. Set-Up Functions

#### 1. *print\_board*

This function prints out the board’s rows and columns into the terminal, and this grid is created by printing dashes (-) and plus signs (+). Moreover, each row is marked with a letter **A-H**, while each column is numbered **1-8**. The positions of disks are also marked on the board with “**W**” for white or “**B**” for black. *print\_board* recognizes any placed disks by examining the two-dimensional array **board** – its parameter – which contains an element for every board position. If an element in **board** is empty, then *print\_board* does not add a letter to that corresponding spot. If a **board** element contains “**W**” or “**B**”, however, that letter will be printed on the grid in the terminal. Overall, the program accomplishes this task by running a nested **for** loop that accounts for each row and column, as well as the position of disks.

```
def print_board(board): # Prints the board to the user
    print ("\n\n\n\n\n\n\n\n      1   2   3   4   5   6   7   8")
    for r in range(0, 8):
        if r > 0:
            print(" ---+---+---+---+---+---+---+---+---")
        for c in range(0, 8):
            if c > 0:
                print(" | ", end = "")
            else:
                print(chr(65+r), end = " ")
            print(board[r][c], end = " ")
        print()
    print()
```

**Figure 3: *print\_board* Function Definition**

## 2. *clear\_board*

The ***clear\_board*** function is critical to the beginning of every game, and it replaces every position on the game board with a space, essentially providing the user with the view of an empty board. To accomplish this task, it utilizes a ***for*** loop in order to iterate through all of its rows. (Reversi contains eight rows on an 8 by 8 board, so the ***for*** loop runs eight times.) The function also uses the ***r*** variable when keeping track of its current iteration. Within the body of the function, it utilizes the ***board*** parameter, a two-dimensional array that represents the game board, which is passed into each call of the function. Various rows of ***board*** are accessed using square brackets and the ***r*** variable, and each of these rows is assigned to a list of eight spaces, representing the eight columns in each row. At the end of all eight iterations, the 8 by 8 board will be filled with spaces, allowing for the start of gameplay.

```
# Clears board by putting a space into every position on the board
def clear_board(board):
    for r in range(0, 8):
        board[r] = [ " ", " ", " ", " ", " ", " ", " ", " "]
```

**Figure 4: *clear\_board* Function Definition**

## B. Play Functions

### 1. *start*

The ***start*** function creates the menu at the beginning of the game, which presents the player with eight game modes: ***Player vs. Computer (Easy)***, ***Player vs. Computer (Hard)***, ***Player vs. Player***, and various

combinations of randomized computer and smart computer intelligence modes. It prints a message that prompts the user to select a mode by inputting the corresponding number **1-8**. If the player's input is equal to a valid string (a number **1-8**), then the **player\_dictionary** will assign each color their role, depending on the game mode. Specifically, the “**player**” role means that the user will control a specific color’s moves, the “**random**” role means that the computer will generate random moves for that color, and the “**smart**” roles mean that the computer will generate the best move for that color, depending on the program’s minimax algorithm. (The number after the “**smart**” roles indicates the level of the **look\_ahead** function that each game mode uses, which is equivalent to how many moves that the smart computer “looks ahead” in order to find its best move. A higher number would specify a “harder” game mode because the computer has looked through more turns and thus more move sets for each color.) Inputting an appropriate number will also break the **while** loop and end the **start** function. However, if the player does *not* input a valid string, the boolean **is\_legal** will be set to **False**, and the program will print a message that alerts the user of invalid input. It will then continue the **while** loop until the player enters a number **1-8**.

```
def start(): # Creates start menu of the game
    while True:
        global player_dictionary
        print("Pick your game mode - Type 1, 2, 3, 4, 5, or 6:\n1. Player vs. Computer (Easy)\n2. Player vs. Computer (Hard)\n3. Player vs. Player\n4. Computer vs. Computer\n5. Computer vs. Smart Computer\n6. Smart Computer Lvl. 0 vs. Smart Computer Lvl. 1\n7. Smart Computer Lvl. 1 vs. Smart Computer Lvl. 2\n8. Smart Computer Lvl. 2 vs. Smart Computer Lvl. 3")
        inp = input()
        is_legal = True

        if inp == "1":
            player_dictionary = {"W" : "player", "B" : "random"}
            is_legal = True
            break
        elif inp == "2":
            player_dictionary = {"W" : "player", "B" : "smart3"}
            is_legal = True
            break
        elif inp == "3":
            player_dictionary = {"W" : "player", "B" : "player"}
            is_legal = True
            break
        elif inp == "4":
            player_dictionary = {"W" : "random", "B" : "random"}
            is_legal = True
            break
        elif inp == "5":
            player_dictionary = {"W" : "random", "B" : "smart3"}
            is_legal = True
            break
        elif inp == "6":
            player_dictionary = {"W" : "smart0", "B" : "smart1"}
            is_legal = True
            break
        elif inp == "7":
            player_dictionary = {"W" : "smart1", "B" : "smart2"}
            is_legal = True
            break
        elif inp == "8":
            player_dictionary = {"W" : "smart2", "B" : "smart3"}
            is_legal = True
            break
        else:
            is_legal = False
            print("Not a valid mode, try again.")
            continue
    if is_legal:
        return
```

**Figure 5: start Function Definition**

## 2. play

Immediately following the **start** function, which the player uses to choose the current game mode, the **play** function runs when the global variable **game\_over** is false, meaning that the game is ongoing. Its first task in the **while** loop is determining the value of the list **position**, a variable responsible for holding the coordinates of the next disk position. For example, if the player wants to place a disk at C4, **position** would

hold the value **[2, 5]** when the player input is passed into **player\_turn**, which returns a set of coordinates. A computer player, on the other hand, would use the set of coordinates returned by the **random\_turn** and **smart\_turn** as the value of **position**. **play** then finds the value of **position** through **if** statements, depending on whether the color is being played by a random, human, or smart player. Any other case corresponds to an invalid outcome, breaking the **while** loop. Additionally, the element **color** is a global variable that corresponds to the white player (“**W**”) or the black player (“**B**”). After the program defines the next disk position, it passes this value to the **make\_move** function, which also requires the current playing board and the color in play. These coordinates are solidified as the next move; the program can now print the board and return the current number of colors for both colors.

After this process, **play** then prints out the “board values” – corresponding to the advantage of their current disk positions – of each color. At this point in the function’s runtime, the variable **pause** is **True**, signifying that the current color has ended its turn. This action prompts a message to the user that asks them to press **Enter** in order for the game to continue to the next color’s turn. Before sending this message, however, the program first sets **color** to the opposite color so that the next player can make a move move. The **while** loop continues this process until the game ends. If **check\_win** does not return the string “**Draw**,” then **game\_over** is set equal to **True**, and **pause** is set equal to **False** — the game is over. It also compares **w\_score** (the white player’s score) to **b\_score** (the black player’s score) in order to find which variable has the largest value. **color** would then be set equal to the winning player, depending on if **w\_score** or **b\_score** is higher, and the player that corresponds to the variable **color** will be congratulated for winning the game through the printed statement. In the case where **check\_win** does return the string “**Draw**,” **game\_over** is still equal to **True** and **pause** is still equal to **False**, but the program will now print a statement that reveals that the game has ended in a draw.

```
def play(): # Begins gameplay
    global game_over
    global color
    global position

    while not game_over:
        if player_dictionary[color] == "player":
            position = player_turn(the_board,color)
        elif player_dictionary[color] == "random":
            position = random_turn(the_board, color)
        # The second parameter of smart_turn determines the level of look_ahead
        elif player_dictionary[color] == "smart0":
            position = smart_turn(the_board, 0, color)
        elif player_dictionary[color] == "smart1":
            position = smart_turn(the_board, 1, color)
        elif player_dictionary[color] == "smart2":
            position = smart_turn(the_board, 2, color)
        elif player_dictionary[color] == "smart3":
            position = smart_turn(the_board, 3, color)
        else:
            print("I'm confused.")
            break

        make_move(the_board, color, position)
        print_board(the_board)
        get_score(the_board)
        pause = True
```

Figure 6a: **play** Function Definition

```

if check_end_of_game(the_board):
    if check_win(the_board) != "Draw":
        game_over = True
        pause = False
        if w_score > b_score:
            color = "W"
        if b_score > w_score:
            color = "B"
        print("Congratulations," + color + ", you are the winner!")
    else:
        game_over = True
        pause = False
        print("The game is a draw.")

if color == "W":
    color = "B"
else:
    color = "W"

if pause:
    input("Press Enter to continue.")

```

**Figure 6b: *play* Function Definition**

### 3. *end*

When the boolean **game\_over** is set to **True**, meaning that the game has ended because one or both of the players have run out of legal moves, the program uses the **end** function to prompt the player with the choice to play again. An input of “**Y**” or “**y**,” which indicates “yes,” sets **game\_over** to **False** and essentially clears and reprints the board. It also calls the **start** and the **play** function in order to restart the game. On the other hand, an input of “**N**” or “**n**,” which indicates “no,” simply terminates the program. An invalid input sets the boolean **is\_legal** to **False** and sends another message to the user that prompts them to retype their response. The **while** loop then continues until the player returns a valid response.

```

def end(): # Either terminates the game or allows for replay
    while True:
        global game_over
        print("Would you like to play again? Type Y/N.")
        inp = input()
        is_legal = True

```

```

if inp == "Y" or inp == "y":
    game_over = False
    is_legal = True
    clear_board(the_board)

    the_board[4][4] = the_board[3][3] = "W"
    the_board[3][4] = the_board[4][3] = "B"

    print_board(the_board)
    start()
    play()
    break

elif inp == "N" or inp == "n":
    sys.exit()
else:
    is_legal = False
    print("Not a valid option, try again.")
    continue

    if is_legal:
        return

```

Figure 7: *end* Function Definition (Both Photos)

## C. Turn Functions (Non-Minimax)

### 1. *player\_turn*

This function is used whenever the player begins their turn. *player\_turn* first prompts the player to type a position on the board where it wants to place its disk. The format of this input should be a letter followed by a number with no spaces in between those two characters. (This input will correspond to the board position's coordinates.)

However, the player's input must be “translated” into the program's coordinates. The terminal grid has a coordinate system of **A-H** for the rows and **1-8** for the columns, while the program numbers both the rows and columns **0-7**. For example, the position of C4 on the terminal grid would correspond to the coordinates **[2, 5]** within the program. *player\_turn* then converts the player-inputted coordinates into coordinates that the program would be able to recognize by converting the letter (or row) into a number with ASCII and changing the user-inputted number into an *int* type. Both of these values are decremented by 1 in order to compensate for the zero-based coordinate system.

The user-inputted row and column values are compared to the row and column values within the *legals* list, and if the coordinates are legal moves, the *while* loop terminates and those coordinates are “set” as the player's move. If that position is not a legal move, *is\_legal* is set to *False*, and the player is prompted

with a message that asks them to input valid coordinates. The **while** loop then continues until the player inputs a legal move.

```
def player_turn(board, color): # Takes player input and determines whether it is a legal move
    while True:
        print("Player's Turn - Type desired location (letter and number with no space and press enter): ")
        inp = input()
        if len(inp) == 0:
            print("Not a legal move, try again.")
            continue

        r = ord(inp[0]) % 32 - 1
        c = int(inp[1]) - 1
        is_legal = True

        for move in legal_moves(board, color):
            if [r,c] == move:
                is_legal = True
                break
            else:
                is_legal = False

        if is_legal == False:
            print("Not a legal move, try again.")
            continue
        elif is_legal:
            return [r,c]
```

**Figure 8: *player\_turn* Function Definition**

## 2. *random\_turn*

This function is used when the player chooses one of the modes with the **Computer (Easy)** as a player. It generates a random legal move that can be used by the color representing the computer. Essentially, this function calls the **legal\_moves** function, providing a list of legal move coordinates for the given color, and uses **random.choice** (from Python's **random** library) to return a randomly chosen legal move from the list.

```
def random_turn(board, color): # Returns a random legal move that can be made by color
    return random.choice(legal_moves(board, color))
```

**Figure 9: *random\_turn* Function Definition**

## D. Gameplay Functions

### 1. *check\_dir*

This function checks all directions of a specific disk position and then returns the number of disks that can be “swapped” (or changed to the opposite color) in each direction. Moreover, **check\_dir** accepts several parameters: the two-dimensional array **board**, an integer **r**, an integer **c**, a string **color**, an integer **dx**, and an integer **dy**. **board** holds all of the board positions, while the integers **r** and **c** mark the current column and row in the iteration, respectively. **color** indicates the color of the player that will make the next move, and the integers **dx** and **dy** represent the distance between the center and the other coordinates of the board. (**dx** and **dy** connect the center to a point on the edge of the border, which generates a direction. For

example, if **dx** equals **-1** and **dy** equals **-1**, the direction is northwest, as shown in Figure 10.) The combination of **dx** and **dy** thus represents the distance that the program is currently “checking.”

**check\_dir** then sets **count** equal to **1** because the disk position must flip at least one disk in order to be a legal position, and **count** represents the number of swaps that will occur if a disk is placed in a given position. If the disk does not flip any disks in the given direction, the function returns **0**. The function also sets **r** equal to **r + dy** and **c** equal to **c + dx** so that **r** and **c** are continuously “moving closer” to the targeted edge throughout the iteration.

**check\_dir** then sets the variable **opp** to the opposite color of the one that is playing. If the current position is within the bounds of the board and that spot contains a disk of an opposite color, it adds the value of **dy** to **r** and the value of **dx** to **c** before beginning a **while** loop that operates under the same parameters as the **if** statement but also increments **count** by 1. (Failing the initial **if** statement means that the token has no swaps in this direction and would return **0**.) After exiting the **while** loop, the program returns the value of **count** if the current position is within the bounds of the board and that position contains a token of the *same* color. Therefore, this function finds the number of disks that would be swapped or added if a move was made in a given direction (if possible).

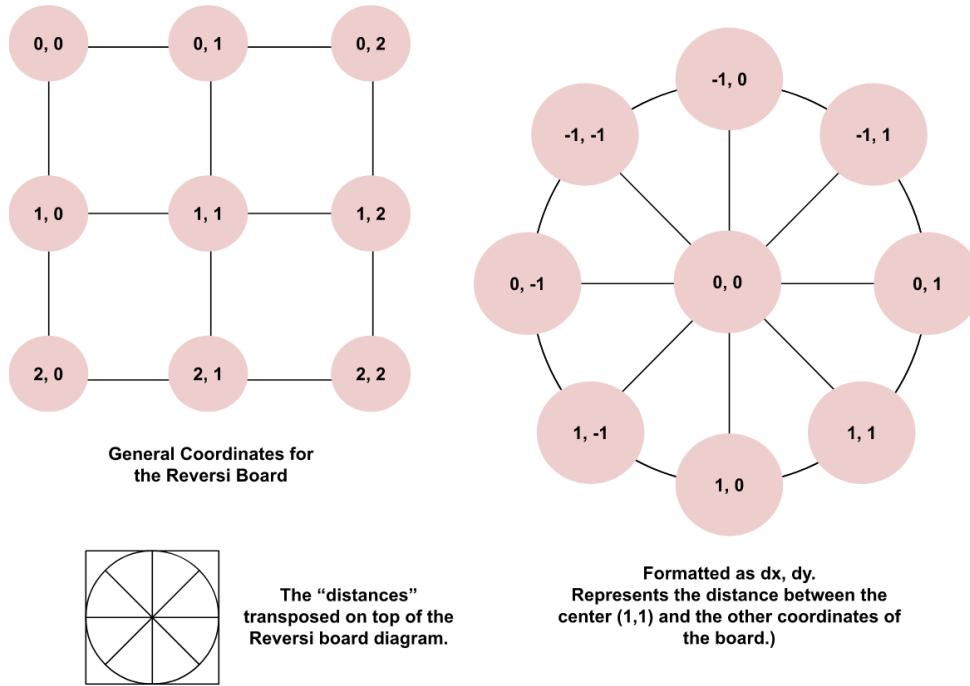


Figure 10: Diagram Explanation of the Correlation Between **dx** and **dy** and Direction

```
# Checks all directions of a specific disk position: returns zero if nothing
should be flipped or returns the number of items to be flipped in this direction
def check_dir(board, r, c, color, dx, dy):
    count = 1
    r = r + dy
    c = c + dx
    opp = "B" if color == "W" else "W"
    if (c > -1 and c < 8 and r > -1 and r < 8 and board[r][c] == opp):
        r = r + dy
        c = c + dx
        while (c > -1 and c < 8 and r > -1 and r < 8 and board[r][c] == opp):
            r = r+dy
            c = c+dx
            count = count+1
        if (c > -1 and c < 8 and r > -1 and r < 8 and board[r][c] == color):
            return count
    return 0
```

Figure 11: *check\_dir* Function Definition

## 2. *legal\_moves*

*legal\_moves* first generates an empty variable called *legals*, a two-dimensional list that contains the coordinates of a player's legal move set. This function then iterates through all of the board. If a position in *board* is empty and a disk placed in that position can swap at least one disk (based on the output of the *check\_dir* function in all directions), then the current element in *board* would be added to the *legals* list. After all iterations, *legal\_moves* returns the *legals* list.

```
def legal_moves(board, color): # Returns a list of legal moves
    legals = []
    for r in range(0, 8):
        for c in range(0, 8):
            if (board[r][c] == " " and (check_dir(board, r, c, color, -1, -1)
                                         or check_dir(board, r, c, color, -1, 0)
                                         or check_dir(board, r, c, color, -1, 1)
                                         or check_dir(board, r, c, color, 0, -1)
                                         or check_dir(board, r, c, color, 0, 1)
                                         or check_dir(board, r, c, color, 1, -1)
                                         or check_dir(board, r, c, color, 1, 0)
                                         or check_dir(board, r, c, color, 1, 1))):
                legals.append([r,c])
    return legals
```

Figure 12: *legal\_moves* Function Definition

### 3. *make\_move*

The ***make\_move*** function takes three parameters: the gameboard, the color in play, and a specific position for the next move of that color. This function first adds a disk of the given color to the board at that position. Then, it iterates through all of the possible positions around the ***position*** coordinates, which are passed as parameters using two nested for loops that each iterate three times. An iteration represents a different set of coordinates with a value of **-1**, **0**, or **1** for the “x” and “y” coordinates, and each iteration represents a different combination of coordinates. Within the body of these nested **for** loops, the conditional, which states “if **i** does not equal **1** or **j** does not equal **1**,” is used to eliminate the placement of the token on the **[0, 0]** node (Figure 11), since the first line of code in the function already accounts for this position.

Aside from this situation, the body of the **for** loop is executed. A variable **n** is first assigned to the number of swaps possible in the given direction. Once this value is assigned to **n**, two new variables are created: **col** and **row**. **col** is assigned to the second number in the **position** coordinate, and **row** is assigned to the first number in the **position** coordinate. Then, the program enters another **for** loop, which will iterate through the list of swappable disks and swap their color. Essentially, the ***make\_move*** function places the **color** on the **board** at a certain **position** and then flips all of the possibly swappable disks with that move.

```
def make_move(board, color, position): # Places the color on the board at a certain position
    board[position[0]][position[1]] = color
    for i in range(0, 3):
        for j in range (0, 3):
            if (i != 1 or j != 1):
                n = check_dir(board, position[0], position[1], color, i-1, j-1)
                row = position[0]
                col = position[1]
                for k in range(0, n):
                    col = col + i - 1
                    row = row + j - 1
                    board[row][col] = color
```

Figure 13: *make\_move* Function Definition

## E. Scoring Functions

### 1. *get\_score*

The ***get\_score*** function uses the global variables **w\_score** and **b\_score** to hold the number of white disks and black disks on the board, respectively. They are set as global variables so that they can be accessed everywhere in the program, and these variables are initially set equal to zero because the score is “refreshed” after every turn due to constant swapping. The function iterates through all of the elements in the two-dimensional array **board**, the parameter of the function, and checks to see if any elements are equal to the strings “**W**” or “**B**.” A “**W**” string results in the addition of one point to **w\_score**, and a “**B**” string results in the addition of one point to **b\_score**. Moreover, the function prints out each color’s current score.

```
# Determines score by summing W and B disks on the board
# Set scores to 0 each time so points lost are not added upon / to new score
def get_score(board):
    global w_score
    global b_score
    w_score = 0
    b_score = 0
    for r in range(0, 8):
        for c in range(0, 8):
            if board[r][c] == "W":
                w_score += 1
            if board[r][c] == "B":
                b_score += 1
    print(f"White Score: {w_score}")
    print(f"Black Score: {b_score}")
```

**Figure 14:** *get\_score* Function Definition

## 2. *check\_end\_of\_game*

The program uses this function to check whether or not the game has ended. Because Reversi ends if one or both of the players has no more legal moves, *check\_end\_of\_game* only returns **True**, meaning that the game *has* ended, if either the length of the *legals* list of the white or black player equals zero. This function thus determines the lengths by calling the *legal\_moves* function twice and checking its length. If neither list is empty, however, *check\_end\_of\_game* returns a **False** statement, meaning that the game is still ongoing.

```
# Checks whether it is the end of the game by evaluating whether there are no legal
moves left for either player
def check_end_of_game(board):
    return (len(legal_moves(board, "B")) == 0) or (len(legal_moves(board, "W")) == 0)
```

**Figure 15:** *check\_end\_of\_game* Function Definition

## 3. *check\_win*

Before this function finds the winner of the game, it first ensures that the game has ended by calling the *check\_end\_of\_game* function. If it has not ended, **None** is returned. *check\_win* then accesses *w\_score* and *b\_score* and compares them to each other. If *w\_score* is greater than *b\_score*, then the function returns the string “**White**,” indicating that the white player has won the game. If *b\_score* is greater than *w\_score*, then the function returns the string “**Black**,” indicating that the black player has won the game. An equivalent black and white score would result in the return of a “**Draw**” – neither player has won.

```
def check_win(board): # Evaluates who is the winner of the game
    global w_score
    global b_score
    if check_end_of_game(board) is False:
        return None
    if w_score > b_score:
        return "White"
    elif b_score > w_score:
        return "Black"
    else:
        return "Draw"
```

Figure 16: *check\_win* Function Definition

## III. Minimax Algorithms and Recursion

### A. Theory

#### 1. Minimax Algorithms

Minimax algorithms can explore all options in a turn-based strategy game through the creation of a tree, where each row or level represents the possible moves of each player. The rows alternate between each player's moves, depending on which player would move during that turn. One of the players would aim to maximize their score in the positive direction, while the other player would aim to minimize their score in the negative direction.

Each move in the game, which is a node in the tree, is assigned a point value based on the board evaluation function. If the tree is in the perspective of the player attempting to maximize their score in the positive direction, a node with a positive number represents an advantageous move for that player (and a disadvantageous move for the opposite player), while a node with a negative number represents a disadvantageous move for that player (and an advantageous move for the opposing player). (Figure 17 shows the nodes as circles with their scores inside of them.) The Reversi program assigns the role of positive score maximization to the white player and the role of negative score maximization to the black player.

Additionally, the tree has “levels,” which are defined by the rows of moves. A row presents all of the possibilities that a given player can move within that turn. The row immediately after that row presents all of the possibilities that the *opposing* player can move during the next turn and these rows alternate based on who is currently playing. Therefore, moving down the tree by a level implies that another piece is placed on the board of the given color and it is now the other player’s turn. If the smart computer plays as black, which always plays first, the algorithm would return the move with the smallest board score after using recursion to look ahead the number of moves that corresponds to the given level and backtracking in order to “return” to its current move. Similarly, if the computer is playing as white, the algorithm would return the move with the highest board score after recursion and backtracking. Overall, the algorithm recursively looks

ahead through all of the possible moves for the given game board and player and returns the move that would be the most advantageous for that player and the most disadvantageous for the opposing player. It is called the Minimax algorithm because it *maximizes* one player's score and *minimizes* the opposing player's score<sup>6</sup>.

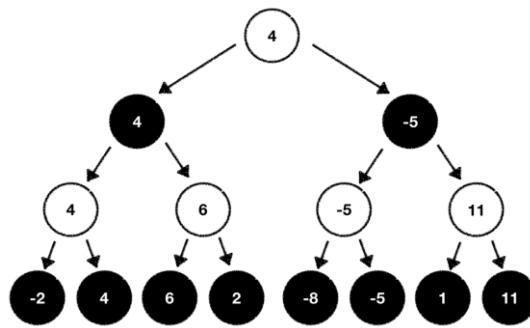


Figure 17: Minimax Tree

## 2. Recursion

Recursion is a widely used technique in the field of computer science. In terms of code, it is defining a function by calling it within itself until the desired result of the algorithm is achieved. This technique is utilized in the program's ***look\_ahead*** function, which allows the computer player to "look ahead" at varying levels to attempt to beat its opponent. This program would not have been able to be constructed without the use of recursive functions, adding to the complexity of the code.

## B. Our Implementation

### 1. Theory Behind the Board Evaluation Function

We split up the board into five regions, based on the strategy discussed earlier in the paper. As shown in Figure 18, those five regions are the corners (green), the corner-adjacent positions, (blue) the danger zone (red), the edges (yellow), and the middle 16 tiles (purple). The corners were given an extremely high positive score because they are strategically most important positions in the game. Furthermore, the edges have a lower, but still fairly high, positive score because they provide an advantage over all of the other positions on the game board, apart from the corner, and a smart computer must still be incentivized to claim these positions. The middle 16 tiles received a low positive score in order to incentivize the algorithm to choose moves with higher numbers of swaps. Additionally, the danger zone received a negative score with a small magnitude so that the smart computer does not lose a valuable corner or edge during the early game. The positions adjacent to the corners, on the other hand, received a negative score of larger magnitude to further penalize a player from potentially losing a corner to the opposing player.

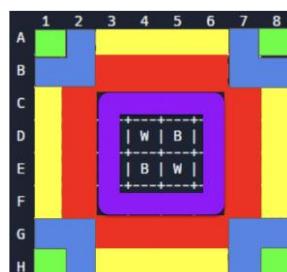


Figure 18: Zone Map

## 2. `calc_board`

The `calc_board` method assigns a numeric board value, or “score,” for the state of the board, based on the position of the disks within each region of the grid. (A position’s region dictates whether or not they have an advantage at that time in the game.) Each position has an integer weight that determines its usefulness, and the entire board of weights is stored in the two-dimensional array `weights`, as displayed in Figure 19. (The rows and columns match the true game board, so `weights` is essentially a copy of the Reversi grid.) A higher value signifies a more advantageous position, while a lower value signifies a more disadvantageous position. For example, the variable `corner`, which represents a corner, has the highest value of **50**. The variable `edge` has a fairly high value of **15**. These two regions are extremely powerful. `other`, on the other hand, represents the “safety zone” that holds positions easily taken by both sides and has a value of **1**. Moreover, the disadvantageous variables `danger` and `corner_risk` refer to the danger zone and positions adjacent to the corner respectively. The former stores a value of **-5**, and `corner_risks` stores a value of **-10**. `calc_board` iterates through `weights` and examines every position, and if that position contains a disk of the color in play, the spot evaluation is added to the color’s “board score” (`board_eval`). A disadvantageous (or negative) position value decreases `board_eval`, while an advantageous (or positive) position value increases its value. After the program has finished iterating through all of the board positions, `calc_board` returns the variable `board_eval` for the given color.

```
# calc_board Function / Board Evaluation Routine: Returns a numeric value for positions on the board
def calc_board(board, color):
    corner = 50
    corner_risk = -10
    edge = 15
    danger = -5
    other = 1

    weights = [[corner, corner_risk, edge, edge, edge, corner_risk, corner],
               [corner_risk, corner_risk, danger, danger, danger, danger, corner_risk],
               [edge, danger, other, other, other, other, danger, edge],
               [corner_risk, corner_risk, danger, danger, danger, danger, corner_risk, corner_risk],
               [corner, corner_risk, edge, edge, edge, corner_risk, corner]]

    board_eval = 0

    for r in range(0, 8):
        for c in range(0, 8):
            if board[r][c] == color:
                board_eval += weights[r][c]
            elif board[r][c] != " ":
                board_eval -= weights[r][c]

    return board_eval
```

Figure 19: `calc_board` Definition

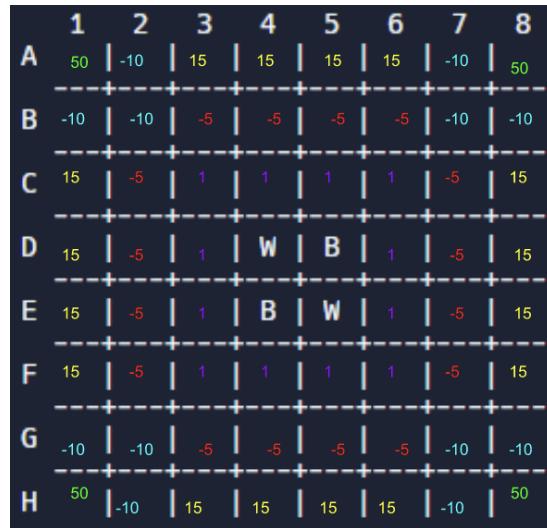


Figure 20: Reward and Penalty Map

### 3. *flip\_board*

**flip\_board** “swaps” the colors of all the disks on the board from one playing color to the other. This function uses a nested **for** loop in order to examine each board position of the two-dimensional array **board**. If an element is marked “**W**” (and has a white disk in the corresponding position), **flip\_board** will swap that element’s color to “**B**”. Similarly, it will also change the element to “**W**” if it is already marked as “**B**”. These actions allow the program to flip the game board so that the minimax algorithm can determine the best move for the smart player.

```
# Method used for swapping (turns every W to B and every B to W)
def flip_board(board):
    for r in range(0, 8):
        for c in range(0, 8):
            if board[r][c] == "B":
                board[r][c] = "W"
            elif board[r][c] == "W":
                board[r][c] = "B"
```

Figure 21: *flip\_board* Function Definition

### 4. *look\_ahead*

The **look\_ahead** function returns the best value for any possible move. The effectiveness of this function is based on the level of the smart computer. If it is level **0**, the program does *not* look ahead – it plays greedily. The computer would try to gain the most points each turn by choosing the position that swaps the most disks to its color. Otherwise, a **look\_ahead** level greater than **0** would set the **best\_val** to **-999**, meaning that the best move available would always be much lower than the *first* move found, and iterate over the list of legal moves for the opposite color currently playing. For each legal move, a deep copy – an

entirely new game board, disconnected from the original board, with the same values – is made of the given game board and assigned to the ***temp\_board*** variable. The ***flip\_board*** function is then called on this copy and flips the white and black disks. Afterwards, the function makes the given move from the list of legal moves, using the given color. If the move ends the game, a value of **-1000** is returned because it creates a loss for the algorithm. The program would otherwise look ahead a given number of levels, find the largest value from the recursive call over the entire list of legal moves, store it in ***best\_val***, and return the negative version of that value because the board evaluation of the color at this level is the negative version of the original color's board evaluation value.

```
def look_ahead(board, level, color): # Returns the best value for any possible move
    if (level == 0):
        return calc_board(board, color)
    else:
        legal_move_list = legal_moves(board, "W" if color == "B" else "B")
        best_val = -999

        for my_move in legal_move_list:
            temp_board = copy.deepcopy(board)
            flip_board(temp_board)

            make_move(temp_board, color, my_move)

            if check_end_of_game(temp_board):
                return -1000

            val = look_ahead(temp_board, level - 1, color)
            if val > best_val:
                best_val = val

    return -best_val
```

Figure 22: *look\_ahead* Function Definition

##### 5. *smart\_turn*

***smart\_turn*** provides the grid coordinates for a smart computer's next disk placement during its turn. The function first creates the list of legal moves for the inputted color and then shuffles the order of elements so that the smart computer does not necessarily choose the same legal move for a turn with equivalent board evaluation. It then instantiates the variable ***best\_val*** to **-1001** and creates a list called ***best\_move***. ***best\_val*** is first set as an extremely low number so that any initial board position value should be higher than the initialized value. (Position values are later tested to see if they are the best value through comparison with the previous highest board position value in the list. However, if the program is examining the first position value in the list, it should automatically be the best value, hence the extremely low initial value for ***best\_val***.) The function then iterates through ***legal\_move\_list*** and creates a temporary full copy of the board through ***deepcopy***. It makes a move on the temporary board with the move specified by the list and then checks if that move generates a win so that it can return that move. If it is not a winning move, ***smart\_turn*** calls the ***look\_ahead*** function and determines the highest position value for a given level of ***look\_ahead***. After finding the best value, it compares that value to the best value yielded by the previous legal moves in the

list (from that turn). If it *is* greater than the best value, the computer would use backtracking to find the value that leads to that “future value.” The turn that acts as a “stepping stone” to the best value in that future turn is now ***best\_val***, and a copy of the list ***my\_move*** is set equal to ***best\_move***. The best move found in the function is returned.

```
def smart_turn(board, level, color): # Returns the best legal move
    if color == "W":
        legal_move_list = legal_moves(board, "W")
    elif color == "B":
        legal_move_list = legal_moves(board, "B")

    random.shuffle(legal_move_list)
    best_val = -1001
    best_move = []

    for my_move in legal_move_list:
        temp_board = copy.deepcopy(board)
        make_move(temp_board, color, my_move)

        if check_win(temp_board):
            return my_move

        val = look_ahead(temp_board, level, color)
        if val > best_val:
            best_val = val
            best_move = copy.copy(my_move)

    return best_move
```

Figure 23: *smart\_turn* Function Definition

## IV. Experimental Data

Once our code was written, we needed a way to ensure that the program was working correctly. Our next step in the process was to run multiple tests and exhaust all possibilities for error. Prior to entering the rigorous testing phase, we performed multiple quick tests of 10 trials each to ensure that the evaluation function was working as intended and then proceeded with the overall data experimentation after these tests yielded the expected results.

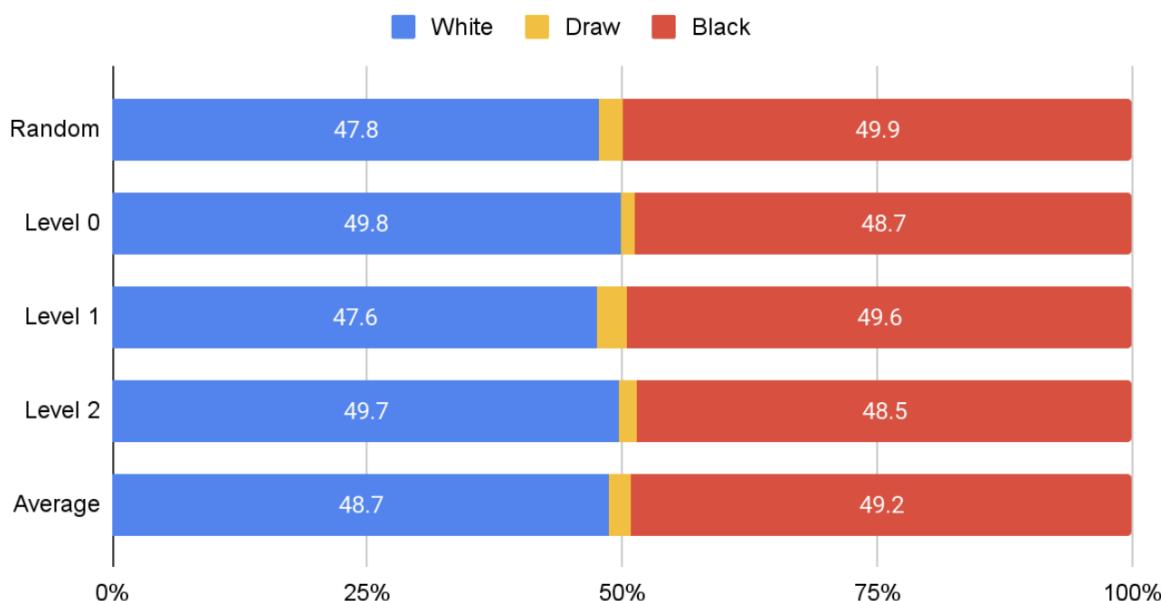
### A. Identical Intelligence Tests

We first needed to determine that the levels of intelligence of our program were working correctly against themselves before pitting them against varying levels. (Smart computers of the same intelligence level should win an equal number of times.) Additionally, we wanted to determine if the program favored one color as a winner over the other color since our program should yield higher win rates due to differences in intelligence, *not* color. We thus ran four different tests before moving onto our next phase of testing:

- Random vs. Random
- Level 0 vs. Level 0
- Level 1 vs. Level 1
- Level 2 vs. Level 2

We expected that white and black would each win about 50 percent of the time since they were using the same level of intelligence against each other. As shown in Figure 24, we received this statistic as the result that our trials yielded. For all levels of intelligence, the trial outcomes were overall split evenly between white and black wins. (A small percentage of trial outcomes were draws.) We then found the averages of the win values and plotted them alongside our levels of intelligence, and the average was also split evenly between each color. Therefore, these tests proved that our levels of intelligence were working properly against themselves.

### Trial Outcomes: Same Intelligence



**Figure 24: Comparison of Trial Outcomes with the Same Level of Intelligence**

## B. Randomized Moves vs. Minimax Algorithms

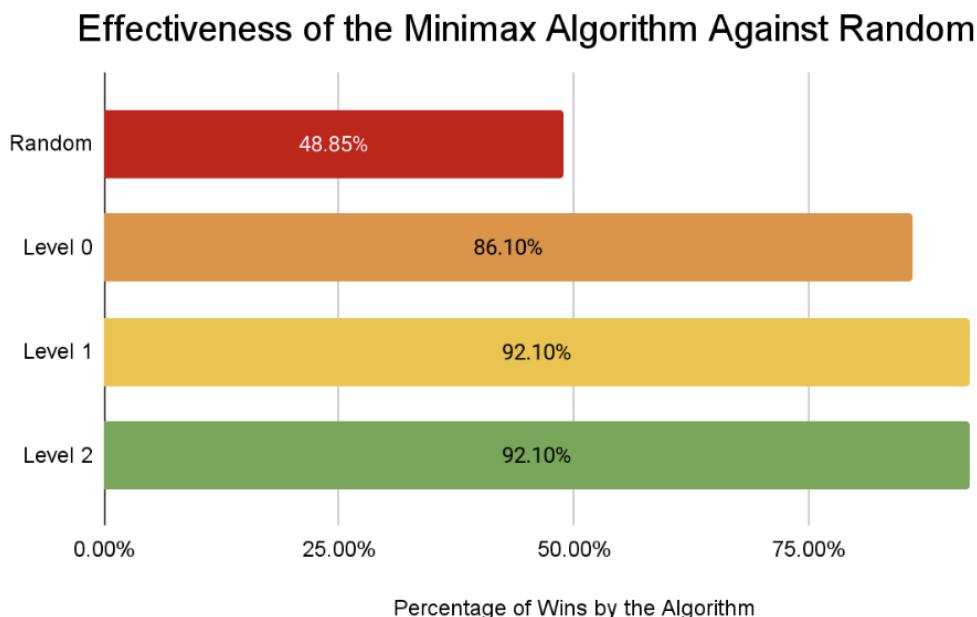
Following our trials that played games between smart computers of the same intelligence level, we decided to pit our random player against the smart computers with varying levels of recursion depth:

- Random (Black) vs. Level 0 (White)
- Random (White) vs. Level 0 (Black)
- Random (Black) vs. Level 1 (White)
- Random (White) vs. Level 1 (Black)
- Random (White) vs. Level 2 (White)
- Random (White) vs. Level 2 (Black)

Our evaluation function was decently effective in comparison to the results between two random players at level 0, which displayed an 86.1% win rate on Figure 25. As expected, the program was even more effective when playing at level 1 and level 2, as evident from its 92.1% win rate for both levels of intelligence. The

graph, however, displays no difference between levels 1 and 2, suggesting that the maximum effectiveness of this evaluation function against random moves occurs at level 1.

Moreover, we held trials with white as the smarter computer vs. black as the smarter computer in order to account for the possibility of parity. Parity is a phenomenon that originates from Othello, where white would generally win a significantly higher percentage of games than black because white would always play last. Since the game ends when one player has no moves remaining in Reversi, the effect of this phenomenon is reduced – if not eliminated – because the game has a relatively high chance of ending before the board is full.



**Figure 25: Comparison of Effectiveness of Minimax at Different Levels to Random Moves**

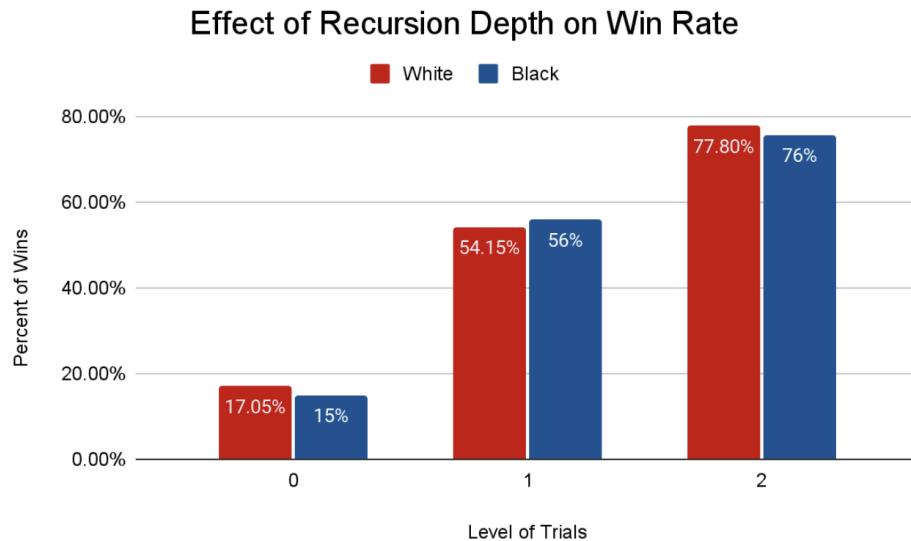
### C. Effect of Recursion Depth on Win Rate

We also ran trials to examine our varying recursion depths. In order to determine whether or not our smart computer player was able to accurately “look ahead” at varying levels and perform well against its lower intelligence opponent, we ran a wide range of these tests:

- Level 0 (White) vs. Level 2 (Black)
- Level 0 (Black) vs. Level 2 (White)
- Level 0 (White) vs. Level 1 (Black)
- Level 0 (Black) vs. Level 1 (White)
- Level 1 (White) vs. Level 2 (Black)
- Level 1 (Black) vs. Level 2 (White)

Each test prompted the white and black players to choose their moves based on different levels of recursion, or varying levels of “looking ahead” and predicting future moves. We ran 5 trials of 200 games each for every recursion depth combination, resulting in 6000 total games. 5913 of those games were conclusive with a clear winner, while 87 of these games resulted in a draw.

Figure 26 then shows that the win rate increased with recursion depth, which was expected. If the “computer” player is looking ahead more turns, it would make smarter moves and therefore win a game of Reversi more frequently. We also tested both black and white as the “smarter” player to avoid parity. Based on the data, however, parity is basically nonexistent in our program because the difference between white and black wins for each level is at a maximum of 0.68% over the 6000 trials created.



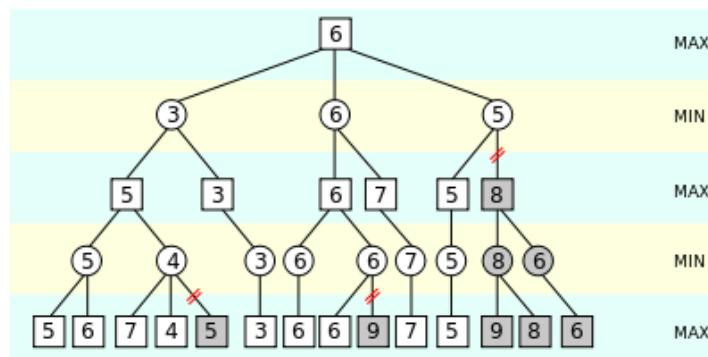
**Figure 26: Comparison of Varying Recursion Depth on Win Rate**

## V. Future Improvements

Given more time and resources, the program could have included further extensions.

### A. Alpha-Beta Pruning

Alpha-Beta Pruning is an optimization technique in the Minimax algorithm that “prunes” the game tree by cutting off branches that do not need to be searched. Examining these moves would be futile because a “better” move has already been determined for that player. For instance, if the smart player is playing as white, and therefore aiming to maximize its score, the program would not need to recurse over a subtree tree that the smart computer knows will not contain its optimal move. (Figure 27 displays this idea by cutting the tree with red lines and shading unused nodes.) This change would drastically improve the memory and time efficiency of the game<sup>7</sup>.



**Figure 27: Alpha-Beta Pruning Visualization**

## B. Experimentation with Weights

Another area of possible research is experimentation with different sets of weights for the evaluation function, which would determine the efficiency of the smart AI. The proper weight values could then be determined and used so that the program would win most games without needing to look ahead a significant number of moves. (This improvement would obviously reduce runtime.)

## C. Neural Networks and Deep Learning

The experiments could be optimized by using Neural Networks and Deep Learning (Figure 28) so that the system can be trained to have the best weights for any given level of recursion.

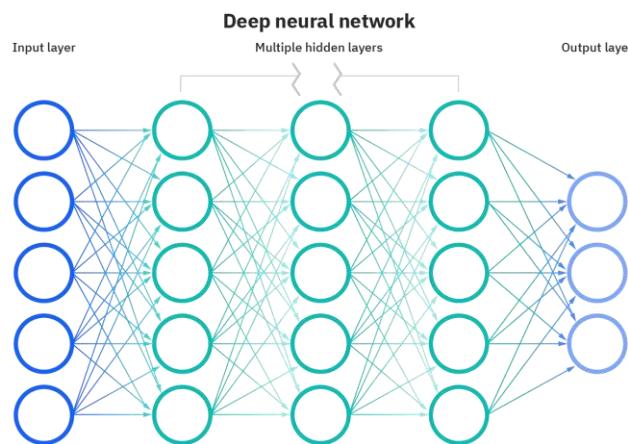


Figure 28: Neural Network Visualization

## VI. Acknowledgments

We would like to extend our gratitude to our team project advisor, Dr. Neil Simonetti, and our Teaching Assistants, Joseph McLaughlin and Michael Rybalkin, for their constant support and encouragement throughout this project's journey.

We would also like to thank Dr. Barry Luokkala, the Pennsylvania Governor's School for the Sciences Alumni Association, and Carnegie Mellon University for providing us with the resources and facilities to conduct this research.

## Appendix A: Links

The code for this project can be found on GitHub:

[https://github.com/irastogi1349342/PGSS\\_2022\\_Reversi](https://github.com/irastogi1349342/PGSS_2022_Reversi)

The experimental data sheet is located on Google Sheets:

[https://docs.google.com/spreadsheets/d/1kltr2aXk7vPtHFHS04VW\\_qG3JXNwQfrFE5jR75nb7BE/edit?usp=sharing](https://docs.google.com/spreadsheets/d/1kltr2aXk7vPtHFHS04VW_qG3JXNwQfrFE5jR75nb7BE/edit?usp=sharing)

## References

- <sup>1</sup> Group, U. C. B. G. C. R. (n.d.). GamesCrafters. Retrieved July 28, 2022, from <http://gamecrafters.berkeley.edu/games.php?game=othello>
- <sup>2</sup> Home. Reversi. (n.d.). Retrieved July 28, 2022, from <http://www.cyningstan.com/game/73/reversi>
- <sup>3</sup> JK Webdesign, www.jk.nl. (n.d.). *Othello News*. World Othello Federation. Retrieved July 28, 2022, from <https://www.worldothello.org/news/47#:~:text=Aesthetic%20differences%3A%20Othello%20is%20always,in%20different%20colors%20and%20styles>
- <sup>4</sup> wikiHow. (n.d.). *How to play othello: 15 steps (with pictures)*. wikiHow. Retrieved July 28, 2022, from <https://www.wikihow.com/Play-Othello>
- <sup>5</sup> Plarium. (2022, May 19). *RTS vs TBS: What are they and how are they different? - plarium*. plarium.com. Retrieved July 28, 2022, from <https://plarium.com/en/blog/rts-vs-tbs/>
- <sup>6</sup> *Minimax algorithm in Game theory: Set 1 (introduction)*. GeeksforGeeks. (2022, June 13). Retrieved July 28, 2022, from <https://www.geeksforgeeks.org/minimax-algorithm-in-game-theory-set-1-introduction/>
- <sup>7</sup> *Minimax algorithm in Game theory: Set 4 (alpha-beta pruning)*. GeeksforGeeks. (2021, August 18). Retrieved July 28, 2022, from <https://www.geeksforgeeks.org/minimax-algorithm-in-game-theory-set-4-alpha-beta-pruning/>
- <sup>8</sup> *Buy reversi ultra+ for PC & xbox - microsoft store en-ke*. Microsoft Store. (n.d.). Retrieved July 28, 2022, from [https://www.microsoft.com/en-ke/p/reversi-ultra-for-pc-xbox/9nztktc7zb5x?cid=msft\\_web\\_gamesforwindowsphone\\_chart](https://www.microsoft.com/en-ke/p/reversi-ultra-for-pc-xbox/9nztktc7zb5x?cid=msft_web_gamesforwindowsphone_chart)
- <sup>9</sup> By: IBM Cloud Education. (n.d.). *What are neural networks?* IBM. Retrieved July 28, 2022, from <https://www.ibm.com/cloud/learn/neural-networks>
- <sup>10</sup> Caprasi, C. (2022, June 20). *The power of recursion*. Medium. Retrieved July 28, 2022, from <https://blog.devgenius.io/the-power-of-recursion-b15910be9ee7>
- <sup>11</sup> *File:visual studio code 1.35 icon.svg*. Wikimedia Commons. (n.d.). Retrieved July 28, 2022, from [https://commons.wikimedia.org/wiki/File:Visual\\_Studio\\_Code\\_1.35\\_icon.svg](https://commons.wikimedia.org/wiki/File:Visual_Studio_Code_1.35_icon.svg)
- <sup>12</sup> Lien, X. (2021, December 14). *Solving a word game with minimax*. Medium. Retrieved July 28, 2022, from <https://towardsdatascience.com/solving-a-word-game-with-minimax-68695c5e963>
- <sup>13</sup> Marcin. (2018, June 4). *Reversi and othello – two different games. do you know their different rules?* Bona Ludo. Retrieved July 28, 2022, from <https://bonaludo.com/2016/02/18/reversi-and-othello-two-different-games-do-you-know-their-different-rules/>
- <sup>14</sup> *Minimax and Monte Carlo Tree Search*. Philipp Muens. (n.d.). Retrieved July 28, 2022, from <https://philippmuens.com/minimax-and-mcts>
- <sup>15</sup> Olivet Nazarene University Digital Commons @ Olivet. (n.d.). Retrieved July 29, 2022, from [https://digitalcommons.olivet.edu/cgi/viewcontent.cgi?article=1014&context=engn\\_stsc](https://digitalcommons.olivet.edu/cgi/viewcontent.cgi?article=1014&context=engn_stsc)
- <sup>16</sup> *Tips to win othello*. UltraBoardGames. (n.d.). Retrieved July 28, 2022, from <https://www.ultraboardgames.com/othello/tips.php#:~:text=Watch%20Out%20For%20The%20%20danger,indicate%20this%20%22danger%20zone%22>
- <sup>17</sup> Wikimedia Foundation. (2022, July 6). *Alpha–beta pruning*. Wikipedia. Retrieved July 28, 2022, from <https://en.wikipedia.org/wiki/Alpha%20Beta%20Pruning>
- <sup>18</sup> Wikimedia Foundation. (2022, July 17). *Replit*. Wikipedia. Retrieved July 28, 2022, from <https://en.wikipedia.org/wiki/Replit>
- <sup>19</sup> Wikimedia Foundation. (2022, July 27). *Python (programming language)*. Wikipedia. Retrieved July 28, 2022, from [https://en.wikipedia.org/wiki/Python\\_\(programming\\_language\)](https://en.wikipedia.org/wiki/Python_(programming_language))

# **Artificial Intelligence Architecture for Turn-Based Strategy Games: Nine Men's Morris**

Nikola Cao, Eric Cui, Samuel Wang, Daniel Zhang

## **Abstract**

Advances in the field of artificial intelligence (AI) in recent decades have allowed computers to solve previously unsolvable problems. AI allows programs to analyze novel situations and determine how best to achieve their goals based on provided information. Naturally, researchers have sought to combine this technology with turn-based games in an effort to create programs capable of analyzing the nearly countless possible moves in such games and play well enough to beat even the best human players. As early as the 1990s, Google had already created a chess AI that defeated world champion Garry Kasparov<sup>1</sup>. Our team project's goal was to create a Nine Men's Morris AI capable of imitating human play. To accomplish this we employed a minimax algorithm, which assigns point values to board states based on how favorable a position is and selects the move that confers the largest advantage.

## **I. Introduction**

### **A. Background**

#### **1. Artificial Intelligence**

Among the most exciting developments in computer science is the field of AI. It is clear why: the potential applications are endless, ranging from chatbots capable of providing instant customer support to the aiding of researchers through information analysis. Although humanity has dreamed about artificial beings capable of intelligent thought since Greek times<sup>2</sup>, scientists only began truly discussing the possibility of artificial intelligence around the 1940s. British computer scientist Alan Turing laid much of the groundwork for modern AI research and proposed a definition for a thinking machine with his famous Turing test: if a human has a conversation with a robot and cannot distinguish them from a human, the robot is intelligent.<sup>3</sup> Later on, in 1956, a Dartmouth Workshop officially coined the term “artificial intelligence” and is considered by many to be the origin of the field. In the following decades, enthusiasm for AI peaked and successful demonstrations of problem solving led to many projects being funded. However, while primitive implementations of AI existed, numerous obstacles, the most significant of which being hardware limitations, stood in the way of truly intelligent programs. By the late 1990s, however, an AI capable of defeating a chess champion in chess had already been created<sup>1</sup>. In the past few years, AI has only become more prevalent and powerful. It has become integrated into almost every industry, and will surely continue to improve lives in the future.

#### **2. Turn-based Games**

Since ancient times, turn-based games have been a prevalent part of every human civilization. They act as important social and cultural bonding events which can also serve a variety of purposes, including recreation and mental development. These games often pit two players against each other and require a mixture of luck and skill. The oldest known board game is Senet, which originated in ancient Egypt and represented the passing from life to the afterlife. It involved two players racing to move all their pawns off the board first using strategy and a little luck.<sup>4</sup> Perhaps the most well known board game is chess, which is played by millions of people around the world. Chess relies solely on skill, so the outcome is completely determined by which player plays better. This has made it an incredibly competitive game with international

tournaments, as well as the focus of numerous studies attempting to maximize the odds of winning. Nowadays, computers can beat even the best human players decisively.<sup>5</sup>

## II. Nine Men's Morris

### A. History

Nine Men's Morris has been popular since Ancient Rome and the medieval period, although it is unknown exactly when and where it was created. Nine Men's Morris boards can be found carved into stone buildings throughout the Roman Empire and in many medieval cathedrals. The game has been played continuously since the Roman period.<sup>6</sup>

### B. The Board

A Nine Men's Morris board consists of three concentric squares with three playable slots on each of their sides. There are eight playable positions on each square resulting in a total of 24 playable slots. All slots on each side of the square are connected with a line with the middle slot of each square's side additionally being connected. A standard board configuration is shown in Figure 1.

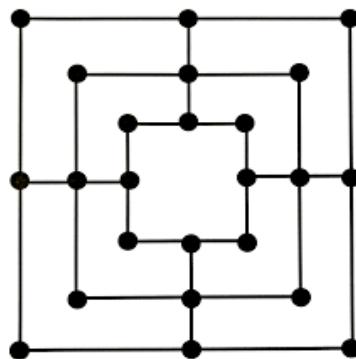


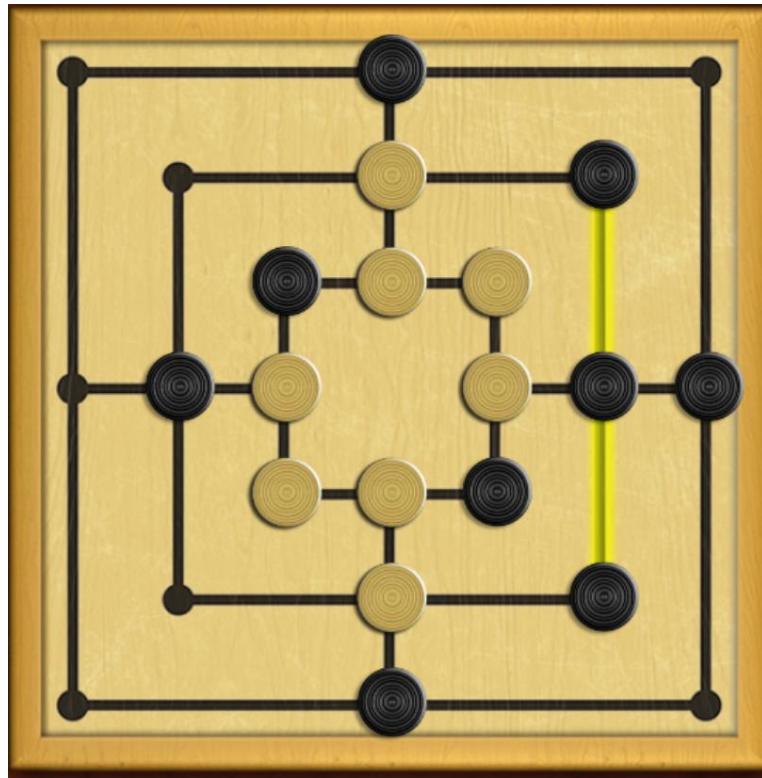
Figure 1: An Empty Nine Men's Morris Board

### C. Rules

Nine Men's Morris is a competitive two-player game and is most commonly played with two game phases. The first phase involves placing pieces on empty board spaces while the second phase involves moving existing pieces to adjacent open spaces on the board. Achieving three pieces in a row allows a player to remove an opposing piece from the board. Players win by either reducing their opponent to two pieces or by trapping the opponent's pieces, ensuring they have no legal moves.

#### 1. Phase 1 (Placing)

In Phase 1, each player starts with nine pieces and alternates placing pieces on the board. The game is traditionally played with black and white pieces, and the player playing white goes first. A player can place their piece anywhere on the board, as long as that position is not already occupied. If a player manages to place three pieces in a row horizontally or vertically (known as creating a mill), they are allowed to remove an opposing piece from the board. Phase 1 concludes when each player has placed all nine of their pieces. It is worth noting as well that players are not given replacements for removed pieces, meaning it is possible a player will end the first phase with less than nine pieces on the board. Figure 2 depicts a possible board state at the end of Phase 1.



**Figure 2: A Populated Nine Men's Morris Board Following Phase 1**

## **2. Phase 2 (Moving)**

In Phase 2, no additional pieces are placed on the board and players are only allowed to move pieces already on the board. Each player takes turns shifting their pieces and can only move their pieces to an adjacent position that is not occupied. Players move one piece per turn. The goal of the game is now to move pieces and create mills to remove enemy pieces. Players can destroy and rebuild mills by shifting a piece out of the mill and moving the piece back in on the next turn. This can be done any number of times, and the player performing this can remove an opposing piece each time the mill is rebuilt. Players win by reducing their opponent to two pieces or by trapping their opponent's pieces so they have no more legal moves.

## **3. Phase 3 (Flying)**

Some Nine Men's Morris games may choose to implement a third phase, known as the "flying" phase. This phase occurs when a player has been reduced to three pieces. Once this happens, that player is no longer restricted to moving their pieces to adjacent positions and can now "fly" them across the board and place them in any unoccupied position. The player who still has more than three pieces is still restricted to only moving their pieces to adjacent positions. The winning conditions of the game slightly change with the inclusion of the third phase. The only winning condition is now reducing a player to two pieces. The game can no longer end when there are no legal moves, since there will always be a legal move if a player is allowed to "fly."

## D. Basic Strategy

### 1. Piece Placement

Proper piece placement in the first phase is vital to developing an advantageous board state in the second phase. While most novice players focus on trying to form mills to reduce the number of pieces the enemy has, it is actually more favorable for a player to place their pieces for movement later in the game. If a player only goes for mills in the first phase, they force the enemy to constantly try to block these mills. As a result, a player's pieces may be restricted as they are surrounded by enemy pieces. Therefore, the most advantageous strategy during Phase 1 is to place pieces in positions that maximize versatility. For example, a corner space on any square has only two adjacent spaces while midpoints have four adjacent spaces. In addition, the player should try to spread their pieces out among the squares.

### 2. Blocking

A popular strategy is to repeatedly move a piece out and back into a mill. This allows players to remove an enemy's piece every two turns. To prevent this, the opposing player should attempt to position their pieces so that they are able to block a broken mill. If none of the player's pieces are near the enemy's mill and they are unable to block it, the player can try to form their own mill and force the enemy to choose whether to block the player's mill or continue reforming theirs. Alternatively, the opposing player can seek to destroy any opportunities for creating a rebuildable mill by forming their own mill and strategically removing their opponent's pieces.

## III. Coding Language and Environment

This study's scripts were coded entirely in Python as the language is dynamically typed, making it extremely quick to code up the project's logical structure. Scripting was done on Replit.com, a free online integrated development environment, for its ease of use and simple real-time collaborative editing capabilities. Version control was natively tracked in Replit, with Github being employed for large-scale code storage.

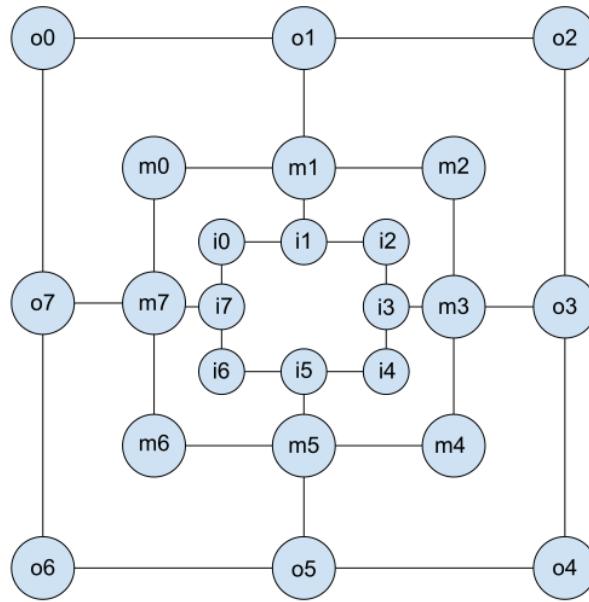
## IV. Creating The Game

### A. Translating The Game into Code

All of the code utilized in this project can be found on Github at the link located under Appendix T.

#### 1. Describing the Board

The first step to intelligently automating Nine Men's Morris was to translate the game into a representative language a computer can understand. To start, we developed a coordinate system for the board. Each of the 24 playable positions can be described by the square they are located on as well as their slot on the square. From this, a two-character notation was developed. The characters "o", "m", and "i" were used to describe the outer, middle, and inner squares respectively, and the integers 0 - 7 were used to describe the eight slots of each square progressing in the clockwise direction. Figure 3 depicts this notation scheme.



**Figure 3: Board Notation Scheme**

The characters denoting the regions for the simple notation were translated into numbers (“o” = 0, “m” = 1, “i” = 2) while being fed into computational methods.

## 2. The Point Class

A **point** class was created to store slot information. Each playable slot was converted into such an object. Points store a “region” (the index of the square they are located on), a “position” (the position on the square they are located on), and an “occupied” state. The occupied variable describes if a position is occupied (0 represents unoccupied, 1 represents occupied by player 1, and 2 represents occupied by player 2). This class was the basis for much of the game’s evaluative and algorithmic logic.

## 3. The Board Class

A **board** class was created to store both the logical structure of a game board and any required supporting methods. This class stored 24 points and tracked game information such as the number of pieces placed by each player and the phase of the game. The class also stored logic to detect mills, return and make legal moves, detect phase changes and win conditions, and “flip” the board by changing all white pieces to black pieces and vice versa.

## B. Moving Pieces

Due to the different phases of the game, the way in which pieces are allowed to be placed or moved changes. Therefore, separate functions were employed to move the pieces in Phase One as opposed to Phase Two. It is worth noting as well that the optional Phase 3 (the “flying” phase) was not implemented due to its scripting and logical complexity.

### 1. Phase One Methods

As mentioned previously, in Phase 1 players are allowed to place their pieces on any unoccupied position on the board. To achieve play in Phase 1 and track phase changes, a few different functions were needed:

- **legal\_move** returns if a position is unoccupied or not.
- **p1\_legal\_moves** iterates through the board and returns a list of all legal Phase 1 moves (in Phase 1, legal moves consist of unoccupied slots). (See Appendix A for an in-depth description)
- **p1\_make\_move** takes in region, position, and the team playing and places a piece in the corresponding slot. (See Appendix B for an in-depth description)

## 2. Phase Two Methods

Phase 2 requires similar but slightly altered functions to continue play:

- **p2\_legal\_moves** returns a list of all legal Phase 2 moves. This method takes in the origin coordinates of a piece a team wants to move and returns a list of adjacent spaces that piece can be moved to. (See Appendix C for an in-depth description)
- **p2\_make\_move** allows players to make a Phase 2 move. This function takes in an origin coordinate (representing the piece being moved) and a destination coordinate (representing the position that piece is being moved to). The legality of all Phase 2 moves are checked against legal\_move and p2\_legal\_moves. (See Appendix D for an in-depth description)

## C. Supporting Methods

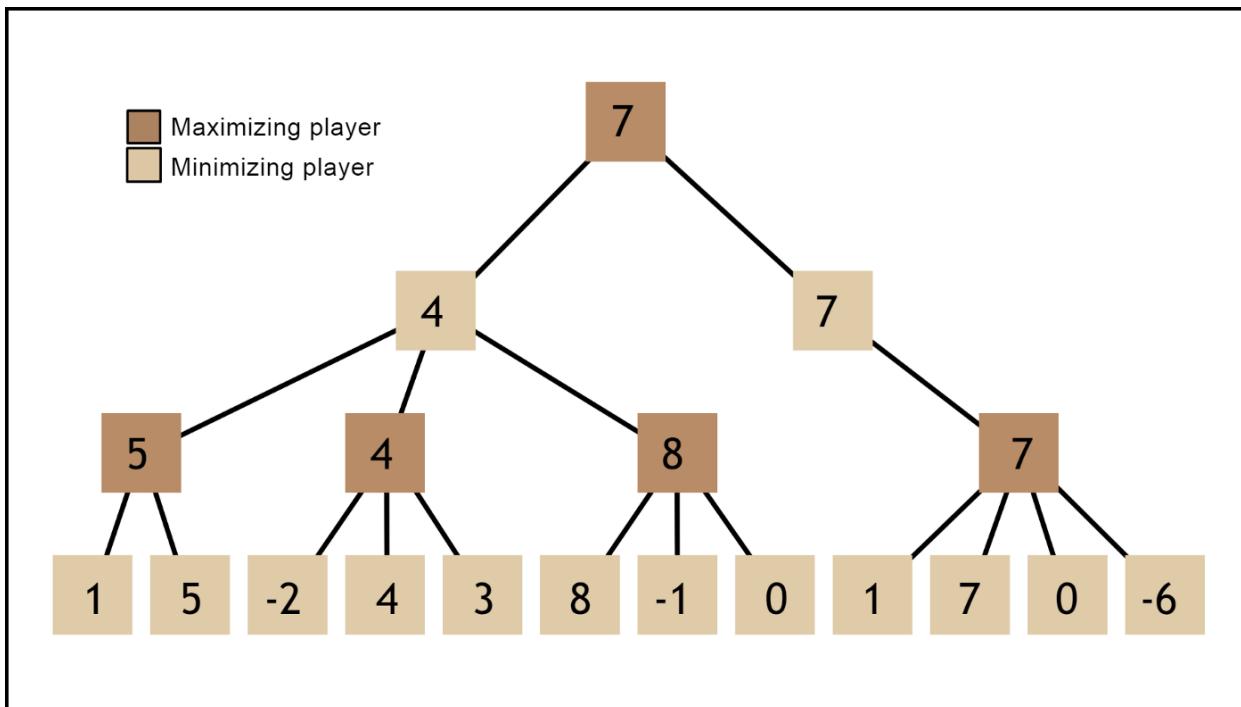
Various supporting methods were created to regulate the state of a game and track important pieces of game data:

- **piece\_counts** tracks the number of pieces each player has placed to regulate transitioning into Phase 2.
- **print\_board** prints a formatted Nine Men's Morris board using ASCII characters for the purposes of game visualization and debugging.
- **mills\_at** utilizes modular arithmetic to determine whether or not a certain occupied position is part of a mill. It returns false for no mills, and true for one or two mills. (See Appendix E for an in-depth description)
- **remove\_piece** allows a player to remove an opposing piece and is called whenever mills\_at evaluates to true. (See Appendix F for an in-depth description)
- **is\_win\_for\_team** tracks the number of pieces the opponent has after each move and iterates over all opposing pieces to see if they have any legal moves left. If the opponent is evaluated to have two pieces or no legal moves, a win is logged for the moving team. (See Appendix G for an in-depth description)

## V. The Minimax Algorithm

A Minimax algorithm is a decision-making algorithm commonly used in two-player turn-based strategy games. It assigns numerical values to board states based on previously-set parameters such as the positions of the pieces. One player is designated as preferring large values (the maximizer), while the other prefers small values (the minimizer), including negative numbers. It is then told to look ahead a certain number of turns and makes a tree of all possible board states at that depth.

## B. Minimax Trees



**Figure 4: Minimax Tree**

Figure 4 depicts a potential Minimax tree. A minimax tree is used to visually display the process that the minimax algorithm executes when running. The algorithm is a recursive backtracking algorithm that chooses the best possible move based on what the opponent will do, assuming that the opponent picks the best possible move for themselves. The algorithm looks at every possible move and assigns the state of each board after said move a score. Assume there are two optimal computers, Max and Min, playing a turn-based game. The higher the score of a board is, the better it is for Max while the lower the score is, the better it is for Min. Figure 4 shown above represents an algorithm that goes three levels deep, in which the deepest level shows all the possible board configurations after three turns. The algorithm looks at these boards and scores each board. Since Max and Min take turns, from Figure 4, we know that Max is the player making the decision that will lead to one of the boards in the third level. The algorithm then scores the four possible boards in the second level by picking the highest values out of the possibilities given in the third level since Max will make whatever move will give it the highest score. Now that the algorithm has scored all the possible boards in the second level, it determines what the scores will be for the two boards in the first level. Since it is a turn based game, we know that Min will be making the decision in the second level. Min prefers low scores so it will choose the lowest possible score in each of the branches. Finally, the algorithm reaches the top of the tree, which is the next turn, and decides to make the move that will go down the right side of the tree. In essence, the algorithm chooses the board with score seven rather than the left side, which could lead to a board with score eight, because it assumes that the other player will not allow it to reach score eight and force it down the branch that goes to a board with score four.

## VI. Nine Men's Morris Minimax

### A. Describing a Turn

Scripting a Nine Men's Morris turn proved initially challenging as there are various turn types, all of which require different amounts of information. There is a Phase 1 turn which only requires the position that a piece is being placed at and a Phase 2 turn which requires both an origin (the position of the piece being moved) and a destination (the new position the piece is being moved to). Additionally, in the event any of these turns create a mill, additional data is needed to describe which opposing piece is being removed. To resolve this, we represented a turn as a list containing three point classes. A completely populated triplet represents a Phase 2 mill-creating move that removes an opposing piece. The first element represents the origin (coordinates of the piece being moved), the second element represents the destination of the piece, and the final element represents the piece being removed. All the possible move types can be constructed from this triplet by leaving elements empty. For example, a Phase 1 move that does not create a mill can be represented by leaving the first and last element of the triplet empty (as this move does not require an origin or data for an opposing piece being removed). Similarly, a Phase 2 move not creating a mill can be represented by leaving the last element empty. This turn object simplified inputting turn data and adjusting board states for the algorithm.

### B. Scoring the Board

Implementing a minimax algorithm into Nine Men's Morris first requires methods to score different board states for each team. To accomplish this, a **scorers** class was created to store the scoring algorithms. A simple, intermediate, and complex scoring algorithm was developed.

### C. Simple Score

A simple scoring algorithm was developed and stored in the **simple\_board\_score** method. This algorithm simply returns the difference between the number of friendly and enemy pieces. As the most common method of winning Nine Men's Morris is to reduce your opponent's pieces, this was the simplest possible form of advantage-based board scoring.

### C. Intermediate Score

Building off the simple scorer, **intermediate\_board\_score** was developed. In addition to the piece differential, this function took into account friendly and enemy premills (defined as horizontal and vertical runs of three slots containing two pieces of the same team and an empty position) and potential mills (defined as horizontal and vertical runs of three slots with one piece and two empty positions). This method utilized the following additional functions:

- **premills\_at** takes a given piece, the team playing, and the board. The function checks all the possible premills (forward, back, across regions) and adds them to a list of premills, if they exist. This list of premills is then returned. (See Appendix H for an in-depth description)
- **team\_premills** takes the team playing and the board. The function loops through all the pieces a team has on the board and sums the number of premills that each piece has. Before returning the sum, it is cut in half since the loop counts each premill twice. The function then returns this number. (See Appendix I for an in-depth description)
- **potential\_premill\_count\_at** takes a given piece and the board. The function checks all the possible potential mills (forward, back, across regions) and adds them to a list of potential mills, if they exist. The list of potential mills is then returned. (See Appendix J for an in-depth description)

- **team\_potemills** takes the team playing and the board. This function loops through all the pieces a team has on a given board and sums the number of potential mills that each piece has. The function then returns the total. (See Appendix K for an in-depth description)

## 6. Intermediate Board Scoring

To create a functional and efficient board scoring system, all the previously mentioned functions are utilized and their results combined into a board score. The program calculates and stores the values for the player's premills and potential mills, the enemy's premills and potential mills, and the difference in number of pieces. The piece differentials are numerically weighted by 10, premills by 3, and potential mills by 2. Friendly values for each of these metrics are multiplied by their respective weights and added to the running score. Enemy product values for these metrics are subtracted from the score. The total score is then returned.

## D. Complex Score

A **complex\_board\_score** was also developed. This method is functionally identical to the intermediate board scorer in that it evaluates piece differential, premills, and potential mills. This function additionally evaluates piece versatility, defined as how mobile a placed piece is (empty adjacent squares). The complex scorer also looks for wins for either team.

## E. Implementing Minimax

Once the different board scoring algorithms were developed, MiniMax was implemented through a recursion-based logical structure. This functionality was implemented in the **mini\_max** function which returns the best board score for a given depth of search. This is fed into the **MM\_move function**, which determines the best turn based on the returned mini\_max value and takes the turn.

# VII. Player Classes

## A. Human Player

A **human\_player** class was developed for the purposes of preliminary game testing, debugging, and rudimentary player versus algorithm testing.

### 1. Player Input

The preliminary function in this class was **input\_position**, and it simply received input from the player. Since we must take player input in many different functions, such as making moves and removing pieces, we made a simple function that we can call in one line to get the player's region and position inputs.

### 2. Using Mills

One of the necessary functions in this class was one that allows the player to use a mill that they've formed.

- **use\_mills** takes the team playing and if a mill was made. The function then allows the player to remove an enemy's piece from the board. (See Appendix L for an in-depth description)

### 3. Phase One Moving

The next function that was created was one that allowed a human player to make a move in Phase 1.

- **make\_p1\_move** takes which team is playing. The function allows the player to input where they would like to place a piece and then places a piece in that space. The function then checks if this move has formed a mill. If so, the function runs `use_mills`. (See Appendix M for an in-depth description)

#### 4. Phase Two Moving

The next function made was one that allowed a human player to move a piece in phase 2.

- **make\_p2\_move** takes which team is playing. The function allows a player to choose which piece they want to move. It then allows the player to choose a space to move the piece to and moves the piece there. (See Appendix N for an in-depth description)

#### B. Random Player

A `randomAI` class was developed, which makes completely random moves by evaluating all of the possible legal moves in a given board position and randomly selecting a move. Phase 1 and Phase 2 random move functions were developed. The code segments for these functions (`random_p1_move` and `random_p2_move`) are stored in the appendix. (See Appendix O and P for in-depth descriptions)

#### C. Smart Player

A `smart_player` class was developed as our artificially intelligent player. This class stored the minimax functions and took in a board scorer as a parameter. This allowed use to have the `smart_player` deploy the minimax algorithm with the simple, intermediate, or complex scoring routine.

## VII. Testing

#### A. Random Vs. Random

RandomAI was pitted against itself for the purposes of preliminary testing and gauging whether or not moving first conferred a significant advantage. Five trials of 1,000 games were run, and the results are tabulated in Table 1.

**Table 1: Random Vs. Random (Average)**

Random Vs. Random (Average)						
Data	Blue Wins (Team 1)	Red Wins (Team 2)	Total Trials	Blue Win %	Red Win %	% Difference
Average	526.8	473.2	1000	52.68	47.32	5.36

On average, blue (the team moving first) appeared to have a 5.36% higher chance of winning than red. This suggested that moving first is advantageous, and this was taken into account in the subsequent experiments by systematically alternating if the testing AI played as blue (moving first) or red (moving second).

## B. Simple Vs. Random

Following preliminary testing, the simple scoring algorithm was fed into smart\_player, and the simple AI was then pitted against randomAI. Tree Depth was varied as well as if the simple AI moved first or second. Five trials of 1,000 (split into 500 with the simple AI moving first and 500 with the simple AI moving second) were run at tree depths of 0, 1, and 2. Due to the large number of game states after each move, runtime at tree depths of 3 and above became impractically high, so 2 was the upper bound for these trials. Additionally, functionality to detect stalemates was implemented. This was determined to be games that ran over 100 turns as once this point was reached, the algorithm and random AI infinitely switched between two identical turns, making it impossible for a win condition. See Appendix Q for the raw data. The averages of these trials are tabulated in Table 2.

**Table 2: Simple Vs. Random (Averages)**

Simple Vs. Random (Averages)												
Move Order	Simple Moving 1st			Simple Moving 2nd			Totals					
Data	Simple Wins	Random Wins	Stalemates	Simple Wins	Random Wins	Stalemates	Simple Total Wins	Random Total Wins	Simple Win %	Random Win %	% Difference	
Depth = 0	497.4	1	1.6	491.2	5.8	3	988.6	6.8	98.86	0.34	98.52	
Depth = 1	498	0.4	1.6	493.8	3.6	2.8	991.8	4	99.18	0.2	98.98	
Depth = 2	494	2.8	3.2	486.4	6.8	6.8	980.4	9.6	98.04	0.48	97.56	

This experiment yielded some interesting results. It can be seen that the simple algorithm's performance increased from depths 0 to 1, but decreased from depths 1 to 2. This is hypothesized to be due to the simple nature of the algorithm's evaluation method. As the simple scorer simply attempts to maximize friendly pieces and minimize enemy pieces, faulty, unoptimized intelligence emerges at a tree depth of 2. The algorithm plays extremely conservatively, prioritizing retaining pieces over sacrificing mills for potentially higher board positions in the long term. This may resolve at higher tree depths. However, it is worth noting that the algorithm was still able to consistently beat the randomAI over 98% of the time regardless of depth, proving the algorithm is effective.

## C. Intermediate Vs. Random

The smart\_player was then fed the intermediate scoring algorithm and was once again pitted against the randomAI. Tree depth and first-move were once again varied. Five trials of 1,000 evenly split amongst the intermediate AI moving first and second were once again run at tree depths of 0, 1, and 2. Stalematting was once again tracked. See Appendix R for the raw data. The averages of these trials are tabulated in Table 3.

**Table 3: Intermediate Vs. Random (Averages)**

Intermediate Vs. Random (Averages)												
Move Order	Intermediate Moving 1st			Intermediate Moving 2nd			Totals					
Data	Int. Wins	Random Wins	Stalemates	Int. Wins	Random Wins	Stalemates	Int. Total Wins	Random Total Wins	Int. Win %	Random Win %	% Difference	
Depth = 0	494.8	0	5.2	494	0.4	5.6	988.8	0.4	98.88	0.02	98.86	
Depth = 1	496.4	0	3.6	496.2	0	3.8	992.6	0	99.26	0	99.26	
Depth = 2	498.6	0	1.4	496.6	0.2	3.2	995.2	0.2	99.52	0.01	99.51	

This experiment yielded incredibly promising results. Firstly, it can be seen that the intermediate AI outperformed the random AI at all depth levels. Secondly, it can also be seen that the intermediate AI consistently improved in performance as depth level increased, confirming the efficacy of the intermediate method of board evaluation.

## E. Intermediate Vs. Simple

Following the AI versus randomAI trials, the intermediate and scoring algorithms were then run against each other. This experiment was run to confirm that the intermediate board evaluation method is superior to the simple scorer. Five trials of 1,000 were once again run with split moves, and tree depth was tested at 0, 1, and 2. The two algorithms maintained the same tree depth for all trials. Stalemating was once again tracked. See Appendix S for the raw data. The averages of these trials are tabulated in Table 4.

**Table 4: Intermediate Vs. Simple (Averages)**

Intermediate Vs. Simple (Averages)												
Move Order	Intermediate Moving 1st			Intermediate Moving 2nd			Totals					
Data	Int. Wins	Simple Wins	Stalemates	Int. Wins	Simple Wins	Stalemates	Int. Total Wins	Simple Total Wins	Int. Win %	Simple Win %	% Difference	
Depth = 0	278.4	200.6	21	284.4	196.8	18.8	562.8	397.4	56.28	19.87	36.41	
Depth = 1	335	154.2	10.8	287	206.4	6.6	622	360.6	62.2	18.03	44.17	
Depth = 2	423.2	52	24.8	434.8	47.4	17.8	858	99.4	85.8	4.97	80.83	

This experiment confirmed the relative effectiveness of the two algorithms. The table shows that at a depth of 0, the intermediate AI had a 56% win rate while the simple AI had a 19.87% win rate (the remaining games were stalemates). Furthermore, the intermediate AI's win rate increased to 62.2% and 85.8% at depth levels of 1 and 2 respectively. This shows that the difference in effectiveness between the two scoring algorithms grew more dramatic as the depth level increased, confirming the superiority of the intermediate scoring algorithm.

## F. Complex Vs. Simple

The simple AI was additionally played against an algorithm utilizing the complex board scoring algorithm. Preliminary trials (about 100) were run at depths of 0 and 1 to evaluate experiment runtime. The complex AI surprisingly did worse than the intermediate AI against the simple AI at these depths. Run at a depth of 2 however, the complex AI performed marginally better than the intermediate AI. The data for the depth two trials of complex versus simple are shown in Table 5.

**Table 5: Complex Vs. Simple (Average)**

Complex Vs. Simple (Average)											
Move Order	Complex Moving 1st			Complex Moving 2nd			Totals				
Data	Complex Wins	Simple Wins	Stalemates	Complex Wins	Simple Wins	Stalemates	Complex Total Wins	Simple Total Wins	Complex Win %	Simple Win %	% Difference
Depth = 2	437.8	42.2	20	435.6	43.2	21.2	873.4	85.4	87.34	4.27	83.07

As shown in the table, the complex AI had a win rate 1.54% more than that of the intermediate AI. This suggests that the complex AI seems to require higher depth to serve as a feasible substitute to the intermediate AI.

## VIII. Conclusion

### A. Discussion

This project shed light on the intricacies of Nine Men's Morris and successfully digitized much of its game theory. The game proved to be an extremely complex turn-based entity, filled with numerous nuances and unique game states. From this a simple, intermediate, and complex game algorithm were developed that, in conjunction with the minimax algorithm, were consistently able to beat a randomly playing AI. Additionally, it was also shown that the intermediate AI easily outperformed the simple AI when pitted against each other, and the complex AI seemed to be more successful than the intermediate AI at a tree depth of 2 at playing the simple AI.

### B. Future Study

#### 1. Human Vs. AI Trials

Future study would include running large-scale AI versus human trials. We did not have a large enough sample of human players to run enough tests to be statistically significant, so no AI versus human experiments were designed and deployed. Small scale play tests were run by colleagues for all three algorithms at tree depths of 0, 1, and 2. Based on subjective observation and player feedback, the algorithms successfully imitated human play and exhibited enough intelligence to competently play against a human and make logical moves, such as blocking an immediate mill threat.

#### 2. Complex Vs. Intermediate Trials

Due to runtime and time constraints, the intermediate AI was not played against the complex AI. This would have been a fascinating experiment to run, especially at the lower depth levels of 0 and 1. This is because the complex AI appeared to perform worse than the intermediate AI against the simple AI at these depths,

with the complex AI slightly outperforming the intermediate AI once it reaches a depth of 2. From this, one would expect the complex AI to outplay the intermediate AI at tree depths of 2 and above.

## 2. Exploring the Board Evaluation

It is worth diving deeper into methods of board evaluation. It is still not fully understood why the complex board scorer was significantly worse than the intermediate board scorer when played against the simple AI. A deeper understanding of the board states and recursive logic would shed light on this issue. Additionally, increased experimentation (such as varying metric weights or testing new metrics) could potentially lead to more effective and optimized board scoring routines.

## 3. Diving Deeper

One of the greatest limitations to this project was computational power. With the computers made available to us (standard computer lab PCs), testing at depths greater than 2 was infeasible due to runtime. However, running experiments at trials greater than 2 would yield deeper insights into the nature of our various algorithms and how they diverge in effectiveness as they are able to search deeper.

## IX. Acknowledgements

We would like to thank our wonderful faculty advisor, Dr. Neil Simonetti, as well as our two TAs, Michael Rybalkin and Joseph McLaughlin, without whom this project never would have been completed. In addition, we'd like to thank Dr. Barry Luokkala, Carnegie Mellon University, and the PGSS faculty and staff for making this incredible program possible.

## XX. References

<sup>1</sup> SITNFlash. (2020, April 23). *The history of Artificial Intelligence*. Science in the News. Retrieved July 27, 2022, from <https://sitn.hms.harvard.edu/flash/2017/history-artificial-intelligence/>

<sup>2</sup> P. G. (2016, October 26). *Talos - giant automation in Greek mythology*. Mythology.net. Retrieved July 27, 2022, from <https://mythology.net/greek/greek-concepts/talos/#:~:text=What%20is%20Talos%3F%20Talos%20is%20a%20giant%20automaton,the%20Cretan%20dialect%2C%20much%20like%20Helios%20in%20Greek>.

<sup>3</sup> Oppy, G., & Dowe, D. (2021, October 4). *The turing test*. Stanford Encyclopedia of Philosophy. Retrieved July 27, 2022, from <https://plato.stanford.edu/entries/turing-test/>

<sup>4</sup> Preskar, P. (2021, March 21). *Senet-the world's oldest known board game*. Medium. Retrieved July 27, 2022, from <https://historyofyesterday.com/senet-f6acbd8f109f>

<sup>5</sup> Encyclopædia Britannica, inc. (n.d.). *Chess and Artificial Intelligence*. Encyclopædia Britannica. Retrieved July 27, 2022, from <https://www.britannica.com/topic/chess/Chess-and-artificial-intelligence>

<sup>6</sup> Eli. (2018, February 28). *Nine men's Morris*. Ancient Games - Playing the Board Games of the Ancient World. Retrieved July 27, 2022, from <https://www.ancientgames.org/nine-mens-morris/>

<sup>7</sup> Group, U. C. B. G. C. R. (n.d.). GamesCrafters. Retrieved July 27, 2022, from <http://gamecrafters.berkeley.edu/games.php?game=ninemensmorris#sec-references>

## Appendix A: p1\_legal\_moves

This function first creates an empty list that stores all the available positions to place a player's piece. It then loops through the board and tests every space on the board to see if it is occupied by a piece or not by calling the legal\_move function on said space. If the legal\_move function returns true for the space that was tested, that space gets added to the list of legal moves. This function was used to print out a list of legal moves for the player to choose from.

## Appendix B: p1\_make\_move

This function takes the parameter of region, position, and the team of whichever player is playing. The function first tests whether the chosen space is a legal move. If it is not, it displays a message that the move is not valid. However, if the chosen space is available, the program will mark the space as occupied and display it on the board by changing the color of the space to the team that placed the piece there. In addition, the number of pieces each team has placed is tracked every time they place a piece, since the first phase ends once both sides have placed nine pieces.

## Appendix C: p2\_legal\_moves

This function takes the parameters of the region and position of a piece the player wants to move. Given a piece on the board, the program checks the adjacent spaces in the same region (square) to see if they are occupied or not. If the spaces are not occupied, they are added to the list of legal moves that that piece can make. If the given piece is in an odd position, the adjacent spaces that are in different regions must also be checked. For these odd positions, there are two if statements that take the input of which region the given piece is in. If the given piece is in either the first or second region, the program checks if the adjacent space in the next region is occupied and adds the space to the list accordingly. If the given piece is in either the second or third region, the program checks if the adjacent space in the previous region is occupied and adds the space to the list accordingly. The function then returns the list of eligible spaces that the given piece can move into.

## Appendix D: p2\_make\_move

This function takes the parameters of the region and position of the space we want to move a piece into, which player's turn it is, and the region and position of the piece that the player wants to move. The program first checks if the space or piece chosen is occupied and displays a message stating that the move is illegal if they are occupied. In addition, a player could possibly input a space that is not adjacent to the piece that they chose. To prevent this, the program checks if the selected space is within the list provided by p2\_legal\_moves. If it is not in the list, the program will display a message stating that the move is illegal. If none of these cases occur, the program will move the piece from its original position to the new position.

## Appendix E: mills\_at

After every turn, the program must always check if one of the player's has made a mill. This function takes the parameters of the region and position of a certain piece on the board and which player's turn it is. A player can make a mill in four different ways: a horizontal row in one region, a vertical row in one region, a horizontal row across regions, and a vertical row across regions. However, the mills that are formed across regions only occur when the selected piece is at an odd position. As a result, if the given piece is in an even position, only the two positions in front or behind the piece need to be checked to determine if they have pieces of the same team in them. If either of these cases are true, the function returns true. If not, the

function returns false. On the other hand, if the piece's position is odd, the space in front and behind the piece are checked to see if they are occupied and of the same team. If they are, we return true. If not, we proceed to check if there is a mill going across all three regions, the final case. It would be rather inefficient to check for a mill based on the given piece, since a piece on the first region would have us check the two spaces in front of it while a piece on one of the other regions would be slightly different. Therefore, we simply take whatever position the given piece is in and assign variables to the space in each region of that same position. We then check if these three spaces are occupied and of the same team and return true or false accordingly.

Some Nine Men's Morris games allow a player to remove two enemy pieces if they make two mills at the same time. However, for the sake of simplicity, this circumstance was omitted. In the case a player forms two mills at once, they can only remove one piece from the opponent.

## **Appendix F: remove\_piece**

After a mill is formed and the program detects a mill, we need to be able to remove the enemy's piece from the board. This function takes inputs of the region and position of the piece that is being removed and which player is removing the piece. The function is very simple as all we need to do is check that the selected piece is on the opposite team and that the space is actually occupied. As long as these conditions are fulfilled, we can remove the piece from the board and make that region and position an empty space again. However, if one of those conditions fails to be true, the program displays a message stating that the player should choose a different piece to remove.

## **Appendix G: is\_win\_for\_team**

After each players' turn, we need to be able to check if that player has won the game. This function takes the parameter of whichever player's turn it is. The function first creates variables to store the number of pieces that the enemy player has and a list of possible moves that the enemy player can make. The function then loops through the entire board and every time it passes over an enemy piece, it adds one to the variable that stores the number of pieces the enemy has. In addition, the program will run the p2\_legal\_moves function on that piece and add each legal move to the list that stores the enemy's possible moves. After looping through the board, if the list is empty, meaning that the enemy has no legal moves, or if the enemy has two or less pieces on the board, the function returns true and returns false otherwise.

## **Appendix H: premills\_at**

Premills were classified as pieces that are in an unfinished, possible mill that need only one more piece. For example, a piece in a corner would be a part of a premill if there is a piece of the same team in the next consecutive position and there is an empty space in the position two ahead of it. Similarly, it is in a premill if there is a piece of the same team in the previous position and there is an empty space two positions behind it. A middle piece is in a premill if there is a piece of the same team in one of the sides of it and there is an empty space on the other side.

The function takes the parameters of region, position, which player's turn it is, and a board.

Suppose the given piece is in an even position on the board. For a mill going in front of the piece, the program checks if the space in the next position has a piece of the same team and the space in two positions is unoccupied. If so, the unoccupied space is added to the list of pieces (this list represents all the different premills). If not, the program then checks if the next position is unoccupied and the space in two positions

has a piece of the same team. If so, the unoccupied space is added to the list of pieces. For a mill going behind the piece, the program checks if the previous space has a piece of the same team and the space two positions behind is unoccupied. If so, the unoccupied space is added to the list of pieces. If not, the program checks if the previous position is unoccupied and the position two behind has a piece of the same team, if so, the unoccupied space is added to the list. The program then returns the list.

Suppose the given piece is in an odd position on the board. For a mill in the same region, the program checks if the position in front of the given piece has a piece of the same team and the position behind the piece is unoccupied. If so, the unoccupied space is added to the list of pieces. If not, the program checks if the position in front of the piece is unoccupied and the position behind it has a piece of the same team. If so, the unoccupied space is added to the list of pieces. For a mill going across regions, the program stores the “state” of each of the three spaces across the regions but with the same position. These “states” tell if the space is occupied, on a player’s team, or on the other player’s team. The way the piece’s states are stored accounts for all 3 combinations of possible premills. Therefore, we simply check if there are two pieces of the same team and an empty space in those three regions and add the space to the list of pieces. This list is then returned.

## **Appendix I: team\_premills**

This function takes the parameters of which team’s turn it is and a board. This function sums the number of premills a given team has on a given board.

The function iterates through the list of pieces that a given team has through the available\_pieces function. For every piece in the list, it gets a list of the premills the piece is in with the premills\_at function and adds the length of that list to a total. However, the function only returns the total divided by two. This is because in every premill, there are already two pieces. The function, though, goes through every piece and checks if it is in a premill. Therefore, the function will count twice for every premill. As a result, we have to halve the number of premills the function counts.

## **Appendix J: potential\_premill\_count\_at**

Potential premills were classified as an unfinished possible mill that needs two more pieces. For example, a piece in the corner would be in a potential premill if the two spaces in front of it are empty. It would be in another premill if the two spaces in the positions behind it were empty. A piece in an odd position and the outermost region would be in a potential premill if the two spaces of the same position but in the next two regions were empty.

This function takes the parameters of a piece’s region, position, and a board. This function returns the number of potential premills that a piece is in.

Suppose that the given piece is in an even position. If the spaces in the next two positions are both empty, the program increments the total. If the spaces in the previous two positions are both empty, the program increments the total by one. The program then returns the total.

Suppose the given piece is in an odd position. If the space in the next position and the space in the previous position are both empty, the program increments the total by one. If the spaces in the next two consecutive regions are unoccupied, the program increments the total by one. For the cases that the given piece is in the second or third region, the program uses the modulus function to correctly account for the next two regions. The program then returns the total.

## Appendix K: team\_potemills

The final function necessary to score a board is the team\_potemills, which takes the parameters of which team's turn it is and a board. This function counts the total number of potential mills that a given team has on a given board. The function iterates through all the pieces a team has on a board and adds the number of potential mills that each piece has using the potential\_premill\_count\_at. The function then returns the total.

## Appendix L: use\_mills

This function takes the parameters of whether a mill exists and which player made it. The program will display a message saying that the player can remove an opposing piece and will also display a list of the enemy pieces on the board. We then call the input\_position function to get which piece the player wants to remove. Finally, we call remove\_piece on that piece.

## Appendix M: make\_p1\_move

This function takes the parameters of the team playing and the board. The function will display the legal moves for phase 1 for the player to choose from with p1\_legal\_moves. The function then takes the user's input with input\_position and runs p1\_make\_move on the chosen space. The function then checks to see if the newly placed piece has formed a mill with the mills\_at function. If so, the program will call use\_mills (described above) and allow the player to remove an opposing piece.

## Appendix N: make\_p2\_move

This function takes the parameters of the team playing and the board. The function loops through all the pieces a team has on the board with available\_pieces. For each available piece, the function will run p2\_legal\_moves on it and if the list of legal moves is not empty, the function will add the piece to a list of eligible pieces. The function then takes the user input to determine which of the eligible pieces the user wants to move. The program will then take user input to determine where the piece should be moved and runs p2\_make\_move on the piece and chosen space. Similar to make\_p1\_move, the function checks to see if the moved piece makes a mill with the mills\_at function. If so, the program will call use\_mills and allow the player to remove an opposing piece.

## Appendix O: random\_p1\_move

The first function within the random class is random\_p1\_move, which executes a random piece placement in the first phase. The function first calls p1\_legal\_moves to get a list of possible moves and then chooses a move from that list at random. The function then determines if the piece just placed resulted in a mill with the mills\_at function. If so, the function will list all the possible enemy pieces to remove by calling available\_pieces on the opposing team. It will then choose a piece at random and call remove\_piece on it.

## Appendix P: random\_p2\_move

The second function in the random class is random\_p2\_move, which executes a random piece placement in the second phase. The function first creates a list of the random team's pieces and chooses a piece at random to move. However, there is a possibility that the piece chosen by the random function doesn't have

any legal moves. Therefore, we use a while loop to choose a different random piece if the piece chosen didn't have any legal moves. After choosing a valid piece, the program creates a list to store the possible moves for that piece using p2\_legal\_moves. The program then chooses a random move from that list and runs p2\_make\_move on the chosen move. Similar to phase 1, the program checks if any mills have been made and randomly removes a piece if so.

## Appendix Q: Simple Vs. Random Raw

Simple Vs. Random												
Depth = 0	Simple Moving 1st			Simple Moving 2nd			Totals					
Trial No.	Simple Wins	Random Wins	Stalemates	Simple Wins	Random Wins	Stalemates	Simple Total Wins	Random Total Wins	Simple Win %	Random Win %	% Difference	
1	495	2	3	491	4	5	986	6	98.6	0.3	98.3	
2	498	1	1	492	6	2	990	7	99	0.35	98.65	
3	497	1	2	493	4	3	990	5	99	0.25	98.75	
4	499	0	1	485	11	4	984	11	98.4	0.55	97.85	
5	498	1	1	495	4	1	993	5	99.3	0.25	99.05	
Average	497.4	1	1.6	491.2	5.8	3	988.6	6.8	98.86	0.34	98.52	
Depth = 1	Simple Moving 1st			Simple Moving 2nd			Totals					
Trial No.	Simple Wins	Random Wins	Stalemates	Simple Wins	Random Wins	Stalemates	Simple Total Wins	Random Total Wins	Simple Win %	Random Win %	% Difference	
1	497	0	3	492	6	2	989	6	98.9	0.3	98.6	
2	497	1	2	494	4	2	991	5	99.1	0.25	98.85	
3	499	0	1	493	3	4	992	3	99.2	0.15	99.05	
4	500	0	0	495	2	3	995	2	99.5	0.1	99.4	
5	497	1	2	495	3	3	992	4	99.2	0.2	99	
Average	498	0.4	1.6	493.8	3.6	2.8	991.8	4	99.18	0.2	98.98	
Depth = 2	Simple Moving 1st			Simple Moving 2nd			Totals					
Trial No.	Simple Wins	Random Wins	Stalemates	Simple Wins	Random Wins	Stalemates	Simple Total Wins	Random Total Wins	Simple Win %	Random Win %	% Difference	
1	494	1	5	486	7	7	980	8	98	0.4	97.6	
2	498	0	2	481	10	9	979	10	97.9	0.5	97.4	
3	491	7	2	485	6	9	976	13	97.6	0.65	96.95	
4	496	2	2	490	5	5	986	7	98.6	0.35	98.25	
5	491	4	5	490	6	4	981	10	98.1	0.5	97.6	
Average	494	2.8	3.2	486.4	6.8	6.8	980.4	9.6	98.04	0.48	97.56	

## Appendix R: Intermediate Vs. Random Raw

Intermediate Vs. Random												
Depth = 0	Intermediate Moving 1st			Intermediate Moving 2nd			Totals					
Trial No.	Intermediate Wins	Random Wins	Stalemates	Intermediate Wins	Random Wins	Stalemates	Intermediate Total Wins	Random Total Wins	Intermediate Win %	Random Win %	% Difference	
1	495	0	5	495	1	4	990	1	99	0.05	98.95	
2	496	0	4	494	0	6	990	0	99	0	99	
3	492	0	8	496	0	4	988	0	98.8	0	98.8	
4	496	0	4	495	1	4	991	1	99.1	0.05	99.05	
5	495	0	5	490	0	10	985	0	98.5	0	98.5	
Average	494.8	0	5.2	494	0.4	5.6	988.8	0.4	98.88	0.02	98.86	
Depth = 1	Intermediate Moving 1st			Intermediate Moving 2nd			Totals					
Trial No.	Intermediate Wins	Random Wins	Stalemates	Intermediate Wins	Random Wins	Stalemates	Intermediate Total Wins	Random Total Wins	Intermediate Win %	Random Win %	% Difference	
1	498	0	2	493	0	7	991	0	99.1	0	99.1	
2	496	0	4	497	0	3	993	0	99.3	0	99.3	
3	494	0	6	496	0	4	990	0	99	0	99	
4	496	0	4	497	0	3	993	0	99.3	0	99.3	
5	498	0	2	498	0	2	996	0	99.6	0	99.6	
Average	496.4	0	3.6	496.2	0	3.8	992.6	0	99.26	0	99.26	
Depth = 2	Intermediate Moving 1st			Intermediate Moving 2nd			Totals					
Trial No.	Intermediate Wins	Random Wins	Stalemates	Intermediate Wins	Random Wins	Stalemates	Intermediate Total Wins	Random Total Wins	Intermediate Win %	Random Win %	% Difference	
1	499	0	1	496	0	4	995	0	99.5	0	99.5	
2	497	0	3	498	1	1	995	1	99.5	0.05	99.45	
3	498	0	2	496	0	4	994	0	99.4	0	99.4	
4	500	0	0	497	0	3	997	0	99.7	0	99.7	
5	499	0	1	496	0	4	995	0	99.5	0	99.5	
Average	498.6	0	1.4	496.6	0.2	3.2	995.2	0.2	99.52	0.01	99.51	

## Appendix S: Intermediate Vs. Simple Raw

Simple Vs. Intermediate											
Depth = 0	Intermediate Moving 1st			Intermediate Moving 2nd			Totals				
Trial No.	Intermediate Wins	Simple Wins	Stalemates	Intermediate Wins	Simple Wins	Stalemates	Intermediate Total Wins	Simple Total Wins	Intermediate Win %	Simple Win %	% Difference
1	303	169	28	277	191	32	580	360	58	18	40
2	262	214	24	291	195	14	553	409	55.3	20.45	34.85
3	270	209	21	276	209	15	546	418	54.6	20.9	33.7
4	275	208	17	303	176	21	578	384	57.8	19.2	38.6
5	282	203	15	275	213	12	557	416	55.7	20.8	34.9
Average	278.4	200.6	21	284.4	196.8	18.8	562.8	397.4	56.28	19.87	36.41
Depth = 1	Intermediate Moving 1st			Intermediate Moving 2nd			Totals				
Trial No.	Intermediate Wins	Simple Wins	Stalemates	Intermediate Wins	Simple Wins	Stalemates	Intermediate Total Wins	Simple Total Wins	Intermediate Win %	Simple Win %	% Difference
1	328	160	12	294	200	6	622	360	62.2	18	44.2
2	326	165	9	283	210	7	609	375	60.9	18.75	42.15
3	332	158	10	276	214	10	608	372	60.8	18.6	42.2
4	351	136	13	295	200	5	646	336	64.6	16.8	47.8
5	338	152	10	287	208	5	625	360	62.5	18	44.5
Average	335	154.2	10.8	287	206.4	6.6	622	360.6	62.2	18.03	44.17
Depth = 2	Intermediate Moving 1st			Intermediate Moving 2nd			Totals				
Trial No.	Intermediate Wins	Simple Wins	Stalemates	Intermediate Wins	Simple Wins	Stalemates	Intermediate Total Wins	Simple Total Wins	Intermediate Win %	Simple Win %	% Difference
1	420	54	26	436	44	20	856	98	85.6	4.9	80.7
2	424	57	19	441	41	18	865	98	86.5	4.9	81.6
3	421	49	30	433	51	16	854	100	85.4	5	80.4
4	423	56	21	425	60	15	848	116	84.8	5.8	79
5	428	44	28	439	41	20	867	85	86.7	4.25	82.45
Average	423.2	52	24.8	434.8	47.4	17.8	858	99.4	85.8	4.97	80.83

## Appendix T: Source Code

<https://github.com/samuelwang23/Nine-Mens-Morris>



**MATHEMATICS  
TEAM PROJECT**



# Applying Stochastic Processes to Decision Making

Devin Bryant, Thomas Gordon, Jack Hasker, Rafael Quinodoz,  
Ben Sparling, Lilyanna Strocko, and Kovid Tandon

## Abstract

Stochastic Processes can be applied to many different fields, from particle physics to finance in order to optimize decisions based on the outcomes of previously occurring events. Inspired by the format of weightlifting competitions, we created a simplified carnival game that used stochastic processes to optimize decisions in order to earn the highest expected value.<sup>5</sup> The basic format of the game involved picking a  $K$  value on the interval 0 to 1. Then an  $X$  value will be selected from the same domain. If it is greater than the  $K$  value, you win  $K$  and if it is not, you win nothing. In order to maximize the expected value of a round, (which can be alternatively thought of as your “winnings”) you have to multiply the optimized gains of that round (equivalent to the best  $K$  value you can select) times the probability of that given  $K$  value occurring. The situation in which you only have to pick one  $K$  in one round was just described, but what if the game is more complex? In this paper, we will look at different scenarios of this game. These scenarios include situations where you must pick one, two, or three  $K$  values before you begin playing, picking  $K$  values as you progress through the rounds, and a few others. We will also show how to select the best  $K$  values in every given situation to maximize the game’s expected value and elaborate on other mathematical techniques (such as Newton’s Method) used to find these  $K$  values and the expected value they produce.

## I. Introduction

A stochastic process is a process in which current decisions are optimized based on events of the past.<sup>1</sup> In this situation, a set of random variables on a common probability space is associated with the value of the variable at previous times. These processes track how random phenomena change throughout time. They can be applied to a variety of real-world situations from gambling and carnival games, to even particle physics and finance. Although the previously mentioned applications are compelling in their own way, the inspiration for this project originated from weightlifting. In our weightlifting scenario, there are a variety of rules that must be followed. First, each participant will be required to lift three separate times. For their first lift, the participant will be allowed to select a desired quantity of weight. If they are not able to lift the weight, they will receive no points. However, if they are able to lift the weight, the number of points they receive will be equivalent to the mass of the weight. In addition, they will be required to lift a second weight that must be at least as heavy as the first regardless of whether they succeeded or failed to lift the first weight. If they succeed in lifting the second weight, they will receive a number of points equivalent to the mass of the second weight. This process will go on up until the third weight. The total amount of points is equivalent to only the heaviest weight lifted through the games.<sup>4</sup> If you apply stochastic processes to this situation, the following question could be asked and answered: what is the optimal amount of weight the lifter should start with to earn the maximum amount of points?<sup>2</sup>

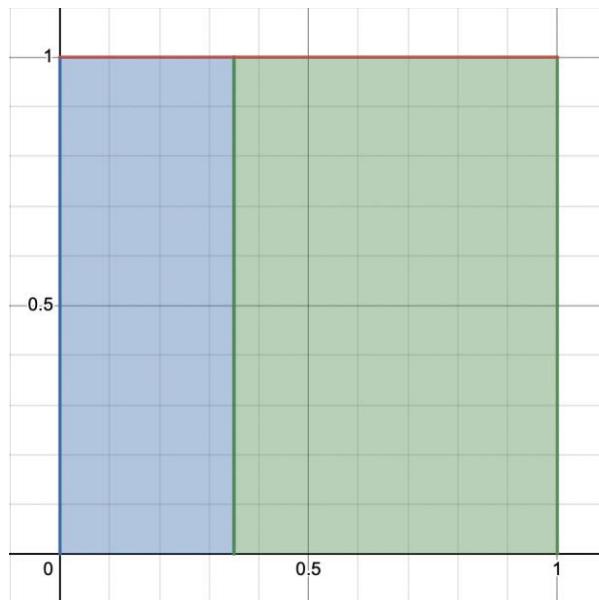
This situation inspired us to apply mathematical methods like stochastic processes to find the optimal strategy for situations with similar premises, following the rule that you cannot go lower than your previously chosen number.

In order to conceptualize this project, our calculations and scenarios will be embedded into the setting of a hypothetical game. This game will make it easier to cover different scenarios and make computations.

## II. Our Game

This hypothetical game will have several variations; however, some aspects will stay consistent throughout. In every variant, the participant will have the opportunity to select one or several  $K$  values. The  $K$  value will be on the interval  $[0, 1]$ . Then, an  $X$  value will be selected uniformly at random on the same interval (some versions of the game will have several  $X$  values). If the  $X$  value is greater than the  $K$  value, then you win  $K$ , but if it is less, then you win nothing.

In each scenario, we will generate an expected gains equation ( $\mathbb{E}[G]$ ) that is dependent on the  $K$  value(s).  $G$  represents the amount gained at the end and  $\mathbb{E}[G]$  represents the expected value of the game. The expected value is found by adding two quantities. The first is the amount gained when you win multiplied by the probability of that happening and the second quantity is the amount gained when you lose times the probability of that occurring. Going back to the weightlifting analogy, if you lift a weight that is  $K$  units, then you will win  $K$  points, and the expression for the win quantity is  $K(1 - K)$ . This is the formula for the winning half of the expected value equation because it is the gains if the weight was successfully lifted ( $K$ ) multiplied by the probability of success ( $1 - K$ ). The probability of success is  $(1 - K)$  because it represents all values greater than  $K$ , which are needed to win. If you do not lift the weight, then you win nothing, and the expression for the lose quantity is  $0 * K = 0$ . The value of 0 represents the points gained if the weight is not successfully lifted, and the value of  $K$  represents all values from 0 to  $K$  (probability of losing), the values for which the lifter loses.



**Figure 1.** This is a graphical representation of an arbitrarily selected  $K$  value (represented by the dark green line). If the selected  $X$  is greater than the  $K$ , you win and that probability is represented by the green area. If it is less, you lose and the probability is shown with the blue area.

For each game situation created, there is a  $K$  value that will yield the best expected gain for the player. In order to find the  $K$  value that optimizes the expected value equation in every situation, we will first deduce the  $\mathbb{E}[G]$  equation. Then we will simplify it and find the derivative of the equation in terms of  $K$ . After differentiating, we will set the derivative equal to zero and solve for  $K$ . This will give us the value of  $K$  where the graph of  $\mathbb{E}[G]$  is at its maximum or minimum value. To figure out whether it is a maxima or minima, we

used the second derivative test. If the second derivative is negative, the graph is concave down and the value is a maximum. If the second derivative is positive, the graph is concave up and the value is a minimum. The value we find in all of our game scenarios will be the optimal  $K$  value or the point where the graph of the expected value equation is at its maximum. Returning to the weightlifter example, the maximum we find represents the value of  $K$  that the participant should select in a given situation to maximize their expected gains from the game.

### III. Scenarios

#### A: Situations Involving Picking 1 K Value

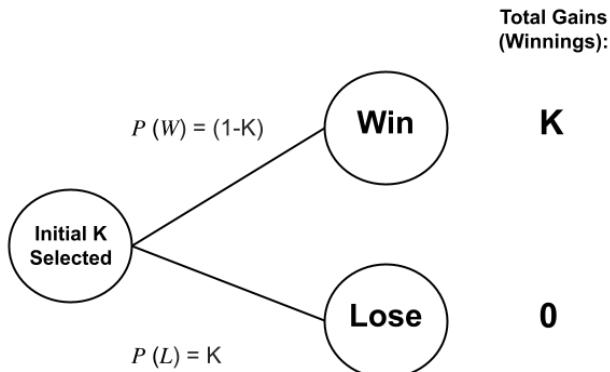
The situations below are represented by the notation  $\mathbb{E}[G^n]$  for the expected value where  $n$  represents the number of rounds for each situation and  $G^n$  represents the gain.

##### 1. The First Situation (Picking 1 K Value, Random X Selected, 1 Round)

In this game, we can choose any value on the interval on  $[0,1]$ ; the number that we choose will be denoted as  $K$ . A random number is then randomly selected uniformly within the same interval. This random number will be denoted as  $X$ . If the selected value  $K$  is less than the randomly selected value  $X$ , then  $K$  is earned.

The expected value shows the average amount a player would win if they played the game over and over again. To find the formula for expected value, we take the earnings if the game is won,  $K$ , and multiply this by the chances of winning,  $1 - K$ . We would then take the value of losing the game,  $0$ , and multiply this by the chances of losing,  $K$ . When these two quantities are added together, it results in the expected value formula for this situation. The expected value formula that we generate can be used to find the optimal value of  $K$  that can then be substituted back into the expected value formula to maximize the expected value. This is done by rewriting the expected value equation as a function of  $K$ , finding the derivative of this function, and setting it equal to 0. The value where the derivative equals 0, will be the optimal  $K$  value to choose in order to maximize the expected value of the game.

The figure shown below (figure 2) assisted us in finding the expected value equation for this situation. We will use  $G^1$  to denote the gain for the initial pick, which in this game, will be the value  $K$  if the game is won and zero if the game is lost. We will use  $P(W)$  to represent the probability of winning, and  $P(L)$  to represent the probability of losing this round.



**Figure 2. Tree showing all of the possible paths with 1  $K$  value and 1  $X$  value**

The computation for the expected value is expressed in the following equation:

$$\mathbb{E}[G^1] = K * (1 - K) + 0 * K$$

By simplifying this equation and calculating the derivative of the expression, we can find the value of  $K$  that would yield the highest expected value:

$$\begin{aligned}\mathbb{E}[G^1] &= K * (1 - K) + 0 * K \\ &= K(1 - K) \\ &= K - K^2.\end{aligned}$$

Now, let  $f(K) = K - K^2$  and take the derivative:

$$f(K) = K - K^2$$

$$f'(K) = 1 - 2K .$$

Set the derivative equal to 0 and solve for  $K$  to maximize the function  $f(K)$ :

$$0 = 1 - 2K$$

$$K = \frac{1}{2}.$$

This shows that selecting the value  $K$  to equal  $\frac{1}{2}$  would yield the highest result for the expected value in the one-round game where one  $K$  value is selected.

The second derivative test on  $f(K)$  can be used to prove that  $\frac{1}{2}$  is indeed the optimal value for  $K$ . The graph of  $f(K)$  is concave down because we found the second derivative to be negative. Because it is concave down, this point is a maximum, making  $\frac{1}{2}$  the best value for  $K$ .

Now that the value of  $K$  which yields the maximum expected value has been calculated, we can find the maximum expected value for one round by replacing  $K$  with the value of  $\frac{1}{2}$  in the function  $\mathbb{E}[G^1]$ :

$$\mathbb{E}[G^1] = \frac{1}{2} \left(1 - \frac{1}{2}\right)$$

$$\mathbb{E}[G^1] = \frac{1}{4}.$$

Hence, the maximum expected value where  $K = \frac{1}{2}$  in a single round game is  $\frac{1}{4}$ .

## 2: The Second Situation (Picking 1 $K$ Value, 2 Random $X$ Values Selected, 2 Rounds)

This situation is similar to the first, but it entails the addition of one more round. Instead of one random  $X$  value being generated, two random  $X$  values are generated;  $X_1$  representing the first round and  $X_2$

representing the second round. Once again, we can choose any value on the interval  $[0, 1]$  to be our  $K$  value, and that will be our  $K$  value for both of the rounds.  $P$  represents the probability of an event happening.

In the first round, the  $X_1$  value will be revealed. If our  $K$  is greater than the  $X$  value, then we lose and gain nothing. If our  $K$  value is less than or equal to the  $X_1$  value, then we gain  $K$ . We can only win the value  $K$  once, so even if both rounds are won, the value  $K$  is only gained once. To yield the best outcome with a singular  $K$  value, only one win is necessary. If the first round is won, it will continue onto the winning node in figure 3. If the first round is lost, it will continue onto the losing node of the tree in figure 3.

In the second round, we will maintain the same  $K$  value that was selected at the beginning, regardless of the result of the first round. If we are in the win node, regardless of what the value of  $X_2$  is, we will gain  $K$  because the first round was won. If we are in the losing node - because we lost the first round - we essentially have a second chance to win  $K$ . If the second  $X$  value is greater than or equal to  $K$ , then we win  $K$ . If  $X$  is less than  $K$ , then we end the game without any winnings because both rounds were lost.

In this version of the game, we will use a similar process described in the initial situation, with appropriate adjustments to account for the two  $X$  values or two rounds of the game. In order to find the value of  $K$  that would maximize the expected gain, we can add the expected gains for all four paths of the tree shown in Figure 3. Each expected gain is obtained by multiplying the amount gained ( $K$  or 0) times the probability of gaining that amount. To represent this game, we will denote the expected value as  $\mathbb{E}[G^2]$ , using  $G^2$  instead of  $G^1$  because the number in the exponent represents the number of rounds:

$$\mathbb{E}[G^2] = \sum_{N \in \{K, 0\}} N * P [gaining N].$$

To represent the different paths that may be followed on the tree illustrated in figure 3, the sequence of wins and losses leading up to the round can be denoted in parenthesis. This is important because depending on the series of wins and loses, there will be a different amount gained at the end of the game, and different probabilities of following each different path. For example,  $(WL)$  represents winning the first round and losing the second round.

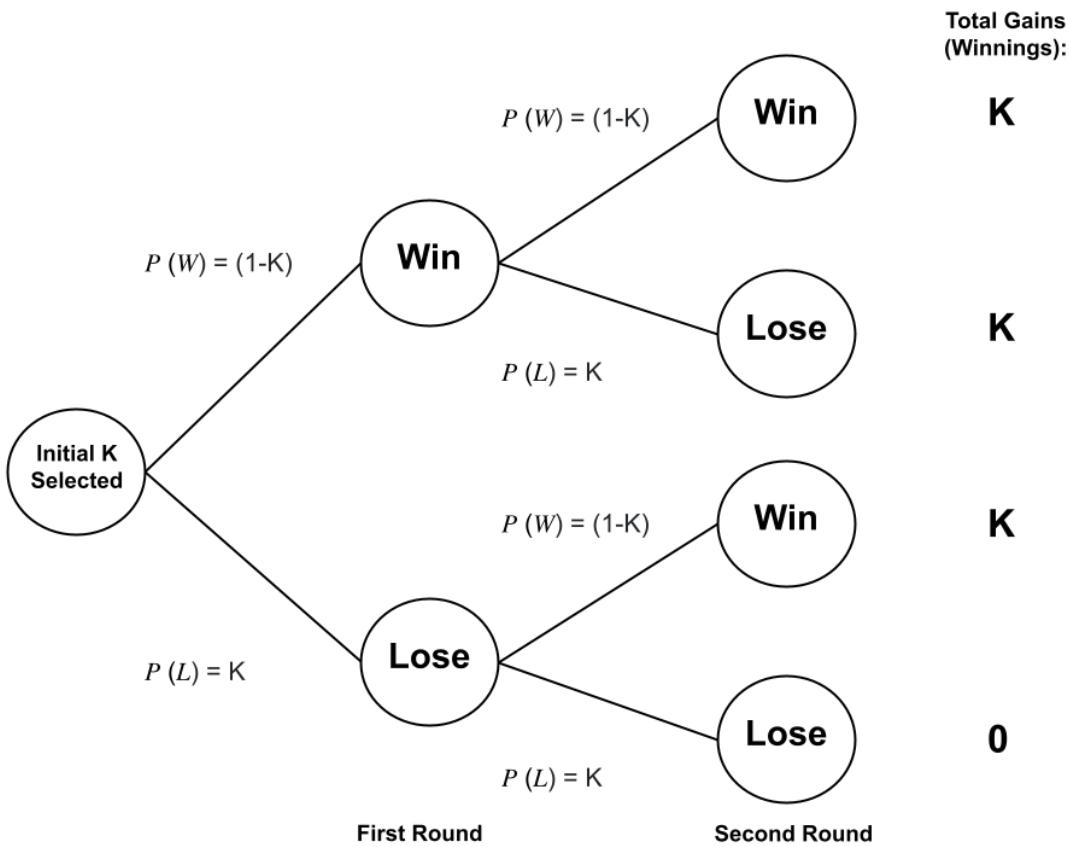
The first path, which is denoted as  $(WW)$ , is where the first two rounds are won. To find the expression to represent this path, we will first multiply  $K$ , the value gained from winning a round, by  $1 - K$  (the chances of winning up to the first layer) and  $1 - K$  again (the chances of winning from the first layer to the second layer). The first path will therefore be represented by the expression  $K(1 - K)(1 - K)$ .

The second path  $(WL)$  involves winning the first round and losing the second one. Therefore, we will multiply  $K$ , the value gained from winning a round, by  $1 - K$  (the chances of winning up to the first layer) and  $K$  (the chances of losing the round from the first to the second layer). This path will be represented by  $K(1 - K)K$  in the expected value equation.

The third path is denoted as  $(LW)$  and it is the situation where we lose the first round and win the second. This means we will multiply  $K$ , the value gained from winning a round, by  $K$  (the chances of losing the round from the first to the second layer) and  $1 - K$  (the chances of winning from the first to the second layer). The third path's component in the expected value equation is  $K(K)(1 - K)$ , which is the same as the second path.

Lastly is the fourth path  $(LL)$  which involves losing both rounds. The expression to represent this path is

found by multiplying 0, the value gained from losing a round, by  $K$  (the chances of losing the round from the initial layer to the first layer) and  $K$  (the chances of losing the round from the first to the second layer). The fourth path is written as  $(0)(K)(K)$ , which then simplifies to 0.



**Figure 3. Tree showing all of the possible paths with 1  $K$  value and 2  $X$  values**

When each path is added together, the expected value equation for this situation is the following:

$$\mathbb{E}[G^2] = (K)(1 - K)(1 - K) + (K)(1 - K)(K) + (K)(K)(1 - K) + (0)(K)(K)$$

By simplifying the expected value equation, calculating the derivative of the expression, and setting it equal to zero, we can find the value of  $K$  that would yield the highest expected value, which is the optimal  $K$ . First, we simplify the expected value equation:

$$\begin{aligned} \mathbb{E}[G^2] &= (K)(1 - K)(1 - K) + (K)(1 - K)(K) + (K)(K)(1 - K) + (0)(K)(K) \\ &= (K)(1 - K)(1 - K) + (K)(1 - K)(K) + (K)(K)(1 - K) \\ &= (K)(K^2 - 2K + 1) + (K)(-K^2 + K) + (K)(-K^2 + K) \\ &= (K^3 - 2K^2 + K) + (-K^3 + K^2) + (-K^3 + K^2) \\ &= -K^3 + K \end{aligned}$$

Let  $f(K) = -K^3 + K$ . Then differentiate:

$$f(K) = -3K^2 + 1$$

We set  $f'(K)$  equal to 0 in order to find the value of  $K$  which maximizes  $f(K)$ . This gives us:

$$0 = -3K^2 + 1$$

$$-1 = -3K^2$$

$$K = \sqrt{\frac{1}{3}}$$

The above calculations show that choosing a  $K$  value of  $\sqrt{\frac{1}{3}} \approx 0.577$  would yield the highest expected value in earnings from this version of the game. This answer deviates from the initial situation because there are two rounds, so there is an additional opportunity to win, resulting in a higher optimal  $K$  value.

The higher optimal  $K$  value means the total earnings will be higher if the game is won, and the probability of winning  $K$  in one round decreases. However, having two rounds compensates for this decreased probability because there are two chances to win  $K$  (which raises the probability of winning  $K$  by the end of the game).

With this calculated  $K$  value, we can find the maximum expected value by replacing  $K$  in our equation with  $\sqrt{\frac{1}{3}}$ , which is shown below:

$$\begin{aligned} \mathbb{E}[G^2] &= \left( \sqrt{\frac{1}{3}} * \left( 1 - \sqrt{\frac{1}{3}} \right) \right) + \left( \sqrt{\frac{1}{3}} * \left( 1 - \sqrt{\frac{1}{3}} \right) \left( \sqrt{\frac{1}{3}} \right) \right) + \left( \sqrt{\frac{1}{3}} * \left( 1 - \sqrt{\frac{1}{3}} \right) \left( \sqrt{\frac{1}{3}} \right) \right) + \\ &\quad \left( 0 \left( 1 - \sqrt{\frac{1}{3}} \right) \left( \sqrt{\frac{1}{3}} \right) \right) \\ &= -\left( \sqrt{\frac{1}{3}} \right)^3 + \sqrt{\frac{1}{3}} \\ &= \frac{2}{3\sqrt{3}}. \end{aligned}$$

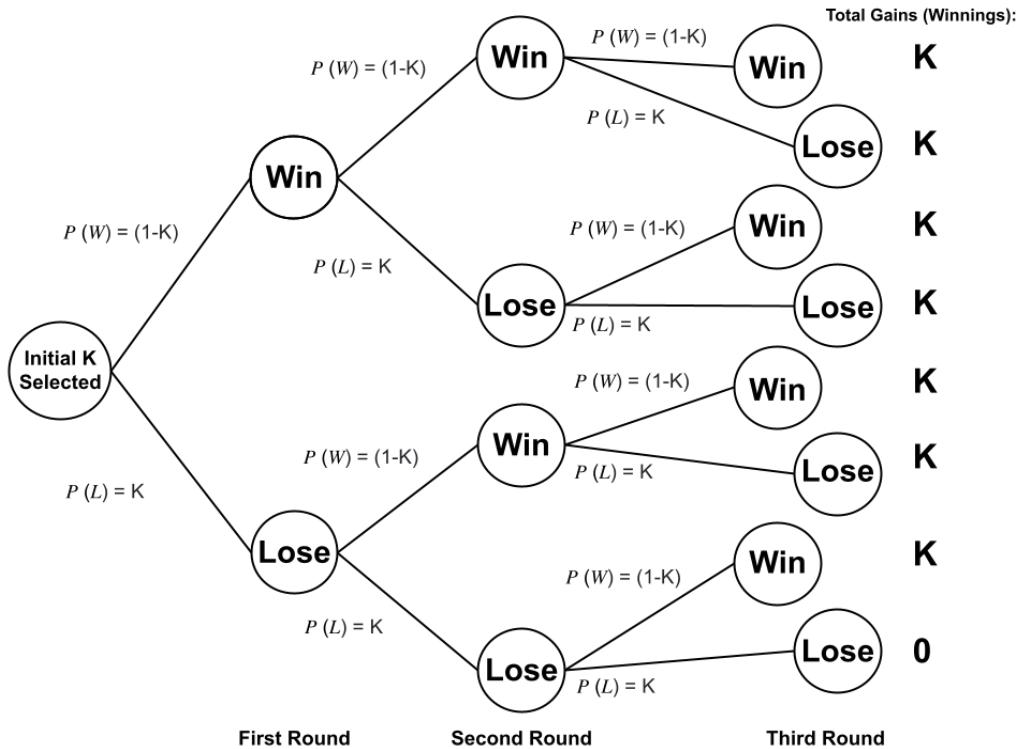
Hence, the maximum expected value happens when  $K = \sqrt{\frac{1}{3}}$ . The expected value is  $\frac{2}{3\sqrt{3}} \approx 0.385$

### 3. The Third Situation (Picking 1 $K$ Value, 3 Random $X$ Values Selected, 3 Rounds)

The difference between this version of the game and the previous version is that there is another  $X$  value chosen, so there is one more round. This expands our game from two rounds to three. One  $K$  value is chosen at the beginning of the game on the interval  $[0, 1]$ , and that value is used for all three rounds. The first random variable will be denoted as  $X_1$ , the second will be denoted as  $X_2$ , and the third will be  $X_3$ . The subscript represents the round number being played.

The first 2 rounds remain unchanged from the second game, and our third round consists of an additional win and lose branch on each of the nodes in round 2. Each path that consists of at least one win node will result in earning the value  $K$  because  $K$  can only be won once. The path that consists of losing every single round will result in earning 0. Therefore, every path other than (LLL) will result in gaining  $K$ , as shown in Figure 4.

The expected value of this game can be shown as the value of the result of each end node multiplied by the probability of earning the result. The following tree (labeled as figure 4) can be used to create the expected value equation of this three-round game.



**Figure 4. Tree showing all of the possible paths with 1  $K$  value and 3  $X$  values**

The expected value equation in this situation is more complex than the previous two situations. The equation now becomes the expected value of specifically  $G^3$  because there are three rounds in this variation. The expected value equation is shown below.

$$\mathbb{E}[G^3] = (K * (1 - K)^3) + 3(K^2 * (1 - K)^2) + 3(K^3 * (1 - K) + (0(K^3))).$$

By simplifying this equation and calculating the derivative of the expression, we can find the value of  $K$  that would yield the highest expected value:

$$\begin{aligned} \mathbb{E}[G^3] &= (K * (1 - K)^3) + 3(K^2 * (1 - K)^2) + 3(K^3 * (1 - K) + (0(K^4))) \\ &= -K^4 + 3K^3 - 3K^2 + K + 3K^4 - 6K^3 + 3K^2 + 3K^3 - 3K^4 \\ &= K - K^4 \end{aligned}$$

To find the highest expected value of this expression, let  $f(K) = K - K^4$ .

We then take the derivative of  $f(K)$ :

$$f'(K) = 1 - 4K^3.$$

Finally, we set the derivative equal to zero to find the optimal  $K$  value:

$$\begin{aligned} 0 &= 1 - 4K^3. \\ -1 &= -4K^3. \\ K &= \sqrt[3]{\frac{1}{4}}. \end{aligned}$$

The calculation above shows that when  $K = \sqrt[3]{\frac{1}{4}}$  ( $\approx 0.630$ ), the expected value equation is at its maximum.

This follows the pattern observed in the previous situation where the value of  $K$  increases as the number of rounds grows.

With this calculated value of  $K$ , we can find the maximum expected value by replacing  $K$  with the value  $\sqrt[3]{\frac{1}{4}}$  in our expected value equation:

$$\begin{aligned} \mathbb{E}[G^3] &= (\sqrt[3]{\frac{1}{4}}) * (1 - (\sqrt[3]{\frac{1}{4}})^3) + 3((\sqrt[3]{\frac{1}{4}})^2 * (1 - (\sqrt[3]{\frac{1}{4}})) + 3((\sqrt[3]{\frac{1}{4}})^3 * (1 - (\sqrt[3]{\frac{1}{4}}) + (0((\sqrt[3]{\frac{1}{4}})^4))) \\ &= \sqrt[3]{\frac{1}{4}} - (\sqrt[3]{\frac{1}{4}})^4 \\ &= \frac{3}{4(\sqrt[3]{4})}. \end{aligned}$$

Hence, when  $K = \sqrt[3]{\frac{1}{4}} \approx 0.630$ , the maximum expected value in this situation is  $\frac{3}{4(\sqrt[3]{4})} \approx 0.472$ .

### i: Synthesizing the Initial Situation, the Second Situation, and the Third Situation

Situations 1 through 3 all follow a distinct pattern. In each situation, you are keeping all of the parameters constant, aside from the addition of an  $X$  value, which can be alternatively thought of as an additional layer to the tree. In the initial situation, we found a simplified expected value equation of  $\mathbb{E}[G^1] = K - K^2$ . In situation 2, we found a simplified expected value equation of  $\mathbb{E}[G^2] = K - K^3$ . In situation 3 we got an expected value equation of  $\mathbb{E}[G^3] = K - K^4$ .

A clear pattern is present in these situations. As the number of rounds played,  $n$ , increases by one, the exponent on the second  $K$  term in the equation does as well. Therefore, the number in the exponent is always one greater than the number of rounds,  $n$  in this situation.

In fact, this pattern continues for all  $n$  as proven in the combinatorial proof below.

**Proposition:**

We have,

$$\mathbb{E}[G^n] = K - K^{n+1}.$$

*Proof:* First, we will show that

$$\mathbb{E}[G^n] = K \sum_{i=0}^n \frac{n}{i} K^i (1-K)^{n-i} - K^{n+1}.$$

When  $n$  rounds of the game are played, there are  $\frac{n}{i}$  ways to lose  $i$  rounds and win  $n - i$  rounds, with probability  $K$  of losing each round and probability  $(1 - K)$  of winning each round.

If at least one round of the game is won, there is a gain of  $K$ , while if all the rounds are lost, there is a gain of 0.

Thus, the total expected value can be written as

$$K \sum_{i=0}^{n-1} \frac{n}{i} K^i (1-K)^{n-i} = K \sum_{i=0}^n \frac{n}{i} K^i (1-K)^{n-i} - K^{n+1}.$$

By the Binomial Theorem,

$$\sum_{i=0}^n \frac{n}{i} K^i (1-K)^{n-i} = 1.$$

Therefore,

$$\mathbb{E}[G^n] = K(1) - K^{n+1} = K - K^{n+1}.$$

This equation is very important because when we differentiate it, set the derivative equal to zero, and solve for  $K$ , we can find the optimal singular  $K$  value for any amount of rounds or  $X$  values chosen. This process is shown below:

$$\mathbb{E}[G^n] = K - K^{n+1}$$

To find the highest expected value of this expression, let  $f(K) = K - K^{n+1}$ .

$$f(K) = K - K^{n+1}.$$

We then take the derivative of  $f(K)$ :

$$f'(K) = 1 - (n+1)K^n$$

Finally, we set the derivative equal to zero and solve for  $K$  to find the optimal  $K$  value:

$$0 = 1 - (n + 1)K^n$$

$$K^n = \frac{1}{n + 1}$$

$$K = \frac{1}{\sqrt[n+1]{n+1}}.$$

The equation above can be used to find the optimal  $K$  value for any number of rounds,  $n$ , or the number of selected  $X$  values in a game. By finding the limit of this equation as the number of rounds,  $n$  in the game increases, we can discover the trend of the optimal  $K$  value. The limit of this expression is solved below:

$$\lim_{n \rightarrow \infty} \frac{1}{(n + 1)^{(1/n)}} = \frac{1}{\lim_{n \rightarrow \infty} (n + 1)^{(1/n)}}.$$

Solving for the limit of the denominator (but keeping in mind that the numerator is 1), we obtain:

$$\begin{aligned} \lim_{n \rightarrow \infty} (n + 1)^{(1/n)} &= \lim_{n \rightarrow \infty} e^{\ln(n + 1)^{(1/n)}} \\ &= e^{\lim_{n \rightarrow \infty} \ln(n + 1)^{(1/n)}} \\ &= e^{\lim_{n \rightarrow \infty} \frac{1}{n} \ln(n + 1)}. \end{aligned}$$

Finding the limit of the exponent using L'Hôpital's rule, we get:

$$\begin{aligned} e^{\lim_{n \rightarrow \infty} \frac{1}{n} \ln(n + 1)} &= e^{\lim_{n \rightarrow \infty} \frac{\ln(n + 1)}{n}} \\ &= e^{\lim_{n \rightarrow \infty} \frac{\frac{1}{n+1}}{\frac{-1}{n^2}}} \\ &= e^{\lim_{n \rightarrow \infty} \frac{n}{n+1}} \\ &= e^{\lim_{n \rightarrow \infty} \frac{1}{n+1}} \\ &= e^0 \\ &= 1. \end{aligned}$$

Keeping in mind the numerator limit was also 1:

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{(n + 1)^{(1/n)}} &= \frac{1}{1} \\ &= 1. \end{aligned}$$

This proves that as the number of rounds being played increases without bound, (or  $X$  values selected) the value of  $K$  that should be selected to maximize your expected gains will approach 1.

This value makes sense, because since you can only win once, a value of  $K$  closer to one will result in a better gain, and as more rounds are played, the probability of winning that higher value  $K$  increases.

## B: Situations Involving Picking More Than 1 $K$ Value at the Beginning

These situations are represented by the notation  $E[\hat{G}]^{(n)}$  for the expected value where  $n$  represents the number of rounds.

### 1. The Fourth Situation (Select $K_1 \leq K_2$ at the beginning, then 2 random variables, $X_1$ and $X_2$ are selected)

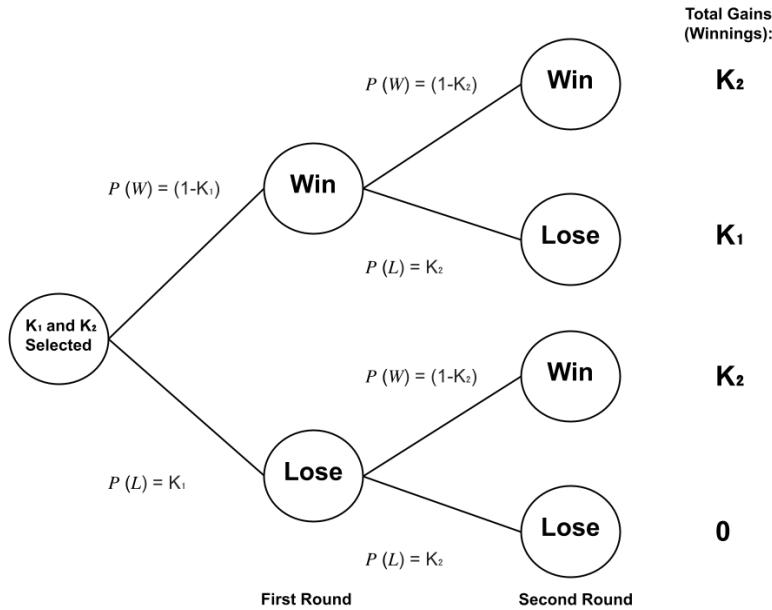
Situations 4 and 5 differ from situations 1 through 3 because in this situation, there will be more than one  $K$  value chosen. More specifically, before the game even begins, the participant will be able to choose two potentially different  $K$  values. In the previous games, there was only one selected value that was constant throughout all of the games, so this game is unique because there are multiple  $K$ 's being chosen at the beginning of the game that correlates to the number of rounds being played in the entirety of the game.

The only rules that must be followed are that the  $K$  values must be in the interval  $[0,1]$  and  $K_1$  (or the  $K$  value you will choose for round one) must be less than or equal to  $K_2$  (the  $K$  value for round two). This restriction relates back to our weightlifting analogy, where the competitor was not allowed to select a second weight that was lower than the initial chosen weight.

In the first round, the  $X_1$  value will be revealed and it will be compared to  $K_1$ . In order to determine whether you won or not, the same rules apply as those in situations 1 through 3, so if the  $X_1$  value is greater than the  $K_1$ , then you win and if it is smaller than  $K_2$ , you lose. In the second round, the same idea applies. The only difference is that you are using  $K_2$  and  $X_2$ .

In the second round, we will use the  $K_2$  value that was chosen at the very beginning of the game, regardless of whether we are on the win or lose node of our tree (shown in figure 5). If we are in the losing node, (because the first round was lost) we essentially have a second chance to win, but this time the  $K_2$  value instead of the  $K_1$  value. If the  $X_2$  value is greater than or equal to  $K_2$ , then we win  $K_2$ . If  $X_2$  is less than  $K_2$ , we end the game without any winnings. We can still only earn the highest  $K$  we win a round with, so either  $K_1$ ,  $K_2$ , or 0 is won at the end.

Therefore, out of the four paths that can be taken in this round, two will end with the participant winning  $K_2$  (WW) or (LW), one with the participant winning  $K_1$  (WL), and one with the participant winning 0 (LL).



**Figure 5. Tree showing all of the possible paths with 2 K values selected at the beginning and 2 random X values selected.**

When all four paths are synthesized together, the expected value equation is expressed in the following equation:

$$\mathbb{E}[\hat{G}]^{(2)} = K_2(1 - K_2) + K_1(K_2)(1 - K_1) + 0(K_1)(K_2).$$

From Figure 5, it is evident that there is a higher chance of winning  $K_2$  than of winning  $K_1$  or zero. This increases our expected value for this specific game because the  $K_2$  value is greater than the other values, so it would be our greatest possible earning. Figure 5 also assists in visualizing the expected value equation.

In order to find the values of  $K_1$  and  $K_2$  that will optimize our expected value, the expected value equation must be simplified:

$$\begin{aligned}\mathbb{E}[\hat{G}^{(2)}] &= K_2(1 - K_2) + K_1(K_2)(1 - K_1) \leq + 0(K_1)(K_2) \\ &= K_2(1 - K_2) + K_1(K_2)(1 - K_1).\end{aligned}$$

The simplified equation can now be differentiated. Since there are two variables in this equation, it is different from the previous situations. We cannot simply differentiate to find the optimal  $K_1$  and  $K_2$  values. To find the optimal  $K$  values, we must use partial differentiation to separate the variables. Then we set the derivative equal to 0 and solve for either  $K_1$  or  $K_2$ .

Let  $f(K_1, K_2) = K_2(1 - K_2) + K_1(K_2)(1 - K_1)$ .

To find the optimal value for  $K_1$ , we take the partial derivative of  $f$  with respect to  $K_1$  and set it equal to 0:

$$\frac{\partial f}{\partial K_1} = K_2 - 2K_2K_1$$

$$0 = K_2 - 2K_2K_1$$

$$-K_2 = -2K_2K_1$$

$$K_1 = \frac{1}{2}.$$

To find the optimal value for  $K_2$ , we take the partial derivative of  $f$  with respect to  $K_2$ :

$$f(K_1, K_2) = K_2(1 - K_2) + K_1(K_2)(1 - K_1)$$

$$\frac{\partial f}{\partial K_2} = 1 - 2K_2 + K_1(1 - K_1).$$

We then set the derivative equal to 0 and solve for  $K_2$ :

$$0 = 1 - 2K_2 + K_1(1 - K_1)$$

Substituting  $\frac{1}{2}$  for  $K_1$ :

$$0 = 1 - 2K_2 + \frac{1}{2}(1 - \frac{1}{2})$$

$$0 = 1 - 2K_2 + \frac{1}{4}$$

$$K_2 = \frac{5}{8}.$$

The partial differentiation steps above separated the expected value equation into parts to make it possible to solve for the optimal value of each  $K$  value. This process can be done for games with more rounds as well. The calculated partial derivatives above show that when  $K_1 = \frac{1}{2}$  and  $K_2 = \frac{5}{8}$ , the expected value of the game is maximized.

Inputting these values of  $K$  into the equation, we can find the maximum expected value:

$$\mathbb{E}[\hat{G}^{(2)}] = \frac{5}{8}(1 - \frac{5}{8}) + \frac{1}{2}(\frac{5}{8})(1 - \frac{1}{2}) + 0(\frac{1}{2})(\frac{5}{8})$$

$$= \frac{15}{64} + \frac{10}{64}$$

$$\mathbb{E}[\hat{G}^{(2)}] = \frac{25}{64}.$$

Hence, the maximum expected value occurs when  $K_1 = \frac{1}{2}$  and  $K_2 = \frac{5}{8}$ . The maximum expected value at this point is  $\frac{25}{64} \approx 0.39$ .

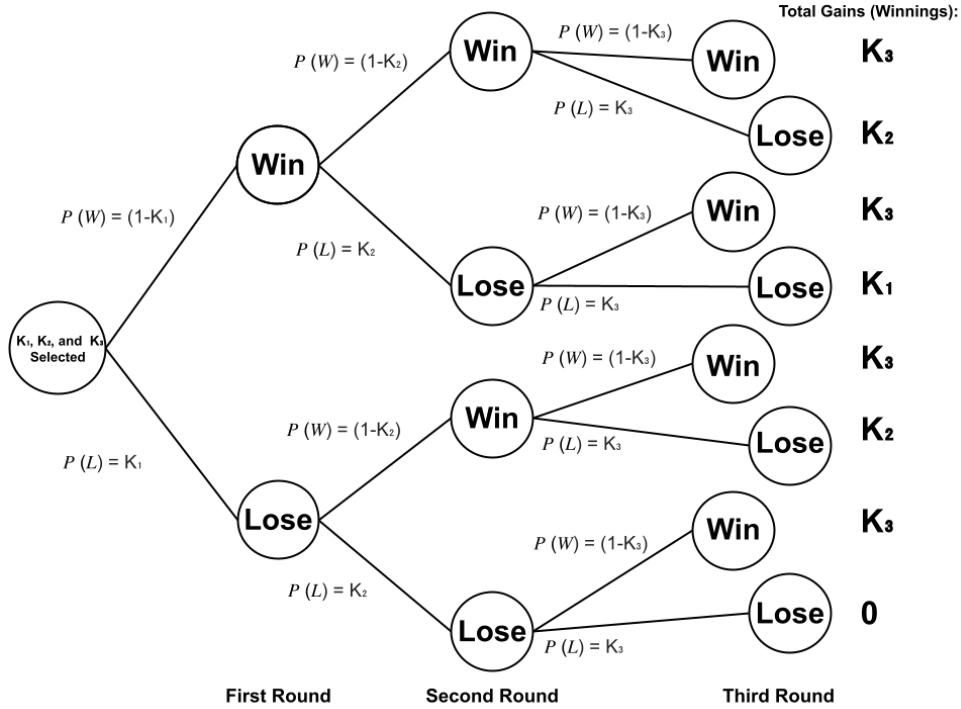
## 2. The Fifth Situation (Select $K_1 \leq K_2 \leq K_3$ , then 3 random variables, $X_1$ , $X_2$ , and $X_3$ are all selected)

In this variation of the game, all of the selected variables ( $K$ ) for each round are chosen before any of the random variables ( $X$ ) are selected. Each selected  $K$  value must be greater than or equal to the  $K$  value

from the previous round. The subscript denotes the round number that is being played, for example,  $K_3$  represents the  $K$  value for round 3.

To find the optimal  $K$  values for rounds 1, 2, and 3, we must first find the expected value equation. We do this in a similar way to the previous situations; by multiplying the probability of an occurrence by the amount gained for each path and adding all of the paths together. The following is the expected value equation for this situation.

$$\mathbb{E}[\hat{G}]^{(3)} = K_3(1 - K_3) + K_2(K_3)(1 - K_2) + K_1(K_2)(K_3)(1 - K_1).$$



**Figure 6. Tree showing all of the possible paths with 3  $K$  values selected at the beginning and 2 random  $X$  values selected.**

After simplification, the expected value equation can be differentiated using partial differentiation with respect to specific  $K$  values.

To find the optimal value for  $K_1$  we take the partial derivative with respect to  $K_1$  and set it equal to zero to maximize the expected value function:

$$\frac{\partial f}{\partial K_1} = K_2(K_3)(1 - 2K_1) = 0$$

$$K_1 = \frac{1}{2}.$$

To find the optimal value for  $K_2$  we take the partial derivative with respect to  $K_2$  and set it equal to zero to maximize the expected value function:

$$\frac{\partial f}{\partial K_2} = K_3(1 - 2K_2) + K_3(K_1 - K_1^2) = 0$$

$$K_3(1 - 2K_2 + K_1 - K_1^2) = 0.$$

Input the  $K_1$  value from the previous step to get the following value:

$$(1 - 2K_2 + \left(\frac{1}{2}\right) - \left(\frac{1}{2}\right)^2) = 0$$

$$2K_2 = \frac{5}{4}$$

$$K_2 = \frac{5}{8}.$$

To find the optimal value for  $K_3$  we take the partial derivative with respect to  $K_3$  and set it equal to zero to maximize the expected value function:

$$\frac{\partial f}{\partial K_3} = (1 - 2K_3) + (K_2 - K_2^2) + K_2(K_1 - K_1^2) = 0.$$

Input  $K_1$  and  $K_2$  values from the previous steps to solve for  $K_3$ :

$$(1 - 2K_3) + \left(\frac{5}{8} - \left(\frac{5}{8}\right)^2\right) + \frac{5}{8}\left(\frac{1}{2} - \left(\frac{1}{2}\right)^2\right) = 0$$

$$1 - 2K_3 = -\frac{25}{64}$$

$$-2K_3 = \frac{-89}{64}$$

$$K_3 = \frac{89}{128}.$$

These values of  $K$  can be substituted into the equation to find the maximum expected value:

$$\mathbb{E}[\hat{G}]^{(3)} = \frac{89}{128}(1 - \frac{89}{128}) + \frac{5}{8}(\frac{89}{128})(1 - \frac{5}{8}) + \frac{1}{2}(\frac{5}{8})(\frac{89}{128})(1 - \frac{1}{2})$$

$$= \frac{3471}{16384} + \frac{1780}{8912}$$

$$\mathbb{E}[\hat{G}]^{(3)} \approx 0.412.$$

Hence, the maximum expected value happens when  $K_1 = \frac{1}{2} \approx 0.5$ ,  $K_2 = \frac{5}{8} \approx 0.625$ , and  $K_3 = \frac{89}{128} \approx 0.690$ . The maximum expected value at this point is approximately 0.412.

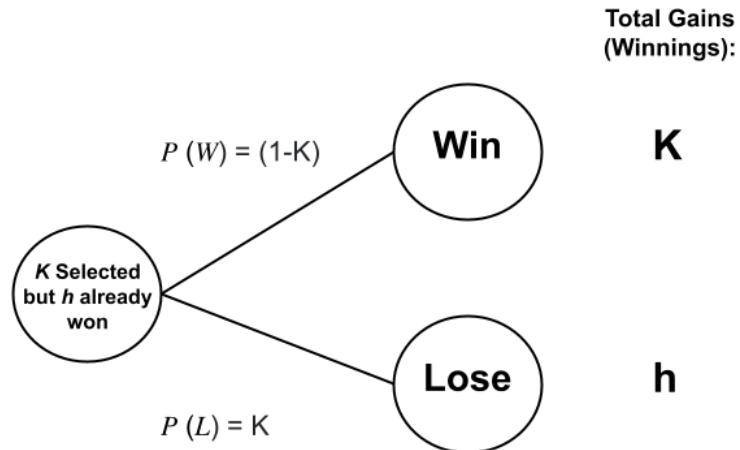
## C: Situations Involving Picking a $K$ Value Based on the Results of a Previous Round

### 1. The 6th Situation: (Select $K > h$ after having already won value $h$ )

This variation of the game is similar to situation 1, in that there is only one round of the game. However, unlike situation 1, the game starts with the participant already having won a value  $h$ . This scenario has the same probabilities as situation 1, where the probability of winning is  $(1 - K)$  and the probability of losing is still  $K$ . Unlike situation 1, we define  $H$  as the gain in this situation, which is still  $K$  when the game is won, but  $h$  when the game is lost, instead of zero.

Therefore the equation for the expected gain is now:

$$\mathbb{E}[H] = K(1 - K) + h(K).$$



**Figure 7. Tree showing the sixth situation involving an  $h$  value that will be won regardless of the following round(s).**

The expected gains equation can be simplified to the following:

$$\mathbb{E}[H] = K - K^2 + hK.$$

This can be differentiated, set equal to zero, and solved for the optimal value of  $K$  while treating  $h$  as a constant:

To find the highest expected value of this expression, let  $f(K) = K - K^2 + hK$

$$f(K) = K - K^2 + hK.$$

We then take the derivative of  $f(K)$ :

$$f'(K) = 1 - 2K + h.$$

Finally, we set the derivative equal to zero to find the optimal  $K$  value:

$$0 = 1 - 2K + h$$

$$K = \frac{1+h}{2}.$$

Thus, the optimal value for  $K$  with a value  $h$  already won is  $\frac{1+h}{2}$ . The maximum expected value can be found by substituting the optimal  $K$  value into the expected value formula:

$$\mathbb{E}[H] = \frac{1+h}{2} - \left(\frac{1+h}{2}\right)^2 + h\left(\frac{1+h}{2}\right)$$

Now, simplify this equation and combine all like terms to get:

$$\begin{aligned} &= \frac{1+h}{2} - \frac{1+2h+h^2}{4} + \frac{h+h^2}{2} \\ &= \frac{1+2h+h^2}{4} \\ &= \frac{(1+h)^2}{4}. \end{aligned}$$

The value above is equivalent to the maximum expected value or gains one can get when they already have won  $h$ . As observed above, the maximum expected value is dependent on the  $h$  value.

## **2. The 7th Situation: Choose $K$ values as rounds are played. $K$ values can change without restriction after each round depending on the result for only 2 rounds.**

This situation is represented by the notation  $\mathbb{E}[\check{G}^{(2)}]$  for the expected value where 2 represents the number of rounds.

In this variation of the game, after each round we can change the value of  $K$ . In this situation, the game has two rounds. After knowing the result of the first round, (meaning the participant has selected a  $K_1$  value and an  $X_1$  value has been revealed determining whether the participant won or lost) the participant can select a  $K_2$  value.

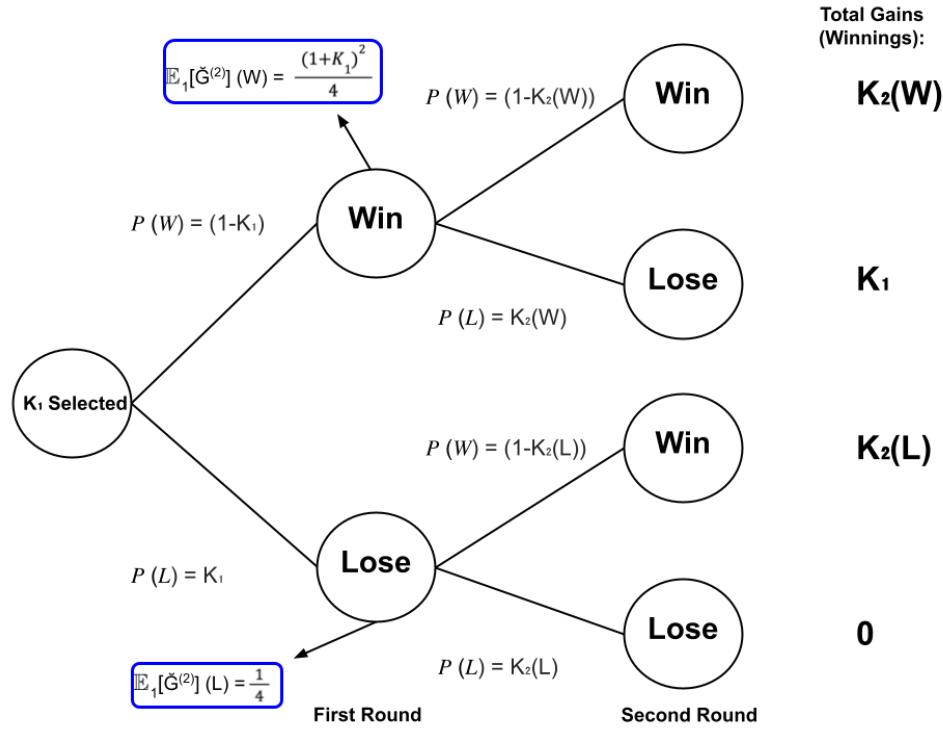
The selection of this  $K_2$  value is determinant of whether the participant won or lost the first round. Therefore the previously mentioned method of notation will be reintroduced. The participant's choice of  $K_2$  is based on the result of the first round. For example,  $K_2(L)$  represents the  $K_2$  value selected for the second round, based on the fact the first round was lost and  $K_2(W)$  represents the  $K_2$  value selected based on the case that the first round was won.

In order to represent this situation, the notation  $\mathbb{E}_\beta[\check{G}^{(n)}]$  will be used to denote the expected gains after  $n$  rounds of the game and based on the results of the previous game,  $\beta$ , where  $K$  is allowed to change based on the results of the previous game.

The expected value of this game can be represented as the sum of the conditional expected value of the path given you won the first round added to the conditional expected value of the path given you lost the

first round (this is what was modeled in the 6th situation and is shown in the blue boxes). This can be alternatively expressed as the sum of the expected gain of the game following winning the first round, and the expected gain of the game following losing the first round:

$$\mathbb{E}[\check{G}^{(2)}] = \mathbb{E}_1[\check{G}^{(2)}](W) + \mathbb{E}_1[\check{G}^{(2)}](L).$$



**Figure 8.** Tree showing the seventh situation involving a  $K$  value that changes depending on the results of the previous round.

Substituting values for these terms yields the following expression:

$$\mathbb{E}[\check{G}^{(2)}] = \frac{(1+K_1)^2}{4}(1 - K_1) + \frac{1}{4}(K_1).$$

The first value, or  $\frac{(1+K_1)^2}{4}$ , is the maximum gains possible based on having won the first round. This is shown in the sixth scenario.

The second value, or  $\frac{1}{4}$ , is the maximum gain of the game given we have lost the first round. The  $\mathbb{E}_1[\check{G}^{(2)}](L)$  tree from this situation is identical to the  $\mathbb{E}[G^1]$  tree from situation 1.

Simplifying the equation above gives us the following.

First, multiply out the terms:

$$\begin{aligned}\mathbb{E}[\check{G}^{(2)}] &= \frac{(1+K_1)^2}{4}(1-K_1) + \frac{1}{4}(K_1) \\ &= \frac{(1+K_1)^2(1-K_1)}{4} + \frac{K_1}{4} \\ &= \frac{(1+2K_1+K_1^2)(1-K_1)+K_1}{4} \\ &= \frac{1-K_1+2K_1-2K_1^2+K_1^2-K_1^3+K_1}{4}.\end{aligned}$$

Combine like terms in the numerator and factor out  $\frac{1}{4}$ :

$$\begin{aligned}&= \frac{-K_1^3 - K_1^2 + 2K_1 + 1}{4} \\ &= \left(\frac{1}{4}\right)(-K_1^3 - K_1^2 + 2K_1 + 1).\end{aligned}$$

Now that the equation has been simplified, the derivative can be found to give us the value of  $K_1$  that will yield the maximum expected value:

Let,  $f(x) = \left(\frac{1}{4}\right)(-x^3 - x^2 + 2x + 1)$ , and differentiate:

$$f'(x) = -3x^2 - 2x + 2.$$

Set this derivative to maximize the expected value, and solve for x:

$$0 = 3x^2 + 2x - 2$$

$$x = \frac{-2 \pm \sqrt{(2)^2 - 4(3)(-2)}}{6}$$

$$x = -\frac{1}{3} \pm \frac{2\sqrt{7}}{6}.$$

Let,  $K_1 = -\frac{1}{3} + \frac{2\sqrt{7}}{6}$

$$K_1 \approx .549.$$

This is the value of  $K_1$  that yields the maximum expected value in this situation. We can find the maximum expected value by substituting the optimal  $K_1$  into the expected value equation.

$$\begin{aligned}
\mathbb{E}[\check{G}^{(2)}] &= \frac{(1+K_1)^2}{4}(1-K_1) + \frac{1}{4}(K_1) \\
&= \frac{(1 + (-\frac{1}{3} + \frac{2\sqrt{7}}{6}))^2}{4}(1 - (-\frac{1}{3} + \frac{2\sqrt{7}}{6})) + \frac{1}{4}(-\frac{1}{3} + \frac{2\sqrt{7}}{6}) \\
&= \frac{1}{4}(-\frac{1}{3} + \frac{\sqrt{7}}{3}) + \frac{1}{4}(\frac{4}{3} - \frac{\sqrt{7}}{3})(1 + (-\frac{1}{3} + \frac{\sqrt{7}}{3})^2) \\
\mathbb{E}[\check{G}^{(2)}] &\approx 0.284.
\end{aligned}$$

This is the maximum expected value when the following three conditions are met: the  $K$  values are chosen as rounds are played, the  $K$  values can change after each round depending on the result, and two rounds are played.

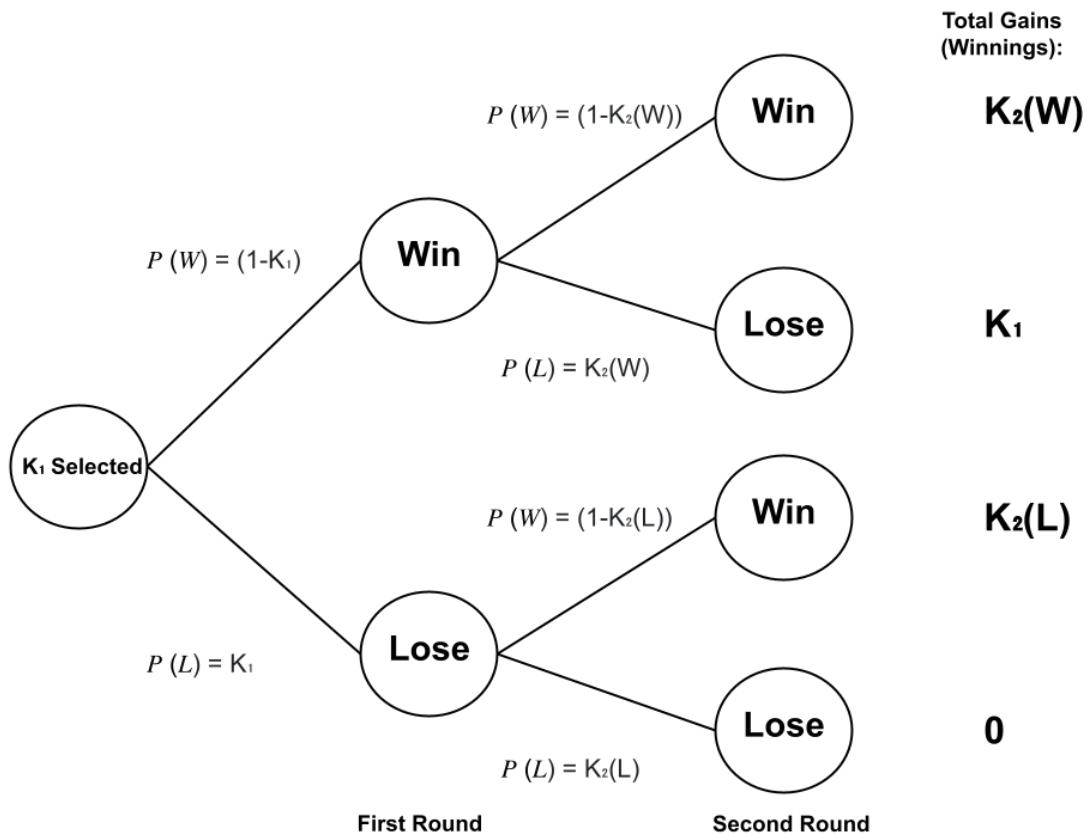
#### **4. The 8th Situation: ( $K$ values chosen after each round depending on the result, for 2 rounds with restriction $K_1 \leq K_2$ )**

This situation is represented by the notation  $\mathbb{E}[\check{G}^{(2)}]$  for the expected value where 2 represents the number of rounds.

This variation of the game is similar to the 7th situation where you choose the  $K$  value after each previous round is played. In this situation, the  $K$  value must be greater than the one before it.

When playing for two rounds, we will find an optimal value for  $K_1$ , and then  $K_2$  will change depending on whether the first round was won or lost. We use  $K_2(W)$  to denote the  $K_2$  value if the first round was won and  $K_2(L)$  to denote the  $K_2$  value if the first round was lost.

In the previous situation, the ideal choice for  $K_2(L)$  was found to be  $\frac{1}{2}$ . However, in the previous situation, the  $K_1$  value was found to be higher than  $\frac{1}{2}$ . Since there is a restriction that  $K_1 \leq K_2$ , the optimal value of  $K_2(L)$  is equivalent to  $\frac{1}{2}$  or  $K_1$ .



**Figure 9. Tree showing the eighth situation involving a  $K$  value that changes depending on the results of the previous round (but the  $K$  value must always be greater than the preceding one).**

If we assume  $K_2(L) = K_1$  and keep  $K_2(W) = \frac{1+K_1}{2}$ , the expected value can be split into the expected gain of the second round after winning the first round multiplied by the probability of winning the first round, added to the expected gain of the second round after losing the first round multiplied by the probability of losing the first round. This is depicted below:

$$\begin{aligned} \mathbb{E}[\check{G}^{(2)}] &= \mathbb{E}[\check{G}^{(1)}](W) * P(W) + \mathbb{E}[\check{G}^{(1)}](L) * P(L) \\ &= \frac{(1+K_1)^2}{4} * (1 - K_1) + K_1(1 - K_1)(K_1). \end{aligned}$$

Like terms in the equation are combined and  $\frac{1}{4}$  is factored out in the following work:

$$\begin{aligned} &= \frac{1 + 2K_1 + K_1^2 - K_1 - 2K_1^2 - K_1^3}{4} + K_1^2 - K_1^3 \\ &= \frac{1}{4}(1 + K_1 + 3K_1^2 - 5K_1^3). \end{aligned}$$

This can be differentiated in terms of  $K_1$  in order to find the optimal value of  $K_1$ .

Let  $f(K_1) = \frac{1}{4}(1 + K_1 + 3K_1^2 - 5K_1^3)$ .

Then, find the derivative:

$$f'(K_1) = \frac{1}{4}(1 + 6K_1 - 15K_1^2).$$

Set the derivative of  $f$  equal to zero in order to find the value of  $K_1$  which maximizes the function:

$$0 = \frac{1}{4}(1 + 6K_1 - 15K_1^2)$$

$$K_1 = K_2(L) = \frac{3+2\sqrt{6}}{15}$$

$$K_2(W) = \frac{(1+K_1)}{2} = \frac{9+\sqrt{6}}{15}.$$

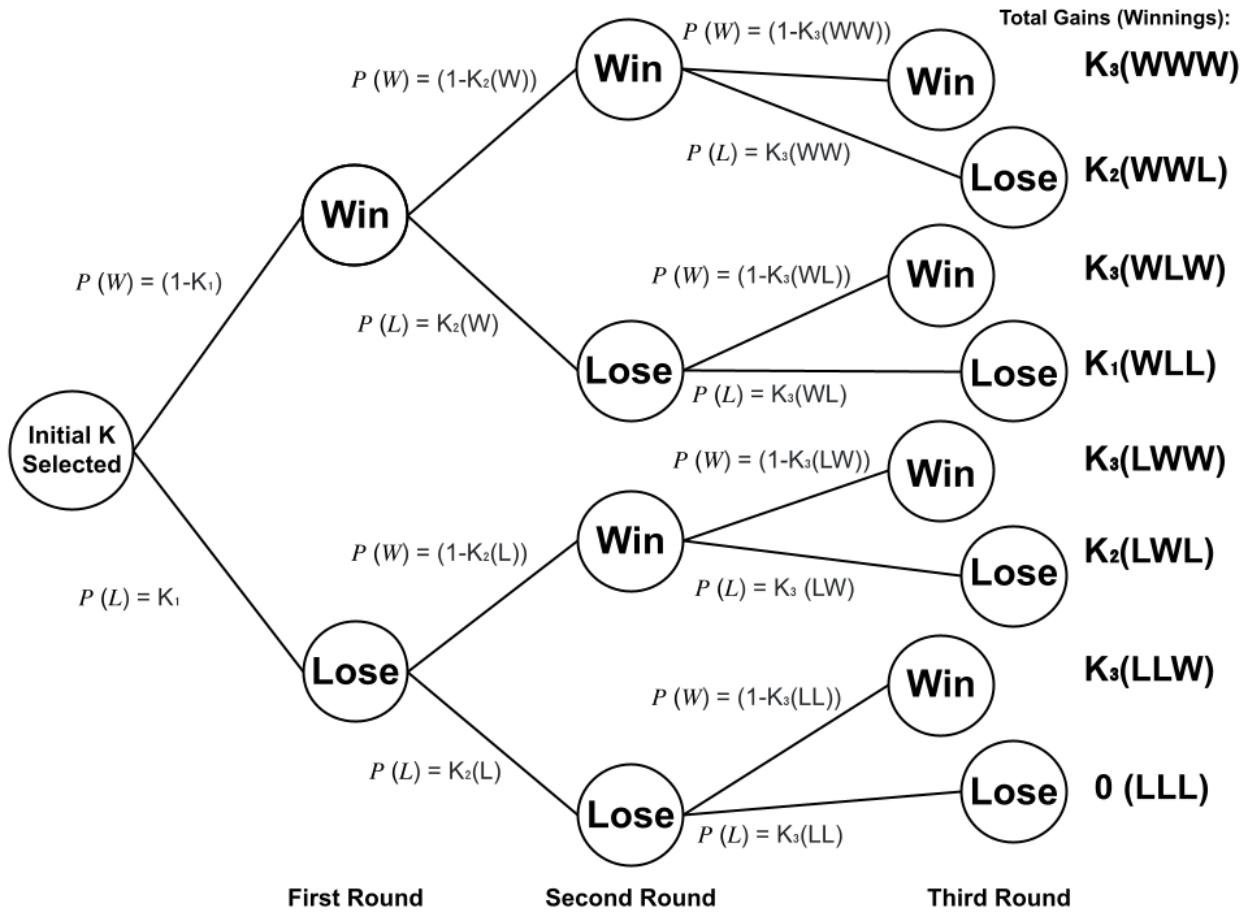
The optimal value for  $K_1$  was found to be  $\frac{3+2\sqrt{6}}{15} \approx 0.5265$  which is greater than  $\frac{1}{2}$ . Therefore, the value of  $K_2(L) = K_1$ . The optimal value for  $K_2(W)$  was found to be  $\frac{9+\sqrt{6}}{15} \approx 0.7633$ .

##### **5. The 9th Situation: ( $K$ values can change after each round depending on the result, for 3 rounds. With restriction $K_1 \leq K_2 \leq K_3$ )**

This situation is represented by the notation  $\check{G}^{(3)}$  which is the random variable that represents the amount we win or our gains.  $E[\check{G}^{(3)}]$  is the expected value when 3 rounds are played in the game. The subscript of  $E$  represents the round at which the expected value is being computed.

The 9th situation is similar to the 8th situation because it has the same restrictions where the new  $K$  values are chosen each subsequent round. This situation differs, however, because there are three rounds rather than two.

There will be a singular optimal  $K_1$  value, then two optimal values for  $K_2$ , because the winning node and losing node will differ from each other. These will be represented in the same way as the last situation, having  $K_2(W)$  denoting  $K_2$  after the first round is won, and  $K_2(L)$  denoting  $K_2$  after the first round is lost. There are four different optimal choices for  $K_3$  because they depend on the sequence of wins and losses before them. These will be represented by  $K_3(WW)$ ,  $K_3(WL)$ ,  $K_3(LW)$ , and  $K_3(LL)$ . The sequence of wins and loses within the parenthesis go in order from left to right from round one to round two. For example, (WL) represents winning the first round and losing the second.



**Figure 10. Tree showing the ninth situation involving 3  $K$  values that are dependent on the results of the previous round (but the  $K$  value must always be greater than the preceding one).**

To generate the expected value equation for this situation, we must represent each path that can be taken. The expected value equation shown below shows the sum of the expected values of the win and lose branches of the tree shown in figure 10.

$$\mathbb{E}[\check{G}^{(3)}] = \mathbb{E}_2[\check{G}^{(3)}](W)(1 - K_1) + \mathbb{E}_2[\check{G}^{(3)}](L)(K_1).$$

To find the expected values after winning or losing a round in this situation, we can further expand our equation for the expected value and take the sum of all expected values of the rounds. The following expression for  $\mathbb{E}[\check{G}^{(3)}]$  consists of all the expected values along all the branches multiplied by all their probabilities.

$$\mathbb{E}[\check{G}^{(3)}] = (\mathbb{E}_3[\check{G}^{(3)}](WW)(1 - K_2 W) + \mathbb{E}_3[\check{G}^{(3)}](WL)(K_2 W)(1 - K_1)) + (\mathbb{E}_3[\check{G}^{(3)}](LW)(1 - K_2 L) + \mathbb{E}_3[\check{G}^{(3)}](LL)(K_2 L)(K_1))$$

In order to find the  $\mathbb{E}_3[\check{G}^{(3)}](WW)$  equation, we multiply the gain,  $K_3(WW)$ , if we win round 3 with the probability of winning round 3,  $(1 - K_3(WW))$ , and add it to the gain if we lose,  $K_2(W)$  times the probability of losing,  $(K_3(WW))$ :

$$\begin{aligned}\mathbb{E}_3[\check{G}^{(3)}](WW) &= K_3(WW) * (1 - K_3(WW)) + K_2(W) * (K_3(WW)) \\ &= K_3(WW) - (K_3(WW))^2 + K_3(WW) * (K_2(W)).\end{aligned}$$

Let  $x = K_3(WW)$ .

To find the highest expected value of this expression, let  $f(x) = x - x^2 + (K_2(W))(x)$ :

$$f(X) = x - x^2 + (K_2(W))(x)$$

We then take the derivative of  $f(x)$ :

$$f'(X) = 1 - 2x + (K_2(W)).$$

Finally, we set the derivative equal to zero and solve for  $x$  to find the optimal  $K_3(WW)$  value:

$$0 = 1 - 2x + (K_2(W))$$

$$x = \frac{1+K_2(W)}{2}.$$

Plug  $K_3(WW)$  back in for  $x$

$$K_3(WW) = \frac{1+K_2(W)}{2}.$$

Plugging in the optimal  $K_3(WW)$  value into the  $\mathbb{E}_3[\check{G}^{(3)}](WW)$  equation:

$$\begin{aligned}\mathbb{E}_3[\check{G}^{(3)}](WW) &= K_3(WW)(1 - K_3(WW)) + K_2(W)(K_3(WW)) \\ &= (\frac{1+K_2(W)}{2})(1 - (\frac{1+K_2(W)}{2})) + K_2(W)(\frac{1+K_2(W)}{2}) \\ &= \frac{(1+K_2(W))^2}{4}.\end{aligned}$$

In order to find the  $\mathbb{E}_3[\check{G}^{(3)}](WL)$  equation, we multiply the gain,  $K_3(WL)$ , if we win round 3 with the probability of winning round 3,  $(1 - K_3(WL))$ , and add it to the gain if we lose,  $K_1$ , and multiply it by the probability of losing,  $(K_3(WL))$ :

$$\begin{aligned}\mathbb{E}_3[\check{G}^{(3)}](WL) &= K_3(WL)(1 - K_3(WL)) + K_1(K_3(WL)) \\ &= K_3(WL) - (K_3(WL))^2 + K_3(WL)(K_1).\end{aligned}$$

Let  $x = K_3(WL)$ .

To find the highest expected value of this expression, let  $f(x) = x - x^2 + (K_1)(x)$ .

$$f(x) = x - x^2 + (K_1)(x)$$

We then take the derivative of  $f(x)$ :

$$f'(x) = 1 - 2x + K_1$$

Finally, we set the derivative equal to zero and solve for  $x$  to find the optimal  $K_3(WL)$  value:

$$0 = 1 - 2x + K_1$$

$$x = \frac{1 + K_1}{2}$$

Plug  $K_3(WL)$  back in for  $x$

$$K_3(WL) = \frac{1+K_1}{2}.$$

Plugging in the found  $K_3(WL)$  into the  $\mathbb{E}_3[\check{G}^{(3)}](WL)$  equation:

$$\begin{aligned}\mathbb{E}_3[\check{G}^{(3)}](WL) &= K_3(WL)(1 - K_3(WL)) + K_1(K_3(WL)) \\ \mathbb{E}_3[\check{G}^{(3)}](WL) &= \left(\frac{1 + K_1}{2}\right)\left(1 - \left(\frac{1 + K_1}{2}\right)\right) + K_1\left(\frac{1 + K_1}{2}\right) \\ \mathbb{E}_3[\check{G}^{(3)}](WL) &= \frac{(1+K_1)^2}{4}.\end{aligned}$$

In order to find the  $\mathbb{E}_3[\check{G}^{(3)}](LW)$  equation, we multiply the gain,  $K_3(LW)$ , if we win round 3 with the probability of winning round 3,  $(1 - K_3(LW))$ , and add it to the gain if we lose,  $K_2(L)$ , and multiply it by the probability of losing,  $(K_3(LW))$ :

$$\begin{aligned}\mathbb{E}_3[\check{G}^{(3)}](LW) &= K_3(LW)(1 - K_3(LW)) + K_2(L)(K_3(LW)) \\ &= K_3(LW) - K_3(LW)^2 + K_3(LW)(K_2(L)).\end{aligned}$$

Let  $x = K_3(LW)$ .

To find the highest expected value of this expression, let  $f(x) = x - x^2 + (K_2(L))(x)$ ,

$$f(x) = x - x^2 + (K_2(L))(x).$$

We then take the derivative of  $f(x)$ :

$$f'(x) = 1 - 2x + K_2(L).$$

Finally, we set the derivative equal to zero and solve for  $x$  to find the optimal  $K_3(LW)$  value:

$$0 = 1 - 2x + K_2(L)$$

$$x = \frac{1+K_2(L)}{2}.$$

Plug  $K_3(LW)$  back in for  $x$

$$K_3(LW) = \frac{1+K_2(L)}{2}.$$

Plugging in the found  $K_3(LW)$  into the  $\mathbb{E}[\check{G}^{(3)}](LW)$  equation:

$$\mathbb{E}[\check{G}^{(3)}](LW) = K_3(LW)(1 - K_3(LW)) + K_2(L)(K_3(LW))$$

$$\mathbb{E}[\check{G}^{(3)}](LW) = \frac{1+K_2(L)}{2}(1 - (\frac{1+K_2(L)}{2})) + K_2(L)(\frac{1+K_2(L)}{2})$$

$$\mathbb{E}[\check{G}^{(3)}](LW) = \frac{(1+K_2(L))^2}{4}$$

Solving for  $\mathbb{E}[\check{G}^{(3)}](LL)$  given  $K_2 \leq \frac{1}{2}$ :

**Note:** This computation is identical mathematically to the one completed in situation 1. The optimal  $K_3$  is  $\frac{1}{2}$  and the expected value is  $\frac{1}{4}$ , just like the optimal value of the initial  $K_3$  and the expected value of the first situation. The computations included below are included for consistency:

$$\begin{aligned}\mathbb{E}[\check{G}^{(3)}](LL) &= K_3(LL)(1 - K_3(LL) + 0(K_3(LL))) \\ &= K_3(LL) - (K_3(LL))^2.\end{aligned}$$

Let  $x = K_3(LL)$ .

To find the highest expected value of this expression, let  $f(x) = x - x^2$ :

$$f(x) = x - x^2.$$

We then take the derivative of  $f(x)$ :

$$f'(x) = 1 - 2x.$$

Finally, we set the derivative equal to zero and solve for  $x$  to find the optimal  $K_3(LL)$  value:

$$0 = 1 - 2x$$

$$x = \frac{1}{2}.$$

Plug  $K_3(LL)$  back in for  $x$ :

$$K_3(LL) = \frac{1}{2}.$$

Plugging the found  $K_3(LL)$  into the  $\mathbb{E}[\check{G}^{(3)}](LL)$  equation given the ideal situation, the  $\mathbb{E}[\check{G}^{(3)}](LL)$  is found to be:

$$\mathbb{E}[\check{G}^{(3)}](LL) = K_3(LL)(1 - K_3(LL) + 0(K_3(LL)))$$

$$\mathbb{E}[\check{G}^{(3)}](LL) = \left(\frac{1}{2}\right)(1 - \left(\frac{1}{2}\right) + 0\left(\frac{1}{2}\right))$$

$$\mathbb{E}[\check{G}^{(3)}](LL) = \frac{1}{4}.$$

Solving for  $\mathbb{E}[\check{G}^{(3)}](LL)$  given  $K_2 \geq \frac{1}{2}$ :

$$K_3(LL) = K_2(L).$$

Plugging the found  $K_3(LL)$  into the  $\mathbb{E}[\check{G}^{(3)}](LL)$  equation given the ideal situation, the  $\mathbb{E}[\check{G}^{(3)}](LL)$  is found to be:

$$\mathbb{E}[\check{G}^{(3)}](LL) = K_3(LL)(1 - K_3(LL) + 0(K_3(LL)))$$

$$\mathbb{E}[\check{G}^{(3)}](LL) = (K_2(L))(1 - (K_2(L)) + 0(K_2(L)))$$

$$\mathbb{E}[\check{G}^{(3)}](LL) = K_2(L) - K_2(L)^2.$$

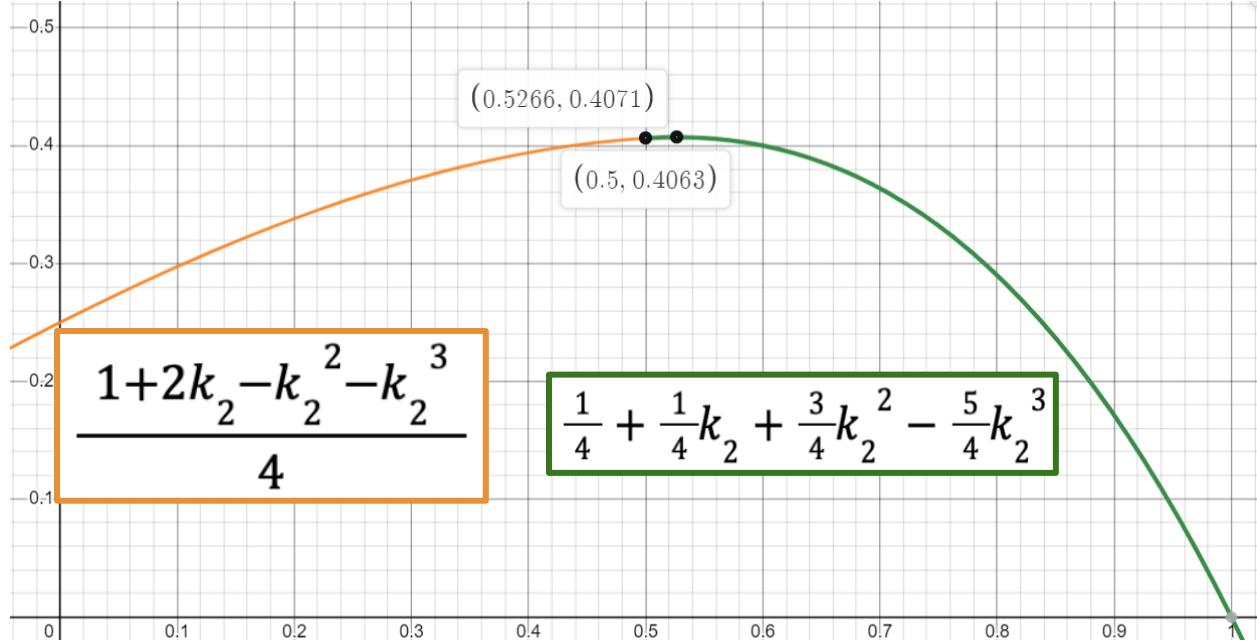
Finding the lose branch  $\mathbb{E}[\check{G}^{(3)}](LL)$  given  $K_2 \geq \frac{1}{2}$ :

$$\begin{aligned} \mathbb{E}_2[\check{G}^{(3)}](L) &= (K_2(L) - K_2(L)^2)(K_2(L) + \left(\frac{(1+K_2(L))^2}{4}\right)(1 - K_2(L))) \\ &= \frac{1}{4} + \frac{1}{4}K_2(L) + \frac{3}{4}K_2(L)^2 - \frac{5}{4}K_2(L)^3. \end{aligned}$$

Finding the lose branch  $\mathbb{E}[\check{G}^{(3)}](LL)$  given  $K_2 \leq \frac{1}{2}$ :

$$\mathbb{E}_2[\check{G}^{(3)}](L) = \left(\frac{1}{4}\right)(K_2(L) + \left(\frac{(1+K_2(L))^2}{4}\right)(1 - K_2(L)))$$

$$= \frac{1+2K_2(L)-K_2(L)^2-K_2(L)^3}{4}.$$



**Figure 11: A graphical representation of the optimal value for  $K_2$  that would result in the maximum expected volume.**

The optimal value of  $K_2(L)$  is 0.5266, but we have to consider the case where  $K_1$  is greater than 0.5266. If this is the case, you would select a  $K_2(L)$  value equal to  $K_1$ , as  $K_1$  be as close to our ideal value of 0.5266 as possible.

Now using the optimal values of  $K_1$  to maximize  $\mathbb{E}_2[\check{G}^{(3)}](WW)$  and  $\mathbb{E}_2[\check{G}^{(3)}](WL)$  we can find the expected value of the win branch.

$$\mathbb{E}_2[\check{G}^{(3)}](W) = \left(\frac{(1+K_1)^2}{4}\right)(K_2(W) + \frac{(1+K_2(W))^2}{4}(1-K_2(W)):$$

$$\begin{aligned} \mathbb{E}_2[\check{G}^{(3)}](W) = & \\ & \frac{1}{4} \left( \frac{2}{3} + \frac{1}{3} \sqrt{3k_1^2 + 6k_1 + 7} \right)^2 \left( \frac{4}{3} - \frac{1}{3} \sqrt{3k_1^2 + 6k_1 + 7} \right) \\ & + \frac{1}{4} (1+k_1)^2 \left( \frac{1}{3} \sqrt{3k_1^2 + 6k_1 + 7} - \frac{1}{3} \right) \end{aligned}$$

Now, we can take our expected gains of the win and lose branches to find the expected gains of the entire three-round game.

$$\mathbb{E}[\check{G}^{(3)}] = \mathbb{E}_2[\check{G}^{(3)}](W)(1-K_1) + \mathbb{E}_2[\check{G}^{(3)}](L)(K_1).$$

Given  $K_1 \leq 0.5266$ :

$$\begin{aligned} & \frac{7}{108} + \frac{239}{2700} \hat{k}_1 + \frac{7}{54} \sqrt{3 \hat{k}_1^2 + 6 \hat{k}_1 + 7} \\ & - \frac{1}{54} \hat{k}_1 \sqrt{3 \hat{k}_1^2 + 6 \hat{k}_1 + 7} - \frac{1}{18} \hat{k}_1^2 \sqrt{3 \hat{k}_1^2 + 6 \hat{k}_1 + 7} \\ & - \frac{1}{18} \hat{k}_1^3 \sqrt{3 \hat{k}_1^2 + 6 \hat{k}_1 + 7} + \frac{1}{12} \hat{k}_1^2 + \frac{1}{12} \hat{k}_1^3 \\ \mathbb{E}[\check{G}^{(3)}] = & + \frac{8}{225} \hat{k}_1 \sqrt{6} \end{aligned}$$

Given  $K_1 \geq 0.5266$ :

$$\begin{aligned} & \frac{7}{108} + \frac{1}{54} \hat{k}_1 + \frac{7}{54} \sqrt{3 \hat{k}_1^2 + 6 \hat{k}_1 + 7} \\ & - \frac{1}{54} \hat{k}_1 \sqrt{3 \hat{k}_1^2 + 6 \hat{k}_1 + 7} - \frac{1}{18} \hat{k}_1^2 \sqrt{3 \hat{k}_1^2 + 6 \hat{k}_1 + 7} \\ \mathbb{E}[\check{G}^{(3)}] = & - \frac{1}{18} \hat{k}_1^3 \sqrt{3 \hat{k}_1^2 + 6 \hat{k}_1 + 7} + \frac{1}{3} \hat{k}_1^2 + \frac{5}{6} \hat{k}_1^3 - \frac{5}{4} \hat{k}_1^4 \end{aligned}$$

Now we can take the derivative of the expected value equations given  $K_1 \geq 0.5266$  or  $K_1 \leq 0.5266$ .

$$\left\{ \begin{array}{l} \frac{1}{2700} \frac{700 - 1500 \hat{k}_1 + 239 \sqrt{3 \hat{k}_1^2 + 6 \hat{k}_1 + 7} - 5700 \hat{k}_1^2 - 4500 \hat{k}_1^3 + 675 \hat{k}_1^2 \sqrt{3 \hat{k}_1^2 + 6 \hat{k}_1 + 7} - 1800 \hat{k}_1^4 + 450 \hat{k}_1 \sqrt{3 \hat{k}_1^2 + 6 \hat{k}_1 + 7} + 96 \sqrt{6} \sqrt{3 \hat{k}_1^2 + 6 \hat{k}_1 + 7}}{\sqrt{3 \hat{k}_1^2 + 6 \hat{k}_1 + 7}} \quad \hat{k}_1 < \frac{1}{5} + \frac{2}{15} \sqrt{6} \\ - \frac{1}{54} \frac{30 \hat{k}_1 + 114 \hat{k}_1^2 + 90 \hat{k}_1^3 - \sqrt{3 \hat{k}_1^2 + 6 \hat{k}_1 + 7} + 270 \hat{k}_1^3 \sqrt{3 \hat{k}_1^2 + 6 \hat{k}_1 + 7} + 36 \hat{k}_1^4 - 135 \hat{k}_1^2 \sqrt{3 \hat{k}_1^2 + 6 \hat{k}_1 + 7} - 36 \hat{k}_1 \sqrt{3 \hat{k}_1^2 + 6 \hat{k}_1 + 7} - 14}{\sqrt{3 \hat{k}_1^2 + 6 \hat{k}_1 + 7}} \quad \frac{1}{5} + \frac{2}{15} \sqrt{6} < \hat{k}_1 \end{array} \right.$$

## 6. Newton's Method

As shown above, the derivative of the expected value function of the three-round game is extremely complicated and it would be very difficult to find the zeros by hand. Instead, in order to approximate the point at which the expected value equation is maximized, we employed a method known as Newton's method. Newton's method helps approximate the x-value at which a function intersects with the x-axis (or when the y value equals 0).<sup>3</sup> We used Newton's Method on the derivative of the expected value function, and by computing where the derivative equals 0, we can find the maximum value of the expected value function. Newton's Method consists of the following formula(for any arbitrary function  $f(x)$ ):

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

We used a computer program called Maple to write a computer program to iteratively use this formula and find an approximate solution for when the derivative of the expected value function equals 0.

A Newton's Method algorithm/function is defined in order to numerically maximize Expected Value functions.

```

newtonsMethod := proc(func, x0, maxDiff, maxIter);
  xn := x0;
  for i from 0 to maxIter do
    #Iterates through maxIter times
    if abs(evalf(eval(diff(func(x), x), x=xn))) < maxDiff then
      #Checks if function value is close to 0
      print("Found solution in iterations:");
      print(i);
      print("The solution is:");
      return xn;
    end if;
    xn := evalf(xn - evalf(diff(func(x), x), x=xn) / evalf(diff(func(x), x$2), x=xn));
    #Executes Newton's Method formula
  end do;
  print("Exceeded maximum iterations");
  return Null;
end proc;

```

**Figure 12. Code of Newton's Method Algorithm.**

Using the aforementioned Newton's Method algorithm, we were able to find the optimal value for the expected value function and  $K_1$ :

$$\begin{aligned} K_1 &\approx 0.551601 \\ \mathbb{E}[\check{G}^{(3)}] &\approx 0.512766 \end{aligned}$$

Using the calculated  $K_1$  value we can also evaluate all following  $K_2$  and  $K_3$  values.

$$\begin{aligned} K_1 &\approx K_2(L) \approx K_3(LL) \approx 0.551601 \\ K_2(W) &\approx K_3(WL) \approx 0.783328 \\ K_3(LW) &\approx 0.775800 \\ K_3(WW) &\approx 0.891664 \end{aligned}$$

## IV. Conclusion

### A: Situations Involving Picking 1 K Value

In this scenario, we are choosing our  $K$ 's before we know the outcomes of the individual rounds. We found that there was an interesting pattern that occurred as we increased the number of rounds in the game. When the game is played  $n$  times, we found that the expected value for this scenario is:

$$\mathbb{E}[G^n] = K - K^{n+1}$$

We found that the equation below can be used to find the optimal  $K$  value for any number of rounds,  $n$ .

$$K = \frac{1}{n\sqrt{n+1}}.$$

As more rounds of this game are played, the optimal  $K$  value approaches one.

$$\lim_{n \rightarrow \infty} \frac{1}{(n+1)^{(1/n)}} = \frac{1}{e}$$

### B: Situations Involving Picking More Than 1 K Value at the Beginning

We found the optimal values for playing three rounds, however, we did not find a general equation for this pattern. If you pick distinct  $K$  values before playing, the optimal choices include  $K_1 = \frac{1}{2}$ ,  $K_2 = \frac{5}{8}$ , and  $K_3 = \frac{89}{128}$ .

We came to the conclusion that these choices will give the highest expected value for this game, and will stay the same as the first three values no matter how many rounds are played. So, no matter how many rounds of this game are played, if all the  $K_n$  values are chosen at the beginning, the best  $K_1$  value will always be  $\frac{1}{2}$ , etc.

Partial differentiation continues to be necessary as the number of rounds increases. The variables found in the smaller number of rounds will stay consistent. The process will get more complex as the number of rounds increases because the number of variables present will increase as well.

### C: Situations Involving Picking a K Value Based on the Results of a Previous Round

The ultimate goal of this research paper was to optimize the expected value for this type of game. This type of game mirrors the original inspiration of weightlifting and uses the 3 attempt format. Since we knew the earnings for each possibility of  $K$  in each round, we were able to work backward to try and find the optimal first choice for  $K_1$ . We combined the final sets of branches into larger ones working backward toward the initial decision. We had to use a piecewise function in the (LL) branch's expected value equations. This is because of the constraint  $K_1 \leq K_2 \leq K_3$ . If the previous choice exceeded the optimal choices we found previously, we would have to choose the previous  $K$  value again despite it not being optimal, since it is closest to the optimal  $K$  value. Then we used Newton's Method to find when the derivative of the expected value function for the three-round game was equal to 0. This gave us a strong approximation for our optimal  $K_1$  value. Finally, we used the optimal  $K_1$  value to determine the optimal  $K$  value for each following round.

## V. Acknowledgments

First, we would like to thank every member of the PGSS program staff, PGSS alumni association, and the PGSS Foundation for putting in countless hours to plan this program and their funding. We are immensely grateful to have this opportunity. We would also like to thank Dr. Barry Luokkala for providing us with the opportunity to come together and learn, Dr. David Handron for assisting us in our inspiration, research, presentation, and paper, and Justine Dell for being our guide and mentor in creating a presentation and report on mathematics. In addition, we would like to thank our friends and family for creating an environment where we are encouraged to apply and complete this program to the best of our abilities. We could not have completed this research without any of you.

## VI. References

- <sup>1</sup> Baudoin, F. (2010). Stochastic processes. Stochastic Processes - an overview | ScienceDirect Topics. Retrieved July 27, 2022, from <https://www.sciencedirect.com/topics/neuroscience/stochastic-processes>
- <sup>2</sup> Russell, J. (2019, January 9). Mathematical modeling of attempt selection: Jeff Russell. Starting Strength. Retrieved July 27, 2022, from <https://startingstrength.com/article/mathematical-modeling-of-attempt-selection>
- <sup>3</sup> Neuhauser, C., & Roper, M. L. (2019). Calculus for biology and medicine (Third). Pearson.
- <sup>4</sup> Goodman, E. (2021, October 8). Weightlifting 101: Rules. Paris 2024 Olympic Games. Retrieved July 28, 2022, from <https://www.nbcolympics.com/news/weightlifting-101-rules>
- <sup>5</sup> Kenton, W (2022, March 10). Expected Value (EV). Investopedia. <https://www.investopedia.com/terms/e/expected-value.asp>

## Appendix: Copy of the Complete Maple Code

A Newton's Method algorithm/function is defined in order to numerically maximize Expected Value functions.

```
newtonsMethod := proc(func, x0, maxDiff, maxIter);
  xn := x0;
  for i from 0 to maxIter do
    #Iterates through maxIter times
    if abs(evalf(eval(diff(func(x), x), x = xn))) < maxDiff then
      #Checks if function value is close to 0
      print("Found solution in iterations.");
      print(i);
      print("The solution is:");
      return xn;
    end if;
    xn := evalf(xn - evalf((eval(diff(func(x), x), x = xn)) / eval(diff(func(x), x$2), x = xn)));
    #Executes Newton's Method formula
  end do;
  print("Exceeded maximum iterations");
  return Null;
end proc;
```

with(Optimization);  
 [ImportMPS, Interactive, LPSolve, LSSolve, Maximize, Minimize, NLPSolve, QPSolve] (1)

2 Games, can change value of k after 1st game, no restriction on  $k_2$

$$e_2 := \left( \frac{(1+x)^2}{4} \cdot (1-x) \right) + \frac{x}{4} \quad (2)$$

$$\frac{1}{4} (1+x)^2 (1-x) + \frac{1}{4} x$$

$$\text{newtonsMethod}\left(e_2, \frac{1}{2}, 0.00000001, 10\right); \quad (3)$$

$$0.5485837702$$

$$\text{Maximize}(e_2(x), \{0 \leq x, x \leq 1\}); \quad (4)$$

$$[0.407782577360224718, [x = 0.548583770354858]]$$

$$\text{evalf}\left(\frac{-1}{3} + \frac{1}{3} \sqrt{7}\right) \quad (5)$$

$$0.5485837703$$

2 Games, can change value of k after 1st game,  $k_2 \geq k_1$

$$E_2 := x \rightarrow \left( \frac{(1+x)^2}{4} \cdot (1-x) \right) + ((x-x^2) \cdot (x)) \\ x \rightarrow \frac{1}{4} (1+x)^2 (1-x) + (x-x^2) x \quad (6)$$

$$\text{newtonsMethod}\left(E_2, \frac{1}{2}, 0.00000001, 10\right); \\ 0.5265986325 \quad (7)$$

$$\text{Maximize}(E_2(x), \{0 \leq x, x \leq 1\}); \\ [0.407092968632290719, [x = 0.526598632371106]] \quad (8)$$

$$\text{evalf}\left(\frac{1}{5} + \frac{2}{15} \sqrt{6}\right) \\ 0.5265986323 \quad (9)$$

3 Games, can change value of k after 1st and 2nd game,  $k_2 \geq k_1, k_3 \geq k_2$

$$E_3 := x \rightarrow \left( \left( \frac{\left(1 + \left(\frac{1}{3} \sqrt{3x^2 + 6x + 7} - \frac{1}{3}\right)\right)^2}{4} \cdot \left(1 - \left(\frac{1}{3} \sqrt{3x^2 + 6x + 7} - \frac{1}{3}\right)\right) \right) \right. \\ \left. + \left(\frac{(1+x)^2}{4} \left(\frac{1}{3} \sqrt{3x^2 + 6x + 7} - \frac{1}{3}\right)\right) \right) \cdot (1-x) + \left(\frac{1}{4} + \frac{1}{4}x + \frac{3}{4}x^2 - \frac{5}{4}x^3\right) \cdot x \\ x \rightarrow \left( \frac{1}{4} \left(\frac{2}{3} + \frac{1}{3} \sqrt{3x^2 + 6x + 7}\right)^2 \left(\frac{4}{3} - \frac{1}{3} \sqrt{3x^2 + 6x + 7}\right) + \frac{1}{4} (1 \quad (10) \\ + x)^2 \left(\frac{1}{3} \sqrt{3x^2 + 6x + 7} - \frac{1}{3}\right) \right) (1-x) + \left(\frac{1}{4} + \frac{1}{4}x + \frac{3}{4}x^2 - \frac{5}{4}x^3\right) x$$

$$\text{newtonsMethod}\left(E_3, \frac{1}{2}, 0.00000001, 10\right); \\ 0.5516013721 \quad (11)$$

$$\text{Maximize}(E_3(x), \{0 \leq x, x \leq 1\}); \\ [0.512766375904309069, [x = 0.551601371735228]] \quad (12)$$



**PHYSICS**  
**TEAM PROJECTS**



# Raman Spectroscopy: The Use of Inelastic Light Scattering to Compare Models in Classical Physics to Quantum Mechanics

Sammi Blocher and Elaine Gombos

## Abstract

Quantum and classical mechanics are established as distinct regimes in physics, but Raman spectroscopy may provide a pathway to compare vibrational modes of molecules on the quantum level to classical mechanical models of those vibrations. Raman spectroscopy is a non-invasive technique for identifying compounds through the inelastic scattering of light. Two experiments were conducted: one assessing the accuracy of a mass-spring model in illustrating the symmetric vibrations of regular benzene and deuterated benzene; the other involving three mathematical models of varying complexity in predicting the vibrational activity of carbon tetrachloride. However, none of the models proved to be accurate representations of quantum vibrational activity, supporting the general notion that there are assumptions in classical mechanics that cannot be applied to the vibration of molecules, which are quantum mechanical systems.

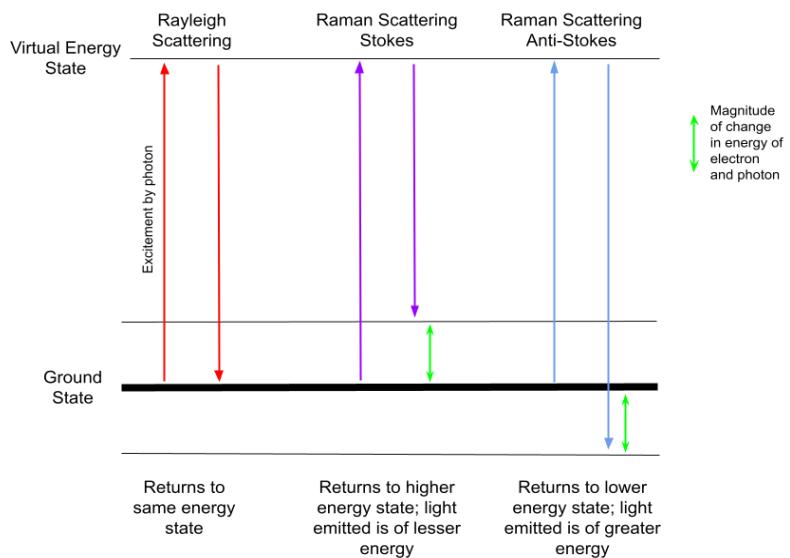
## I. Introduction

### A. Light Scattering

Light scattering refers to the re-emitting of light through the process of temporarily exciting electrons to a virtual energy state with photons. The photon is then released when the electron falls to a lower energy level<sup>[1]</sup>. Light scattering differs from phenomena such as emission and fluorescence because electrons are not excited to a discrete quantum mechanical energy state, and its position in the virtual energy state is very short-lived.<sup>[2]</sup> There are two ways for light to scatter: elastic (Rayleigh) and inelastic (Raman)<sup>[3]</sup>.

#### 1. Elastic Light Scattering

In elastic or Rayleigh scattering, all the energy from a light source that interacts with a molecule is emitted back into the environment and every excited electron returns to its original ground state (see Figure 1).



**Figure 1: Types of Light Scattering and Energy States**

This releases a photon at the frequency of the initial incident light wave, due to the polarization of electrons which induces a dipole. The dipole oscillates with the light, which makes it a moving charge, and in turn causes the molecule to radiate electromagnetic waves<sup>[4]</sup>. Rayleigh scattering occurs at high optical frequencies, around  $10^{15}$  hertz, so as to only affect movement in the electrons and not the nuclei. In classical electromagnetism, this effect is described through Equation 1:

$$\mu_{ind} = \alpha E \quad (1)$$

where  $\mu_{ind}$  is the induced dipole in the molecule,  $\alpha$  is the polarizability (which remains constant in Rayleigh scattering), and  $E$  is the external electric field from the light.

## 2. Inelastic Light Scattering

In inelastic or Raman scattering, lower optical frequencies are used to cause vibrations in the nuclei as well, which alters the electron distribution and the molecule's polarizability. When the electron returns to a lower energy state<sup>[5]</sup> and releases photons, it does not return to its original ground state, resulting in a net difference in energy as shown in the Raman scattering in Figure 1. In other words, the incident photon causes a change in the overall state of the molecule<sup>[6]</sup>. Thus, the light wave the electron emits during its descent also carries a different amount of energy than the incident light wave, resulting in a different frequency as well. It is this difference in the frequency of light waves that the Raman spectrometer detects and depicts<sup>[7]</sup>. Further, there are two types of Raman scattering: stokes and anti-stokes. Stokes occur when there is a net absorption of energy by the electrons, and the photons are emitted with a higher energy and frequency. Anti-stokes are where light particles leave at a lower energy state and the wavelengths are shorter<sup>[8]</sup>. The Raman spectrometer used will only examine the stokes, since they are more likely to occur than anti-stokes and thus return a stronger signal.

## B. Vibrational Modes and Bond Movements Within Molecules

Raman spectroscopy, which uses inelastically scattered light and induced dipole activity, essentially permits the detection of vibrational modes in molecules, which occur through intramolecular bond movements. Such vibrational modes are unique to each molecule, as they are dependent on atomic arrangement, the masses of the atoms, and chemical bond strength. This also returns a unique Raman spectrum that can be used for identification purposes.

In total, a linear molecule has  $3N-5$  and a non-linear molecule has  $3N-6$  normal vibrational modes, with  $N$  being the number of atoms in the molecule<sup>[9]</sup>. Vibrational modes result from bonds periodically bending or stretching, as well as out-of-plane deformations of a planar molecule. This experiment focuses on the bending and stretching of bonds, and it is important to note that bond bending requires less energy than stretching. Among all normal vibrational modes in a molecule, not all of them result in a significant change in polarizability and will not all be visible on a Raman spectrum, but each peak seen on a Raman spectrum (excluding phenomena such as harmonics mentioned in Section V.b) corresponds to a vibrational mode of the molecule. Ultimately, the subset of vibrational modes that do result in a peak are sufficient to produce a unique spectrum<sup>[10]</sup>.

Vibrational modes can be further classified into symmetric or asymmetric, which will be further discussed in Subsection E. Symmetric vibrational modes, such as the symmetric stretching of the molecule in Figure 4.a, allows the molecule to dilate and preserve its overall shape. Asymmetric vibrational modes do not preserve the shape of the molecule, as seen in Figure 4.b.

## C. Raman Spectrum: Shift and Intensity

A Raman spectrum displays the Raman shift ( $\text{cm}^{-1}$ ) on the horizontal axis, as well as the signal's intensity (count) on the vertical axis. The Raman shift is given by

$$\omega = \frac{1}{\lambda_{inc}} - \frac{1}{\lambda_{scat}} \quad (2)$$

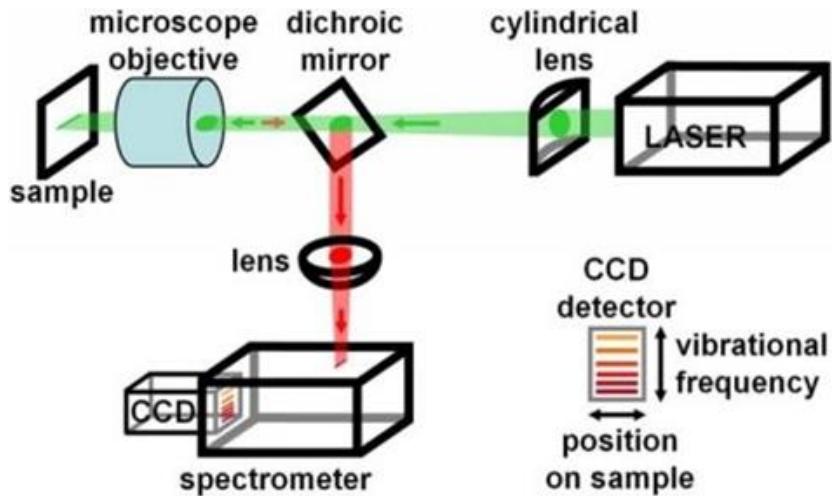
where  $\omega$  is the Raman shift in  $\text{cm}^{-1}$ ,  $\lambda_{inc}$  is the wavelength of the incident light waves released from the laser, and  $\lambda_{scat}$  is the wavelength of the light waves scattered by the sample detected by the photoreceptor, in cm. Intuitively, Raman shift is a wavenumber, which describes the number of cycles a wave is able to travel per centimeter. In other words, the Raman shift corresponds to the change in frequency or change in energy between the incident light and the light scattered by the sample. The larger the Raman shift, the more energy the molecule took from the incident photon to result in a vibration in the nuclei of the atoms in a molecule. Thus, the stretching of a bond would have a larger Raman shift than bending because stretching requires more energy.

The intensity of each peak on the spectrum is dependent on the concentration of the sample as well as the magnitude of the induced dipole, which is proportional to the polarizability of the molecule as seen in Equation 1. In short, the more photons that a molecule inelastically scatters, the taller the peak<sup>[11]</sup>.

## D. Characteristics of the Raman Spectrometer

A Raman Spectrometer measures Raman shift and intensity by passing polarized, intense light from a laser through a cylindrical lens. The light then goes through a dichroic mirror, into a microscope objective and into the sample. The sample reflects light back through the microscope objective. The elastic light goes through the mirror again, while the inelastic light bounces off perpendicularly. This happens because the dichroic mirror only allows light of a certain wavelength to pass through it. Because Raman scattering

occurs to approximately 1 in  $10^7$  photons and is a much rarer phenomenon than Rayleigh scattering, the mirror is needed to filter out the Rayleigh scattering, and the detector would only receive the Raman photons. The inelastic light goes through another lens and into the spectrometer where wavelength is measured<sup>[12-13]</sup> (see Figure 2). In this experiment, the Raman spectrometer by BaySpec was used, which has a laser releasing light of wavelength 532nm.

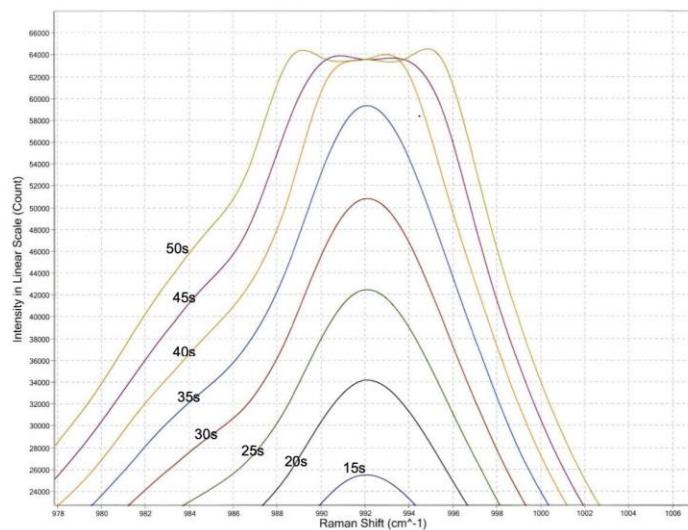


**Figure 2: Raman Spectrometer Schematic**

Several preliminary parameters set by the user dictate the intensity of the Raman peaks. In order to ensure that the Raman spectrometer will produce the maximal, clearest signals prior to the official experimentation, acquisition time and the light polarization angle were optimized.

### 1. Acquisition Time Optimization

Acquisition time refers to the amount of time the spectrometer is collecting data and receiving photons. If the acquisition time is too short, the signals would not be as clear and strong; whereas if the acquisition time is too long, the detector would reach its upper limit, or in other words, become saturated. The detector is then unable to detect additional scattered photons on that peak, resulting in an inaccurate spectrum. Trials were conducted with the benzene sample under acquisition times of 15 seconds to 50 seconds, at 5 second intervals (see Table 1, Appendix A for specific data). At first, the strength of the signal increased steadily with the increase in acquisition time. However, beginning with 40 seconds of acquisition time, the tip of the strongest peak began to plateau due to overexposure of the detector (see Figure 3). Thus, an acquisition time of 35 seconds was used in following experimentation, because that provided the strongest signal without saturating the detector.



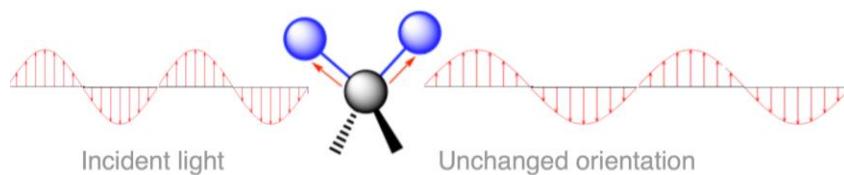
**Figure 3: Shape of Benzene Peak over Different Acquisition Times Experiment**

## 2. Light Polarization Angle Optimization

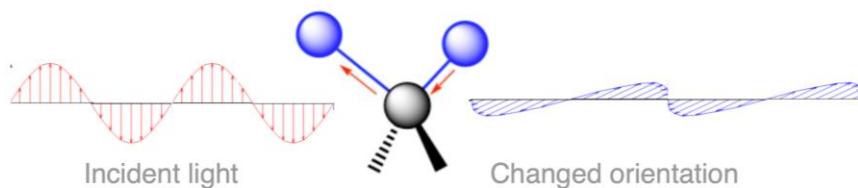
Polarized light refers to light waves that have their electromagnetic fields oriented in the same planes. Light polarization angle on the Raman spectrometer refers to the angle of light orientation that the detector receives. All the scattered light from the sample is filtered with a polarization analyzer that is located before the detector, which only accepts a certain orientation of light relative to the incident light waves. The polarization analyzer can be rotated in both directions, and spans angles of  $0^\circ$  to  $90^\circ$ , with  $0^\circ$  being light that is in the same orientation as the incident light released from the laser, and  $90^\circ$  being perpendicular to that orientation. To find the optimal angle that produces the strongest and clearest signals, light polarization angles of  $-30^\circ$ ,  $0^\circ$ ,  $30^\circ$ ,  $60^\circ$ , and  $90^\circ$  were tested with the same benzene sample as the acquisition time optimization trials. A light polarization angle of  $0^\circ$  yielded the spectrum with the strongest peak, so it was concluded that  $0^\circ$  was the optimal angle (see Table 2 and Figure 13, Appendix A for data supporting this). This indicates that most of the scattered light from the samples preserves the orientation of the incident light.

## E. Symmetrical and Asymmetrical Bond Vibrations in Raman Spectroscopy

In addition to altering the intensity of the signals, measurements at the  $90^\circ$  and  $0^\circ$  light polarization angles on the Raman spectrometer also serve to differentiate between symmetric and asymmetric bond vibrations (see Figure 4).



**Figure 4.a: Symmetrically Vibrating Molecule and its Orientation of Raman Scattered Light**



**Figure 4.b: Asymmetrically Vibrating Molecule and its Orientation of Raman Scattered Light**

Symmetrically vibrating molecules, as seen in Figure 4.a, tend to release light in the same orientation as the incident light. Thus, symmetric vibrations return a strong signal when the light polarization angle is set to  $0^\circ$  compared to when it is  $90^\circ$ . Asymmetrically vibrating molecules, on the other hand, radiate light in any configuration, thus releasing depolarized light that is oriented at all angles. Consequently, the difference in signal strengths between light polarization angles of  $0^\circ$  and  $90^\circ$  would be smaller compared to symmetric vibrations. Such a phenomenon is used in the following depolarization ratio:

$$\rho = \frac{I_{\perp}}{I_{\parallel}} \quad (3)$$

where  $\rho$  is the depolarization ratio,  $I_{\perp}$  is the intensity of a peak when the Raman spectrum is produced with a light polarization angle of  $90^\circ$ , and  $I_{\parallel}$  is the intensity of the same peak with  $0^\circ$  light polarization.

It is accepted that if  $\rho = \frac{3}{4}$ , then the vibrational mode is asymmetric; if  $\rho < \frac{3}{4}$ , then it is at least partially symmetric; if  $\rho \approx 0$ , then it is totally symmetric<sup>[14]</sup>.

## II. Goals

Differences between the quantum and macroscopic worlds have become an increasing area of study by physicists. Since Raman spectroscopy permits the observance of the quantum world, the technique can be used to answer the questions:

1. Can classical physics be used to model molecular vibrations and intramolecular bond movements on the quantum level?
2. How accurate are those models?

To accomplish this goal, two experiments were performed. The first experiment analyzed the ratio of vibrational energies between benzene and deuterated benzene, in comparison to a classical mass-spring system. The second experiment used the different vibrational modes of carbon tetrachloride to assess the accuracy of three different classical mathematical models.

## III. Methods

Both experiments involved similar procedures in operating the Raman spectrometer. First, a clear, empty vial was placed into the Raman spectrometer to establish a baseline. The sample was measured for 35 seconds. A new background was conducted every time a new angle or substance was used. When the background was complete, the sample was placed into the Raman spectrometer. The samples included pure benzene, deuterated benzene, and carbon tetrachloride. No additional preparation of the samples was needed.

### A. Experiment One: Benzene and Deuterated Benzene

For the comparison between regular benzene and deuterated benzene, the samples were first recorded once each so that the two spectra were superimposed on the same axis. This made comparing the two spectra easier. Additionally, one trial for each compound was run with a light polarization angle of 90° to further confirm which peaks correspond to symmetric vibrational modes. Each compound was then put through the Raman spectrometer six additional times with 0° light polarization to assess the reproducibility and precision of the spectra. This totals seven trials per compound. The Raman shifts of the two major peaks on each spectrum were located by zooming in on the peak and recording the horizontal coordinate of the highest point on that peak. For the Raman shifts of each peak, the average and standard deviation values were calculated across the seven trials. The average and standard deviation values were then used in calculations to compare the results to predictions from classical models.

### B. Experiment Two: Carbon Tetrachloride

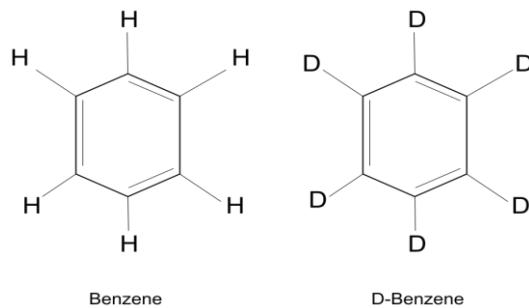
First, two superimposed scans of carbon tetrachloride were performed—one with 0° and another with 90° of light polarization. Five additional trials were then taken at 0° light polarization to assess reproducibility. The four peaks were analyzed and matched to their corresponding vibrational modes. The count and Raman shift values of each peak were located using the same method as in the first experiment. The average and standard deviation values across the six trials were calculated as well and used to compare the data to classical predictions.

## IV: Producing Classical Models of Molecular Vibrations

Classical models for intramolecular bond movements were based on the molecule's structure and the specific activities of each bond in each vibrational mode.

### A. Classical Modeling of Benzene and Deuterated Benzene

Benzene and deuterated benzene (abbreviated d-benzene) are very similar in structure, as seen in Figure 5. The only difference is that the hydrogens in benzene are replaced by deuterium in d-benzene. Deuterium is an isotope of hydrogen, with one more neutron and twice its mass<sup>[15]</sup>.



**Figure 5: Molecular Structures of Benzene and Deuterated Benzene**

Benzene structures have a total of thirty vibrational modes, two of which are totally symmetric. The first involves the periodic stretching and compression of the C=C bonds that form the benzene ring seen in

Figure 5. The second symmetric vibrational mode involves the stretching and compression of all the C-H bonds (in the case of benzene) or C-D bonds (in the case of deuterated benzene) surrounding the ring.

The stretching and compression of these bonds seem strikingly similar to the mass-spring system in classical mechanics. The two atoms involved in the bond stretch would be the masses, and the spring that connects them would represent the bond. A spring-mass system in classical mechanics is simple harmonic motion and abides by Hooke's law:

$$F = -kx \quad (4)$$

where  $F$  is the restoring force from the spring,  $x$  represents displacement of the oscillating mass, and  $k$  is the spring constant, which describes the stiffness of the spring. In the case of the model,  $k$  would represent the stiffness of the bond. Using Hooke's law, the following equation for the frequency of the oscillating spring-mass system can be derived:

$$f = \frac{1}{2\pi} \sqrt{\frac{k}{m}} \quad (5)$$

where  $f$  is the frequency of oscillation,  $m$  is the oscillating mass, and  $k$  is the spring constant. Because deuterium has twice the mass of hydrogen, Equation 5 can be used to find the ratio between the frequencies of vibrations in benzene and d-benzene that would be affected by that doubling in mass. This classical model yields a frequency ratio of  $\sqrt{2}$  from benzene to d-benzene (see Appendix B for full derivation). This model assumes that the effective spring constant of the carbon-hydrogen bond,  $k$  is not affected by the change in mass between hydrogen and deuterium.

## B. Classical Modeling of Carbon Tetrachloride

Carbon tetrachloride is a tetrahedral molecule with a total of 9 vibrational modes. However, some of these vibrational modes cannot be detected on Raman spectra, while others are degenerative, meaning that their vibrations have the same Raman shift and thus only produce one peak. Ultimately, only four types of vibrational modes are visible, which are shown in Figure 6. Out of the four, only the  $\nu_1$  vibrational mode is symmetric, as it has all four bonds stretching in a synchronized manner.

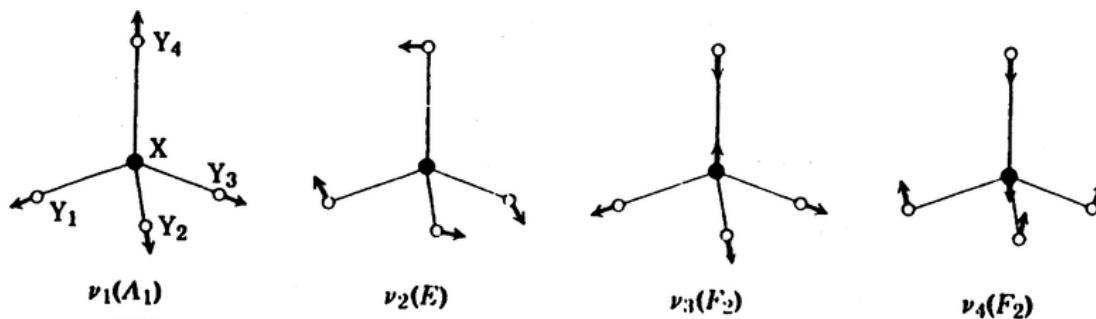


Figure 6: Bond Vibrations of Carbon Tetrachloride<sup>[16]</sup>

An analysis of three mathematical models based in classical physics, each with their own set of equations, was used to determine if classical physics can accurately model quantum activity. The three mathematical models used were the valence force model, the central force model, and the group theory model.

### 1. Valence Force Model

The valence force model in this experiment assumes that there are spring-like interactions between the carbon and chlorine, but no other forces exist. Another assumption is that the chlorines do not attract each other. The valence force model follows the set of equations in Figure 7.

$$Eq. 1 : 4\pi^2 c^2 \omega_1^2 = \frac{k}{m_{cl}}$$

$$Eq. 2 : 4\pi^2 c^2 \omega_2^2 = \frac{3k_\delta}{m_{cl}(l^2)}$$

$$Eq. 3 : 4\pi^2 c^2 (\omega_3^2 + \omega_4^2) = \frac{k}{m_{cl}} \left(1 + \frac{4m_{cl}}{3m_c}\right) + \frac{2k_\delta}{m_{cl}(l^2)} \left(1 + \frac{8m_{cl}}{3m_c}\right)$$

$$Eq. 4 : 16\pi^4 c^4 \omega_3^2 \omega_4^2 = \frac{2kk_\delta}{m_{cl}^2(l^2)} \left(1 + \frac{4m_{cl}}{m_c}\right)$$

**Figure 7: Equation Set for Valence Force Model**

In this set of equations,  $k$  represents the stiffness of the carbon-chlorine bond when stretching and is described in units of Newton per meter stretched and  $k_\delta$  represents the spring constant for when the bond is bent and is described in units of Newton per radian bent.

### 2. Central Force Model

The central force model is similar to the valence force model except it accounts for interactions between the chlorines, not just the bond between carbon and chlorine. It incorporates a variable to describe this chlorine-chlorine interaction. The model is demonstrated by the following set of equations in Figure 8.

$$Eq. 1 : 4\pi^2 c^2 \omega_1^2 = \frac{k_2}{m_{cl}} + 4 \frac{k_1}{m_{cl}}$$

$$Eq. 2 : 4\pi^2 c^2 \omega_2^2 = \frac{k_1}{m_{cl}} - \frac{k'}{m_{cl}}$$

$$Eq. 3 : 4\pi^2 c^2 (\omega_3^2 + \omega_4^2) = \frac{2k_1}{m_{cl}} + \frac{4m_{cl} + 3m_c}{3m_c m_{cl}} k_2 - \frac{2(3m_c + 16m_{cl})}{3m_c m_{cl}} k'$$

$$Eq. 4 : 16\pi^4 c^4 \omega_3^2 \omega_4^2 = \frac{2(4m_{cl} + m_c)}{3m_c m_{cl}^2} (k_1 k_2 - 8k_1 k' - 5k_2 k' - 8(k')^2)$$

$$\begin{aligned} Combined Eqs. 1, 2 & \& 3 : 4\pi^2 c^2 (\omega_3^2 + \omega_4^2) - \frac{4m_{cl} + 3m_c}{3m_c} (4\pi^2 c^2 \omega_1^2) - \frac{2(3m_c + 16m_{cl})}{3m_c} (4\pi^2 c^2 \omega_2^2) \\ & = \frac{-4(m_c + 4m_{cl})}{m_c m_{cl}} \end{aligned}$$

**Figure 8: Equation Set for Central Force Model**

In this equation set, the first, second, and third equations had to be combined and solved using systems of equations in order to find all variables.  $k' = \frac{-a}{t}$ , where  $\alpha$  is a constant of proportionality and  $t$  is the length of the tetrahedron edge. The parameter  $k_1$  represents the spring constant associated with the Chlorine-Chlorine interaction and  $k_2$  represents the spring constant associated with the Carbon-Chlorine interaction.

### 3. Group Theory Models

The group theory model builds on the valence force model, but also accounts for bond-bond interactions by adding a constant. The model is demonstrated by the following set of equations in Figure 9.

$$Eq. 1 : 4\pi^2 c^2 \omega_1^2 = \frac{k_r + 3k_{rr}}{m_{cl}}$$

$$Eq. 2 : 4\pi^2 c^2 \omega_1^2 = \frac{3k_\delta}{m_{cl}(l^2)}$$

$$Eq. 3 : 4\pi^2 c^2 (\omega_3^2 + \omega_4^2) = \frac{k_r - k_{rr}}{m_{cl}} \left(1 + \frac{4m_{cl}}{3m_c}\right) + \frac{2k_\delta}{m_{cl}(l^2)} \left(1 + \frac{8m_{cl}}{3m_c}\right)$$

$$Eq. 4 : 16\pi^4 c^2 \omega_3^2 \omega_4^2 = \frac{2(k_r - k_{rr})k_\delta}{m_{cl}^2(l^2)} \left(1 + \frac{4m_{cl}}{m_c}\right)$$

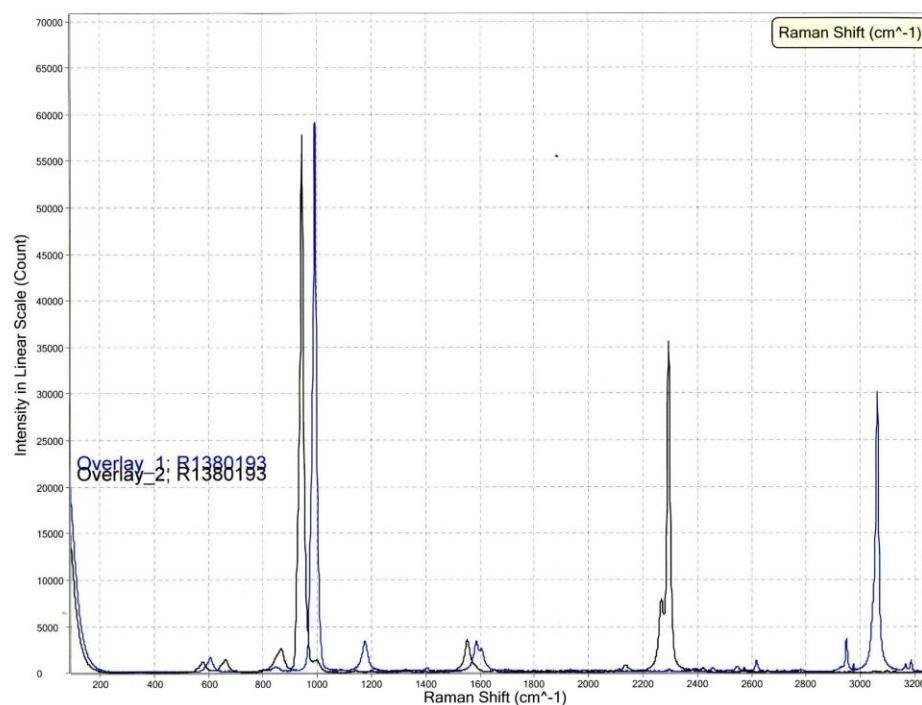
**Figure 9: Equation Set for Group Theory Model**

$k_{rr}$  represents the constant that was added to account for the interactions from the chlorine-chlorine bonds.  $k_r$  represents the stiffness of the bond between C-Cl.

## V. Results

### A. Experiment One: Benzene vs. D-Benzene

Even though benzene and d-benzene are similar molecules, they still produce unique spectra<sup>[17]</sup>. The classical model predicts that there will still be a significant difference in frequency — at a ratio of  $\sqrt{2}$ , to be exact. Since the Raman shift is a wavenumber, it is proportional to frequency. Thus, the accuracy of the classical model can be assessed by comparing it to the ratio between the Raman shifts of peaks in the spectra of benzene and d-benzene. The two spectra are superimposed in Figure 10.



**Figure 10: Comparison of Benzene and D-Benzene's Raman Shifts**

In Figure 10, the benzene spectrum is shown in blue, while d-benzene is shown in black. Each spectrum has two strong peaks, representing the two symmetrical modes of vibration. Stretching C=C bonds requires less energy than stretching C-H or C-D bonds, so the tallest peaks located on the left must correspond to the symmetric vibrational mode with C=C bond stretching. The black peak on the right would then correspond to the symmetric vibrational mode with C-D stretching, and the blue peak on the right with C-H stretching.

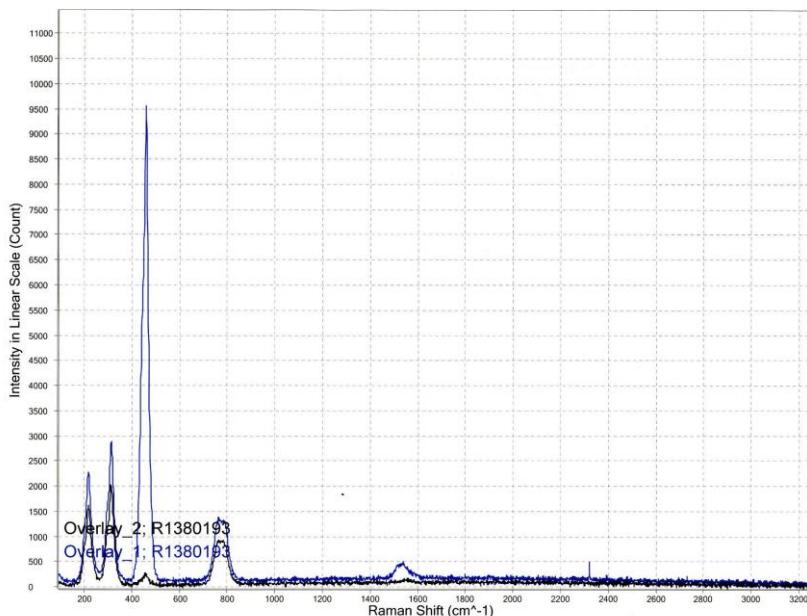
There is a very significant difference in Raman shift between the C-D and C-H stretch peaks, with the C-D bond stretch having an average Raman shift of  $3062.21\text{ cm}^{-1}$ , and the C-H bond stretch occurring at  $2293.96\text{ cm}^{-1}$ , on average. Meanwhile, there is a smaller but still significant shift between the C=C stretch peaks of the two compounds, with benzene's average peak at  $3062.21\text{ cm}^{-1}$  and d-benzene's at  $944.83\text{ cm}^{-1}$ . Although hydrogen and deuterium were not directly involved in the C=C bond stretches, the difference in mass still influenced the Raman shift. The Raman shifts over multiple trials also yielded very small standard deviations (see Table 3, Appendix A).

To compare classical physics to quantum mechanics, the average Raman shifts of the two peaks in each compound were multiplied together. The product of benzene was divided by the product of d-benzene to determine the ratio of the frequencies. The classical model predicted a ratio of  $\sqrt{2}$  or roughly 1.41421. The actual value found was 1.40250. Factoring in the standard deviation of the Raman shift ratio, the model had 40 degrees of error.

## B. Experiment Two: Carbon Tetrachloride Modeling

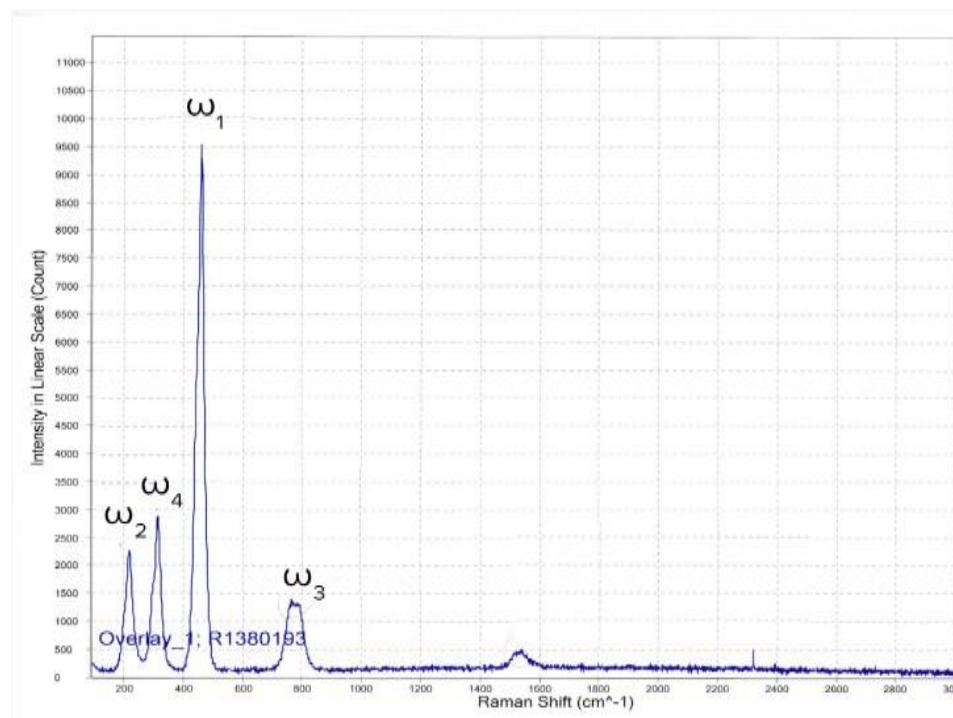
To begin assessing the accuracy of the classical models of the vibrational modes in carbon tetrachloride, the Raman spectra must first be analyzed and each of its four peaks matched to its corresponding vibrational mode illustrated in Figure 6.

Figure 11 shows the spectra for carbon tetrachloride at two different light polarization angles,  $0^\circ$  and  $90^\circ$ . The blue overlay indicates the spectrum taken with a  $0^\circ$  angle, and the black indicates a  $90^\circ$  angle. Thus, these overlaid spectra reveal information about whether each peak depicts symmetric or asymmetric modes of bond vibration. Note that there is a fifth peak at approximately  $1530\text{ cm}^{-1}$ . This is a harmonic and is located at twice the Raman shift of the fourth peak from the left. Thus, it is not a representation of distinct bond behavior and is not included in the calculations.



**Figure 11: Effect of Angle of  $\text{CCl}_4$  Peaks**

Starting from the left, the four peaks yielded depolarization ratios of 0.75, 0.73, 0.03, and 0.73. Therefore, of the four peaks, only the third peak from the left is totally symmetric and must correspond to  $\nu_1$ . In Figure 12, the peaks are labeled with  $\omega$  and a subscript that indicates its matching vibrational mode since  $\omega$  represents Raman shift. The remaining peaks were determined through the order of their energies. The peak farthest to the left is  $\omega_2$  because  $\nu_2$  only involves the bending of bonds. The next peak is mode  $\omega_4$  because  $\nu_4$  contains three bonds that bend and one that compresses. The compression increases the change in energy more than a bend would. Finally, the peak furthest to the right is labeled  $\omega_3$  because  $\nu_3$  involves the stretching or compression of all four bonds, thus taking more energy from the incident photon than  $\nu_2$  and  $\nu_4$ .



**Figure 12: Labeled Peaks of  $\text{CCl}_4$**

Note that the  $\omega_3$  peak does not have a distinctive tip and is actually made of two overlapping peaks due to resonance. The Raman shift of the  $\omega_3$  peak used in the calculations is an average of these two peaks.

Once again, multiple trials were conducted to identify the reproducibility of the experiment (see Table 4, Appendix B). The reproducibility trials were averaged. These means were used as  $\omega_1$ ,  $\omega_2$ ,  $\omega_3$ , and  $\omega_4$  in the force model equations.  $\omega_1$ ,  $\omega_2$ ,  $\omega_3$ , and  $\omega_4$  were averaged to be  $459.23 \text{ cm}^{-1}$ ,  $217.97 \text{ cm}^{-1}$ ,  $775.60 \text{ cm}^{-1}$ , and  $314.23 \text{ cm}^{-1}$ , respectively.

### 1. Valence Force Model

The stretching spring constant,  $k$ , was found to be  $44.05 \text{ N/m}$ . The bending spring constant,  $k_\delta$ , was found to be  $10.36 \text{ N/rad}$ . After finding all variables from the first and second equations, the numbers were plugged into the third and fourth equations. The left side of equation number three, which describes the experimental data, was equal to  $2.48473 \times 10^{24}$ , while the right side, which describes the classical model, was equal to  $4.6901 \times 10^{24}$ . The left side of equation number four was equal to  $7.4779 \times 10^{47}$  and the right side was equal to  $1.0769 \times 10^{48}$ . Factoring in the standard deviation of the left side of the equation, the third equation had around 750 degrees of error and the fourth equation had an error of around 65 degrees of error. Ultimately, since the other two models only used the fourth equation to assess the model's overall accuracy, this model was concluded to have 65 degrees of error (see Table 5, Appendix A).

### 2. Central Force Model

The set of equations used in this model are shown in Figure 8. A system of equations was made between the first three equations to find  $k_1$ . The value of  $k_1$  was determined to be  $7.8754 \text{ N/m}$ . The value of  $k_2$  was found to be  $12.55 \text{ N/m}$ . The value of  $k'$  was found to be  $-2.05 \text{ N/m}$ . The first three equations were equal on both sides. The left side of the fourth equation was  $7.4779 \times 10^{47}$  and the right side was  $7.9533 \times 10^{47}$ . The fourth equation returned around 9.5 degrees of error (see Table 5, Appendix A).

### 3. Group Theory Model

The set of equations used in this model are shown in Figure 9. The second equation can be solved for  $k_\delta$ . A system of equations using the first and third equation, and the value of  $k_\delta$  can find the values of  $k_r$  and  $k_{rr}$ .  $k_\delta$  was the same as it was in the valence force diagram at 10.36 N/rad. The values of  $k_r$  and  $k_{rr}$  were found to be 24.32 N/m and 6.5763 N/m, respectively. Once again, the fourth equation served as a check on the model, with the left side equaling  $7.4779 \times 10^{47}$  and the right side equaling  $4.3384 \times 10^{47}$ . This yielded 62 degrees of error for this model (see Table 5, Appendix A).

### 4. Overall Comparison of Models

Overall, the most effective model was the Central Force Model that had the least error between classical and quantum mechanics at 9.5 degrees of error. The next most accurate was the Group Theory Model with 62 degrees of error. Finally, the least accurate model was the Valence Force Model with an error of 65 degrees. See Appendix A for a table directing comparing the error of equation four (see Table 5, Appendix A).

## VI. Conclusion

In both experiments, classical physics failed to model the experimental data; according to the error, there was too great of a disparity between what was modeled and what was observed. In the first benzene and d-benzene experiment, the experimental data revealed a ratio of less than  $\sqrt{2}$ , the value predicted by the classical mass-spring model. The classical model only considered the change in mass and assumed that  $k$ , the stiffness of the bond, would remain constant between benzene and d-benzene, when in reality, deuterium could interact with the rest of the molecule differently than hydrogen does. This is also similar in the force models of  $CCl_4$ ; the models could not accurately predict the molecular vibrations and the interactions between atoms as a whole. Despite having models of varying complexity and that took many variables into account, none of the models could accurately describe the quantum vibrational modes. Thus, this supports the overall notion that assumptions made in classical mechanics cannot be applied to quantum mechanics.

## VII. Future Directions

There are many future applications to pursue with Raman spectroscopy. Firstly, Raman spectroscopy can be paired with IR spectroscopy to receive more comprehensive information. The two spectroscopy techniques are very similar in that they both use electromagnetic waves to look at bond movements and vibrational modes. The two often provide complementary information about molecules; Raman spectroscopy returns stronger signals for asymmetric vibrational modes, whereas IR spectroscopy favors asymmetrical vibrational modes. By pairing the two techniques, the majority of the bond vibrations can be studied to a greater degree. This can assist the identification of unknown compounds, in addition to comparing spectra to previously established databases. The purity of a sample can also be assessed with spectroscopy techniques. Furthermore, there are many variations of Raman spectroscopy that aim to enhance each signal, including resonance and surface-enhanced Raman spectroscopy. Resonance Raman spectroscopy uses incident light of a similar frequency to a particular electron transition, and needs less time and less amounts of the sample to function. Surface-enhanced Raman spectroscopy uses molecules adsorbed to rough metal surfaces or nanostructures and is so sensitive that it can detect single molecules. In conclusion, techniques such as Raman spectroscopy and its future applications can help physicists gain further insight into the quantum world, bettering our understanding of the world and universe.

## VII. Acknowledgements

We would like to thank everyone who made PGSS and our project experience possible including the PGSS Alumni Association, corporate sponsors, our State Legislators, Carnegie Mellon University, Mellon College of Science, and CMU Department of Physics. A special thank you to Dr. Barry Luokkala, Josh Kipiller, and Beka Modrekiladze for their guidance and support throughout the process. Another special thank you to Ms. Melissa Lessure for helping with the logistics of the program and our journals.

## VIII. References

1. *Background: Atoms and Light Energy*. (20202–04-20). National Aeronautics and Space Administration - Goddard Space Flight Center. Retrieved July 22, 2022, from [https://imagine.gsfc.nasa.gov/educators/lessons/xray\\_spectra/background-atoms.html#:~:text=When%20an%20electron%20temporarily%20occupies,a%20nearby%20atom%20or%20particle](https://imagine.gsfc.nasa.gov/educators/lessons/xray_spectra/background-atoms.html#:~:text=When%20an%20electron%20temporarily%20occupies,a%20nearby%20atom%20or%20particle).
2. Clarke, R. J., & Oprysa, A. (2004). Fluorescence and Light Scattering. *Journal of Chemical Education*, 81(5), 705–707. <https://doi.org/10.1021/ed081p705>
3. *What is Raman Spectroscopy?* (n.d.). Horiba Scientific. Retrieved June 30, 2022, from <https://www.horiba.com/ind/scientific/technologies/raman-imaging-and-spectroscopy/raman-spectroscopy/>
4. *Carbon tetrachloride*. (n.d.). California Office of Environmental Health Hazard Assessment. <https://oehha.ca.gov/chemicals/carbon-tetrachloride>
5. Britannica, T. Editors of Encyclopaedia (2011, December 27). Energy State. Encyclopaedia Britannica. <https://www.britannica.com/science/energy-state>
6. Graves, P. R. G. D. J., & Gardiner, D. (1989). Practical Raman Spectroscopy. Springer, 10, 978-3.
7. *Experiment 35: Raman spectroscopy: Vibrational spectrum of CCl<sub>4</sub>*. (n.d.). <http://chemwiki.cloud.ccas.gwu.edu/PhysicalChemistryLab/images/5/5f/CCL4Raman.pdf>
8. Edinburgh Instruments Ltd. (2021, November 25). *What is Raman Spectroscopy? | Raman Spectroscopy Principle*. Edinburgh Instruments. Retrieved July 25, 2022, from <https://www.edinst.com/blog/what-is-raman-spectroscopy/>
9. Chapter 5. IR spectroscopy and Raman scattering. (n.d.). In *Polymer analysis*. <https://www.eng.uc.edu/~beaucag/Classes/Analysis/Chapter5.pdf>
10. University of Michigan Advanced Physics Laboratory. (2006, February 15). *Raman spectroscopy*. [http://instructor.physics.lsa.umich.edu/adv-labs/Raman\\_Spectroscopy/Raman\\_spect.pdf](http://instructor.physics.lsa.umich.edu/adv-labs/Raman_Spectroscopy/Raman_spect.pdf)
11. Long, D. A. (2002). Introduction to Theoretical Treatments of Incoherent Light Scattering. *The Raman Effect*, 21. <https://doi.org/10.1002/0470845767>
12. Downes, Andrew & Elfick, Alistair. (2010). Raman Spectroscopy and Related Techniques in Biomedicine. Sensors (Basel, Switzerland). 10. 1871-89. 10.3390/s100301871.
13. Hammes, G. G. (2005). Vibrations in Macromolecules. *Spectroscopy for the Biological Sciences*, 92. <https://doi.org/10.1002/0471733555>
14. *Raman Spectroscopy*. (n.d.). Raman Spectroscopy. Retrieved July 24, 2022, from <https://gustavus.edu/chemistry/pchem/Raman%20Spectroscopy.html>
15. Garoff, S., Luokkala, B., & Schumacher, R. (2017, April 14). *Characterizing molecular vibrations using Raman spectroscopy*.
16. Chiriu, Daniele. (2007). Optical and structural characterization of crystalline oxides for laser applications.
17. Demming, A. (2019, October 2). *My favourite Nobel prize: Raman scattering is a universal fingerprint* – Physics World. Retrieved July 26, 2022, from <https://physicsworld.com/a/my-favourite-nobel-prize-raman-scattering-is-a-universal-fingerprint/>

## Appendix A: Raw Data

This appendix contains the raw data from the trials conducted during the experiment.

**Table 1: Effect of Time on Raman Shift**

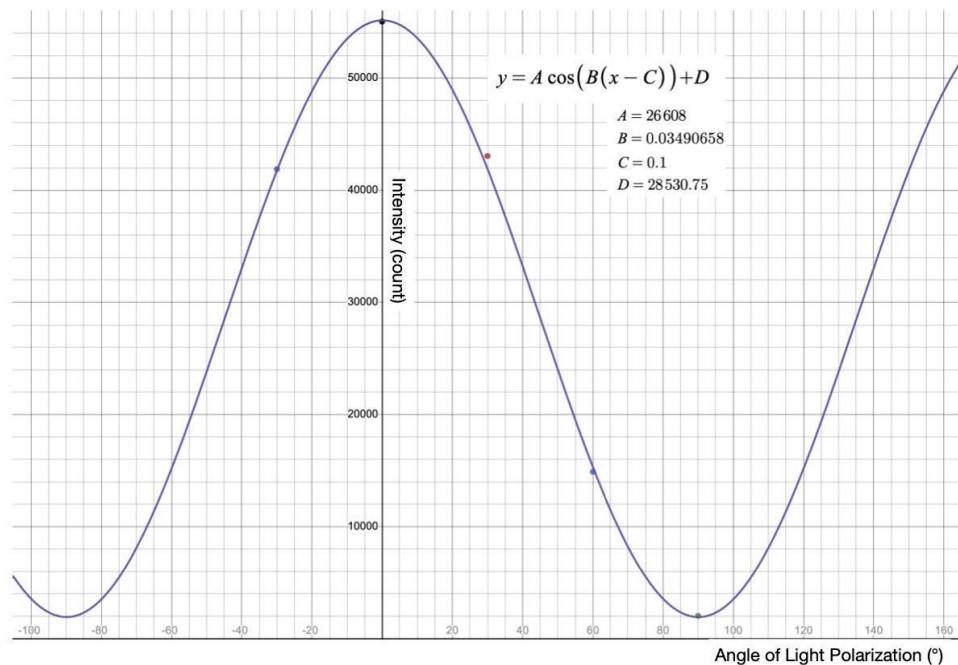
	Time: 15 sec.	20 sec.	25 sec.	30 sec.	35 sec.	40 sec.	45 sec.	50 sec.
Raman Shift* (cm <sup>-1</sup> )	992.04	992.12	992.11	992.11	992.10	992.98	990.87	994.85
Count	25506.5	34145.9	42473.5	50790.6	59348.8	63992.3	63883.4	64500.0

\*This Raman Shift represents the first peak in C<sub>6</sub>H<sub>6</sub> at an angle of 0°. See Figure X to see the visible effect of the oversaturation of the Raman Spectrometer. Figure X is taken from the Time reproducibility trials of benzene and zoomed into the first peak.

**Table 2: Effect of Light Polarization Angle on Raman Shift**

	Angle: -30°	Angle: 0°	Angle: 30°	Angle: 60°	Angle: 90°
Raman Shift* (cm <sup>-1</sup> )	992.315	992.3	992.3	992.24	992.0
Count	41873.1	55028.5	43046.4	14887.4	2033.0

\*This Raman Shift represents the first peak in C<sub>6</sub>H<sub>6</sub> at a time of 35 seconds.



**Figure 13: Sinusoidal Regression for Light Polarization Angle Optimization**

**Table 3: Comparing the Raman Shifts of Benzene and D-Benzene**

Trials	Benzene		Deuterated Benzene	
	Peak #1 (cm <sup>-1</sup> )	Peak #2 (cm <sup>-1</sup> )	Peak #1 (cm <sup>-1</sup> )	Peak #2 (cm <sup>-1</sup> )
1	992.83	3062.31	944.90	2294.09
2	992.78	3062.24	944.94	2293.98
3	992.76	3062.26	944.79	2293.94
4	992.85	3062.46	944.84	2293.92
5	992.63	3062.12	944.77	2293.91
6	992.58	3062.09	944.69	2293.88
7	992.34	3062.00	944.86	2294.02
<b>Mean (cm<sup>-1</sup>)</b>	992.68	3062.21	944.83	2293.96
<b>SD</b>	0.18	0.15	0.084	0.073

This table shows the Raman shifts of benzene and d-benzene over seven trials. Referring to Figure 10, Peak #1 of benzene is the blue peak on the Raman spectrum that is on the left and is higher. Peak #2 of

benzene is the other major blue peak on the right. Peak #2 of benzene compared to benzene's first peak is slightly lower. Peak #1 and #2 of d-benzene are the same corresponding peaks, but in black.

**Table 4: Carbon Tetrachloride Reproducibility**

Trials	Peak $\omega_2$	Peak $\omega_4$	Peak $\omega_1$	Peak $\omega_3$
1 ( $\text{cm}^{-1}$ )	219.54	314.20	459.52	775.65
2 ( $\text{cm}^{-1}$ )	217.23	316.17	459.28	775.86
3 ( $\text{cm}^{-1}$ )	217.58	313.29	459.08	775.96
4 ( $\text{cm}^{-1}$ )	219.24	313.34	459.31	775.035
5 ( $\text{cm}^{-1}$ )	216.96	314.36	459.13	775.545
6 ( $\text{cm}^{-1}$ )	217.25	314.04	459.06	775.53
SD ( $\text{cm}^{-1}$ )	1.12	1.05	0.18	0.32
Mean ( $\text{cm}^{-1}$ )	217.97	314.23	459.23	775.60

This table shows the reproducibility trials for carbon tetrachloride.

**Table 5: Carbon Tetrachloride Mathematical Model Calculations**

	Equation 4 Left:	Equation 4 Right:	Standard Deviation:	Error:
Valence Force Model:	$7.4779 \times 10^{47}$	$1.0769 \times 10^{48}$	$5.03 \times 10^{45}$	65 degrees
Central Force Model:	$7.4779 \times 10^{47}$	$7.9533 \times 10^{47}$	$5.03 \times 10^{45}$	9.5 degrees
Group Theory Model:	$7.4779 \times 10^{47}$	$4.3384 \times 10^{47}$	$5.03 \times 10^{45}$	62 degrees

Each system of equations was able to be solved with the first three equations. This allowed the forth equation to act as a “test” to determine the accuracy of the model. This “test” is what is listed above. Left refers to the results from the experimental data and the right refers to the predictions from the classical model.

## Appendix B: Classical Prediction of $\sqrt{2}$ for Benzene Experiment

This is the derivation of the ratio between the energies of benzene and d-benzene.

First, a ratio (R) was established between the frequencies of benzene ( $f_{s_b}$ ) and d-benzene ( $f_{s_d}$ ):

$$R = \frac{f_{s_b}}{f_{s_d}}$$

Frequencies of mass spring oscillators are equal to the following equation:

$$f_s = \frac{\sqrt{k}}{2\pi(\sqrt{m})}$$

Using that equation and assuming d-benzene has twice the mass, one gets:

$$R = \frac{\frac{\sqrt{k}}{2\pi(\sqrt{m})}}{\frac{\sqrt{k}}{2\pi(\sqrt{2m})}}$$

Simple algebra leads to the cancelation of  $\sqrt{k}$ ,  $2\pi$ , and  $\sqrt{m}$ , and one gets:

$$R = \frac{1}{\frac{1}{\sqrt{2}}} = \sqrt{2}$$

Therefore, the ratio between the energies of benzene and d-benzene should be  $\sqrt{2}$ .



# Determining Particle Diameter With Elastic Light Scattering and Lasers

Hannah Gong, Gabe Kenning, Sarah Meigs

## Abstract

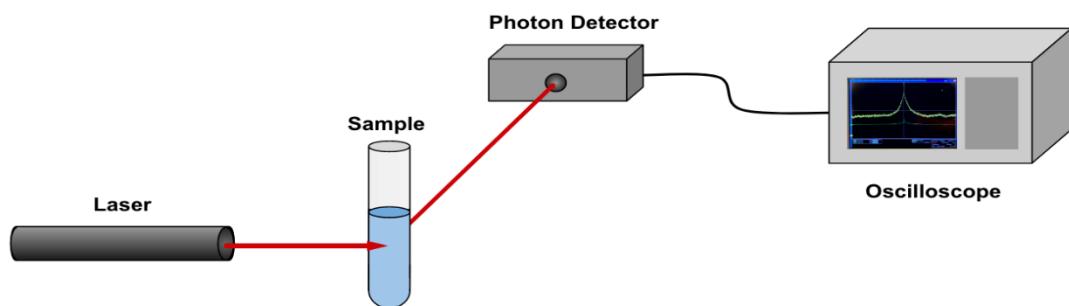
Quasi-elastic light scattering (QELS), otherwise known as dynamic light scattering, is a process by which particles of light are scattered by a sample of particles in a solution. This process can be used to determine the size of the particles in the sample by measuring the time constant of light intensity decay due to the diffusion of particles caused by Brownian motion. Using a laser, a photon detector, and a digital phosphor oscilloscope, this intensity decay can be recorded, processed, and analyzed. This paper investigates the accuracy of QELS in determining particle size and introduces a method of finding an optimal angle range, detailing procedures for processing data in a similar QELS set up. Using exponential decay equations and a derivation of the Stokes-Einstein equation, multiple trials run at different angles resulted in an average diameter for a sample of polystyrene latex spheres. While this average diameter was not within the manufacturer-given range, information about optimal angle measurement ranges and QELS data processing was discovered. The results of this paper can be applied to other research within biology, chemistry, and physics that require effective small particle size measurement.

## I. Introduction

### A. Quasi-Elastic Light Scattering

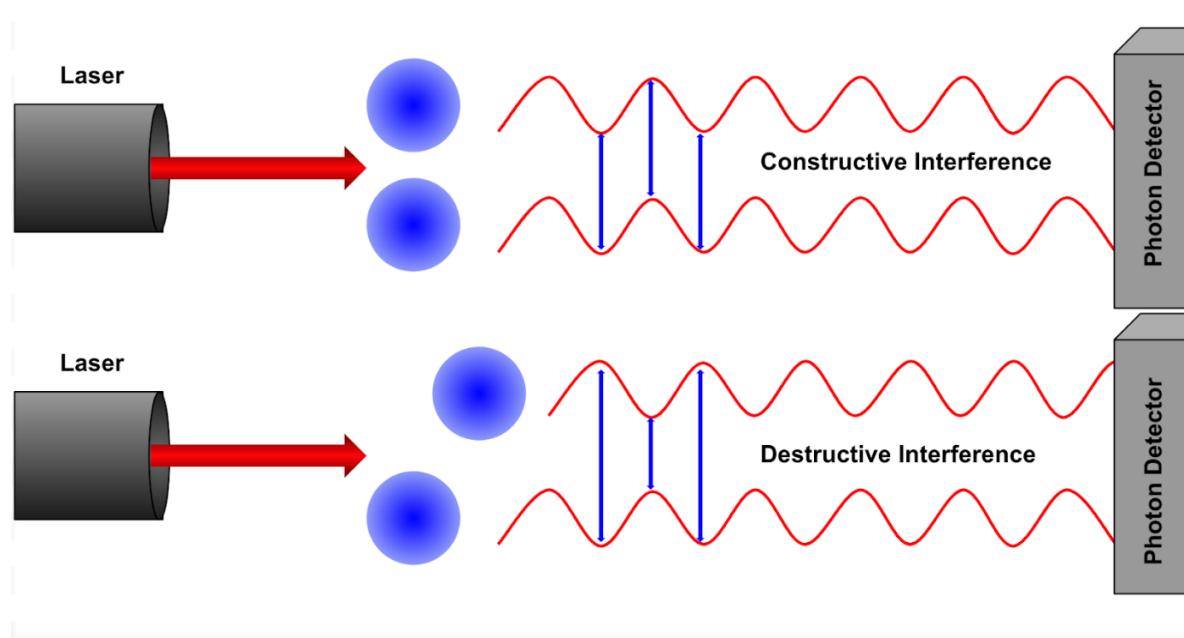
Quasi-elastic light scattering is a means of measuring the size of particles in an ideal fluid solution. It is an efficient way to determine particle size because it is noninvasive, has an expansive range of measurable sizes, and is relatively quick to perform.<sup>i</sup>

During QELS, a laser beam is shone into a sample of particles. The light waves of the laser beam collide with the particles and scatter at different angles. A detector is placed at a well-defined angle, relative to the incident beam. A photomultiplier within the detector then produces a voltage value proportional to the intensity of the scattered light.<sup>ii</sup> In Figure 1, the set up for laser light detection is shown.



**Figure 1: QELS Photon Detection Set Up**

After the coherent light beam of the laser hits the sample, both constructive and destructive interference occur. Collisions of light waves with particles cause the fluctuations in intensity that the photon detector records. The particles in the sample will scatter the light in varying paths, and will, in turn, disrupt light coherence. As seen in Figure 2, after colliding with sample particles the light waves exit the sample either in phase or out of phase with each other. When the waves reach the detector, they will cause varying measurements of intensity. If the photon detector detects waves with constructive interference, it will record a high intensity of light.



**Figure 2: Constructive and Destructive Interference**

In Figure 2 the effect of Brownian motion on QELS is shown. Brownian motion explains the random motion of particles in fluid resulting from collisions with solvent molecules. As particles undergo Brownian motion in fluid, they get further and further away from each other. This is due to the fact that when microscopic particles collide, they do not lose kinetic energy, they only change direction.<sup>iii</sup> As time passes and collisions continually occur, particles move randomly in new directions, exhibiting diffusion. Because these molecules are spreading apart, the average intensities per unit of time recorded during QELS decays. When the two particles move away from each other, the light path for the top particle changes, and therefore the wave phases are different. Thus, the change in scattered light intensity versus time depends on the rate of diffusion of the particles, due to Brownian motion. Furthermore, when measuring smaller particles, it should be noted that smaller particles will diffuse faster in liquid, meaning that the smaller the particle, the faster the rate of change in scattered light intensity.

## B. Stokes-Einstein Equation

### 1. Derivation of Time Constant Equation

Below is the equation for the diffusion coefficient, otherwise known as the Stokes-Einstein Equation.<sup>iv</sup>

$$D = k_b T * \frac{1}{C}$$

This essentially means that the molecules will diffuse as a function of the temperature of the fluid in which they are suspended over their frictional resistance. Boltzmann's constant is used because it directly relates the energy of one molecule as a function of the energy of an entire, ideal fluid system.

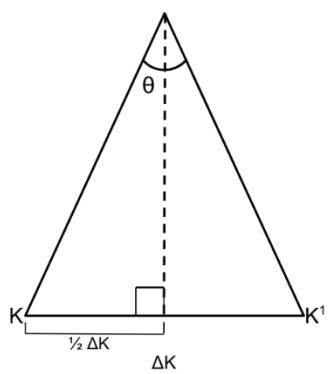
$$D = \frac{k_b T}{3\pi\eta d}$$

The previous equation breaks the diffusion coefficient down further by stating that frictional resistance is a function of the diameter of the molecule and the viscosity of the fluid in which the molecules are suspended. In essence, molecules that are larger and have more surface area will experience more friction because they have more area for the same amount of pressure. The equation also states that friction on a molecule will also increase if pressure on that molecule (that pressure being from other particles in the solution) is increased due to viscosity of the fluid being higher.

$$D = \frac{RT}{N} \cdot \frac{1}{C} \quad K_B = \frac{R}{N} \quad C = 3\pi\eta d$$

$$D = K_b T \cdot \frac{1}{3\pi d}$$

The Stokes-Einstein equation can be simplified by understanding that Boltzmann's Constant is equal to the Ideal Gas Constant divided by Avogadro's number (it represents the energy of a single molecule in relation to the energy of an ideally even system). We also know that C, the frictional resistance, is equal to the viscosity multiplied by the size of the molecule, so we can instead write the Diffusion Coefficient as the above equation.



**Figure 3: Deflection of Light's Momentum After Scattering**

In Figure 3, K and K' are two momentums of light before its collision with the particles and after its collision with those particles. It is possible to create a vector triangle in order to find how the change in momentum,  $\Delta K$ , relates to the angles. The angle formed to make a triangle with  $\Delta K$  as a side is equivalent to the angle at which the light is deflected. The diagram shows that if a right triangle is made by splitting  $\theta$  in half,  $\Delta K$  and be equated to K, so an equation using K can be used rather than  $\Delta K$ .

$$\tau_c = \frac{1}{2D\Delta k^2} \quad (\#1) \quad \#2) \quad nk = \frac{2\pi}{\lambda}, \quad \lambda = \frac{2\pi}{k}$$

$$(\#3) \quad \Delta k = 2k \cdot \sin\left(\frac{\theta}{2}\right)$$

Here, equation #3 is plugged into equation #1 and equation #2 is then plugged into equation #1.

$$\tau_c = \frac{1}{2D \left(2nk \cdot \sin\left(\frac{\theta}{2}\right)\right)^2}$$

$$\tau_c = \frac{1}{2D \left(2 \cdot \frac{2\pi}{\lambda} \cdot n \cdot \sin\left(\frac{\theta}{2}\right)\right)^2}$$

$$\tau_c = \frac{1}{2D \cdot 4n^2 \cdot \sin^2\left(\frac{\theta}{2}\right) \cdot \frac{4\pi^2}{\lambda^2}}$$

Using the equation that defines D, a final and useful form for finding the time constant comes forth.

$$\tau_c = \frac{1}{2\left(K_B T \cdot \frac{1}{3\pi\eta d}\right) \cdot 4n^2 \cdot \sin^2\left(\frac{\theta}{2}\right) \cdot \frac{4\pi^2}{\lambda^2}} \quad \tau_c = \frac{3\pi\lambda^2\eta d}{32\pi^2 n^2 \sin^2\left(\frac{\theta}{2}\right) k_B T}$$

$$\tau_c = \frac{3\lambda^2\eta d}{32\pi n^2 \sin^2\left(\frac{\theta}{2}\right) k_B T}$$

This is the final equation that values were plugged into.

## II. Methods

### A. Materials

This QELS spectroscopy experiment required a laser, a sample tube with particles in liquid, a photon detector, a photomultiplier, and an oscilloscope. A hard drive for collecting data and a computer for graphing and data analysis were also used. The laser had a wavelength of 632.8 nm. The sample consisted of monodispersed polystyrene microspheres in water. The samples were Polysciences Latex Spheres with a manufacturer given diameter of  $200.8 \pm 9.7$  nm. These samples were in a dual layered glass tube, which ensured that the light would not be diffracted by the glass.

### B. Data Collection

#### 1. Using the Apparatus

All the light scattering trials were performed in a dark room. A piece of cloth was placed over the apparatus, leaving only a slit for the laser. This was to ensure that there were no interfering lights that could cause noise, and to generally prevent as much background signal as possible. The detector was set at a certain angle in relation to the tube for each trial, ranging from  $15^\circ$  to  $90^\circ$ . The range did not go farther than these angles because light scattering becomes random at higher angles (something that was tested), and not enough scattering occurs at lower angles, which means that no meaningful results can be attained from a low angle data set. The apparatus allowed control of this angle through a manual but precise adjustment knob. Data was collected by a computer connected to the apparatus, in which a voltage threshold was set so that only intensities above a certain value would be detected and recorded. The reasoning behind this threshold was to limit the amount of detected light considered to be significant. The voltage threshold was largely dependent on the angle of the scattering due to the fact that the light at lower angles is inherently more intense since most of the light will be only slightly scattered. It was important to keep the threshold

low enough so that enough data could be collected to show a decay in intensity, but high enough so that the rate of acquisition of intensity values was at 1 acquisition per second, which would provide an accurate view of intensity decay during the trial. To strike this balance consistently, the threshold was adjusted for each trial. Each trial took around 20-30 minutes for the acquisition of 1,000 trigger voltages (voltages above the threshold).

## 2. Changing Angles

As mentioned in subsubsection 1 above, the angle of the detector to the tube could be changed. Furthermore, the angle of detection can be mathematically related to the hydrodynamic diameter. The adjusted equation is shown below.

$$d = \tau \cdot \frac{32\pi n^2 \sin^2\left(\frac{\theta}{2}\right) k_b T}{3\lambda^2 \eta}$$

Therefore, the angle of the detector was changed for each trial to determine an optimal range for particle size measurement for this certain sample size. While this range would change based on the size of the particle, the procedures used to determine the optimal range for the given Polysciences Latex Spheres can be applied to other samples.

## C. Calculations

### 1. Defining a New Constant

While equation 2 seems to have many variables, many of the values are constants. To streamline the calculation of diameter, these constants are combined to create one constant to be used in each calculation. Table 1 below shows all the constant values to be plugged into the equation. The constant 'diffraction index' refers to the ratio between the original wavelength of the laser light and the altered wavelength when the light is in the sample solution. The purpose of Boltzmann Constant in this equation is essentially to settle units.

**Table 1: Constants Used**

Wavelength of Laser ( $\lambda$ )	632.8 nm
Viscosity of Solution ( $\eta$ )	1.0016 mPa·s
Diffraction Index (n)	1.333
Boltzmann Constant (K <sub>B</sub> )	1.35049x10-23 $\frac{m^2 kg}{s^2 k}$

Then, the above values were plugged into equation 2,

$$\frac{32\pi n^2 k_B}{3\lambda^2 \eta} = \frac{32\pi(1.333)^2(1.35049 \times 10^{-23} \text{ m}^2 \text{kg s}^{-2} \text{ K}^{-1})}{3(632.8 \text{ nm})^2(1.0016 \text{ mPa} \cdot \text{s})}$$

$$\approx 2.051 \cdot 10^{-6} \frac{J}{\text{mPa} \cdot \text{s} \cdot \text{m}^2 \text{K}}$$

Therefore, equation 3 can be rewritten as:

$$d = 2.051 \times 10^{-6} \cdot \left( \tau \cdot \sin^2\left(\frac{\theta}{2}\right) \cdot T \right)$$

## 2. Finding the Time Constant Graphically

The final graph of an intensity curve after a QELS spectroscopy trial followed a simple exponential decay equation due to Brownian motion and diffusion,

$$I(t) = I_{max} e^{-\frac{t}{\tau}}$$

Therefore, it could be assumed that when one time constant was reached,

$$I(\tau) = \frac{I_{max}}{e}.$$

This means that if the maximum intensity was found in each trial, then simply dividing that intensity by e would find the intensity value after one time constant. Then, the actual value for the time constant could be found graphically.

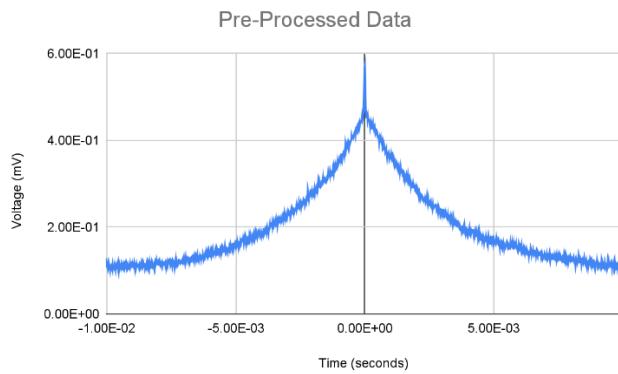
Because the photomultiplier within the apparatus converts light intensity to a voltage, it can be assumed that the voltage pattern on the oscilloscope's would follow the same pattern as the intensity's decay, or that the equations to describe both are equivalent.

## 3. Hydrodynamic Diameter

One thing to note is that equation 4 does not actually find the size of the particles. It merely measures the diameter of a perfect solid sphere that would diffuse at the same rate as the particles in fluid. This measure is known as the hydrodynamic diameter.

#### 4. Data Processing Methods

The raw data were transferred from the digital oscilloscope into Google Sheets for analysis. An example data set is shown in Figure 4. Because a large number of numerical values were transferred from the oscilloscope to the spreadsheet in each trial (approximately 1000), and the graph was symmetric about the y-axis (see graph 1 below), any voltages with negative time values were removed as all values were reflected in positive time as well. Additionally, the long plateau at the end of the decay line was cut because it was not needed. Plugging in the time constant, set angle, and recorded temperature of the sample into equation 4 calculates the diameter of the particle. The time constant and diameters were calculated for angles between 15 degrees to 90 degrees at 15-degree increments. 35 degrees was also tested, as will be explained later in Section 4A.



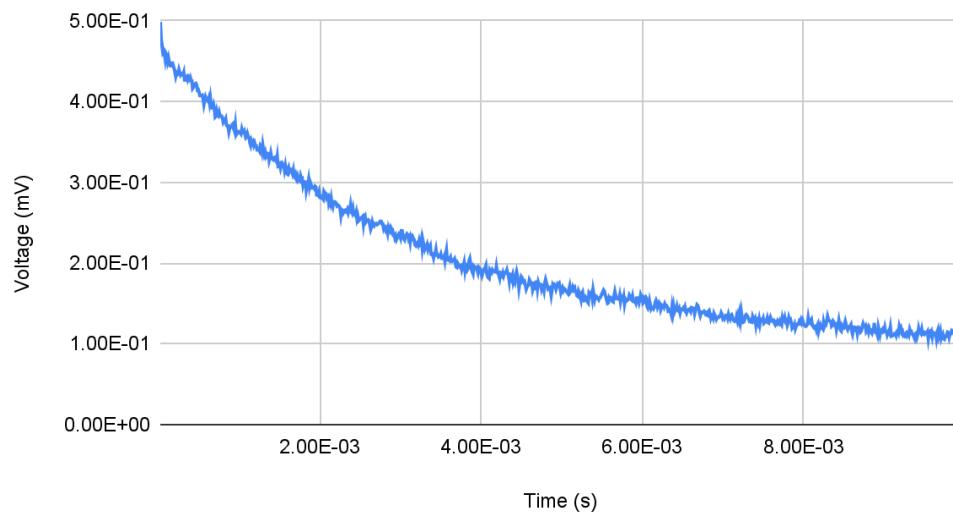
**Figure 4: Raw Oscilloscope Data for 35 Degrees**

### III. Results

#### A. Determination of Exponential Time Constants

Figure 5 is an example plot of the processed data for a trial at 35 degrees. The curve starts off at a maximum intensity and follows an exponential decay until it plateaus. An exponential fit is performed to find the time constant, which is proportional to the particle diameter.

### 35 Degrees: Voltage vs Time



**Figure 5: Processed Light Scattering Data taken at 35 Degree**

## B. Final Diameter Values

After conducting the experiments, data values were collected, processed, and calculated. The final diameters found were found, organized, and listed. For the purposes of this experiment, 15, 75, and 90 degrees were not included in the table because they were not sound predictions.

Table 2 shows the exponential time constants from fits to light scattering data obtained at various angles of scattering and the resulting calculated particle diameters.

**Table 2: Exponential Time Constants and Calculated Particle Diameters**

$\theta$	$I_{max}$ (mV)	$I_{max}/e$ (ms)	$T_c$ (ms)	Diameter (nm)
30°	0.607	0.223	4.76	194.80
35°	0.571	0.210	3.79	207.95
45°	0.395	0.145	2.56	227.61
60°	0.264	0.0971	1.63	247.39

## IV. Discussion

After reviewing the results, it can be concluded that only a range of 30 to 60 degrees had consistent results. This could be because of the sensor picking up too much noise from non-scattered light sources.

30 and 35 degrees seemed to be the best angle for scattering light, and the overall average of the diameters ended up being 219 nm. It is believed that this value could get closer to the true diameter if more data points were collected.

## A. Analysis of Results

### 1. Determining the Optimal Range

During experimentation, it was found that angles outside of the range of 30 to 60 degrees had inconsistent results. While both 15 degrees and 90 degrees were tested, after calculating diameter values for those two trials it was found that the values were very far off the given manufacturer diameter. Furthermore, during data collection in these trials, the thresholds for light had to be very high or very low, respectively, in order to get 1,000 triggers. The reasons for this are explained in Section 4B.

However, ignoring the outlier angles, within the range of 30 and 60 degrees, 35 degrees was tested to see if it would result in a value in the manufacturer's range. This did occur, proving that angles within the range are effective for measurement. Between the four angles in the optimal range, an average diameter of 219.44 nm was calculated. This falls above the manufacturer provided  $200.8 \pm 9.7$  nm range. However, both 30 and 35 degrees provided measurements in the range of error.

### 2. Statistical Testing

The standard deviation was calculated for the angle diameters using the equation below.

$$\sigma = \frac{\sqrt{\sum(x_i - \mu)^2}}{N}$$

A standard deviation of 19 was calculated for the sample size. The standard error of mean was also calculated to be 9.96. That means for a 68.3% confidence interval, there was a margin of error of  $219.4375 \pm 9.96$ . For a 95% confidence interval, there was a margin of error of  $219.4375 \pm 19$ .

## B. Experimental Errors

Due to the nature of QELS, if the sample is contaminated with other particles such as dust before experimentation, the diameter measurement could be much less accurate as dust particles would affect the average diffusion rate of particles. Furthermore, outside of the optimal range of angles, there are errors for diameter measurement.

However, at lower angles, a multiple scattering effect can be seen (especially in samples of higher concentrations). During multiple scattering, photons can deflect off neighboring particles and make it seem that particles are moving faster, hitting the detector more often as well. Since smaller particles are faster, this will make size measurements at lower angles smaller than the actual diameter.

At higher angles, the detector is less likely to be hit as well, due to the fact that it is less likely for a particle to deflect light at a large angle. While large angle deflections are possible, they occur less, meaning that less light reaches the detector and therefore, the particles seem larger.

Finally, it is possible that light from non-laser sources (such as electronics) could have been detected by the photodetector during trials, affecting the intensity measurements.

## V. Conclusions

### A. Future Directions

It is clear from the statistical testing that the data was varied. While the average diameter found was close to the manufacturer range, more trials in the optimal range of angles would be beneficial to lower the standard deviation and determine a more accurate diameter. QELS only works within a specific range of angles, though, as explained in Section 4B. Therefore, it is important for future experiments to set the angle of the photomultiplier to at least 15 degrees and at most 75 degrees.

Testing different viscosities, molecular diameters, and temperatures would be interesting routes to pursue in further confirming the equation that defines the time constant. Does the equation hold up at its extremes, or is it more complex and in need of further addition? Looking forward, there are many opportunities to test the equation and its relationship with its variables and light.

### B. Applications

Some methods are more effective than DLS in some cases. For example, Atomic Force Microscopy is more effective at smaller size ranges (20-100 nm), whereas DLS is more useful for size ranges that are slightly larger.<sup>v</sup> DLS tends to be more effective than Transmission Electron Microscopy, though, when analyzing microscopic particles.<sup>vi</sup>

Therefore, dynamic light scattering can be used for analysis of small particles, especially those that are too small for physical methods to find. This process can also be used to see if there is a relatively even distribution of a molecule within a solution because the process only works when the solution is “ideal” and therefore works with Boltzmann’s constant.

QELS can also be used to measure the viscosity of ideal fluids in limited situations, especially if there is not enough of the fluid to measure the viscosity with other methods. Also, it is possible to measure the composition and energy of any fluid, granted that the detector is strong and accurate enough to establish a true relationship between time and voltage.

QELS can be used to confirm chemical compositions. For example, QELS has been used to determine the sizes of proteins and other macromolecules fairly easily and efficiently. Beyond this, QELS has been used to find the size of microbial bacteria.<sup>vii</sup> Overall, the applications of QELS in industry and research across multiple fields are immense.

## References

E. J. Davis et al., *The Airborne Microparticle* (Springer-Verlag Berlin Heidelberg, New York, 2002)

Malm, A.V., Corbett, J.C.W. "Improved Dynamic Light Scattering using an adaptive and statistically driven time resolved treatment of correlation data." *Sci Rep* **9**, 13519 (2019). <https://doi.org/10.1038/s41598-019-50077-4>

Malvern Panalytical. (2018, May 18). *A basic introduction to dynamic light scattering (DLS) for particle size analysis* [Video]. Youtube. A basic introduction to Dynamic Light Scattering (DLS) for particle size analysis

Neher, R. (2018, September 27). Brownian motion, diffusion, and Fick's law. Neher Lab. [https://neherlab.org/20180927\\_theoretical\\_biophysics.html](https://neherlab.org/20180927_theoretical_biophysics.html)

---

<sup>i</sup> George D. J. Phillies, "Quasielastic Light Scattering", *Analytical Chemistry* **62**(20), 1049A-1057A (1990).

<sup>ii</sup> David W. Ball, *Field Guide to Spectroscopy*, (SPIE Press, 2006), 97

<sup>iii</sup> Heather MacNeill Falconer, Gina Battaglia, Anthony Carpi, "Diffusion I" Visionlearning Vol. CHE-3 (4), 2015.

<sup>iv</sup> Christina Cruickshank Miller, "The Stokes-Einstein law for diffusion in solution", Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character, **106**(740), 724–749, 1924. <https://doi.org/10.1098/rspa.1924.0100>

<sup>v</sup> Hoo, C.M., Starostin, N., West, P. et al. A comparison of atomic force microscopy (AFM) and dynamic light scattering (DLS) methods to characterize nanoparticle size distributions. *J Nanopart Res* **10**, 89–96 (2008). <https://doi.org/10.1007/s11051-008-9435-7>

<sup>vi</sup> T. G. F. Souza et al 2016 *J. Phys.: Conf. Ser.* 733 012039

<sup>vii</sup> Stetefeld, J., McKenna, S. A., & Patel, T. R. (2016). "Dynamic light scattering: a practical guide and applications in biomedical sciences." *Biophysical reviews*, **8**(4), 409–427. <https://doi.org/10.1007/s12551-016-0218-6>

# Optical Tweezers

Jason Weller, Jan Mejía-Toro, and Edison DeWilde

## Abstract

"Optical tweezers" are devices that focus a laser down to capture and hold a small particle. The laser is focused so that any free-floating object that impacts the light is pushed towards the laser's focal point. When the object reaches the focal point, the force of the laser comes from all sides, trapping the bead in place. At the beginning of the experiment, the Optical Tweezers system was not functioning properly. After identifying sources of probable error and introducing measures to counteract them, the system was able to successfully trap a free-floating silica bead. Multiple calibration tests were performed on the system to both ascertain the extent to which it was functional and to determine the strength of the optical trap. The trap performed correctly, with the spring constant increasing as the laser power increased. Further testing with organic materials revealed the constraints of an Optical Tweezers system. During testing with an orange sample, excess moisture inside the sample blocked the laser from being able to access the cells, revealing the tweezers' inability to work on inherently wet samples. A second major limitation was discovered while testing the leaf sample, where the laser encountered an unstable object, which promptly exploded from the added heat of the laser, demonstrating the difficulty of experimenting on heat-sensitive samples.

## I. Hardware Design

An optical trap is a method utilizing a highly focused infrared laser and a microscope to manipulate and observe silica microspheres or beads. The term Optical Tweezers refers to the physical device and components that are used to create an optical trap.

More specifically, the laser traps beads of silica by altering the momentum of the beads in a way that traps them. This is possible because of the way the laser is focused; rather than one cylindrical beam, the laser is spread out in a conical beam that causes the beads to be drawn to the tip of the cone. The beads do this because, as the laser passes through the molecules, the beam is scattered in a direction that causes the bead to experience a force in the direction of the beam propagation (the tip of the conic beam). The beads stay in the center of the beam of the laser because they experience these forces in all directions, effectively trapping them in a position of stasis (Figure 1).<sup>i</sup>

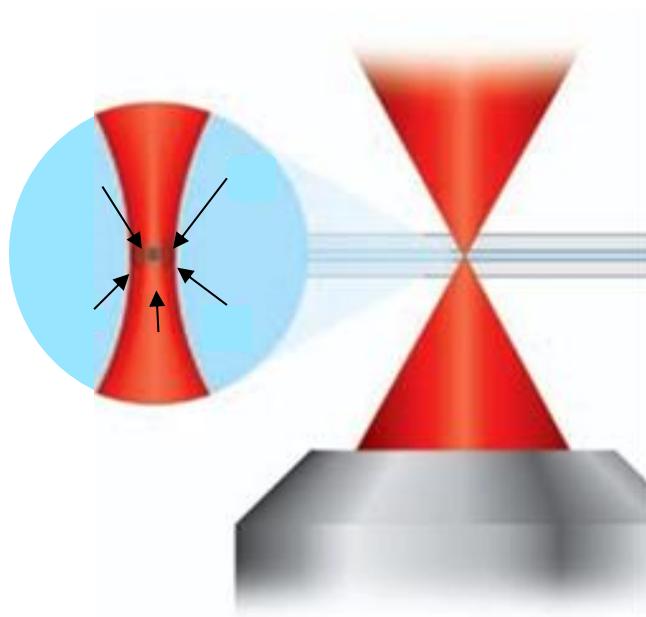
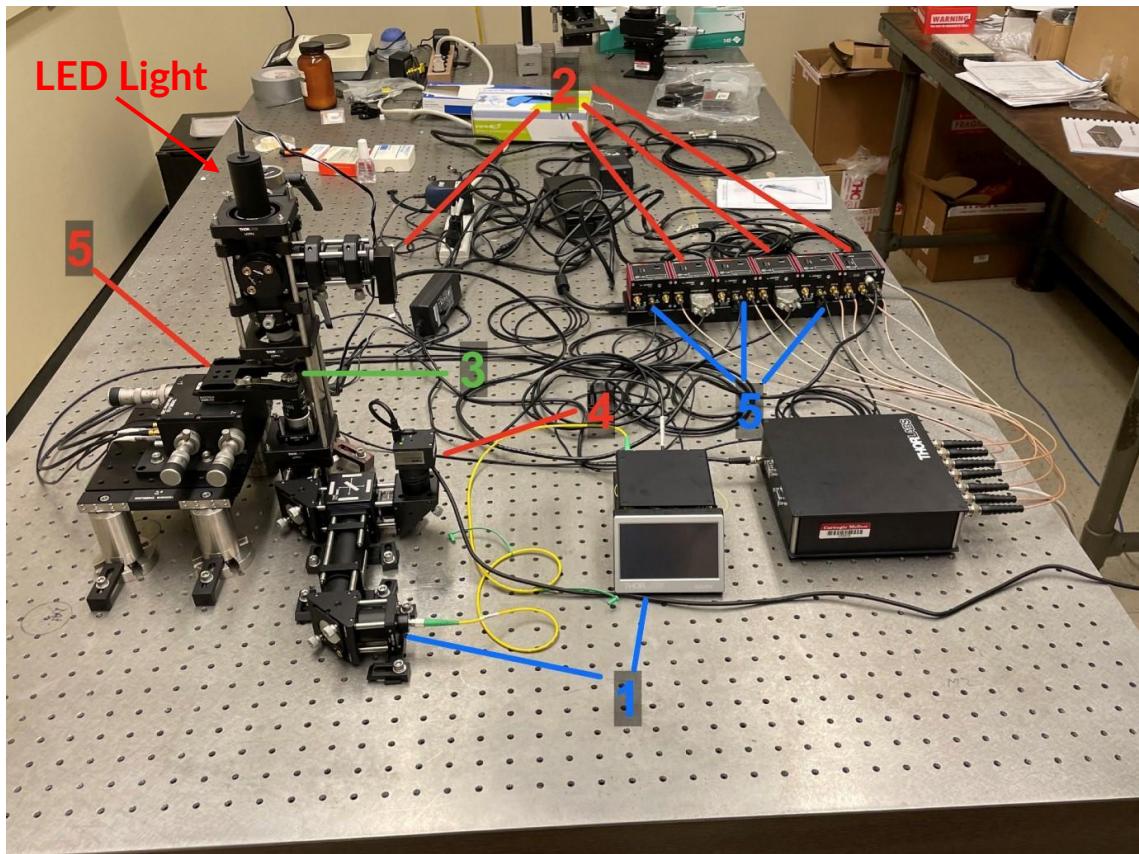


Figure 1: Diagram of a Trapped Bead at the Center of Propagation of the Laser<sup>ii</sup>

### A. Physical Components

Both the hardware and software components of the optical tweezer system were sourced from Thorlabs. This allowed for consistent connections between the components and ensured each component was of high quality.



**Figure 2: Diagram of Numerically Labeled Optical Tweezers Setup Corresponding to Subsections**

### 1. Infrared Laser

The laser that was used in the Optical Tweezers setup was a 976 nm infrared laser. The input was a single diode hermetically sealed in a 14-pin butterfly package. The package provides constant power to the laser, keeping the overall power output stable. The output of the laser was then connected to a collimator, and then a fiber optic cable to the microscope.<sup>iii</sup>

### 2. Quadrant Position Detector (QPD) and QPD Sensors

The Quadrant Position Detector (QPD) is a sensor that is able to detect the force, stiffness, and position of the Optical Tweezers. The force sensor is utilized in the process of optical trapping to measure the forces being applied to the silica beads. The QPD has to be placed in a plane that is connected to the back of the focal plane of the condenser to function properly. The QPD Sensors allow the data collected by the QPD to be directly sent to the computer and interpreted by the computer software.

### 3. 100x Oil Immersion Objective Lens Microscope with Condenser

The microscope used to magnify the sample for viewing from the camera as well as focus the beam of the laser down to a spot size of 1.1  $\mu\text{m}$  was a 100x oil immersion Nikon objective lens microscope. Also, the focusing of the laser by the lens was important because it was what caused the laser to point in the conical

shape that is crucial for optical trapping. This is because, without the conical shape, the beads have the possibility of being pushed out of the trap.

The condenser on the side opposing the objective lens was a 10x air condenser from Nikon with a numerical aperture of 0.25. The condenser is important because it collimates the light after it passes through the sample, which allows the QPD to relay an accurate force reading.

#### **4. 1440 x 1080 Pixel Color Camera and LED Light**

The camera used to view the sample while under the microscope and laser was a 1440 x 1080 pixel color CMOS camera. There is a dichroic mirror in the light path as well as a short-pass filter that prevents any backscattered light from the laser from entering the camera sensor. The LED light on the opposite side of the sample (Figure 2) is there to illuminate the sample, so it is possible to have a clear image with the camera.

#### **5. Sample Positioning Stage**

The sample positioning stage is the microscope slide holder mounted onto a stage capable of X,Y, and Z translations. The stage is controlled on these by 3 piezoelectric motors connected to the computer. These motors are connected to the breadboard, which is connected to three compact Kinesis K-Cube controllers that are then connected to the computer, which allows the motors to be controlled by software from the computer.

### **B. Identification of the Problem**

Our set up had some issues that needed to be addressed before the apparatus could be calibrated and used. The notes authored by Michael Murphy, a sophomore physics major at Carnegie-Mellon University, and the published report from MIT about their study with Optical Tweezers were especially helpful in identifying any issues.<sup>iv,v</sup>

#### **1. Cover Slip**

The first problem that was identified was that it seemed from the notes of the previous students working on the “tweezers” that they had been putting the slide with the beads in the incorrect orientation. To elaborate, the cover slip that is applied on one side of the sample slide to cover the sample was facing upwards in the previous students’ tests, but the coverslip had to be facing down for the laser to work correctly (Figure 4). This error was most likely the issue that was the cause of the device not working properly.

#### **2. Level of Focus of the Laser**

Another variable that was identified as a possible problem affecting the functionality of the laser was that the ability of the laser to perform a trap might be affected by the focusing of the microscope. This was because to focus the microscope, the z-axis of the microscope would have to be adjusted, which could move the sample out of the location most optimal for the laser to trap a bead. Due to the way the laser is focused, there is a point where there is maximum focusing of the laser, and any point with a greater or lesser z-value will experience a decreased focusing of the laser, which could lead to a less effective trapping result in tests.

### 3. Physical Setup Error

Also identified was an error in the physical setup of the laser that could have had an effect on the functionality of the optical trapping (Figure 3). This error was that the beam expander and subsequent parts of the system that were connected to the laser were slightly off axis. This misalignment meant that instead of allowing the laser to cleanly pass through the beam expander and mirrors, the beam is reflected slightly off-center, which could have caused problems with the focusing of the laser and the ability to generate accurate measurements.

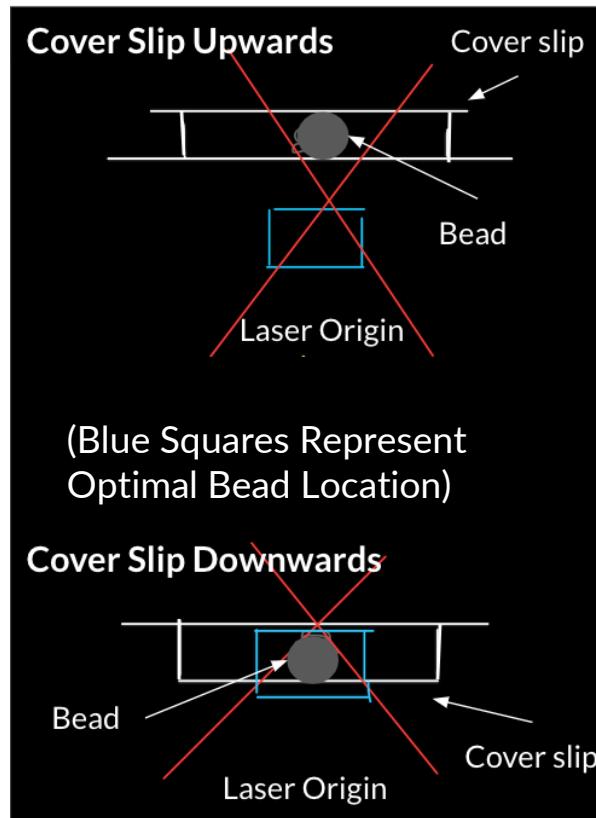


**Figure 3: Image of Physical Misalignment in Laser Setup**

Additionally, when the system was turned on to test and determine possible issues, it was clear that the location that the laser was affecting on the slide was not the point at the center of where the camera was focused, but rather on the edge of its field of view. This error in the physical setup of the laser could have an effect on the results of the tests, but it still seemed functional, so other errors were tested first.

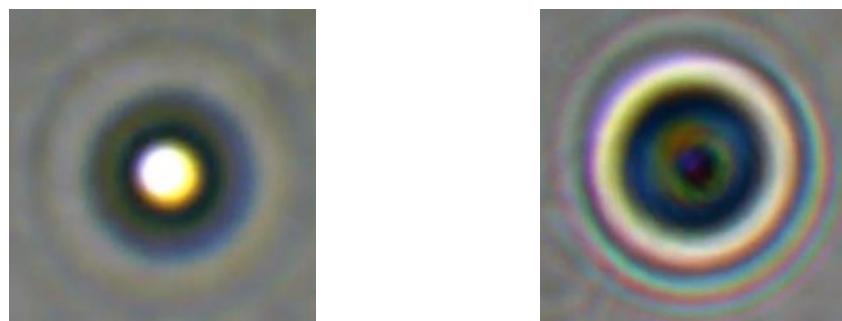
## C. Prospective Solutions and Conclusions

The first problem worked on was flipping the sample slide into the orientation where the coverslip was on the bottom, because it was the simplest issue to test, and it would determine if there were any other issues or software related malfunctions. Additionally, the reason the coverslip had to be on the bottom of the slide was because the point of propagation of the laser beam would be at too low of a z-value if the coverslip was on the top, but when it was on the bottom, the point of propagation would be higher, allowing molecules to be trapped (Figure 4).



**Figure 4: Diagram of Optical Trap with Orientations of Cover Slip**

In addition to this change in the orientation of the slide, there were also a few adjustments made to verify everything was in working order to isolate the problem to just the orientation of the slide. This was done by focusing the camera with the z-axis controls. This was to verify that the beads were in the correct z-position to be trapped by the laser. This was indicated by either a white or black dot in the center; when the black dots were in focus on the camera, the optical trap was most effective, but when the dot in the center was white, the optical trap was non-functional, meaning that this change in the z-axis to focus the camera was also a factor in making the system functional (Figure 5).



**Figure 5: Focused Beads at Varying Z-Axis Values**

In our tests, the QPD was aligned and centered before every trial. If the QPD was not centered, the force readings that would be sent to the computer would be inaccurate and cause incorrect conclusions to be drawn.

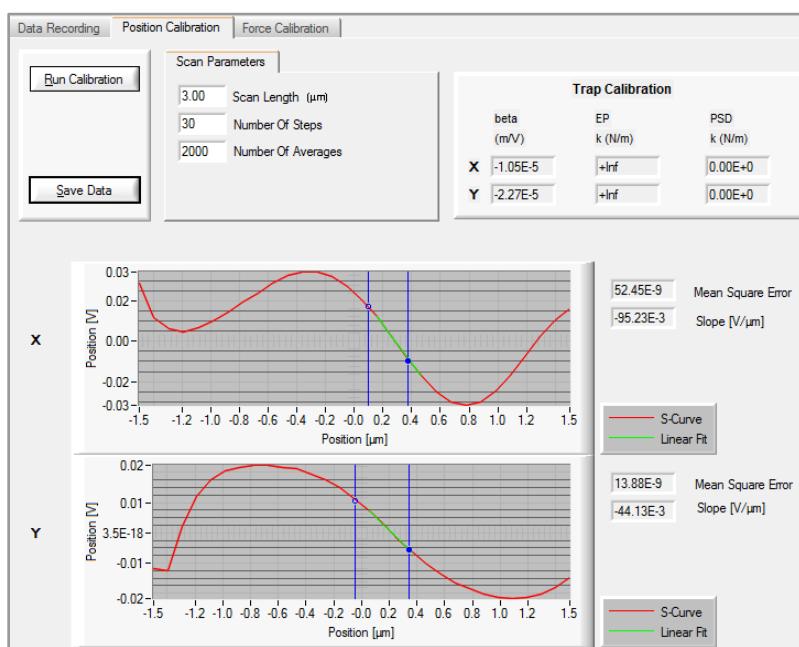
No physical components of the Optical Tweezers setup were changed because when the tests were run after making only the changes stated previously, the device was able to fully trap a bead without changing anything with the physical laser components.

## II. Calibration and Spring Constant

Three calibration tests were run to gather the necessary data to determine the location of the laser's focal point relative to the camera's field of view and the strength of the laser trap. Two main methods were used to find the laser's focal point, the first using a stuck bead (a silica bead attached to the cover slip so it cannot be moved by the laser), and the second using a free bead (a silica bead floating freely through the solution).

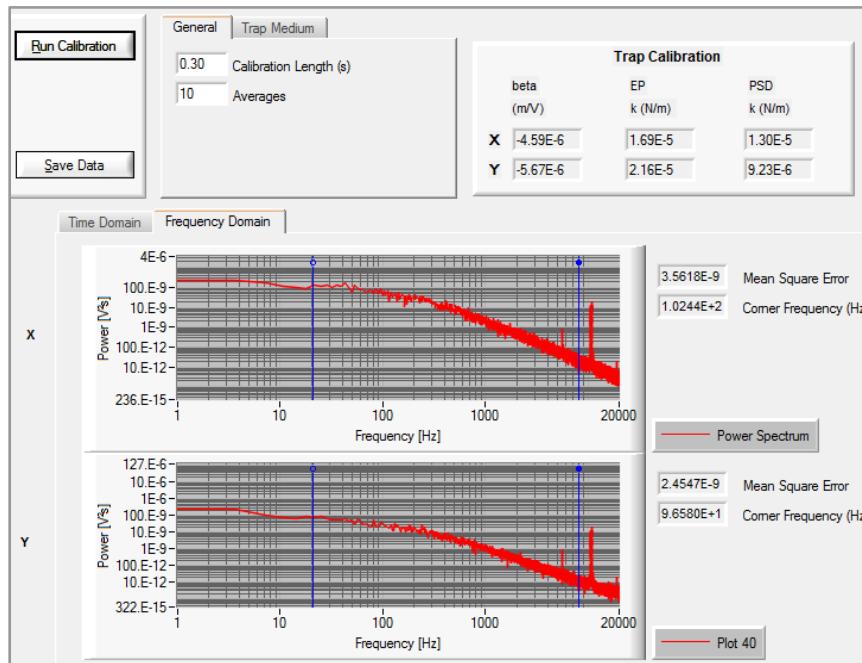
For the first method, a stuck bead is manually translated through the camera's field of view. When the bead is moved through the laser's focal point, the QPD output spikes. With the bead centered on this point, the Data Recording calibration is able to further specify the center of the trap. This calibration is run by oscillating the stage (and thus the stuck bead) through the trap, allowing the software to record the point at which the laser's path is diverted the most by the bead, thereby revealing the exact trap location.

For the second method, a free bead is trapped as normal, with the laser at high power. A free-floating bead will gravitate to the exact center of the trap, allowing for a more rapid determination of the focal point of the laser. One of the major flaws in this calibration technique is that it assumes that the Optical Tweezers hardware is functioning properly. Since this specific device was nonfunctional upon beginning this experiment, the first calibration method was required in order to verify that the trap was working properly. For most experiments, including our later trials, the second calibration method is both faster and more accurate.



**Figure 6: Results of Position Calibration**

The second of the three calibrations is the Position Calibration, which is run using a stuck bead centered on the laser's focal point. Utilizing the OTKBEM-CAL software, an automatic calibration algorithm can be run on the bead. This algorithm runs the slide (including the stuck bead) through a series of minor movements and records the position conversion factor of the laser (as shown in figure 6).



**Figure 7: Results of Force Calibration at 400 mA**

The final calibration test is the Force Calibration, which is run on a trapped free-floating bead. This trial is also run automatically through the OTKBEM-CAL software. This test runs the slide through a series of oscillations similar to those of the Position Calibration. However, since the bead is stuck inside the trap, it does not move relative to the camera, so it may look like nothing is moving. The results of this trial are the Power Spectral Analysis (PSD) and the Corner Frequency. Both of these values are dependent on the current running through the laser, so gathering multiple values at various currents is required.

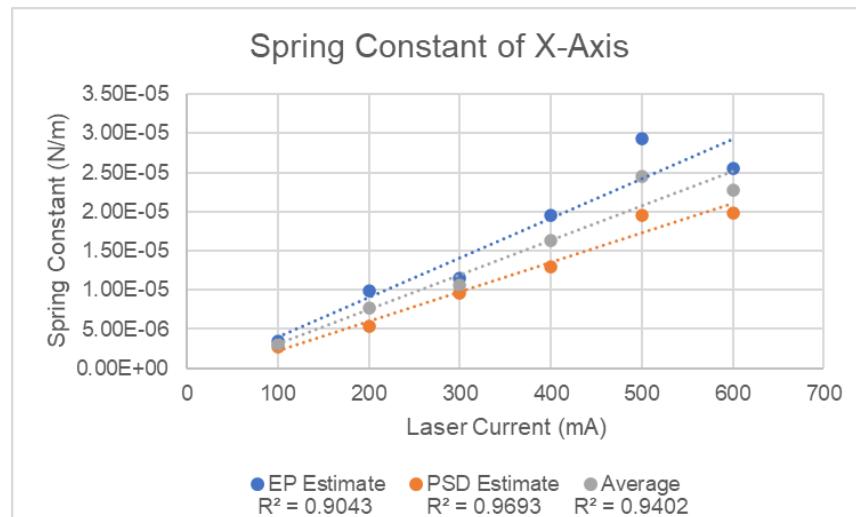
It is necessary to run the tests roughly in order, with one exception. The method chosen to find the center of the trap will determine whether the Position Calibration or the Force Calibration should be run second. If the first method of data calibration is used, the Position Calibration should be run second since both tests use a stuck bead. Similarly, the Force Calibration should be run second if the other method of Data Calibration is chosen, as both use a free-floating bead.

All of these tests together provide enough information to determine the force the trap exerts on a trapped bead. Since the bead is stuck inside the trap due to the force the laser light is exerting on it, the system acts similarly to a mass on a spring, meaning that the force of the trap can be represented by  $k$ , the spring constant of the system.

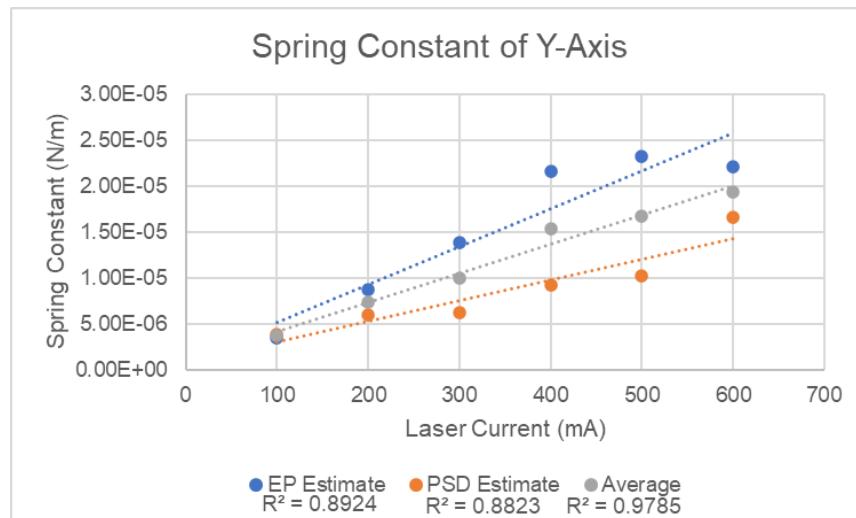
The Power Spectral Density (PSD) is a method of normalization and streamlining of the varied particle vibrations and motions. This normalization allows for more precise measurements of the energy of the trap and its effect on the bead.

The Electrostatic Potential Energy (EP) is the amount of energy that the bead has while it is stuck in the laser. When the bead is translated towards the edge of the trap, part of this energy is converted to kinetic energy. By combining these readings with the oscillation velocity, the spring constant of the optical trap can be determined.

The OTKBEM-CAL software estimates the spring constant twice, once using the PSD readings,<sup>vi</sup> and once using the EP readings.<sup>vii</sup> By averaging these two values together, a value of the spring constant at various laser amperages can be calculated, showing the relationship between the strength of the trap and the laser current.



**Figure 8: Spring Constant of X-Axis**



**Figure 9: Spring Constant of Y-Axis**

As shown in figures 8 and 9, the OTKBEM-CAL software's EP estimates for the spring constant were consistently higher than its PSD estimates. On both graphs, however, the explained sample variance ( $R^2$  value) was greatest in the average of the two methods, proving its statistical viability.

For all trials of both the  $x$ -axis and  $y$ -axis, increasing the current, and thus the power, flowing through the laser system caused a direct linear increase in the spring constant of the trap. The spring constant is a direct representation of the strength of the trap, showing that the trap becomes stronger as the power flowing through it is increased. These values are similar to other Optical Tweezers systems of the same design, reaffirming that the system is fully functioning.

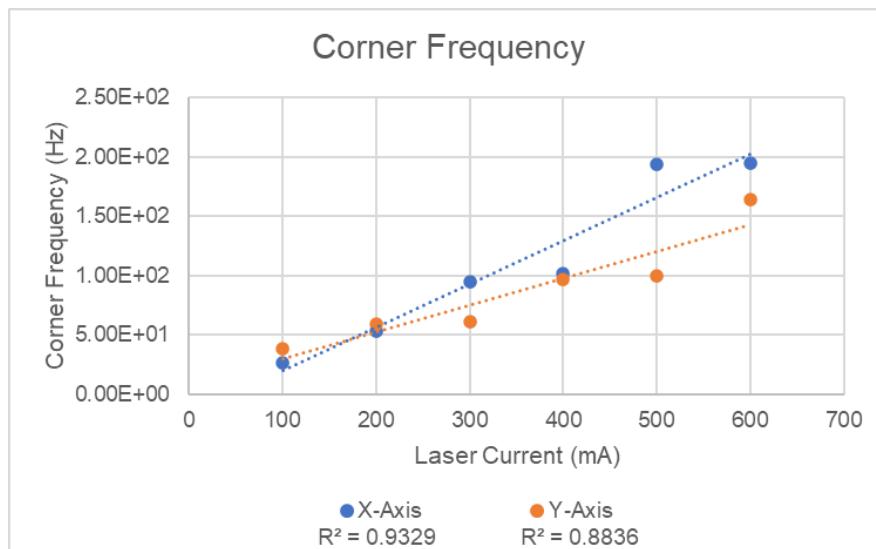
This data is also consistent with qualitative observations taken during the trials. During the lower-current trials, the bead was clearly lagging behind the trap center, and if the slide was moved too rapidly, the bead would slip completely out of the trap. At the higher currents, the bead was kept very rigid inside the trap and did not slip out even when the slide was translated rapidly.

Difference Between X-Axis and Y-Axis			
Current (mA)	X Average	Y Average	Difference (%)
100	3.07E-06	3.70E-06	-17.027
200	7.66E-06	7.33E-06	4.505
300	1.06E-05	1.00E-05	5.653
400	1.63E-05	1.54E-05	5.741
500	2.45E-05	1.68E-05	45.970
600	2.27E-05	1.94E-05	17.313
Average			10.359

**Figure 10: Percent Difference Between the X-Axis and Y-Axis**

The optical trap, by its nature, is designed to be a circular trap, with a constant, equal force pushing at the bead from all sides. With this specific device, the spring constants for the  $x$ -axis at varying currents average 10.359% stronger than their  $y$ -axis counterparts. The difference in the intensity of the forces pushing on the bead would cause the physical shape of the trap to distort, transforming it into an ellipsoidal shape.<sup>viii</sup>

However, as described in section I, one of the major issues with the tweezers apparatus was the misaligned laser pathway. This misalignment caused the laser to travel into the focusing lens at an angle, which may be the reason that the trap is in the shape of an ellipse. Even with this caveat, the trap is still able to perform near its full functionality.



**Figure 11: Corner Frequency of X-Axis and Y-Axis**

The corner frequency (also known as the cut-off frequency) is the frequency at which the power transferred through the laser trap begins to decrease. This occurs as the bead is being translated towards the outside of the trap. Essentially, it describes the area in the trap where the bead will stay stuck in the laser.<sup>ix</sup> The data in figure 11 supports the assertion that this system has an ellipsoidal trap, since the y-axis corner frequencies are significantly lower than the x-axis frequencies. This shows that it takes less energy for a bead to escape through the y-axis, meaning that it has a smaller trap area than the x-axis, confirming the trap's ellipsoidal nature. Additionally, Figure 11 shows that as the current increases, the effective trap area increases as well.

### III. Biological Specimens and Further Applications

Within the scope of the experimentation, the Optical Tweezers were used to manipulate various biota, in accordance with the system's primary use in research currently. In total, the system was used on cooked and uncooked onion cells, a leaf, an orange, a sample from a "Mike and Ike's" candy, and the membrane of a hard-boiled egg. This section will further present some of this qualitative data obtained after the system was calibrated, and will further examine previous uses of the system in a biological setting, as well as some potential future applications.

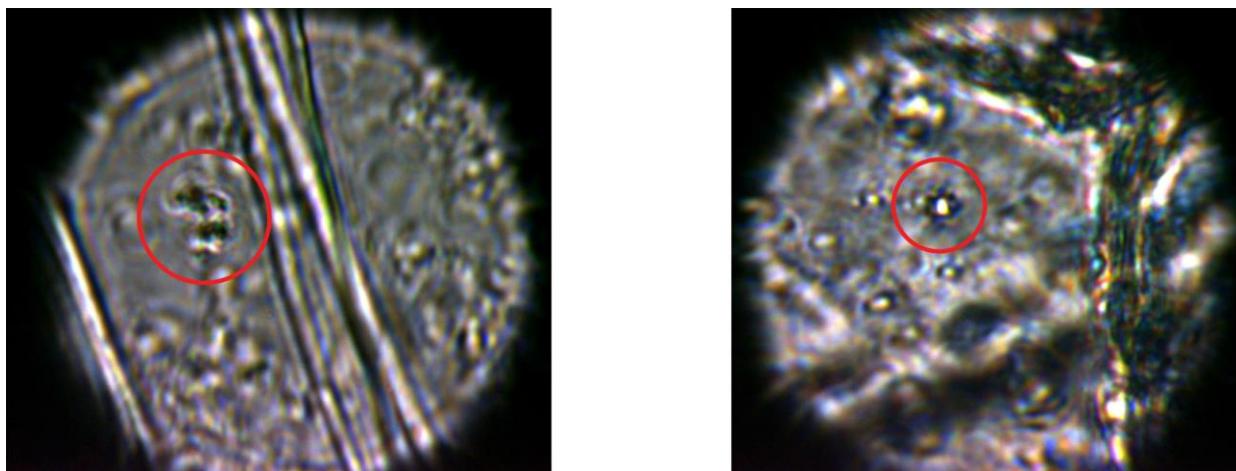
#### A. Qualitative Data

To manipulate organelles later in the experiment, trap calibration needed to be performed with the silica beads. This allowed for a more controlled trap, as there was less interference from the cells' cytoplasm by virtue of suspending the silica beads in distilled water, and in addition, it was easier to see how a trapped organelle, modeled by the bead, would behave in comparison to a stuck one. This created the grounds for the observation that a trapped bead would travel with the laser as the microscope stage moved beneath it, whereas a stuck bead would move with the stage while the laser moved independently.

Furthermore, when a bead was trapped by the laser, and the laser was subsequently turned off, the stabilizing net force provided by the laser disappeared, as expected, and the bead drifted away from where

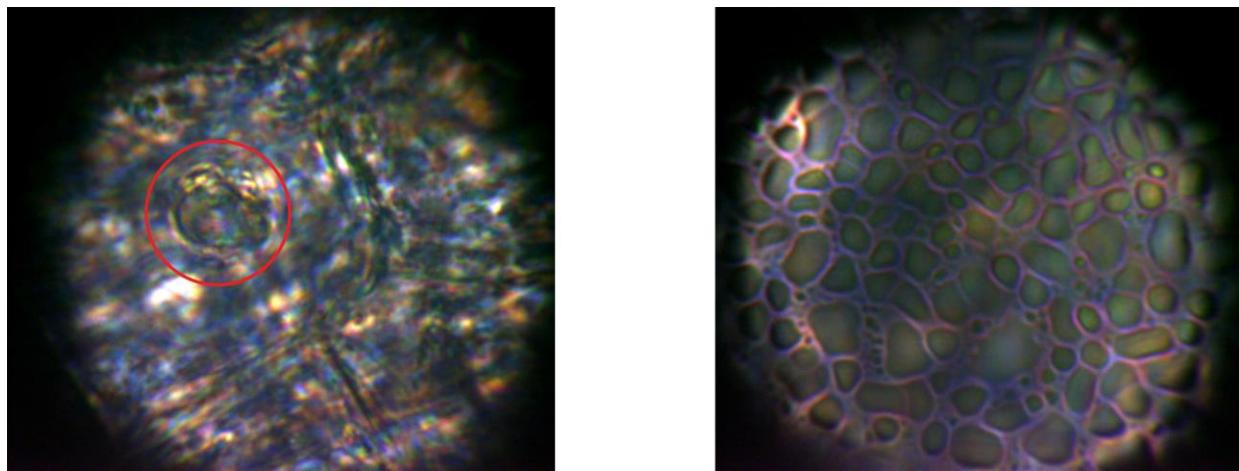
the laser had trapped it. After the bead had drifted away, the laser was turned back on, and when the laser's focal point was below the silica bead, the bead was attracted to the laser and was pulled toward the focal point until it eventually snapped back to where it once was. As suggested earlier, this effect was inverted when the focal point of the laser was situated above the silica bead due to the force of the laser pushing away from the focal point below it. Extending this, when the other microbiota were placed under the microscope, certain organelles were more attracted to the laser as a function of how close they were to the laser, while others were repulsed by the laser when they were at a lower z-height.

The cooked and uncooked onions displayed significantly different behaviors under the microscope, affecting the ability of the laser to trap their organelles in a similar vein. Because the uncooked onion skin cells were more recently alive, they occasionally showed a more greenish color in their organelles (Figure 12a), whereas the cooked cells showed more free-flowing particles outside of the cells, likely the seasoning and oil cooked with the onions (Figure 12b), and very little motion inside of the cells. In addition, a rainbow effect appeared frequently throughout the cooked onion cells, further suggesting the incorporation of oil. As a result, capturing organelles within the uncooked onion cells was relatively simple, but more difficult within cooked cells. The other organelles that were observed in the cooked cells were likely stuck to the slide, as there was little to no observed motion from them, and they were largely unreactive to the laser, even though they were about the size that the laser generally traps.



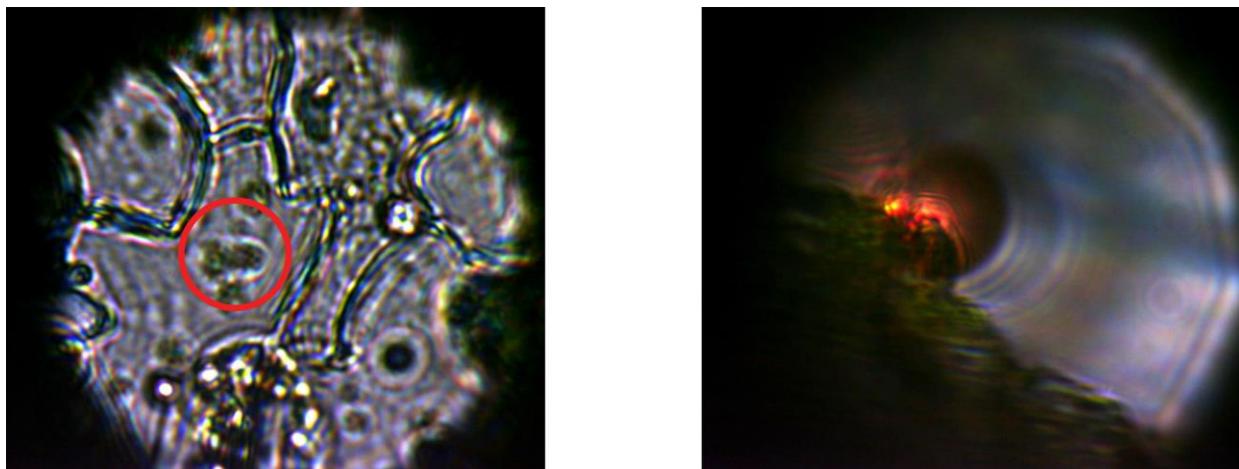
**Figure 12a, 12b: The frozen (left) and cooked (right) onions, with cell walls and trapped particles.**

The cell cross-section of the orange fiber proved to demonstrate the full power of the laser; here, an orange cell's nucleus was completely caught within the laser beam (Figure 13a), highlighting a valuable application of the laser in manipulating the largest organelle of the cell. Despite this, the trap strength showed its limitations, as any attempt to force the nuclear envelope outside of the cell resulted in the envelope bouncing off of the cell wall. In the water layer above the cell cross-section, the laser was able to break the surface tension of a large water droplet that had formed, which quickly reformed into a single droplet after the laser had been turned off (Figure 13b).



**Figure 13a, 13b: An orange cell, with a trapped nucleus (left), and with a shattered droplet (right).**

The laser's variable power on a microscopic scale was further highlighted in its interactions with the leaf. It had been collected directly from a living tree, so the living cells and organelles were more active and less likely to be stuck. In examining the leaf, the laser beam came into contact with a chemically fragile part of the leaf structure, and upon contact with the laser, that part combusted (Figure 14b). After four more such sparks, the region of the slide that contained the explosion and the blackish bubble was covered with what appeared to be the remainder of the bubble. The laser was also able to capture what appeared to be a mitochondrion in a cell when moved to a part of the surface of the leaf (Figure 14a), but as was the case with the orange cell's nucleus, the trapped mitochondrion was unable to break through the cell wall.



**Figure 14a, 14b: A leaf with a trapped mitochondrion (left), and with the exploding bubble (right).**

Finally, the egg shell membrane and Mike and Ike candy showed the limitation of the laser in dealing with more macroscopic objects. The shell membrane, hypothesized to contain a great number of free-floating particles to examine, showed no more than mostly rough surfaces and one isolated crystalline structure. Similarly, the Mike and Ike's candy mostly had a uniform surface from the perspective of the laser, but there were occasional bright spots and ridges. Neither sample elicited a response to the laser being activated –

while there was an isolated part of the candy that had broken off, it appeared to be stuck to the slide and was not trapped by the laser.

## B. Further Applications

With applications and uses for nanotechnology increasing greatly in recent years, having devices that are capable of manipulating small objects without coming into physical contact with them is almost a necessity. Optical Tweezers, especially devices capable of trapping multiple objects (either through multiple lasers or by switching the trap focus quickly between multiple objects), are able to build and maneuver small particles and sections of the nanomachine.

Optical Tweezers are also capable of forcing molecules to form bonds with each other in situations where they otherwise would have stayed separate. This can extend to forcing molecules that are normally stable and do not make new bonds to pair with another compound, creating endless possibilities for new combinations and substances.

### 1. Biology

Within the realm of biology, the Optical Tweezers system has found use as a means to manipulate cells and the organelles within them. As Polimeno et al. noted in 2018,<sup>x</sup> the system has been used to calculate the force required to break apart red blood cells and analyze their elastic properties. To do this, researchers first attached silica beads akin to the ones used in the above experimentation to red blood cells, then used two optical traps, moving them in opposite directions, in order to attempt to elastically deform the cell.<sup>xi</sup> It was also found that the optical trap does not inflict damage on the trapped cell, and that the tweezers only had a distortive effect when the force applied by the separating tweezers was greater than 10-15 picoNewtons.

### 2. Chemistry

Optical trapping techniques have made it possible to analyze single nano-particles during Raman spectroscopy, fluorescence, and absorption spectroscopy. During Raman spectroscopy, a laser of a single wavelength impacts the particle, sending the light back as scattered, frequency-shifted photons. An optical tweezer apparatus is capable of both trapping the particle and separating it from other molecules, as well as scattering the light that is gathered as data.

Microrheology studies also benefit from including Optical Tweezers to move particles. By trapping a particle and moving it through a solution, the viscoelastic properties of the fluid can be determined by studying the forces the liquid exerts on the trapped particle.

### 3. Nanotechnology

The Optical Tweezers system also finds a strong use case in nanotechnology – both more obviously as a means of manipulating and creating nanodevices, but also as a novel way to power them. The angular momentum of the light rays that stabilize trapped particles can be imparted onto the particle, causing it to rotate.<sup>xii</sup> In addition, a novel approach has produced a steam engine on a scale of one micrometer. The use of the Optical Tweezers system to superheat a liquid with a bead below the waist of the laser causes the

bead to become attracted toward the waist further. In turn, the surface of the bead closest to the waist becomes heated by the laser as it comes closer. This effect magnifies the superheating effect of the fluid at the waist of the laser, and to stabilize this situation, the fluid spontaneously becomes a vapor bubble, subsequently collapsing in on itself in a small explosion, finally pushing away the bead.<sup>xiii</sup>

## IV. Acknowledgements

We would like to thank our project advisor and PGSS Director, Dr. Barry Luokkala, for giving us the generous opportunity to experiment on the optical trapping system. We would also like to thank our project TA, Josh Kipiller, for the guidance he provided.

## V. References

- 
- <sup>i</sup> Lumicks, "What are optical tweezers?", <https://lumicks.com/knowledge/what-are-optical-tweezers/> (2022).
- <sup>ii</sup> ThorLabs, "Optical Tweezers Tutorial," [https://www.thorlabs.com/newgroupage9.cfm?objectgroup\\_id=10774](https://www.thorlabs.com/newgroupage9.cfm?objectgroup_id=10774)
- <sup>iii</sup> ThorLabs, "Modular Optical Tweezers," [https://www.thorlabs.com/NewGroupPage9.cfm?ObjectGroup\\_ID=3959](https://www.thorlabs.com/NewGroupPage9.cfm?ObjectGroup_ID=3959)
- <sup>iv</sup> MIT Department of Physics, "Optical Trapping," <http://web.mit.edu/8.13/www/JLExperiments/JLExp51.pdf> (2017).
- <sup>v</sup> Michael Murphy, "Modern Lab: Optical Tweezers," Physics Department, Carnegie Mellon University (2021).
- <sup>vi</sup> Pedro Almendarez-Rangel, Beatriz Morales-Cruzado, Erick Sarmiento-Gómez, and Francisco G. Pérez-Gutiérrez, "Finding trap stiffness of optical tweezers using digital filters," *Applied Optics* 57(4), 652-658, (2018).
- <sup>vii</sup> Joseph Junio, Jack Ng, Joel A. Cohen, Zhifang Lin, and H. Daniel Ou-Yang, "Ensemble method to measure the potential energy of nanoparticles in an optical trap," *Optics Letters* 36(8), 1497-1499 (2011).
- <sup>viii</sup> Antti Kauppila, Matti Kinnunen, Rashes Karmenyan, Risto Myllylä, "Measurement of the trapping efficiency of an elliptical optical trap with rigid and elastic objects," *Applied Optics* 51(23), 5705-5712 (2012).
- <sup>ix</sup> Erwin J. G. Peterman, Frederick Gittes, and Christoph F. Schmidt, "Laser-induced heating in optical traps," *Journal of Biophysics* 84(2), 1308-1316 (2003).
- <sup>x</sup> Paolo Polimeno, Alessandro Magazzù, Maria Antonia Iati, Francesco Patti, Rosalba Sajja, Cristian Degli Esposti Boschi, Maria Grazia Donato, Pietro G. Gucciardi, Philip H. Jones, Giovanni Volpe, Onofrio M. Maragò, "Optical Tweezers and their applications," *Journal of Quantitative Spectroscopy and Radiative Transfer* 218, 131-150 (2018).
- <sup>xi</sup> Sylvie Hénon, Guillaume Lenormand, Alain Richert, François Gallet, "A new determination of the shear modulus of the human erythrocyte membrane using Optical Tweezers," *Biophysical Journal* 76(2), 1145-1151 (1999).
- <sup>xii</sup> Miles Padgett and Roberto Di Leonardo, "Holographic optical tweezers and their relevance to lab on chip devices," *Lab on a Chip* 7, (2011).
- <sup>xiii</sup> Pedro A. Quinto-Su, "A microscopic steam engine implemented in an optical tweezer," *Nature Communications* 5, 5889 (2014).



# The Behavior of Magnetic Moments at Low Temperatures

Grace Krakauskas, Brian Olsen, Pradyun Solai, Julia Vizza, Francis Wehbe, and Matthew Wehler

## Abstract

This experiment focused on the magnetic susceptibility of different sample materials and how their electron configurations in conjunction with the temperature of the system affected their magnetic properties. To assess the magnetic susceptibility at different temperatures, the samples were measured in two different environments. One environment was room temperature and involved the suspension of the materials above a magnetic field as the apparent change in weight was measured on a balance scale. The other environment was immersed in liquid nitrogen which has a temperature of 77K. That method involved suspending the samples inside a chamber that was thermally in contact with the substances while still physically isolated from it by a thin container of silicon dioxide. The results revealed that with a lower entropy (caused by a decrease in temperature), the magnetic properties were more pronounced. Accurate calculations of varied materials' magnetic susceptibilities could aid high-precision scientific equipment such as particle accelerators and magnetometers used to measure planetary magnetic fields.

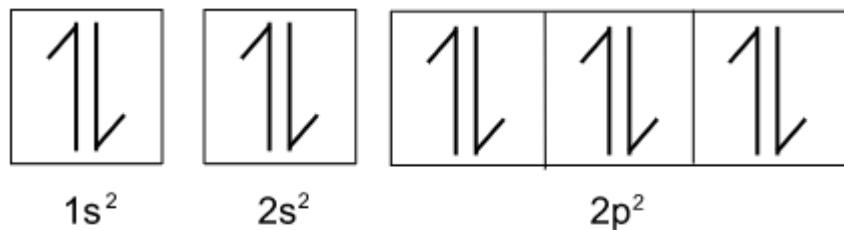
## I. Introduction

### A. Types of magnetic materials

Different materials have different responses to an external magnetic field, depending upon the arrangement of their valence electrons. Materials with unpaired electrons have a tendency to be drawn into a magnetic field gradient, while those with no unpaired electrons tend to be pushed out of a field gradient.

#### 1. Diamagnetism

Materials that are repelled by a strong magnetic field are considered diamagnetic. In diamagnetic materials, all the electrons are paired; this results in none of the electrons polarizing to line up with the magnetic field. The electrons move slightly and generate small electric currents that oppose the direction of the magnetic field; these effects are small. It is expected that these materials will experience an electromagnetic force to move the material out of the magnetic field. Most organic compounds and heavy metals, as well as all noble gases, show diamagnetic properties<sup>1</sup>.

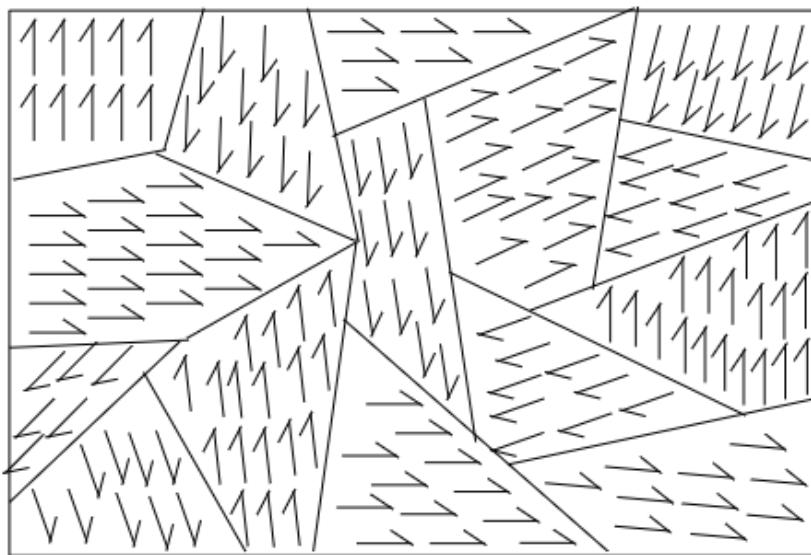


**Figure 1(A): Diamagnetic electron configuration of neon gas**

Neon gas exhibits diamagnetic properties due to having only paired electrons.

## 2. Ferromagnetism

Materials that are strongly attracted to a magnetic field are ferromagnetic. Ferromagnetic materials have unpaired electrons that are organized into large groups of atoms that have unpaired electrons aligned in spin known as electron domains. This formation results in a strong magnetic attraction of the electrons. In this experiment iron, cobalt, and nickel are ferromagnetic<sup>ii</sup>.



**Figure 1(B): ferromagnetic electron configuration of Iron**

Figure 1(B) shows the electron spin directions. All the arrows are in separate directions to cancel each other out, preventing a magnetic field. Iron has unpaired electrons which allow it to exhibit ferromagnetic properties.

## 3. Paramagnetism

Materials that are weakly attracted to a magnetic field are paramagnetic. When a material has unpaired electrons and fails to orient itself into magnetic domains, it reacts well with strong magnetic fields but has a lower susceptibility in comparison to ferromagnets. Magnets with susceptibility (or the measurement of the degree of attraction or repulsion that a material has to a magnetic field) between  $1.0 \times 10^{-5} \text{ m}^3/\text{kg}$  and  $1.0 \times 10^{-4} \text{ m}^3/\text{kg}$  are classified as paramagnetic<sup>iii</sup>.

## 4. Antiferromagnetism

When materials are similar to paramagnetic, but drastically change with a low temperature are antiferromagnetic<sup>4</sup>. This change occurs at a specific temperature known as the Néel temperature, which falls within the range of 300°C to 400°C. At temperatures above this range, the material's ions surrounding it will behave as "micromagnets," making the material exhibit the aforementioned paramagnetic properties<sup>iv</sup>. These magnetic moments will then align antiparallel to their neighbors at or below the Néel temperature. Due to their antiparallel alignment, the material will then exhibit a magnetic moment of 0, appearing nonmagnetic<sup>4</sup>.

## 5. Superconductivity

Superconductive materials possess little to no magnetic properties at room temperature; however, when they reach a critical temperature, they begin to exert a magnetic field that expels all magnetic forces applied to it back out. When the critical temperature is reached, the electrons of the superconductors align their spins anti-parallel creating a perfect diamagnet. There are two types of superconductors classified by how they break down when exposed to a strong external magnetic field. Type I superconductors break down abruptly due to a first-order phase transition when a material changes to another form of matter as the strength of the external magnetic field exceeds the critical value. As opposed to type I, type II superconductors have two critical magnetic field strengths. The first critical magnetic field strength creates small vortexes, forming circulating currents of the external magnetic field within the material; it still retains its ability to superconduct in areas that are not affected. The second strength is built by vortexes that overcome the conductor and eradicate its ability to conduct the magnetic field present<sup>v</sup>.

## B. Random movement of electrons

Different materials are sorted into separate magnetic categories because of the locations and configurations of their electrons. Electrons behave more erratically in groups of atoms with higher average kinetic energy than in groups with lower average kinetic energy. Average kinetic energy is measured as temperature because the temperature of the substance is just a factor of the random movement of the atoms that comprise the substance.

## C. The relationship between temperature and magnetism

As temperature increases, a substance's total entropy increases, which means that the movement of its particles and electrons becomes more random. In ferromagnetic materials, a strong attraction to a magnetic field will be exhibited and as the temperature of that system is decreased, the magnetic attraction will increase. However, in antiferromagnetic substances, the attraction will slightly increase as temperature decreases, and then form no attraction at all once it has hit the critical "Néel" temperature<sup>vi</sup>. Conversely, paramagnetic materials will exhibit very little attraction at room temperature, but once that temperature decreases, they will become significantly more attractive. A superconductor will experience some magnetic attraction at room temperature, but at an exceptionally low temperature, it will begin to expel its own magnetic field through the area.

The change in the system's entropy leads to a change in the magnetic properties a substance exhibits. Therefore, if a substance is exposed to low temperatures, then its exhibited magnetic properties will be enhanced.

## D. Equations

The force on an object is equal to the change in energy divided by the change in distance. This is the definition of energy.

$$F = \frac{dE}{dy} \quad (1)$$

The energy of a magnetic moment  $\vec{\mu}$  within an external magnetic field  $\vec{H}$  is equal to the dot product  $E = -\vec{\mu} \cdot \vec{H}$ . Plugging this into the earlier equation:

$$F = \frac{d}{dy} (-\vec{\mu} \cdot \vec{H}) \quad (2)$$

Pulling the  $-\vec{\mu}$  out of the derivative:

$$F = -\vec{\mu} \cdot \frac{d\vec{H}}{dy} \quad (3)$$

The magnetic moment can be rewritten in terms of the magnetization  $\vec{M}$  of the object, which is defined as magnetic moment per unit volume, or  $\vec{M} = \frac{\vec{\mu}}{dV}$ <sup>vii</sup> where  $dV$  is an infinitesimal volume element. Using  $\vec{\mu} = \vec{M}dV$ :

$$F = -\vec{M}dV \cdot \frac{d\vec{H}}{dy} \quad (4)$$

The magnetization within the sample is caused by the external magnetic field, and can be calculated using the magnetic susceptibility  $\chi$ , which measures how strongly the sample magnetizes due to an external magnetic field. They are related by  $\vec{M} = \chi\vec{H}$ .

$$F = -\chi\vec{H}dV \cdot \frac{d\vec{H}}{dy} \quad (5)$$

Because the cross-sectional area is constant,  $dV$  can be rewritten as  $Ady$  where  $dy$  is an infinitesimal height element.

$$F = -\chi\vec{H}Ady \cdot \frac{d\vec{H}}{dy} \quad (6)$$

The  $d\vec{H}$  and  $\vec{H}$  are in the same direction, so they can be treated them as scalars, and the  $dy$  cancels out.

$$F = -\chi HAdH \quad (7)$$

This gives the force upon the infinitesimal volume element. To find the total force on the object one must take the integral from the bottom of the sample to the top of the sample.  $H_B$  and  $H_T$  are the external magnetic fields at the bottom and top of the sample, respectively.

$$F_{Total} = \int_{H_B}^{H_T} -\chi HAdH \quad (8)$$

Evaluating the integral and distributing the negative:

$$F_{Total} = \frac{1}{2}\chi A(H_B^2 - H_T^2) \quad (9)$$

Since the force on the object is being measured by the apparent change in mass on a scale, the apt equation for this is given by  $F_{Total} = \Delta mg$ , where  $\Delta m$  is the change in mass, and  $g$  is gravitational acceleration.

$$\Delta mg = \frac{1}{2}\chi A(H_B^2 - H_T^2)<sup>viii</sup> \quad (10)$$

Solving for  $\chi$ :

$$\chi = \frac{2\Delta mg}{A(H_B^2 - H_T^2)} \quad (11)$$

The magnetic susceptibility calculated above is the volume susceptibility. To compare

this collected data with existing data, it was crucial to change this to the mass susceptibility instead, which can be done by dividing by the density  $\rho = \frac{m_s}{V}$ , where  $m_s$  is the mass of the sample.

$$\chi_m = \frac{2\Delta mg}{A(H_B^2 - H_T^2) \frac{m_s}{V}} \quad (12)$$

Using  $V = Ah$  where  $A$  is the cross-sectional area and  $h$  is the height of the sample:

$$\chi_m = \frac{2\Delta mgh}{m_s(H_B^2 - H_T^2)} \quad (13)$$

## II. Methods

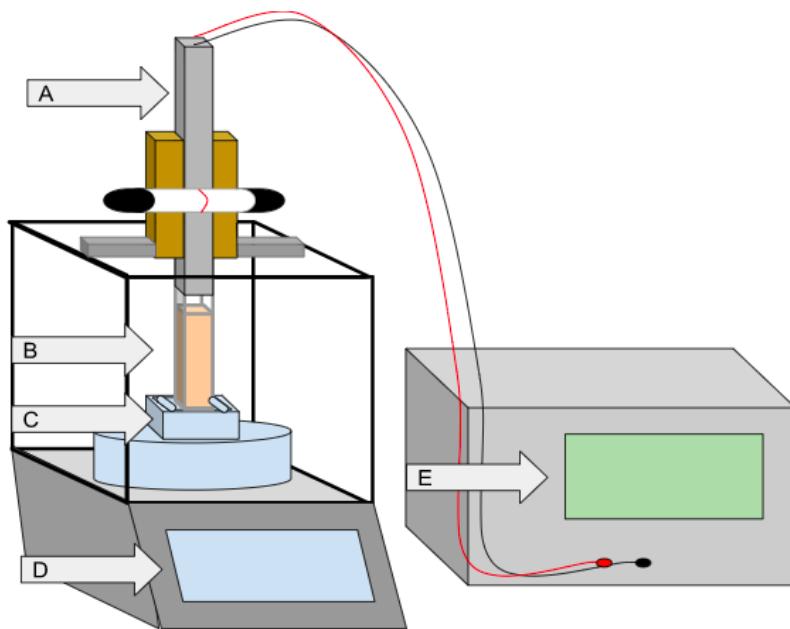
To investigate the relationship between electron configuration, temperature, and magnetic moments, materials with different electron configurations were placed in magnetic fields at two different temperatures. Each method explored a separate aspect of the system; the tabletop method investigated a simple model of the change in mass exhibited at room temperature to establish the relationship between the change in weight of the magnets with the magnetic susceptibility of the sample. Additionally, a “low temperature” model offered the changes in mass at 77K with a room temperature basis as well to compare the entropies of the samples at different temperatures and the resulting effect on the magnetism. In addition, an Ising model offered a theoretical interpretation of how the alignment of the atomic dipoles of the substances affected their magnetic attraction.

### A. Samples Used

The two experiments sampled Mohr’s salt, copper, titanium, manganese carbonate, zinc, silicon dioxide, nickel sulfate, and aluminum. Additionally, the low-temperature experiment utilized the superconductor Yttrium Barium Copper Oxide.

### B. Tabletop method

The tabletop method provides a baseline understanding of how the materials used will react to a magnetic force so that a comparison can be made to their behavior at low temperatures. Experiment 1 was performed using two alkaline earth magnets at room temperature.



**Figure 2(A): Room temperature apparatus**

- A. A sample holder apparatus consisting of an adjustable slider that locks into the sample cuvette attached to a mainframe that locks onto the top of the scale for stability
- B. The cuvette attaches to A and contains the samples.
- C. The magnet system consists of an open-topped box with a uniform magnetic field inside.
- D. The entire apparatus is placed onto a scale
- E. The system is connected to an ammeter which can read the magnetic field going through a wire in a cuvette. This makes it possible to read the magnetic field in between the magnets.

### 1. The magnetic field between the magnets

First, the magnetic field strength was calculated using the relationship  $B = F/Il$ .  $B$  represents the magnetic field being measured;  $F$  is the magnetic force acting on the object in the field;  $I$  is the current moving through the object, typically a wire;  $l$  is the length of the object being acted upon. This equation is derived from the dot product  $F = ILB$ , where  $F$ ,  $I$ , and  $B$  are vectors. Due to the magnetic field being directly perpendicular to force and current, it is appropriate to represent the relationship as  $B = F/Il$ .

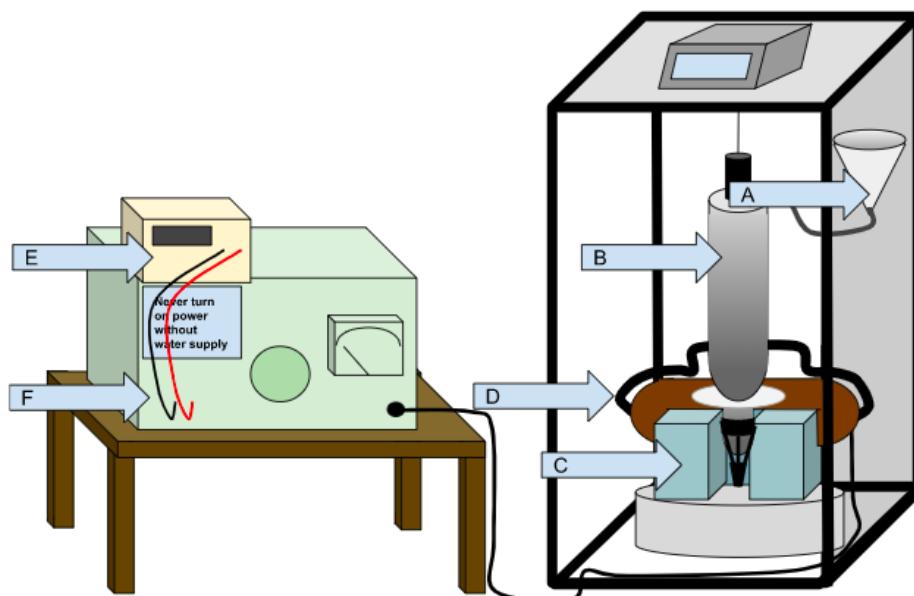
A cuvette was inserted into the apparatus containing copper foil to apply a specific amount of current to the magnetic field and induce a change in the weight of the system. Next, a cuvette containing a sample was lowered into the chamber with its bottom located at the center of the magnets. The initial mass with no current added was recorded; then, amperage was applied in increments between 0 and 1.

### 2. The susceptibility at room temperature

The voltmeter adapter was replaced with the sample adapter; the voltmeter was turned off. Mohr's salt was attached to the sample adapter, and the sample was lowered into the magnetic field. Then, the bottom of the cuvette was placed at the center of the magnets. Next, the sample was lowered so that the top is between the magnets. Finally, the change in weight was calculated as shown on the electronic precision balance that measures the apparent mass of the magnet. This measures the difference between the force

of gravity and the force that the magnetic field exerts on the magnet. This was repeated using nickel sulfate, manganese carbonate, aluminum, titanium, and copper.

### C. Low-temperature method



**Figure 2(B): low-temperature apparatus**

- F. The funnel leads into the nitrogen insulation in the sample chamber.
- G. The sample chamber houses the  $N_{2(l)}$  and the sample tube. The sample tube is physically isolated from the  $N_{2(l)}$  chamber.
- H. The electromagnets are used to create the magnetic field; they are cooled by a water-cooled system.
- I. An ammeter measures the current flowing through the magnet.

### D. Electromagnet method at room temperature

First, the water-cooling system was activated to cool the electromagnet, and the room temperature was recorded. Next, a Gauss meter measured the magnetic field at both the top and bottom positions of the samples (a difference in height of about 14.3 cm). After removing the Gauss meter, one of the samples was lowered on a string between two of the electromagnets and upon rest its initial mass was recorded. Then, the magnet was turned on and the mass was recorded again at various levels of current. These methods were tested on Mohr's salt, manganese carbonate, zinc, nickel sulfate, yttrium barium copper oxide, and silicon dioxide.

### E. Electromagnet method at 77 K

To avoid contact between the liquid nitrogen and the air, the apparatus was purged with helium before the start of the experiment. Because liquid oxygen is paramagnetic, its liquid form could change the apparent mass of the sample. First, the valves of the helium tank were opened to induce the input and output pressures. Next, the chamber surrounding the electronic precision balance and the thermal chamber were attached to the tubes. After five to ten minutes, the areas were completely flushed. While waiting, two thermoses of liquid nitrogen were collected. A sample was then inserted into the thermal chamber and the

air access was closed off. Then, the entire sample was covered in liquid nitrogen from inside the thermal chamber. After ten minutes, the system obtained thermal equilibrium and the measurements were calculated. The process of changing the current exerted on the substance described in the room temperature experiment was imitated. Then, a Gauss meter was used to record the resulting change in mass in a similar fashion to the prior method. This was done to ensure the compatibility of the magnetic field and the results of the experiment in calculations. Any remaining liquid nitrogen was poured out onto an empty floor away from all lab participants.

## F. Ising Model

In addition to doing the experiment, we also created a computational model called the Ising Model.<sup>ix</sup>

The Ising Model is a computational model that simulates a lattice of magnetic moments and spins. It is a type of Monte Carlo simulation that uses Markov Chains to simulate the changing of the lattice. This model operates on the fundamental idea of the Pauli exclusion principle in which the minimum energy state of two electrons within adjacent atoms is to have aligned spins for ferromagnets (or anti-aligned for antiferromagnets). Thus, trending towards homogenous spin states would indicate the lowest energy states of the given system<sup>x</sup>.

The energy of a lattice is defined by this equation:

$$H = -J \sum_{\langle i, j \rangle} \sigma_i \sigma_j - h \sum_i \sigma_i \quad (20)$$

This is the Hamiltonian equation that describes the total energy in a lattice with a net magnetic moment. Through the sum of the interaction energy between neighboring spins and the spins' alignments within a given magnetic field, you can determine the overall energy of the system.

More specifically it is defined by the two negative summations. The first summation refers to the sum of the multiplications between neighboring atoms (left, right, up and down in a two-dimensional lattice). The second summation represents the sum of the spins with respect to the external magnetic field, where the spins are in binary states of -1 and 1.

Another important constant to note is  $\beta$  because it is what is used to describe the temperature of the system.

$$\beta = \frac{1}{k_b T} \quad (21)$$

In the simulation, the Boltzmann constant  $k_b$  is defined by one as the calculations are working in dimensionless units. The constant  $\beta$  itself is commonly described as the thermodynamic temperature of a given system, and in this particular case it is a canonical ensemble. For clarification, a canonical ensemble is an ensemble of systems in which an idealized infinitely large reservoir of heat exchanges energy with a much smaller system. This assumption allows is vital to accurately describe the phase transitions within the simplistic model.

### 1. Boltzmann Distribution and Partition Function

The probability that the state of the lattice is in a specific state  $i$  is given by the Boltzmann Distribution:

$$p_i = \frac{e^{-\beta E_i}}{Z} \quad (22)$$

$Z$  is the partition function that normalizes the Boltzmann Distribution. This makes it so that the sum of all of the probabilities adds up to one.

$$Z = \sum_k e^{-\beta E_k} \quad (23)$$

The ratio  $\frac{p_j}{p_i}$  is the probability that the lattice changes from state  $i$  to state  $j$ . This is called the Boltzmann Factor.

$$\frac{p_j}{p_i} = e^{-\beta(E_j - E_i)} \text{ or } \frac{p_j}{p_i} = e^{-\beta\Delta E} \quad (24)$$

Notice how the partition function  $Z$  cancels out of the equation, meaning there is no need to go through every single possible lattice configuration to calculate it

The equation above is commonly known as the Boltzmann Distribution function, and is central to calculating the transition probabilities for this model. This function returns a probability in which a system that is contact with a heat reservoir has some energy  $E$  at some temperature  $T$ .

## 2. Heat Capacity

Heat Capacity is defined as the derivative of energy with respect to temperature.

$$C = \frac{dE}{dT} \quad (25)$$

This is important to the Ising Model because a phase transition occurs whenever the heat capacity trends towards infinity. Applying physical logic to this, one can state that the energy does not contribute to the change in temperature, but rather the change in phase.<sup>xi</sup>

However, due to the random nature of Monte Carlo simulations, a more useful equation to calculate heat capacity is this:

$$C = \frac{\sigma_E^2}{k_B T^2} \quad (26)$$

It would frequently happen that after many steps, the system would reach an equilibrium around a value, but would include a nonzero error margin above and below. To account for this statistical noise, the standard deviation of the energy,  $\sigma_E$ , is evaluated for every step and used in equation 26 to calculate the heat capacity. Some crucial background for the validation of the collected data would be using the previously established exactly solved Ising model value of  $\frac{2J}{k_B \ln(1 + \sqrt{2})}$  or approximately  $\sim 2.269$  when  $J = 1$ . A phase transition is expected to occur at this temperature.

## 3. Antiferromagnetism

Antiferromagnetism occurs when adjacent spins prefer to misalign with each other. This is call anti alignment. This is signified with a negative  $J$  value. This differs from ferromagnetism because in ferromagnetism, spins prefer to align parallel with each other and  $J$  is positive.<sup>xixii</sup>

## 4. The General Structure of the Model

The Metropolis Algorithm describes how the lattice of spins evolves to equilibrate at the lowest possible energy. First, a random lattice (a 50 x 50 2D grid) is initialized with 75% spin up or 75% spin down. Spin up is represented by a 1, and spin down is a -1. Next, the energy of the lattice is calculated using the

Hamiltonian shown above. Then, a random element is chosen to change spin (down to up or up to down). The difference in energy due to this change is calculated. If the change is negative (the new lattice has lower energy than the first), the new lattice is accepted. If the change is positive, then the probability of accepting the change is given by the Boltzmann Factor shown above. The energy and mean spin of the lattice is saved at each iteration.

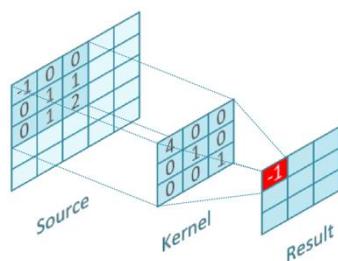
The Metropolis Algorithm is run at temperature values ranging from 0.1 to 5. For each temperature, after running the simulation for 1 million iterations, the average energy of the final 200,000 iterations (when the model is sure to have reached equilibrium) is calculated. The mean spins and the standard deviations of the energy and mean spin is also calculated.

To calculate the heat capacity of the material at each temperature, equation 26 is used. This is the square of the standard deviation of the final 200,000 iterations at the temperature divided by the Boltzmann Constant (which is equal to one) times the temperature squared.

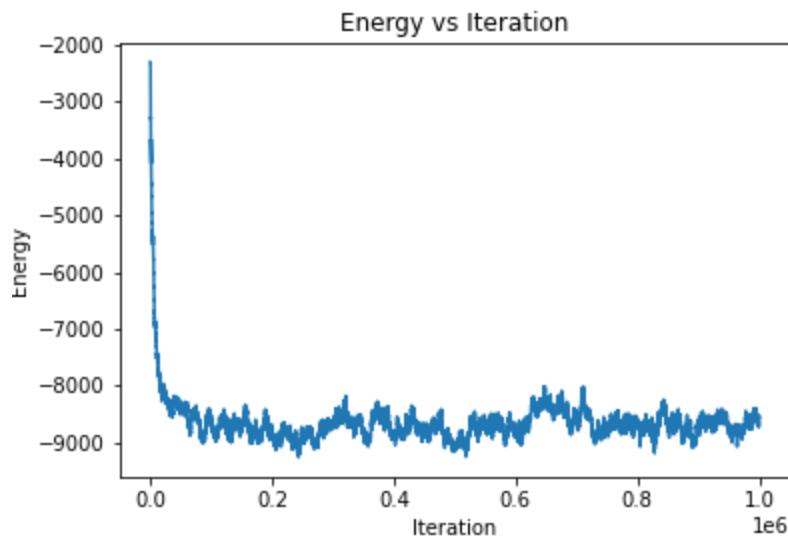
For antiferromagnetism, the Metropolis Algorithm was run again but with J being rather set to -1. The negative J value signifies that the sample is antiferromagnetic and the spins prefer to anti-align. The final state of the lattice after one million iterations.

## 5. Optimization of the Algorithm

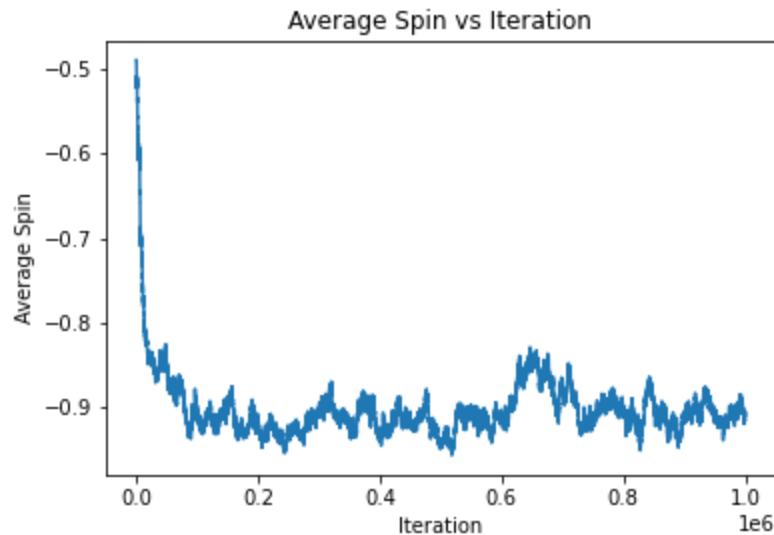
Before running the algorithm for a multitude of trials it was essential to reduce the time complexity of the energy calculations. Initially the simplest possible solution was applied, which was two nested for loops, providing a summation for each cell's identity value (-1 or 1) as well as the multiplication of two neighbors to avoid double counting. This evolved into a more parallelized version, where one would vectorize the loops by picking every other index of each value for columns and vectors respectively. You then effectively do the exact same matrix-multiply operations as before but with an "unrolled loop". The most efficient method that was implemented in time was that of the convolution (technically Fast Fourier Transform Convolution). The convolution is a commonly used operation in the regime of signal processing. The algorithm essentially utilizes a 3x3 defined matrix (or a kernel) for the neighbors that need to add together and multiply by the position of the center value [see Fig. 3]. This reduces the time complexity from  $O(n^2)$  to  $O\left(\left(\frac{n}{m}\right)^2\right)$  where m corresponds to the size of the kernel. One can extrapolate that with greater kernels sizes the efficiency of the algorithms scale to minimize computational time. Note that in the code, an alternatively function could have been `fftconvolve()` which utilizes the Fast Fourier Transform to perform mathematically equivalent operations, but in  $O(n \log n)$  time.



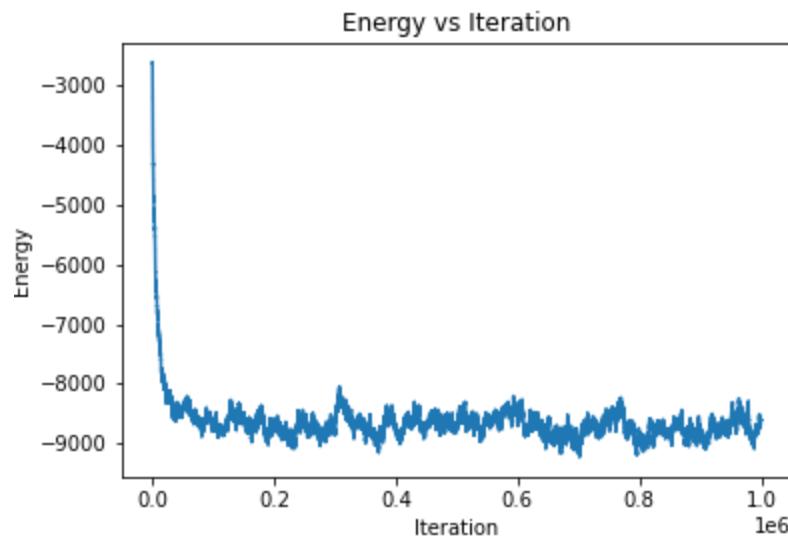
**Figure 3:** An example of a convolution filter for that was used for encoding. <sup>xiv</sup>



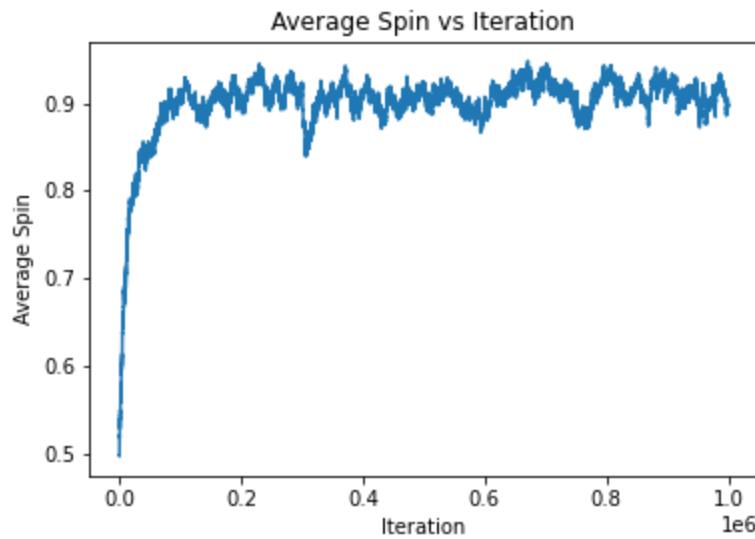
**Figure 4: Energy vs Iterations or Steps.** After few step the energy equilibrates around the arbitrary value of  $-9000$ . The plot was made under a preferential spin direction, with 25% of the spins aligning upward.



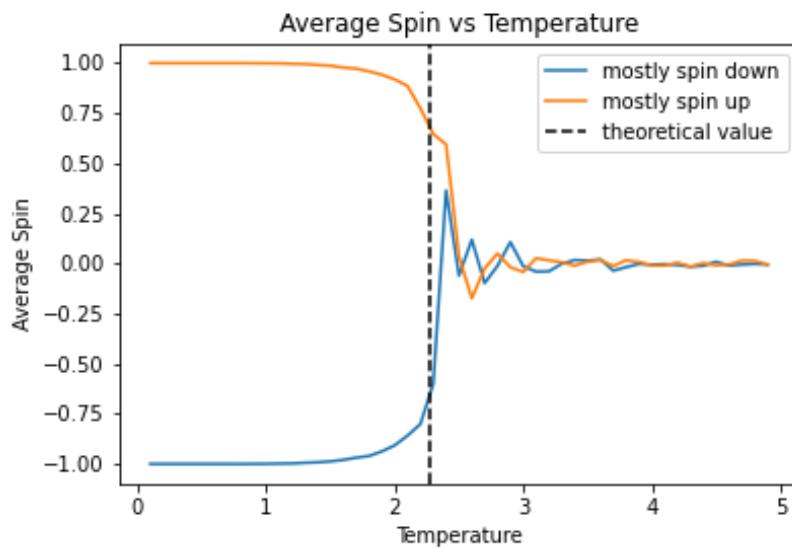
**Figure 5: Average Spin of the system vs Steps.** The plot follows a similar pattern to Figure 1, in which the average spin of the system deviates around a fixed range from a given value (-0.9).



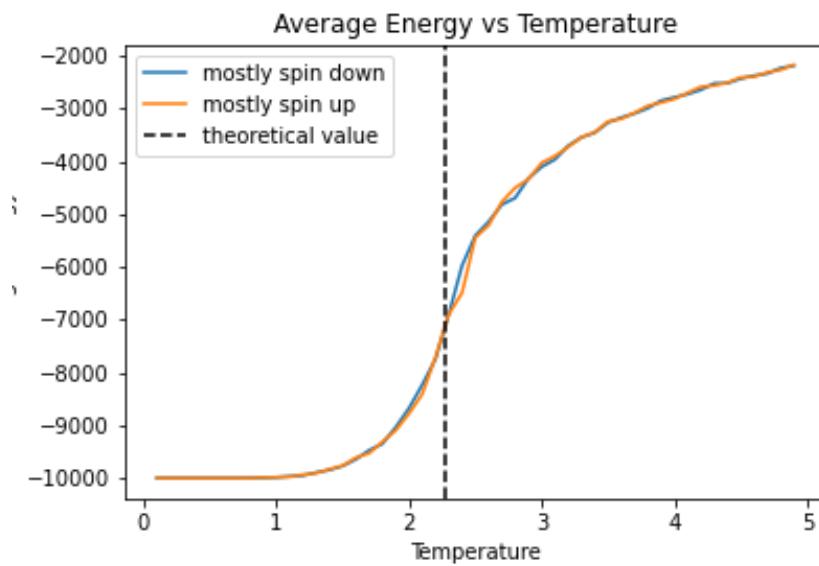
**Figure 6: Energy vs Iterations or Steps.** Effectively the same results being exhibited when compared to Figure 4. The plot was made under a preferential spin direction of 75% of the spins aligning upward.



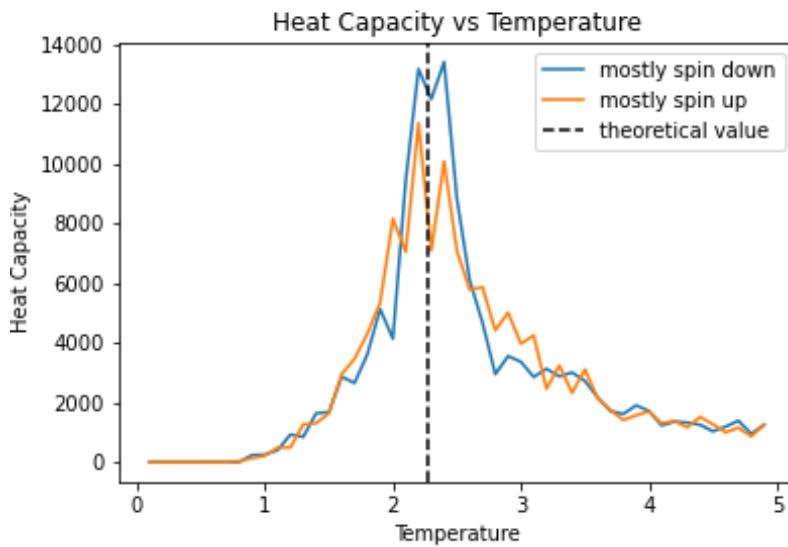
**Figure 7: Average Spin of the system vs Steps.** The symmetry around the average spin of 0.5, or truly random between this figure and Figure 5 indicates the aforementioned 75% spin upward differential.



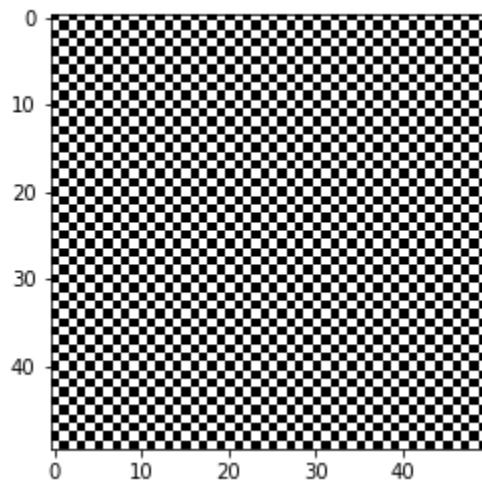
**Figure 8: Average Spin vs Temperature.** Illustrates slowly increasing/decreasing graphs (respective to spin up and spin down) that morph behavior suddenly at some dimensionless temperature  $\sim 2.5$ . Both plots later stabilize at 0, indicating that the magnetization of the given material has been completely lost. The theoretical value is corroborated within this graph as it is clearly seen that the respective spin graphs dramatically decrease after this value.



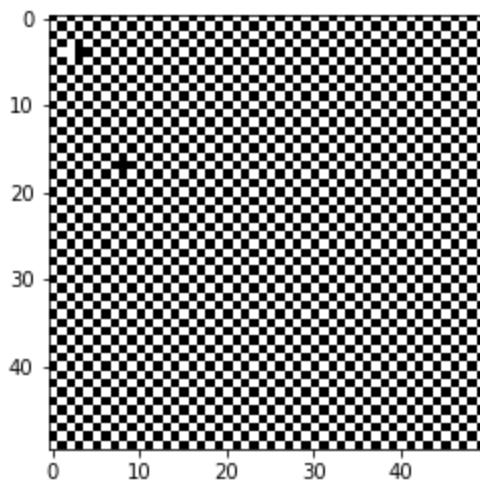
**Figure 9: Average Energy vs Temperature.** There is a seemingly logically regressive path being expressed, where the average energy of the system is meta-stable around this recurrent temperature  $\sim 2.5$ . This meta-stability implicates that the energy of the system will converge quickly on low or high extremes in temperature, or is unstable at some modeled Curie temperature. The analytically solved value represented by the dashed line designates approximately where the greatest partial derivative of energy with respect to temperature ( $\frac{dE}{dT}$ ) exists, although because this was calculated through finite means, it is not infinite.



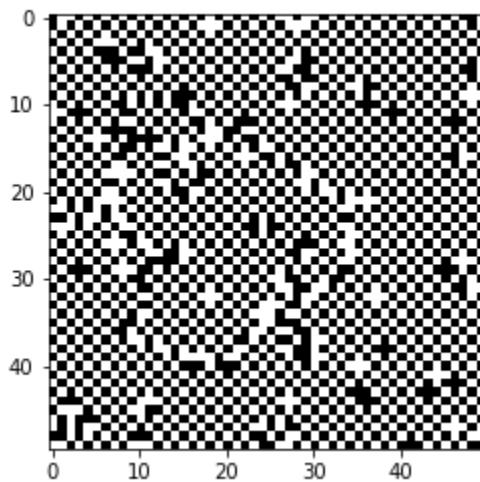
**Figure 10: Heat Capacity vs. Temperature.** This final graph is the most insightful into fulfilling the goal of the model. It contains highly similar plots of preferentially spin up and spin down systems converging at an identical, infinite peak. The plots correlate that when a phase transition occurs, or the magnetization drastically changes, the heat capacity nearly approaches infinity in its rapid growth rate on either side of the limit.



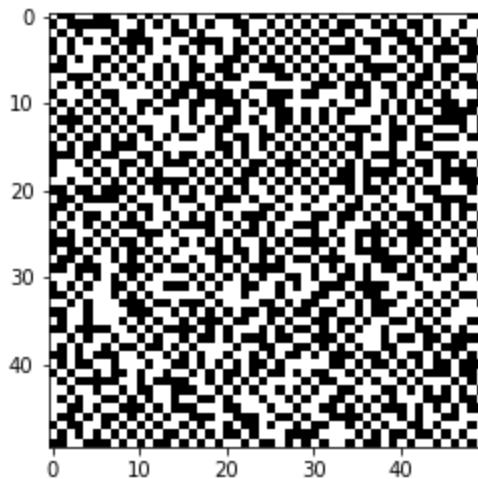
**Figure 11: Final lattice for antiferromagnetic trial with  $T = 0.1$  and  $J = -1$ .** The checkerboard pattern shows how adjacent spins tend to misalign in antiferromagnets, especially when temperature is very low.



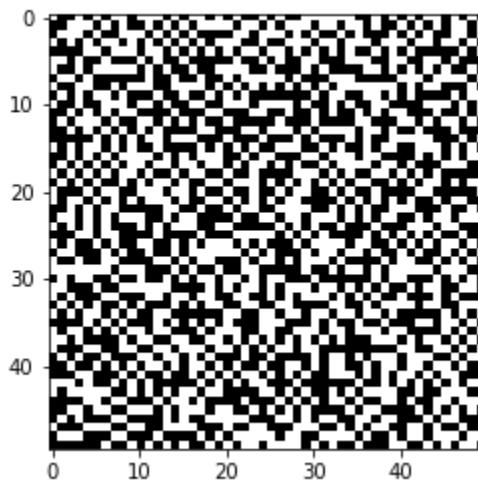
**Figure 12: Final lattice for antiferromagnetic trial with  $T = 1$  and  $J = -1$ .** There are a few impurities, but the lattice is almost entirely still checkerboard.



**Figure 13: Final lattice for antiferromagnetic trial for  $T = 2$  and  $J = -1$ .** The lattice is still almost entirely checkerboard, but there is more randomness throughout the lattice.



**Figure 14: Final lattice for antiferromagnetic trial with  $T = 3$  and  $J = -1$ .** The lattice is now mostly random, but there are a few pockets of checkerboard patterns throughout.



**Figure 15: Final lattice for antiferromagnetic trial with  $T = 4$  and  $J = -1$ .** The lattice is now almost entirely random, but there are still a few small areas of checkerboard patterns. This shows that as the temperature of the system increases, the randomness increases.

## 6. Final Results Summary for the Ising Model

Initially the model showed that the energy stabilizes at certain equilibria dependent on which orientation of spins they preferentially biased to converge to. This fact in tandem with similar behavior of the spins for each type iteration shows that this model is reasonably accurate in approximating some greater tendency of the system. Less vaguely, the system consistently deviates around the same average spins and energies. Later on, a rapid convergence of spin up and down graphs towards 0 were observed at some given temperature ( $T = 2.5$ ).

This is the first sign of a phase transition being simulated by this model, as it indicates that the material, when increasing temperature, rapidly lost its ability to magnetize, or align its spin with any given external magnetic field. The results are corroborated by the simple physically based phenomena of which ferrous metals lost virtually all magnetization ability when heated past their Curie (or Neel for antiferromagnetic materials) temperature.

Additionally, the Heat Capacity vs Temperature curve delineates a clear spike around this same temperature, looping this discussion back to the project's original assumptions. It was assumed that it was feasible to model a ferromagnetic lattice with a finite grid of randomly initiated spins, all according to a canonical ensemble-based heat exchange. Meaning more thoroughly that it theoretically takes an infinite heat capacity to perform a phase transition of an infinite grid of crystalline Fe atoms. However, since this simulation inherently contains a finite limit, a mere peak is shown rather than an idealistic asymptotic behavior.

Furthermore, in a qualitative analysis of the antiferromagnetism graphs, we can see that for comparative values in temperature the structures are much more stable. This corroborates what is expected by theory in that at low temperatures the interaction energy between neighboring spin states should dominate the self-magnetized alignment of the entire material. An interesting consequence of this property is that it allows for multiple equally minimal ground states.

### III. Results

#### A. Experiment 1 (Room Temp: 294.71 K)

##### 1. Magnetic Field (Copper wire)

To begin the tabletop experiment, a copper wire was used to establish a baseline of the magnetic field of the current itself through all samples. These values would then be used to calculate the final magnetic susceptibility of each material from this trial. The magnetism of the copper strip was found for the top and the bottom of the sample. After using the equation:

$$9.81 * (0.4313423653) / (1000 * (.9/100)) = 0.4701631782 \text{ T}$$

It was found that the magnetism of the bottom was 4701.728839 G  $\pm$  1111.509202 and the top was -0.000006874389334 G. This value for the top is such a minuscule number, that it was deemed negligible and was not used for any further calculations.

**Table 3(A)**

Sample	Change in Measured Mass (g)	Measured Magnetic Susceptibility (T)
Mohr's Salt	0.3438 $\pm$ 0.001	$2.70 \times 10^{-5} \pm 1.28 \times 10^{-5}$
Nickel Sulfate ( $\text{NiSO}_4$ )	0.1752 $\pm$ 0.001	$1.31 \times 10^{-5} \pm 4.96 \times 10^{-6}$
Manganese Carbonate ( $\text{MnCO}_3$ )	0.5036 $\pm$ 0.001	$8.36 \times 10^{-5} \pm 3.17 \times 10^{-5}$
Titanium (Ti)	0.1292 $\pm$ 0.001	$2.65 \times 10^{-6} \pm 1.01 \times 10^{-6}$
Copper (Cu)	-0.0142 $\pm$ 0.001	$-1.61 \times 10^{-7} \pm 6.13 \times 10^{-8}$
Aluminum (Al)	0.0064 $\pm$ 0.001	$-5.31 \times 10^{-7} \pm 9.16 \times 10^{-8}$

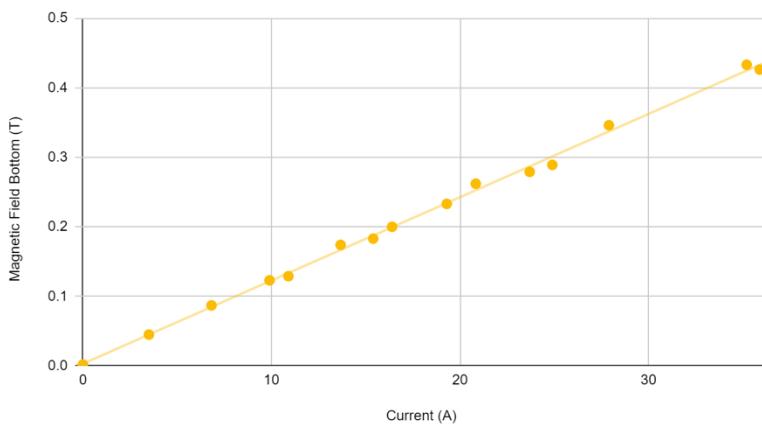
This data table shows the changes in mass due to the magnetic field being applied to the sample materials. The change in magnetic susceptibility shows how much the substances reacted to the applied magnetic field.

### Experiment 2 (Room Temp: 294.71 K)

Constants: Tube Length of Samples = 15.7 cm

#### 1. Base Measurements of Electromagnet Used:

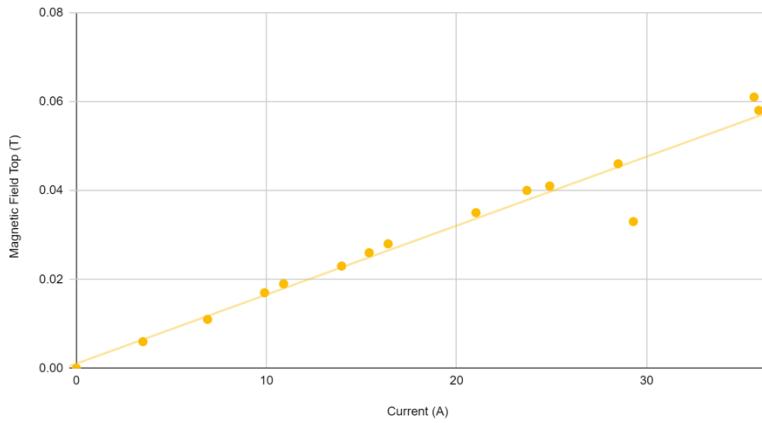
The Effect of Current on Magnetism (Bottom)



**Figure 16: Magnet Calibration on Axis of Poles**

Figure 16 shows the strength of the electromagnet (bottom) vs the applied current of the D.C PSU to the magnet for this section.

The Effect of Current on Magnetism (Top)



**Figure 17: Magnet Calibration at Top of Sample**

Figure 17 shows the strength of the electromagnet (top) vs the applied current of the D.C PSU to the magnet for this section.

#### 2. Silicon Dioxide ( $\text{SiO}_2$ ):

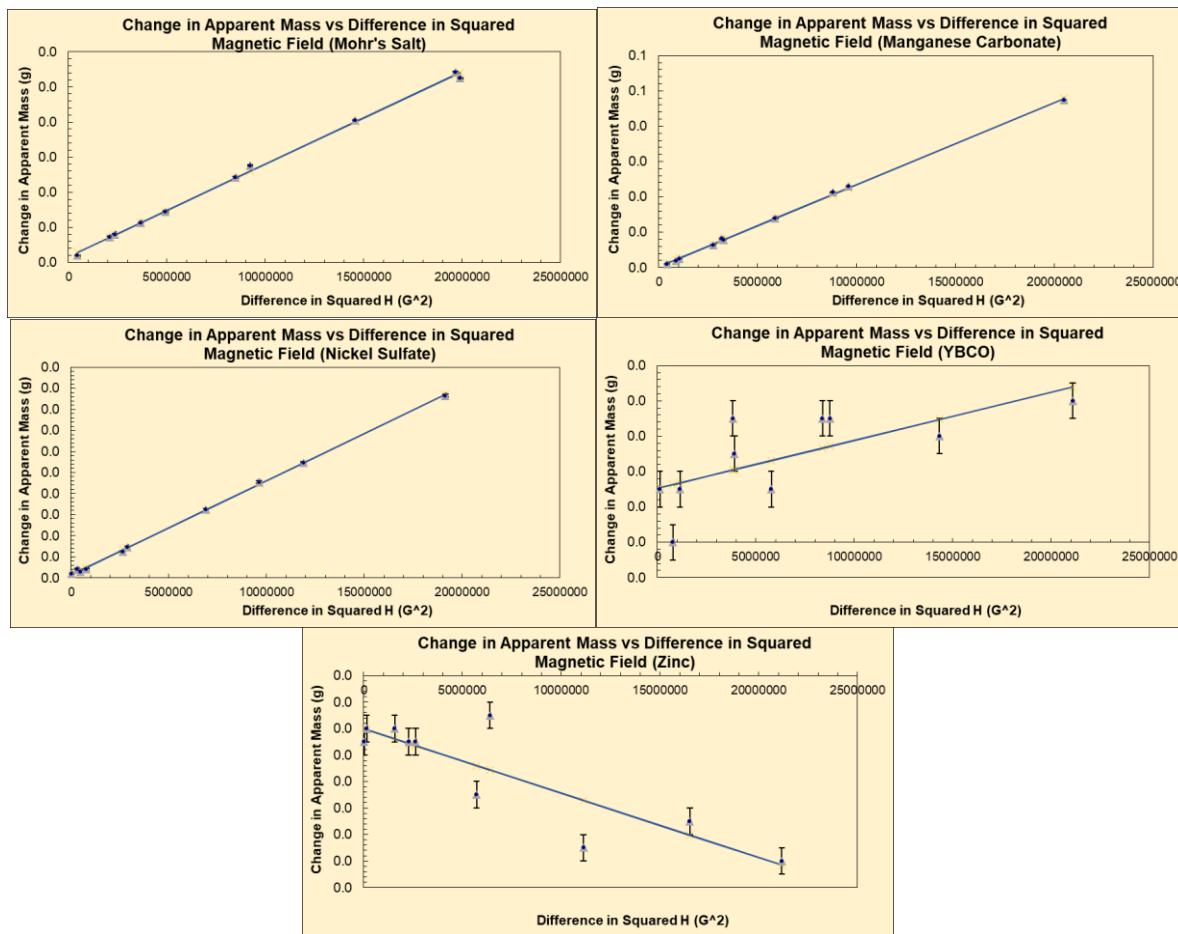
Silicon dioxide forms the glass encasings for all samples used within this experiment. It was then tested for magnetic properties to find any potential factors that could derange the data found within this experiment.

The average difference in the mass and the magnetic field strength were compared using a linear regression. The slope of that regression ( $y$ ) was used in the equation:

$$2 * 15.7 * 981 * (y) / 7.0627$$

The magnetic susceptibility ( $\chi$ ) of the silicon dioxide was then found to be  $-2.39 \times 10^{-7} \pm 4.20 \times 10^{-8}$ . As this number is so close to zero, it can be labeled as negligible in any further calculations involving the other samples.

### Data for All Samples at Room Temperature:



**Figure 18: Determination of Magnetic Susceptibilities at Room Temperature**

The apparent difference in the mass and the difference in squared  $H$  were compared using a linear regression. Then, the slope of that regression ( $y$ ) was used in the equation  $2*15.7*981*(y)/ms$ . According to these graphs, Mohr's salt is ferromagnetic, but because its value is close to zero, it may also function as a paramagnetic material. It can be understood that nickel sulfate functions the same way. Similarly, the manganese carbonate also has a very close proximity to zero; however, because its  $y$ -intercept is negative, this experiment shows that it is antiferromagnetic. This essentially means that these substances can function as paramagnets until they reach their respective critical temperatures and its behavior changes. Yttrium Barium Copper Oxide acts very ferromagnetically which indicates that within the temperature range

of the experiment it acts as a superconductor. Finally, Zinc has a negative slope, indicating that the substance is diamagnetic.

**Table 3(B)**

Sample	Measured Magnetic Susceptibility (T)
Mohr's Salt	$2.74 \times 10^{-5} \pm 4.85 \times 10^{-7}$
Manganese Carbonate ( $\text{MnCO}_3$ )	$8.17 \times 10^{-5} \pm 1.02 \times 10^{-6}$
Nickel Sulfate ( $\text{NiSO}_4$ )	$1.43 \times 10^{-5} \pm 1.60 \times 10^{-7}$
Zinc (Zn)	$-3.02 \times 10^{-7} \pm 6.17 \times 10^{-8}$
Yttrium Barium Copper Oxide ( $\text{YBa}_2\text{Cu}_3\text{O}_7$ )	$2.16 \times 10^{-7} \pm 7.62 \times 10^{-8}$

## A. Experiment 2 (Low Temperature: 77K)

Constants: Tube Length of Samples - 15.7 cm

### 1. Base Measurements of Electromagnet Used:

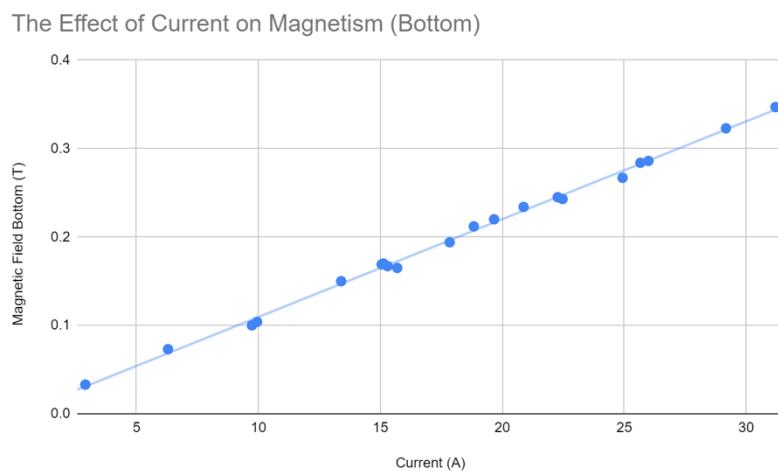
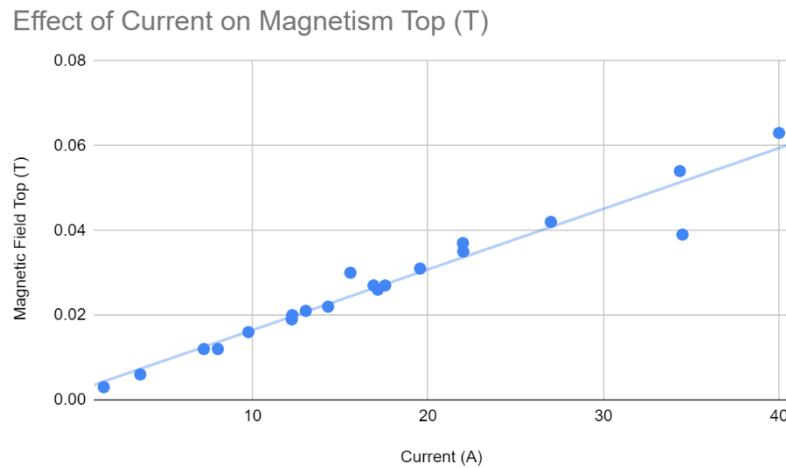
**Figure 19: Magnet Calibration on Axis of Poles**

Figure 19 shows the strength of the electromagnet (bottom) vs the applied current of the D.C PSU to the magnet for this section.



**Figure 20: Magnet Calibration at Top of Sample**

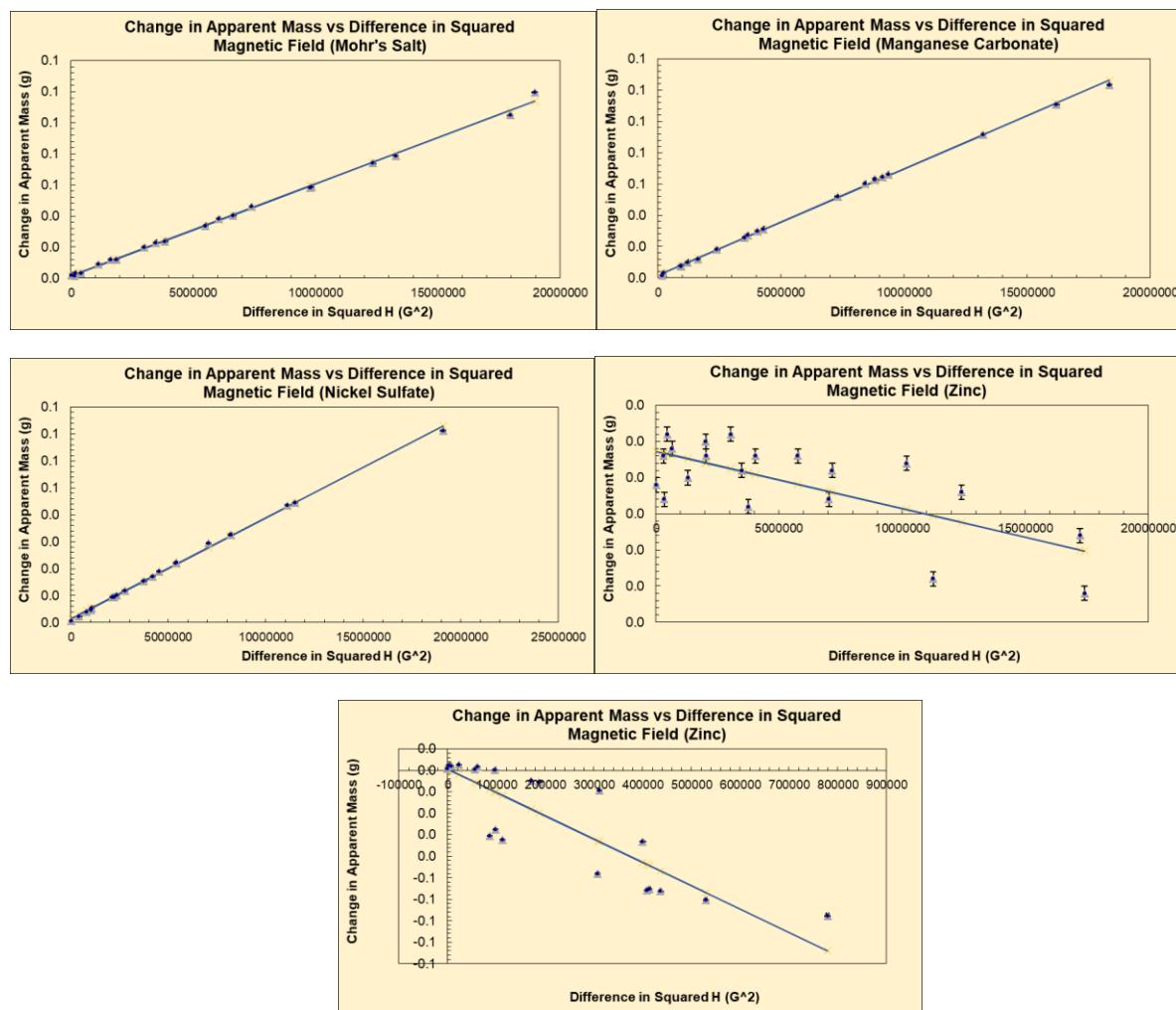
Figure 20 shows the strength of the electromagnet (top) vs the applied current of the D.C PSU to the magnet for this section.

## 2. Silicon Dioxide ( $\text{SiO}_2$ ):

Silicon dioxide forms the glass encasings for all samples used within this experiment. It was then tested for magnetic properties to find any potential factors that could derange the data found within this experiment. The average difference in the mass and the magnetic field strength were compared using a linear regression. The slope of that regression (y) was used in the equation:

$$2 * 15.7 * 981 * (y) / 7.0627$$

The magnetic susceptibility ( $\chi$ ) of the silicon dioxide was then found to be  $-4.56 \times 10^{-7} \pm 1.79 \times 10^{-7}$ . As this number is so close to zero, it is able to be labeled as negligible in any further calculations involving the other samples.

**Data for All Samples at Liquid Nitrogen Temperature:****Figure 21: Determination of Magnetic Susceptibilities at 77K**

The apparent difference in the mass and the difference in squared H were compared using a linear regression. Then, the slope of that regression (y) was used in the equation  $2 \times 15.7 \times 981 \times (y)/ms$ .

**Table 3(C)**

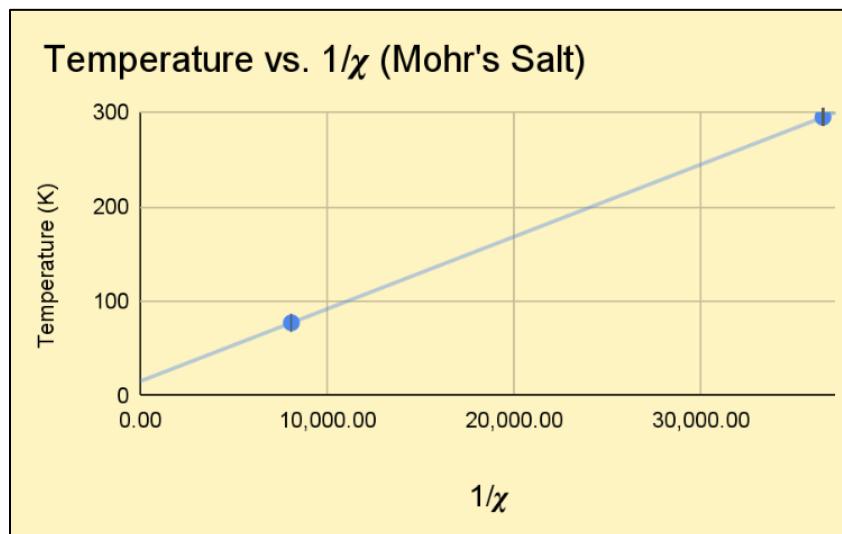
Sample	Measured Magnetic Susceptibility (T)
Mohr's Salt	$1.24 \times 10^{-4} \pm 1.65 \times 10^{-6}$
Manganese Carbonate ( $MnCO_3$ )	$2.38 \times 10^{-4} \pm 2.21 \times 10^{-6}$
Nickel Sulfate ( $NiSO_4$ )	$6.00 \times 10^{-5} \pm 6.61 \times 10^{-7}$
Zinc (Zn)	$-4.94 \times 10^{-7} \pm 1.12 \times 10^{-7}$
Yttrium Barium Copper Oxide ( $YBa_2Cu_3O_7$ )	$-8.51 \times 10^{-4} \pm 1.15 \times 10^{-4}$

This table shows the calculated magnetic susceptibilities of the samples at 77K.

## B. Final Results

**Table 3(D)**

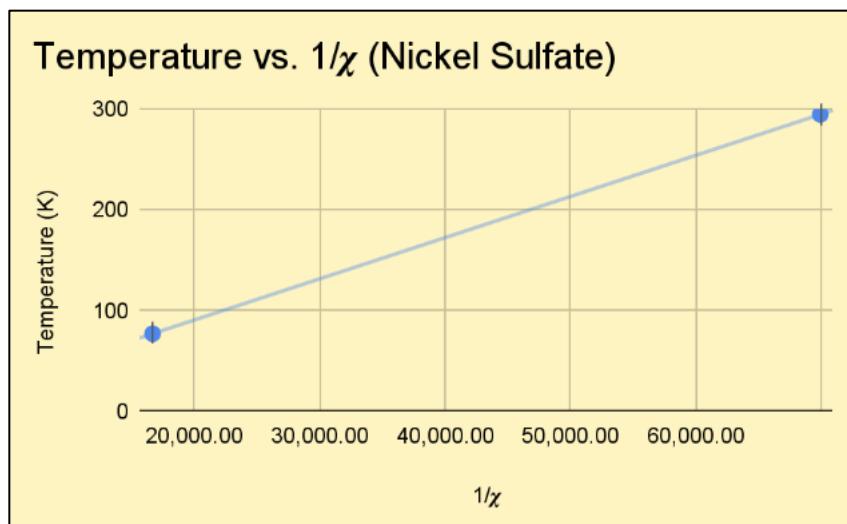
Average Magnetic Susceptibility per Unit Volume at Different Temperatures			
Sample	Tabletop Method (294.71 K)	Room Temperature (294.71 K)	Low Temperature (77 K)
Silicon Dioxide	N/A	$-2.39 \times 10^{-7} \pm 0.42 \times 10^{-7}$	$-4.56 \times 10^{-7} \pm 1.79 \times 10^{-7}$
Mohr's Salt	$2.7 \times 10^{-5} \pm 1.3 \times 10^{-5}$	$2.74 \times 10^{-5} \pm 5.57 \times 10^{-12}$	$1.240 \times 10^{-4} \pm 0.017 \times 10^{-4}$
Manganese Carbonate	$8.4 \times 10^{-5} \pm 3.2 \times 10^{-5}$	$8.16 \times 10^{-5} \pm 8.73 \times 10^{-5}$	$2.380 \times 10^{-4} \pm 0.022 \times 10^{-4}$
Nickel Sulfate	$1.31 \times 10^{-5} \pm 0.50 \times 10^{-5}$	$1.430 \times 10^{-5} \pm 0.016 \times 10^{-5}$	$6.00 \times 10^{-5} \pm 0.07 \times 10^{-5}$
Zinc	N/A	$-3.02 \times 10^{-7} \pm 0.62 \times 10^{-7}$	$-4.94 \times 10^{-7} \pm 1.12 \times 10^{-7}$
YBa <sub>2</sub> Cu <sub>3</sub> O <sub>7</sub>	N/A	$2.16 \times 10^{-7} \pm 0.75 \times 10^{-7}$	$-8.51 \times 10^{-4} \pm 1.15 \times 10^{-4}$
Titanium	$2.6 \times 10^{-6} \pm 1.0 \times 10^{-6}$	N/A	N/A
Copper	$-1.61 \times 10^{-7} \pm 0.61 \times 10^{-7}$	N/A	N/A
Aluminum	$-5.31 \times 10^{-7} \pm 0.92 \times 10^{-7}$	N/A	N/A



**Figure 22: Determination of Critical Temperature for Mohr's Salt**

Figure 22 shows that Mohr's Salt weakly exhibits the property of ferromagnetism, since the critical temperature is small but positive, according to the equation:

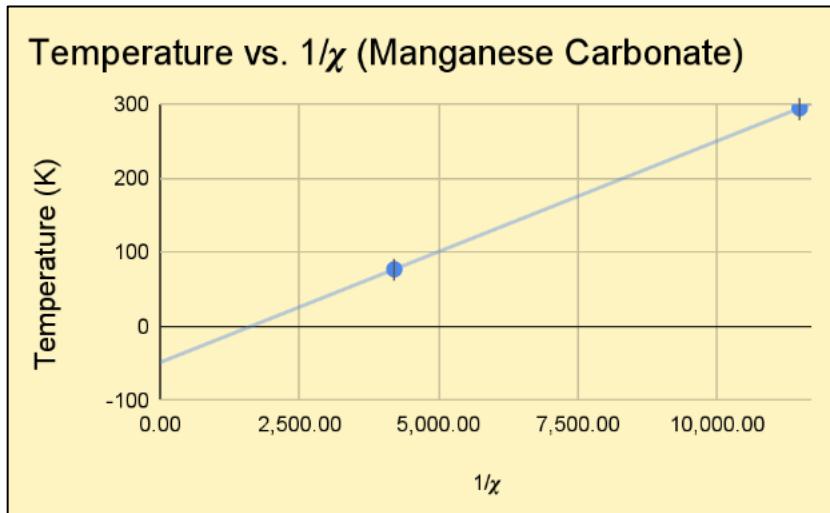
$$T = (c/\chi) + T_c$$



**Figure 23: Determination of Critical Temperature for Nickel Sulfate**

Figure 23 shows that nickel sulfate also exhibits the property of ferromagnetism, but considerably more strongly than Mohr's Salt (Fig.23). The y-intercept is clearly positive, according to the equation:

$$T = (c/\chi) + T_c$$



**Figure 24: Determination of Critical Temperature for Manganese Carbonate**

Figure 24 shows that manganese carbonate exhibits the property of antiferromagnetism, since the y-intercept is clearly negative.

## IV. Conclusion

In this series of experiments, it was shown that a decrease in temperature increased the respective magnetic properties of the various materials. At room temperature, the entropy of the system was higher because of the increased temperature and higher kinetic energy of the electrons. The higher entropy left the magnetic moments in disarray, and the net magnetism of the substance became weaker. It is then assumed that the inverse is true.

The data reflected this concept as many of the substances reacted to low temperature environments by becoming stronger in their own respective ways. The diamagnetic substance zinc began to express the property of repelling most magnetic fields as their magnetic susceptibilities decreased at a lower temperature than expected. Materials that exhibit ferromagnetic properties, such as Mohr's salt and nickel sulfate, increase in magnetic attraction or repulsion when exposed to an external magnetic field; they are further enhanced when exposed to low temperatures. A similar set of principles can be applied to paramagnetic substances on a much smaller scale. Their magnetic moments align when exposed to a magnetic field as well, and the decrease in entropy could then further align these moments. An antiferromagnet could then function like the two former types, but at a certain temperature, known as the Néel temperature. At this temperature its magnetic moments will align antiparallel and opposite to its neighboring moments, causing a net magnetic force of 0, which was true for manganese carbonate. Instead of an alignment to create a magnetic field, a diamagnetic material will create antiparallel alignments and repel an opposing magnetic field. This effect will become more evident as temperature decreases, which held true for silicon dioxide and zinc. Superconductors, like Yttrium Barium Copper Oxide, would create antiparallel alignments much like a more perfect iteration of a diamagnetic substance. The values recorded for the superconductor within this experiment then expressed this very principle by changing to an extremely negative magnetic susceptibility.

The results of this lab show that a decrease in temperature increased the magnetic susceptibilities of various materials. Every magnetic material is explained through the theory and statistics found within this set of experiments.

## V. Further research

Following this experiment, there are still many avenues of knowledge that can be researched. Using liquid helium instead of liquid nitrogen would provide a lower temperature for experimentation and a purer pool of results. This would provide more information about type I superconductors, and it would further isolate the magnetic susceptibility because the lower temperature allows for more dramatic readings.

Additionally, there were only two data points per material in Experiment 2 outside of taking multiple trials. Using different temperatures (a middle ground between room temperature and liquid nitrogen temperature) would further explore the relationship between temperature and magnetic susceptibility. Finally, testing antiferromagnetic materials and other diamagnetic, paramagnetic, and ferromagnetic substances would solidify the inverse relationship between magnetic susceptibility and the entropy of a system.

## VI. Acknowledgments

This research would not have been made possible without the help of Ms. Sophia Korono, Dr. Barry Luokkala, Mr. Dipto Mukherjee, and Ms. Ayda Ozdoganlar. Thank you to the Carnegie Mellon University Department of Physics, the PGSS Campaign, and PGSS faculty for granting access to the resources that facilitated this project.

## References

- <sup>1</sup>The Editors of Encyclopaedia Britannica . (n.d.). *Magnetism*. Encyclopædia Britannica. Retrieved July 18, 2022, from <https://www.britannica.com/science/magnetism>
- <sup>2</sup>Sandeman, K. (2004). *Ferromagnetic Materials*. Ferromagnetic materials (all content). Retrieved July 18, 2022, from <https://www.doitpoms.ac.uk/tplib/ferromagnetic/printall.php>
- <sup>3</sup>The Editors of Encyclopaedia Britannica. (2006, December 20). *paramagnetism*. Encyclopedia Britannica. Retrieved July 18, 2022 from <https://www.britannica.com/science/paramagnetism>
- <sup>4</sup>The Editors of Encyclopaedia Britannica. (n.d.). *Antiferromagnetism*. Encyclopædia Britannica. Retrieved July 18, 2022, from <https://www.britannica.com/science/antiferromagnetism>
- <sup>5</sup>Hilke, M. (n.d.). Quantum Nanoelectronics Laboratory (Michael Hilke). Retrieved July 21, 2022, from <https://www.physics.mcgill.ca/~hilke/>
- <sup>6</sup>Khan, H. R. (2003). *Néel temperature*. Néel Temperature - an overview | ScienceDirect Topics. Retrieved July 18, 2022, from <https://www.sciencedirect.com/topics/engineering/neel-temperature#:~:text=It%20is%20that%20temperature%20at,300%20to%20400%C2%B0C>.
- <sup>7</sup>Khan, S. K. (2011). *What is magnetic flux? (article)*. Khan Academy. Retrieved July 21, 2022, from <https://www.khanacademy.org/science/physics/magnetic-forces-and-magnetic-fields/magnetic-flux-faradays-law/a/what-is-magnetic-flux#:~:text=Magnetic%20flux%20is%20a%20measurement,something%20occupying%20a%20given%20area>
- <sup>8</sup>Pavarini, E. (2013). *Magnetism: Models and Mechanisms* (Vol. 3). Institute for Advanced Simulation. Retrieved July 18, 2022, from <https://www.cond-mat.de/events/correl13/manuscripts/pavarini.pdf>.
- <sup>9</sup>Wellmon, K. L. (2021, September 8). *Antiferromagnetism*. Engineering LibreTexts. Retrieved July 18, 2022, from [https://eng.libretexts.org/Bookshelves/Materials\\_Science/Supplemental\\_Modules\\_\(Materials\\_Science\)/Magnetic\\_Properties/Antiferromagnetism#:~:text=Antiferromagnets%20are%20typically%20ceramics%20of,field%20below%20a%20Néel%20temperature](https://eng.libretexts.org/Bookshelves/Materials_Science/Supplemental_Modules_(Materials_Science)/Magnetic_Properties/Antiferromagnetism#:~:text=Antiferromagnets%20are%20typically%20ceramics%20of,field%20below%20a%20Néel%20temperature).
- <sup>10</sup>*The Ising Model* | Lucas V. Schuermann. (n.d.). Retrieved July 29, 2022, from <https://lucasschuermann.com/writing/ising-model>
- <sup>10b</sup>Simon, M. (n.d.). *Diamagnetic Levitation*. Diamagnetic levitation. Retrieved July 18, 2022, from <https://www.physics.ucla.edu/marty/diamag/#:~:text=Diamagnets%20repel%2C%20and%20are%20repelled,oppose%20the%20external%20magnetic%20field>
- <sup>11</sup>*The Ising model*. (n.d.). Retrieved July 29, 2022, from <https://farside.ph.utexas.edu/teaching/329/lectures/node110.html>
- <sup>12</sup>Lipowski, A., Lipowska, D., & Ferreira, A. L. (2017). Phase transition and power-law coarsening in Ising-doped voter model. *Physical Review E*, 96(3), 032145. <https://doi.org/10.1103/PhysRevE.96.032145>
- <sup>13</sup>*2D Ising model*. (n.d.). Retrieved July 29, 2022, from <https://www.ibiblio.org/e-notes/Perc/ising.htm>
- <sup>14</sup>Huang, R., & Gujrati, P. D. (2019). Phase transitions of antiferromagnetic Ising spins on the zigzag surface of an asymmetrical Husimi lattice. *Royal Society Open Science*, 6(3), 181500. <https://doi.org/10.1098/rsos.181500>
- <sup>15</sup>*2D Ising model*. (n.d.). Retrieved July 29, 2022, from <https://www.ibiblio.org/e-notes/Perc/ising.htm>
- <sup>16</sup>Wentworth, C. D., & Wang, Y. L. (1993). The antiferromagnetic Ising model in a magnetic field: Linked-cluster expansion analysis. *Journal of Applied Physics*, 73(10), 5485–5487. <https://doi.org/10.1063/1.353674>

## VII. Appendix:

Ising Model Source Code:

```

import numpy as np
import random as r
import math
from scipy import ndimage

size = 50
J = 1
h = 0
T = 2
k_b = 1
beta = 1 / (k_b * T)

np.random.seed(1)
def generate_lattice(size):
    matrix_initial = np.random.rand(size, size)
    matrix_final = np.zeros((size, size))
    matrix_final[matrix_initial <= .25] = 1
    matrix_final[matrix_initial > .25] = -1
    return matrix_final

def energy_total(n, matrix, h, J):
    basic_kernel = ndimage.generate_binary_structure(2,1)
    basic_kernel[1][1] = False
    _matrix = -1 * J * matrix * ndimage.convolve(matrix, basic_kernel, mode='wrap')
    spin = -np.sum(matrix)
    H = np.sum(_matrix) + spin*h

    return H

def diff_energy_point(j,k,matrix,size):
    E = -((-1*matrix[j][k] * (matrix[(j+1)%size][k] + matrix[j][(k+1)%size])
e]
        + matrix[(j-1)%size][k] + matrix[j][(k-1)%size]))
        - (matrix[j][k] * (matrix[(j+1)%size][k] + matrix[j][(k+1)%size]
        + matrix[(j-1)%size][k] + matrix[j][(k-1)%size] )))
    return E

def metropolis(size, n, J, h, T):
    spins = []

```

```

energies = []
lattice = generate_lattice(size)
H = energy_total(size, lattice, 0, 1)
for i in range(n):
    x = np.random.randint(0, lattice.shape[1])
    y = np.random.randint(0, lattice.shape[0])
    lattice2 = np.matrix.copy(lattice)
    lattice2[y, x] *= -1
    diff_E = diff_energy_point(y, x, lattice, lattice.shape[1])

    if diff_E <= 0:
        lattice = np.matrix.copy(lattice2)
        H += 2*diff_E
    elif np.exp(-1/T * (diff_E)) > np.random.uniform():
        lattice = np.matrix.copy(lattice2)
        H += 2*diff_E

    energies.append(H)
    spins.append(np.mean(lattice))

return energies, spins

energy_avg = []
energy_std = []
spins_avg = []
spins_std = []

for t in temperatures:
    _energies, _spins = metropolis(size, 1000000, J, h, t)
    energies = np.array(_energies)
    spins = np.array(_spins)

    energy_avg.append(np.mean(energies[-200000:]))
    energy_std.append(np.std(energies[-200000:]))

    spins_avg.append(np.mean(spins[-200000:]))
    spins_std.append(np.std(spins[-200000:]))

np.save("energy_avg.npy", energy_avg_np)
np.save("energy_std", energy_std_np)
np.save("spins_avg", spins_avg_np)
np.save("spins_std", spins_std_np)

heat_capacity = energy_std**2 / (k_b * temperatures**2)

```

# **Behind the Beauty of the Carina Nebula: Star Formation, Ionization Fronts, and Interstellar Reddening**

Parimala Rajesh, Emma Wilson, Emily Zhang

## **Abstract**

Utilizing data from the Hubble Space Telescope, we studied the star-forming regions of the Carina Nebula and the O-type star system Eta Carinae to coordinate efforts to better understand the birth, life, and death of stars. Furthermore, we investigated the high amount of ultraviolet radiation in the Nebula and how it affects star formation and molecular composition. Through this survey and data analysis, we identified ionization fronts and the elements present in them. An exploration of dark pillars present in the Carina Nebula and how they are related to Eta Carinae was also conducted, which demonstrated which ionization fronts were in closest proximity to the star system. Reddening and extinction was concluded by dividing two images to show the ratio of H-alpha radiation to Paschen-beta radiation. Additionally, colored images were created by layering three grayscale images with different filters in order to provide us with the context needed to investigate these phenomena.

## **I. Introduction**

### **A. Introduction to the Carina Nebula**

The Carina Nebula, scientifically known as NGC 3372, is an immense, complex nebula located in the Carina-Sagittarius arm of the Milky Way galaxy. In our night sky, it can be seen from the southern hemisphere in the Carina constellation. The Carina Nebula was first discovered by astronomer Nicolas Louis de Lacaille in 1752, is around 7,500 light years away from the Earth,<sup>1</sup> and stretches across approximately 230 light years. It contains many bright, massive stars that are estimated to be 50 to 100 times the mass of our sun. At least two of these massive stars belong to the star system Eta Carinae, a very frequently studied stellar system due to its extreme luminosity and size. These stars emit large amounts of ultraviolet radiation, which contribute to the high rates of star formation in the Nebula. An image of the Carina Nebula is shown in Figure 1.



**Figure 1: Image of the Carina Nebula taken by the Hubble Space Telescope<sup>1</sup>**

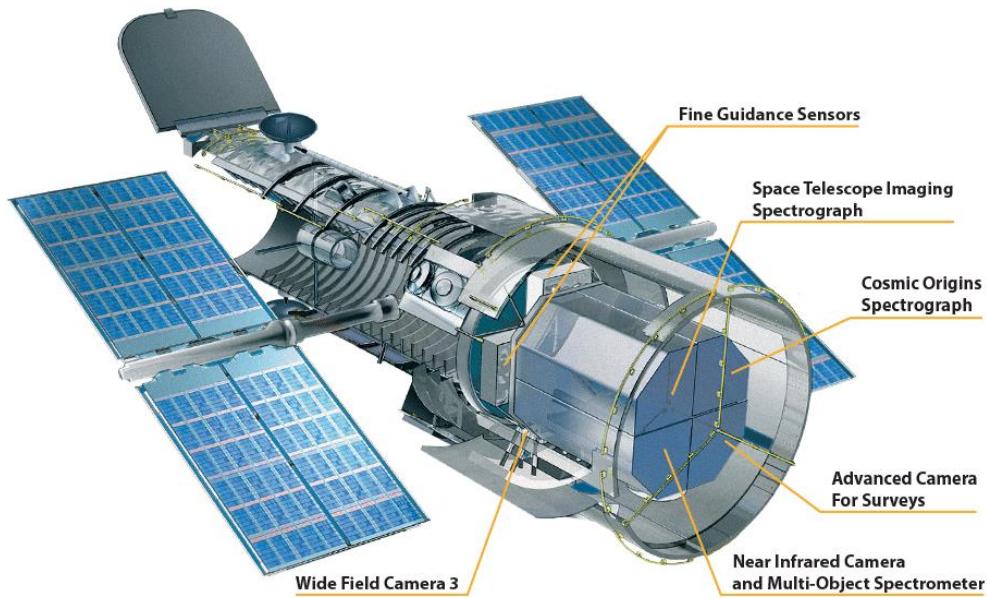
The birth of stars in the Carina Nebula began around three million years ago when the first newborn stars condensed and ignited in their parent clouds made of cold hydrogen and dust. The radiation from these ignited stars sculpted bubbles of hot gas and dust, creating the first star-forming regions of the Nebula. One of these star-forming regions is shown in Figure 2.



**Figure 2: Star Formation Region in the Carina Nebula taken by the James Webb Telescope<sup>2</sup>**

## B. Hubble Space Telescope

The Hubble Space Telescope (HST, Hubble) is a satellite that was deployed by the Space Shuttle Discovery in 1990. It observes and studies the universe from above the distortion and without the effects of the terrestrial environment. Over the years of its operation, there have been several new parts and instruments that have been mended or added. The cameras on the HST can detect wavelengths of light from the ultraviolet (UV) to the infrared (IR).<sup>3</sup>



**Figure 3: Hubble Space Telescope Select Instruments**

There are three types of instruments that analyze light on the HST: cameras, spectrographs, and interferometers, as shown in Figure 3. The two cameras are the Advanced Camera for Surveys (ACS) and the Wide Field Camera 3 (WFC3), which take images over a large range of wavelengths. The HST has two spectrographs, the Cosmic Origins Spectrograph (COS) and the Space Telescope Imaging Spectrograph (STIS). Spectroscopy is used for breaking the light from celestial objects into the component parts of the light. The spectrum can then be used to determine temperature, density, chemical composition, and velocity, among other characteristics of the object. There are three interferometers on the HST, called the Fine Guidance Sensors, which measure the position and brightness of stars. Past instruments that recorded data also contribute to this study and current studies over various wavelengths. These include the Wide Field Planetary Camera (WFPC) and the Wide Field Planetary Camera 2 (WFPC2).<sup>4</sup>

### C. James Webb Space Telescope

The James Webb Space Telescope (JWST, Webb) is a satellite telescope launched in December 2021, whose main mission is deep space infrared imaging. Some of its goals will be to see the origins of the universe, such as the first stars and galaxies, and compare them to newer stars and galaxies. The usage of infrared imaging will allow for clearer observations through dust.

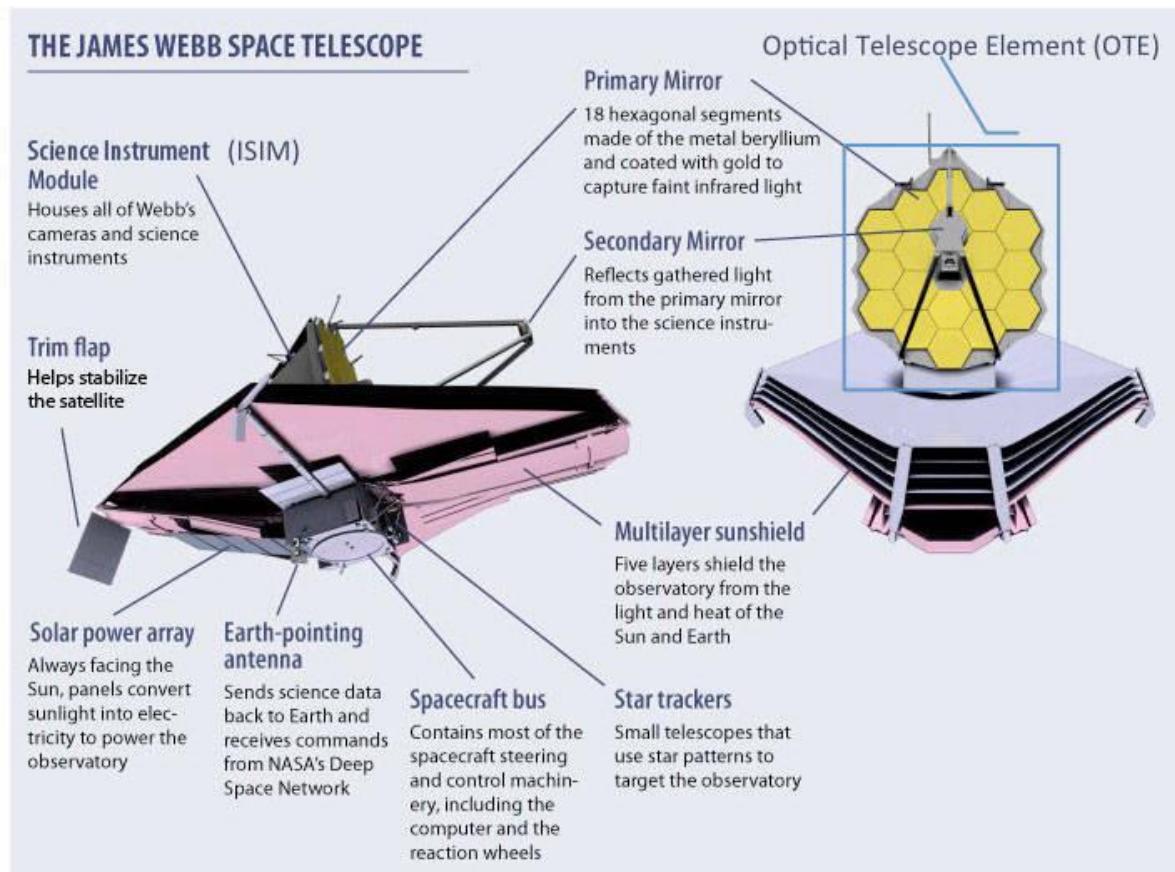


Figure 4: JWST Select Parts<sup>5</sup>

The Integrated Science Instrument Module (ISIM) contains the cameras and scientific instruments, as shown in Figure 4. These include the Near-Infrared Camera (NIRCam), Near-Infrared Spectrograph (NIRSpec), Mid-Infrared Instrument (MIRI), and Fine Guidance Sensor/Near Infrared Imager and Slitless Spectrograph (FGS/NIRISS). The NIRCam is the primary imager of JWST and is equipped with a coronagraph that allows for imaging of dim objects. The NIRSpec is a spectrograph that allows for the analysis of the physical properties of an object. The MIRI is a combination of a camera and spectrograph that allows for the detection of redshifted light of distant and faint objects. The FGS allows for precision of the camera and points the telescope. The NIRISS investigates first light and exoplanets<sup>5</sup>.

#### D. Filters

Both the Hubble Space Telescope<sup>6</sup> and the James Webb Space Telescope<sup>7</sup> have various filters for imaging that are classified by their wavelength and bandwidth, as shown in Figure 5 and Table 1. Each filter is named for the wavelength that it measures. However, the name of the filters is customarily a three-digit number, so the actual wavelength is often abbreviated. Each filter name starts with "F" denoting that it is a filter. The filter name ends with a letter that describes the bandwidth of the filter. The ones used in this study are "N" for narrow, "M" for medium, and "W" for wide.

Each filter also measures or detects something specific. Each narrowband filter measures the presence of a certain element from the wavelength of radiation that is emitted by the object or region and detected by the filter. Wideband filters are used for a variety of reasons, such as image creation and reducing background noise in narrowband measurements.

**Table 1: HST and JWST Select Filter Specifications**

Filter Name	Element/Usage	Throughput (%)	Wavelength	What mission?
<b>F502N</b>	[O III]	26	UVIS	HST
<b>F657N</b>	Wide H $\alpha$ + [N II]	25	UVIS	HST
<b>F658N</b>	[N II]	25	UVIS	HST
<b>F673N</b>	[S II]	25	UVIS	HST
<b>F775W</b>	Sloan Digital Sky Survey i'	23	UVIS	HST
<b>F090W</b>	General purpose	29	IR	JWST
<b>F110W</b>	YJ	56	IR	HST
<b>F126N</b>	[Fe II]	46	IR	HST
<b>F128N</b>	Paschen- $\beta$	50	IR	HST
<b>F164N</b>	[Fe II]	46	IR	HST
<b>F187N</b>	Paschen- $\alpha$	37	IR	JWST
<b>F200W</b>	General purpose	47	IR	JWST
<b>F335M</b>	PAH, CH4	42	IR	JWST
<b>F444W</b>	General purpose	47	IR	JWST
<b>F770W</b>	PAH, broadband imaging	37	IR	JWST
<b>F1130W</b>	PAH, broadband imaging	34	IR	JWST
<b>F1280W</b>	Broadband imaging	34	IR	JWST
<b>F1800W</b>	Silicate, broadband imaging	33	IR	JWST

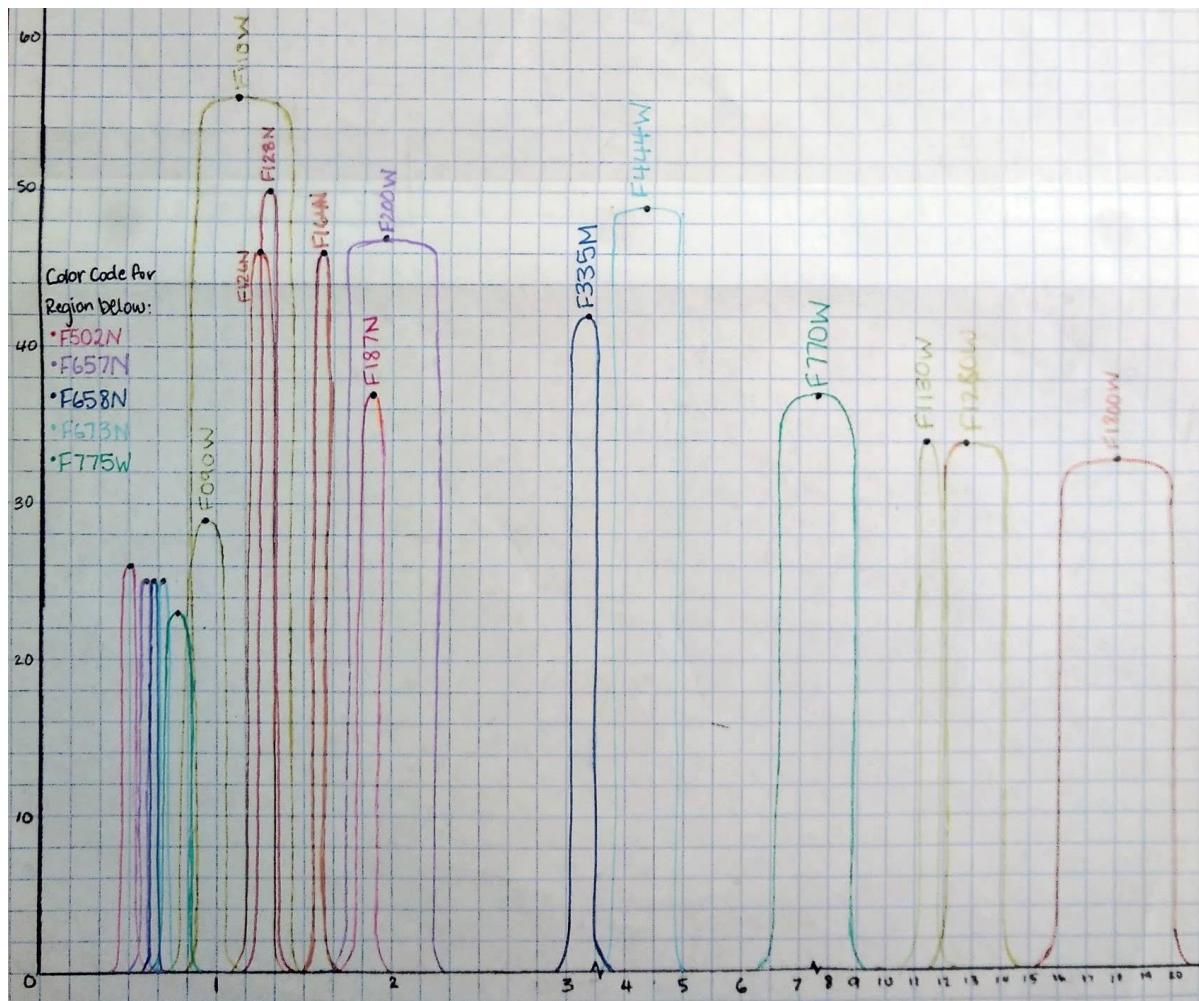


Figure 5: HST and JWST Select Filter Wavelength ( $\mu\text{m}$ ) and Throughputs (%)

## II. Background on Star-Forming Nebula

### A. Star Formation

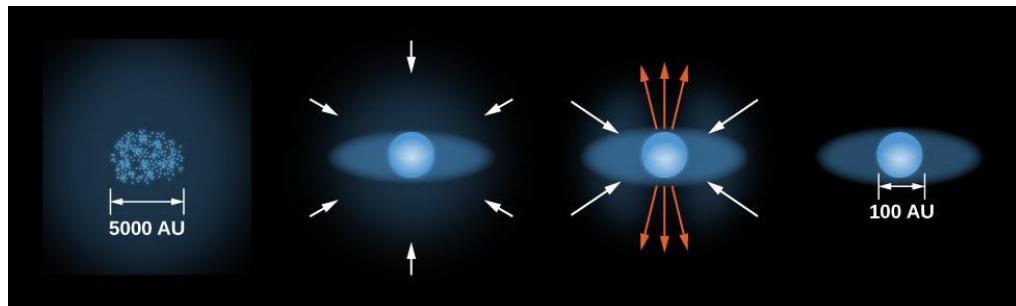
#### 1. Molecular Clouds and Stellar Cores

Star formation is initiated inside of parent molecular clouds. These clouds are extremely dense and cold, with temperatures ranging from 10-20 K. They are made primarily of hydrogen gas and dust, with the addition of other gasses such as oxygen and sulfur. However, hydrogen is the main building block of stars. Clumps, also known as Bok globules, reside inside of these clouds and contain cores which are the eventual birthplace of stars. All cores must satisfy the conditions of having low temperatures and high densities in order to successfully initiate star formation.

A star is subjected to two important forces: gravity and internal pressure. Gravity pulls the star inward on itself, whereas internal pressure—the result of interacting gas atoms within the star—causes the star to counteract gravity. These two interacting forces are what drive stars to become active within their respective molecular clouds. The low temperature of the core is directly correlated with a low internal pressure, so it aids gravity in rapidly collapsing core matter into a very dense, hot ball by a factor of  $10^{20}$  in size. This resulting body is known as a young stellar object (YSO), which initiates the formation of a true star<sup>8</sup>.

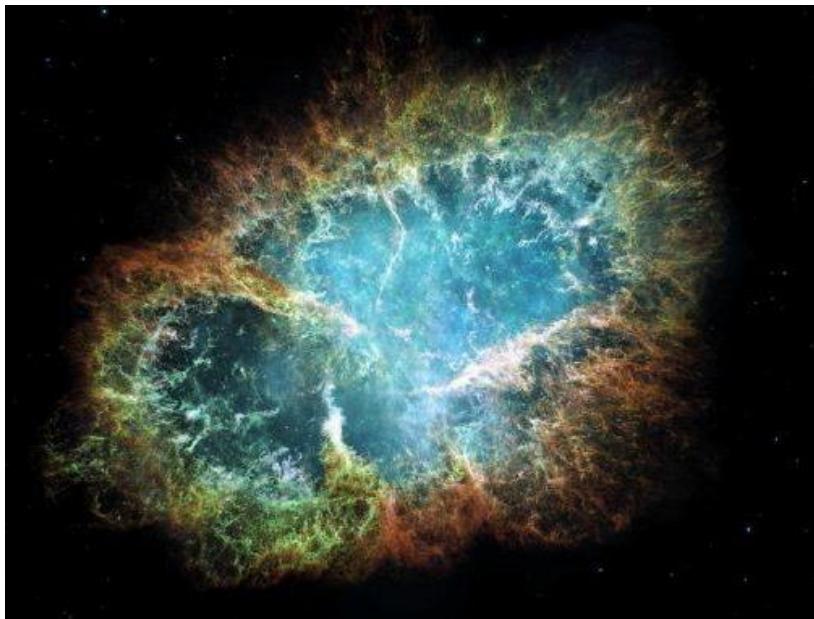
## 2. Birth and Death of a Star

Once a core has formed inside the molecular cloud, it begins contracting into a star. This stage of star formation is known as a protostar, which is a type of YSO. During this stage, protons have not yet started fusing together to produce helium. Protostars are characterized by having a dense core with material surrounding it. They also exhibit a spinning motion—as the star continues to contract, the radius decreases and it starts to spin faster, obeying the principle of conservation of angular momentum. Over time, collapsing material of dust and matter will spin into a disk around the protostar, supported by centrifugal force.<sup>8</sup> Because of the heavy amount of extinction during this stage, only infrared radiation can be detected. Extinction is the process of stars appearing dimmer or failing to appear at all in images of the star. This happens due to the dust that surrounds the protostar scattering and absorbing its light. Infrared radiation, though, is able to penetrate through the dust and is therefore detectable. Over tens of thousands of years, the gas particles in the protostar collide and the protostar eventually becomes hot enough to start fusing hydrogen into its core,<sup>9</sup> a process shown in Figure 6. The now true star will continue collapsing because of this and emit ultraviolet light, shedding its surrounding birth cloud along the way. Wind will also sweep away the excess matter, decreasing extinction and thereby allowing us to see the star on the visible light spectrum.



**Figure 6: Star Formation Progression from Core to Star**

Eventually, a star's core will run out of hydrogen and begin fusing its helium into carbon. Once that helium is exhausted, however, the star will fuse its remaining carbon into other heavier elements such as oxygen, neon, etc., eventually culminating in iron. Fusion halts at iron because it is so tightly bound by atoms that it cannot be fused into any other element without losing energy in the process.<sup>10</sup> An iron core then forms and starts to increase in temperature as it begins to collapse under gravity. The core becomes extremely dense as the electrons and protons inside merge to make it a neutron core. Finally, once the temperature becomes high enough, the core will explode and release large amounts of energy into space. This phenomenon is known as a supernova explosion, and it marks the death of an active star.<sup>11</sup> These explosions often leave behind supernova remnants, which contain interstellar material and copious amounts of energy, an example of which is shown in Figure 7.



**Figure 7: SN 1054 Supernova Remnant (Crab Nebula)<sup>12</sup>**

The remains of the star core will become either a neutron star, white dwarf, or a black hole, depending on the mass of the original star. If the original star is more massive, it will either become a neutron star or a black hole. Less massive stars like the Sun will turn into white dwarfs after they collapse.<sup>13</sup> Furthermore, the shockwave of energy emitted during the supernova explosion can initiate star formation in other molecular clouds. When massive stars—especially O-type stars—die, they emit large amounts of UV radiation and matter known as stellar winds. This radiation heats the surrounding gas, causing it to expand and promote star formation in molecular clouds.

### 3. Other Features: Jets, Pillars, and Star Clusters

When a young, very hot star (presumably O-type Star) has a ring of dust around it and a powerful magnetic field, jets are shot out of its poles in the form of energy when the field interacts with the dust and surrounding gas. These jets create a cloud-like formation around the star's poles and collide with other clumps, causing the emission of visible light. They are denoted as Herbig-Haro (HH) objects in the solar system. Eventually, the jets become photoionized by the star's ultraviolet radiation and are cooled, resulting in ionization fronts—areas of ionization of gas and dust—in the surrounding interstellar medium.<sup>14</sup> One of these jets is shown in Figure 8.



**Figure 8: Stellar Jet in the Carina Nebula taken with Hubble's WFC3 Detector<sup>15</sup>**

When O-type stars have strong enough winds that can sculpt their surrounding gas and dust, pillars of material form. Furthermore, the intense ultraviolet radiation from these stars heats up surrounding gas to form expanding bubbles within the pillar. The “peaks” of the pillars contain newly forming stars, while jets of material that shoot out from the peaks contain newborn stars inside.<sup>16</sup> These pillars also house many star clusters, a region where multiple stars are held together by mutual gravitational attraction.<sup>17</sup> One notable star cluster in the Carina Nebula is Trumpler 14, which contains some of the hottest and most massive stars in our galaxy. It is defined as an open star cluster, which contains stars in an asymmetrical arrangement. The stars in Trumpler 14 will send strong winds outwards into the Nebula, which also helps sculpt gas and dust into pillars. A famous pillar in the Carina Nebula, known as the “Mystic Mountain,” is shown in Figure 9.



**Figure 9: Collection of Pillars in the Carina Nebula named “Mystic Mountain<sup>18</sup>”**

## B. Classifying Stars

In astronomy, main-sequence stars are classified using the Hertzsprung-Russell diagram, a graph that plots stars' luminosity and color against their absolute magnitude, the observed brightness of a star. These types of stars are fueled by nuclear fusion where hydrogen turns into helium, and are in the most stable stage of their lifespans. Main-sequence stars encompass most of the stars in our galaxy, with the exception of giants, supergiants, white dwarfs, and neutron stars. Astronomers use the Hertzsprung-Russell diagram to determine a star's color, temperature, luminosity, spectral type, and evolutionary stage. Typically, the hotter the star, the more luminous (brighter) it is.<sup>19</sup>

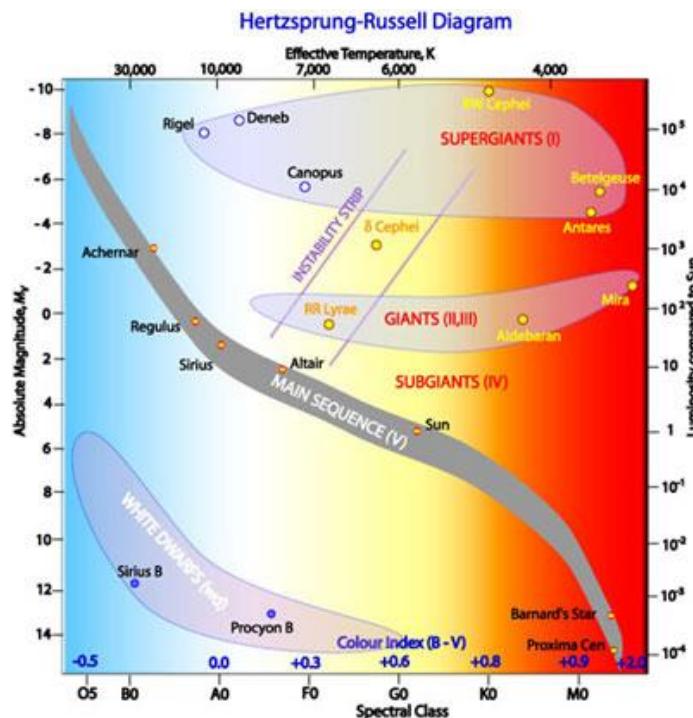


Figure 10: The Hertzsprung-Russell Diagram<sup>20</sup>

As seen in Figure 10, stellar luminosity is plotted on the vertical axis, and stellar surface temperature is plotted on the horizontal axis. The diagonal band that runs from the upper left corner to lower right corner is where main-sequence stars are located. Because high mass stars are subjected to very high gravitational forces, their core temperatures are much hotter. On top of that, their nuclear fusion rates are faster than that of lower mass stars. Therefore, most high mass stars are located near the upper left corner (high luminosity and temperature), and low mass stars are located near the lower right corner (low luminosity and temperature).

Main-sequence stars are classified in order of decreasing temperature: O, B, A, F, G, K, and M. O-type stars are the most massive, luminous, and have the highest surface temperature. The Carina Nebula contains many of these very massive but rare O-type stars, which heavily contribute to the high amounts of ultraviolet radiation in the Nebula. Table 2 shows the different star types and their characteristics.<sup>21</sup>

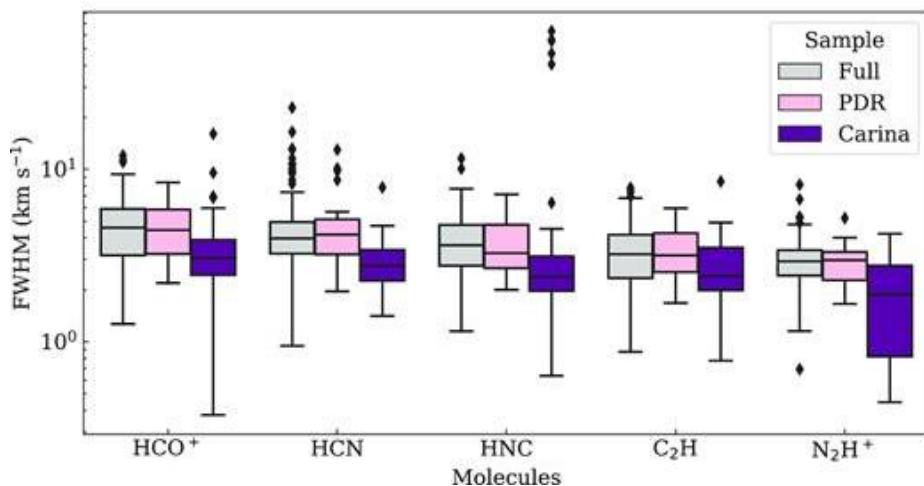
**Table 2: Spectral Classes and Star Characteristics**

Star Type	Color	Approximate Surface Temperature	Average Mass compared to The Sun (The Sun = 1)	Average Radius compared to The Sun (The Sun = 1)	Average Luminosity compared to The Sun (The Sun = 1)
O	Blue	>25,000 K	60	15	1,400,000
B	Blue	11,000 - 25,000 K	18	7	20,000
A	Blue	7,500 - 11,000 K	3.2	2.5	80
F	Blue/White	6,000 - 7,500 K	1.7	1.3	6
G	White/Yellow	5,000 - 6,000 K	1.1	1.1	1.2
K	Orange/Red	3,500 - 5,000 K	0.8	0.9	0.4
M	Red	<3,500 K	0.3	0.4	0.04

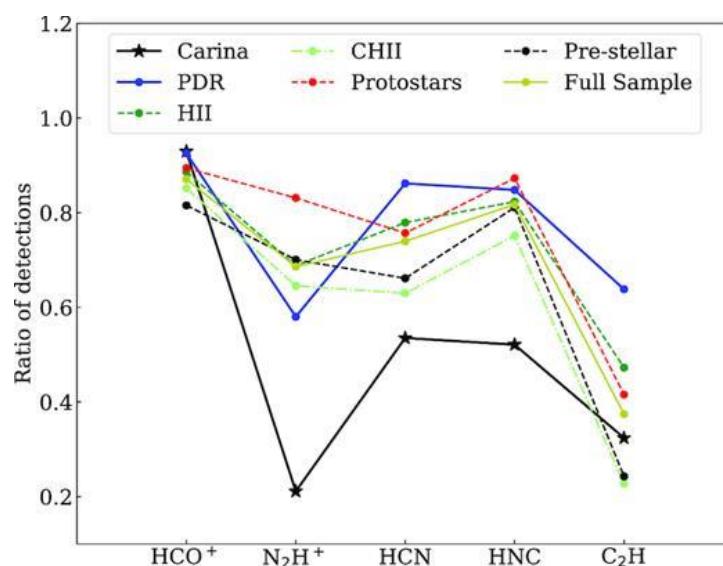
### C. Ultraviolet Radiation in the Carina Nebula and its Significance

Since being discovered in 1752, the Carina Nebula has long been a prominent research topic for scientists when investigating the process of star formation. Many of its stars have similar masses and characteristics to our very own sun, and the Nebula contains plenty of regions where stars are in varying stages of formation. With its relatively low extinction and surplus of hot, luminous stars, the Carina Nebula is also ideal for studying the effect that high-mass (O) stars have on their surrounding interstellar medium (ISM). O-type stars are characterized by being very hot and luminous due to the immense amounts of ultraviolet light they emit, thereby making them appear blue or white in the visible spectrum. This intense ultraviolet light emission is vital during the beginning of star formation.

The copious amount of ultraviolet radiation in the Carina Nebula makes star clumps have different chemical compositions and higher temperatures than others observed in the galactic plane. There are lower detection rates for common molecules in the Carina Nebula as compared to other star-forming nebulae, suggesting that ultraviolet emission is so intense that it has altered the chemistry of the Nebula's star clumps despite the Nebula being closer to Earth than the other regions studied.<sup>22</sup> In Figure 11 and 12, the ratio of detections for compounds HCO<sup>+</sup>, N<sub>2</sub>H<sup>+</sup>, HCN, HNC, and C<sub>2</sub>H are exhibited. The ratios for N<sub>2</sub>H<sup>+</sup>, HCN, and HNC are significantly lower compared to other regions and objects in the galactic plane, showing that ultraviolet radiation-induced ionization likely occurred in those compounds.



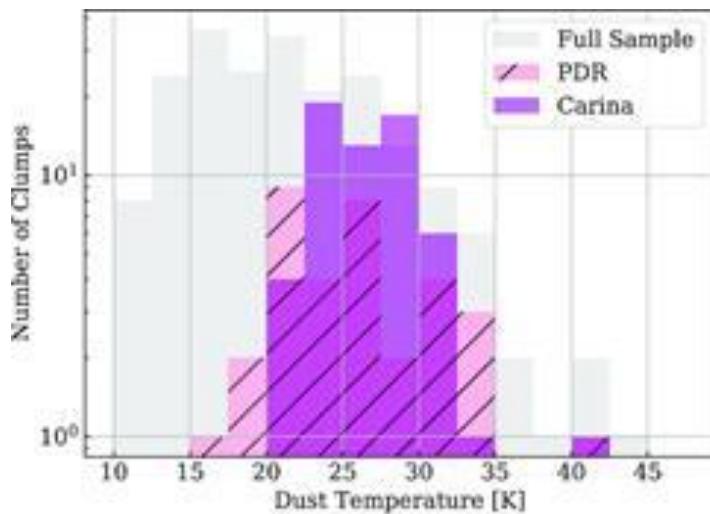
**Figure 11: Molecule Detections in the Carina Nebula vs. Other Galactic Regions (Box-and-Whisker Format)<sup>22</sup>**



**Figure 12: Ratio of Detections in the Carina Nebula vs. Other Galactic Regions (Line Graph Format)<sup>22</sup>**

These detection ratios were found using molecular line maps, which are astronomical maps that detect the absorption of certain elements and compounds in molecular clouds. The ratios were compared to the detection rates of the MALT90 (Millimeter Astronomy Legacy Team 90GHz) survey, a project that made 90Ghz molecular line maps for over 2,000 dense cores in the galactic plane.<sup>23</sup> Furthermore, the low detection of N<sub>2</sub>H<sup>+</sup> can also be explained by the extremely high temperatures of the gas surrounding massive O-type stars. Higher temperatures cause an increase in CO which destroys the N<sub>2</sub>H<sup>+</sup> gas; thus, as CO increases around these stars, N<sub>2</sub>H<sup>+</sup> decreases. These high temperatures also catalyze reactions that destroy the HNC compound.

Star clumps in the Carina Nebula have higher than average temperatures ( $26.6 \pm 0.1$ K) as compared to other MALT90 clumps ( $21.2 \pm 0.4$  K). However, the temperatures of these clumps do not vary far from that of photodissociation regions (PDRs) in the surrounding interstellar medium, likely because PDRs require high amounts of ultraviolet photons during the heating and dissociation process. This similarity is illustrated by Figure 13.

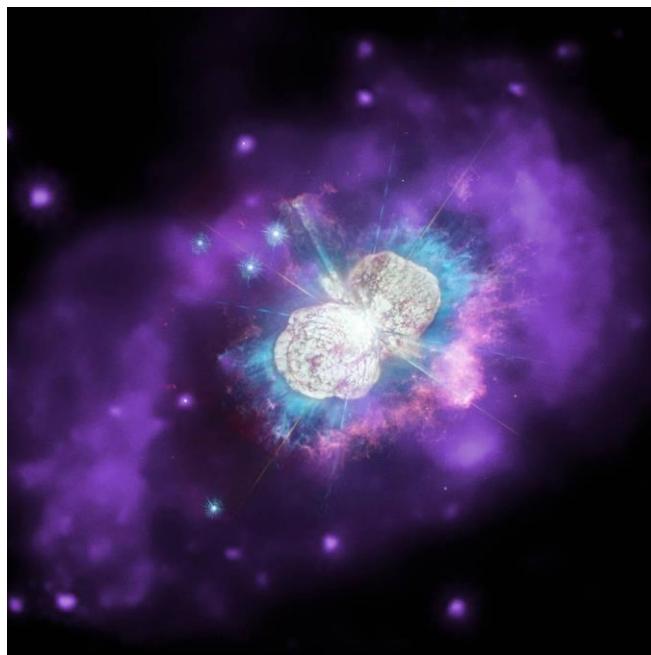


**Figure 13: Dust Temperature (K) in the Carina Nebula vs. Other Galactic and Photodissociation Regions**

Because of high ultraviolet radiation in the Carina Nebula, star clumps are chemically and physically different from many of those found in the galaxy.

#### D. Eta Carinae

Eta Carinae is a major stellar system consisting of at least two stars, Eta Carinae A and Eta Carinae B, located in the Carina Nebula. Both stars are associated with the open star cluster Trumpler 16, and they orbit each other approximately every 5.54 years. Additionally, Eta Carinae is an extremely luminous star system with a luminosity five million times greater than the sun, and is the only known star that emits ultraviolet laser emissions. Its immense ultraviolet radiation is very influential to star formation in the Carina Nebula as well. This makes Eta Carinae a very unique and important set of stars that is important in research about the life and death of massive stars. A dramatic image of Eta Carinae is shown in Figure 14.



**Figure 14: Eta Carinae taken by The Hubble Space Telescope<sup>24</sup>**

Despite being initially classified as a 4th magnitude star in the 1820s, Eta Carinae brightened to the 1st magnitude in the 1840s. This event is known as the “Great Eruption,” and it created as much visible light as a supernova explosion. During the Great Eruption, Eta Carinae exploded and created two polar lobes and a large, thin, equatorial disk. It is known as an imposter supernova because it failed to destroy itself during the explosion. Instead, it is now surrounded by a cloud of gas and dust known as the Homunculus Nebula which contributes to the variations in brightness that the star system now exhibits. For the next fifty years after its explosion, Eta Carinae began to fade. However, its visual magnitude has differed over the ages due to its unusual eruptions of brightness. These variations in brightness were first identified as having a 5.54 year period after 1996 when Eta Carinae had already gone through multiple brightening-fading cycles, but this was refuted after it broke the pattern in 2014.

Eta Carinae will explode as a supernova at any given time between now and one million years in the future. Once it finally destroys itself, it will likely collapse into a black hole.<sup>25</sup>

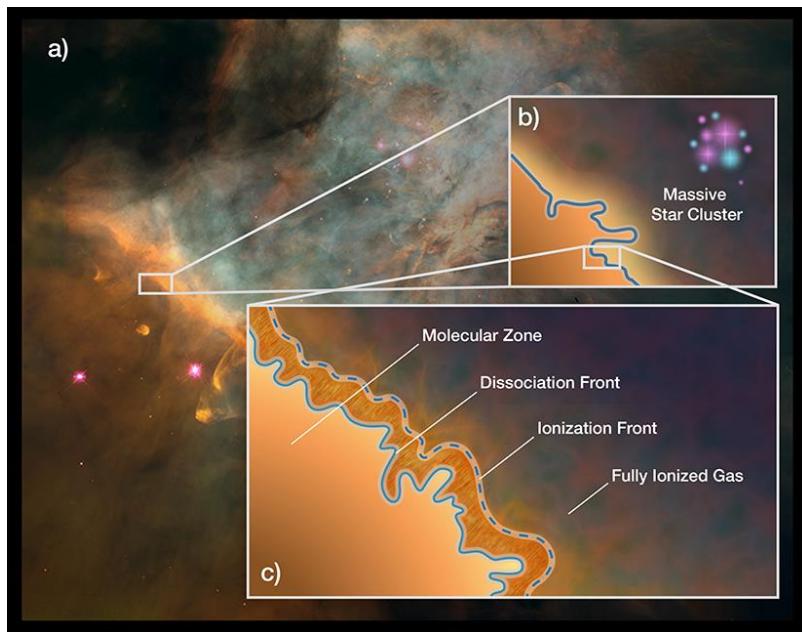
### **III. Ionization Fronts**

#### **A. Background Information**

##### **1. Photodissociation Regions**

Photodissociation regions (PDR) are areas where the heat source is from electromagnetic waves, often ultraviolet photons from surrounding hot stars or clusters, that strongly influence the gas chemistry of the region.<sup>26</sup>

The region is characterized by the unique structure of ionization shown in Figure 15. The surrounding stars, possibly from a star cluster, provide the ionizing UV radiation. The PDR essentially moves outwards as gas that is closer to the star is ionized. The leading side of the region is the dissociation front and the trailing side is the ionization front. Beyond the dissociation front is the molecular zone of un-ionized gas.



**Figure 15: Parts of a Photodissociation Region<sup>27</sup>**

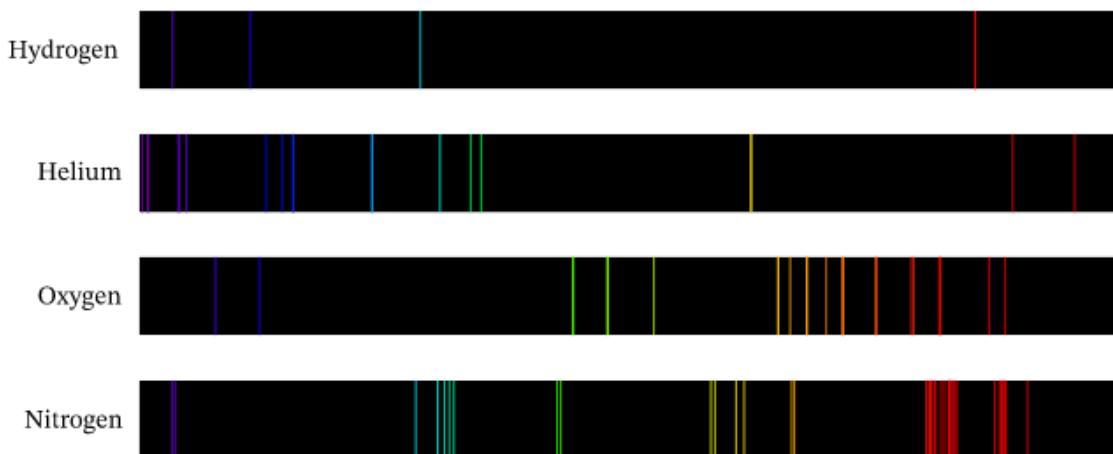
The main processes in the region are the chemical reactions that are caused by the increased amount of energy and radiation in the region. This includes both reformation and dissociation reactions, typically in dense clumps.

When hydrogen atoms chemically adsorb to graphitic and silicate surfaces, they are called physisorbed hydrogen, and as the surfaces get hot due to the ionizing radiation, the hydrogen atoms evaporate.<sup>28</sup> At this point impinging hydrogen in the surrounding interstellar medium reacts with the chemisorbed hydrogen, which is the hydrogen held separate by chemical forces, from the evaporation. The formation of hydrogen molecules also aids in the formation of carbon monoxide and other trace compounds, such as HCN and HCO<sup>+</sup>.

Hydrogen dissociation is when the hydrogen molecules break down into atoms. This is a function of the effective temperature of the star, so the dissociation rate can be calculated based on star temperature and distance.<sup>29</sup> The amount of time that a hydrogen molecule keeps that form is dependent on the temperature.

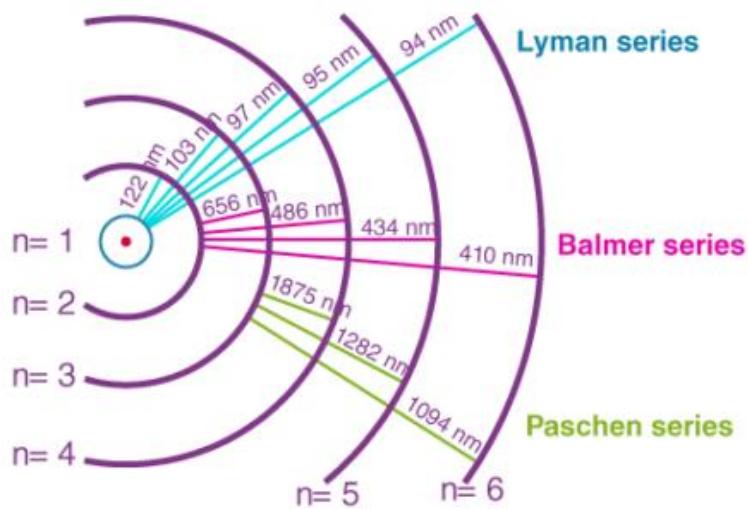
## 2. Emission Spectra

When elements are charged with energy from radiation such as UV photons or from collisions, whose energy is dependent on temperature, the electrons in the atoms increase their energy level. However, when they drop back down to lower energy levels, the energy is lost. This can be observed as the emission spectra, as shown in Figure 16, that is released by both elements and ions. Each element and ion has a unique spectrum that can be used to identify the elements and ions present.<sup>30</sup>



**Figure 16: Select Emission Spectra<sup>31</sup>**

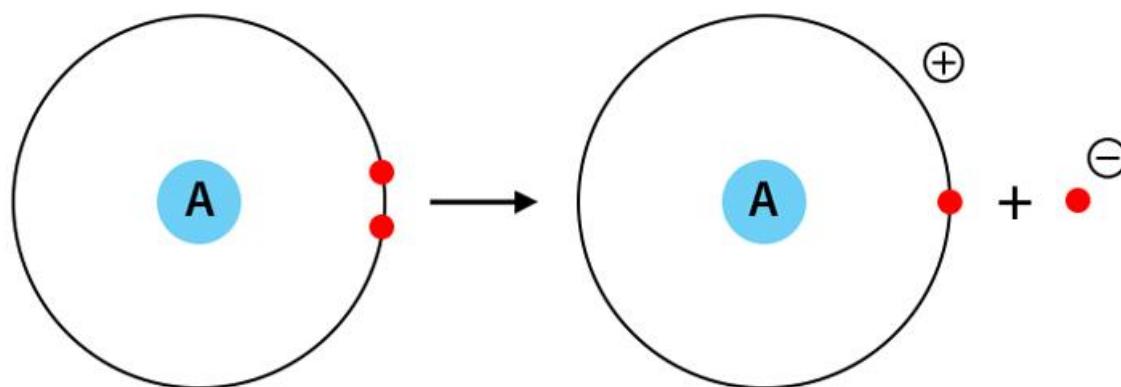
The example of this jump and drop in energy, referred to as a spectral series, analyzed in this study is the Paschen series of hydrogen as shown in Figure 17. Paschen-beta ( $\text{Pa}\beta$ ) is when the electron goes from the fifth to third orbital and releases energy. The  $\text{Pa}\beta$  luminosity can be used to estimate the temperature of the nebula from various ionization ratios.<sup>32</sup>



**Figure 17: Spectral Series of Hydrogen<sup>33</sup>**

### 3. Ionization

Ionization is the process by which electrons are removed from an atom as shown in Figure 18. The amount of energy needed to ionize an element depends on the element itself: if it is an element that is to the top right of the periodic table, the ionization energy is higher. The ions also provide emission spectra that can be analyzed.



**Figure 18: Ionization of an Atom<sup>34</sup>**

Doubly ionized oxygen, denoted [O III], is created by the double ionization of an oxygen atom. It has a unique spectrum, as every element does, that allows for its identification. Ionized nitrogen, denoted [N II] often has a faint spectrum, but can contribute to the calculation of temperature of a nebula<sup>32</sup>. The ionization of iron, [Fe II], is often in high density zones and is most often prompted by electron collisions.<sup>35</sup>

#### 4. Ionization Fronts

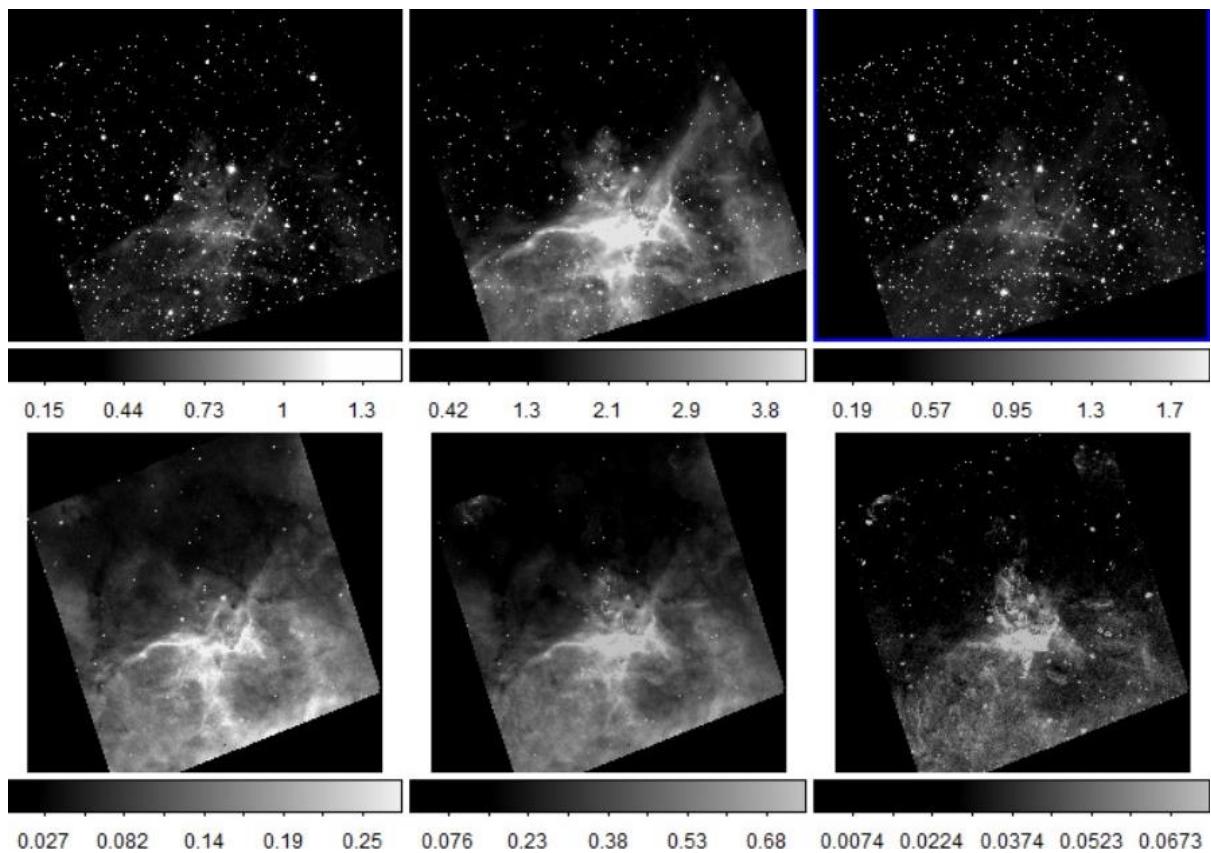
An ionization front exists where the radiation emitted from a star increases the energy of elemental particles in the area of space surrounding the star. When this happens, the electrons in an element increase their energy levels. However, when they lose some energy to their environment or to any other interaction, the energy is released in the form of observable electromagnetic radiation, picked up by detectors such as those on the HST. This observable radiation can be indicative of the locations where different elements ionize, giving information on the temperature and density of the specific areas.

Photoionization is a main source of this ionization, as the ionizing radiation from stars, such as UV radiation, increases the energy of the valence electrons, which are forced out of the atom and turn it into an ion. The energy is then taken with the electron as kinetic energy.

Ionization fronts form around hot stars that release energy in the form of electromagnetic radiation. O-type stars are a type of this hot star. They are characterized by their bluish white color and surface temperature of 25,000 to 50,000 K. They have lines of ionized helium in the spectra.<sup>36</sup> They are extreme population I objects, meaning they are thought to have originated from interstellar gas that has undergone various processes, so they contain iron, nickel, carbon, and various other heavy elements; though the majority of their constituents is still 90% hydrogen and 9% helium.<sup>37</sup>

#### B. Methods

Raw images of the section of the Nebula where the areas of interest are located were acquired from the Hubble Legacy Archive and the Barbara A. Mikulski Archive for Space Telescopes (MAST) databases. The analysis tools used were SAOImageDS9 (DS9) and data tabulation software, Google Sheets. For the identification and analysis of the ionization fronts, images from six different narrowband filters were used, F126N, F128N, F164N, F502N, F657N, and F673N, as shown in Figure 19. Each filter represents a different atom or ion: these represent [Fe II], Paβ (H), [Fe II], [O III], wide H-alpha (Hα) and [N II], and S II respectively. The first three are in the infrared section of the spectrum (IR), and the last three are visible and ultraviolet (UVIS).



**Figure 19:** The six images of the ionization fronts, ordered by wavelength<sup>38</sup>

DS9 was used to create analysis regions to determine the ionization location along the region, as shown in Figure 20. This allowed for a dataset to be extracted from the image.



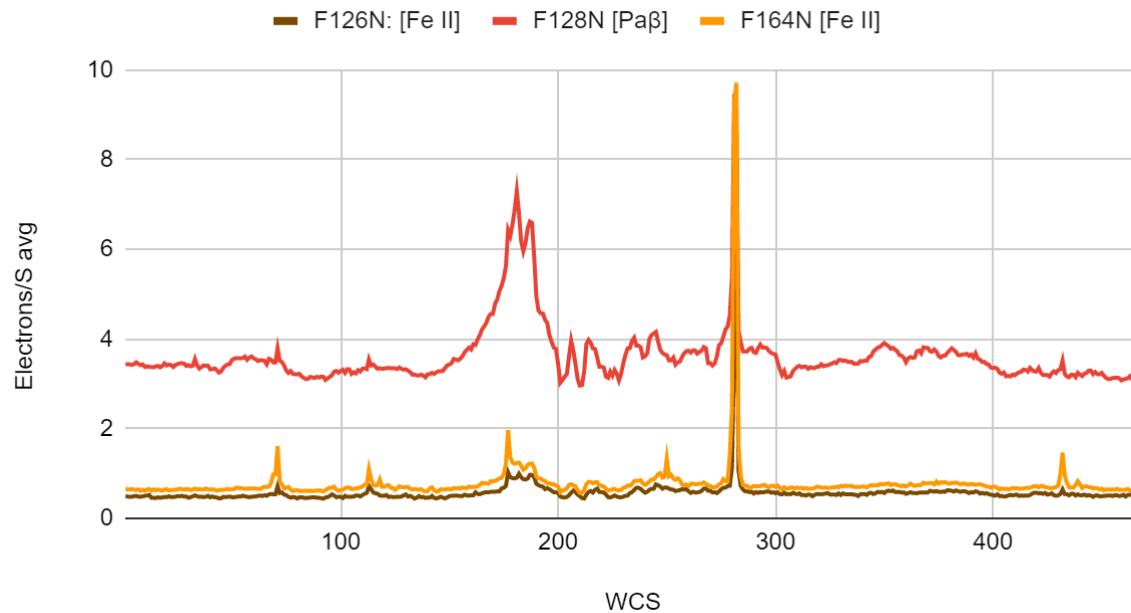
**Figure 20:** The region is shown in green, perpendicular to the fronts that are being analyzed<sup>38</sup>

Based on this data, graphs were made through Google Sheets. The peaks of the graphs were identified to find various ionization fronts of different elements. The difference in the peak location of the graphs of each filter represents the separation in the fronts for various elements, which was used to calculate the distance between the fronts in light years. In order to find this distance in light years, the World Coordinate System (WCS) embedded in the image was used to convert to arcseconds and then to degrees. Using the trigonometric function of tangent and the distance to the nebula, 7,500 light years, the distance was calculated in light years.

### C. Results

There were two sets of ionization fronts that were analyzed for separation. Both are indicated in Figure 24. Each ionization front was analyzed by six images from six different filters.

Three of the six images, F126N, F128N, and F164N, showed a peak for the two ionization fronts, representing the presence of their element, as shown in Figure 21. The other three filters showed no peak, representing the lack of presence of their element in that area.



**Figure 21: Ionization Front Peaks**

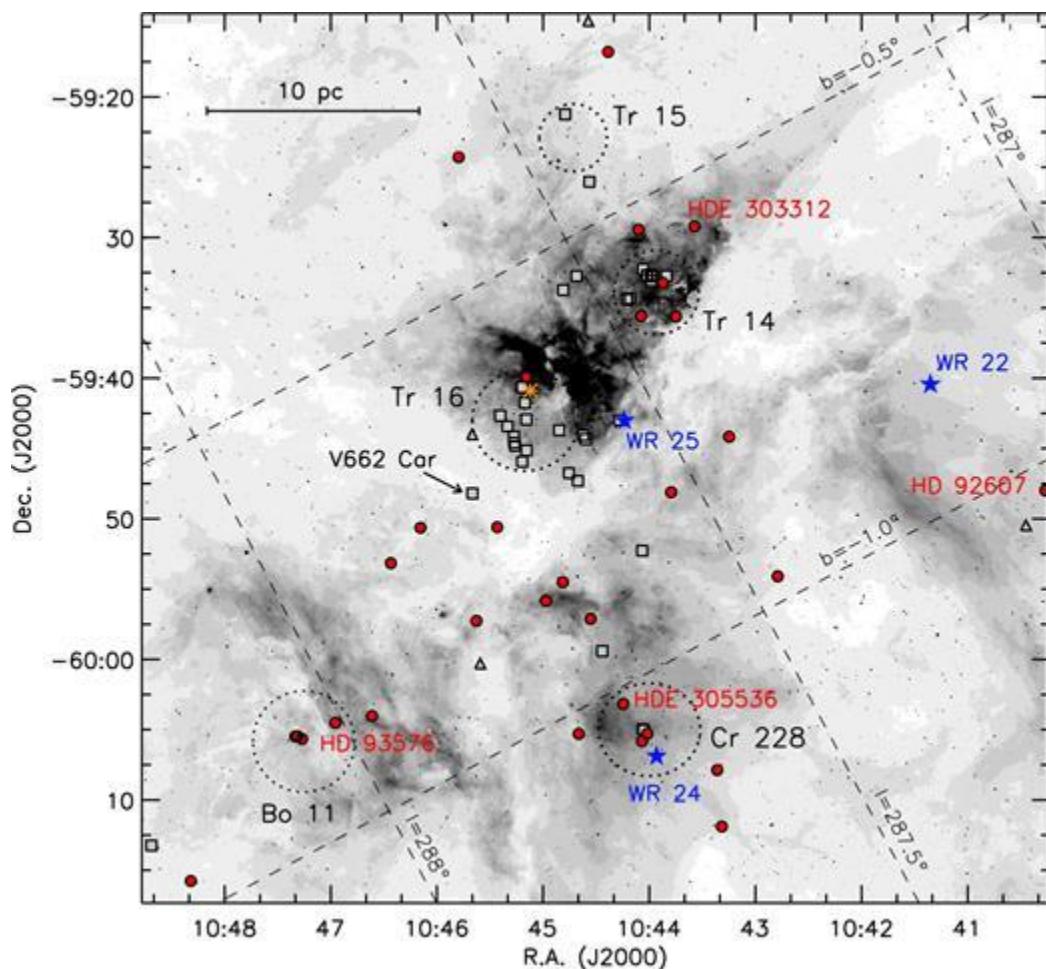
The lower ionization front is the right peak on Figure 21. It is characterized by a separation between the [Fe II] fronts and the Pa $\beta$  front, where the [Fe II] fronts are lower. The distance between the two fronts was calculated to be 0.27 light years.

Ionization temperatures are based on the ionization energy of the element, which is the amount of energy required to remove one electron from the atom. A higher ionization energy corresponds to a higher ionization temperature. Based on this principle, the [Fe II] fronts are cooler, and further from the star that facilitates the ionization, implying the presence of a star above the fronts.

The higher ionization front is the left peak on Figure 21. It is characterized by the same type of separation between the [Fe II] fronts and the Pa $\beta$  front, however the cooler [Fe II] fronts are toward the top of the image. Thus, the star that facilitates ionization should be below the fronts. The separation distance is 1.07 light years.

## D. Conclusions

There are O-type stars throughout the Nebula, and several of these stars emit energy in the form of UV radiation that causes ionization of various elements in the surrounding dust in the Nebula,<sup>39</sup> as shown in Figure 22. The orange star is Eta Carina and the area labeled “Tr 15” is the Trumpler 15 star cluster.

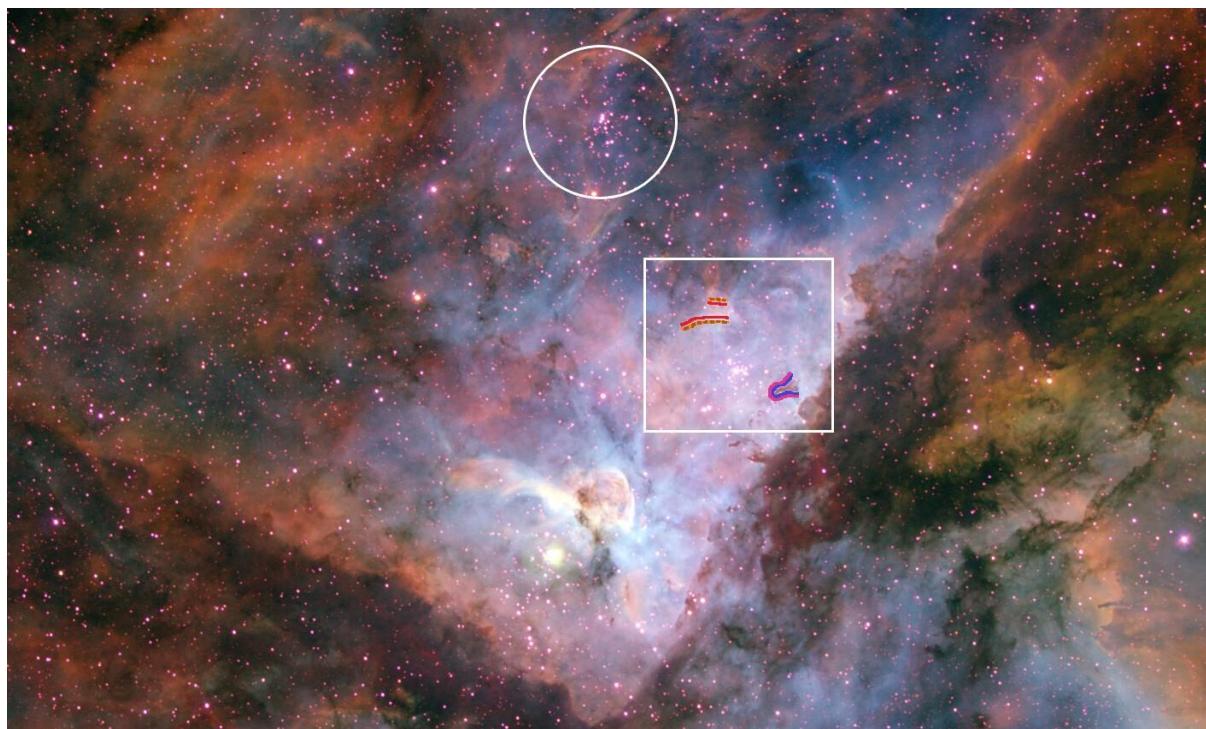


**Figure 22: A map of O-type Stars in the Carina Nebula<sup>39</sup>**

In Figure 23, the circled region is the Trumpler 15 star cluster. The ionization fronts are indicated by red, orange, and brown fronts, which represent the filters F128N, F126N, and F164N respectively. The bright white star in the lower center of the image is the Eta Carina star. The boxed area is expanded in Figure 24 to better represent the separation of the fronts.

The two fronts are in opposite orders, almost a mirror image of each other. This suggests that the two ionization fronts are likely caused by two different stars in two different planes of the image, as the image is a two-dimensional representation of the three-dimensional Nebula.

Some important features of the area of the Nebula surrounding the studied ionization fronts includes the star cluster above the ionization fronts, the dust surrounding the fronts, and the main stars of the Nebula, Eta Carinae, as shown in Figure 23.



**Figure 23: Carina Nebula<sup>40</sup>**



**Figure 24: Specific Study Region Expanded<sup>40</sup>**

The area of dust below the lower ionization front along with the order of the ionization fronts suggests that the source is from an area above the fronts. The Trumpler 15 star cluster above the ionization fronts could be the source for the radiation that ionizes these areas.

The opposite is true for the smaller, higher ionization front. The hotter, Pa $\beta$  front is toward Eta Carinae and the area of cool dust is toward the top of the image. Thus, the source of the ionizing radiation could be Eta Carina itself. However, the three-dimensional nature of the nebula must be considered when analyzing the locations of the fronts. As the radiation is from two different stars, the two fronts must be in different planes, but appear to be in line with each other in the two-dimensional perspective of this image.

The areas of dust surrounding the fronts in certain areas further suggest a similar conclusion. The lower ionization front has an area of dust below it, indicating the direction of the energy to be from above the location of the fronts. The higher ionization front has an area of dust above it, indicating the direction of the energy to be from below the location of the fronts.

## IV. Dark Pillars

### A. Background Information

Dark pillars are formed when a photodissociation region hits an area of dust or gas that is un-ionized. This initially forms a wall, where there is uniform ionization along the wall of dust and gas, however some areas are more easily ionized due to their composition and are ionized first, which leaves behind the dark pillars of un-ionized material,<sup>41</sup> as shown in Figure 25.

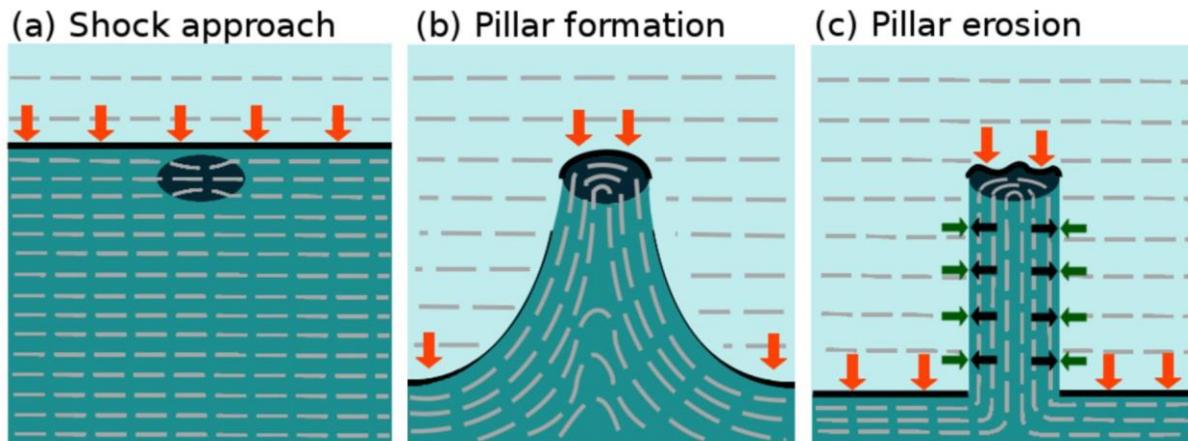
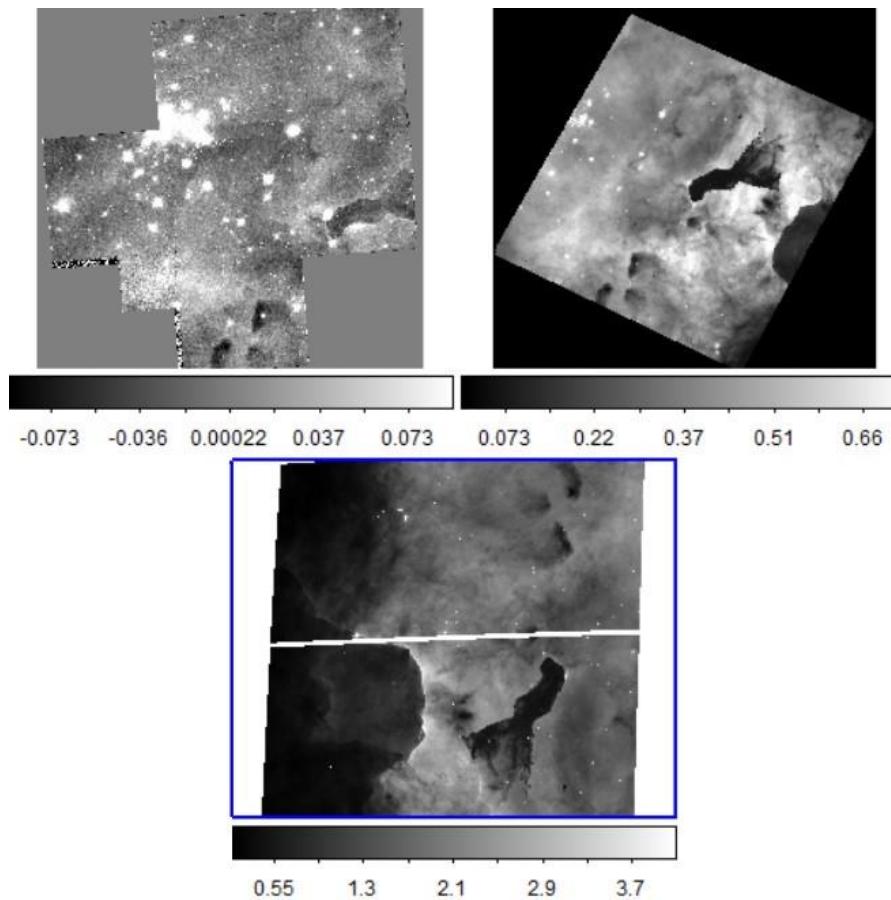


Figure 25: Formation of Pillars<sup>41</sup>

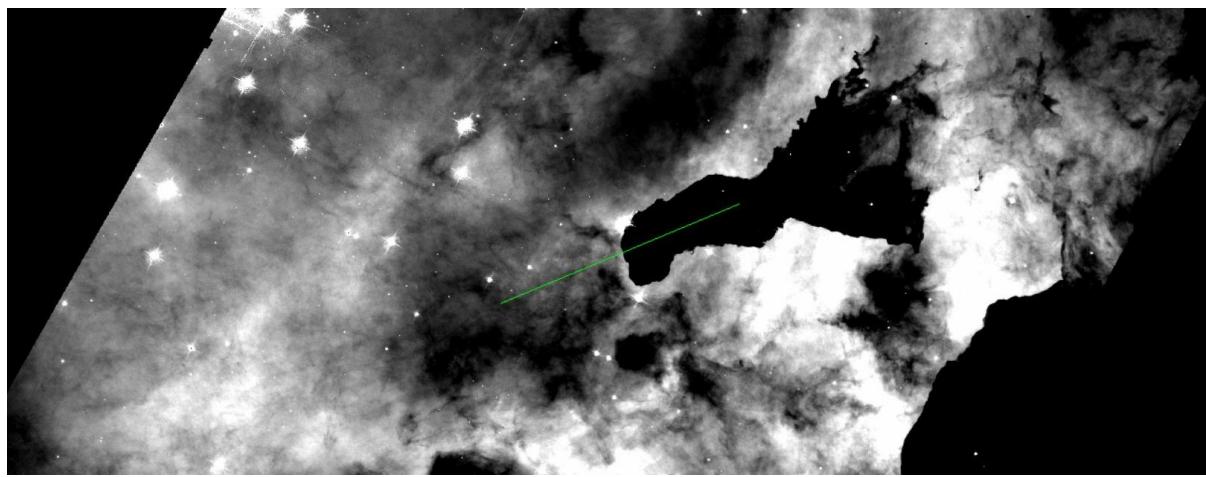
### B. Methods

A similar process was used for the analysis of the dark pillars as the ionization fronts, where three different narrowband filters, F502N, F658N, and F673N, were used, as shown in Figure 26.



**Figure 26: The three images of the dark pillars. Order: F673N, F502N, F658N<sup>38</sup>.**

All the filters were in the UVIS section of the electromagnetic spectrum. DS9 was used to create a region of analysis as shown in Figure 27.

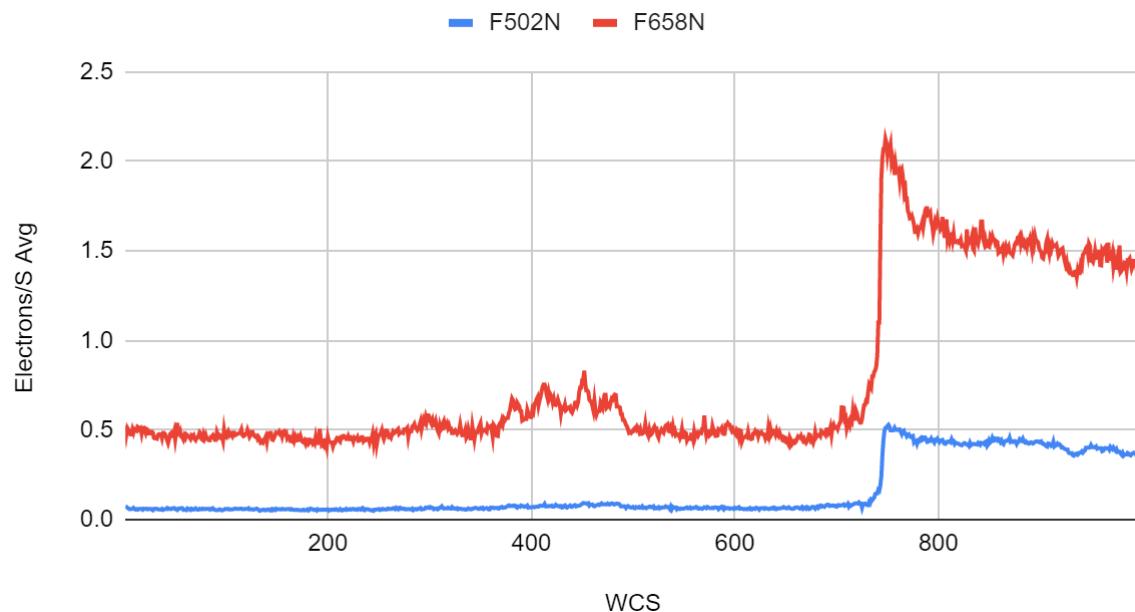


**Figure 27: The region is shown in green, perpendicular to the fronts that are being analyzed<sup>38</sup>.**

The data from the image was then analyzed for peaks of ionization. In the dark pillars, these represent the outer edge of the pillar, and the variation of the location of the peaks represents the separation of the fronts, calculated to light years.

## C. Results

From the three images used to analyze the dark pillar, two filters, F502N and F658N, representing [O III] and [N II] respectively, indicated presence of the element, as shown in Figure 28.



**Figure 28: Dark Pillars Peaks**

Between the two fronts, a separation of 0.80 light years was observed.

## D. Conclusions

In Figure 24, the ionization around the dark pillar is represented by the purple and blue fronts, which represent the filters F502N and F658N respectively.

Between the two fronts around the dark pillar, the [O III] ionization front is closer to the star than the [N II] front. This is explained by the higher ionization energy of oxygen due to being doubly ionized in contrast to nitrogen being once ionized. Thus, oxygen would need a higher temperature and pressure to ionize than nitrogen, so it ionizes closer to the star.

## V. Interstellar Reddening

### A. Background Information

The Carina Nebula consists primarily of gas, dust, and stars. When observing the Nebula, this gas and dust is in between Earth and the stars. This has a variety of effects on how the Nebula is perceived. The dust scatters and absorbs light emitted from the stars which alters how well the stars can be seen and what the stars look like. These processes are called reddening and extinction. Extinction occurs when the light

emitted from the stars is absorbed and scattered enough that the star appears dim or disappears entirely in our image. Reddening is the process by which stars appear more red because the dust is scattering the blue light produced by the star, causing the stars to appear more red than they actually are. This happens because blue wavelengths are shorter than red wavelengths and interstellar dust grains are the perfect size for scattering these shorter wavelengths away.<sup>42</sup>

A familiar example of a kind of reddening in our everyday lives is the sunset. During the day, the sky appears to be blue because the molecules in the atmosphere scatter light from the sun. The molecules scatter blue light much more than other colors, so the sky looks blue and much of the light that comes directly from the sun is missing some blue. The result of this is that the sun appears more yellow during the day. At sunset, the light has to travel through more atmosphere to reach our eyes and because of this, more blue light is scattered away. This is why as the sun gets lower in the sky, its appearance and the appearance of the area of sky around the sun becomes redder. So, provided it is not cloudy, the sun appears red at sunset but at noon the sky is bright blue.<sup>43</sup>

Understanding interstellar dust is a crucial part to understanding how and why reddening occurs. First, nebulae consist of gas and dust, so why does gas not cause reddening? Gas doesn't cause the reddening seen in nebulae because gas does not absorb very much light. If gas were the culprit behind reddening, the amount of gas in nebulae would be so massive that it would cause gravitational effects that would be able to be seen. So, it can be concluded that interstellar dust causes reddening. As demonstrated in Figure 29, each particle of dust consists of two parts: the core and the mantle. Cores are typically rich in carbon or silicates. The mantles are icy and are made of water, methane, and ammonia. These particles of dust are just a little bit smaller than the wavelength of visible light. Because of their size, they are able to effectively block light of shorter wavelengths but allow light of longer wavelengths to pass through. If the dust was slightly smaller, it would hardly block any light and reddening and extinction would not occur. If the dust was slightly bigger, it would block light of short and long wavelengths, causing reddening to not occur and extinction to occur to an even greater degree.<sup>43</sup>

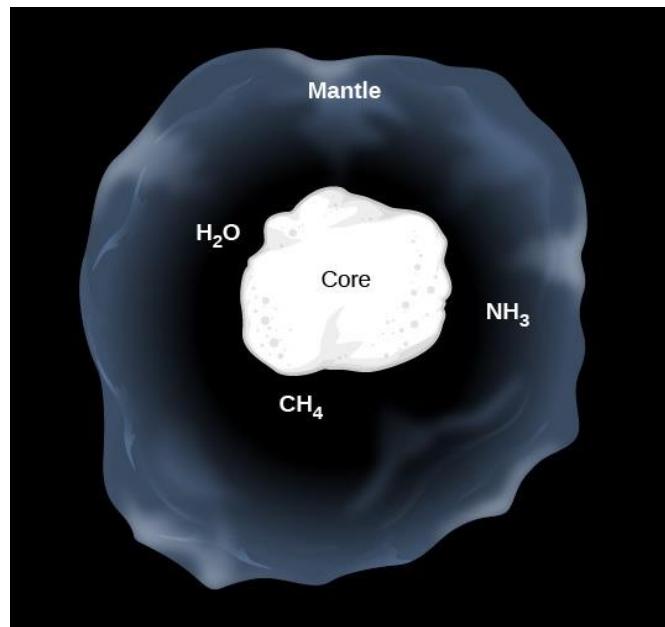
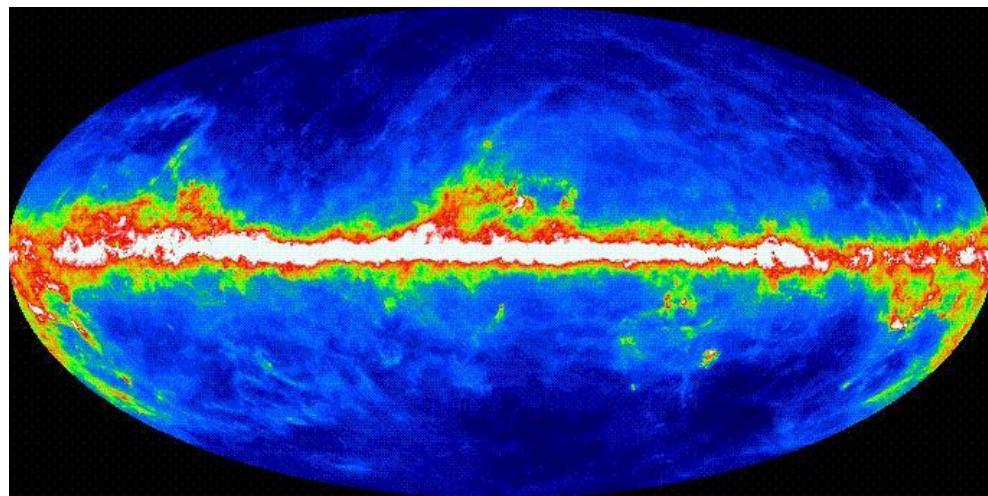


Figure 29: Illustration of Interstellar Dust Particle<sup>43</sup>

It is important that astronomers take reddening into account when they are interpreting images lest they misinterpret something. Reddening and extinction can cause astronomers to overestimate the distance of a star or underestimate luminosity. Reddening especially can cause astronomers to underestimate the temperature of a star, which can cause issues when trying to compare that star to other stars.<sup>44</sup> Due to this, many astronomers are working to develop better reddening maps of significant parts of the sky. For example, Figure 30 shows a reddening map of the entire sky made by astronomers at Caltech. On this map, blue represents low amounts of dust and red and white indicate high amounts of dust. The primary purpose of this map is to help astronomers gauge galactic extinction.<sup>45</sup>



**Figure 30: All-Sky Reddening Map<sup>45</sup>**

## B. Method

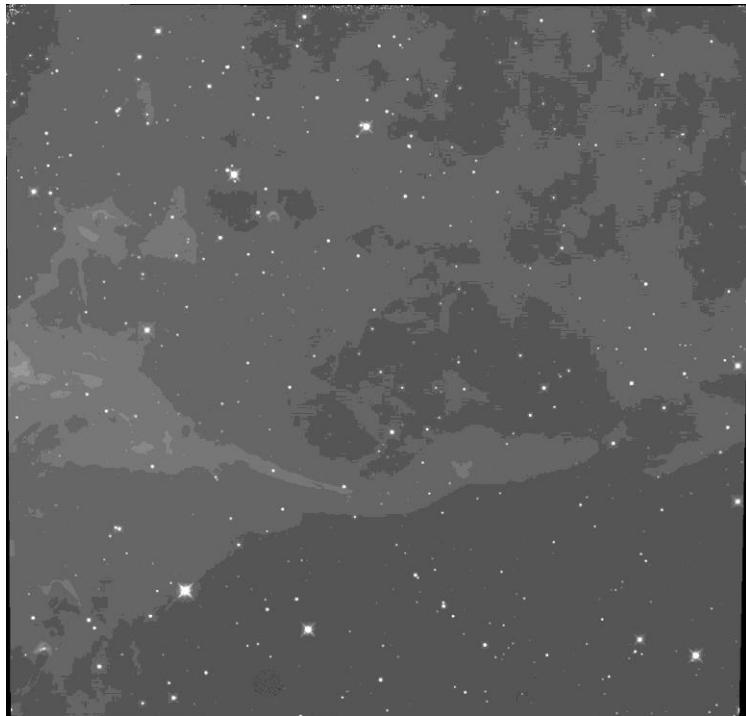
To conduct reddening analysis, images of the HH901 region of the Carina Nebula were selected. The HH901 region is commonly called “Mystic Mountain” and is shown in Figure 31. Two images with narrowband filters were selected and specific filters for each of these images were chosen because they correspond to particular wavelengths of light that are necessary for determining how much reddening is occurring.



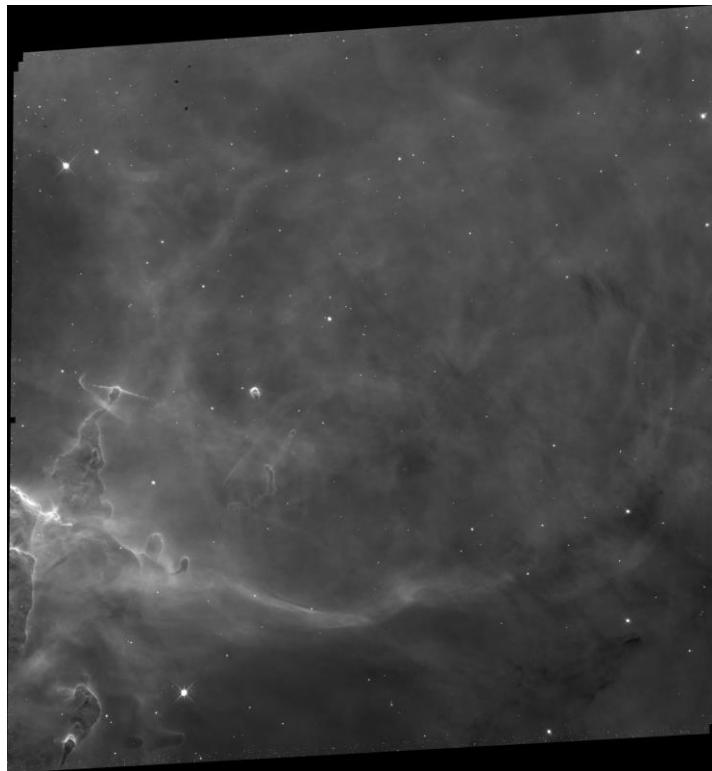
**Figure 31: Region HH901 of Carina Nebula<sup>46</sup>**

To understand the specific images that were chosen, a brief explanation of emission spectra is required. As described above, when atoms absorb or emit energy, electrons move between energy levels. Two of these emission lines, H $\alpha$  and Pa $\beta$ , are important to reddening analysis. Pa $\beta$  is also described above and H $\alpha$  is part of the Balmer series. H $\alpha$  represents a jump from the third to the second orbital of a hydrogen atom.<sup>47</sup> The H $\alpha$  line has a wavelength of 656 nm and Pa $\beta$  has a wavelength of 1281 nm. H $\alpha$  is in the visible spectrum and Pa $\beta$  is in the infrared spectrum. Thus, H $\alpha$  has a shorter wavelength than Pa $\beta$  meaning that the dust between an observer and the stars in the Nebula scatters more H $\alpha$  than Pa $\beta$ . This causes the actual value of the H $\alpha$  to Pa $\beta$  ratio to be smaller than the expected value if there were no dust.

For this reddening analysis, two narrowband images of the same area of the Carina Nebula were chosen and they are shown in Figures 32 and 33. The image with F657N filter corresponds to the H $\alpha$  line and the image with F128N filter corresponds to the Pa $\beta$  line. After images were selected, it was necessary to scale them properly. The F657N image was multiplied by 4.523 so that the size of the pixels and the exposure time were aligned for both images.

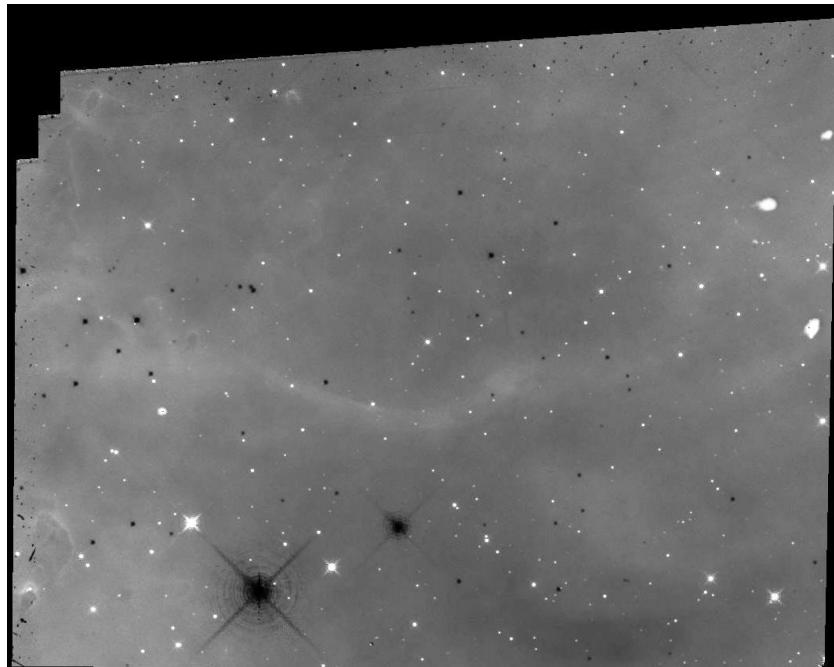


**Figure 32: F657N Image of HH901<sup>38</sup>**



**Figure 33: F128N Image of HH901<sup>38</sup>**

Next, the images were divided. Using ImageJ, both images were selected and F657N divided by F128N was applied. The resulting image, Figure 34, shows the reddening of this portion of the Carina Nebula.



**Figure 34:  $H_{\alpha} / Pa_{\beta}$  Resulting Image Used for Reddening Analysis**

### C. Conclusion

Figure 34 shows the resulting divided image. The F657N image was divided by the F128N image to give us the ratio of  $H_{\alpha}$  to  $Pa_{\beta}$ . Dividing the image takes each pixel in the first image and divides it by the corresponding pixel in the second image. The resulting image, then, is the ratio of each value of each pixel. This means that the darker regions have an absence of  $H_{\alpha}$  radiation meaning that more scattering of light is occurring there. Therefore, this comparison will show that there is the most dust in regions of the resulting image that are darkest. Since the presence of more dust will cause more reddening and extinction to occur, it can be concluded that the stars in Figure 34 that are dark, there is a significant amount of reddening and extinction occurring. In this region of the Carina Nebula, the two big stars in the lower half of the image are affected by reddening and extinction. Additionally, there are a variety of other smaller stars that have a significant amount of reddening occurring. This entire portion of the Carina Nebula has dust present as the entire background is gray. Finally, the star slightly above and to the left of the biggest star, along with various other smaller stars in the image, do not seem to have very much dust around them and, as a result, are not being reddened.

There were various steps that were unable to be done in this reddening analysis. The most prominent of these is that background radiation was unable to be subtracted. Had this been done, the resulting image would give us an even clearer idea of the extent of reddening and extinction that is occurring in the Carina Nebula. Regardless, the current image still gives us valuable information. The next steps for this reddening analysis would be to better quantify the data and then apply this to other images to correct them for the reddening that is occurring. This will allow the inner workings of the Carina Nebula to be better interpreted to investigate precisely what is happening.

## VI. Image Creation

### A. Hubble Space Telescope Archive

#### 1. Carina Nebula

To create an image of the Carina Nebula, three images of the Nebula were extracted from the Hubble Space Telescope Archive that were taken at different wavelengths. All images were taken using the WFC3/IR camera, with wavelengths F126N, F128N, and F164N. These images were taken on 2020-02-02 at approximately the same time.

The images were opened and placed into the PhotoPea software as different layers after they were downloaded from the Hubble database. Each layer was screened and the images' alignments were checked before their respective color channels were chosen. The image with the shortest wavelength (F126N) was assigned to the color blue, whereas the longest wavelength was assigned to red. The middle wavelength was green. After adjusting the curve feature for each layer in order to bring out the most detail possible, a final image of the Carina Nebula was created as shown in Figure 35.



**Figure 35: Image of Carina Nebula created using PhotoPea and Hubble Space Telescope Archive**

#### 2. Eta Carinae

In order to create an image of Eta Carinae, three images of the stars were also chosen from the Hubble Legacy Archive. All three images used the WPC2 camera. The filters of the images were F502N, F375N, and F658N. All three images were taken on December 31, 1993.

After the images were downloaded, they were transferred to FITS Liberator where their black and white levels were adjusted in order to display as much detail as possible. Then, the images were transferred to PhotoPea, where they were layered on top of each other and then the appropriate colors were assigned to each wavelength (F375N is blue, F502N is green, and F658N is red). After adding an adjustment layer to change the curve of each color, the final image of Eta Carinae is presented in Figure 36.

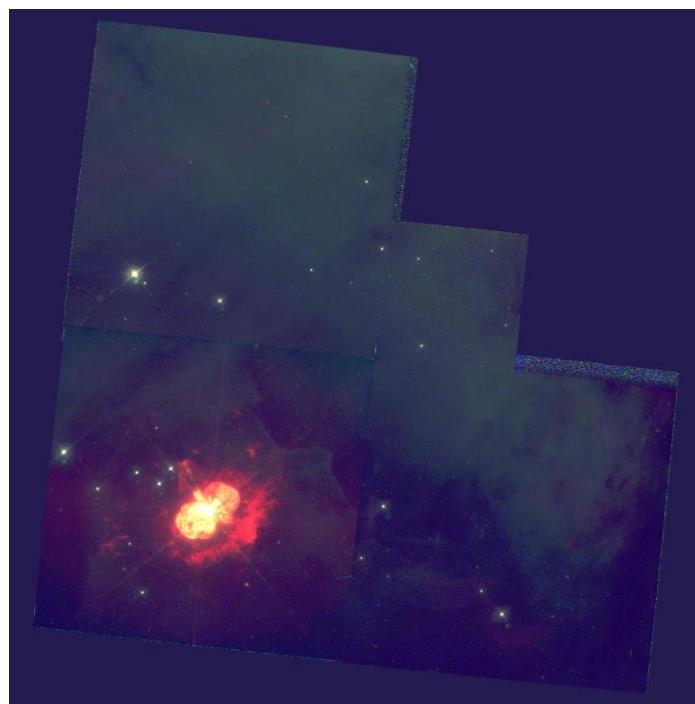


Figure 36: Image of Eta Carinae created using PhotoPea and FITS Liberator

## B. James Webb Telescope Released Images

In light of the newly released James Webb Telescope Image of a star-forming region in the Carina Nebula on July 12, 2022, the released files were compiled to create an image as well. The released files contained images collected at wavelengths 090N, 187N, 200W, 335M, and 444W. However, only the images from 200W, 335N, and 444W were used because all five files together were too large to be uploaded into one composite image. All three images were uploaded into the GIMP software and adjusted from there.

Many adjustments had to be made because the images were relatively new and weren't completely aligned on top of each other. First, they were cropped to a significantly smaller region to make it easier to handle. Then, the lasso warp feature was used to align each star in the cropped region on top of each other so that the layers seemed coherent together. Finally, the paint feature was used to color in the dark discrepancies inside the stars. The colors red, blue, and green were assigned to the wavelengths 200W, 335M, 444W respectively, and the curve feature was used to bring out detail in each layer. The final image is shown in Figure 37.



**Figure 37: Image of Star Formation Region in Carina Nebula created using GIMP and FITS Liberator**

## VII. Conclusion

The Carina Nebula is a very beautiful object in our galaxy. It is a complex structure that acts as a high temperature physics lab where many processes are taking place. This includes numerous star forming regions, significant ultraviolet radiation influence from massive stars, photodissociation regions including ionization fronts and dark pillars, as well as reddening and extinction.

The copious amounts of ultraviolet radiation in the Carina Nebula stems from the abundance of rare O-type stars in the Trumpler 15 region and the Eta Carinae star system. This ultraviolet radiation is extremely influential during the initiation of star formation, which explains why the Carina Nebula contains so many star forming regions. Furthermore, many compounds such as  $\text{N}_2\text{H}^+$ , HCN, and HNC found in star clumps have lower detection rates due to ionization from this radiation. These characteristics make the unique Carina Nebula resemble a high temperature physics lab with many processes inside.

The intricate features and processes in the Carina Nebula, especially in photodissociation regions, are clear representations of the interactions between stars, clusters, dust, and gas in nebulae. The identification of ionization fronts in various locations including around dark pillars shows the processes that are caused by the high energy UV emissions from O-type stars. This also identifies the interactions between the various elements present in a complex nebula.

Interstellar reddening and extinction are vital parts of astronomy that allow astronomers to better understand and investigate stars and structures in nebulae. In order to continue the study of the Carina Nebula, the creation of more reddening maps will be necessary.

Both the James Webb Telescope and Hubble Space Telescope are incredible pieces of technology that have been vital to a multitude of discoveries that have advanced modern science. Additionally, nearly all the data collected by these telescopes is made freely available, which makes studies such as this one accessible to conduct and easy to be expanded upon.

In conclusion, it is important to understand and study the various processes that occur in our galaxy, especially that of a nebula as unique as the Carina Nebula. Research into these processes will give astronomers more insight into the overall characteristics of our galaxy, and allow them to further discover the beauty behind outer space.

## **Appendix A: Tools Used**

### **A. SAOImageDS9 (DS9)<sup>48</sup>**

DS9 is an image analysis software that allows for the analysis of astronomical images in the FITS format, among other data. It was created by Mike Van Hilst, at the Smithsonian Astrophysical Observatory, Center for Astrophysics, Harvard University, under the direction of Eric Mandel in 1990. It is made possible by funding from the Chandra X-ray Science Center (CXC), the High Energy Astrophysics Science Archive Center (HEASARC) and the JWST Mission office at Space Telescope Science Institute.<sup>49</sup>

### **B. FITS Liberator**

FITS liberator allows for the processing of FITS image files, among others. This includes creating color images and adjusting features of an existing FITS image, as well as converting the file type of a FITS file.<sup>50</sup>

### **C. PhotoPea**

PhotoPea is an image editor that allows for the creation of false and true color images by the coloration and layering of obtained image files.<sup>51</sup>

### **D. Google Sheets**

Google Sheets is a data tabulation software that allows for the analysis of and calculations from raw data. This also aids in the creation and formatting of tables and graphs.<sup>52</sup>

### **E. ImageJ**

ImageJ is a public domain software that is also available for use in the browser. It is meant to be used to process and analyze scientific images. Wayne Rasband is the primary creator of ImageJ and he started this in 1997 where it was initially a cross-platform version of NIH Image. It provides the platform where we are able to subtract and divide images.<sup>53</sup>

## F. GIMP

GIMP stands for GNU Image Manipulation Program and is a free and open source image editor. The software was originally created by Spencer Kimball and Peter Mattis. This software opened FITS files and converted them to a format that allowed them to be opened in ImageJ.<sup>54</sup>

## References

- <sup>1</sup> *The Carina Nebula: Star Birth in the Extreme.* (n.d.). Hubble Space Telescope. Retrieved July 26, 2022, from <https://hubblesite.org/contents/media/images/2007/16/2099-Image.html>
- <sup>2</sup> *NASA's Webb Reveals Cosmic Cliffs, Glittering Landscape of Star Birth.* (2022, July 12). NASA. Retrieved July 26, 2022, from <https://www.nasa.gov/image-feature/goddard/2022/nasa-s-webb-reveals-cosmic-cliffs-glittering-landscape-of-star-birth>
- <sup>3</sup> Belleville, M. (2022, May 26). *About The Hubble Space Telescope.* NASA. Retrieved July 26, 2022, from [https://www.nasa.gov/mission\\_pages/hubble/about](https://www.nasa.gov/mission_pages/hubble/about)
- <sup>4</sup> Garner, R. (2022, May 26). *Observatory - Instruments.* NASA. Retrieved July 26, 2022, from <https://www.nasa.gov/content/goddard/hubble-space-telescope-science-instruments>
- <sup>5</sup> Garner, R. (2018, February 6). *The James Webb Space Telescope Observatory.* NASA. Retrieved July 26, 2022, from [https://www.nasa.gov/mission\\_pages/webb/observatory/index.html](https://www.nasa.gov/mission_pages/webb/observatory/index.html)
- <sup>6</sup> NASA. (n.d.). *Hubble Space Telescope User Documentation.* HST User Documentation - HDox - HST User Documentation. Retrieved July 26, 2022, from <https://hst-docs.stsci.edu/hom>
- <sup>7</sup> NASA. (n.d.). *JWST User Documentation.* JWST User Documentation Home - JWST User Documentation. Retrieved July 26, 2022, from <https://jwst-docs.stsci.edu/>
- <sup>8</sup> Kohler, S. (2017, March 1). *The Birth of Disks Around Protostars.* AAS Nova. Retrieved July 26, 2022, from <https://aasnova.org/2017/03/01/the-birth-of-disks-around-protostars/>
- <sup>9</sup> Palma, C. (n.d.). *Nuclear Fusion in Protostars | Astronomy 801: Planets, Stars, Galaxies, and the Universe.* John A. Dutton e-Education Institute. Retrieved July 28, 2022, from [https://www.e-education.psu.edu/astro801/content/l5\\_p4.html](https://www.e-education.psu.edu/astro801/content/l5_p4.html)
- <sup>10</sup> (n.d.). Stellar Death. Retrieved July 26, 2022, from <http://abyss.uoregon.edu/~js/ast122/lectures/lec18.html>
- <sup>11</sup> Freudenrich, C. (n.d.). *The Death of a Star - How Stars Work | HowStuffWorks.* Science | HowStuffWorks. Retrieved July 26, 2022, from <https://science.howstuffworks.com/star6.htm>
- <sup>12</sup> Miller, J. (2017, September 10). *Deep-Sky Objects: Crab Nebula (M1) -.* Astronomy Trek. Retrieved July 26, 2022, from <https://www.astronomytrek.com/deep-sky-objects-crab-nebula/>
- <sup>13</sup> *Q and A of the Day: White Dwarfs vs. Neutron Stars? | ChandraBlog | Fresh Chandra News.* (2010, January 29). Chandra X-ray Observatory. Retrieved July 26, 2022, from <https://chandra.harvard.edu/blog/node/182>
- <sup>14</sup> Petersen, C. C. (2019, December 8). *Exploring the Carina Nebula.* ThoughtCo. Retrieved July 26, 2022, from <https://www.thoughtco.com/carina-nebula-4149415>

<sup>15</sup> *Jet in Carina*. (n.d.). Hubble Space Telescope. Retrieved July 26, 2022, from <https://hubblesite.org/contents/media/images/2009/25/2594-Image.html?news=true>

<sup>16</sup> Higgins, N., & Moskowitz, C. (2014, September 1). *How the Iconic Pillars of Creation Arose*. Scientific American. Retrieved July 26, 2022, from <https://www.scientificamerican.com/article/how-the-iconic-pillars-of-creation-arose/>

<sup>17</sup> Fernie, J. D. (n.d.). *star cluster | Definition & Facts | Britannica*. Encyclopedia Britannica. Retrieved July 26, 2022, from <https://www.britannica.com/science/star-cluster>

<sup>18</sup> (2017, February 28). Hyperwall: Mystic Mountain: Pillars in the Carina Nebula from Hubble. Retrieved July 26, 2022, from <https://svs.gsfc.nasa.gov/30860>

<sup>19</sup> *Star Classification - Types of Stars*. (n.d.). Enchanted Learning. Retrieved July 26, 2022, from <https://www.enchantedlearning.com/subjects/astronomy/stars/startypes.shtml>

<sup>20</sup> *Hertzsprung-Russell Diagram | COSMOS*. (n.d.). Centre for Astrophysics and Supercomputing. Retrieved July 26, 2022, from <https://astronomy.swin.edu.au/cosmos/h/hertzsprung-russell+diagram>

<sup>21</sup> *Star Classification - Types of Stars*. (n.d.). Enchanted Learning. Retrieved July 26, 2022, from <https://www.enchantedlearning.com/subjects/astronomy/stars/startypes.shtml>

<sup>22</sup> Contreras, Y., Rebollo, D., Breen, S. L., Green, A. J., & Burton, M. G. (2018, December 03). Environmental conditions shaping star formation: the Carina Nebula. *Royal Astronomical Society*, 483(2), 1437-1451. <https://academic.oup.com/mnras/article/483/2/1437/5228753>

<sup>23</sup> (2014, January 20). MALT90. Retrieved July 26, 2022, from <http://malt90.bu.edu>

<sup>24</sup> *Eta Carinae*. (2020, September 3). NASA. Retrieved July 26, 2022, from [https://www.nasa.gov/mission\\_pages/chandra/images/eta-carinae.html](https://www.nasa.gov/mission_pages/chandra/images/eta-carinae.html)

<sup>25</sup> *Eta Carinae (η Carinae) | Facts, Information, History & Definition*. (2020, February 4). The Nine Planets. Retrieved July 26, 2022, from <https://nineplanets.org/eta-carinae-%CE%B7-carinae/>

<sup>26</sup> Wolfire, M. G., & Kaufman, M. J. (2011). Photodissociation Regions. *Encyclopedia of Astrobiology*, 1236–1243. [https://doi.org/10.1007/978-3-642-11274-4\\_1197](https://doi.org/10.1007/978-3-642-11274-4_1197)

<sup>27</sup> *Anatomy of a Photodissociation Region*. (2021). WebbTelescope.org; Q Starter Kit. <https://webbtelescope.org/contents/media/images/2021/024/01F5KKSNNM6YWR7MNMDJ2NEBCN?news=true>

<sup>28</sup> Tielens, A. G. G. M. (1993). Photodissociation Regions and Planetary Nebulae. *Symposium - International Astronomical Union*, 155, 155–162. <https://doi.org/10.1017/s0074180900170330>

<sup>29</sup> *Ionization fronts in interstellar gas and the expansion of HII regions | Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences*. (2016). Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences. <https://royalsocietypublishing.org/doi/10.1098/rsta.1961.0001>

<sup>30</sup> *Spectroscopy 101 – How Absorption and Emission Spectra Work*. (2021). WebbTelescope.org; Q Starter Kit. <https://webbtelescope.org/contents/articles/spectroscopy-101--how-absorption-and-emission-spectra-work>

<sup>31</sup> *Lesson Worksheet:Emission and Absorption Spectra | Nagwa.* (2022). Nagwa.com; Nagwa. <https://www.nagwa.com/en/worksheets/658148208480/>

<sup>32</sup> *III Ionized Hydrogen (HII) Regions.* (n.d.). Retrieved July 27, 2022, from <https://www.astronomy.ohio-state.edu/pogge.1/Ast871/Notes/Ionized.pdf>

<sup>33</sup> Admin. (2017, January 20). *Hydrogen Spectrum - Balmer Series, Definition, Diagram, Spectrum.* BYJUS; BYJU'S. <https://byjus.com/chemistry/hydrogen-spectrum/>

<sup>34</sup> *Ionization Energy: Definition, Chart & Periodic Table Trend.* (2021, December 21). Chemistry Learner. <https://www.chemistrylearner.com/the-periodic-table/ionization-energy>

<sup>35</sup> Bautista, M. A., & Pradhan, A. K. (1998). Ionization Structure and Spectra of Iron in Gaseous Nebulae. *The Astrophysical Journal*, 492(2), 650–676. <https://doi.org/10.1086/305061>

<sup>36</sup> Stellar Classification | Astronomy | Britannica. (2022). In *Encyclopædia Britannica*. <https://www.britannica.com/science/stellar-classification#ref288668>

<sup>37</sup> Populations I and II | Definition, Examples, & Facts | Britannica. (2022). In *Encyclopædia Britannica*. <https://www.britannica.com/science/Population-I#ref141711>

<sup>38</sup> STScI | Mikulski Archive for Space Telescopes (MAST) Portal. (2022). Stsci.edu. <https://mast.stsci.edu/portal/Mashup/Clients/Mast/Portal.html>

<sup>39</sup> Kiminki, M. M., & Smith, N. (2018). A radial velocity survey of the Carina Nebula's O-type stars. *Monthly Notices of the Royal Astronomical Society*, 477(2), 2068–2086. <https://doi.org/10.1093/mnras/sty748>

<sup>40</sup> Schmidt, Curtis, Smith N., CTIO, NOIRLab, NSF, AURA, & University of Minnesota. (2022). *Carina Nebula (NGC 3372)*. www.noirlab.edu. <https://noirlab.edu/public/images/noao0136a/>

<sup>41</sup> Kohler, Susanna. *Peering Inside the Pillars of Creation.* (2018, June 27). AAS Nova. <https://aasnova.org/2018/06/27/peering-inside-the-pillars-of-creation/>

<sup>42</sup> *Interstellar Reddening.* (n.d.). COSMOS - The SAO Encyclopedia of Astronomy. Retrieved July 26, 2022, from <https://astronomy.swin.edu.au/cosmos/II/Interstellar+Reddening>

<sup>43</sup> Fraknoi, A., Morrison, D., & Wolff, S. C. (n.d.). *Astronomy*. OpenStax. <https://pressbooks.online.ucf.edu/astronomybc/chapter/20-3-cosmic-dust/>

<sup>44</sup> *Comparing stars: 2.4 The effect of interstellar dust.* (2018, October 10). The Open University. Retrieved July 26, 2022, from <https://www.open.edu/openlearn/science-maths-technology/comparing-stars/content-section-2.4>

<sup>45</sup> *Galactic Dust Extinction Service.* (n.d.). IRSA/IPAC. Retrieved July 27, 2022, from <https://irsa.ipac.caltech.edu/applications/DUST/docs/background.html>

<sup>46</sup> NASA, ESA, & Hubble 20th Anniversary Team. (2018, April 11). *Pillars in the Carina Nebula.* Hyperwall. Retrieved July 26, 2022, from <https://svs.gsfc.nasa.gov/30940>

<sup>47</sup> LibreTexts. (2022). *Physical Chemistry*. LibreTexts. [https://chem.libretexts.org/Bookshelves/Physical\\_and\\_Theoretical\\_Chemistry\\_Textbook\\_Maps/Physical\\_Chemistry\\_\(LibreTexts\)/01%3A\\_The\\_Dawn\\_of\\_the\\_Quantum\\_Theory/1.04%3A\\_The\\_Hydrogen\\_Atomic\\_Spectrum](https://chem.libretexts.org/Bookshelves/Physical_and_Theoretical_Chemistry_Textbook_Maps/Physical_Chemistry_(LibreTexts)/01%3A_The_Dawn_of_the_Quantum_Theory/1.04%3A_The_Hydrogen_Atomic_Spectrum)

<sup>48</sup> Joye, W. A., & Mandel, E. (2012). New Features of SAOImage DS9. *Astronomical Data Analysis Software and Systems XII*, 295, 489. <https://ui.adsabs.harvard.edu/abs/2003ASPC..295..489J/abstract>

<sup>49</sup> SAOImageDS9 - About. (2022). Google.com. <https://sites.google.com/cfa.harvard.edu/saoimageds9/about?authuser=0>

<sup>50</sup> NOIRLab. (2022). Application: FITS Liberator 4. [Www.noirlab.edu.  
https://noirlab.edu/public/products/applications/app001/](https://noirlab.edu/public/products/applications/app001/)

<sup>51</sup> Photopea | Online Photo Editor. (2022). Photopea.com. <https://www.photopea.com/>

<sup>52</sup> Google Sheets: Online Spreadsheet Editor | Google Workspace. (2022). Google.com. <https://www.google.com/sheets/about/>

<sup>53</sup> ImageJ Software. (n.d.). ImageJ. Retrieved July 28, 2022, from <https://imagej.net/software/imagej/>

<sup>54</sup> A Brief (and Ancient) History of GIMP. (n.d.). GIMP. Retrieved July 28, 2022, from [https://www.gimp.org/about/ancient\\_history.html](https://www.gimp.org/about/ancient_history.html)



## **APPENDIX**



Abigail Barlow 1623 Mount Davis Rd Fort Hill PA 15540 abbybarlow1805@icloud.com Turkeyfoot Valley Area School IU 8	Eric Cui 5069 Westbury Farms Drive Erie PA 16506 ericcui2005@gmail.com NW PA Collegiate Academy IU 5	Korina Goldsmith 199 Farmington Place Greensburg PA 15601 korinagoldsmith@gmail.com Hempfield Area High School IU 7
Samantha Blocher 13 Turmeric Drive Mechanicsburg PA 17050 sblocher@outlook.com Cumberland Valley High School IU 15	Charlotte Devlin 66 Colony Oaks Drive Pittsburgh PA 15209 CharlotteDevlin8@gmail.com Oakland Catholic High School IU 3	Elaine Gombos 6301 Riverfront Dr. Pittsburgh PA 15238 goelaine314@gmail.com Shady Side Academy IU 3
Devin Bryant 467 5th Avenue New Kensington PA 15068 devinb8294@gmail.com Valley High School IU 7	Edison DeWilde 107 Chandler Road Chadds Ford PA 19317 e2gbrothers@gmail.com Unionville High School IU 24	Hannah Gong 1473 Catlin Way Dresher PA 19025 hannahcgong@gmail.com Upper Dublin High School IU 23
Nikola Cao 1638 Bridgeton Rd Breinigsville PA 18031 cao.nikola.a@gmail.com Parkland High School IU 21	Brea Fennick 60 Paul Revere Rd Oil City PA 16301 breaf11@gmail.com Oil City Senior High IU 6	Thomas Gordon 5464 Wilkins Avenue Pittsburgh PA 15217 thomas.mackay.gordon@gmail.com Winchester Thurston School IU 2
Krishnakumar Chinnasamy 2320 Middlegreen Court Lancaster PA 17601 krishnakchinnasamy@gmail.com Manheim Township HS IU 13	Owen Fick 400 Sensen Road Wernersville PA 19565 otfick@gmail.com Wilson High School IU 14	Aidan Green 103 Woodcrest Place Venetia PA 15367 aidang626@gmail.com Peters Township High School IU 1

Quiana Guo 2417 Prairie Rose Ln. State College PA 16801 qzguo11@gmail.com State College Area High School IU 10	Gabriel Kenning 891 White Farm Road Indiana PA 15701 gabekdagreat@gmail.com Indiana Area Senior High School IU 28	Graciela León 360 Winesap Dr Pittsburgh PA 15201 gracielapleon@icloud.com CAPA 6-12 IU 2
Jack Hasker 181 N. Fairview St Nazareth PA 18064 jkhasker@gmail.com Nazareth Area High School IU 20	Harris Khan 2101 Northway Road, Apt 1101 Williamsport PA 17701 rayykhan78@gmail.com Loyalsock Township High School IU 17	Daniel Levin 1344 Squirrel Hill Ave Pittsburgh PA 15217 daniellevin44@gmail.com Allderdice High School IU 2
Alena Hemminger 82 Lakeview Drive West Middlesex PA 16159 alenahemmi10@gmail.com West Middlesex High School IU 4	Peter Ko 314 Scarlet Peak Court Cranberry PA 16066 petersko230@gmail.com Seneca Valley High School IU 4	Xi Lin 2213 Bressler Ct Wyomissing PA 19610 xilin.0720@gmail.com Wilson High School IU 14
Diya Hundiwala 600 Addison Way Warrington PA 18976 diyahundiwala@gmail.com Central Bucks HS South IU 22	Grace Krakauskas 1112 Annin St Philadelphia PA 19147 geKrakauskas@gmail.com Little Flower Catholic High School IU 26	Rachel Lyn-Sue 1061 Knoll Dr Hummelstown PA 17036 rlynsue843@gmail.com Hershey High School IU 15
Sahil Jain 6285 Stonebridge Drive Fairview PA 16415 krish.jain0107@gmail.com NW PA Collegiate Academy IU 5	Olivia Krimin 1593 Philadelphia Street Indiana PA 15701 omkrimin@gmail.com Indiana Area Senior High School IU 28	Miranda Lynch 1011 Kent Gardens Lititz PA 17543 mirandalynchh@hotmail.com Manheim Township High School IU 13

Sarah Magdeburg 3 Sajer Road Pottsville PA 17901 sarahmag05@icloud.com Blue Mountain High School IU 29	Jan Mejía-Toro 110 Applewood Lane Lansdale PA 19446 johannomen@gmail.com Solebury School IU 23	Abhinav Palle 4 Chatham Hill Circle S Abington Twp PA 18411 abhinav.palle@gmail.com Abington Heights High School IU 19
Mason Matich 2925 Whitetail Ct Doylestown PA 18902 mcmatich@gmail.com Central Bucks HS East IU 22	Annaliese Niebauer 475 Tanner St. Ebensburg PA 15931 annalieseniebauer@cencam.org Central Cambria High School IU 8	Neel Pandula 1815 JFK Boulevard #2411 Philadelphia PA 19103 23npandula@gmail.com Julia R. Masterman IU 26
Andrew Mauriello 402 Blueberry Drive Duryea PA 18642 drewmauriello12@gmail.com Wyoming Seminary Upper School IU 18	Della Noon 513 Goers Hill Archbald PA 18403 della513618@gmail.com Valley View High School IU 19	Amanda Qu 971 Countess Drive Yardley PA 19067 amandajjqu@gmail.com Pennsbury High School IU 22
Lindsay McBride 714 Millbrook Lane Haverford PA 19041 lindsaymcbride6@gmail.com Academy of Notre Dame IU 25	Brian Olsen 437 Wicker Avenue Bensalem PA 19020 brolsen2232@gmail.com Bensalem High School IU 22	Rafael Quinodoz 123 Misty Meadow Lane Lansdale PA 19446 rcq2005@icloud.com St. Joseph's Preparatory School IU 23
Sarah Meigs 134 W Highland Ave Philadelphia PA 19118 sarahhmeigs@gmail.com Germantown Friends School IU 26	Lanee Olson 103 Washington Road Scranton PA 18509 laneeocm@gmail.com Scranton High School IU 19	Parimala Rajesh 6351 Mercury Drive Mechanicsburg PA 17050 parimala1125@gmail.com Cumberland Valley HS IU 15

Jai Rastogi  
 709 Kent Rd  
 Wallingford PA 19086  
 jai1rastogi@gmail.com  
 Strath Haven High School  
 IU 25

Pradyun Solai  
 9819 Sumner Drive  
 Allison Park PA 15101  
 pradyunsolai@gmail.com  
 North Allegheny High School  
 IU 3

Siming Tang  
 5241 Forbes Ave  
 Pittsburgh PA 15217  
 jelliphishi@gmail.com  
 Allderdice High School  
 IU 2

Malina Reber  
 325 North Stratton Street  
 Gettysburg PA 17325  
 mcreber3277@gmail.com  
 Gettysburg Area High School  
 IU 12

James Sparling  
 129 Mansion Drive  
 Media PA 19063  
 ben4.sparling@gmail.com  
 Penncrest High School  
 IU 25

Gianna Tout-Puissant  
 114 Big Buck Dr  
 Saylorburg PA 18353  
 toutpuissantg@mamail.net  
 Moravian Academy  
 IU 20

Bea Angela Ricafort  
 17168 Mt. Airy Road  
 Shrewsbury PA 17361  
 moiraangiebee@gmail.com  
 Susquehannock High School  
 IU 12

Emma Strickland  
 1605 Heather Lane  
 Williamsport PA 17701  
 emmastrickland2023@gmail.com  
 Loyalsock Township High School  
 IU 17

Tianna Tout-Puissant  
 114 Big Buck Dr  
 Saylorburg PA 18353  
 toutpuissantt@mamail.net  
 Moravian Academy  
 IU 20

Emma Riddell  
 306 Trinity Drive  
 Washington PA 15301  
 ridem1204@gmail.com  
 Trinity High School  
 IU 1

Lilyanna Strocko  
 30 Greenbrier Ave  
 Selinsgrove PA 17870  
 lmstrocko@gmail.com  
 Selinsgrove Area High School  
 IU 16

Daniel Tu  
 34 Independence Pl  
 Chesterbrook PA 19087  
 daniel.rr.tu@gmail.com  
 Conestoga High School  
 IU 24

Catherine Smith  
 3038 Bethel Church Rd  
 Warfordsburg PA 17267  
 smith.cgs99@gmail.com  
 Southern Fulton Jr Sr HS  
 IU 11

Kovid Tandon  
 104 Agnes Grace Ln  
 Wayne PA 19087  
 23tandonk@gmail.com  
 Conestoga High School  
 IU 24

Julia Vizza  
 332 Vanderbilt Lane  
 Harleyville PA 19438  
 juliaevizza@gmail.com  
 Souderton Area High School  
 IU 23

Amy Wang  
217 Evans Ave  
Blairsville PA 15717  
wangamy2023@gmail.com  
River Valley High School  
IU 28

Jason Weller  
671 Pheasant Run Rd.  
Hummelstown PA 17036  
jasonweller1@yahoo.com  
Hershey High School  
IU 15

Kurtis Zhang  
901 Caspian Dr.  
York PA 17404  
kurtiszhang200@gmail.com  
Central York High School  
IU 12

Daniel Wang  
1492 Main Street  
Pittsburgh PA 15241  
danielwangusc23@gmail.com  
Upper St. Clair High School  
IU 3

Emma Wilson  
21 Tallgrass Path  
Lancaster PA 17602  
emmagwilson2005@gmail.com  
Lancaster Mennonite High School  
IU 13

Christopher Zhou  
5128 S Jarrod Ct.  
Erie PA 16506  
chrischriszhou@gmail.com  
McDowell High School  
IU 5

Samuel Wang  
4 Bradford Circle  
Fort Washington PA 19034  
swang2023@gmail.com  
Germantown Academy  
IU 23

Toby Zapotocky  
546 Johnson St.  
Freeland PA 18224  
zappy051@gmail.com  
Hazleton Academy of Sciences  
IU 18

Francis Wehbe  
422 Fern Rd.  
Orwigsburg PA 17961  
protoscan@icloud.com  
Blue Mountain High School  
IU 29

Daniel Zhang  
338 Majestic Circle  
Dallastown PA 17313  
z331317@gmail.com  
Dallastown Area High School  
IU 12

Matthew Wehler  
536 Charles Street  
Saint Marys PA 15857  
mgwehler@gmail.com  
Saint Marys Area High School  
IU 9

Emily Zhang  
2005 Country Glenn Lane  
State College PA 16801  
emilyruohanhang33@gmail.com  
State College Area High School  
IU 10



John Alison  
Department of Physics  
Carnegie Mellon University

Benjamin Campbell  
Department of Engineering  
Robert Morris University

Hael Collins  
Department of Physics  
Carnegie Mellon University

Carrie Doonan  
Department of Biological Sciences  
Carnegie Mellon University

Lynley Doonan  
Department of Biological Sciences  
Carnegie Mellon University

Emily Drill  
Department of Biological Sciences  
Carnegie Mellon University

David Handron  
Department of Mathematics  
Carnegie Mellon University

Kerry Handron  
Physics  
Public VR

Barry Harris  
Department of Biological Sciences  
Ringgold High School

Barry Luokkala, PGSS Program Director  
Department of Physics  
Carnegie Mellon University

Brooke McCartney  
Department of Biological Sciences  
Carnegie Mellon University

Andrew McGuier  
Computer Science  
Carnegie Mellon University

Natalie McGuier  
Department of Biological Sciences  
Carnegie Mellon University

Prahlad Menon  
Swanson School of Engineering  
University of Pittsburgh

Beka Modrekiladze  
Department of Physics  
Carnegie Mellon University

Diptajyoti Mukherjee  
Department of Physics  
Carnegie Mellon University

Diane Nutbrown  
Chemistry Instructor  
The Neighborhood Academy

Jiben Roy  
Department of Chemistry  
Mississippi University for Women

Neil Simonetti  
Computer Science and Mathematics  
Bryn Athyn College

John Urbanic  
Pittsburgh Supercomputing Center  
Carnegie Mellon University

Russell Walker  
Department of Mathematics  
Carnegie Mellon University



**Co-Directors of Resident Life**

Ayda Ozdoganlar (First 4 weeks)  
Carnegie Mellon University

Cameron Blackwood-Short (last week)  
Carlow University

Justine Dell (last week)  
Haverford College

**Teaching Assistants/Counselors/Lab Assistants**

Matthew Abode  
Biology  
Bucknell University

Sophia Korono  
Physics  
Carnegie Mellon University

Cameron Blackwood-Short  
Biology  
Carlow University

Joseph McLaughlin  
Computer Science  
Carnegie Mellon University

Justine Dell  
Mathematics  
Haverford College

Michael Rybalkin  
Computer Science  
Stanford University

Clara (Wen) Dou  
Chemistry  
Carnegie Mellon University

Chenrui Shao  
Chemistry Storeroom Attendant  
Carnegie Mellon University

Gus Eberlein  
Physics  
Penn State University

Michael Sobol  
Biology  
University of Pittsburgh

Yu (Mars) Hong  
Chemistry  
Carnegie Mellon University

Shiyang Tian  
Computer Science/Mathematics  
Carnegie Mellon University

Joshua Kipiller  
Physics  
Penn State University

Jing-Yi Jennifer Zhou  
Chemistry  
Carnegie Mellon University