JOURNAL OF
EMERGING
INVESTIGATORS

VOLUME 3, ISSUE 6 | JUNE 2020
emerginginvestigators.org

The exoplanet search

An investigation of Transiting Exoplanet Surveying Satellite candidates

Capturing air pollutants

Zinc-based nanocrystals may be useful air filter components

Reducing tumor growth

Can these natural products help treat cancerous tumors?

Pollination patterns

How hummingbird pollination varies by season, age and sex

Will your book be a best-seller?

A new method for predicting success on the NYT Bestsellers List

JOURNAL OF EMERGING INVESTIGATORS

The Journal of Emerging Investigators is an open-access journal that publishes original research in the biological and physical sciences that is written by middle and high school students. JEI provides students, under the guidance of a teacher or advisor, the opportunity to submit and gain feedback on original research and to publish their findings in a peer-reviewed scientific journal. Because grade-school students often lack access to formal research institutions, we expect that the work submitted by students may come from classroom-based projects, science fair projects, or other forms of mentor-supervised research.

JEI is a non-profit group run and operated by graduate students, postdoctoral fellows, and professors across the United States.

Contents

VOLUME 3, ISSUE 6 | JUNE 2020

JOURNAL OF
EMERGING INVESTIGATORS

Article

The Effect of Different Fructose Diets on the Lifespan of *C. elegans*

Annabella Chen[1], Anne Fu[1], Ralph Zhang[1], Emily Xu[1], Aaron Mathieu[1]
[1]Acton-Boxborough Regional High School, Acton, MA, USA

SUMMARY

Sugar, such as fructose, is widely known to be a dietary cause for many health complications including diabetes, heart disease, and even cancer, and is an enticing yet harmful substance if consumed in large quantities. Fructose, typically in the form of high-fructose corn syrup, is the sugar most commonly found in fast foods. Thus, we chose to study the health effects of a high fructose diet on humans through experimentation on *Caenorhabditis elegans*, which are effective model organisms due to their rapid reproduction rate and stability despite variations in environmental conditions. We hypothesized that increasing sugar intake in *C. elegans* will reduce *C. elegans* survival, though moderate amounts of sugar may increase survival. The results show that the concentration of fructose had a significant influence on the survival rate of *C. elegans*. The *C. elegans* receiving 0% and 5% fructose concentration treatments had much higher survival rates than the 15% plates, which had zero surviving *C. elegans* after six days. After statistical analysis, the 5% and 15% plates were determined to yield significantly different survival rates. Thus, there is sufficient data to conclude that diets containing high levels of fructose negatively impact *C. elegans* life, suggesting that diets high in sugars such as high fructose corn syrup are harmful to humans. However, it was not possible to discern any significant difference between the 0% and 5% treatments from the data generated. Further experimentation would be needed to investigate the effects of diets containing a moderate amount of sugar.

INTRODUCTION

Overconsumption of sugar is known to reduce human life expectancy by causing a variety of diseases such as type 2 diabetes, heart attacks, and hypertension, posing a major risk to public health, especially in the United States (1, 2). *Caenorhabditis elegans* is a suitable model organism because 40% of human disease genes are homologous to genes in the *C. elegans* genome (3, 4). As a result, *C. elegans* have been well documented under many controlled conditions and are used in various studies to model systems in the human body, including issues regarding sugar-induced toxicity (5, 6) (5, 6). Thus, examining the effects of sugar on *C. elegans* gives insight on how it affects the human body, providing an ethical way to study human diseases. Previous studies found that high glucose diets in concentrations above 2% or fructose diets in concentrations at or above 10% shorten the lifespan of *C. elegans* (7). Zheng *et al.* discovered that while 555 mM (10% w/v) of fructose decreased lifespan, low amounts of fructose at 55 mM (1%) and 111 mM (2%) actually increased lifespan (8). High doses of fructose increase intestinal fat deposition (IFD) which disrupts the balance of hormones in *C. elegans*, resulting in a shorter lifespan and impaired ability to maintain homeostasis (8). Therefore, high fructose diets may also have harmful effects humans, since the human digestive system shares similarities with that of *C. elegans* (9). These include similar processes of lipid metabolism due to homologous genes for fat storage, making *C. elegans* a good model for human energy homeostasis and metabolic pathways (10, 11). Based on the conclusions made by Zheng *et al.*, we hypothesized that with a diet of 5% (278 mM) fructose concentration, the *C. elegans* lifespan increases in comparison to the worms exposed to the 0% fructose concentration, because the amount of sugar in a 5% concentration is not high enough to generate excessive IFD, instead providing extra energy for the worms to grow (8). Furthermore, we hypothesized that fructose concentrations of 10% (555 mM) or above at 15% (833 mM) would cause the lifespan of the *C. elegans* to decrease due to excessive IFD. In our experiment, we found that, in general, increasing concentrations of fructose decreased *C. elegans* rate of survival. Thus, our results indicate that fructose may have similar effects on humans, supporting claims that higher amounts of sugar can be detrimental to bodily health, generating excessive body fat.

RESULTS

In this study, we tested the effects of different concentrations of fructose on the survival rate of *C. elegans*. This was tested by subjecting the *C. elegans* to 0%, 5%, 10%, and 15% fructose concentrations in a total of twenty separate petri dishes, the concentration of fructose being the independent variable. The experimental control was the 0% fructose dishes which acted as a negative control, establishing the baseline of *C. elegans* survival when they are not treated with fructose.

To account for external variables such as temperature and light exposure, all of the dishes were kept together in the same place so as to be exposed to the same environmental conditions and were measured at approximately the same time every day. To ensure controlled group numbers, only five *C. elegans* were put into each petri dish so that the fluctuations in *C. elegans* population could be better observed and compared each day.

When gathering results, each petri dish was observed

Days After Plating	0% fructose	5% fructose	10% fructose	15% fructose
0	5/5	5/5	5/5	5/5
1	7/36	9/20	8/21	9/18
2	10/22	8/11	12/23	2/21
3	19/22	11/14	14/24	0/20
6	139/146	469/502	41/57	0/8

Table 1. Number of *C. elegans* alive over time over the total number of *C. elegans*. The data graphed in **Figure 1**, which shows how many *C. elegans* were observed to be alive each day (0, 1, 2, 3, and 6 days) for each test (0%, 5%, 10%, and 15%). The individual data points were calculated by finding the sum of the number of *C. elegans* alive on a particular plate in the five trials for each fructose concentration.

Days After Plating	0% fructose	5% fructose	10% fructose	15% fructose
0	100	100	100	100
1	19.4	45	38.1	50
2	45.5	72.7	52.2	9.5
3	86.4	78.6	58.3	0
6	94.5	93.4	71.9	0

Table 2. Average percent of *C. elegans* left alive days after start of experiment. The average percentage of C. elegans alive for each fructose group on each day (0, 1, 2, 3, and 6 days), calculated by dividing the number of C. elegans moving by the total number counted in each plate and averaging the resulting values for each test (0%, 5%, 10%, and 15%) (n = 5).

under a microscope every day to count the number of living *C. elegans* and the total number of *C. elegans* found on each plate, dead or alive. To determine which worms were alive we took advantage of the fact that when stimulated by light emitted by the microscope, living *C. elegans* display negative phototaxis movement, while dead *C. elegans* do not move. We counted the total number of live *C. elegans* for each concentration of fructose daily (**Table 1**, **Figure 1**); the total number of *C. elegans* counted each day was also recorded (**Table 1**). Using the ratio of *C. elegans* alive to the total number of *C. elegans* observed on each dish, the percentage of *C. elegans* alive on each dish was calculated. For each concentration, we calculated the average of the percentage of *C. elegans* alive and the standard error of the percentages each day (**Table 2**, **Figure 2**). We display the fluctuations in population growth and survival of *C. elegans* over time for each fructose concentration (**Figure 1**, **Figure 2**).

Further investigating the statistical meaning of our data, we conducted an ANOVA test. The null hypothesis states that there is no statistically significant difference between all of the fructose groups; the alternate hypothesis states that there is a statistically significant difference between the fructose groups. The test yielded an f-statistic value of 4.1643 and a *p*

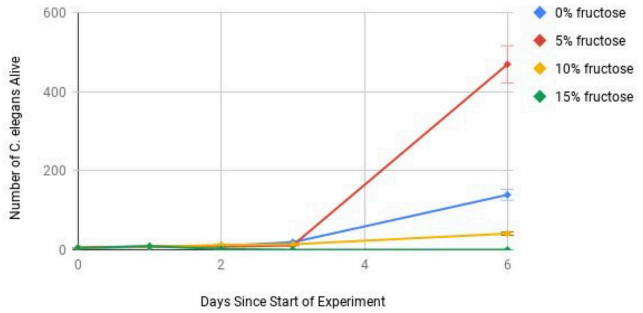


Figure 1. Effect of different fructose concentrations on *C. elegans* population size. The total population of C. elegans for each fructose concentration test (0%, 5%, 10%, and 15%) was observed over time (0, 1, 2, 3, and 6 days).

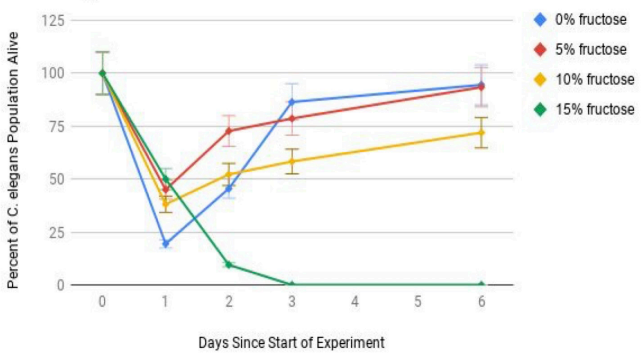


Figure 2. Effect of different fructose concentrations on percent of C. elegans alive. The percent of C. elegans alive in each test over time (0, 1, 2, 3, and 6 days). For each test (0%, 5%, 10%, and 15%), five trials were done and error bars represent standard error.

-value of 0.0309 (**Figure 3**). The calculated *p-value* is less than the critical value of 0.05, rejecting the null hypothesis. Thus, there was a statistically significant difference between the four fructose groups.

To identify which pairs of groups had statistically significant differences, a Tukey honestly significant difference (HSD) test was done (**Table 3**). Pairs 0% and 5%, 0% and 10%, 5% and 10%, and 10% and 15% had *Q-values* 0.8922, 0.5142, 1.4064, and 3.2722 respectively. With values less than the $Q_{critical\ value}$ of 3.46, the Tukey test indicated that there was no significant difference within these pairs. However, the 0% and 15% fructose pair showed a possibility of significant difference with a *Q-value* of 3.7864, and the 5% and 15% fructose pair yielded a significant difference with a *Q-value* of 4.6786.

After the first day, the percentage of *C. elegans* alive in the 5% fructose group was higher (45.0%) than in the 0% group (19.4%) (**Table 2**). However, when the 5% fructose line dipped lower than the 0% line after three days (**Figure 2**), the 0% and 5% fructose tests resulted in survival rates of 94.5% and 93.4% respectively on Day 6, which, according to the Tukey test, are not statistically different. Thus, it appears that adding a moderate amount of fructose to the *C. elegans'* diet may not have a beneficial effect on the survival rate of the *C. elegans*. However, the numbers of living *C. elegans* in the 0% fructose test and 5% fructose test are worth noting and investigating in the future; the 0% fructose test yielded a total of 139 *C.*

ANOVA Summary					
Source	Degrees of Freedom DF	Sum of Squares SS	Mean Square MS	F-Stat	P-Value
Between Groups	3	7561.0419	2520.3473	4.1643	0.0309
Within Groups	12	7262.772	605.231		
Total:	15	14823.8139			

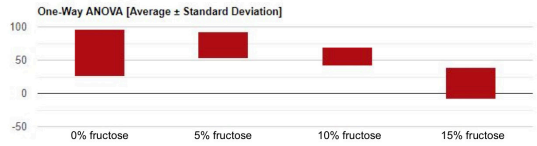


Figure 3. ANOVA test between all groups (0%, 5%, 10%, and 15% Fructose). Displays the ANOVA test done on the data for all of the groups (0%, 5%, 10%, and 15%). The null hypothesis states that there is no difference between the means of the groups. The calculated *p*-value is 0.0309.

Fructose groups	0% & 5%	0% & 10%	0% & 15%	5% & 10%	5% & 15%	10% & 15%
Q-value	0.8922	0.5142	3.7864	1.4064	4.6786	3.2722

Table 3. Tukey HSD Test Between All Groups (0%, 5%, 10%, 15% Fructose). *Q*-values from the Tukey test done on the data for all groups (0%, 5%, 10%, and 15%). The *Q*-critical value is 3.46.

elegans alive, and the 5% fructose test yielded a total of 469 *C. elegans* alive, the 5% fructose group clearly having a much higher resulting population. Looking at the other test groups, the 15% fructose group had zero surviving *C. elegans* by the sixth day, which was quite different from the 71.5% survival rate of the 10% fructose group and was a stark contrast with the near 100% survival rates of the 0% and 5% fructose groups. According to the results of the ANOVA and Tukey tests, there was a significant difference between the 15% fructose diet and both the 0% and 5% fructose diets, supporting that a higher fructose diet leads to lower survival rates. Therefore, the results partly align with our hypothesis; increasing the concentration of fructose in the diet of *C. elegans* decreases survival, though there is insufficient evidence to conclude that moderate amounts of fructose (5%) increase *C. elegan* life.

DISCUSSION

The experiment was designed to test how different concentrations of fructose in the diet of *C. elegans* would affect their survival rate, which was measured by counting the number of worms alive on each plate every day. Prior research indicated that the lifespan of *C. elegans* is correlated with the concentration of fructose in their diet, because the sugar builds intestinal fat in the worms, disrupting homeostasis in their bodies (8). With an impaired ability to maintain homeostasis, the *C. elegans* may be unable to carry out bodily functions effectively, subsequently decreasing their rate of survival. The results of this experiment support our hypothesis that increasing fructose concentrations above 10% lowers survival rate. Low concentrations of fructose were hypothesized to improve survival by providing a low damage energy boost though the results do not support that a 5% fructose diet would necessarily increase survivability.

The ANOVA test confirmed that there was a significant difference in survival rates among the groups. The post-hoc Tukey test indicated that between the groups there was a statistically significant difference in survival rate between the 0% or 5% fructose groups and 15% fructose group, concluding that when fed diets containing over 10% fructose, *C. elegans* survival significantly decreases. On the other hand, there is not enough evidence to indicate that there was a statistically significant difference between the 0% and 5% groups or between the 0% and 10% groups, both with *Q-values* below the $Q_{critical\ value}$. As a result, we cannot conclude that 5% fructose diets increase survival or that 10% fructose diets necessarily result in a significant decrease in survival. These results suggest that high fructose diets, which are lethal to *C. elegans*, may also have detrimental impacts on human health due to the high degree of conserved genes.

However, there are facets of the experimental design that could have caused error. One of the potential sources of error in this experiment was the method of placing *C. elegans* into the petri dishes. By using the inoculating loops, which are too big for accurate worm selection, it is possible that *C. elegans* eggs were carried over into the agar plates, causing unintended variances in the initial number of *C. elegans* in each plate. Furthermore, it was difficult to single out worms of the same size, so worms at different stages of their life cycle were used. However, trouble plating the worms was uniform for all of the experimental groups, so this error does not create bias and does not alter the results of the experiment. This would, however, be a possible explanation for the error bars on the data points that are indicative of variation between the groups. Finally, a factor that could have caused errors in the data is that Days 4 and 5 of data collection were skipped due to our inability to collect data over the weekend, resulting in an incomplete data set.

Another problem arose when analyzing the data regarding the number of *C. elegans* observed at the end of the experiment. The vast difference in the number of *C. elegans* present and the number alive begs the question of how to factor these differences into the statistical analysis. Since there is not one consistent total number of *C. elegans* in each group, the ANOVA and Tukey tests could only be conducted on the percentage of *C. elegans* alive. However, there is quite a large difference between having 57 *C. elegans* left in the 10% group and having 146 left in the 0% group, which is also very small in comparison to the 502 left in the 5% group (**Table 1**). Though both the 0% and 5% groups had near 100% survival rates, the difference between the numbers of living *C. elegans* observed in these two groups leads to uncertainty regarding whether a 5% fructose diet does indeed increase growth rates.

Therefore, future experiments can benefit from ensuring that the stated variables are securely controlled so as to minimize bias in the resulting data. It is worth investigating how to keep the total number of C. elegans constant and how to compare survival rates in a statistical analysis while taking

into consideration the number of C. elegans in total at the end of the experiment. Thus, this experiment could be revised to better gauge the effects of different fructose diets on *C. elegans*. Furthermore, the optimal fructose concentration in the *C. elegans* diet for the highest rate of survival could also be tested by using smaller increments of fructose concentration, for example testing one percent instead of five percent, so as to have more precise observations. Due to the homology between *C. elegans* and humans, this information can be used to advance research in the medical field, allowing for a better understanding of what levels of sugar and fructose are beneficial or harmful to humans.

MATERIALS AND METHODS

To set up the experiment, we repeatedly microwaved nematode agar (Fisher Science) for approximately 30 second intervals until a liquid solution was obtained. While the agar was melting, we acquired 20 small petri dishes and split them into 4 groups of 5 dishes, each group labeled with the concentrations being tested: 0%, 5%, 10%, and 15% fructose (Scholar Chemistry). We labeled the five petri dishes in each group individually with the numbers 1 through 5 to keep track of trials. Then we measured 2.5 grams, 5 grams, and 7.5 grams of fructose for the 5%, 10%, and 15% solutions respectively. After measuring out the fructose, we mixed each amount of fructose with 50 mL of water and 1 gram of nematode agar to create four diluted solutions, one solution containing only water and agar for the 0% test group. We then poured these solutions into their respectively labeled plates and left the dishes to cool for one day with the lids on to prevent contamination. We obtained the *C. elegans* through Dr. Gabel at Boston University School of Medicine. After the agar solutions cooled and solidified, we placed five live *C. elegans* onto each plate using inoculating loops under microscopes, and sealed the petri dishes with parafilm to be stored at room temperature on a lab bench. We distinguished live *C. elegans* from dead *C. elegans* by their movement after slightly agitating the plates, movement indicating life. In the following days, we counted and recorded the number of live and dead *C. elegans* on each plate daily. After the experiment, we conducted an ANOVA test and post-hoc Tukey test on the data for statistical analysis.

ACKNOWLEDGEMENTS

We would like to acknowledge Mr. Mathieu, our AP Biology teacher, for guiding us through the research and experiment process, helping to provide the resources we needed to conduct the experiment and giving us helpful feedback on our work. With his guidance, we learned a lot about biology and were able to improve our understanding of lab procedure and design. Furthermore, we wish to thank the reviewers and editors for giving us valuable constructive feedback, so we could improve our report.

Received: October 06, 2019
Accepted: April 25, 2020
Published: May 10, 2020

REFERENCES

1. Taubes, Gary. “Is Sugar Toxic?” *The New York Times*, www.nytimes.com/2011/04/17/magazine/mag-17Sugar-t.html.
2. Chalasani N, Younossi Z, Lavine JE, Diehl AM, Brunt EM, Cusi K, Charlton M, Sanyal AJ. “The diagnosis and management of non-alcoholic fatty liver disease: practice guideline by the American Gastroenterological Association, American Association for the Study of Liver Diseases, and American College of Gastroenterology.” *Gastroenterology*, vol. 142, no. 7, 2012, pp. 1592-609.
3. Riddle DL, Blumenthal T, Meyer BJ, Priess JR. “C. elegans II.” *Cold Spring Harbor Mongraph Series*, 2nd ed., vol. 33, Cold Spring Harbor Laboratory Press, 1997.
4. Corsi, Ann K. “A biochemist's guide to Caenorhabditis elegans.” *Analytical biochemistry*, vol. 359, no. 1, 2006, pp. 1-17, doi.org/10.1016/j.ab.2006.07.033.
5. Zhang, Jingyan et al. “C. elegans and its bacterial diet as a model for systems-level understanding of host-microbiota interactions.” *Current opinion in biotechnology*, vol. 46, 2017, pp. 74-80, doi.org/10.1016/j.copbio.2017.01.008.
6. Alcántar-Fernández J., Navarro RE, Salazar-Martínez AM, Pérez-Andrade ME, Miranda-Ríos J. “Caenorhabditis elegans respond to high-glucose diets through a network of stress-responsive transcription factors” *Plos One*, 10 Jul. 2018, doi.org/10.1371/journal.pone.0199888.
7. Choi, Shin Sik. “High glucose diets shorten lifespan of Caenorhabditis elegans via ectopic apoptosis induction.” *Nutrition research and practice*, vol. 5, no. 3, 2011, pp. 214-8. doi.org/10.4162/nrp.2011.5.3.214.
8. Zheng, Jolene *et al*. “Lower Doses of Fructose Extend Lifespan in Caenorhabditis elegans.” *Journal of dietary supplements*, vol. 14, no. 3, 2017, pp. 264-277, doi.org/10.1080/19390211.2016.1212959.
9. McGhee, James D. “The C. elegans intestine.” *WormBook: The Online Review of C. elegans Biology* [Internet], 3 May. 2007, www.ncbi.nlm.nih.gov/books/NBK19717.
10. Jones, Kevin T, and Kaveh Ashrafi. “Caenorhabditis elegans as an emerging model for studying the basic biology of obesity.” *Disease models & mechanisms*, vol. 2, no. 5-6, 2009, pp. 224-9, doi:10.1242/dmm.001933.
11. Hashmi S, Wang Y, Parhar RS, *et al.* “A C. elegans model to study human metabolic regulation.” *Nutrition and Metabolism*, vol. 10, no. 3, BMC, 4 Apr. 2013. doi.org/10.1186/1743-7075-10-31.
12. Lai CH, Chou CY, Ch’ang LY, Liu CS, Lin W. “Identification of novel human genes evolutionarily conserved in Caenorhabditis elegans by comparative proteomics.” *Genome research*, vol. 10, no. 5, May 2000, www.ncbi.nlm.nih.gov/pubmed/10810093.

JOURNAL OF
EMERGING INVESTIGATORS

Article

Capturing harmful air pollutants using an electrospun mesh embedded with zinc-based nanocrystals

Hansa Doppalapudi, Saunak Badam, Michael Tomlinson
Academy of Science, Academies of Loudoun, Leesburg, Virgina

SUMMARY

Metal-organic frameworks (MOFs) are compounds consisting of metal ions supported by organic ligands and may have the potential to combat the growing issue of air pollution. Zeolithic imidazolate framework-8 (ZIF-8) is a specific MOF that has favorable qualities for use in an air filter including high porosity, high surface area, and a high positive surface charge. ZIF-8 is known to be capable of adsorbing particulate matter. Therefore, the objective of this experiment was to determine the effectiveness of ZIF-8 in adsorbing polar, gaseous air pollutants, specifically nitrogen dioxide and hydrogen sulfide. In order to determine effectiveness, the percent change in concentration for each gas after the application of ZIF-8 crystals was measured via Fourier-transform infrared spectroscopy (FTIR). This is a technique used to obtain high resolution infrared absorption or emission spectra of solids, liquids or gases over a wide spectral range. Both experimental gases, nitrogen dioxide and hydrogen sulfide, were adsorbed by an average of ~32% over just 5 minutes, and the leakage within the capsule was an average of ~3.5%. These results were confirmed with the Kruskal-Wallis tests (p-value = 0.05). In the future, ZIF-8 could be embedded into existing air filters or electrospun on its own and tested for its effectiveness in adsorbing other polar pollutants such as carbon monoxide and ammonia. Our work highlights crystals as a potentially promising alternative or addition to current filter materials to reduce atmospheric pollution.

INTRODUCTION

Rapid industrialization and economic growth have caused a massive increase in the amount of harmful air pollution. Polluted air alone causes 3.3 million premature deaths every year worldwide (1). The concentration of photochemical and industrial pollution is visible in the brown and gray curtain looming over city buildings, termed “smog.” Smog consists of particulate matter and other harmful gases. Particulate matter (PM) is a form of industrial pollution, composed mostly of organic matter, nitrates, sulfates, ammonium, chloride, and elemental carbon (1). The emissions of millions of cars and factories concentrated in a tiny area contribute to this toxic smog. Smog is not effectively removed by current filters installed in buildings or cars creating a dangerously toxic environment in densely populated areas (1). Activated carbon filters are the most modern forms of filters, but are not effective in these high-smog areas, where the immense amounts of smog render filters unusable after a certain point (2). While efforts are being made to decrease pollution all over the world, a cheap, effective, and reusable way to reduce the toxic chemicals in the air has yet to be found.

Currently, most households and buildings implement activated carbon filters, which are proficient in removing certain toxins from the air (3). However, they cannot remove carbon monoxide, hydrogen sulfides, or certain nitric oxides, among others (3). These filters are also not reusable, as they must be replaced every few months. Although they can form strong chemical bonds with certain chemicals such as odors, noxious fumes, particulate matter, and organic compounds, they function ineffectively in areas where the air pollution contains too many contaminants (2). When pollution is very concentrated, filters that typically last for a long period of time will instead stop functioning in less than half of its lifetime, which is generally about one month (2). An ideal replacement for activated carbon filters should be reusable and adsorb a wide range of pollutants at concentrations at least as high as those commonly found in heavily polluted areas (3).

Metal-organic frameworks (MOFs) may potentially combat issues faced by activated carbon filters. MOFs are compounds consisting of metal ions or clusters of metal ions coordinated by organic linkers. The bond between organic and inorganic molecules yields a crystalline structure with strong bonds. The metal ions and organic linkers constitute the two components that dictate the structure and properties of the compound. MOFs have a much higher surface area and porosity than previously known porous materials, and also possess a generally high chemical and thermal stability. Their surface area ranges from 1,000 to 10,000 m^2/g, and generally, open space accounts for over 50% of the volume. There are over 20,000 types of MOFs to date, each with different properties and uses. The specific metal organic framework that has qualities that would be useful for a filter is a zeolitic imidazolate framework (ZIF-8), a zinc based metal organic framework with high porosity, high surface area, and high surface charge, which has a potential to adsorb certain air pollutants effectively (**Figure 1**). ZIF-8 is also known to be chemically and thermally stable and has been applied

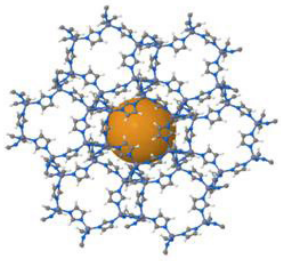


Figure 1. ZIF-8. A zinc-based metal organic framework with high porosity, high surface area, and high surface charge, which has a potential to adsorb certain air pollutants effectively.

in gas separation, catalysis, electronic devices, and drug delivery. Its use in gas separation with its high surface charge shows a high potential of bonding to harmful air pollutants (4). Furukawa et al examined the potential applications of MOFs and showed that MOFs also have a unique capability to selectively adsorb and store certain gases, which makes them useful for decreasing environmental pollution (5).

Previous research by Zhang et al, tested the effects of different MOFs electrospun into a filter on PM in the air (6). In this study, four different MOFs were spun via an electrospinner with one of two different polymers: polyacrylonitrile (PAN) and polystyrene (PS), creating a total of eight different filters. The different MOFs were: ZIF-8 (**Figure 1**), UIO-66-NH2, MOF-199, and Mg-MOF-74. Each metal organic framework was chosen for a different specific quality that would theoretically enable it to adsorb toxins from the air. The pollutants can be adsorbed by the MOF through chemisorption, a chemical reaction in which two substances react together, and the resultant chemical is trapped on the filter material (3). The researchers highlighted ZIF-8 as the most efficient MOF because of its large surface area, which enables it to be exposed to more air pollutants and its high surface charge, which is the primary reason it was the most efficient MOF. The ZIF-8 filter spun with a PAN polymer exhibited the highest removal efficiency for PM of various sizes. The polarity of the PM enhanced the adhesion with the unbalanced metal ions on the surface of the MOF, whereas the net positive charge of the filter attracted the partial negative charges of the polar molecules. After the test, the structure of the MOF remained intact, suggesting some degree of reusability, but the extent of this reusability remains unknown (6). These results could be explored in a different experiment, by testing a similar MOF for different pollutants. Nonpolar gases in the future have also been found to be able to be adsorbed with the use of various types of ZIF crystals such as ZIF-95 and ZIF-100. (7)

Other studies have also shown promise for the reusability of ZIF-8 filters, such as a study done by Chen et al in Beijing, where the ZIF-8 filter showed > 90% particulate matter capture even after 30 days of continual use and exposure to pollutants (8). It can also be reused after washing with water and ethanol. The ZIF-8 is able to adsorb molecules due to the structural deformation of “gate-like” openings which is induced by characteristics of the gas, such as polarity and molecular shape (9).

ZIF-8 can be nanospun as a filter on its own or eventually fabricated onto a variety of substances. As shown in this study, spinning onto plastic meshes yielded the best results in terms of the performance of the filter, rather than metal filters (8). The electrospinning process with metal organic frameworks allows the dispersion of the structures onto a surface. When a high enough voltage is applied to the solvent solution with the ZIF-8, the body of the liquid becomes charged, and electrostatic repulsion between the surface charges counteracts the surface tension, creating a cone. This forces the ejection of a liquid jet from the nozzle, which undergoes a stretching and whipping process, causing the creation of a long, thin thread. Once the electrospinning process is complete, the substrate should only have the polymer and the additive (ZIF-8) on it, as the solvent does not get electrospun (10).

Over the past year, research has shown that ZIF-8 crystals in pure powder form are able to selectively adsorb polar air molecules; the strong positive charge of the zinc ions in ZIF-8 created strong electrostatic interactions with the partial negative side of polar air molecules. Previous research has indicated that these polar molecules will be further attracted to the active sites of the ZIF-8 crystals, when electrospun, because these active sites will be further exposed through the fiber creation process (11). This research showed that nitrogen dioxide and hydrogen sulfide were significantly adsorbed by the crystals. In this study we were electrospinning these ZIF-8 crystals into a filter to test if the properties of binding air pollutant crystals still remain after electrospinning. Building on this knowledge, in this study we tested whether the properties of ZIF-8 remained after electrospinning into an air filter. The filter was made with polyethylene oxide (PEO) as the polymer and ethanol as the solvent. Previous research has shown that PEO is a polymer which works well with electrospinning ZIF-8, but this combination has yet to be tested in the context of gases (12).

RESULTS

We hypothesized that zeolitic imidazolate framework-8 (ZIF-8) will adsorb polar air pollutants, more efficiently than current heating, ventilation, and air conditioning (HVAC) filters. Current HVAC filters are presently the best suited for preventing particulate matter from entering the human airway. Some of these current filters even have activated carbon, known as a universal adsorbent and have been proven to be effective in eliminating the chemical pollution found in HVAC systems. However, these carbon-activated filters still lack in required characteristics for an ideal filter (7).

To test whether ZIF-8 filters could adsorb polar gases, we electrospun ZIF-8 crystals into a filter and then measured the absorbance efficiency of two gases : nitrogen dioxide and hydrogen sulfide. ZIF-8 loading was done for the experimental group, and its gas adsorption of nitrogen dioxide and hydrogen sulfide was measured against the control group, which would

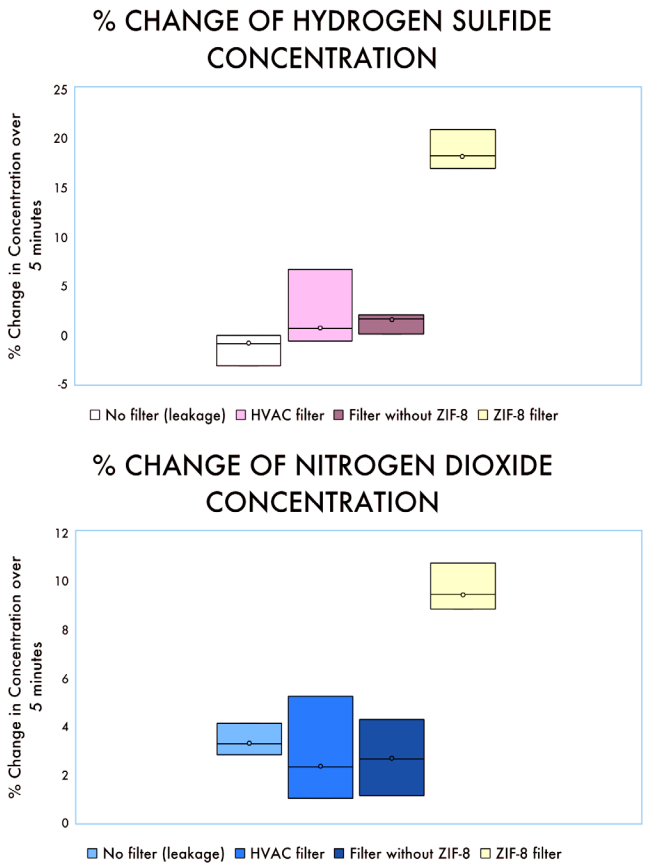


Figure 2. Percentage change in concentration of Hydrogen Sulfide during ZIF-8 filter usage. The change from the ZIF-8 filter was more than any other experimental filter tested. (Non ZIF-8 filter : 5.33% change; ZIF-8 filter : 22.58% change). $p = 0.05$. Similarly the percentage change in concentration of Nitrogen dioxide was greater when ZIF-8 filter was used, more than any other experimental filter tested. (Non ZIF-8 filter : 1.73% change; ZIF-8 filter : 33.20% change). $p = 0.05$. Each filter was tested three times.

be just the electrospun filter without ZIF-8 loading. All of the data were compiled into two tables and tested for significant differences using a Kruskal Wallis test. The percentage change in concentration of Nitrogen dioxide was greater when ZIF-8 filter was used, more than any other experimental filter tested. (Non ZIF-8 filter : 1.73% change; ZIF-8 filter 33.20% change). $p = 0.05$. Similarly, the percentage change in concentration of Hydrogen Sulfide was greater when ZIF-8 filter was used, more than any other experimental filter tested. (Non ZIF-8 filter : 5.33% change; ZIF-8 filter 32.58% change). $p = 0.05$ (**Figure 2**). That could be believed to be one of the lowest possible p-values for that size data set because both had the same values indicating that the change caused by ZIF-8 could be significant. By testing the efficiency of ZIF-8, a reusable filter was devised that can adsorb harmful air pollutants, which cannot be effectively adsorbed by current filters such as toxic carbon monoxide and volatile organic compounds. Removing harmful pollutants from the air is crucial, especially as this becomes an increasingly larger problem as the state of air quality around the world worsens.

DISCUSSION

We originally hypothesized that the zeolitic imidazolate framework-8 (ZIF-8) crystals would adsorb polar pollutants better than current filters. We used FTIR (Fourier-transform infrared spectroscopy), a technique used to obtain an infrared spectrum of absorption or emission of a solid, liquid or gas. The hypothesis was tested by comparing absorption of nitrogen dioxide and hydrogen sulfide between electrospun filters embedded with ZIF-8 crystals, electrospun filters without ZIF-8 crystals and store-bought grade 2 HVAC filters. The data gathered in this investigation supports the claim that ZIF-8 crystals are able to effectively and selectively remove polar air pollutants from the air. The filter with ZIF-8 crystals in it decreased the concentration of the pollutant in the capsule, and much more so than any of the control groups (**Figure 2**). The tests themselves were also done over only five minutes, so it is possible that over a longer period of time, even more gas would be adsorbed. Additionally, the way the experiment was designed, the air had no way of actually flowing through the filter, it was only on top of it. This suggests that electrospun meshes embedded with ZIF-8 crystals not only can adsorb these harmful air molecules, but selectively attract them down to its surface and trap them. This increases the likelihood of ZIF-8 preventing air pollution, because the air only has to be in contact with, not flowing through it.

This experiment also paves the way for the possibility of ZIF-8 being able to adsorb many other polar gases. For example, carbon monoxide is a polar gas, that is very dangerous causing thousands of deaths each year. Due to carbon monoxide's strong polar bonds and partial charges, ZIF-8 has the potential to adsorb carbon monoxide, in addition to certain volatile organic compounds (VOCs). Both of these gases currently have no viable filtration method and are very to the environment and public health. In the future, its thermal stability could be tested to determine the level of its reusability. Due to ZIF-8's propensity to adsorb polar molecules, there are numerous possible applications. ZIF-8 can be used as its own filter and used as a substitute for traditional filters currently used in air vents as it will be able to prevent a greater number of dangerous gases to go through. It can also be used in conjunction with catalytic converters in automobiles to adsorb dangerous pollutants that are not able to be catalyzed. Metal-organic frameworks could be the future of gas absorption, as ZIF-8 shows the possibility of a working, cost-effective method of adsorption that may be reusable.

MATERIALS AND METHODS

We originally hypothesized that the ZIF-8 crystals would adsorb polar pollutants better than current filters. The hypothesis was tested by comparing absorption of nitrogen dioxide and hydrogen sulfide between electrospun filters embedded with ZIF-8 crystals, electrospun filters without ZIF-

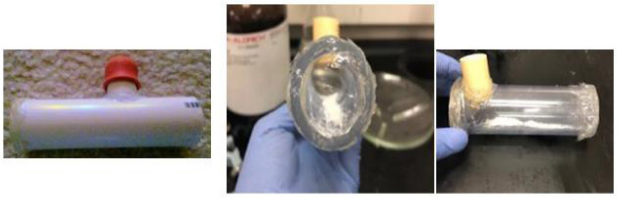


Figure 3. Fabrication of a gas capsule using a PVC pipe. A PVC pipe of 18 cm length and 2.5 cm bore was acquired by cutting from a longer PVC pipe. A smaller PVC pipe of 2 cm length and 1 cm bore was also acquired. A hole of about 1.3 cm diameter was cut into the side of the former PVC pipe using a drill press. The latter PVC pipe was glued into the hole using hot glue. Then saran wrap was hot glued onto the ends of the larger PVC pipe to ensure gas could not escape through the sides. Finally, a septum stopper was placed on the open end of the smaller PVC pipe to create an airtight seal. Gases would be injected via hypodermic needle through the septum stopper.

8 crystals and store-bought grade 2 HVAC filters.

To complete this research, nitrogen dioxide and hydrogen sulfide had to be synthesized. Fourier-transform infrared spectroscopy (FTIR) was used to determine concentration changes, and a gas capsule was made to put into the FTIR. The percent change in concentration of each gas was then recorded on the FTIR machine, calculated as the area under the IR spectra curve at five minutes as a percentage of the area under the IR spectra at zero minutes.

Gas capsule creation

A set up consisting of PVC, saran wrap and a septum stopper was used (**Figure 3**). Gases were injected via hypodermic needle through the septum stopper.

Synthesis of gases

Nitrogen dioxide

Nitrogen dioxide was synthesized using the reaction that occurs between tin and nitric oxide. 0.2 g of tin was placed in a small Erlenmeyer flask and a septum stopper was placed on top of it (**Figure 4**). Then, with a syringe and a hypodermic needle, 5 mL of nitric acid was inserted into the flask. Nitric acid was used in excess. An open needle was placed into the septum stopper to prevent extreme pressure buildup. The reaction was allowed to occur until the expulsion of heat was easily detected. Then, a second hypodermic needle was used to intake the desired volume of gas and then place into the gas capsule. The concentration of the gas in the capsule was determined through stoichiometric calculations, with an assumption of 100% yield.

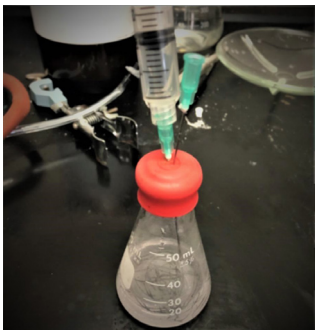


Figure 4. Nitrogen dioxide synthesis by reacting tin and nitric oxide. 0.2 g of tin was obtained, placed in a small Erlenmeyer flask and a septum stopper was placed on top of it.

Hydrogen sulfide

Hydrogen sulfide was synthesized with the reaction that occurs between sodium sulfide and hydrochloric acid. 3.0 mL of 2.0 M sodium sulfide was placed in a small Erlenmeyer flask, and a septum stopper was placed on top of it. With a syringe and a hypodermic needle, 6.0 mL of 3.0 M hydrochloric acid was inserted into the flask. An open needle was placed into the septum stopper to prevent extreme pressure buildup. The reaction was allowed to complete, and transferred to the gas capsule in a manner similar to previously described for nitrogen dioxide.

Electrospinning

The electrospinning parameters remained constant for all of the filters. The voltage was 15 kV, the flow rate was 1.2 mL/hr, the spinneret diameter was 21 gauge, and the distance of the needle to the plate was 15 cm. Two types of filters were made: -ZIF-8 and +ZIF-8. For both types of filters, polyethylene oxide (PEO) was used as the polymer and ethanol as the solvent. 0.200 g of PEO and 10 mL ethanol were used in both filter types.

The +ZIF-8 filter also had 0.1 g of ZIF-8 crystal dissolved in the ethanol via sonication. Polymer solutions were made by mixing PEO and ethanol solutions for three hours. The solutions were spun with the electrospinning parameters aforementioned onto an aluminum foil substrate about 3 cm by 3 cm.

Testing ZIF-8

For the no filter control trials, 6.0 mL of the gas was inserted into the gas capsule, and the initial area under the IR spectra was recorded. For the ZIF-8 control trials, a 1.5 cm by 7 cm strip of ZIF-8 filter was placed in the capsule. 6.0 mL of the gas was inserted into the gas capsule, and the initial area under the IR spectra was recorded. Finally, for the HVAC control trials, a MERV level 2 HVAC filter was bought for this experiment. A 1.5 cm by 7 cm strip of HVAC filter was placed in the capsule. 6.0 mL of the gas was inserted into the gas capsule, and the initial area under the IR spectra was recorded. For each of these experiments, the area was recorded again after five minutes, and the change in area was used to determine the percentage change in concentration of the gas due to leakage. This was done three times for both gases, for a total of six HVAC control trials.

Experimental trials

A 1.5 cm by 7 cm strip of +ZIF-8 filter was placed in the capsule. 6.0 mL of the gas was inserted into the gas capsule, and the initial area under the IR spectra was recorded. After five minutes, the area was recorded again, and the change in area was used to determine the percentage change in concentration of the gas due to leakage. This was done three times for both gases, for a total of six experimental trials. A 100% yield could be assumed because there is a density difference between nitrogen dioxide, hydrogen sulfide and normal air; nitrogen dioxide and hydrogen sulfide are both denser than air, and thus are at the bottom of the flask. When the syringe was used to pull out the gas, the syringe was put at the bottom, to make sure that it was only picking up the synthesized gas.

The expulsion of heat was measured relatively compared to the original temperature of the glassware. Because it was done by the same person every time, the expulsion of heat was mostly a way to make sure the gas itself was created and that the reaction was complete.

Statistical Analysis

To test the significance of ZIF-8 on absorbance of just polar molecules, nonparametric Kruskal-Wallis tests were performed on control trials without ZIF-8 and experimental trials with ZIF-8. Gas absorption between without ZIF-8 and with ZIF-8 filters was tested.

A nonparametric, Kruskal-Wallis test was run to determine the effect of the filter on each of the gases. This was done to see if a significant amount of gas was removed by the filter with ZIF-8, by comparing it to one without ZIF-8. The Kruskal-Wallis test was used as the primary test rather than its corresponding parametric test, an ANOVA test, due to the limited sample size. The Kruskal-Wallis compared the data without a comparison to a normal curve, and with the use of medians and ranks rather than with the use of means. Nonparametric tests are less affected by any outliers or skewed data, and work best for data with small sample sizes which may not necessarily fall under a normal curve. For these reasons, a Kruskal-Wallis was the primary statistical test for this experiment.

Received: September 11, 2018
Accepted: October 19, 2019
Published: May 12, 2020

REFERENCES

1. Lelieveld, J., *et al*. "The Contribution of Outdoor Air Pollution Sources to Premature Mortality on a Global Scale." *Nature*, vol. 525, no. 7569, 2015, pp. 367–371., doi:10.1038/nature15371.
2. Fox, A.R. & Lise, J.M. "Air filter with sorbent crystals". *Google Patents* US Patent 9,539,532. 2017.
3. Chambre, A. "Effects of Carbon Filtration Type on Filter Efficiency and Efficacy: Granular Loose-Fill vs. Bonded Filters". *Air Science Corporation LLC, (1), 1-6.* 2014.
4. Chen, B., *et al*. "Zeolitic Imidazolate Framework Materials: Recent Progress in Synthesis and Applications." *J. Mater. Chem. A*, vol. 2, no. 40, 2014, pp. 16811–16831., doi:10.1039/c4ta02984d.
5. Furukawa, H., *et al*. "The Chemistry and Applications of Metal-Organic Frameworks." *Science*, vol. 341, no. 6149, 2013, pp. 1230444–1230444., doi:10.1126/science.1230444.
6. Zhang, Y., *et al*. "Preparation of Nanofibrous Metal–Organic Framework Filters for Efficient Air Pollution Control." *Journal of the American Chemical Society*, vol. 138, no. 18, Feb. 2016, pp. 5785–5788., doi:10.1021/jacs.6b02553.
7. Wang, B., Cote, A.P., *et al*. "Colossal cages in zeolitic imidazolate frameworks as selective carbon dioxide reservoirs". *Nature, 453, 2017-211.*
8. Chen, Y., *et al*. "Roll-to-Roll Production of Metal-Organic Framework Coatings for Particulate Matter Removal." *Advanced Materials*, vol. 29, no. 15, 2017, p. 1606221., doi:10.1002/adma.201606221.
9. Ania, C.O., *et al*. "Understanding Gas-Induced Structural Deformation of ZIF-8." *The Journal of Physical Chemistry Letters*, vol. 3, no. 9, 2012, pp. 1159–1164., doi:10.1021/jz300292y.
10. Li, D., and Y. Xia. "Electrospinning of Nanofibers: Reinventing the Wheel?" *Advanced Materials*, vol. 16, no. 14, 2004, pp. 1151–1170., doi:10.1002/adma.200400719.
11. Ostermann, R., *et al*. "Metal–Organic Framework Nanofibers Via electrospinning." *Chem. Commun.*, vol. 47, no. 1, 2011, pp. 442–444., doi:10.1039/c0cc02271c.
12. Armstrong, M.R., *et al*. "Hierarchical Pore Structures and High ZIF-8 Loading on Matrimid Electrospun Fibers by Additive Removal from a Blended Polymer Precursor." *Industrial & Engineering Chemistry Research*, vol. 55, no. 37, Aug. 2016, pp. 9944–9951., doi:10.1021/acs.iecr.6b02479.

JOURNAL OF
EMERGING INVESTIGATORS

Article

Factors influencing muon flux and lifetime: An experimental analysis using cosmic ray detectors

Finsam Samson[1], Andrew Du[2], Michael Niedballa[3]
[1]Troy High School, Troy, Michigan
[2]Cranbrook School, Bloomfield Hills, Michigan
[3]Wayne State University, Detroit, Michigan

SUMMARY

Muons, one of the fundamental elementary particles, originate from the collision of cosmic rays with atmospheric particles and are also generated in particle accelerator collisions. Muon flux and lifetime are usually analyzed to obtain important insights regarding time dilation and to ascertain the characteristics of the particles and processes that occur during collision. In this study, we analyzed the factors that influence muon flux and lifetime using Cosmic Ray Muon Detectors (CRMDs). We hypothesized that the positioning of cosmic ray detector scintillator plates and depth of water above such scintillator plates would affect muon flux and lifetime. Muon flux and lifetime would be affected as an increase in matter in the path of a muon would slow or stop muons passing through it. Altering liquid shielding indicated that muon flux was inversely related to the depth of water that it passed through. Altering scintillator positioning indicated that detection rates of muon decay increased with the distance between the top and bottom of the scintillator plates resulting from the scintillator plate orientation. Unlike previous studies, the effects of the shielding were noticeable at smaller scales. Overall, this study suggests that water can be used to decrease muon flux and that the scintillator orientation is a potential determinant of the volume of data collected in muon decay studies.

INTRODUCTION

The Standard Model, a particulate model of how the universe works, purports the existence of two classes of elementary particles: fermions and bosons (1). Bosons include many of the force carrying particles in the universe, while fermions include leptons and quarks, which make up much of the matter in the universe. Leptons consist of three categories of particles with charge – the electron, the muon, and the tau, as well as particles with no charge – the neutrinos. Neutrinos are significantly smaller than electrons, and correspond to each charged lepton.

The Muon

The muon is a fundamental elementary particle that was discovered in 1937 by C.W. Anderson and S.H. Neddermeyer (2). The muon is about 207 times larger than an electron, and has the same charge as the electron, -1 e. Muons are unstable particles with a mean lifetime of 2.2 μs, after which they decay into lighter particles with the same net charge as the muon because charge must be conserved when decay occurs (3). Each muon decays into an electron (with the same charge as the muon), an electron antineutrino, and a muon neutrino. This is conserves charge. The mean-life of the muon is significantly longer than most other unstable elementary particles.

The Source of Muons on Earth

Muons are a product of the collision of cosmic rays with molecules in the upper atmosphere and other atmospheric particles (4). Cosmic rays are high-energy charged particles traveling near the speed of light, originating from supernovae and other sources beyond the solar system. When a cosmic ray proton hits Earth's upper atmosphere, it collides into a particle in the Earth's atmosphere. As this occurs, the cosmic ray forms a shower of various particles resulting from the collision. These new particles then travel and decay as well, cascading down through the atmosphere. Muons are generated when the cosmic ray proton collides with atoms in the atmosphere, creating short-lived pions. The pions soon decay into muons that travel towards the Earth's surface (**Figure 1**). Often, the muon, having a longer lifetime than many other unstable subatomic particles, reaches Earth's surface.

Cosmic Ray Muon Detectors

Cosmic Ray Muon Detectors (CRMDs) are devices that investigators use to study muons originating from cosmic rays. CRMDs use a set of counters, each comprised of a scintillator plate and a photomultiplier tube (PMT) (5). The scintillator plate luminesces when exposed to muons. This effect is amplified by the PMT, which allows for an electrical signal to be produced when a single muon passes through the plate. Photons can interfere with the signaling from scintillator counters, so each scintillator plate must be shielded from light.

CRMD signals can be read and analyzed by a specialized QuarkNet software (5). The software can characterize the

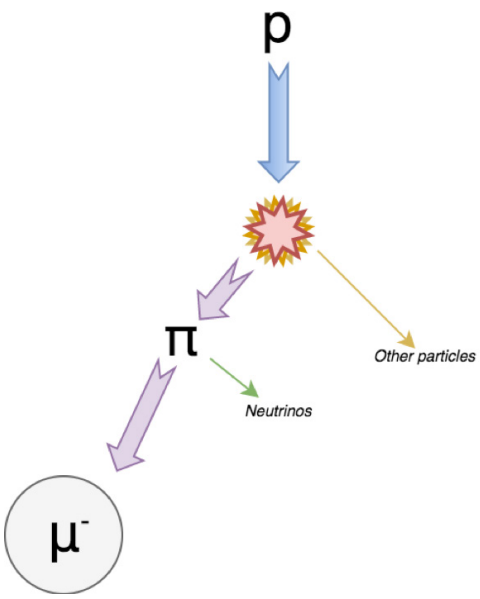


Figure 1. Cosmic ray showers. The collision of a cosmic ray proton (p) into an atmospheric nuclide produces pions (π). Pions, which are very unstable, quickly decay into muons (μ^-), which can be detected at the Earth's surface.

hits for single muons that have passed through two, three, or all four of the counters as two-fold, three-fold, or four-fold coincidences to allow for differential analysis of muons from specific directions or angles. It also allows for muon flux and lifetime data analysis. In a flux study, the rate of muon hits in a given area can be viewed over time, so that the effect of changes in the CRMD's surroundings can be detected. In a lifetime study, data on muons that decay within the scintillator plate is collected. Additionally, previously collected global data from other cosmic ray detectors can be accessed through the software. CRMD systems can have variations in their setup and hardware that result in inaccurate data, so calibration is required.

Purpose of this Study

Cosmic ray showers are often analyzed because they are natural examples of high-energy collisions often simulated in particle accelerators. Moreover, analysis of muon speed and lifetime gives researchers more evidence for the existence of time dilation, according to the theory of special relativity (6). Therefore, the purpose of this study was to identify and analyze the factors that influence muon flux and lifetime. We hypothesized that the positioning of scintillator plates and depth of water above such scintillator plates would affect muon flux and lifetime as an increase in matter in the path of a muon would slow or stop muons passing through it. In order to test this hypothesis, two experiments were conducted: 1) the liquid shielding flux experiment, and 2) the scintillator positioning lifetime experiment. CRMDs were used to conduct this study because these detectors enable the collection of data on muon presence and decay. The outcome of this

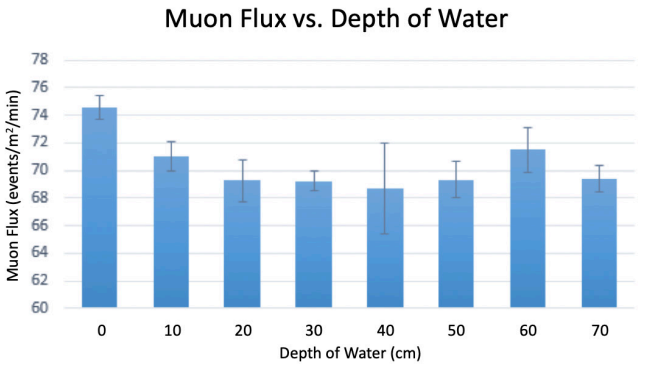


Figure 2. Liquid shielding flux experiment results. Muon flux measured as events/m^2/min was recorded for depths of water ranging from 0 to 70 cm. Error bars represent standard error of the mean.

research is significant because the information obtained from this experiment would enhance our understanding of the processes that can affect muon decay as well as detection. Additionally, this study aims to detect impacts of shielding at scales smaller than previous work. Understanding the path of muons after a collision allows researchers to ascertain the characteristics of these particles and processes that occur during the collision.

RESULTS

Muon Flux Decreases with an Increased Depth of Water

To measure the effect of water depth on muon flux, flux data was collected at water depths from 0 to 70 cm. Muon flux decreased as the depth of the water in the bins increased, particularly when the depth was between 0 and 30 cm. However, higher variability in the latter half of the experiment (depths greater than 30 cm.) did not reflect the trend observed with lower depths of water. Nevertheless, a trend exists in the first 30 cm. of depth, implying muon flux decreases as water depth increases at this depth range (**Figure 2**).

Detection of Muon Decay Increases with Scintillator Plate Orientation Height

To determine the optimal orientation of scintillator plates for data collection, lifetime data at a one-fold coincidence was collected for each of three distinct orientations; the plates were set on their long edge, on their short edge, or flat (**Figure 3**). When the software collects lifetime data, it only counts

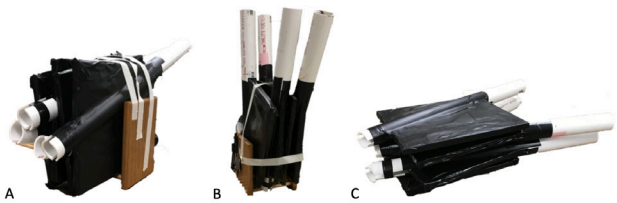


Figure 3. Scintillator positioning experimental setup. The three images present the actual orientations of the scintillator plates. It is important to note that position C has the shortest height, B the tallest, and A has an intermediate height.

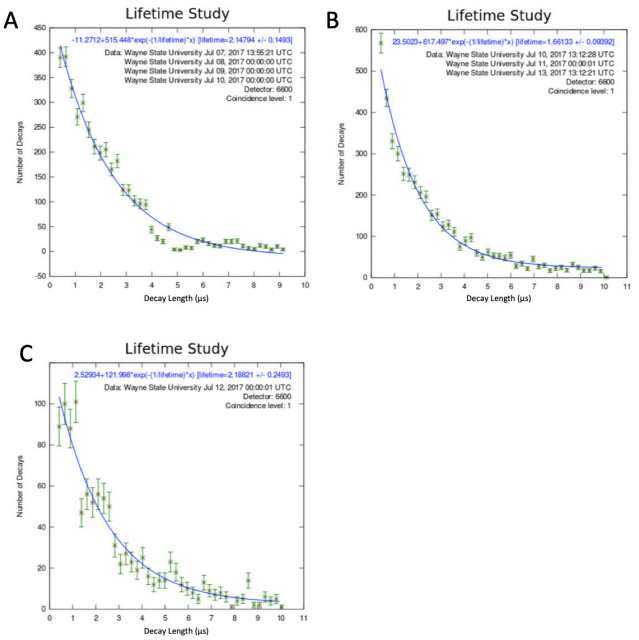


Figure 4. Scintillator positioning lifetime experiment results. Muon lifetime measured in μs was recorded for each scintillator plate orientation. Error bars represent standard error of the mean.

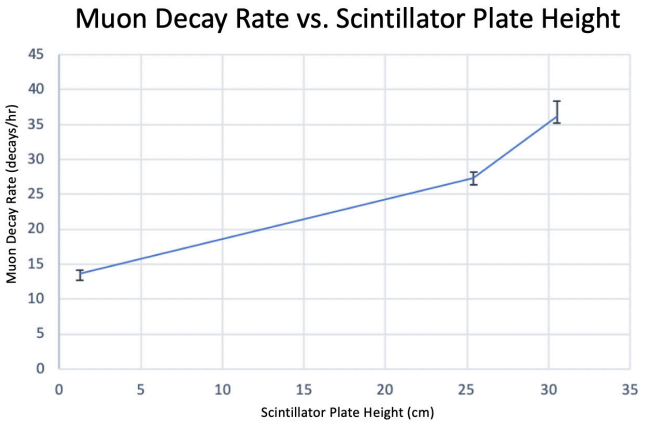


Figure 5. Scintillator positioning experiment synopsis. Muon decay rate in decays/hr. was calculated for each scintillator plate orientation corresponding to a vertical height in cm. Error bars represent standard error of the mean.

decay events that occur within a scintillator plate. The graph of lifetime data generated can then be integrated to provide a count of total hits during the data collection period for that orientation, and then a rate of hits for that orientation can be generated.

In the first, flat orientation of scintillator plates, we recorded 984 lifetime hits over a 72-hour timespan, so the rate of detected muon decay in this setup was 13.7 decays per hour. The standard error of the mean calculated for this position was 0.44 decays per hour (**Figure 4C**). In the second orientation, with the scintillator plates stacked on the long edge, we recorded 1258 lifetime hits over a 46-hour timespan, or 27.3 decays per hour, with a calculated standard error of the mean of 0.77 decays per hour (**Figure 4A**). In the final orientation, with the counters stacked on the short edge, we recorded 289 lifetime hits over a time span of 8 hours, so the rate of decay was 36.2 decays per hour. The calculated standard error of the mean in this orientation was 2.1 decays per hour (**Figure 4B**).

This data demonstrates an upward trend that corresponds to the vertical height of the scintillator plate orientations (**Figure 5**). When the orientation results in a greater distance between the top and the bottom of the scintillator plates, more muons can decay in the scintillator plates. Then, these decays can be measured.

DISCUSSION

We hypothesized that greater vertical height from scintillator positioning and less shielding would increase muon decay detection rate and muon flux.

In the liquid shielding flux experiment, we demonstrate a measurable effect of low-depth water shielding on muon flux. However, this relationship was not observed at depths greater than 30 cm. Confounding variables that may have led to this include power outage-associated voltage fluctuations that occurred on the day of data collection for the 40 cm depth, as well as on subsequent days, as changes in voltage affect the calibration and thus overall data collection integrity of the scintillator plates. Despite this potential inconsistency, the data in the initial periods of the experiment suggest measurable water shielding of muons. We know that this miscalibration was not present for the water depths below 40 cm because the data was collected sequentially; first 0 cm, then 10 cm, and so forth until 70 cm. However, if we presume that the calibration was unaffected by the power outage, our data suggests an initial decrease in muon flux when any depth of water is present but does not suggest a clear trend relating water depth and muon flux. Therefore, the shielding experiment should be repeated without confounding variables such as voltage fluctuations to ensure that the results are reproducible.

Researchers have analyzed muon flux, comparing seafloor readings underwater to surface readings (7). Additionally, the impact of shallower depths of water on muon flux have also been investigated (8). However, this present research shows that shielding from water can be detected at markedly smaller scales. Moreover, if water shielding could selectively block low-energy muons, our findings will be of value because researchers working on particle accelerators can selectively analyze high-energy muons, allowing for more specific insight during collisions and more specialized research into higher-energy particles because lower-energy muons may be shielded by water if implemented. Therefore, additional studies are needed to determine whether water shielding can selectively block lower-energy muons, or non-selectively blocks muons of any energy value.

In the scintillator positioning lifetime experiment, the observed rate of muon decay increased substantially when the vertical height of the scintillator plates increased. There

was a 100% increase in the rate of muon decay when the scintillator plates were placed on their long edges, then an additional 32.2% increase as they were placed on their short edges. This is likely due to the fact that more muons approach Earth's surface vertically rather than horizontally (9). If a muon approaches horizontally, or at an angle, it has to travel through more atmospheric matter, which in turn slows down the muon, and therefore, it decays before reaching the surface (10). When muons descend vertically, they travel through less atmosphere, and therefore they maintain more of their velocity. As a result, incoming muons would be most likely to enter a scintillator plate from the top, vertically. Because of this, scintillator plates with an orientation that lends to a larger vertical height will provide denser matter for muons to decay in. Understanding this concept, based on our experimental results, provides new insights for better design of future muon lifetime detectors, and other future applications of muons.

Future studies can build on the research conducted in this study. For example, more data is necessary for depths of 40 cm and greater in the water shielding flux experiment to analyze the impact of greater depths. Additionally, liquids other than water or solids could be analyzed using a similar experimental design. Factors such as density could be analyzed to further confirm the theorized factors that affect muon velocity through matter. Furthermore, setting up additional quantitative heights in the scintillator positioning experiment would benefit future studies. The conducted study compared the heights of the scintillator plates in reference to each other, in three distinct heights. The scintillator positioning experiment can also be further expanded by incorporating new angles and setting up coincidences so that muon flux and lifetime can be analyzed from the zenith to the horizon, and possible cosmic ray muon generation in the water can also be accounted for.

The findings of this study confirmed our hypothesis that muons either slow down or remain in the detector for longer periods and decay when they have more matter to travel through. This study is novel because the effects of the shielding are noticeable at much smaller scales than those applied in previous studies. The results of the water shielding experiment suggested that water can be used to decrease the muon flux by allowing for greater decay. The results of the scintillator positioning lifetime experiment suggested that the height of the scintillator in its orientation is a potential determinant of the volume of data collected for muon decay studies.

METHODS

Cosmic Ray Muon Detector Setup

This study used the QuarkNet Model 6000 Cosmic Ray Muon Detector and was conducted in a university physics research laboratory (5). Four counters, each comprised of a scintillator plate and a photomultiplier tube (PMT), were used in the setup of the detector for both experiments. Black tape was used to fully cover each scintillator plate. The orientation and position of the counters varied depending on the specific needs of the study. Each of the four counters were wired to a data acquisition (DAQ) board. This board registered and counted all the data incoming from the four counters that were set up. The information was sent to a connected computer, in which the data was analyzed using specialized software from QuarkNet, generating readable flux and lifetime data (5).

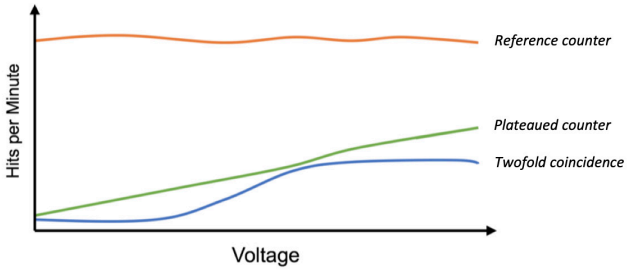


Figure 6. Plateauing graph. This graph shows, theoretically, how a typical plateauing curve should appear. The orange trace represents the reference counter. The green trace represents the counter being plateaued, which increases with voltage, while the reference counter remains at the same level. The blue trace represents the twofold coincidence consisting of the two aforementioned counters.

Cosmic Ray Muon Detector Calibration

Appropriate calibration of the CRMD was conducted so that a hit from a muon was neither over-represented nor under-represented in the generated electrical signal. The optimal voltage applied to each counter was determined by graphing the hits per minute of each counter versus a reference counter.

The PMTs were linked to a device known as the power distribution unit (PDU). The PDU allowed users to check the voltage of a counter with a voltmeter and alter the voltage if need be.

A plateauing method was used to calibrate each counter. One reference counter was set to a voltage that generated about 40-60 hits per second. The counter to be calibrated was stacked under the reference counter, and a two-fold coincidence was set. Data was then collected for about two minutes at a time, with a gradually increasing voltage in each period. The hit rates of the individual counters and the two-fold coincidence rates were graphed against voltage. In the graph, the reference rate stayed constant, much higher the other two traces in the graph. The two-fold coincidence trace produced a logistic function, and eventually plateaued. The individual trace of the counter being calibrated rose linearly, similar to the rate of the two-fold coincidence, until the two-fold coincidence graph plateaued. The voltage where the plateaued counter's graph diverged from the two-fold coincidence graph is the optimal voltage for the counter (**Figure 6**).

Plateauing was completed for the calibration of all four counters in this study. For counters 0, 1, 2, and 3, the values of optimal voltage were 770, 850, 820, and 910 mV, respectively. Counters 0, 1, 2, and 3 in the QuarkNet hardware correspond to counters 1, 2, 3, and 4 in the QuarkNet software, respectively.

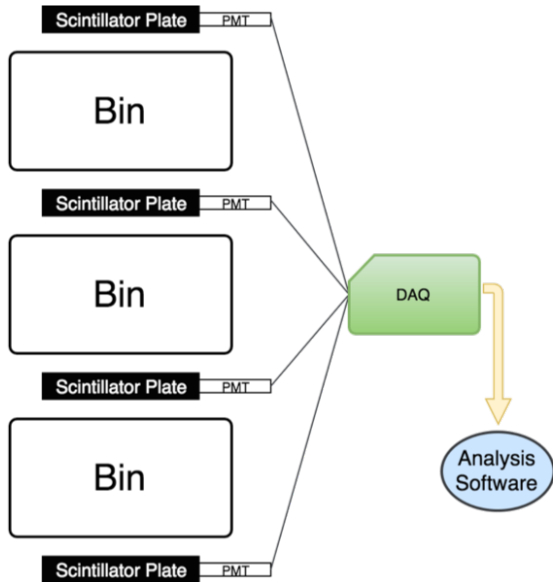


Figure 7. Liquid shielding flux experiment schematic. This diagram shows the setup of the counters and bins used to hold the water in the experiment. The counters were wired to the DAQ, which registered muon hits and compiled the raw data. That data was then uploaded and analyzed by QuarkNet data analysis softwares.

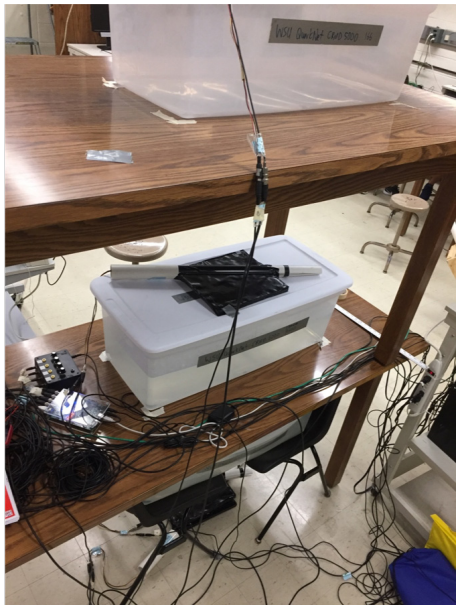


Figure 8. Detector positioning for shielding experiment. The image illustrates the actual setup of the liquid shielding flux experiment. Wire lengths and GPS positioning of each counter is entered into the data analysis software, as they affect the timing of signal reception from counter to DAQ.

Liquid Shielding Flux Experiment

The flux study was set up with four counters in a stacked configuration (11), with bins to hold water between each of them (**Figures 7 and 8**). First, a scintillator plate (counter) was placed flat on the ground. A bin that can hold a depth of up to 26 cm of water was placed 20 cm above the counter. This was repeated to have three bins, with four counters; one above the bins, one below the bins, and two in between the bins. As such, the bins can be filled to a total depth of 78 cm. As this study was set to a four-fold coincidence, only muons that passed through all four counters were counted during each data collection period. This was so that any detected muons are muons that have passed through the designated water depth, ensuring the validity of data collected in this study.

Control runs were conducted using the same configuration with no water. Data was collected for 0, 10, 20, 30, 40, 50, 60, and 70 cm of cumulative water depth. For each depth, data was collected for at least four hours.

Scintillator Positioning Lifetime Experiment

Muon lifetime data was collected when the paddles were positioned flat, in a stacked orientation (**Figure 3C**). A one-fold coincidence was applied to this setup. Data was collected for 72 hours. The length of time for which the data was collected does not impact the quantitative values of the muon lifetime data.

Muon lifetime data was then collected when the paddles were positioned at the longer edge, stacked side-by-side (**Figure 3A**). A one-fold coincidence was applied to this setup. Data was collected for 46 hours.

In the final orientation of the counters, the paddles were positioned on the shorter edge, stacked side-by-side, when the muon flux data was collected (**Figure 3B**). A one-fold coincidence was also applied to this configuration. Data was collected for eight hours.

Received: August 25, 2019
Accepted: May 11, 2020
Published: May --, 2020

REFERENCES

1. Aurenche, Patrick. “The Standard Model of particle physics.” *Paris-Sud* University, 1997, pp. 1-25.
2. Street, J. C., and E. Stevenson. “New Evidence for the Existence of a Particle of Mass Intermediate Between the Proton and Electron.” *Physical Review*, vol. 52, no. 9, 1937, pp. 1003-4.
3. Gribov, V. and B. M. Pentecorvo. “Neutrino astronomy and lepton charge.” *Physics Letters B*, vol. 28, no. 7, 1969, pp. 493-6.
4. Dzikowski, T., *et al*. “Possible mechanisms for the origin of the cosmic-ray showers coming from the direction of the Crab pulsar.” *Journal of Physics G: Nuclear Physics*, vol. 9, no. 4, 1983, pp. 459-65.
5. Rylander, J., *et al*. “Cosmic Ray Muon Detector User’s Manual.” *QuarkNet*, 2010, pp. 2-63.
6. Lämmerzahl, C. “Special relativity: A matter of time.” *Nature Physics*, vol. 3, no. 12, 2007, pp. 831-2.

7. Aiello, S., *et al*. "Measurement of the atmospheric muon flux with the NEMO Phase-1 detector." *Astroparticle Physics*, vol. 33, no. 4, 2010, pp. 263-73.
8. Aguayo, E., *et al*. "Cosmic Ray Interactions in Shielding Materials." *U.S. Department of Energy*, 2011, pp. 42-50.
9. Patrignani, C., *et al*. "Review of Particle Physics" *Chinese Physics C*, vol. 40, no. 10, 2016, pp. 421-9.
10. Procureur, S. "Muon imaging: Principles, technologies and applications." *Nuclear Instruments and Methods in Physics Research*, vol. 878, 2018, pp. 169-79.
11. Aartsen, M.G., *et al*. "Characterization of the atmospheric muon flux in IceCube." *Astroparticle Physics*, vol. 78, 2016, pp. 1-27.

JOURNAL OF
EMERGING INVESTIGATORS

Article

Antibacterial Effects of Copper Surfaces

Srimayi Mulukutla, Ronald Kinser
Sewickley Academy High School, Sewickley, Pennsylvania

SUMMARY

Healthcare-associated infections (HAIs) are a major problem affecting 2 million people and cost $30 billion annually. However, most HAIs are preventable. Ancient civilizations used copper to help purify water and treat wounds. During the cholera epidemic, copper workers were not as affected which suggests that copper might be a bacteria-fighting element. This study examines the ability of copper and copper alloy surfaces to inhibit bacterial growth. We developed three operating hypotheses. First, copper surfaces will inhibit bacterial growth more than aluminum. Second, a copper alloy surface, brass, will inhibit the growth of bacteria more than aluminum but to a lesser extent than copper. Third, the longer the bacteria are in contact with a copper or brass surface, the greater the extent of inhibition of bacterial growth. Two non-pathogenic strains of bacteria, *Escherichia coli* (*E. coli*) and *Staphylococcus epidermidis* (*S. epidermidis*), were exposed to different metal plates, copper, brass, and aluminum, for varying degrees of time. The bacteria were then transferred to an agar plate to allow the bacteria to grow. Copper surfaces significantly inhibited the growth of *E. coli* and *S. epidermidis* more than the aluminum surfaces. Brass significantly inhibited the growth of bacteria more than aluminum, but to a lesser extent than copper. The longer the bacteria were in contact with brass or copper, the greater the extent of inhibition of bacterial growth. Overall, copper, and to some extent brass, may be good options to help prevent bacterial growth and to prevent HAIs in healthcare settings.

INTRODUCTION

Healthcare-associated infections (HAIs) are infections that patients obtain while receiving medical care in a hospital environment. They annually affect almost 2 million people and cost the healthcare system approximately $28-45 billion each year (1, 2). They increase the length of hospitalizations and are associated with a higher mortality rate than cancer, traffic accidents, or HIV/AIDS (3, 4). Some of the most common bacteria for HAIs include Staphylococci, *Escherichia coli* and *Salmonella*, and Streptococci. These bacteria can cause various types of problems such as wound infections, food poisoning, urinary infections, and respiratory illnesses; however, most HAIs are preventable.

While long hospitalizations and poor nutrition contribute to HAIs, the most important cause of HAIs is poor hand hygiene (5, 6). HAIs are a result of multiple people, including patients and healthcare professionals, coming in contact with many common objects such as bathroom fixtures, intravenous poles, doorknobs, and bed railings. While handwashing has become imperative in hospitals, it is difficult to monitor (7). Therefore, it is crucial to find other solutions to kill the harmful bacteria that cause HAIs.

This study explores the possibilities of using copper to prevent bacterial growth, which may then be the basis in future experiments to prevent HAIs. Copper, one of the world's oldest metals used by humans, was used in ancient civilizations for drinking water, treating wounds, aiding skin conditions, and healing leg ulcers (8, 9). Interestingly, the cholera epidemic of the 1800s did not affect copper workers as much as others suggesting that copper may be a bacteria-fighting element (10). Copper may fight bacteria through a variety of mechanisms. One hypothesis is that copper ions come in contact with the bacteria and damage the bacterial cell wall (11). Another hypothesis is that copper causes the formation of reactive oxygen molecules that damage bacterial DNA (12). However, while the mechanisms may not fully be understood, the effect of a particular metal such as copper or brass in comparison to aluminum upon bacterial growth has not been thoroughly studied.

There are several important unanswered questions regarding copper's potential impact upon bacterial growth. While there have been studies suggesting that copper hinders the growth of bacteria, it is not clear how quickly copper surfaces can inhibit bacterial growth (13, 14). Additionally, it is unknown whether copper alloys have the same potential as copper in inhibiting the growth of the bacteria. This study has three main objectives. First, it will investigate the effects of copper surfaces on bacterial growth compared with a control group, an aluminum surface. Second, the study will examine the effectiveness of a copper alloy surface, brass, and its impact upon bacterial growth in comparison to aluminum and copper surfaces. Finally, the study also observes the effects that time has on the inhibition of bacterial growth when in contact with the three surfaces.

RESULTS

Visual Inspection of Agar Plates

The growth of bacteria on agar plates was assessed visually after the bacteria were incubated on the various metals for different durations of time. Visually, the 60-minute copper exposure reduced the bacterial growth more than the 15-minute exposure (**Figure 1**). This difference in bacterial

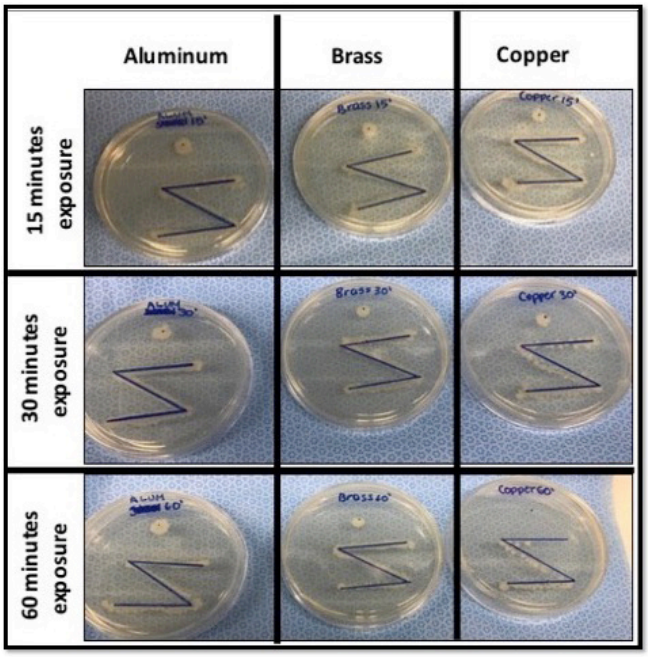


Figure 1: Representative images of the growth of bacteria on agar plates after the bacteria were exposed to various metals for varying durations of time. Note the varying amounts of growth based on which metal (aluminum, brass, or copper) the bacteria were on and the duration of time (15, 30, or 60 minutes) that the bacteria were on the different metals.

growth can be appreciated on the "Dot" as well as on the "Z-patterns." In addition, the growth of the bacteria is much thicker on the aluminum samples compared at all time exposures compared to the brass or copper samples. While visual inspection is useful, it is important to quantify the results. By measuring the diameter of the bacterial dot, we can estimate how many bacteria were transferred to the agar plates. The raw data for *E. coli* (**Table 1**) and *S. epidermidis* (**Table 2**) show that for both bacteria, the mean dot size was largest when bacteria were exposed to aluminum and smallest when exposed to copper.

Effect of Time of Contact with Metal Upon Bacterial Growth

The mean results for the growth of the *E. coli* bacteria on each of the metals for the various time points were calculated and plotted graphically (**Figure 2**). For the bacteria collected from the aluminum plates, there was no significant difference in growth at 15, 30, or 60 minutes (student's t-test, $p = 0.31$). On brass plates, the growth of the bacterial dot was significantly smaller after 60 minutes of contact compared with 15 minutes of contact (student's t-test, $p = 0.006$). Similarly, for copper, the growth of the bacterial dot was significantly smaller after 60 minutes of contact compared with 15 minutes (student's t-test, $p < 0.001$). Overall, we found that increasing contact time with brass and copper, but not aluminum, increasingly inhibited *E. coli* growth. The mean results for the growth of *S. epidermidis* on each of the metals for the various time

Trial number	Aluminum	Brass	Copper
	15-minute exposure period		
1	11.0	7.5	9.0
2	12.0	8.5	7.5
3	11.5	9.0	7.0
4	12.0	8.5	6.5
5	10.5	8.5	8.5
Average	11.4	8.4	7.7
SD	0.7	0.5	1.0
	30-minute exposure period		
1	9.5	9.5	7.0
2	10.5	8.5	7.5
3	12.0	8.0	7.0
4	12.5	10.0	6.5
5	10.5	8.5	7.0
Average	11.0	8.9	7.0
SD	1.2	0.8	0.4
	60-minute exposure period		
1	11.5	7.0	1.5
2	11.0	5.5	2.0
3	10.5	5.0	0.0
4	10.5	7.5	0.0
5	11.5	5.5	1.0
Average	11.0	6.1	0.9
SD	0.5	1.1	0.9

Table 1: Raw data for each of the five trials for each of the metals for every exposure time period for *E. coli* bacteria (diameter of dot after 72 hours, mm).

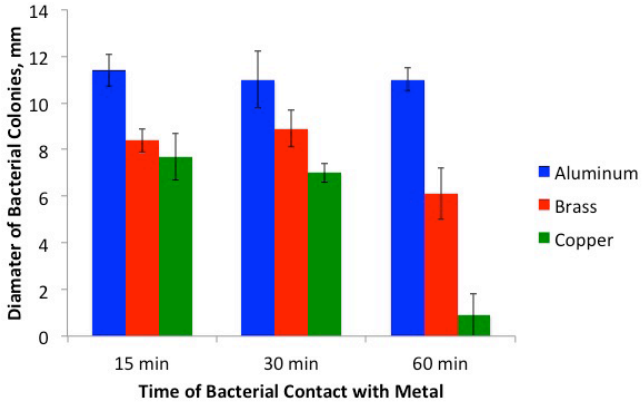


Figure 2: *E. coli* growth was inhibited at all time points after exposure to brass and copper, and greater inhibition occurred with longer exposure to brass and especially copper. *E. coli* growth was measured after 15, 30, or 60 minutes of exposure to aluminum (blue), brass (red), or copper (green). Growth was measured as the size (mm) of bacterial colonies. Error bars represent standard deviation of 5 trials.

Trial number	Aluminum	Brass	Copper
	15-minute exposure period		
1	8.5	7.5	7.5
2	10.0	7.0	8.0
3	8.5	7.0	6.0
4	8.0	7.5	6.0
5	8.0	7.0	7.5
Average	8.6	7.2	7.0
SD	0.8	0.3	0.9
	30-minute exposure period		
1	9.0	4.5	2.0
2	7.0	5.0	2.5
3	8.5	4.0	2.0
4	8.5	4.5	1.5
5	9.5	4.5	0.0
Average	8.5	4.5	1.6
SD	0.9	0.4	1.0
	60-minute exposure period		
1	8.5	3.5	0.5
2	8.0	3.0	0.0
3	9.0	5.0	0.0
4	9.0	3.0	0.5
5	7.5	3.0	0.0
Average	8.4	3.5	0.2
SD	0.7	0.9	0.3

Table 2: Raw data for each of the five trials for each of the metals for every exposure time period for *S. epidermidis* bacteria (diameter of dot after 72 hours, mm).

points were calculated and plotted graphically (**Figure 3**). For the bacteria collected from the aluminum plates, there was no significant difference in bacterial growth at 15, 30, or 60 minutes (student's *t*-test, $p = 0.68$). For brass, the growth of the bacterial dot was significantly smaller after 60 minutes of contact compared with 15 minutes of contact (student's *t*-test, $p < 0.001$). Similarly, for copper, the growth of the bacterial dot was significantly smaller after 60 minutes of contact compared with 15 minutes (student's *t*-test, $p < 0.001$). Overall, similar to E. coli, we found that increasing contact time with brass and copper, but not aluminum, increasingly inhibited *E. coli* growth.

Bacterial Growth on Different Metals

The same data can be evaluated in a slightly different way to assess differences in the growth of bacteria when in contact with the different metals. At the 15-minute time point, there was significantly less bacterial growth on copper and brass compared to aluminum (aluminum vs. brass or copper, student's *t*-test, $p < 0.001$), but the growth on the brass and copper was similar (brass versus copper, student's *t*-test, $p = 0.22$). At the 30- and 60-minute marks, there was less growth of bacteria on the copper compared to either aluminum or brass. For example, at the 60-minute time point, the growth on the copper was very low and significantly less than the growth on either brass or aluminum (student's *t*-test, $p < 0.001$). This data shows that brass did inhibit the growth of the *E. coli* compared to aluminum, but copper inhibited bacterial growth even more than brass. For *S. epidermidis*, at the 15-minute time point, there was significantly less bacterial growth on copper and brass compared with aluminum (aluminum vs. brass or copper, student's *t*-test, $p < 0.05$), but the growth on the brass and copper was similar (brass versus copper, student's *t*-test, $p = 0.66$). At the 30- and 60-minute marks, there was less growth of bacteria on the copper compared to either aluminum or brass. For example, at the 60-minute time point, the growth on the copper was very low and significantly less than the growth on either brass or aluminum (student's *t*-test, $p < 0.001$) (**Figure 3**). This data, similar to that of *E. coli*, shows that brass did inhibit the growth of the *S. epidermidis* compared to aluminum, but copper inhibited bacterial growth even more than brass.

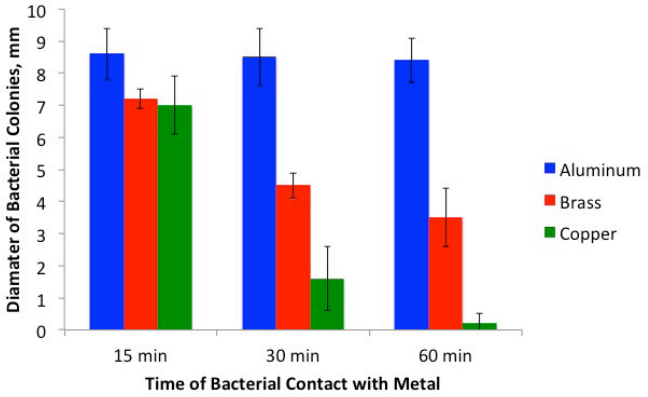


Figure 3: *S. epidermidis* growth was inhibited at all time points after exposure to brass and copper, and greater inhibition occurred with longer exposure to brass and especially copper. *S. epidermidis* growth was measured after 15, 30, or 60 minutes of exposure to aluminum (blue), brass (red), or copper (green). Growth was measured as the size (mm) of bacterial colonies. Error bars represent standard deviation of 5 trials.

DISCUSSION

This study investigated bacterial growth upon different metals. There were three major findings in this study. First, copper surfaces significantly inhibited the growth of *E. coli* and *S. epidermidis*, as compared to aluminum surfaces. Second, brass significantly inhibited the growth of bacteria more than aluminum, but to a lesser extent than copper. Finally, the study showed that the longer the bacteria were in contact with brass or copper, the greater the extent of bacterial growth inhibition. Furthermore, the bacterial growth was not affected based upon the duration of contact with the aluminum.

This study could have important implications, but this

essential knowledge may have been known for centuries. It has been used in India and China to cure various medical problems (15), and even today, it is used in the ancient art of Ayurvedic medicine (16). For instance, in ancient and present-day India, the practice of Ayurvedic medicine has advocated for drinking water from copper vessels because copper kills harmful bacteria. The use of copper for drinking water has been in practice for thousands of years. Additionally, there is evidence of its use by the Greeks and Aztecs in treating wound infections and ear infections (17). In these cultures, copper was applied to wounds to prevent infections or the spread of infections. In actual hospital settings, there is some evidence to support the hypothesis that HAIs can be decreased with copper surfaces (18-20).

This study is important because it raises awareness of an enormous problem around the world. Most healthcare facilities do not use copper surfaces despite the benefits. One reason may be the higher perceived costs, although the reduction in HAIs would likely offset that expense. In addition, because much of the hospital infrastructure is already built, replacing everything with copper may be time-consuming and expensive. In addition, not everyone accepts that copper may truly reduce infections and there are no long-term studies about the efficacy of copper to prevent HAIs. Therefore, more data are needed in practical clinical settings.

An important additional element from this analysis is that there is relatively limited data about the impact of time of bacterial contact with bacterial growth. Significantly, the bacterial growth was not inhibited immediately. Our observation that bacterial growth was most limited after the bacteria were in contact with the surface for a longer duration of time might be important since HAIs may still occur even with copper surfaces if the bacteria were only on the surface for a short duration. Still, it is encouraging that if the bacteria were in contact for longer durations, bacterial growth was significantly lessened.

There are multiple potential surfaces in a hospital setting that may be sources for HAI, including computer keyboards, computer screens, wooden railings, among others. Aluminum is a common and prevalent surface in hospitals because it is on bed railings, doorknobs, intravenous infusion pump poles, and sinks. Future studies may focus on alternatives to the surfaces that were not studied in this experiment.

There were a few limitations to this study. First, only two common types of bacteria were tested, and disease-causing bacteria from hospitals were not tested. Second, only the aluminum surface was used as a control. In making the bacterial "dot" on the agar plate, the method used was to transfer the bacteria using a sterile swab. This technique may have resulted in some differences in the amount of bacterial transfer based on how hard the swab was pressed. However, given the results showed relatively little variation, the results likely can be trusted. Finally, counting the number of colony-forming units (CFUs) was initially planned but was impossible to measure given the density of bacterial growth. Therefore, it was decided to use the diameter of bacterial growth instead of the number of colony-forming units. Typically, CFUs are measured through a series of bacterial sample dilutions, which would have helped define a more precise number of individual bacteria. Quantitatively, the dilution strategy would have been more accurate in assessing differences across the metals. In the method used in this experiment, the bacterial spread in the "dot" may have been inhibited by growth limits, and this could have potentially underestimated the amount of bacterial growth. Counting actual CFUs could have provided a more quantitative insight.

In conclusion, copper and brass are effective alternatives to aluminum to help prevent bacterial growth and to prevent HAIs in healthcare settings. These findings suggest that hospitals should further investigate this issue in order to promote the safety of patients and healthcare workers.

MATERIALS AND METHODS

Preparing the Bacteria

Cultures of live, non-pathogenic, freeze-dried strains of *E. coli* and *S. epidermidis* were purchased from homesciencetools.com. The bacteria arrived in the form of pellets and were kept frozen. When ready for experimentation, each pellet was dissolved in 10 mL of tryptic soy broth (homesciencetools.com) and incubated for 2 hours at 37 °C.

Placing Bacteria on Metal Plates

Metal plates (5.4x8.6 cm) each made of aluminum, brass, or copper were laid out across the laboratory bench. For each trial, a total of six plates of each type of metal were used, with three being used for E. coli and three for S. epidermidis. For each bacterial strain, one plate was used for the 15-minute exposure, one for the 30-minute exposure, and one for the 60-minute exposure. All plates were cleaned with 70% ethanol and allowed to air dry. To ensure even spreading, 500 uL of the appropriate bacterial solution was pipetted onto each plate and spread evenly across the plate. The bacteria were allowed to stay on the metal plate for the allotted time (i.e., 15, 30, or 60 minutes) prior to transferring the bacteria to the agar plates.

Plating Bacteria onto Agar Plates

Since bacteria cannot be seen on the metal plates, they must be transferred to agar plates so that we can measure the number of bacteria that were present on the metal. Agar plates were purchased from homesicencetools.com and were first prepared prior to experimentation by labeling the plates with the bacterial type, metal type, and the time of bacterial contact with the metal. After the pre-planned duration of exposure with the metal plate, a sterile swab was swiped across a 5 cm length of the metal plate and then transferred to the agar plate in a "Z-pattern." Another sterile swab was pressed onto the metal plate and the bacteria were transferred to the agar plate in a "dot" form at the top of the agar plate. All plates were placed in an incubator at 37 °C for a total of 72

hours, with measurements made at every 24-hour interval.

Measuring Bacterial Growth

Day 0 was counted as the date the bacteria were plated onto the agar plates. On days 1-3, visual assessment of bacterial growth density was assessed on the "Z-pattern." The diameter of the bacterial "dot" was measured. All data were recorded onto an Excel spreadsheet.

Statistical Analyses

Tests were repeated for a total of five trials. Means and standard deviations were calculated for each metal (i.e., aluminum, brass, and copper) at each time interval (i.e., 15, 30, and 60 minutes). Student's *t*-test was used to compare means between groups.

Received: December 18, 2019
Accepted: March 18, 2020
Published: May 19, 2020

REFERENCES

1. Stone, Patricia W. "Economic Burden of Healthcare-Associated Infections: an American Perspective." *Expert Review of Pharmacoeconomics & Outcomes Research*, vol. 9, no. 5, 2009, pp. 417–422., doi:10.1586/erp.09.53.
2. Klevens R. Monina, et al. "Estimating Health Care-Associated Infections and Deaths in U.S. Hospitals, 2002." *Public Health Reports*, vol. 122, no. 2, 2007, pp. 160–166., doi:10.1177/003335490712200205.
3. Haque, Mainul, et al. "Health Care-Associated Infections; an Overview." *Infection and Drug Resistance*, Volume 11, 2018, pp. 2321–2333., doi:10.2147/idr.s177247.
4. Anderson, Robert N. "Deaths: Leading Causes for 1999." *National Vital Statistics Report*, vol 49, 2001, pp. 1-87.
5. McLaws, Mary-Louise. "The relationship between hand hygiene and health care-associated infection: it's complicated." Infection and drug resistance vol. 8 7-18. 29 Jan. 2015, doi:10.2147/IDR.S62704.
6. WHO. "WHO Guidelines on Hand Hygiene in Health Care: First Global Patient Safety Challenge Clean Care is Safer Care." Geneva: World Health Organization; 2009.
7. Carling, Philip C., and Judene M. Bartley. "Evaluating Hygienic Cleaning in Health Care Settings: What You Do Not Know Can Harm Your Patients." *American Journal of Infection Control*, vol. 38, no. 5, 2010, doi:10.1016/j.ajic.2010.03.004.
8. Borkow, Gadi. "Using Copper to Improve the Well-Being of the Skin." *Current Chemical Biology*, vol. 8, no. 2, June 2015, pp. 89–102., doi:10.2174/2212796809666150227223857.
9. Sudha, V. B. Preethi, et al. "Storing drinking-water in copper pots kills contaminating diarrhoeagenic bacteria." *Journal of Health, Population, and Nutrition*, vol. 30,1 (2012): 17-21. doi:10.3329/jhpn.v30i1.11271.
10. Reddaway, Rebecca. "Metals in Medicine and the environment: copper the anti-inflammatory healer." *Metallurgy 101 (by Popular Request)* - Faculty.virginia.edu. faculty.virginia.edu/skrutskie/astr512.s06/pdf/lec07.pdf, 2009.
11. Villapún, Victor M., et al. "Antibacterial Metallic Touch Surfaces." *Materials (Basel, Switzerland)*, vol. 9,9 736. 29 Aug. 2016, doi:10.3390/ma9090736.
12. Grass, Gregor, et al. "Metallic copper as an antimicrobial surface." *Applied and Environmental Microbiology*, vol. 77, no. 5, 2011, pp. 1541-7. doi:10.1128/AEM.02766-10.
13. Casey Anna L., et al. "Role of copper in reducing hospital environment contamination." *Journal of Hospital Infection*, vol. 74, no. 1, 2010, pp. 72– 77. doi: 10.1016/j.jhin.2009.08.018.
14. Noyce, J. O. et al. "Use of Copper Cast Alloys To Control Escherichia Coli O157 Cross-Contamination during Food Processing." *Applied and Environmental Microbiology*, vol. 72, no. 6, Jan. 2006, pp. 4239–4244. doi:10.1128/aem.02532-05.
15. Keevil, Bill. "Hospitals should learn about copper's unsuspected health properties." *Newsweek*. Feb 27, 2017.
16. Galib, Ruknuddin, et al. "Therapeutic Potentials of Metals in Ancient India: A Review through Charaka Samhita." *Journal of Ayurveda and Integrative Medicine*, vol. 2, no. 2, 2011, p. 55, doi:10.4103/0975-9476.82523.
17. Keevil, Bill. "Copper is great at killing superbugs – so why don't hospitals use it?" *The Conversation*. Feb 24, 2017.
18. Salgado, Cassandra D., et al. "Copper Surfaces Reduce the Rate of Healthcare-Acquired Infections in the Intensive Care Unit." *Infection Control & Hospital Epidemiology*, vol. 34, no. 5, 2013, pp. 479–486., doi:10.1086/670207.
19. Schmidt, Michael G., et al. "Copper Continuously Limits the Concentration of Bacteria Resident on Bed Rails within the Intensive Care Unit." *Infection Control & Hospital Epidemiology*, vol. 34, no. 5, 2013, pp. 530–533., doi:10.1086/670224.
20. Dessauer, Bettina Von, et al. "Potential Effectiveness of Copper Surfaces in Reducing Health Care–Associated Infection Rates in a Pediatric Intensive and Intermediate Care Unit: A Nonrandomized Controlled Trial." *American Journal of Infection Control*, vol. 44, no. 8, 2016, doi:10.1016/j.ajic.2016.03.053.

JOURNAL OF
EMERGING INVESTIGATORS

Article

A Novel Model to Predict a Book's Success in the New York Times Best Sellers List

Matthew Lee[1], Siddhant Arora[2], Johann Lee[3], and Rohan Vaidya[4]

[1]San Marino High School, San Marino, CA , [2]Clements High School, Sugar Land, TX, [3]The Lawrenceville School, Lawrenceville, NJ, [4]Dougherty Valley High School, San Ramon, CA

SUMMARY

The popularity of media, such as books and music, has historically been considered difficult to forecast. This popularity is important in determining the success that can be achieved once the media is published. Therefore, we aim to evaluate the extent to which this fact holds true, as we propose that these public opinion trends are quite deterministic in nature. We investigated the important non-textual attributes determining a book's popularity, including but not limited to a book's previous ranking in the New York Times Best Sellers list and its popularity in searches. We then constructed two models: a generalized classifier of a successful Best Seller and a predictor of a book's weekly rank on the New York Times Best Sellers list. The reasonable accuracy of our classification and regression models suggest that book popularity is indeed deterministic. These findings point towards definitive characteristics that can help creators produce successful works.

INTRODUCTION

Even with the advent of technology, reading remains a popular pastime for Americans as the average American reads 12 books every year. Last year, 695 million books were sold in the United States alone (1). Even with so many books to read, only a select few become extremely popular. The exclusivity of book popularity made us wonder whether it is possible to predict which books will be popular. In the past, studies have generated algorithms to predict the popularity of other media such as research papers and songs with a high degree of precision.

FutureRank, an algorithm that uses citation information, authors, and publication time to predict the future ranking of a scientific paper, was able to rank the first 25% of retrieved articles with 100% precision (2). Another model plotted the path a song takes through the Billboard Top 100 by predicting the $n+1^{th}$ rank from the nth previous rankings with very low error (3).

Given the success in predicting popularity in other forms of media, it is likely that by considering the right attributes, book popularity can be predicted as well. Several of these attributes have already been highlighted in past research. Six literary features – action, measure of emotion, personalities of major characters, themes, romaticization, and simplicity – can determine what makes a best-selling novel to an accuracy of 82% (4). Characteristics like the season published or genre, which intuitively appear crucial to book success, do indeed play a large role in a book's New York Times Best Seller list rank; however, other more subtle properties like initial rank also heavily influence the book's future ranks (5). Such features have been used to manually classify Best Sellers against non-Best Sellers with an accuracy of 86% (6).

Model	Accuracy
Random Forest	98.4560570%
SVM (Linear Kernel)	92.3396674%
Naive Bayes	90.1425178%
SVM (Sigmoid Kernel)	71.6745843%

Table 1. Accuracy of the different models tested.

Despite identification of contributing features, there has yet to be a model that predicts a book's success. Based on past studies, we believed that the various characteristics of a book have a strong correlation to its success and popularity. We hypothesized that if a book hits the New York Times Best Sellers, its relative weekly performance will strongly correlate with future performance. Therefore, our research was divided into two subsections: the classification model and the regression model. Our classification model determined whether a specified book would make the New York Times Best Sellers or not. If the book did indeed make it, our regression model predicted the path of the book's weekly ranking. We believe this two-fold method best accounted for how most books never make it onto the New York Times Best Sellers. Our classifier and predictor are the first to focus specifically on the success of literature based on nonliterary features. Our models are also a proof of concept that prediction of the New York Times Best Sellers is possible.

RESULTS

We built a classifier to classify Best Sellers from non-Best Sellers and a forecasting model to predict the path of the book on the New York Times Best Sellers in order to determine what non-literary features impacted popularity most. For the classification model, we used non-literary characteristics of a novel as features for several classification algorithms. For the forecasting model, we used three time series as features for a time-series based regression model. Our example book was *Winter of the World* (ISBN:0525952926), a New York Times Best Seller.

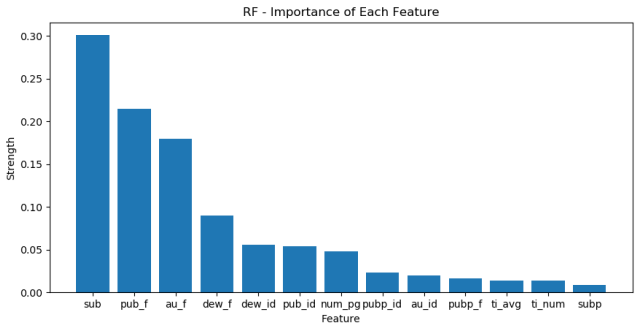


Figure 1. Random Forest Model-Relative Importance Scale and Visualization. The strength of each factor in the random forest model (a measure of how often the factor was considered in the model's nodes). To train this model, 6,000 books were used. The features are from left to right: subject, publisher frequency, author frequency, Dewey Decimal number frequency, Dewey Decimal number, publisher, number of pages, publishing location, author, publish place frequency, title average word length, title number of words, and subject place.

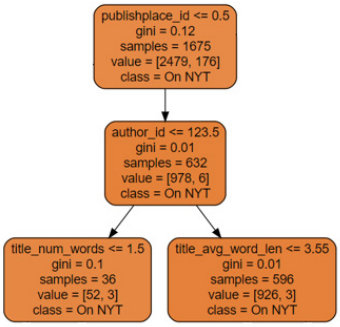


Figure 2. Sub-section of the RF Model Decision Tree. Decision tree framework can be represented using a graph-like model of yes/no statements. The RF model used in this paper contained several hundreds of nodes.

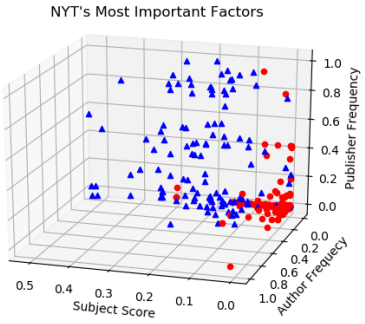


Figure 3. 3-D Plot of the Most Important Features. Reducing the k-dimension data to 3, this 3-dimensional plot constructs a visualization of the high separability between books that made the list and books that did not.

Classification

Winter of the World was correctly classified as a Best Seller. Several models were used to classify this book. Random forest classification with 100 estimators seemed to produce the best fit to the training data with an accuracy of around 98.46%, indicating that of the 1684 testing entries, only 29 were misclassified (**Table 1**). Each classification model utilized all of the 13 features: subject score, publisher frequency, author frequency, Dewey Decimal frequency, Dewey Decimal ID, publisher ID, number of pages, publish place ID, author ID, publishing place frequency, title average word length, title number of words, and subject place score. Frequencies measured the occurrence of their respective categories among all the books surveyed. Other IDs and scores represented features of a book in a numerical format. The relative importance of each of these features (how much they impacted a book's classification) was determined (**Figure 1**). A portion of our Random Forest Decision Tree nodes and the if-statements that determine how a book will be classified was also generated (**Figure 2**).

Surprisingly, we found that the textual features we considered, title average word length and title number of words, played little effect on whether a book was classified as a Best Seller or not. Non-literary features dominated in terms of relative importance (**Figure 1**). Therefore, we took the three most important features (subject score, publisher frequency, and author frequency) and constructed a 3D plot to better visualize the distinction between Best Sellers and non-Best Sellers (**Figure 3**). Finally, a confusion (error) matrix was presented to view the false positive/negative rate of the random forest model (**Figure 4**). Type I and II error rates for the RF model were quite low, with approximately 50 false positive and 200 false negatives out of the entire 6,000 book dataset.

Regression

We defined the best model as the model with the lowest cumulative mean absolute error (MAE) at week 10. The four models we tested were determined by the different combinations of (m, k) we considered. Each model considered past rank, weekly percent genre (the ratio of the number of books in that book's genre to the total number of books), and Google Search Ranks index.

The MAEs of the various possible combinations per week were plotted (**Figure 5**). An (m, k) of (2, 2) consistently has the lowest MAE at week 10 (4.449), making it the best (m, k) by our metric. By running our model with various combinations of features, we generated three different predicted paths (**Figure 6**). Overall, considering all three features provided the best fit prediction.

We ran a Student's t-test on the β coefficients (2, 2) model to determine whether our features had a statistically significant relationship with the ranks and generated a p-value of $3.55 * 10^{-8}$. Given that our value was much lower than the commonly accepted alpha of 0.001, we rejected the null hypothesis, allowing us to conclude that there was a relationship between previous rank, genre, and Google Trends index with current rank.

DISCUSSION

For all models, we accounted for overfitting by performing cross validation. The sample size of our data was small enough that overfitting was possible, so cross validation

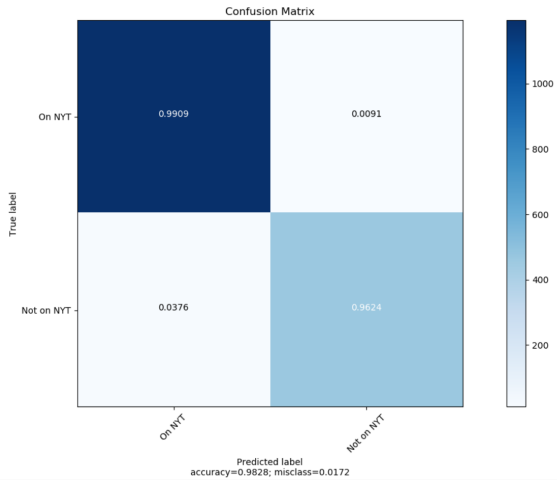


Figure 4. Confusion (Error) Matrix. Confusion matrix of the random forest model. Rows represent the true label and columns represented the predicted label. Color intensity represents number of books in that category.

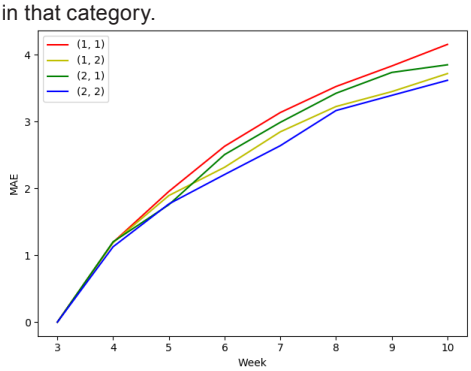


Figure 5. MAE per Week. Mean absolute error of our training model while considering multiple combinations for the degrees of both the regression and time series models. The x-axis represents the week on the Best Sellers and the y-axis represents of the mean absolute error of the ranks predicted.

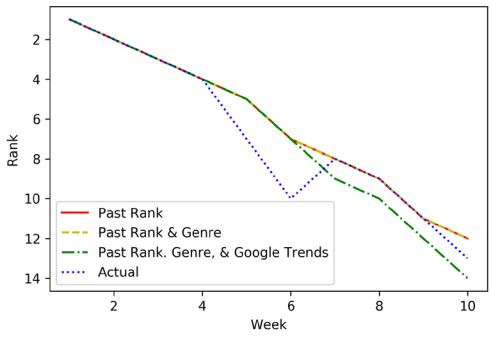


Figure 6. Actual vs Predicted Path for Winter of the World. An example of the predicted path against actual path for Winter of the World, taken with past rank only, past rank and genre, and past rank, genre, and Google Trends. The x-axis represents the week on the Best Sellers and the y-axis represents the rank predicted on the Best Sellers list.

ensured that our models were still able to generalize. We implemented five-fold cross validation, which split our dataset into five randomly assigned groups. Each unique group was used as a validation data set while the model trained on the other four. The model was then evaluated based on prediction performance on the validation data set. From this process, we were able to determine how extensively we needed to train to optimize error and generalization capabilities. By using classification models like random forest and support vector machines (SVMs), we were able to predict the chance of a book reaching the New York Times Best Seller List with high probability given a set of certain features. Unlike prior approaches using textual analysis and NLP algorithms, we introduced a novel approach of classification where we only considered external discrete features. Our results suggest that if a book includes certain characteristics, it is able to reach the New York Times Best Sellers without regard to its actual content. By using multilevel quadratic regression, we were able to create a model that predicted the path of a book through the New York Times Best Sellers with a MAE of 3.449 when predicting 10 weeks ahead. Our analysis and statistical significance provide strong evidence suggesting that book popularity is deterministic in nature and can be represented as a function of past rank, weekly percent genre, and search popularity. We are not aware of any previous studies utilizing our methodology to predict book ranking and our experiment suggests such prediction is indeed possible.

The limitations our research faced can be split into two main categories: application programming interface (API) restrictions and data restrictions. Many of the APIs (interfaces used to collect data) we considered using were unusable or were heavily rate limited. On the data restriction side, several considered sources were either hidden behind paywalls or defunct. Apart from overcoming the limitations we faced, future research could implement other models of time series regression such as conditional random fields and hidden Markov models. In addition, our research could be applied to other forms of media in order to recognize whether similar properties apply and whether future rank is similarly deterministic. We can also make our models more accurate by incorporating more features.

The two models suggest that book success has little to do with the actual content of the book; if the book fits into certain categories, it is likely to become popular. Although cultural phenomena like book popularity are generally regarded as random, our models indicate that the underlying mechanisms for success are more mathematical than they appear.

MATERIALS AND METHODS

Technologies

Data manipulation was done with the Python library pandas, which allowed for data parsing and imputations for any missing data (8). For the book rankings regression, we made use of both scikit-learn's built-in ridge regression and numpy's polyfit; likewise, scikit-learn's classification tools gave us great ease in analyzing and constructing model (9).

Data

The New York Times Best Sellers dataset, which returned the top 20 Best Sellers each week from 2008-2018, made up the core of our data (3). Additional data came from the

Goodreads-10k book dataset, which gave us random assorted books (10). From our datasets, we were able to obtain two critical details for our time series: the path it took through the New York Times Best Sellers in terms of rank and the ratio of genres each week. The third detail of our time series, search relevance, was obtained using the Google Trends API. We found the first and last week each ISBN appeared on the New York Times Best Sellrs, and by using this interval and the book title, we were able to compile the search relevance of each book from one month before the first week to one month after the last week. The final data was split as follows: 70% to train, 15% to validate, and 15% to test.

For our classification model, we created one list of ISBNs that made the New York Times Best Sellers and another list of ISBNs that did not. We then collected details about the books, ranging from book weight to cover. One issue we encountered was the large number of categories some books belonged to. Our solution was to generate a hash map that mapped the categories present in the New York Times Best Sellers to their respective appearance frequency. We then represented the categorical data as the average of the frequencies of all the categories. Alongside this, we also added original IDs to the author as a measure of mapping authors to their respective titles. Through this data discretization method, we were able to represent categorical data as numerical data. The final data was split in the same ratio as the regression data.

Classification

Although the initial dataset was composed of nearly 14,000 verifiable books, valid information was only accessible for 6,000 books. These books could be described by 13 features: subject, publisher frequency, author frequency, Dewey Decimal number frequency, Dewey Decimal number, publisher, number of pages, publishing location, author, publish place frequency, title average word length, title number of words, and subject place. Contrary to a previous approach. we instead applied less importance on the 'textual values' of a novel and used categorical external factors purely for classifying the books onto the New York Times List (7).

We were able to numericize our frequencies of data by applying a tweaked version of the Natural Language Processing bag-of-words process, in which the encapsulating dictionary was built only from information of books that have been on the New York Times's Best Seller List in the past decade. Thus, a frequency score was given as follows:

$$F_i = \frac{f_i}{max(f)}$$

where f is the frequency of a certain term in the New York Times list. Frequencies scores, ranging from 0 to 1, were calculated for the author, publisher, Dewey Decimal, and publishing place columns. Next, in order to account for large-scale authors, publishers, and genres, an ID for each of category served as an efficient way of recognizing prevalent authors in the New York Times list. This helped predict with high probability that books by these big-name columns would likely achieve success, providing an alternative to the frequency score.

One area of interest was the subject field, which provided a list of anywhere from zero to a hundred unique words and phrases to identify the books by. Once again, we applied the alternate bag-of-words scheme to obtain a total dictionary of unique subject counts strictly from the New York Times books. After eliminating extraneous data fields like “Fiction” and “Protected DAISY” (as they were close to two orders of magnitude more common than other subjects and were not specific descriptions of the book), we calculated a subject score from a book's subject list S as:

$$S_i = \frac{\sum_0^n f_i}{n}$$

where the subject list's length is given as n and the frequency of a subject in the New York Times's dictionary is given as f_i.

This approach is also applied to calculate a subject place score (which is a data field that relays additional information about the setting of a book).

For the classification, we tested SVMs, naive Bayes, and random forests with a 70/15/15 train/validate/test split and compared their accuracies. Support vector machines with kernel trick raise the data by a dimension, then separate the classes with a hyperplane. Naive Bayes models utilize the Bayes Theorem with assumption of independence between the features. Random forest models are a collection of decision trees with feature bagging.

Regression

Initially, we turned the weekly top charts data into a hash map that mapped from each book to a list of its rankings over the weeks in which it traversed the top charts. In addition to that, we took the genres of the books on the New York Times Best Sellers and compiled a 2-D array that described the percentage of books of each specific genre in the Best Seller list per week. For search trends, due to the Google Search Trends index being relative to other searches at the same time period, many books had relative search strengths of 0. To counteract this, we took the data of the books which had more than 25% 0s and generated cumulative weekly searches instead. This helped mitigate the effect of sparse datasets, as seven occurrences of 0 would translate to one occurrence while concurrent zeros with at least one positive value would translate to a positive integer. As a result, sparse data was condensed while retaining relative search interest.

In order to run regression on all three of our time series, we needed to convert them into non-temporal numeric values. Hence, we ran regression on each time series and used the coefficients of each time series regression as features in the overall regression model. Due to the Central Limit Theorem, for n number of samples where $n > 30$, the distribution is a close approximation of the normal distribution (11). Since we have 400 samples for our error terms, which is more than 30,

we pass normality.

Initially, we modeled each of our features - past rank, weekly percent genre, and Google Search Trends index - by the function:

$$f(x) = \sum_{i=0}^{k} a_i$$

where *k* is the degree of regression, and *x* is the past data. We set random initial a_i and applied gradient descent until we reached a local minimum cost. We then took each ai from each model, and fed it as parameters into the level 2 model:

$$f(x) = \sum_{i=1}^{m} \sum_{j=1}^{k} b_i a_j^i + c$$

where *m* is the degree of regression for the level 2 model, *k* is the degree of regression for the time series model, and *c* is a constant. After, we ran gradient descent on the level 2 model with 5-fold cross validation to avoid overfitting. To determine which *m* and *k* were the best, we tried (*m*, *k*) combinations: (1, 1), (1, 2), (2, 1), and (2, 2). We did not test values of *m* and *k* beyond 2 because of a strong possibility of overfitting our data. Error was the difference between predicted rank and actual rank, giving us the mean absolute error:

$$MAE = \frac{\sum_{i=1}^{n} |\hat{y}_i - y_i|}{n}$$

A better MAE indicates a more accurate predictor and a better (*m*, *k*) as well. We suspected that certain features were not impactful and ran models with the following feature combinations: past rank, past rank and weekly percent genre, past rank, weekly percent genre, and Google Search Trends index. The change in MAE between features was an intuitive indicator for which features mattered, while the collection of bij was a more technical measure of importance.

To predict the entire path of the book, we took the prediction of the rank in the $n+1^{th}$ week from the previous n weeks of data, starting with $n = 3$ (as 3 points is the minimum number of points needed for a quadratic line), assuming that it is the actual next value, and appended it to the data. We then re-ran the models again, with the updated data set, and got a new prediction for the $n+2^{th}$ week. We repeated this until we have predicted *w* weeks ahead, and the predictions of ranks from 3 to *w* weeks trace a path that the book rank will take. To determine whether our model was statistically significant, we ran a Student's t-test to determine whether our explanatory variables had a statistically significant impact on the response.

ACKNOWLEDGEMENTS

We gratefully acknowledge Lina Kim, Director of Pre-College Programs at University of California Santa Barbara for her administration of the Science and Engineering Research Academy program, which supported this research. We also express gratitude towards Shadi Mohagheghi, Rachel Redberg, and Angela Zhang for their advice during this research and their review of this manuscript. We would also like to thank the Journal of Emerging Investigators for giving us the feedback needed to improve our manuscript and the opportunity to publish.

Received: October 13, 2019
Accepted: May 06, 2020
Published: May 23, 2020

REFERENCES

1. Watson, Amy. "Unit sales of printed books in the United States from 2004 to 2018 (in millions)." *Statista*, 14 Jan. 2019. https://www.statista.com/statistics/422595/print-book-sales-usa/.
2. Sayyadi, Hassan, and Lise Getoor. "FutureRank: Ranking Scientific Articles by Predicting their Future PageRank." *Proceedings of the 9th SIAM International Conference on Data Mining*, 2009, pp 533-544.
3. Cibils, Cristian, *et al*. "Predicting a Song's Path through the Billboard Hot 100." Stanford CS229, 2015, pp 1-3.
4. Harvey, John. "The Content Characteristics of Best-Selling Novels." *Public Opinion Quarterly*, vol. 17, no. 1, 1953, pp. 91-114.
5. Yucesoy, Burcu, *et al*. "Success in Books: A Big Data Approach to Bestsellers." *EPJ Data Science* 7, 2018, pp 2-3.
6. Archer, Jodie. Reading the Bestseller: An Analysis of 20,000 Novels. 2014. *Stanford University*, PhD dissertation.
7. Archer, Jodie, and Matthew L. Jockers. The Bestseller Code: Anatomy of the Blockbuster Novel. *St. Martins Griffin*, 2017.
8. McKinney, Wes, *et al*. "Data Structures for Statistical Computing in Python." I*n Proceedings of the 9th Python in Science Conference*, 2010, Vol. 445, pp. 51–56.
9. Pedregosa, Fabian, *et al*. "Scikit-learn: Machine Learning in Python." *Journal of Machine Learning Research*, 2011.
10. Zając, Zygmunt. Goodbooks-10k. *Github*, 27 Nov. 2017. https://github.com/zygmuntz/goodbooks-10k.
11. "The Central Limit Theorem." *Florida State College at Jacksonville*, 2019, https://guides.fscj.edu/Statistics/centrallimit/. Accessed: 14 July 2019.

JOURNAL OF
EMERGING INVESTIGATORS

Article

Analyzing the Relationships Between Internet Usage, Social Skill, and Anxiety Severity in Adults with Autism Spectrum Disorder

Celina You[1], Daina Tagavi[2], Anthony Osuna[2], Ty Vernon[2]
[1]Westlake High School, Thousand Oaks, California [2]University of California, Santa Barbara, California

SUMMARY

The use of social media in adults with Autism Spectrum Disorder (ASD) is not very well understood. Online communication may provide these individuals with the opportunity to interact with others without having to engage in face-to-face conversation or interpret nonverbal social cues, which is often a challenge for this population. The goal of the Socialization Education and Learning for the Internet (SELFI) study is to understand how adults with ASD utilize social media platforms to build relationships and to teach them how to act appropriately online. Specifically, this project compares anxiety levels and social skills between adults with ASD who spend different amounts of time on the Internet to observe the potential relationships between online communication frequency, anxiety, and social interaction. Based on current literature, we hypothesized that Internet usage would have beneficial effects, reducing anxiety and improving social skills in these adults. We found that a higher frequency of Internet usage correlated with less severe anxiety symptoms but did not have a significant relationship with the social skills of adults with ASD. This research furthers the SELFI project, which will observe the consequences of social media usage in those with autism. Depending on the results of this study, we also aim to determine whether social media could possibly become a new intervention technique used to decrease anxiety in adults with ASD.

INTRODUCTION

Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder that affects social communication and behavior. Individuals diagnosed with ASD may experience numerous challenges with social interaction that can hinder their ability to succeed in social situations. As a result, people with ASD have a difficult time forming and maintaining friendships, which can then lead to significant loneliness. Many intervention techniques, such as cognitive behavioral therapy (CBT) and applied behavior analysis (ABA), have been established to enhance social development in this population. However, none of these intervention tactics explicitly target the realms of online communication and social media usage, which can enhance the quality of existing relationships in individuals with autism and decrease loneliness (1).

Individuals with ASD also struggle with nonverbal social cues, such as recognizing facial expressions and emotions, resulting in difficulty making friends and interacting with others. However, it has been previously demonstrated that the use of social media was significantly associated with high friendship quality in adolescents with autism, which was moderated by the adolescent's anxiety levels. As individuals with ASD get older and more aware of social difficulties, they are at an increased risk for social anxiety. Research results indicate that social media usage is not related to perceived loneliness (2). However, having more friends and closer relationships is significantly associated with a decrease in loneliness (2). Research has also suggested that adults with ASD who were active on social networking sites had a higher chance of developing close friendships and improving their social functioning, making electronic social media a particularly attractive form of social interaction for those with ASD (1).

Anxiety, in particular, is commonly experienced by individuals with ASD. It is hypothesized that anxiety in these individuals might stem from impaired emotional regulation (ER), a characteristic that is ubiquitous in ASD (3). ER is a person's ability to effectively manage and respond to an emotional experience, and having poor ER could make it harder for one to control intense, usually negative, feelings (4). Impaired ER may explain many common behavioral problems faced by children and adults with ASD, such as aggression, irritability, and anxiety (5).

The present study specifically uses data collected from those with high functioning autism (HFA) or Asperger syndrome (AS). These individuals usually possess high verbal ability and average or above average IQs but have difficulty with nonverbal aspects of social communication. In past studies, those with HFA or AS felt more comfortable interacting over the Internet rather than in person (6). This form of communication seemed to be less stressful for these individuals, and they felt less pressured because of the increased amount of time they had to formulate a proper response. They also experienced having more control over the situation and how they wanted the conversation to progress.

By comparing anxiety levels and social skills between adults with ASD who spent different amounts of time online, we hoped to understand the relationships between online communication frequency, anxiety, and social interaction. This study will contribute to furthering the Socialization Education

and Learning for the Internet (SELFI) study and will observe the consequences of social media usage. Depending on the results of this study, we also aim to determine whether social media could possibly become a new intervention technique used to decrease anxiety in adults with ASD. If social media can be used as an intervention, it will be easier, cheaper, and more accessible for those with autism. Current intervention options, like Applied Behavioral Analysis can cost anywhere from $15,000 to $50,000 according to the Centers for Disease Control. We hypothesize that spending time online would lessen anxiety severity and improve social skills in adults diagnosed with ASD.

RESULTS

Internet Use and Anxiety Levels

Is an increased usage of the Internet associated with a decrease in the anxiety levels of adults with HFA? The Beck Anxiety Inventory (BAI) (7) was administered to participants prior to receiving treatment from the center. After recording the total hours spent online and the corresponding BAI score for each participant, the data was analyzed using SPSS analysis. Results indicated that anxiety levels and hours spent on the Internet were inversely correlated **(Figure 1).** A bivariate correlation, which determines the existence of relationships between two different variables, was then run and revealed that the correlation was significant **(Table 1).**

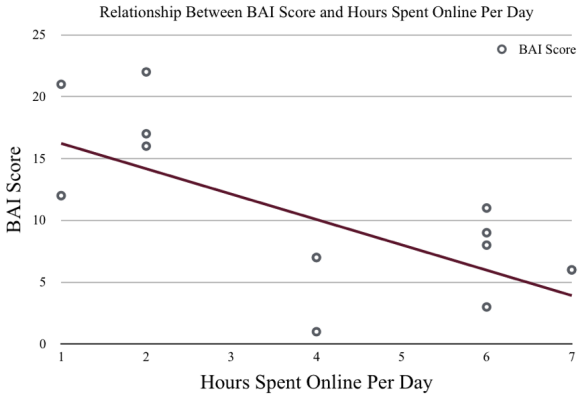


Figure 1: Relationship between Beck Anxiety Inventory (BAI) score and hours spent online. An increase in Internet usage is associated with a decrease in anxiety severity.

Bivariate Correlation Between Internet Usage and Anxiety Severity

		BAI Score	*Hours Spent on Internet*
Hours Spent on Internet	*Pearson Correlation*	*-.560**	*1*
	Sig. (2-tailed)	*0.024*	
	N	*16*	*17*

Table 1: Pearson's correlation between Internet usage and anxiety severity. The correlation between Internet usage and decreased anxiety severity is significant (P < 0.05), meaning that an increased amount of time spent online may be associated with a decreased anxiety symptoms. * Correlation is significant at the 0.05 level (2-tailed).

Internet Use and Social Skill

To determine whether an increased amount of time spent online resulted in better social skills for adults with ASD, we ran a bivariate correlation with hours spent on the Internet, the percentage of 30-second intervals during which the participant asked a question, the percentage of questions the participant responded to, the percentage of 30- second intervals with a silence of four seconds or longer, and the percentage of 30-second intervals during which the participant expressed an initiation, all of which were documented from watching videos of the participant conversing with a center clinician **(Table 2)**. The results indicated that the correlation between hours spent online and social skill was not significant (P > 0.05) **(Table 2)**.

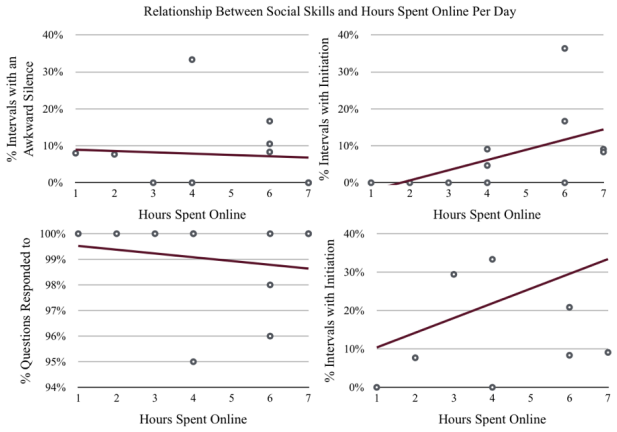


Figure 2: Relationship between social skills (questions responded to, intervals with a question, awkward silences, and intervals with initiation) and hours spent online. There is no significant relationship between increased Internet usage and improved social skills.

Bivariate Correlation Between Internet Usage and Social Skills

		Hours Spent on Internet	*% Intervals with a Question*	*% Questions Responded to*	*% Intervals with a Prolonged Silence*	*% Intervals with Initiation*
Hours Spent on Internet	*Pearson Correlation*	*1*	*0.367*	*-0.161*	*-0.070*	*0.510*
	Sig. (2-tailed)		*0.267*	*0.636*	*0.838*	*0.090*
	N	*17*	*11*	*11*	*11*	*12*

Table 2: Pearson's correlation between Internet usage and social skills. The correlations between Internet usage and improved social skills is not significant (P > 0.05), meaning that there is no strong association between an increased amount of time spent online and an improvement regarding questions asked, questions responded to, prolonged silences, or initiations.

DISCUSSION

This was the first study to examine the possible relationships between Internet use, anxiety levels, and in-person social interaction skills in adults with ASD. The results of this study suggested that use of the Internet is associated with decreased anxiety severity in adults with ASD but not with an ability to socialize appropriately. As anxiety is one of the most common co-occurring psychiatric conditions in

individuals with ASD, results from this study suggested that the Internet could be a tool utilized to combat that experience (8). The conclusions drawn from this study also supported the exploration of a social media-based intervention in adults with ASD. Eleven out of the 17 participants (64.71%) in this study reported some type of social media use (Twitter, Instagram, Snapchat, YouTube, etc.). This finding is consistent with usage in the general adult population as described in recent reports (11-13). Social media might be a more convenient way for those with ASD to interact with their peers in a less socially challenging environment. Social media and electronic communication tools can be used to enhance the quality of existing relationships in typical adolescents without reducing the amount of face-to-face interaction, thus decreasing loneliness (2).

The results of this study did not provide evidence of a relationship between Internet usage and social skill improvements. These results may be because our sample size was relatively small and did not encompass the full spectrum of ASD. Furthermore, the social skills that we observed may not have been the ones affected by the Internet. These results may seem to undermine the benefits that the Internet can have on adults with autism, but the overarching goal of the project is to improve the quality of life of these individuals rather than targeting one specific aspect of improvement. Because the dependent variable encompassed all forms of Internet usage and wasn't limited strictly to using the Internet for communication, the results from this study prompt further investigation of the social effects of increased online communication. While increased Internet usage did not have a significant correlation with individuals' social skills, social media shows promise in making conversation more comfortable for those with ASD.

Although the current study provides insight into the relationships between Internet use, anxiety severity, and social skill in adults with ASD, there are a few limitations that should be noted. The small sample size and cross-sectional nature of the study were two major limitations. The correlational nature of this study, however, makes it impossible to draw any conclusions about how these variables are related causally. Longitudinal and experimental studies would need to be conducted to confirm the effect of Internet usage on adults with ASD.

Additionally, the aspects of social interaction that we coded may not have been the skills that were actually affected by Internet use. We only examined four characteristics of social skills (questions asked, questions responded to, prolonged silences, initiations), but higher frequencies of Interest use may have had an effect on other social behaviors.

Finally, because only a subsample of the ASD spectrum was examined in this project, it should be noted that the results may not generalize to the entire population of adults with ASD. Only individuals with high functioning autism participated in this study, so we cannot yet generalize these findings to the entire autism spectrum.

The current study expands the minimal literature on the relationships between Internet usage and autism symptoms. The results suggest that a higher frequency of Internet usage correlated with lower anxiety levels but did not have a significant relationship with the social skills of adults with ASD. The cross-sectional relationship between Internet usage and anxiety suggests a need for further research to confirm the results of the present study. Testing the effects of more Internet use on other social behaviors like eye contact or empathy may also be a future direction for this study. Furthermore, the data revealed that over half of the sample population used some form of social media to communicate and interact with others. With further research, social media platforms may be utilized as new forms of intervention designed to decrease anxiety severity for adults with ASD. Eventually, we want these individuals to have in-person interactions; teaching them proper online etiquette can be an intermediary step to getting them to be comfortable meeting new people and eventually translating these skills to real life.

MATERIALS AND METHODS

Participants

The participants were 17 adults ranging from ages 19 to 54. Participants were previous or current clients receiving services at the Koegel Autism Center who have been diagnosed with Autism Spectrum Disorder, specifically high functioning autism. Of the 17 participants, 23.53% were female and 76.47% were male. 58.8% of the clients were Caucasian, 23.5% were Asian, 5.9% were Hispanic/Latino, and 11.8% were other. The mean age was 27.71 years with a standard deviation of 10.89.

Measures

Internet use: Participants completed the Koegel Autism Center Adult Questionnaire before their first session at the center. Regarding Internet use, participants were asked: "On average, how many minutes/hours do you spend each day on the computer? How is this time usually spent?" The answers to these questions were self-reported and were not exclusive to interacting online via social media sites.

Anxiety: The Beck Anxiety Inventory (BAI) (7) was administered to participants prior to receiving treatment from the center. The BAI is a multiple-choice self-report inventory used to measure the severity of anxiety in children and adults. Symptoms of anxiety are rated on a four-level scale: Not at All; Mildly (It did not bother me much); Moderately (It was very unpleasant, but I could stand it); and Severely (I could barely stand it). The BAI is widely used and has shown validity and internal consistency across a variety of languages, cultures, and age ranges (7).

Behavioral coding system

To assess social skill level, baseline videos of the participants

conversing with a Koegel Autism Center clinician were coded based off what we characterized as acceptable social behaviors, which are detailed below. These interactions were done in person to simulate a natural conversation. It is important to note that some participants did not have baseline videos, so not all participants had social skills coded for.

Number of questions asked: The questions asked by both the participant as well as the person they were conversing with were also tracked. For each 30-second interval, if the participant asked a question, we marked that down and did the same if the peer asked a question. Asking questions can demonstrate interest and engagement in conversation.

Silences: A study conducted by researchers in Holland reports that brief silences, specifically silences that are four seconds long, can promote negative feelings of anxiety and rejection (9). For every 30-second interval, we kept track of any silences four seconds or longer.

Vocalizations: Live social interactions were coded for "Responding and Initiating" based off of a study conducted by McMahon, Vismara & Solomon (10). We classified vocalizations directed toward another person in the absence of a conversation as "initiating" and vocalizations directed toward another person in the presence of a conversation (i.e. within approximately 10 seconds of a previous vocalization) as "responding" (10). To find the percentage of questions the participant responded to, we divided the number of responses by the number of questions asked by the stranger and multiplied by 100. For every 30-second interval, we noted if the participant exhibited any "initiations." This could suggest more back-and-forth conversations and fewer choppy conversations with pauses of four or more seconds (10).

Calculations

To calculate the percentage of 30-second intervals with a question asked by the participant, we divided the number of intervals with a question by the total number of intervals present and multiplied by 100. We then performed the same calculation for prolonged silences and initiations. To determine the significance of our results, we ran bivariate correlations between Internet usage, anxiety severity, and social skills.

ACKNOWLEDGEMENTS

The first author would like to acknowledge her mentors, Daina Tagavi and Anthony Osuna, for supplying her with resources, answering her countless questions, and supporting her throughout this research journey. This work would not have been possible without their advice, guidance, and encouragement. The first author would also like to recognize Michael Hughes for his feedback on the structure of her work, and the UC Santa Barbara Koegel Autism Center for its facilities. Last, but not least, she would like to thank Dr. Lina Kim for providing her with this incomparable opportunity to learn and grow.

Received: August 12, 2019
Accepted: May 8, 2020
Published: May 23, 2020

REFERENCES

1. Mazurek, Micah O. "Social Media Use Among Adults With Autism Spectrum Disorders." *Computers in Human Behavior*, vol. 29, no. 4, 2013, pp. 1709-1714. doi: 10.1016/j.chb.2013.02.004.
2. van Schalkwyk, Gerrit I., et al. "Social Media Use, Friendship Quality, and the Moderating Role of Anxiety in Adolescents With Autism Spectrum Disorder." *Journal of Autism and Developmental Disorders* vol. 47, no. 9, 2017, pp. 2805-2813. doi: 10.1007/s10803-017-3201-6.
3. Mazefsky, Carla A., et al. "The Role of Emotion Regulation in Autism Spectrum Disorder." J*ournal of the American Academy of Child & Adolescent Psychiatry* vol. 52, no. 7, 2013, pp. 679-688. doi: 10.1016/j.jaac.2013.05.006.
4. White, Susan W., et al. "Social-cognitive, Physiological, and Neural Mechanisms Underlying Emotion Regulation Impairments: Understanding Anxiety in Autism Spectrum Disorder." *International Journal of Developmental Neuroscience* vol. 39, 2014, pp. 22-36. doi: 10.1016/j.ijdevneu.2014.05.012.
5. Mazefsky, Carla A., and Susan W. White. "Emotion Regulation: Concepts & Practice in Autism Spectrum Disorder." *Child and Adolescent Psychiatric Clinics of North America* vol. 23, no. 1, 2014. doi: 10.1016/j.chc.2013.07.002.
6. Benford, Penny, and P. Standen. "The Internet: A Comfortable Communication Medium For People With Asperger Syndrome (AS) and High Functioning Autism (HFA)." *Journal of Assistive Technologies* vol. 3, no. 2, 2009, pp. 44-53. doi: 10.1108/17549450200900015.
7. Steer, Robert A., et al. "Further Evidence For the Validity of the Beck Anxiety Inventory with Psychiatric Outpatients." *Journal of Anxiety Disorders* vol. 7, no. 3, 1993, pp. 195-205. doi: 10.1016/0887-6185(93)90002-3.
8. Vasa, Roma A., and Micah O. Mazurek. "An Update on Anxiety in Youth with Autism Spectrum Disorders." *Current Opinion in Psychiatry* vol. 28, no. 2, 2015, pp. 83. doi: 10.1097/YCO.0000000000000133.
9. Koudenburg, Namkje, Tom Postmes, and Ernestine H. Gordijn. "Disrupting the Flow: How Brief Silences in Group Conversations Affect Social Needs." *Journal of Experimental Social Psychology* vol. 47, no. 2, 2011, pp. 512-515. doi: 10.1016/j.jesp.2010.12.006.
10. McMahon, Camilla M., Laurie A. Vismara, and Marjorie Solomon. "Measuring Changes in Social Behavior During a Social Skills Intervention for Higher-Functioning Children and Adolescents with Autism Spectrum Disorder." *Journal of Autism and Developmental Disorders* vol. 43, no. 8, 2013, pp. 1843-1856. doi: 10.1007/s10803-012-1733-3.

JOURNAL OF
EMERGING INVESTIGATORS

Article

The Effects of Different Aquatic Environments on the Rate of Polyethylene Biodegradation by *Bacillus subtilis*

Aditya Tadimeti and Jeff Sutton
The Harker School, San Jose, California

SUMMARY

The proliferation of plastics in consumer products results in the release of countless tons of plastic waste into the environment. Because of their chemical composition, plastics take hundreds of years to decompose. The world today is overrun by plastic waste that occupies landfills, makes its way into rivers and oceans, and overwhelms several ecosystems. Current methods of handling the waste namely recycling, landfills and incineration, are inadequate, ineffective or harmful. In recent years, scientists have discovered that certain bacteria can degrade and assimilate polyethylene. Microbes have been shown to have the ability to "eat" petroleum-based products like natural gas and light sweet crude oil from oil spills. The ability of bacteria to decompose certain types of plastic makes them a potential bioremediation option. Current research indicates that the bacterium *Bacillus subtilis* (*B. subtilis*) is proficient at degrading polyethylene utilizing a biosurfactant called surfactin. Studies show that the bacterium is versatile and can thrive in various environments, while surfactin can withstand high concentrations of salinity. The aim of this study is to test the ability of *B. subtilis* to degrade high-density and low-density polyethylene in aquatic environments. Rates of degradation were studied and compared across fresh water, brackish water, and ocean water samples. Degradation occurred across all samples, although it was the highest in fresh water and lowest in ocean water. This study supports the hypothesis that *B. subtilis* can potentially be used to help degrade plastic in aquatic environments.

INTRODUCTION

Only 9% of plastic produced in the US is recycled – some of it is incinerated, a process that emits toxic gases into the atmosphere, and most of it is dumped in landfills (1, 2). Over 8 million metric tons of plastic waste enters the world's oceans annually (3). Marine animals are mistaking plastic waste for food; plastic has been found in more than 60% of all seabird and in 100% of sea turtle species (4). Besides being life threatening to animals and birds, this affects humans because we consume seafood that can contain chemicals from plastics. Diethylhexyl phthalate (DEHP), contained in some plastics, is a toxic carcinogen and other toxins in plastics are directly linked to cancers, birth defects, immune system problems, and childhood developmental issues. Bisphenol A (BPA), a known chemical that interferes with human hormonal function, is used in plastic bottles and food packaging materials. Over time the polymer chains of BPA break down and can enter the human body through drinking contaminated water or eating a fish that is exposed to the broken down toxins (5, 6).

Plastics are semi-organic materials that come from oil or petroleum. They are routinely labeled as polymers, as they are comprised of polymers which are large molecules made with a massive amount of smaller, identical molecules. Polymers have a different physical and chemical makeup than their monomers, and more uniquely, their properties can be tailored depending on their main purpose. Polyethylene is a polymer that is exceptionally versatile. About 80 million tons of this compound are produced each year for widespread usage in consumer products (7). It is a thermo-plastic created from the polymerization of ethylene, a process that produces very long, very straight chains of hydrocarbon monomers. By adjusting the polymerization process, the long chains can be made to branch, creating different kinds of polyethylene. The degree of branching determines what kind is produced. **(Figures 1A, 1B, 1C)**. Low-density polyethylene (LDPE) and high-density polyethylene (HDPE) are two of the most common types of polymers in commercial use. LDPE has the most long- and short-chain branching of any form of polyethylene, resulting in its lower density. LDPE is particularly useful for a range of applications, from rigid products like plastic bottles, buckets, and bowls, to filmy ones like plastic grocery bags and plastic cling-wrap. HDPE is characterized by minimal branching of the polymer chain, making it much denser and rigid. The added tensile strength of HDPE makes it useful for rigid applications such as milk and detergent jugs, garbage cans, water pipes and children's toys.

Plastics cannot dissolve in water and are estimated to require hundreds of years to naturally decompose (8). An environmentally-friendly solution to potentially help tackle plastic waste is biodegradation, which is the decay or breakdown of materials that occurs when microorganisms use an organic substance as a source of carbon and energy (9). Without releasing toxic byproducts into the atmosphere, these organisms can decompose the polyethylene present within plastics at rates significantly higher than natural degradation (10). Biodegradation is a two-step process. The first step is the cleavage of the large molecular chains, and the second step is the mineralization. External enzymes are

responsible for the first step, because the sizes of the polymer chains are considerably greater than the majority of the microorganisms. Once sufficiently reduced to a small size, monomeric fragments are transported into the cells where they mineralize. The products of this process are water, salts, minerals, gases and biomass (11).

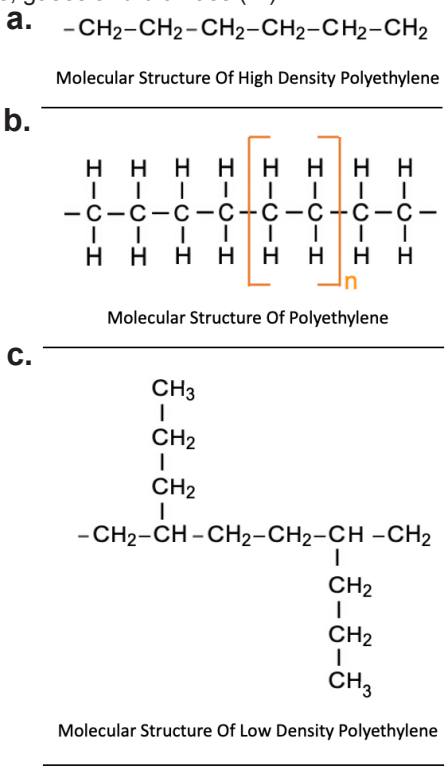


Figure 1: Visual depiction of the molecular structure of (A) HDPE, (B) polyethylene, and (C) LDPE.

All living organisms are able to produce diverse antimicrobial compounds (12). Soil bacteria and fungi produce antibiotics to gain an advantage when competing for food, water, and other limited resources in a particular habitat, as the antibiotics kill off competition (13). Lipopeptides, one category of such compounds, represent a unique class of cyclic peptides and exhibit remarkable therapeutic and biotechnological properties (14). Surfactin, one of the principal representatives of the lipopeptide family, is produced by *Bacillus subtilis* (*B. subtilis*), a rod-shaped gram-positive bacterium typically found in the upper layers of soil as well as in oceans. Gram-positive bacteria are important to the field of biotechnology since these bacteria produce enzymes critical for industry such as proteases, amylases, and lipases, both reliably and at low cost (16). *B. subtilis* secretes proteins in the range of grams per liter, making *B. subtilis* an excellent candidate to study for this application (17).

Surfactants are substances that reduce the interfacial tension between water and oil and adsorb at the interface to stabilize emulsions (18). They have an amphipathic molecular structure – having both hydrophilic and hydrophobic parts – and exhibit properties like formation of emulsions, foaming and detergency. Because of these properties, surfactants lower the surface energy and tension of a medium and oxidize the polyethylene. This oxidation converts carbonyl groups into unsaturated hydrocarbons, breaking down the structure of polyethylene. Through this process, described as depolymerization, the large polymer chains are broken into smaller, water-soluble fragments that can pass through microbial membranes, where they are biodegraded by cellular enzymes and used as carbon and energy sources (19).

Microorganisms, possessing a large surface area-to-volume ratio, produce a variety of surfactants referred to as biosurfactants (20). Biosurfactants have advantages over their synthetic counterparts because they are biodegradable, less toxic, and effective at extreme temperatures and pH (21). One of the most powerful biosurfactants, surfactin lowers the surface tension of water, even at very low concentrations (22). Surfactin is a highly potent agent with diverse commercial applications, including subsurface pollution remediation and the enhancement of the availability of hydrophobic compounds, thus increasing the potential for biodegradation by microbes (23).

Current research indicates that certain biosurfactants are proficient at degrading polyethylene, utilizing surfactin, which isn't degraded itself, to initiate plastic decomposition (24). *B. subtilis* has been found in different oceans; it is known to be versatile and can thrive in various environments. Its polyethylene-degrading biosurfactant can withstand high concentrations of salinity, supporting the theory that the bacteria can survive in underwater environments and decompose the polyethylene, even in situations where it is exposed to high concentrations of chlorine. We therefore utilized *B. subtilis* in this study to determine if it would biodegrade polyethylene in different aquatic environments because of its production of a biosurfactant and ability to withstand extreme environments.

The tests were conducted across three different levels of salinity of water: fresh water, brackish water and ocean water. Bacteria grow slower in ocean water than in freshwater (25, 26). Accordingly, we hypothesized that the reduction of plastic will be the lowest in ocean water and highest in freshwater, with the mass reduction in brackish water falling in between the two.

RESULTS

To test the effect of salinity on the ability of *B. subtilis* to degrade plastic, we placed bacteria and a square piece of polyethylene into test tubes mimicking different aquatic environments. Specifically, the test tubes contained a growth medium consisting of nutrients that support bacterial growth, and we filled each test tube with freshwater, brackish water or ocean water to model the conditions present in various bodies of water. We placed the test tubes in an incubator to allow the bacteria to grow. After 30 days of incubation, we measured the mass of each polyethylene piece. We decided

to run the experiment over a 30 day period as we deemed it to be a reasonably long period for observable degradation to occur. Future iterations of the experiment may be run using different time periods based on results.

We cut all the HDPE and LDPE pieces into similar shapes with a uniform mass of roughly 17 mg. We set up control samples of HDPE and LDPE by placing the HDPE and LDPE pieces in test tubes containing water representing the different aquatic environments, the growth medium, and no bacteria. We ensured that *B. subtilis* was the only agent causing the degradation of the polyethylene. As expected, the control samples, which were not exposed to any bacteria, remained at approximately the same mass, having an average of a 0.04 percent decrease in mass over the 30 day period across all three types of water environments. Some control samples increased in mass over the 30 day period, which could be due to either mass measurement errors or improper cleaning of the samples to remove excess chemicals or bacteria.

To test whether the 0.1M HCL wash and water caused any degradation of the plastic, additional LDPE and HDPE pieces of shape, size, and mass similar to the test tube samples were weighed before and after a HCl wash; no significant amount of mass degradation was noticed. No significant polyethylene degradation occurred during the heating (verified by a separate, dry, control).

Samples exposed to *B. subtilis* in freshwater had the highest amounts of degradation relative to the controls. The average mass of the HDPE pieces had a 5.79% decrease, while the average mass of the LDPE pieces had a 5.77% decrease over a period of one month, when compared to the controls **(Figure 2)**.

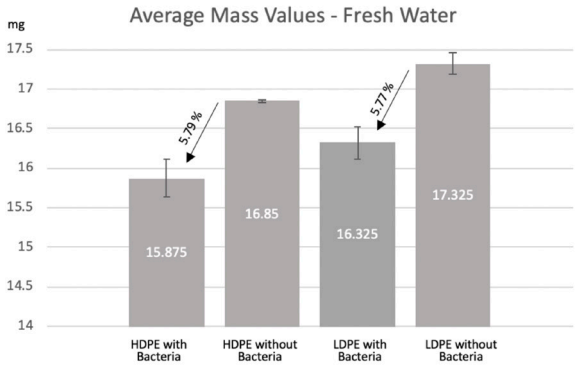


Figure 2: The HDPE and LDPE samples exposed to *B. subtilis* in freshwater over a 30 day period showed degradation when compared to the samples that were not exposed to the bacteria.

Samples exposed to *B. subtilis* in brackish water underwent a smaller decrease in mass. The average mass of the HDPE pieces had a 5.08% decrease, while the average mass of the LDPE samples had a 5.30% decrease over one month, compared to their controls **(Figure 3)**.

Samples exposed to *B. subtilis* in ocean water experienced the least amount of degradation. The average mass of the HDPE pieces decreased 3.61%, while the average mass of the LDPE samples decreased 2.47% over one month, compared to their controls **(Figure 4)**.

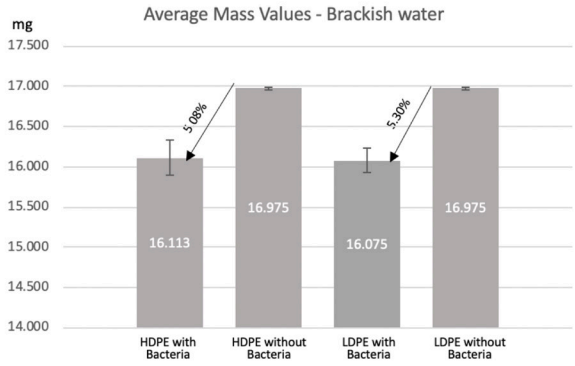


Figure 3. The HDPE and LDPE samples exposed to B. subtilis in brackish water over a 30 day period showed degradation, although it was less than the bacteria-exposed samples in freshwater.

The amount of degradation between HDPE and LDPE pieces was not consistent. We did not explore the implication of this outcome.

The results appear to support the hypothesis that the presence of *B subtilis* accelerates the rate of degradation of HDPE and LDPE. Based on available research findings, this outcome can be attributed to the depolymerization of the large molecules into smaller monomers with the help of the biosurfactant surfactin, and the subsequent passing of the monomers through the cell membranes of the bacteria to be converted into energy.

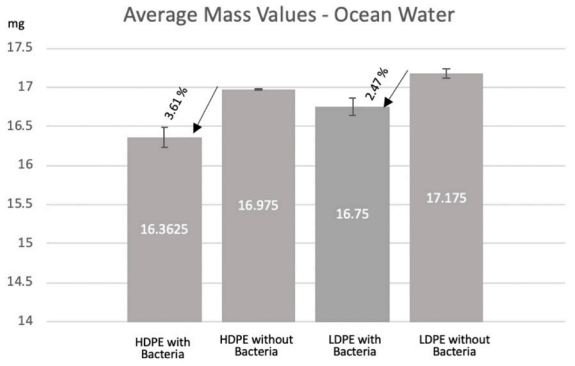


Figure 4: The HDPE and LDPE samples exposed to B. subtilis in ocean water over a 30 day period showed degradation, although it was the lowest when compared to samples in fresh or brackish water.

DISCUSSION

The main goal of our study was to test whether *B. subtilis* can degrade polyethylene in aquatic environments and to compare the rate of degradation across varying levels of salinity that simulate real-world bodies of water. The degradation as measured by the reduction in mass of the

HDPE and LDPE samples after 30 days of incubation was found to be the highest in freshwater samples and the lowest in ocean water samples. This supported the hypothesis that *B. subtilis* can degrade polyethylene in salt water and that the amount of degradation decreases with increase in salinity; our findings are consistent with at least one previous study that identified *B. subtilis* as an agent of degradation in freshwater (27).

The ability of *B. subtilis* to degrade polyethylene in aquatic environments offers the possibility that this bacteria may be deployed as a solution to plastic pollution in various bodies of water. The superior degradation in freshwater suggests that bacterial degradation of polyethylene can tackle plastic pollution in lakes and rivers. The degradation in brackish water suggests that biodegradation can be used in areas like estuaries, large lakes, and at the mouths of rivers. Finally, degradation in more ocean water suggests that bioremediation can be a viable solution to tackle plastic pollution in our oceans. The long-term goal of this study is to develop a cost effective and environmentally safe solution to degrade plastic at scale, using natural organisms like *B. subtilis* at temperature ranges where it is active.

Research shows that bacterial growth is inhibited with increases in salinity, which corresponds to the decreased degradation in the ocean water samples (28). It is not clear from this experiment whether the reduced degradation is due entirely to the slower growth of bacteria or if surfactin's effectiveness is also reduced in ocean environments. The rates of degradation observed may correspond to the temperature and salinity at which the experiment was conducted. To further validate the differences in the rates, the experiment should be repeated at varying temperature and salinity levels.

Further research is required in several areas to establish *B. subtilis* as a practical option for the degradation of plastic waste in our water bodies. The lower rate of degradation implies either that the bacteria are not as productive in ocean environments or that surfactin is less effective in the presence of salts and other chemicals seen in ocean water. There is adequate research to show that *B. subtilis* is a common inhabitant of ocean habitats, so the bacteria's viability in natural, marine environments is well established (29,30). Additional research into where in the oceans the bacteria thrive the most may suggest the regions of the oceans where *B. subtilis* can be deployed. Future studies should also focus on the effects of ocean water on the detergency of surfactants with the end goal of creating conditions that may enhance their effectiveness.

Studying the productivity of *B. subtilis* in expressing surfactin as well as absorbing the degraded smaller monomer molecules when in ocean environments would be key next steps for this research. The optimal temperature for the growth of *B. subtilis* is 25-35°C, while the optimal conditions for high surfactin production in the laboratory are a slightly acidic pH (6.5–6.8) and an incubation temperature of 30°C (31). Ocean temperatures vary from -2°C to 35°C which is well within the range for *B. subtilis* to thrive, though colder areas may see less productivity from bacteria (32). However, the pH of oceans is typically at an average of 8.1 and therefore less acidic than ideal incubation conditions (33). We need to find ways to overcome the pH difference and encourage bacterial growth; otherwise, *B. subtilis* may only be a viable option in fresh and brackish water, and not in oceans.

Under laboratory conditions, research has shown that pretreatment of polymer films with UV radiation aids its accessibility as food for the microorganisms, enabling a much faster rate of biodegradation (27). Therefore, there is reason to believe that in natural conditions, the UV radiation from the sun can aid biodegradation by *B. subtilis*. However, to be a commercially available solution, the rate of degradation by the bacteria may have to be much faster than what was observed in the experiment.

So far, much of the laboratory research points to the use of microbes in situ to produce surfactants. An alternative method to harness *B. subtilis* could be to produce surfactin under optimal industrial conditions, transport it to the site of plastic pollution, and spray it to stimulate biodegradation. However, it is important to note that studying the effectiveness of surfactants to fight hydrocarbon pollution in the open sea remains a challenge, and research is still ongoing in this area (34, 35). Studying formulations that are inherently more vulnerable to biodegradation and promoting the use of such plastics can help with the natural degradation process. Research shows that additives such as pro-oxidants and starch make plastics biodegradable. Starch-blended polyethylene makes the material hydrophilic and allows it to be catalyzed by amylase enzymes; microorganisms are then able to attack and remove this section thus degrading the plastic (36).

For *B. subtilis* to be a practical option to tackle the enormous amount of plastic waste, it should be deployable at scale. Further experimentation is required to study the optimal methods to incubate, transport and deploy large amounts of this bacterium, while also evaluating any potential adverse impact on the aquatic ecosystems when doing so. Surfactin, produced by *B. subtilis*, is indeed commercially produced and sold, although at present it is not cost effective to produce at scale. One reason for the high cost is due to the substrates employed for their production and the purity level required for application in the fields of pharmaceutics and medicine as well as the small batch sizes (37). Recent research into industrial scale production of surfactin at lower cost has been driven by the desire to clean up oil spills using biosurfactants and shows promise (38, 39).

Further research and experimentation is required to assess whether the introduction of additional surfactin can enhance the ability of *B. subtilis* which is naturally present in oceans to act as a viable agent of polymer degradation. As an alternative, the possibility of introducing adequate amounts of *B. subtilis* to ocean environments to help degradation has to

be evaluated along with a study of the impact of introducing such vast amounts of the bacteria on marine ecosystems.
The experiment showed that the presence of *B. subtilis* accelerates the rate of degradation of HDPE and LDPE in aquatic environments. The results offer the promise that microbial biodegradation of plastics can be an option to pursue in our efforts to mitigate the vast amounts of plastic pollution in water bodies.

MATERIALS AND METHODS

Bacterial Culture

B. subtilis (sourced from Carolina Biological by the Harker School wet lab) was obtained in a sterile, 20 x 150 mm test tube, approximately one-third filled with agar. Colonies of *B. subtilis* were visible as small bumps on the surface of the agar **(Figure 5A)**. Bacteria were streaked onto a petri dish containing agar using a cotton swab **(Figure 5B)**. Bacteria in petri dishes were left at room temperature to culture and placed in a refrigerator for preservation.

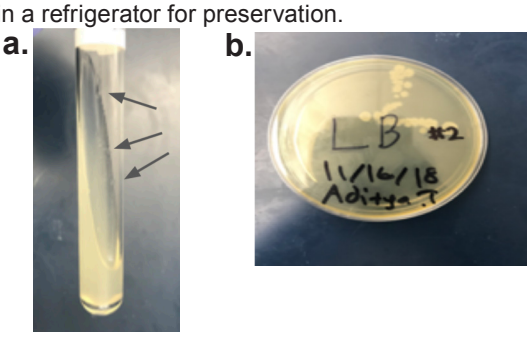


Figure 5. A) *B. subtilis* is visible as small bumps in the test tube with agar. B) *B. subtilis* colonies are visible as small white spots in the petri dish with agar.

Test Tube Setup

Out of the 72 test tubes, 36 contained LDPE and 36 contained HDPE. Each polyethylene piece was cut in roughly the same square shape, with a uniform mass of approximately 17 mg. Within both groups, 24 test tubes contained bacteria, while the remaining 12 tubes served as controls without bacteria. Both control groups and bacteria groups were split into three subgroups, containing freshwater, ocean-water, or brackish water.

Water Setup and Incu-shaker

Sterilized water was added to each test tube. For test tubes in the ocean water group, a premade powder called Instant Ocean was added to achieve a salt concentration of roughly 35 parts per thousand, which is the average salinity of oceans (40). For test tubes in the brackish water group, Instant Ocean powder was added to achieve a salt concentration of roughly 15 parts per thousand. An identical amount of salt medium of very low concentration was added to all test tubes to support the bacteria's growth. All concentrations remained the same with the addition of the growth medium (2 g $NaNO_3$, 0.5 g $MgSO_4$, 0.5 g KCl, 0.01 g $Fe_2(SO4)_3$, 0.14 g KH_2PO_4, 1.2 g K_2HPO_4, 0.02 g yeast extract, 1 L water). All test tubes were then placed inside a shaking incubator set at 35°C.

Mass Measurement and Analysis

After 4 weeks, each polyethylene piece was removed with tweezers and cleaned of bacteria/salt using a 0.1M HCl wash and water. Each piece was placed on a paper towel and heated to 40-50ºC. Then, the mass of the piece was measured on a scale and recorded. The data was analyzed utilizing the statistical and graphing functions of Microsoft Excel.

Received: October 22, 2019
Accepted: April 6, 2020
Published: May 24, 2020

REFERENCES

1. "Plastics: Material-Specific Data." EPA, E*nvironmental Protection Agency*, 30 Oct. 2019, www.epa.gov/facts-and-figures-about-materials-waste-and-recycling/plastics-material-specific-data.
2. Hopewell, Jefferson, et al. "Plastics Recycling: Challenges and Opportunities." *Philosophical Transactions of the Royal Society B: Biological Sciences,* vol. 364, no. 1526, 2009, pp. 2115–2126., doi:10.1098/rstb.2008.0311.
3. Notten, Philippa. "United Nations Environment Programme (2018). Addressing Marine Plastics: A Systemic Approach - Stocktaking Report. Notten, P. United Nations Environment Programme." *United Nations Environmental Program*, wedocs.unep.org.
4. Stegman, Robyn, et al. "Plastics in the Ocean." *Ocean Conservancy*, oceanconservancy.org/trash-free-seas/plastics-in-the-ocean/.
5. Barboza, Luís Gabriel Antão, et al. "Marine Microplastic Debris: An Emerging Issue for Food Security, Food Safety and Human Health." *Marine Pollution Bulletin,* vol. 133, 2018, pp. 336–348., doi:10.1016/j.marpolbul.2018.05.047.
6. "Plastics in the Ocean Affecting Human Health." *Case Studies*, 13 Jan. 2020, serc.carleton.edu/NAGTWorkshops/health/case_studies/plastics.html.
7. "Poly(Ethene) (Polyethylene)." *Poly(Ethene) (Polyethylene)*,www.essentialchemicalindustry.org/polymers/polyethene.html
8. Glaser, John A. "Biological Degradation of Polymers in the Environment." *IntechOpen*, 13 May 2019, www.intechopen.com/books/plastics-in-the-environment/biological-degradation-of-polymers-in-the-environment.
9. "Biodegradation." *Biodegradation - an Overview | ScienceDirect Topics,* www.sciencedirect.com/topics/materials-science/biodegradation.
10. Siracusa, Valentina. "Microbial Degradation of Synthetic Biopolymers Waste." *Polymers,* vol. 11, no. 6, 2019, p. 1066., doi:10.3390/polym11061066.
11. Ray, Suprakas Sinha. "Environmentally Friendly

Polymer Matrices for Composites." *Environmentally Friendly Polymer Nanocomposites,* 2013, pp. 25–40., doi:10.1533/9780857097828.1.25.
12. Zhang, Ling-Juan, and Richard L. Gallo. "Antimicrobial Peptides." *Current Biology*, vol. 26, no. 1, 2016, doi:10.1016/j.cub.2015.11.017.
13. Society, Microbiology. "Microbes and the Human Body: What Is Microbiology?" *What Is Microbiology?* | *Microbiology Society*, 12 Nov. 2019, microbiologysociety.org/why- microbiology-matters/what-is-microbiology/microbes-and- the-human-body.html.
14. Meena, Khem Raj, and Shamsher S. Kanwar. "Lipopeptides as the Antifungal and Antibacterial Agents: Applications in Food Safety and Therapeutics." *BioMed Research International,* vol. 2015, 6 Jan. 2015, pp. 1–9., doi:10.1155/2015/473050.
15. Sizar, Omeed. "Gram Positive Bacteria." StatPearls [Internet]., *U.S. National Library of Medicine*, 20 Nov. 2019, www.ncbi.nlm.nih.gov/books/NBK470553/.
16. Sizar O, Unakal CG. Gram Positive Bacteria. [Updated 2019 Dec 20]. *In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing;* 2020 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK470553/
17. Borriss, Rainer, et al. "*Bacillus Subtilis,* the Model Gram-Positive Bacterium: 20 Years of Annotation Refinement." *Microbial Biotechnology*, John Wiley and Sons Inc., Jan. 2018, www.ncbi.nlm.nih.gov/pmc/articles/PMC5743806/.
18. "Surfactant." *Surfactant - an Overview* | *ScienceDirect Topics,* www.sciencedirect.com /topics/chemistry/surfactant.
19. Alshehrei, Fatimah. "Biodegradation of Synthetic and Natural Plastic by Microorganisms." *Journal of Applied & Environmental Microbiology* 5.1 (2017): 8-19.
20. Santos, Danyelle Khadydja F, et al. "Biosurfactants: Multifunctional Biomolecules of the 21st Century." *International Journal of Molecular Sciences,* MDPI, 18 Mar. 2016, www.ncbi.nlm.nih.gov/pmc/articles/PMC4813256/.
21. Banat, I M, et al. "Potential Commercial Applications of Microbial Surfactants." *Applied Microbiology and Biotechnology, U.S. National Library of Medicine*, May 2000, www.ncbi.nlm.nih.gov/pubmed/10855707.
22. Wu, Yuan-Seng, et al. "Anticancer Activities of Surfactin and Potential Application of Nanotechnology Assisted Surfactin Delivery." *Frontiers in Pharmacology,* Frontiers Media S.A., 26 Oct. 2017, www.ncbi.nlm.nih.gov/pmc/articles/PMC5662584/.
23. Seydlová, Gabriela, and Jaroslava Svobodová. "Review of Surfactin Chemical Properties and the Potential Biomedical Applications." *Open Medicine, Versita,* 22 June 2015, www.degruyter.com/view/j/med.2008.3.issue-2/s11536-008-0002-5/s11536-008-0002-5.xml.
24. Yang, Jun, et al. "Evidence of Polyethylene Biodegradation by Bacterial Strains from the Guts of Plastic-Eating Waxworms." *Environmental Science & Technology,* vol. 48, no. 23, 2014, pp. 13776–13784., doi:10.1021/es504038a.
25. Simpson, D. Randall, et al. "Biosurfactant-Producing Bacillus Are Present in Produced Brines from Oklahoma Oil Reservoirs with a Wide Range of Salinities." *Applied Microbiology and Biotechnology,* vol. 91, no. 4, 2011, pp. 1083–1093., doi:10.1007/s00253-011-3326-z.
26. Painchaud, J et al. "Assessment of salinity-related mortality of freshwater bacteria in the saint lawrence estuary." *Applied and environmental microbiology* vol. 61,1 (1995): 205-8.
27. Vimala, P.p., and Lea Mathew. "Biodegradation of Polyethylene Using *Bacillus Subtilis*." *Procedia Technology*, vol. 24, 2016, pp. 232–239., doi:10.1016/j.protcy.2016.05.031.
28. Peters, Adrian C., et al. "Effects of Salt Concentration on Bacterial Growth on Plates with Gradients of PH and Temperature." *FEMS Microbiology* Letters, vol. 77, no. 2-3, 1991, pp. 309–314., doi:10.1111/j.1574-6968.1991.tb04367.x.
29. Fan, L., et al. "Genome Sequence of *Bacillus Subtilis* Subsp. Spizizenii gtP20b, Isolated from the Indian Ocean." *Journal of Bacteriology*, vol. 193, no. 5, 2010, pp. 1276–1277., doi:10.1128/jb.01351-10.
30. Ivanova EP, et al. "Characterization of Bacillus Strains of Marine Origin." *NCBI*, Dec. 1999, pp. 267–271.
31. Makkar, et al. "Optimization of Surfactin Production by *Bacillus Subtilis* Isolate BS5." *Applied Biochemistry and Biotechnology,* Humana Press Inc, 1 Jan. 1999, link.springer.com/article/ 10.1007/s12010-008-8155-x.
32. US Department of Commerce, and National Oceanic and Atmospheric Administration. "How Does the Temperature of Ocean Water Vary?" *Ocean Exploration Facts: NOAA Office of Ocean Exploration and Research*, 5 Mar. 2013, oceanexplorer.noaa.gov/facts/temp-vary.html.
33. "Understanding the Science of Ocean and Coastal Acidification." *EPA, Environmental Protection Agency,* 23 Aug. 2019, www.epa.gov/ocean-acidification/understanding-science-ocean- and-coastal-acidification.
34. Freitas, Bruno G, et al. "Formulation of a Commercial Biosurfactant for Application as a Dispersant of Petroleum and By-Products Spilled in Oceans." *Frontiers in Microbiology, Frontiers Media* S.A., 18 Oct. 2016, www.ncbi.nlm.nih.gov/pmc/articles/PMC5067411/.
35. Zouari, Raida, et al. "Optimization of *Bacillus Subtilis* SPB1 Biosurfactant Production Under Solid-State Fermentation Using By-Products of a Traditional Olive Mill Factory." *Achievements in the Life Sciences,* 29 Apr. 2015, www.sciencedirect.com/ science/article/pii/S2078152015000188.
36. Bonhomme, S, et al. "Environmental Biodegradation of Polyethylene." *Polymer Degradation and Stability,* Elsevier, 11 June 2003, www.sciencedirect.com/science/article/pii/ S0141391003001290.
37. Wisjnuprapto, N.a., et al. "Strategies toward Commercial Scale of Biosurfactant Production as Potential Substitute

for It's Chemically Counterparts." *International Journal of Biotechnology,* vol. 12, no. 1/2, 2011, p. 66., doi:10.1504/ijbt.2011.042682.

38. Freitas, Bruno G, et al. "Formulation of a Commercial Biosurfactant for Application as a Dispersant of Petroleum and By-Products Spilled in Oceans." *Frontiers in Microbiology,* Frontiers Media S.A., 18 Oct. 2016, www.ncbi.nlm.nih.gov/pmc/articles/PMC5067411/.
39. Parthipan, Punniyakotti, et al. "Biosurfactant and Degradative Enzymes Mediated Crude Oil Degradation by Bacterium *Bacillus Subtilis* A1." *Frontiers in Microbiology,* Frontiers Media S.A., 9 Feb. 2017, www.ncbi.nlm.nih.gov/pmc/articles/PMC5299021/.
40. US Department of Commerce, and National Oceanic and Atmospheric Administration. "Why Is the Ocean Salty?" *NOAA's National Ocean Service*, 14 Nov. 2008, oceanservice.noaa.gov/ facts/whysalty.html.

JOURNAL OF
EMERGING INVESTIGATORS

Article

Banana-based biofuels for combating climate change: How the composition of enzyme catalyzed solutions affects biofuel yield

Hermann F. Klein-Hessling Barrientos, Rachel Ingram
United World College South East Asia, Dover Campus, Singapore

SUMMARY

This investigation was inspired by the ever so precarious nature of climate change and its detrimental implications on our planet and society. One alternative form of energy that shows promise amongst others in alleviating the climate crisis, is biofuel. An extensive experiment was conducted to investigate whether amylase or yeast had a more prominent role in determining the bioethanol concentration and bioethanol yield of banana samples. Given that amylase facilitates the breakdown of complex polysaccharides into simple sugars, we hypothesized that amylase would have the most significant impact on the bioethanol yield and concentration of the samples. This was determined to be the case when the fermentation batch samples received enzyme compositions of yeast and amylase independently. However, when added concurrently, the optimal enzyme composition for maximizing the bioethanol concentration and yield was the addition of amylase and yeast in a 3:2 ratio. Thus, while yeast is an essential component for producing bioethanol, the proportion of amylase supplied through a joint amylase-yeast mixture has a more significant impact on the bioethanol yield. This provides some insight into what industrial engineers would need to take into consideration when mass-producing bioethanol on the industrial scale, namely that of the optimal enzyme composition with respect to cost and the bioethanol yield. Moreover, this study provides a greater understanding of the mechanisms and implications involved in enzyme-based biofuel production, specifically of those pertaining to amylase and yeast.

INTRODUCTION

Climate change and how to address its detrimental implications have recently become the center of discussion. The Intergovernmental Panel on Climate Change (IPCC) published its most recent report emphasizing the necessity of reforming the way we obtain our energy and the reliance we have on non-renewable fossil fuels (1). The IPCC has stressed the fact that if we do not prevent the global temperature from rising 1.5°C, there will be ravaging ramifications on us and our planet (1). Biofuels derived from biomass — any matter originating from plants or animals — show particular promise for combating climate change. They are advantageous because they emit, on average, up to 65% fewer greenhouse gases than conventional petroleum. Additionally, the crops grown for biofuels absorb CO_2 from the surrounding atmosphere and mitigate the carbon footprint created during cultivation, potentially making it carbon neutral (2-3).

Industrially produced biofuel requires two stages, breakdown and fermentation. First, the biomass must be pre-treated either through acid, base, or enzymatic hydrolysis to disrupt the hemicellulose and lignin and facilitate the breakdown of complex sugars to simple sugars (4). Once this process is complete, the biomass is subjected to fermentation, and bioethanol is consequently formed. We chose bananas for this investigation since they are easy to manage as a plantation crop, they readily undergo fermentation and have an average yield of 40-50 tons per hectare (5). In addition to this, 10-15% of the dry mass of bananas consists of simple sugars, which proves especially favourable for fermentation (6). The peels of bananas are commonly discarded and also possess potential for bioethanol production, proving especially advantageous given that banana peels are not relied on as a food source (7).

The biological conversion of biomass to fuels can offer the high yields necessary to be economically viable by lowering economic cost. Extensive literature exists on chemical and physical conditions for optimizing bioethanol yields. For instance, a pH of 4.5-5.5 during enzymatic pre-treatment by amylase increases simple sugar production and a pH of 6.7-7.0 for yeast facilitates fermentation (8). There also exists considerable research into the various methods of pre-treatment for the banana to be broken down into glucose and consequently form bioethanol (9). Biological means of facilitating these processes like enzymatic hydrolysis are particularly attractive. Enzymatic hydrolysis, which promotes the conversion of biomass to fermentable sugars, is arguably the most complex type because the enzymatic interactions and mechanisms are not yet well understood. Therefore, we will investigate these biological processes as they offer greater potential yields, lower energy costs and milder

operation conditions than other chemical processes such as acid-base hydrolysis (10).

This manuscript attempted to provide insight and a better understanding of the mechanisms involved specifically with enzymatic hydrolysis. We tested varying compositions of amylase and yeast to optimize the bioethanol yield through the fermentation of banana samples. We believe that these results will provide important considerations for other fruits and biomaterials which are not used as a food source, such as banana peels or corn husks.

Given that the rate of breakdown and fermentation is dependent on the amount of substrate (glucose) available, greater amounts of both amylase and yeast being supplied to the banana samples should increase the rate. We hypothesize the increase in both amylase and yeast compositions of the batch samples will lead to the proportional increase in glucose concentration post enzymatic hydrolysis.

RESULTS

The breakdown of the banana batch samples through enzymatic hydrolysis and the consequent fermentation of bioethanol would allow us to determine the glucose and bioethanol concentration of the samples. In order to do so the banana batch samples independently received varying compositions of amylase (A1-A6) and yeast (S1-S6) (**Figure 1A-B**). A combined amylase-yeast batch (C1-C6) was also assembled to allow for greater scrutiny in ascertaining the impact of either amylase or yeast in the bioethanol yield of banana samples (**Figure 1C**). All samples received pre-established amounts (**Figure 1**) of their conjugate solutions to ensure enzymatic hydrolysis and fermentation occurred and also so that differences in glucose and bioethanol concentration could be scrutinized appropriately (**Figure 1**). The samples were subjected to enzymatic pre-treatment by the addition of their respective quantities of amylase for a duration of 15 minutes. Afterwards, we took 1 cm³ aliquots of the samples and measured their absorbance by spectrophotometer. We compared these to a pre-established standardised glucose calibration curve to determine their glucose concentrations. (**Figure 2E**). For the amylase (A1-A6) and yeast (S1-S6) fermentation batches, the greater the proportion of amylase and yeast supplied the higher the consequent glucose formation post-enzymatic hydrolysis (**Figure 2A-B**).

The interaction between the yeast and amylase supplied to the banana fermentation batch samples (the substrate) took place in aqueous solution. Therefore, the breakdown of complex sugars into glucose and the consequent conversion into bioethanol was a result of homogenous catalysis (11). The general reaction scheme of an enzyme-catalysed reaction is as follows:

$$E + S \overset{k_1}{\rightarrow} [ES] \overset{k_2}{\rightarrow} E + P \quad (11).$$

In this investigation, the amylase interacted with the banana samples (substrate) by binding to its active sites to form the enzyme substrate complex *[ES].* This is followed by the decomposition of *[ES]* to regenerate the amylase and to form the new product glucose (*P).* This same process occurs during the fermentation of the banana samples but instead bioethanol (*P*) is formed as the result of yeast catalysis. Enzymatic hydrolysis occurred for fifteen minutes, and we assumed that the correlation between the enzyme composition of the banana samples and the glucose concentration to be akin to that of a first order reaction. This assumption would also hold true for the bioethanol concentration of the banana samples since the more glucose available, the more that could undergo fermentation. Modelling the expected rate equation for the enzymatic hydrolysis of the banana samples and for the bioethanol consequently formed after fermentation can be represented as:

$$r = k[\![Yeast]\!]\,[\![Amylase]\!].$$

This is supported by a variation in the Michaelis-Menten equation which employs the steady-state approximation. It states that for initial enzyme concentrations and low values of substrate (*S*) as employed in this study, the rate of product

a)

Enzyme Batch/Samples	Yeast (cm³)	Water (cm³)
S1	10	0
S2	8	2
S3	6	4
S4	4	6
S5	2	8
S6	0	10

b)

Enzyme Batch/Samples	Amylase (cm³)	Water (cm³)
A1	10	0
A2	8	2
A3	6	4
A4	4	6
A5	2	8
A6	0	10

c)

Enzyme Batch/Samples	Amylase (cm³)	Yeast (cm³)
C1	10	0
C2	8	2
C3	6	4
C4	4	6
C5	2	8
C6	0	10

Figure 1. The varying enzyme compositions of the fermentation batch samples. The three enzyme induced fermentation batches were categorized as follows: **(a)** yeast, S1-S6, **(b)** amylase (A1-A6) and **(c)** amylase-yeast combined (C1-C6). To ensure the successful breakdown of complex sugars and fermentation batch samples S1-S6 received 2 cm³ of amylase, A1-A6 received 2 cm³ of yeast and C1-C6 2 cm³ of water. The percentage concentration of both the amylase and yeast solutions were 1%.

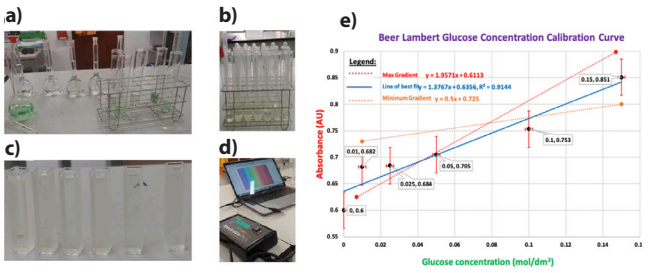
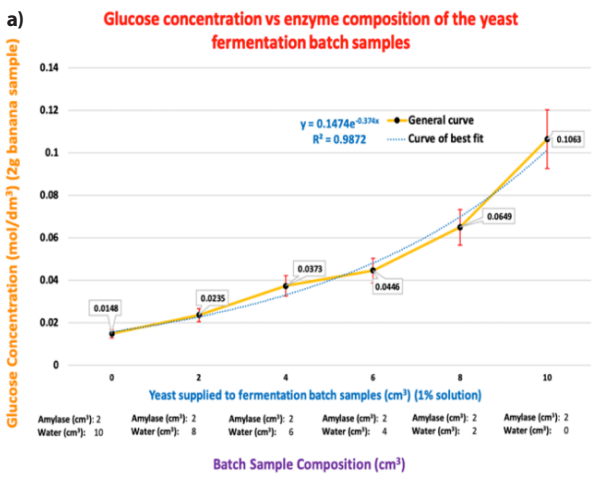
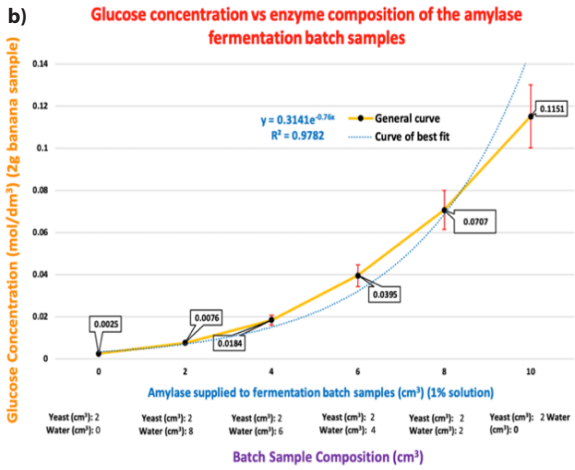


Figure 2. Establishment of the Beer-Lambert's glucose calibration curve. Max-Min gradients are extrapolations of the potential discrepancies in the absorbance values as a result of the ± 13% error in the spectrophotometer. **(a)** Preparation of the Quantitative Benedict's Solution (QBS) used for quantitatively determine the glucose concentrations of the varying fermentation batch samples by spectrophotometer. **(b)** 10 cm³ of QBS was mixed with 1 cm³ of solution of the fermentation batch samples. **(c)** The range in the shade of color emerging from the samples indicated the relative presence of glucose. **(d)** These in turn corresponded to different absorbance values recorded by spectrophotometer. **(e)** Standardised Beer-Lambert calibration curve of glucose which was established.





(a, i)

Sample	Yeast (cm^3)	Water (cm^3)	Glucose concentration (mol/dm^3)	% Glucose yield of total banana mass (2g)
A1	10	0	0.0446	16.07
A2	8	2	0.0395	14.23
A3	6	4	0.0381	13.73
A4	4	6	0.0373	13.44
A5	2	8	0.0235	8.47
A6	0	10	0.0184	6.63

(b, i)

Sample	Yeast (cm^3)	Water (cm^3)	Glucose concentration (mol/dm^3)	% Glucose yield of total banana mass (2g)
S1	10	0	0.017	6.13
S2	8	2	0.0148	5.33
S3	6	4	0.0076	2.74
S4	4	6	0.0061	2.20
S5	2	8	0.0025	0.90
S6	0	10	0.0003	0.11

Figure 3. Correlation between the compositions of the amylase and yeast fermentation batches with respect to their glucose concentration and yield. (a, b) The increase in glucose concentration as a result of the increasing proportion of amylase and yeast in the fermentation batch samples. **(a, i)** The % glucose yields of the yeast samples with respect to the initial banana mass. **(b, i)** The % glucose yields of the amylase samples with respect to the initial banana dry mass.

formation (glucose and bioethanol) is directly proportional to the amount of substrate initially present (*12*). This can be represented as follows:

$$r = \frac{k_2 [E_0][S]}{K_m + [S]}$$

where *[E0].* is the total enzyme concentration and Km the Michaelis constant, a measure of binding affinity which is a way to determine the compatibility of an enzyme with the active site of a given substrate (13).

However, the relationship between the batch compositions of the banana samples and their respective glucose concentrations post-enzymatic hydrolysis was not first order and hence not linear (**Figure 3**). In fact, we found a strong exponential correlation between the glucose concentration post enzymatic hydrolysis of the amylase (R^2 = 0.9872) and yeast (R^2 = 0.9782) fermentation batch samples and their respective compositions. This discrepancy from our hypothesis might be due to the addition of yeast to the samples. Although the yeast is supposed to catalyze fermentation, adding it concurrently with amylase rather than separately may have impacted the concentrations of glucose post enzymatic hydrolysis (**Figure 4A**) The non-linear proportional increase in glucose concentration therefore could be a result of the interaction or interference between the yeast and amylase, also known as enzyme inhibition.

The banana fermentation batch samples were then incubated within water baths at the optimal pH (yeast 4.5-5.5 and amylase 6.7-7.0) and temperature conditions (36.8°C) conducive to fermentation (**Figure 4B**). The samples were then transferred to conical flasks and suspended over dichromate solution to undergo reduction overnight (**Figure 4C**). This was done since alcoholic beverages contain other oxidizable substances that could interfere with the titration (14). The ethanol produced as a result of fermentation oxidized to ethanoic acid by reacting with an excess of potassium dichromate in acid. The unreacted excess was determined by the addition of potassium iodide solution which was oxidized by the potassium dichromate, forming iodine. Then, the iodine was titrated against a standard solution of sodium thiosulfate to calculate the original bioethanol content present in the banana fermentation batch samples (**Figure 4D-Eiii**).

The titrations soon revealed that more amylase and yeast yielded higher bioethanol amounts since more glucose was readily available to undergo fermentation. (**Figure 5A-B)**. The graphs had a strong adherence to the data points plotted and R^2 values of 0.984 (amylase) and 0.994 (yeast), which indicated a strong linear correlation between the two variables. Unlike with the glucose, the bioethanol concentration of the amylase and yeast fermentation batch samples was directly proportional to their compositions as stipulated by the Michaelis-Menten equation. The concurrent addition of amylase and yeast in the amylase-yeast fermentation batch samples allowed for the direct gauging

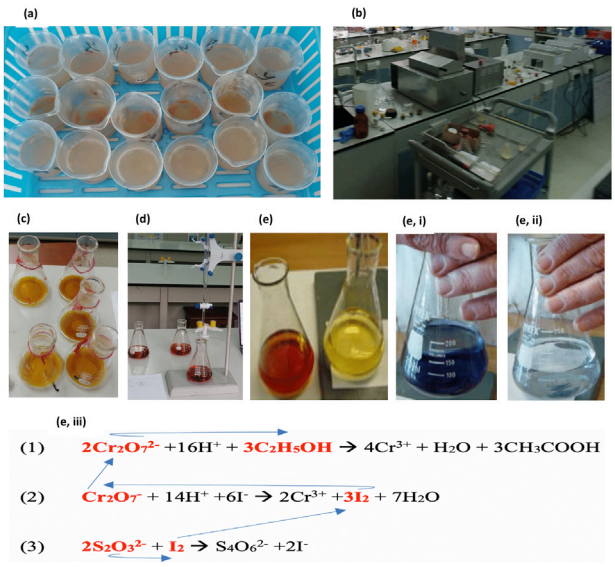


Figure 4. Enzymatic Pre-treatment, fermentation and bioethanol formation. (a) Fermentation batch samples post enzymatic pretreatment being transferred for overnight fermentation. **(b)** The incubation of the banana batch samples for optimal fermentation conditions. **(c)** The batch samples were suspended over dichromate solution and left to reduce overnight. Variations in the extent of reduction arose, indicating tangible differences in the bioethanol concentrations of the banana batch samples. **(d)** The extent of ethanol oxidation indicates the initial concentration which can be determined by redox back titration. **(e)** Left flask shows the brown coloured solution resulting from the formation of iodine. The other indicates the color transition which occurs as the iodine is titrated with thiosulfate. **(e, i)** Upon the addition of starch the analyte solution takes becomes blue due to a formation of a starch-iodine complex. **(e, ii)** The endpoint of the titration occurs when the solution becomes colorless. **(e, iii)** Chemical pathway of the redox titration for determining the bioethanol concentration of the banana fermentation batch samples.

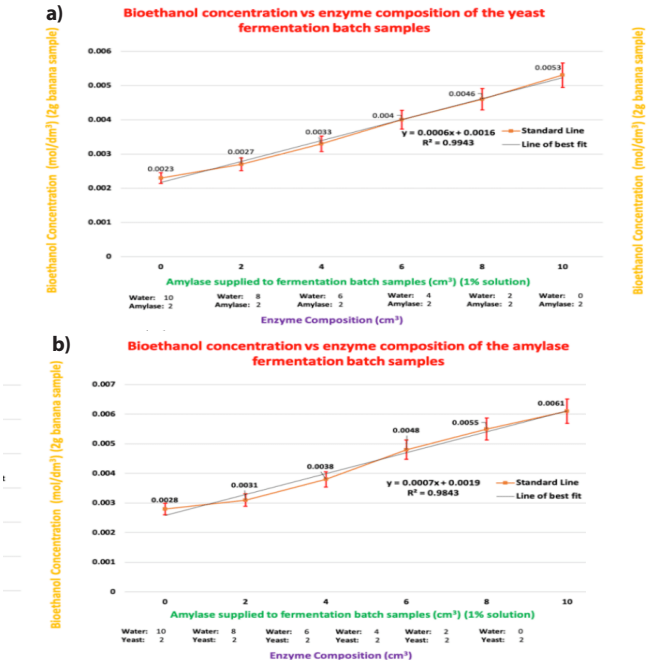
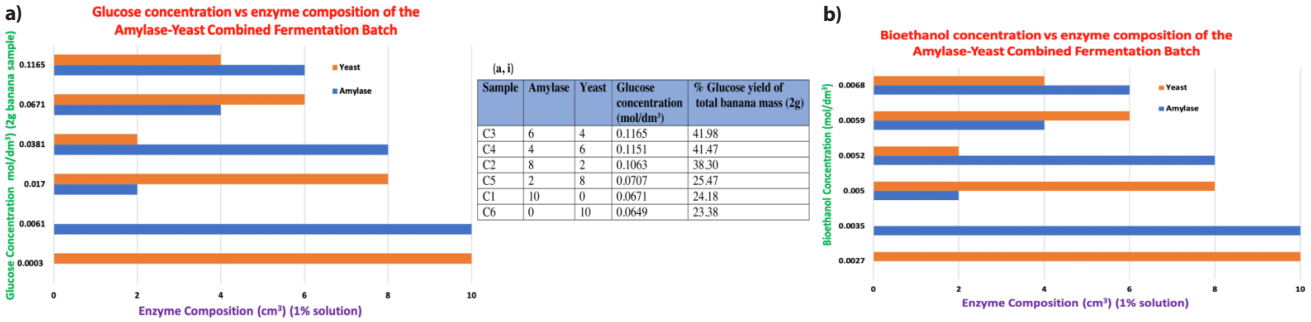


Figure 5. The correlation between the compositions of the amylase and yeast fermentation batch samples and their bioethanol yields. (a, b) The increasing prominence of amylase and yeast as they compromise more of the total batch sample composition and the associated linear increase in bioethanol concentration. This is to be expected given the prior increase in glucose concentration. A greater proportion of glucose available allowed the samples to undergo fermentation and form bioethanol more effectively.

of their relative prominence with regards to their glucose and bioethanol yields. The trend for this fermentation batch was more complex. Sample C1 (**Figure 1**), with 10 cm³ of amylase, did not produce the greatest concentration of glucose, but rather sample C3, with a composition of 6 cm³ of amylase and 4 cm³ of yeast followed by samples C4, C5, C1, and C6 (**Figure 6A**). Not only did this combination outperform the other samples of its fermentation batch with 0.1165 mol/dm³, but also that of the amylase and yeast batch samples (**Figure 3**). Consequently, the same compositions of amylase and yeast produced the greatest bioethanol concentrations. It is clear that for the combined amylase-yeast fermentation batch samples, a greater percentage of amylase yielded greater glucose concentrations after enzymatic hydrolysis and ultimately bioethanol post fermentation (**Figure 5A-B**).

DISCUSSION

To better understand the implications of the results pertaining to the combined fermentation batch samples, it is worthwhile to compare the enzyme composition of samples C3 and C5 in particular. Sample C5 is the conjugate in terms of composition, of sample C3 with 6 cm³ of yeast and 4 cm³ of amylase. Regardless, C3 still yielded a greater glucose concentration after pre-treatment, as well as a higher bioethanol concentration. This is most likely the case because the amylase acts as the 'solvent' and breaks down the complex sugars — disaccharides and polysaccharides — into simple sugars like glucose of our substrate, which is our varying banana fermentation batch samples. This process is known as glycolysis and is an essential step in the production of bioethanol from organic matter (15). A greater amylase concentration would ensure a more efficient glycolysis process resulting in a greater amount of complex sugars being broken down, resulting in a greater amount of glucose readily available to undergo fermentation. (**Figure 7A iii**). Essentially, amylase facilitates the glycolysis process, the most important step in the lead up to ethanol formation. Since amylase determines the amount of substrate that will be available for the conversion to ethanol to take place, compositions favouring yeast like that of C5 and C6 would certainly facilitate the oxidation of NADH to NAD+ allowing for further glycolysis to occur (15). In addition to this, amylase enhances the efficiency of the conversion of acetaldehyde to ethanol. However, given the limited quantity of the substrate in the first place, the formation of ethanol would be curtailed, and the re-initiation of glycolysis would not remain feasible for long. Given these direct comparisons between the varying



(a, i)

Sample	Amylase	Yeast	Glucose concentration (mol/dm³)	% Glucose yield of total banana mass (2g)
C3	6	4	0.1165	41.98
C4	4	6	0.1151	41.47
C2	8	2	0.1063	38.30
C5	2	8	0.0707	25.47
C1	10	0	0.0671	24.18
C6	0	10	0.0649	23.38

Figure 6. Correlation between the ratio of amylase to yeast enzyme in sample composition and their glucose/bioethanol yields. (a) The increase in glucose concentration as a result of the changing proportions of amylase and yeast in the fermentation batch samples. **(b)** The % glucose yields of the amylase-yeast combined samples with respect to the initial banana dry mass. **(c)** Sample C3 with 4 cm³ of yeast and 6 cm³ of amylase supplied formed the greatest glucose concentration of 0.1165 mol/dm³ post enzymatic pre-treatment. Naturally, it yielded the greatest bioethanol concentration given that a greater proportion of glucose was available to undergo fermentation and form bioethanol.

compositions of the amylase-yeast combined fermentation batch, we soon realised that amylase in conjunction with yeast, even when in lower concentrations, takes precedence in determining the concentration and yield of bioethanol for a given fermentation batch sample. Moreover, the optimal batch composition consists of a greater proportion of amylase to yeast to ensure the sufficient breakdown of complex sugars and conversion of acetaldehyde to ethanol, resulting in greater bioethanol yields. In this experiment, the optimal composition was revealed to be C3 consisting of 6 cm³ of amylase and 4 cm³ of yeast, which is analogous to a 3:2 ratio. Lastly, the glucose yields with respect to theoretical presence of glucose in the 2 g banana for the varying samples was impressive with some yielding over 30%, twice that of the most optimistic yield predicted (**Figure 6B**).

Given the extensive experimental procedure and use of equipment for the investigation, there is room for potential error. We used the redox-titration to determine the ethanol concentration of the fermentation batch samples and the standardized Beer-Lambert calibration curve to measure the glucose concentrations. During the fermentation process, Parafilm strips were used to cover the flasks allowing for slight expansion due to an increase in CO_2 produced as a result of anaerobic respiration. Because of the expansive nature of the parafilm, some oxygen could have seeped into the flasks reducing the efficacy of the fermentation process, especially given its anaerobic nature. Additionally, due to time constraints, the fermentation batch samples only underwent fermentation for 24 hours as opposed to the optimal 72 hours. This reduced the confidence in the results as not all fermentation parameters were optimized, therefore the optimal enzyme composition may have differed slightly.

The preliminary redox titration was conducted with beer that contained 5% ethanol. There was a discrepancy between the calculated ethanol concentration and the actual of 13% which was not too disconcerting. However, the potential error in employing this method for the fermentation batch samples would be far greater given the smaller ethanol concentrations of the banana samples. For the redox titrations of the samples, the final colour transition from blue to colourless did not occur at all. The addition of H^+ ions (by concentrated H_2SO_4) to the dichromate solutions had no bearing on this, and the nature of the occurrence remained elusive. However, the penultimate color transition remained mostly consistent, which was reassuring (**Figure 8A**). Lastly, some of the aliquots taken from the fermentation samples contained particle contamination (**Figure 8B**), which influenced the recorded absorbance values compared to the standardized glucose calibration curve that was established.

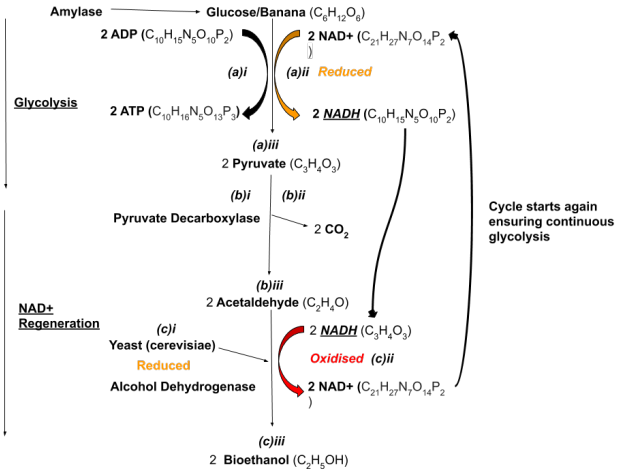


Figure 7. Chemical pathway of ethanol fermentation. Glycolysis of the substrate occurs, in this case the banana samples. **(a, i)** Two molecules of ATP are formed. **(a, ii)** NAD+ is reduced to NADH by the addition of a hydride ion. **(a, iii)** Two molecules of Pyruvate are formed. **(b-c)** NAD+ regeneration. **(b, i)** Pyruvate loses its carboxyl group, facilitated by the enzyme pyruvate decarboxylase. **(b, ii)** CO_2 is released due to carboxyl group leaving. **(b, iii)** Results in the formation of two molecules of acetaldehyde. **(c, i)** Addition of yeast catalyses the reduction of acetaldehyde to bioethanol. **(c, ii)** The acetaldehyde is reduced to ethanol, in this case bioethanol by the addition of a hydride ion. This conversion is facilitated by the enzyme alcohol dehydrogenase. **(c, iii)** by the addition of a hydride ion. This conversion is facilitated by the enzyme alcohol dehydrogenase. **(c, iii)** Whilst this occurs the NADH from earlier is becoming oxidised this ensures a plentiful regeneration of further NAD+ which allows glycolysis to re-initiate and the further formation of ethanol.

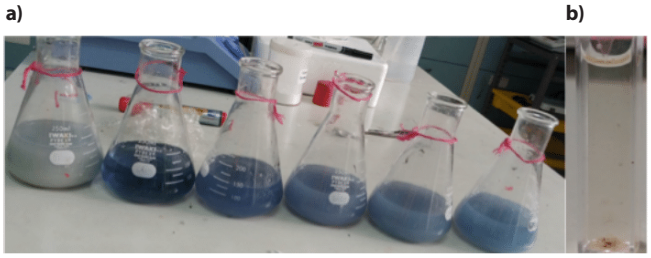


Figure 8. Discrepancy in color transition of redox titration and particle contamination during colorimetry. (a) Samples turned navy blue during the penultimate transition however, shortly afterwards different shades of blue began to emerge. **(b)** Cuvette sample after it was mixed with QBS which turned slightly orangered indicating a relatively strong presence of glucose. The sample contained some particle contaminate potentially stemming from banana remnants which were not completely broken down as a result of enzymatic pre-treatment.

Interestingly, it was observed that the particle contamination was more prominent for fermentation samples that were diluted in terms of enzyme composition (concentration) for both the amylase and yeast fermentation batches. In the combined batch samples, particle contamination became increasingly prevalent the further it deviated from the optimal enzyme composition observed in sample C3 (6 cm^3 of amylase and 4 cm^3 of yeast). If time had allowed, the fermentation batch samples could have been filtered to reduce the particle contamination, providing more accurate absorbance readings. The spectrophotometer used had an error deviation of 13%, which was unfavourable. More accurate equipment would have lowered the error deviation and re-affirmed confidence in the results obtained. Overall, these occurrences contributed the greatest to the systematic error pertaining to the bioethanol produced.

To better understand the use of banana-based biofuel, the mechanisms of glycolysis and fermentation which are catalyzed by amylase and yeast, respectively, need to be analysed and substantiated. Ascertaining the Michaelis-Menten constant of both reaction processes would determine the binding affinity between the substrate's active sites and that of the amylase and yeast used in this investigation. This could provide further insight into their 'compatibility.'

Food waste is the largest greenhouse gas contributor after China and the United States. Our experiments were conducted with fresh bananas but repeating these experiments with a decomposing banana would be worthwhile given the significant potential greenhouse gas savings in preventing fruit from decomposing and releasing methane in landfills (16). It is likely that there would be differences in the bioethanol and glucose concentrations of the fermentation samples depending on the extent of decomposition of the banana. Moreover, determining whether the efficacy or efficiency of fermentation changes based on the extent of said decomposition would be noteworthy. Given the potential benefits of algae-based biofuels, specifically its high energy content and CO_2 absorption, it would be worthwhile to investigate the optimal enzyme composition for the ethanol fermentation of algae. (17-18) Testing various strains of algae and investigating which strain produced the greatest amount of bioethanol in combination with the optimal enzyme composition would be valuable.

The high costs of amylase and yeast would make it important to determine the optimal enzyme composition based on the efficiency of ethanol fermentation with respect to cost, on an industrial scale. While the variations in enzyme compositions of this investigation were rather limited, they could be extended or scaled up to simulate fermentation on the industrial scale. Lastly, it would be worthwhile to ascertain why the titrations of the fermentation batch samples did not transition to colorless. If this were resolved, greater confidence in the experimental results and conclusion would be established.

METHODS

Banana sample preparation

A fresh set of bananas (from a local grocery store) was procured and the peel was removed. Bananas were pulverized using a standard kitchen blender at maximum speed for one minute and then medium speed for an additional minute. Once this was complete, eighteen 100 cm^3 beakers were collected and zeroed on analytical balances. The beakers were pre-emptively labelled according to the fermentation batch they belonged to and with regards to their compositions of amylase and yeast. Then, 2 g of blended banana was measured into each of the beakers.

Redox titration

Blank titrations (those without any beer) were prepared by adding 10 cm^3 of acid dichromate solution to a 250 cm^3 conical flask. This was followed by the addition of 100 cm^3 of water and 1 cm^3 of potassium iodide solution. The resulting solution was thoroughly mixed. The solutions were then titrated against sodium thiosulfate. As the solution faded to yellow, 1 cm^3 of starch solution was added, resulting in the solution turning dark blue. This solution was further titrated against the sodium thiosulfate until the solution became colorless. The titre values were recorded, and the same process was repeated with acidified dichromate solution, which had been reduced overnight by malt beer containing 5% ethanol. The difference in titres between the blank titrations and those of the fermentation batch samples in conjunction with the redox equations previously mentioned were used to determine the ethanol concentration of the malt beer (14).The experimental procedure was repeated with all fermentation batch samples.

Establishment of the standardized glucose calibration curve

A range of known glucose concentrations was prepared to establish a standardized calibration curve. The following concentrations were prepared: 0.15 mol/dm^3, 0.1 mol/dm^3, 0.05 mol/dm^3, 0.025 mol/dm^3, and 0.01 mol/dm^3. The 0.15

mol/dm³ concentration of glucose as a percentage mass of 100 g would be commensurate to 15%, 0.1 mol/dm³ to 10%, 0.025 mol/dm³ to 2.5%, and that of 0.01 mol to 1%. This range would account for the most optimist glucose concentration (around 0.15 mol/dm³) obtained from the banana samples to a more pessimistic 1% (around 0.01 mol/dm³). 1 cm³ aliquots of the glucose samples were mixed with 10 cm³ of Quantitative Benedict's Solution (QBS) and then were placed in boiling tubes within a water bath at 80°C for 5 minutes. The solutions were then placed within cuvettes and their respective absorbances recorded by spectrophotometer and substantiated against a negative control of QBS (0.6 AU). These values were then used to establish a Beer-Lambert curve of the known glucose concentrations vs. their absorbance values.

Fermentation and determining bioethanol concentration

The previously mentioned banana sample solutions were then transferred to their respective conical flasks contingent on their amylase and yeast compositions. As much as possible of the remaining solution was transferred using a miniature spatula. The conical flasks were lightly sealed with paraffin strips to allow for any expansion that could occur due to the increase in CO_2 as a result of fermentation. The conical flasks were then categorically submerged based on the fermentation batch and sample composition within three water baths set to 36.8°C. The samples were left to ferment for 24 hours and were then removed and brought to room temperature. Eighteen 250 cm³ volumetric flasks with modified bungs to make use of sample holders were prepared. Then, 1 cm³ from each of the fermentation batch samples were transferred to the sample holders followed by the addition of 10 cm³ of potassium dichromate (0.1 mol) to each of the conical flasks. The samples were then suspended over the dichromate solution and the conical flasks were sealed by the modified bungs and left overnight in water baths at 36.8°C for the dichromate to reduce. The following morning, the conical flasks were brought to room temperature and the samples were removed. The dichromate solution was then titrated against sodium thiosulfate (0.3 mol) to determine the ethanol concentration of the fermentation batch samples (14).

Enzymatic hydrolysis and determining glucose concentration

Fermentation batch samples received 10 cm³ of pH 7 buffer solution followed by the addition of the respective amylase amounts for 15 minutes. Then, 10 cm³ of pH 4 buffer was supplied to all the samples, as well as their respective amounts of yeast and water. All samples were then shaken and stirred to ensure the contents of the solution that developed were mixed thoroughly. Next, 1 cm3 of aliquots were extracted from each of the samples and mixed with 10 cm3 of QBS and transferred to corresponding boiling tubes. The boiling tubes were then submerged within a water bath at 80°C for 5 minutes. The samples were then removed and their absorbance values were standardized against the negative control of the absorbance value of the QBS and the pre-established standardized glucose calibration curve.

ACKNOWLEDGEMENTS

I would like to thank Mr. Armstrong, Mrs. Walker, and Mr. Gillies for giving up their time to supervise me whilst I was doing my lab work. Another thank you to Mr. Gillies for procuring the beer necessary for my experimental methodology as I was not of legal age. I would also like to thank our schools chemistry technicians Wahida and Joy for helping me make up for lost time by preparing my experiments chemicals when they were spilled accidentally by classmates, and also for their company which made the long hours in the lab more bearable. Last but not least, I would like to thank my supportive girlfriend Debora for encouraging me to submit my manuscript and for sticking with me through the entire challenging process of publication, from start to finish.

Received: February 17, 2020
Accepted: May 20, 2020
Published: May 27, 2020

REFERENCES

1. Business Council for Sustainable Development in Hungary. "We Have 12 Years to Limit Climate Change Catastrophe, Warns UN." *BCSD Hungary*, 15 Oct. 2018, bcsdh.hu/we-have-12-years-to-limit-climate-change-catastrophe-warns-un/. Accessed 18 Feb. 2020.
2. "Biofuels versus Gasoline: The Emissions Gap Is Widening." *EESI*, 2 Sept. 2016, www.eesi.org/articles/view/biofuels-versus-gasoline-the-emissions-gap-is-widening.
3. Crisp, J. "Report: EU Underestimates Biofuels' Greenhouse Gas Savings by 50%." Www.euractiv.com, 14 Nov. 2014, www.euractiv.com/section/transport/news/report-eu-underestimates-biofuels-greenhouse-gas-savings-by-50/
4. Robak, K., and M. Balcerek. "Review of Second-Generation Bioethanol Production from Residual Biomass." *Food Technology and Biotechnology*, vol. 56, no. 2, 2018, pp. 174–187., doi:10.17113/ftb.56.02.18.5428.
5. Food and Agriculture Organization of the United States. "EST: Banana Facts." *Food and Agriculture Organization of the United Nations*, http://www.fao.org/economic/est/est-commodities/bananas/bananafacts/en/#.XrE_JpMzZQI Accessed 18 Feb. 2020.
6. Arnarson, A. "Bananas 101: Nutrition Facts and Health Benefits." *Healthline*, 7 May 2019, www.healthline.com/nutrition/foods/bananas#vitamins-and-minerals.
7. Gebregergs, A., *et al*. "Industrial ethanol from banana peels for developing countries: Response surface

methodology." *Pacific Science Review A: Natural Science and Engineering*, vol. 18, no. 1, 2016, pp. 22-29.
8. C.O.B., Okoye, *et al*. "Optimization of the Fermentation Conditions for the Production of Bioethanol from Cane Sugar Molasses using Saccharomyces cerevisiae." *The Pacific Journal of Science and Technology*, vol. 18, no. 2, Nov. 2017, pp. 48-56.
9. Sun, S., *et al*. "The role of pretreatment in improving the enzymatic hydrolysis of lignocellulosic materials." *Bioresource Technology*, vol. 199, 2016, pp. 49-58.
10. Yang, B., *et al*. "Enzymatic hydrolysis of cellulosic biomass." *Biofuels*, vol. 2, no. 4, 2011, pp. 421-449.
11. Tan, E., *et al*. "Michaelis-Menten Kinetics." *Chemistry LibreTexts*, 26 Oct. 2019, chem.libretexts.org/Bookshelves/Biological_Chemistry/Supplemental_Modules_(Biologica_Chemistry)/Enzymes/Enzymatic_Kinetics/Michaelis-Menten_Kinetics.
12. Burrows, A., *et al*. *Chemistry3: Introducing Inorganic, Organic and Physical Chemistry*. 3rd ed., Oxford UP, 2017. Enzyme Kinetics P. 437-441
13. "Enzyme Kinetics." *UW-Madison Department of Chemistry*, www.chem.wisc.edu/deptfiles/genchem/netorial/modules/biomolecules/modules/enzymes/enzyme4.htm
14. Pham, L. "Determination of Ethanol Concentration in Aqueous Solutions." *Academia.edu*, www.academia.edu/6314752/Determination_of_Ethanol_Concentration_in_Aqueous_Solutions.
15. Clark, J. "Fermentation." *Chemistry LibreTexts*, 5 June 2019, chem.libretexts.org/Bookshelves/Biological_Chemistry/Supplemental_Modules_(Biological_Chemistry)/Metabolism/Catabolism/Fermentation.
16. Hanson, C., *et al*. "What's Food Loss and Waste Got to Do with Climate Change? A Lot, Actually." *World Resources Institute*, 11 Dec. 2015, www.wri.org/blog/2015/12/whats-food-loss-and-waste-got-do-climate-change-lot-actually.
17. Gies, E. "Landfills have a huge greenhouse gas problem. Here's what we can do about it." *Ensia*, 25 Oct. 2016, ensia.com/features/methane-landfills/.
18. Offei, F., *et al*. "Seaweed Bioethanol Production: A Process Selection Review on Hydrolysis and Fermentation." *Fermentation*, vol. 4, no. 4, 2018, p. 99.

JOURNAL OF
EMERGING INVESTIGATORS

Article

Pollination patterns by Green-backed Firecrown hummingbirds

Martin Freeland[1], Juan G. Navedo[2], Tyler N. McFadden[3]

[1] Sacred Heart Schools, Atherton, California

[2] Bird Ecology Lab, Instituto de Ciencias Marinas y Limnológicas, Universidad Austral de Chile, Valdivia, Chile

[3] Department of Biology, Stanford University, Stanford, California

SUMMARY

The Green-backed Firecrown is the only hummingbird species found throughout the temperate rainforests of southern South America and is among the region's most important pollinators. A diverse assemblage of the region's flora appears to rely on the Green-backed Firecrown for pollination, which makes the firecrown vital for ecosystem function. However, relatively little is known of firecrown ecology. We examined the foraging patterns of Green-backed Firecrowns of various ages and both sexes, using pollen samples taken from birds in southern Chile. We hypothesized that males and females would carry distinct pollen communities (sets of pollen morphospecies) and would differ in the number of pollen morphospecies carried. We found a clear difference between pollen communities carried by hummingbirds in summer and winter, but little differentiation by age and sex. We also found that females, on average, carried more pollen morphospecies than do males, and juveniles carried more morphospecies than do adults. All Green-backed Firecrowns carried significantly higher numbers of pollen grains in summer than in winter. Our results show intriguing differences in foraging behavior and pollination roles across ages and between sexes, which support anecdotal observations that adult males are territorialists while females and juveniles forage opportunistically. Since Green-backed Firecrowns are the only hummingbird species throughout most of their range, sexually divergent foraging patterns suggest that males and females may fill the ecological roles of different species. Our results add to the field's understanding of the foraging patterns of Green-backed Firecrowns and help us better appreciate the complexity of its ecological role.

INTRODUCTION

Pollinators play vital roles in ecosystem function. Seventy-eight percent of flowering plants in temperate regions worldwide depend on animal pollinators for reproduction, while a mean of 94% of angiosperms in tropical regions do so (1). In all, at least 299,200 known plant species, including most major food crops, rely on animal pollinators (1, 2). It is, therefore, very important to have a clear understanding of the ecology of pollinators and of the threats that face them. It is especially important to understand the ecology of keystone pollinators, species upon which a disproportionately large number of plant species depend for pollination.

The Temperate Rainforests of Southern South America (TRSSA) in southern Chile and Argentina contain a relatively simple pollinator community, in which the region's only hummingbird seems to play a particularly important role. The Green-backed Firecrown (*Sephanoides sephaniodes*) is endemic to the region and is the southernmost-ranging species of hummingbird in the world, occurring in areas where there are few, if any, sympatric animal pollinators (3, 4). However, the TRSSA's diverse flora is highly dependent on animal pollination (5). While long-proboscid insects (e.g., *Bombus dahlbomii* and Nemestrinidae spp.) may contribute to the pollination of some plant species, much of the floral community appears to be entirely dependent on the Green-backed Firecrown for pollination; in fact, about 20% of the region's woody plant genera feature ornithophilous (bird-pollinated) flowers (3, 5, 6). The reliance of such a large portion of the floral community on a single species is exceptional. In most ecosystems where hummingbirds are found, the ratio of hummingbird-pollinated plant species to hummingbird species is well below 5:1 (3). For instance, in temperate rainforests in the state of Washington, U.S.A., which are located about as far from the equator as are the TRSSA, only one species of plant depends on a hummingbird pollinator (7). The diverse floral community of the TRSSA relies, in large part, upon the Green-backed Firecrown alone, and consequently, the Green-backed Firecrown is often considered a keystone species, a species disproportionately important in maintaining the diversity and structure of its community (3, 5). Therefore, it is essential that the Green-backed Firecrown's ecological role be well studied and understood.

Adult Green-backed Firecrowns are sexually dimorphic. Males are larger than females, have relatively shorter bills, and exhibit bright iridescent crown feathers (8). While males are highly territorial, females are not known to defend discrete territories (9). Despite these morphological and behavioral differences, little information exists on sex-related differences in diet or contributions to plant pollination. We therefore examined pollen samples collected from wild Green-backed Firecrowns (**Figure 1**) in the Los Ríos Region, Chile, to investigate patterns of pollination with respect to differences in floral visitations between males and females, between adults and juveniles, and between birds captured in summer versus in winter. We hypothesized that morphological and behavioral differences between the sexes would cause males and females to differ in diet and to feed from distinct floral communities. Under this hypothesis, we predicted that females would carry a greater diversity of pollen, due in part to their longer bills, which likely enable them to visit a more diverse suite of plant species, and due also to the fact that females are not known to

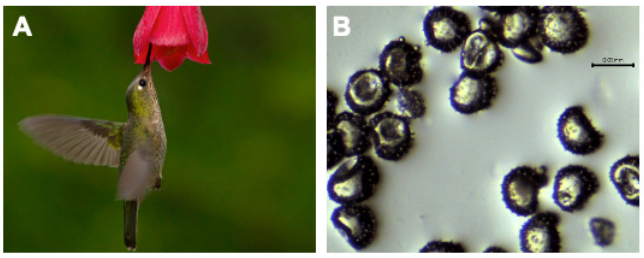


Figure 1: The Green-backed Firecrown and pollen examined. A) A female Green-backed Firecrown visiting a *Lapageria rosea* flower. Photo credit: Juan Andrés Vara Braun. B) Pollen from a *Lapageria rosea* flower viewed under the microscope. The black scale bar measures 0.05 mm.

defend discrete territories, and so are not confined to visiting the plant species present within a small area but rather are able to utilize a wider variety of flowers. Also, given that most hummingbird-pollinated plant species in the TRSSA bloom during the austral spring and summer (6), we hypothesized that hummingbirds would carry more pollen morphospecies in the summer months.

RESULTS

We collected pollen samples from 38 Green-backed Firecrowns (**Table 1**). Twenty-eight of the birds were successfully aged and sexed in the field in Chile and were therefore included in the analyses of differences in pollen richness (the number of unique morphospecies carried by an individual bird) and number of pollen grains.

During the process of pollen examination and identification, we identified 15 distinct morphospecies, based on pollen shape, size, and texture (**Table 2**). Five morphospecies corresponded directly to a plant species or group of species with pollen similar in appearance. The five known plants were *Tristerix corymbosus* (Loranthaceae), *Lapageria rosea* (Philesciaceae), *Greiga landbeckii* (Bromeliaceae), *Embothrium coccineum* (Proteaceae), and species from the Gesneriaceae family (including *Asteranthera ovata*, *Sarmienta repens*, *Mitraria coccinea*). *Lapageria rosea* and *Greiga landbeckii* were observed almost entirely in samples from the austral winter, while eight morphospecies, including *Embothrium coccineum*, were observed only during the austral summer. A non-metric multidimensional scaling (NMDS) plot showed clear differentiation in pollen community across the seasons, but little differentiation in the pollen community between juveniles and adults, or between males and females (**Figure 2**).

Birds in summer carried many more pollen grains than did birds in winter, with an average of nearly four times as many grains per bird in summer than in winter (p = 0.0067; **Figure 3**; **Table 3**). In summer, when juveniles and adults were reliably distinguishable, the two age stages carried similar numbers of pollen grains. There was no detectable difference in average grain counts between males and females. We used General Linear Models (GLMs) to further explore the drivers of grain counts and richness. The best model for explaining grain counts included season as the only predictor variable (**Table 4**). Sex, capture time, bill length, and body mass were not significant or included in the preferred model.

Differences in pollen richness between seasons and sexes were marginally significant (**Table 3**). Hummingbirds carried an average of 1.0 morphospecies more in the summer than in the winter (p = 0.0516; **Figure 4**) and females carried an average of 0.9 morphospecies more than did males (p = 0.0662). Summer juveniles carried an average of 1.4 more morphospecies than did summer adults, though this difference too was only marginally significant (p = 0.0872; **Table 3**). The best model for explaining morphospecies richness included both season and sex as significant (p < 0.05) predictor variables (**Table 4**). Capture time, bill length, and body mass were not significant or included in the preferred model.

DISCUSSION

We observed clear seasonal differences in pollen community composition, suggesting that hummingbirds feed from different flowers in summer than in winter. We also observed compelling patterns in pollen richness and grain counts related to sex, season, and age stage. However, as we analyzed variations in our data across these categories, we noted that these three factors were interrelated in our dataset – juvenile hummingbirds are only distinguishable in summer (though they are likely present year-round), and 86% of the juveniles we sampled were female. This correlation among the categories juvenile, female, and summer, makes it difficult to discern which of these variables are driving the patterns we

Hummingbird ID	Season	Age	Sex	Pollen richness	Number of pollen grains
A01501	Summer	Juvenile	U	2	738
A01508	Summer	Juvenile	F	5	1662
A01509	Summer	Unknown	M	2	14422
A01562	Summer	Adult	F	3	104
A01563	Summer	Unknown	F	2	195
A01564	Summer	Juvenile	U	7	4924
A01565	Summer	Juvenile	F	4	855
A01566	Summer	Adult	M	2	85
A01567	Summer	Juvenile	U	4	8972
A01568	Summer	Juvenile	M	3	4096
A01569	Summer	Unknown	M	1	3847
A01570	Summer	Unknown	M	0	0
A01572	Summer	Juvenile	F	7	4515
A01573	Summer	Juvenile	F	2	726
A01574	Summer	Juvenile	F	5	2611
A01575	Summer	Unknown	U	6	521
A01576	Summer	Adult	M	2	1000
A01577	Summer	Juvenile	U	2	4255
A01579	Summer	Juvenile	U	1	736
A01580	Summer	Adult	M	3	3084
A01620	Summer	Adult	M	2	145
A01631	Summer	Juvenile	U	5	784
A01634	Summer	Adult	M	3	3230
A01635	Summer	Adult	F	3	1821
A01636	Summer	Juvenile	F	2	1521
A01697	Winter	Adult	M	0	0
A01733	Winter	Adult	F	3	35
A01734	Winter	Adult	F	1	10
A01735	Winter	Adult	F	2	173
A01737	Winter	Adult	M	2	192
A01742	Winter	Adult	M	2	116
A01744	Winter	Adult	M	2	29
A01750	Winter	Adult	M	4	2959
A01766	Winter	Adult	F	2	1176
A01768	Winter	Adult	F	3	765
A01781	Winter	Adult	F	4	330
A01782	Winter	Adult	F	3	451
A01783	Winter	Adult	M	2	384

Table 1. Metadata for all 38 Green-backed Firecrown hummingbirds (Sephanoides sephaniodes) captured (M = Male, F = Female, U = Unknown).

Morphospecies	Summer incidence	Winter incidence	Plants included in morphospecies (if known)
A	0.73	0.29	Gesneriaceae (including *Asteranthera ovata*, *Sarmienta repens*, *Mitraria coccinea*)
N	0.50	0.00	
D	0.46	0.57	*Tristerix corymbosus*
G	0.38	0.14	
M	0.23	0.00	
H	0.19	0.00	
Q	0.19	0.00	
K	0.12	0.00	
O	0.12	0.00	
B	0.04	0.57	*Lapageria rosea*
F	0.04	0.07	
I	0.04	0.00	
J	0.04	0.00	*Embothrium coccineum*
C	0.00	0.64	*Greiga landbeckii*
L	0.00	0.00	

Table 2. Pollen morphospecies collected from Green-backed Firecrown hummingbirds (Sephanoides sephaniodes) in the Los Ríos Region, Chile in summer (Jan-Feb) and winter (July-Aug) 2018, ordered by summer incidence. Incidence is the proportion of hummingbirds carrying ≥5 grains of that morphospecies in summer or winter. Morphospecies were assigned a letter ID (A through O) in the order they were described. Corresponding plant species (if known) are reported to the highest taxonomic unit possible.

see in pollen richness and grain counts. For each comparison, we have attempted to identify the underlying drivers of our observed patterns by critically examining our data, drawing on other studies, and considering the life history of our study species. In addition, we use GLMs to attempt to further disentangle the effects of season and sex.

In the case of pollen grains, we noted that grain counts may have been affected both by the number of flowers a hummingbird visits and by the pollen characteristics of the flowers visited (e.g., number of pollen grains produced per flower and tendency of the pollen grains to adhere to the hummingbird). Given the clear differences in the composition of pollen carried in summer versus in winter (**Figure 2**), the most likely explanation for the greater number of pollen grains in summer is that these hummingbirds were visiting different plants in summer and that at least some of these plant species produced sticky or abundant pollen that increased summer pollen counts. Age and sex did not appear to be important drivers of pollen grain counts.

Morphospecies richness was greater in summer than in winter, among juveniles than among adults, and among females than among males (though each of these differences was marginally significant; see **Table 3**). The best supported GLM explaining pollen richness (year-round) included both season and sex as significant predictor variables, providing further evidence indicating that these both are important factors affecting morphospecies richness. All of these patterns seem plausible given our knowledge regarding the foraging behavior of Green-backed Firecrowns. Most hummingbird-pollinated flower species in our study area bloom in spring or summer (6), and so our data confirm our expectation that hummingbirds would carry more pollen morphospecies in the summer. Sex- and age-related differences in pollen richness likely reflect morphological and behavioral differences between juveniles, adult males, and adult females.

Our results suggest that juvenile Green-backed Firecrowns carry more diverse arrays of pollen than do adults. Juveniles are likely subordinate to adults, especially to territorial males, and might therefore be denied access to dense flower patches, which would likely force them to forage from a more dispersed and diverse suite of flowers, and therefore carry a more varied array of pollen. Alternatively, it is possible that juveniles have simply not yet developed preferences for certain flower groups, and so forage from a larger array of plants. Adults display a distinct preference for sucrose-rich solutions, but discrimination between sugar solutions has not been documented in juveniles (10); this suggests that juveniles may be less discriminating in their choice of floral food sources. Additionally, most juveniles sampled were female, and this also may contribute to the perceived disparity in the number of morphospecies characters between juveniles and adults.

Male Green-backed Firecrowns have, on average, bills that are significantly shorter than those of females (4, 8, 11). The average mass of male Green-backed Firecrowns in our study region was 6.2 g, in the center of the range for hummingbirds (8, 11, 12). These qualities—short bills and medium mass—align with those of territorialist hummingbirds (13, 14), and there is some behavioral evidence to suggest that males are also territorial also (11, 15). We found that males generally carried fewer pollen morphospecies than did females, and this too supports the idea that male Green-backed Firecrowns are territorial. It seems likely that territorialist hummingbirds would be more limited in their array of pollen than would generalists or hummingbirds with other foraging strategies, as territorialism allows hummingbirds to monopolize a narrow section of the available resources, limiting the flowers available—both to themselves and to other territorial hummingbirds—and eliminating the need to access a more diverse, less concentrated array of flora. Our finding that male firecrowns carry less diverse arrays of pollen provides additional evidence indicating that male firecrowns are territorial.

In contrast, we found that females carry, on average, more morphospecies of pollen, and other studies have found that females have lower body mass and longer bills (4, 8). A number of authors have suggested that the possession of longer bills by females may be favored as a result of intersexual food competition, as male hummingbirds of many

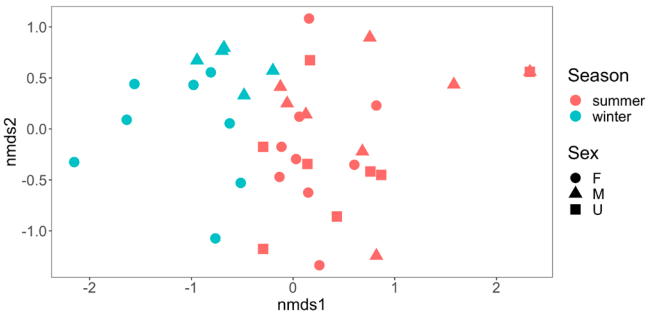


Figure 2. NMDS (non-metric multidimensional scaling) plot of community composition of pollen morphospecies collected from 38 Green-backed Firecrown hummingbirds. Each point represents the community composition of pollen collected from a single hummingbird. Points that are closer together indicate hummingbirds carrying more similar communities of pollen. Point shapes and colors indicate the sex of the hummingbird and the season in which it was captured. NMDS stress = 0.1383.

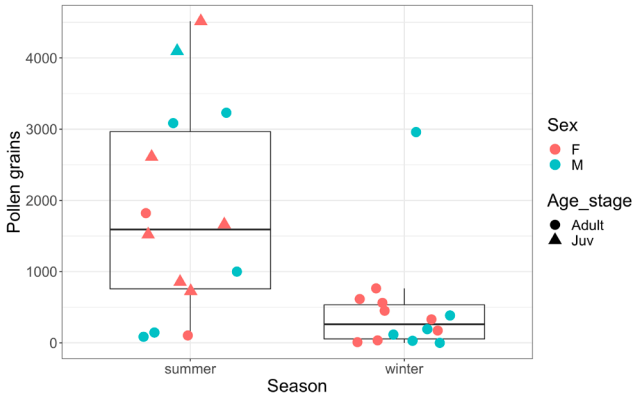


Figure 3. Number of pollen grains carried by 28 Green-backed Firecrown hummingbirds. Number of grains was greater in summer than winter (p = 0.0067), but similar among juveniles and adults (p = 0.2524) and among males and females (p = 0.6622). Points are jittered horizontally to facilitate interpretation.†).

	Response variable	Group	Mean ± SE (n)	p-value
Summer	Pollen richness	Adult Juvenile	2.6 ± 0.7 (7) 4.0 ± 0.2 (7)	0.0872†
	Number of grains	Adult Juvenile	2284 ± 573 (7) 1353 ± 523 (7)	0.2524
Year-round	Pollen richness	Summer Winter	3.3 ± 0.4 (14) 2.3 ± 0.3 (14)	0.0516
		Male Female	2.3 ± 0.3 (12) 3.2 ± 0.4 (16)	0.0662
	Number of grains	Summer Winter	1818 ± 395 (14) 473 ± 202 (14)	0.0067†
		Male Female	1277 ± 453 (12) 1047 ± 295 (16)	0.6622

Table 3. Summary statistics of between group comparisons of pollen richness and number of pollen grains collected from 28 Green-backed Firecrowns. Differences between groups were compared using two-sample t-tests or Welch's t-tests (indicated by †).

species defend flower patches and force females to forage more opportunistically among more scattered resources, occasionally filching from the territories of males, using a strategy referred to by some authors as facultative traplining (16, 17). A long bill is helpful when foraging in this manner, as it tends to allow visits to flowers to be briefer and more efficient. It appears that the Green-backed Firecrown exemplifies this phenomenon. While sexual dimorphism in bill length and body mass is fairly common among hummingbirds, especially in distinctly sexually dichromatic species such as the Green-backed Firecrown (12, 16), the way in which this dimorphism relates to foraging behavior and ecological niche has been documented thoroughly in relatively few species of hummingbird (18).

Our study contributes to a growing number of studies documenting sexual differences in foraging behavior and morphology in Green-backed Firecrowns (8, 9, 11, 19). Our results regarding sex- and age-related foraging patterns are similar to the range of values reported by Fraga *et al.* (11), the only other study of Green-backed Firecrowns that differentiated among males, females, and juveniles (although their study only reported approximate values). As the only hummingbirds throughout nearly all of their range, Green-backed Firecrowns seem to have adapted to occupy multiple pollination niches by developing different strategies for males and females—males seem to be territorialists, and females seem to be facultative trapliners. Interestingly, simple hummingbird communities on islands often consist of two species occupying two niches: facultative trapliners and territorialists (13). Here, male and female Green-backed Firecrowns evidently take on these two roles; in doing so, they seem to almost occupy the roles of two different species. Surprisingly however, despite differences in pollen richness, bill morphology, and apparent foraging behavior, we did not observe clear differences in pollen community composition between males and females as we had hypothesized.

The few other hummingbird species with well-documented sexual dimorphism in foraging strategy (e.g., Purple-throated Carib—see Temeles, *et al.* (18)) co-occur with other hummingbird species. However, throughout nearly all of its range, the Green-backed Firecrown is the only hummingbird present, at least in its habitat. The effects of interspecific competition on the foraging roles of males and of females would therefore be interesting to study more closely. For example, might pressure from interspecific competition result in a change in the foraging strategies of the sexes? Only on the Juan Fernandez Islands does the Green-backed Firecrown co-occur with an ecologically similar bird, its sister species, the critically endangered Juan Fernandez Firecrown *Sephanoides fernandensis* (12, 20). There, male Juan Fernandez Firecrowns appear to be the dominant territorialists, forcing female conspecifics and Green-backed Firecrowns of both sexes into more opportunistic roles (15, 20). It appears that interspecific competition does indeed affect the foraging strategy of males, forcing them into a role more similar to the role of female Green-backed Firecrowns elsewhere. Whether this pressure will result in any morphological changes remains to be seen.

The situation of the Green-backed Firecrown and the Juan Fernandez Firecrown provides an interesting case study that could shed light on niche partitioning, territoriality, and foraging ecology in other hummingbirds. However, the situation between the firecrowns is of special interest as the Juan Fernandez Firecrown is in imminent danger of becoming extinct (20, 21). Should this unfortunate event occur, resulting

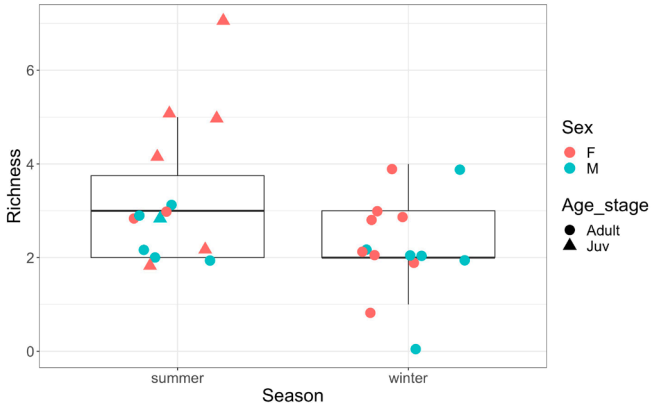


Figure 4. Richness (number) of pollen morphospecies carried by 28 Green-backed Firecrown hummingbirds. Richness was marginally greater in summer than winter (p = 0.0516), in juveniles than adults (p = 0.0872), and in females than males (p = 0.0662). Points are jittered vertically and horizontally to facilitate interpretation.

Response variable	Season	Sex	AIC_c	ΔAIC_c	AIC_c weight	Multiple R^2
Richness	1.2 ± 0.5 t = 2.404 p = 0.025	-1.1 ± 0.5 t = -2.277 p = 0.033	88.38	0	0.79	0.33
Pollen grains	1321 ± 441 t = 2.996 p = 0.006	–	425.91	0	1	0.28

Table 4. Explanatory models (parameter estimates and standard errors) of morphospecies richness and number of pollen grains derived from General Linear Models and selected by Akaike Information Criterion for small samples (AICc) with season, sex, capture time (minutes after sunrise), bill length (mm), body mass (g) evaluated as predictive variables. Only highly supported models (ΔAICc < 2) are shown. See text for details.

changes in foraging patterns of the Green-backed Firecrown would be very interesting to observe.

Our study has identified a number of interesting and ecologically relevant patterns in the foraging behavior of Green-backed Firecrowns. However, due to our relatively small sample size, we are unable to fully document, delineate, and understand many of these trends. In particular, future studies should examine 1) the surprising similarity of pollen communities carried by male and female Green-backed Firecrowns, despite their differences in bill morphology and foraging behavior; 2) the ways in which sexual dimorphism corresponds to sexual differences in ecological roles (especially as this may relate to competition with the endangered Juan Fernandez Firecrown); and 3) the potential similarity in foraging roles between juveniles (of both sexes) and adult female Green-backed Firecrowns. As our knowledge of important pollinators like the Green-backed Firecrown continues to develop, we will be better able to understand the roles they play in the maintenance of the plant community and the ways in which these roles affect ecosystem function.

MATERIALS AND METHODS

Study Area

We captured hummingbirds during the austral summer (January-February) and winter (July-August) of 2018 in native forests and exotic timber plantations in and adjacent to the Llancahue Forest Reserve in Los Ríos Region, Chile (39°50' S, 73°08' W, 1,300 ha). This reserve is located in the Intermediate Depression from 50 to 360 m a.s.l. and contains ~700 ha of old-growth Valdivian Temperate Rainforest and an additional ~400 ha of secondary native forest composed primarily of mixed evergreen broadleaf and *Nothofagus dombeyi*-dominated forest types (22). Annual rainfall averages 2,100 mm, occurring mostly from April–October (22). Temperatures are moderate, with maximum and minimum temperatures at the weather station in Valdivia (25 m a.s.l.) averaging 22.9 and 10.4 ˚C in January and 11.2 and 4.5 ˚C in July. Surrounding the reserve is a heterogeneous mosaic of exotic timber plantations (primarily *Pinus radiata* and *Eucalyptus* spp.), pastureland, and remnant patches of native forest.

Field Methods

Hummingbirds were captured with mist nets (9 x 2.5 m nets, 4–11 nets per day) in the understory of old-growth and secondary native forests, and both *Pinus radiata* and *Eucalyptus* spp. plantations. Nets were opened one hour after sunrise, operated for 5 hours, and checked every 15 minutes. Captures only took place on mornings without precipitation or high winds. Hummingbirds were removed from the nets and a numbered aluminum band was placed on the right tarsus. Once banded, we measured the following morphometric variables: mass (±0.1 g), wing chord (±0.5 mm), bill length (exposed culmen; ±0.01 mm), and tail length (±0.5 mm). Pollen samples were obtained by gently dabbing the bird's bill, head, and neck with a small piece of clear Scotch tape, which was subsequently adhered to a microscope slide for later examination. All birds were provided 20% sucrose nectar (v/v in water) from a glass dropper prior to release.

Pollen Identification

To count and identify the pollen grains sampled from each of the Green-backed Firecrowns, all pollen sample slides were carefully examined under a microscope (Micromaster, Fisher Scientific) at 80x magnification, and each grain of pollen assigned to a morphospecies based on pollen size, shape, and texture (**Figure 1**). Pollen grains for which the morphospecifies identity was unclear were examined more closely under 200x magnification. Tallies of individual grains belonging to each morphospecies were kept using manual clicker counters. Wherever possible, morphospecies were identified to plant species by comparison with standard pollen samples taken directly from flowers of known plants. To avoid potential bias due to sample contamination, morphospecies with < 5 pollen grains in a given sample were excluded from subsequent analyses. All pollen examination was performed at Stanford University.

Statistical Analysis

We evaluated seasonal, sex, and age-related differences in pollen community composition using non-metric multidimensional scaling (NMDS) performed in the R package 'vegan' (23). NMDS is a statistical ordination technique that takes information from many different dimensions (in this case, each dimension represented one pollen morphospecies) and collapses that information into a two-dimensional plot so that it can be visualized and interpreted. Each data point represents the community composition of pollen collected from a single hummingbird. Points that are close together indicate similar community composition, while points that are far apart indicate more distinct communities. Data points were then colored by category (e.g., summer vs. winter) to examine the drivers of community composition.

We used two-sample t-tests to statistically evaluate seasonal, sex, and age-related differences in pollen richness, calculated by counting the number of morphospecies in each sample and the number of pollen grains carried by each hummingbird. Specifically, we compared pollen richness and number of grains in summer vs. winter, male vs. female, and adult vs. juvenile hummingbirds. For these analyses, we excluded samples for which the hummingbird's age and sex were unknown, resulting in a sample size of 28 birds. Comparisons of adult and juvenile hummingbirds were made only for summer samples, since adults and juveniles in winter are indistinguishable. Normality was assessed using probability plots. We used Welch's t-tests for two comparisons in which groups had unequal variance (adult vs. juvenile richness, and summer vs. winter number of grains).

We used GLMs to further explore the factors that explain variation in pollen richness and grain counts. We included season, sex, capture time (minutes after sunrise), bill length, and body mass as predictor variables and used Akaike Information Criterion for small samples (AICc) to select the best models. We did not include age as a predictor variable because juveniles could only be distinguished during the summer season. We calculated AICc weights to infer the relative support of each model, as well as the multiple R2 value, which describes the proportion of variance explained by the predictor variables. T-statistics and p-values reported from the GLMs vary slightly from those reported from the two-sample t-tests because we only included hummingbirds (n = 25) in the GLMs for which we had information on all five predictor variables used. All statistical analyses were performed in RStudio (Version 1.1.383, ©2009-2016 RStudio, Inc.).

ACKNOWLEDGEMENTS

We thank Professor Rodolfo Dirzo for his invaluable support throughout the project. We thank Gabriela Biscarra, Jorge Ruiz, Enzo Basso, Claudio Navarrete, Valeria Araya, Johannes Horstmann, Ahzha McFadden, and Yerko Rivera Mendoza for assistance in the field and with logistics. Banding was conducted with permission from the Chilean Servicio Agrícola y Ganadero (Resolution 688/2018). Access to capture sites was provided by the Universidad Austral de Chile (FONDECYT 1150496) and Arauco SA. Funding for field work (to TNM) was provided by a Stanford Graduate Fellowship, National Science Foundation GFRP and GROW awards, and the Stanford Center for Latin American Studies. We are also very grateful to my parents, Drs. Deborah and Thomas Freeland, for their encouragement and support.

Received: November 25, 2019
Accepted: April 7, 2020
Published: May 28, 2020

REFERENCES

1. Ollerton, Jeff, *et al*. "How Many Flowering Plants Are Pollinated by Animals?" *Oikos*, vol. 120, no. 3, 2011, pp. 321–26.
2. Klein, Alexandra-Maria, *et al*. "Importance of Pollinators in Changing Landscapes for World Crops". *Proc. R. Soc. B* vol. 274, 2007, pp. 303-313.
3. Armesto, Juan J., *et al*. "The Importance of Plant-Bird Mutualisms in the Temperate Rainforest of Southern South America." *High-Latitude Rainforests and Associated Ecosystems of the West Coast of the Americas*, edited by R.G. Lawford, Springer, New York, NY, 1996, pp. 248–65.
4. Rozzi, Ricardo, and Jaime E. Jiménez. *Magellanic Sub-Antarctic Ornithology : First Decade of Long-Term Bird Studies at the Omora Ethnobotanical Park, Cape Horn Biosphere Reserve, Chile*. Edited by Ricardo Rozzi and Jaime E. Jiménez, University of North Texas Press, 2013.
5. Aizen, Marcelo A., and Cecilia Ezcurra. "High Incidence of Plant-Animal Mutualisms in the Woody Flora of the Temperate Forest of Southern South America: Biogeographical Origin and Present Ecological Significance." *Ecología Austral*, vol. 8, 1998, pp. 217–36.
6. Smith-Ramirez, Cecilia. "Hummingbirds and their floral resources in temperate forests of Chiloé Island, Chile." *Revista Chilena de Historia Natural*, vol. 66, 1993, pp. 65–13.
7. Franklin, Jerry F., *et al*. "The Forest Communities of Mount Rainier National Park." *National Park Service Scientific Monograph Series*, vol. 19, 1988, pp. 1–222.
8. McFadden, Tyler N., et al. "Recommended Band Sizes and a Novel Technique for Sexing Immature Green-Backed Firecrown Hummingbirds (*Sephanoides sephaniodes*)". *Ornitología Neotropical,* no. 30, 2019, pp. 179–84.
9. González-Gómez, Paulina L., and Cristian F. Estades. "Is Natural Selection Promoting Sexual Dimorphism in the Green-Backed Firecrown Hummingbird (*Sephanoides sephaniodes*)?" *Journal of Ornithology*, vol. 150, 2009, pp. 351–56.
10. Chalcoff, Vanina R., *et al*. "Sugar Preferences of the Green-Backed Firecrown Hummingbird (Sephanoides Sephaniodes): A Field Experiment." *The Auk*, vol. 125, no. 1, pp. 60–66.
11. Fraga, Rosendo M., *et al*. "Interactions between the Firecrown Hummingbird *Sephanoides sephaniodes* and plants of the Nothofagus forests at Nahuel Huapi National Park, Argentina" *Hornero*, vol. 014, no. 04, 1997, pp. 224–34.
12. Fogden, Michael, *et al*. *Hummingbirds: A Life-Size Guide to Every Species*. Harper Design, 2014.
13. Feinsinger, Peter, and Robert K. Colwell. "Community Organization among Neotropical Nectar-Feeding Birds." *American Zoologist*, vol. 18, no. 4, 1978, pp. 779-795.
14. Tinoco, Boris A., *et al*. "Effects of Hummingbird Morphology on Specialization in Pollination Networks Vary with Resource Availability." *Oikos*, vol. 126, no. 1, 2017, pp. 52–60.
15. Wolf, Coral, and Erin Hagen. "Aggressive Interactions of Firecrowns (*Sephanoides* spp.; Trochilidae) during the Breeding Season of Robinson Crusoe Island, Chile." *Ornitología Neotropical*, vol. 23, 2012, pp. 545-553.
16. Bleiweiss, Robert. "Joint Effects of Feeding and Breeding Behaviour on Trophic Dimorphism in Hummingbirds." *Proceedings of the Royal Society B: Biological Sciences*, vol. 266, 1999, pp. 2491–97.
17. Temeles, Ethan J., and W. Mark Roberts. "Effect of Sexual Dimorphism in Bill Length on Foraging Behavior: An Experimental Analysis of Hummingbirds." *Oecologia*, vol. 94, 1993, pp. 87–94.
18. Temeles, Ethan J., Irvin L. Pan, *et al*. "Evidence for Ecological Causation of Sexual Dimorphism In a Hummingbird." *Science*, vol. 289, no. 5478, 2000, pp. 441–43.
19. González-Gómez, Paulina L., *et al*. "Aggression, Body Condition, and Seasonal Changes in Sex-Steroids in Four Hummingbird Species." *Journal of Ornithology*, vol. 155, no. 4, 2014, pp. 1017–25.
20. Roy, Michael S., *et al*. "Evolution and History of Hummingbirds (Aves: Trochilidae) from the Juan Fernandez Islands, Chile." *Ibis*, vol. 140, no. 2, 2008, pp. 265–73.
21. BirdLife International. "*Sephanoides fernandensis*." *The IUCN Red List of Threatened Species*, 2018.
22. Donoso, Pablo J., *et al*. "Balancing Water Supply and Old-Growth Forest Conservation in the Lowlands of South-Central Chile through Adaptive Co-Management." *Landscape Ecol*, vol. 29, 2014, pp. 245–60.
23. Oksanen, Jari, *et al*. *"vegan" Community Ecology Package.*

R Package Version 2.5-6. 2019.

JOURNAL OF
EMERGING INVESTIGATORS

Article

Ground-based Follow-up Observations of TESS Exoplanet Candidates

Sarah Tang[1] and William Waalkes[2]
[1]Fairview High School, Boulder, Colorado
[2]University of Colorado, Boulder, Colorado

SUMMARY

The goal of this study was to further confirm, characterize, and classify LHS 3844 b, an exoplanet detected by the Transiting Exoplanet Survey Satellite (TESS). Additionally, we strove to determine the likeliness of LHS 3844 b and similar planets as qualified candidates for observation with the James Webb Space Telescope (JWST). We accomplished these objectives by analyzing the stellar light curve, theoretical emission spectroscopy metric (ESM), and theoretical Planck spectrum of LHS 3844 b. We remotely obtained pre-reduced ground-based data of LHS 3844 b from the El Sauce Observatory. We hypothesized that LHS 3844 b and similar target TESS candidates are qualified for future JWST follow-up. Through AstroImageJ and the Python programming language, we converted the calibrated data images into normalized and fitted flux light curves. Through our best-fit light curve model, we classified LHS 3844 b as a terrestrial planet. The calculated ESM of LHS 3844 b surpassed the projected threshold for simulated planets deemed qualified for JWST spectroscopic follow-up, and the Planck spectrum of LHS 3844 b revealed that the observed wavelengths between 6,000 and 10,000 nanometers would produce the highest signal-to-noise spectroscopic observations of LHS 3844 b and like planets. These findings will improve the accuracy and efficiency of spectroscopic follow-ups performed by the JWST, and we intend to apply these methods to study a variety of exoplanets.

INTRODUCTION

The search for similar solar systems and life beyond Earth's solar system has fascinated human beings for centuries (1). The study of exoplanets, planets that orbit a star other than the Sun, provides us with crucial information regarding planetary population characteristics and the differences between Earth's solar system and other solar systems (2). Exoplanets orbit their host stars for measurable periods, and can therefore be detected through predicted planetary transits. Planetary transits occur when a planet crosses the disk of its host star and causes a decrease in the observed amount of light emitted from the host star. A transiting exoplanet can be detected by observing the light flux emitted from a star over a given period of time, typically hours (3).

The recent development of new technologies and equipment has allowed for further exploration and discovery of exoplanets using the transit method. A clear example of this advancement is the Transiting Exoplanet Survey Satellite (TESS), an Explorer mission launched by the National Aeronautics and Space Administration (NASA) in April of 2018 with the goal of detecting exoplanets specifically using the transit method (4). TESS utilizes the Science Processing Operations Center (SPOC) pipeline to convert raw data images into calibrated pixels and transit light curve graphs (5). In the span of two years, TESS will monitor at least 200,000 stars for intervals lasting from a month to a year for signs of planetary transits (4). The goal of TESS is to detect planets that orbit M dwarf stars, as these stars are geometrically more likely to host Earth-like and potentially habitable planets because of their relatively small sizes and low effective temperatures (6). Additional follow-up analyses are performed on TESS exoplanet candidates to confirm the existence of the planets, rule out false positives, and establish precise transit parameters. Scientists around the world complete follow-up analyses through observatories such as the Las Cumbres Observatory Global Telescope Network (7).

A target TESS candidate is an Earth-like planet orbiting an M dwarf star. TESS is targeting these planets because there is a greater chance they orbit within the habitable zone of their host star (4). LHS 3844 b orbits LHS 3844, an M dwarf star with 0.151 $M_\odot$ and 0.189 $R_\odot$, and is therefore considered a target TESS candidate. A relatively large fraction of planets detected by TESS will likely hold similar properties as LHS 3844 b. A projected thermal emission spectroscopic signal-to-noise or ESM value of 7.5 is required for model simulated planets to qualify for James Webb Space Telescope (JWST) follow-up (8). JWST spectroscopic analysis concerns the emission spectrum of a planetary system, and consequently, the JWST will observe secondary transits, which is the passage of a planet behind the disk of its host star. During a secondary transit, there is a slight decrease in the detected amount of emitted light from the system due to the disappearance of the planet's emitted light, which allows for the precise analysis of the planet's individual emission (9).

A Planck spectrum represents the spectral density of electromagnetic radiation emitted from an object in thermal equilibrium. The theoretical Planck spectra of a planet and its host star can be used to determine the wavelength range in which to observe the highest signal-to-noise of the

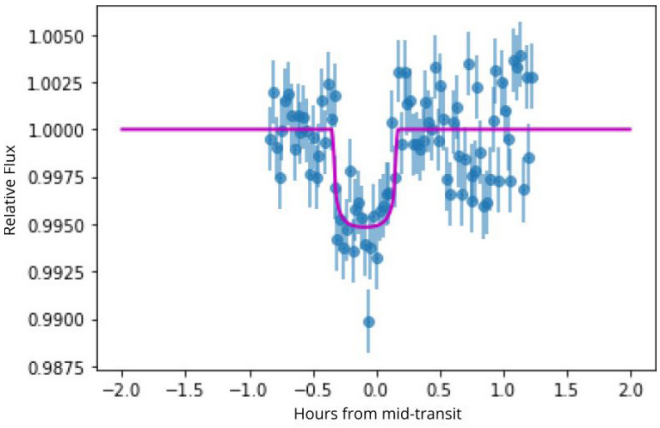


Figure 1. Normalized and modeled transit light curve graph of LHS 3844 b. Error bars represent the 1σ error bar ranges of the host star's flux values, as generated by AstroImageJ.

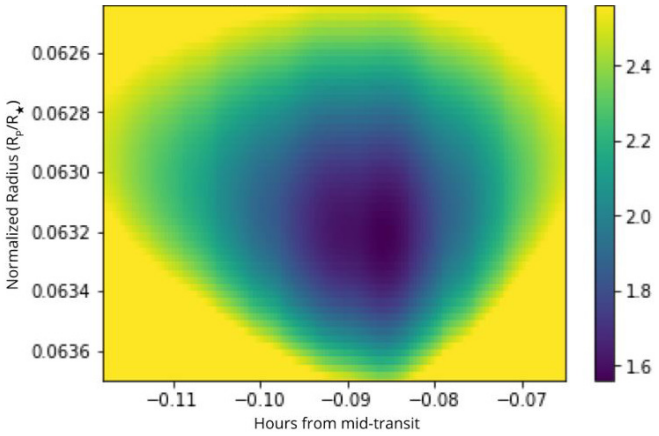


Figure 2. Chi-squared map used to determine the 1σ error bar range (68% Confidence Intervals) for all planetary parameter calculations.

planet's emission. It is crucial to identify the exoplanets that would produce the highest signal-to-noise spectroscopic observations in order to improve the methodology of these spectroscopic follow-ups and account for the substantial amount of exoplanet candidates to be observed during the JWST timeline.

The objective of this study was to analyze pre-reduced ground-based follow-up data on LHS 3844 b to further confirm, characterize, and classify the planet. We will also study the preferred spectroscopic observation conditions of target TESS candidates. We hypothesized that LHS 3844 b is qualified for future spectroscopic observations with the JWST.

RESULTS

Light Curve Analysis & Classification of LHS 3844 b

To analyze the radius and time of mid-transit of LHS 3844 b, we created the normalized and fitted transit light curve graph for the planet (**Figure 1**). We found a reduced chi-squared value of 1.5 for the data set of LHS 3844 b. The chi-squared map for these data of LHS 3844 b (**Figure 2**) was used to determine the 1σ error bars (68% confidence intervals) for all of the planetary calculations.

We calculated the transit depth (RP/R*) of LHS 3844 b to be 0.0632 ± 0.0008, which is consistent with the previously published literature value of 0.0640 ± 0.0007. However, the detected mid-transit time for LHS 3844 b (-0.087 ± 0.031 hours from the expected time) differed from the literature with statistical significance (10).

Along with the radius, the mass of an exoplanet is essential in characterizing and classifying the planet. The mass of a planet is typically determined by analyzing the amount of gravitational force the orbiting exoplanet exerts on its host star through the radial velocity method (11). Since radial velocity measurements have not been taken for LHS 3844 b, we used an empirical relationship between the radius in Earth radii (RΘ) and mass in Earth mass (M_Θ) of archived exoplanets to roughly project the mass of LHS 3844 b as 4.0 ± 1.8 M_Θ (**Figure 3**) (12, 13).

With our calculated radius and projected mass values, we used standards from the Planet Habitability Laboratory to classify LHS 3844 b as a terrestrial planet orbiting an M dwarf star (14). The planetary parameters of LHS 3844 b are summarized to compare them to the literature values (**Table 1**).

Expected Spectroscopic Quality of LHS 3844 b

To determine the validity of LHS 3844 b for spectroscopic observation with the JWST, we calculated the planet's ESM. A higher ESM value represents a larger S/N, and therefore, a more qualified planetary candidate for JWST follow-up. The ESM value for LHS 3844 b was calculated to be 43, which greatly surpasses the projected ESM threshold of 7.5 for model simulated planets that were deemed qualified for JWST follow-up (15).

Planck Spectrum of LHS 3844 b

To determine the wavelength region that would provide the most accurate spectroscopic observations of LHS 3844 b, we plotted the theoretical Planck spectra of LHS 3844 b and its host star (16). The theoretical Planck spectra of LHS 3844 b and its host star approach each other and exhibit similar slopes between the wavelengths of 6,000 and 10,000 nm (**Figure 4**).

DISCUSSION

The planetary parameters of LHS 3844 b classify it as a rocky planet orbiting an M dwarf star. The detected mid-transit time for LHS 3844 b (-0.087 ± 0.031 hours from the expected time) differed from the literature with statistical significance, indicating that the period of this planet can be constrained further with future analysis. This variance in mid-transit time may be the result of the gravitational pulls of an additional planetary body orbiting LHS 3844 or a large uncertainty on

Parameter	Value	Source
Orbital Period (days)	0.4629279 ± 0.0000016	Vanderspek et al., 2018
Inclination (degrees)	88.22 ± 0.30	Vanderspek et al., 2018
Equilibrium Temperature (K)	805 ± 20	Vanderspek et al., 2018
RP/R*	0.0632 ± 0.0008	This work
Semi-major Axis (AU)	0.0062 ± 0.0002	Vanderspek et al., 2018
Radius (R_Θ)	1.30 ± 0.06	This work
Projected Mass (M_Θ)	4.0 ± 1.8	Roques, 1995

Table 1: Parameters of LHS 3844 b.

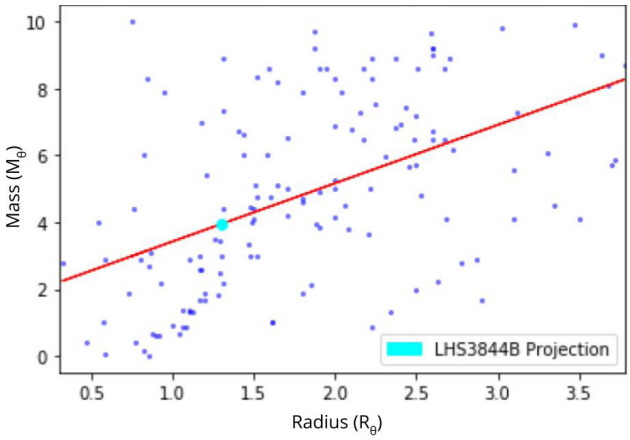


Figure 3. Empirical relationship between the mass in Earth mass (M_Θ) and radius in Earth radii (R_Θ) of archived exoplanets. This relationship was used to project the mass of LHS 3844 b. The r2 value for this data set is 0.25, signifying that the best-fit linear model is 25% matched to the data.

the cited period (10). Additionally, since only one transit was observed, this variance in mid-transit time may be caused by stochastic M dwarf star pulsations, although this cannot be concluded until more transits are observed (17). The normalized radius of LHS 3844 b (0.0632 ± 0.0008) was calculated to be consistent with the literature value.

Its theoretical ESM value and Planck spectrum support our hypothesis that LHS 3844 b and similar target TESS planets will likely be considered qualified candidates for future JWST spectroscopic follow-up. Furthermore, the ESM and Planck spectrum suggest that the JWST should observe target TESS systems in the mid-infrared wavelengths between 6,000 and 10,000 nm to obtain the most accurate information regarding the emissions and atmospheres of these exoplanets. The general wavelength region between 6,000 and 10,000 nm provides the greatest S/N of the planet's emission and therefore indicates the observed wavelengths that would produce the most detectable thermal emission spectra of LHS 3844 b. In this wavelength region, the light emitted from the host star does not overpower the emission from the planet, making it easier to accurately isolate and study the thermal emission of LHS 3844 b. These factors propose that the most accurate spectroscopic observations of LHS 3844 b will be obtained by observing the planetary system's emissions between the wavelengths of 6,000 and 10,000 nm. The JWST will conduct spectroscopic observations in the wavelengths between 600 and 25,000 nm; constraining the wavelength region in which to analyze this target TESS exoplanet class will greatly improve the efficiency and reduce the cost of the JWST data analysis process.

This information will aid in the process of proposing methodology to study atmospheric biosignatures and the molecular compositions of Terran planets in general. LHS 3844 b was the first planet orbiting a Red Dwarf star detected by TESS that was released to the public; more analysis of other detections must be completed before drawing conclusions about the general population of TESS objects and their viabilities for further analysis with the JWST. This classification indicates that LHS 3844 b is a target TESS candidate. Consequently, a relatively large fraction of TESS-detected exoplanets will likely be similar in size and period to LHS 3844 b and will produce similar spectroscopic results. The methods used in this study can be applied to other exoplanetary systems to gain a better understanding of the measures that can be taken to improve the JWST data analysis methodology for a larger variety of TESS exoplanet candidates. In addition, radial velocity observations of LHS 3844 may be taken to constrain the measurements of LHS 3844 b's mass and orbital period. The limitations of this study lie in the lack of repeated observations regarding the transit of LHS 3844 b. This lack of available data may result in large uncertainties in our measured values.

With the recent launch date delays and prolonged testing state, the JWST is now at an anticipated total cost of $8.835B, which exceeds the $8B maximum that Congress set for the mission in 2011 (18). The JWST has allocated $837M towards five years of operations and two additional years of data analysis (19). Knowledge of the precise wavelength regions to study will increase the efficiency of the data analysis process, and may aid in reducing the cost of the JWST.

METHODS

TESS utilizes the SPOC pipeline to convert raw data images into calibrated pixels and transit light curve graphs. The stages of the SPOC pipeline consist of data acquisition, calibration, photometry, presearch data conditioning, transiting planet search, and data validation (5). In this study, we developed a data processing pipeline modeled off of the SPOC pipeline with the stages of data acquisition and calibration, photometric analysis, light curve modeling, and light curve analysis.

Data Acquisition & Calibration

We analyzed pre-reduced ground-based data of the

TESS exoplanet candidate, LHS 3844 b; these data were remotely obtained through the Las Cumbres Observatory Global Telescope Network. This full transit of LHS 3844 b was observed through the IC band on UT 2018 September 06 by the El Sauce Observatory Planewave Corrected Dall-Kirkham (CDK) 0.36-meter telescope located in El Sauce, Chile (20).

Photometric Analysis

We used the photometry software AstroImageJ to stack the data images of LHS 3844 b's star field and perform aperture photometry, light measurement within a fixed size, on the target star (LHS 3844) and neighboring comparison stars (21). The flux measurements of the target star were normalized in regards to the flux of the comparison stars to rid the data of atmospheric disturbance errors. These normalized flux measurements of the target star were then converted into numerical values of relative flux:

$$relfluxT = \frac{F_T}{\sum_{i=1}^{n} F_{Ci}}$$

where FT is the differential light flux count of the target star, n is the number of comparison stars, i is the measurement of the current comparison star, and FCi is the net integrated counts of the current comparison star.

The measures of relative flux for each image were stored in text files along with the corresponding barycentric Julian date and 1σ error bar value:

$$\sigma_{relfluxT} = \frac{F_T}{F_E}\sqrt{\frac{N_T^2}{F_T^2} + \frac{N_E^2}{F_E^2}}$$

where FT is the net integrated counts of the target star, FE is the sum of net integrated counts of all comparison stars, NT is the noise in the target star aperture, and NE is the ensemble noise in all comparison star apertures.

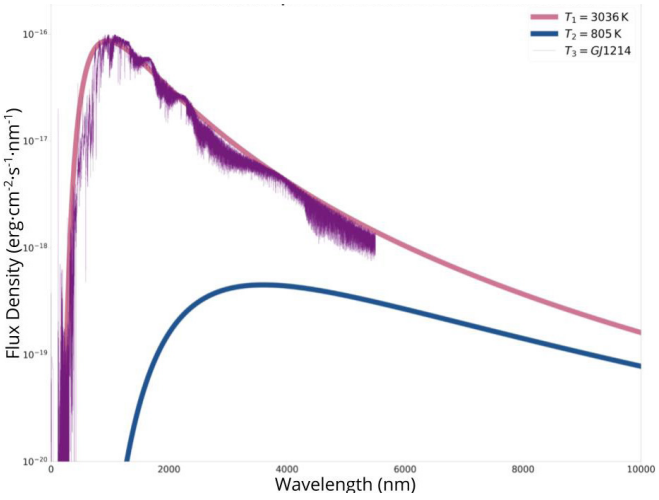


Figure 4. The theoretical Planck spectra of LHS 3844 and LHS 3844 b are plotted over the spectrum of GJ 1214, a star that has the same spectral type (M4.5) as LHS 3844.

Light Curve Modeling

Utilizing the Python programming language, we used the numerical flux and 1σ error bar values generated by AstroImageJ to create the transit light curve graph of LHS 3844 b. To study the light curve in terms of relative flux and transit depth, we normalized the baseline of the transit light curve to a value of one. We proceeded to use a Python package titled BATMAN to fit multiple light curve models to our data set of LHS 3844 b; these models were varying based on a given range of values for the radius and time of mid-transit parameters within the Python code (22). All other parameters used to create the best-fit light curve were based off of the previously published literature results from the initial TESS detection of LHS 3844 b (10).

Chi-squared goodness of fit tests were then performed on each data to model pairing:

$$\chi^2 = \sum_{i=1}^{k} \frac{(O_i - \mu_i)^2}{E_i^2}$$

where k is the number of sets of data points, i is the current set of data points, Oi is the observed transit data point, μi is the point on the predicted light curve that correlates with Oi, and Ei is the associated error (23). The purpose of this statistical test was to find the chi-squared value closest to one, which would signify a strong data to model fitting.

Since the transit light curve graph of LHS 3844 b is in terms of relative flux, we used the detected decrease in stellar flux to calculate the planet's normalized radius (RP/R*):

$$\frac{\Delta F}{F} = \left(\frac{R_p}{R_*}\right)^2$$

where ΔF is the observed change in flux and F is the stellar flux (24).

Emission Spectroscopy Metric

Kempton et al. (2018) recently developed an analytic emission spectroscopy metric (ESM) to quantify the expected signal-to-noise ratio (S/N) of a given planet's thermal emission spectroscopic measure:

$$ESM = 4.29\times10^6 \times \frac{B_{7.5}(T_{day})}{B_{7.5}(T_*)} \times \left(\frac{R_p}{R_*}\right)^2 \times 10^{-\frac{m_K}{5}}$$

where 4.29×10^6 is a scaling factor for the JWST infrared bandpass, B7.5 is Planck's function evaluated for a given temperature at a representative wavelength of 7500 nanometers (nm), Tday is the planet's dayside temperature in Kelvin ($1.10\times T_{eq}$), T* is the host star's effective temperature in Kelvin, RP/R* is the normalized planetary radius, and mK is the apparent magnitude of the host star in the K band. A predicted ESM threshold was determined from model simulated TESS exoplanets such that the population would yield approximately 300 high-quality candidates for follow-up with the JWST (25).

We used the ESM and predicted threshold to determine whether or not LHS 3844 b is likely qualified for future JWST

spectroscopic observation.

Planck Spectra

Planck's Law is given at a certain wavelength (λ) and Kelvin temperature (t), and it assumes that the given object behaves as a blackbody, which is an entity that hypothetically emits radiation in all wavelengths:

$$L(\lambda, t) = \frac{2\pi hc^2}{\lambda^5\left(e^{\frac{hc}{\lambda t}-1}\right)}$$

where h is Planck's constant (6.63×10^{-34} Js) and c is the speed of light (2.99×10^{8} ms^{-1}) (16).

We plotted the spectrum of GJ 1214 to act as a guideline for the theoretical Planck spectrum of LHS 3844; these two stars are the same spectral type (M4.5), and will therefore produce similar spectra (26). In addition, we overplotted the theoretical Planck spectrum of LHS 3844 b to assess the observed wavelength range that would produce the highest quality emission spectroscopic measure of LHS 3844 b and like planets.

ACKNOWLEDGEMENTS

We would like to thank Dr. Paul Strode (Department of Science, Fairview High School) for guiding us through this project, providing us with necessary resources, and teaching us about the wonders of science.

REFERENCES

1. Dotson, Renée, et al. "Exoplanets." University of Arizona Press, 2010, JSTOR:www.jstor.org/stable/j.ctt1814jx6.
2. Fridlund, Malcolm, et al. "The way forward." Astrophysical Journal, vol. 205, no. 4, Dec 2016, pp. 349-372., doi:10.1007/s11214-016-0247-2.
3. Wilson, Paul Anthony. "The Exoplanet Transit Method." 2015, www.paulanthonywilson.com/exoplanets.
4. Ricker, George R., et al. "The Transiting Exoplanet Survey Satellite." Journal of Astronomical Telescopes, Instruments, and Systems, 28 Oct 2014, doi:10.1117/1.JATIS.1.1.014003.
5. Tenenbaum, Peter, et al. "Detection of Transiting Planet Candidates in Kepler Mission Data." National Aeronautics and Space Administration, 6 June 2012, www-conf.slac.stanford.edu/statisticalissues2012.
6. Heath, Martin J., et al. "Habitability of planets around red dwarf stars." Origins of Life and Evolution of the Biosphere29.4 (1999): 405-424.
7. Brown, T.M., et al. "Las Cumbres Observatory Global Telescope Network." Publications of the Astronomical Society of the Pacific, vol. 125, no. 931, 10 Sept 2013, pp. 1031-1055., doi:10.1086/673168.
8. Clampin, M., et al. "Detection of Planetary Transits with the James Webb Space Telescope." European Space Agency, 15 Nov 2007, www.sci.esa.int/jwst/47523-clampin-m-et-al-2007/#.
9. Line, Michael R., et al. "A Systematic Retrieval Analysis of Secondary Eclipse Spectra." The Astrophysical Journal, vol. 779, no. 1, 14 Nov 2013, pp. 3-9., doi:10.1088/0004-637X/779/1/3.
10. Vanderspek, Roland K., et al. "TESS Discovery of an Ultra-short-period Planet Around the Nearby M Dwarf LHS 3844." The Astrophysical Journal Letters, vol. 871, no. 2, 19 Sept 2018, pp. 24-33., doi:10.3847/2041-8213/aafb7a.
11. Wright, Jason T. "Radial Velocities as an Exoplanet Discovery Method." Astro-ph, 25 July 2017, doi:10.1007/978-3-319-55333-7_4.
12. Weiss, Lauren M., Marcy, Geoffrey W. "The Mass-radius Relation for 65 Exoplanets Small than 4 Earth Radii." The Astrophysical Journal Letters, vol. 783, Issue 1, March 2014, doi: 10.1088/2041-8205/783/1/L6.
13. Roques, Francoise, et al. "The Extrasolar Planets Encyclopaedia." 1995, www.exoplanet.eu.
14. Mendez, Abel. "A Mass Classification for both Solar and Extrasolar Planets." Planetary Habitability Laboratory, University of Puerto Rico at Arecibo, 2011, www.phl.upr.edu.
15. Sullivan, Peter W., et al. "The Transiting Exoplanet Survey Satellite: Simulations of Planet Detections and Astrophysical False Positives." The Astrophysical Journal, vol. 809, no. 1, 10 Aug 2015, pp. 99-128., doi:10.3847/1538-4357/837/1/99.
16. Marr, Jonathan. "A Better Presentation of Planck's Radiation Law." Publications by Haverford College, Department of Physics, 2012, www.scholarship.haverford.edu/physics_facpubs.
17. Rodríguez-López, C., J. MacDonald, and A. Moya. "Pulsations in M dwarf stars." Monthly Notices of the Royal Astronomical Society: Letters 419.1 (2012): L44-L48.
18. Chaplain, Cristina T., et al. "James Webb Space Telescope: Integration and Test Challenges Have Delayed Launch and Threaten to Push Costs Over Cap". Feb 2018, www.gao.gov/assets.
19. JWST Team. "James Webb Space Telescope: Technical FAQ on a Variety of Mission Issues, Aspects, and Capabilities." www.jwst.nasa.gov/faq_scientists.
20. Observatorio El Sauce. "Observatorio El Sauce Telescope Hosting". 2018. https://obstech.cl.
21. Collins, Karen A., et al. "AstroImageJ: Image Processing and Photometric Extraction for Ultra-Precise Astronomical Light Curves (Expanded Edition)." AstroImageJ: ImageJ for Astronomy, 19 Jan. 2017, www.astro.louisville.edu/software/astroimagej.
22. Kreidberg, Laura. "batman: BAsic Transit Model cAlculatioN in Python." Publications of the Astronomical Society of the Pacific, vol. 127, 11 Aug. 2015, pp. 1161-1165., www.cfa.harvard.edu/~lkreidberg/batman.
23. Gould, Andrew. "Chi2 and Linear Fits." Astro-ph, 20 Oct 2003, arXiv:astro-ph/0310577v1.

24. Wright, Jason T., et al. “Exoplanet Detection Methods.” Astro-ph, 10 Oct 2012, arXiv:1210.2471.
25. Kempton, Eliza M.-R., et al. “A Framework for Prioritizing the TESS Planetary Candidates Most Amenable to Atmospheric Characterization.” Publications of the Astronomical Society of the Pacific, vol. 130, no. 993, 27 Sept 2018, doi:10.1088/1538-3873/aadf6f.
26. Allen, Jesse S. “The Classification of Stellar Spectra.” University College London, Department of Physics and Astronomy, 2001, www.star.ucl.ac.uk/~pac/spectral_classification.

Article submitted: September 30, 2019
Article accepted: October 30, 2019
Article published: May 10, 2020

JOURNAL OF
EMERGING INVESTIGATORS

Article

Comparing the effects of different natural products on reducing tumor growth in a *Drosophila* model

Aryan Ganesh, Ganesh Vanamu
Solorsano Middle School, Gilroy, California

SUMMARY

Cancer is one of the main causes of death in the present day and due to side-effects on current chemotherapies there is a need for testing natural products in cancer studies. In this work, we compared the effects of common natural products, including sesame, cinnamon, garlic, moringa and turmeric. The tumor model was a comparison of "rough-eye" versus "smooth-eye" studied in *Drosophila* eye, which provides easy detection of the phenotypes via simple microscopy. We measured the size of the eye, the number of flies with a smooth-eye phenotype, the number of flies with uniform ommatidia, and the lifespan of the flies. Our data showed that these natural products cannot be used to reduce tumors once it has completely formed. However, our data confirmed that these natural products can be used to reduce cancer cell growth when treated early. Moringa and sesame helped reduce tumor size when used in right concentrations. Our study also revealed a 15% moringa solution is the best candidate, of the ones tested, to prevent tumors. Flies treated with moringa showed longer eye lengths and widths compared to the control sample and close to healthy wild flies. For the flies treated with moringa, the number of flies with smooth-eye and uniform ommatidia increased with natural product concentration until 15% and decreased at 20%. Garlic, turmeric and cinnamon products did not show much efficacy on the tumor reduction. Though turmeric treatment did not prevent cancer, it helped increase the lifespan compared to control samples and other natural products.

INTRODUCTION

Cancer is one of the most prevalent diseases in the world and is one of the main causes of death in the present day. It has been estimated that deaths due to cancer will rise from 8.2 million to 14.6 million annually in 20 years (1). In 2018, there were around 1.8 million cancer cases in the United States, while 0.6 million cases led to death, resulting in a case mortality rate of ~ 30%. Among females, lung cancer is the leading cause of death followed by breast cancer. Among males, lung cancer is the primary reason for most deaths followed by prostate, colon, and pancreas cancer (2).

Chemotherapy is effective in many cases in the initial stages of many types of cancer, but the side effects are a major concern amongst patients. Patients' discomfort sometimes increases greatly due to side-effects. Both medical providers and patients are hoping to find a better and easier solution to treat cancer and above all prevent cancer if possible (3). One of the popular chemotherapy drugs called Cytoxan (cyclophosphamide) has several major side-effects like decreased blood cell counts, reversible hair loss, bladder damage, fertility impairment, and lung or heart damage at high doses (4).

Cancer occurs when cells do not go through the process of apoptosis, which is the death of cells which occurs as a normal and controlled part of an organism's growth or development (5). Instead of dying, the cells multiply uncontrollably. In the process of multiplying, the cells consume a lot of energy (6). A potential anticancer agent should cause selective apoptosis of tumor cells, with minimal damage to the normal cells. Natural products that are readily available such as garlic (*Allium sativum*), sesame (*Sesamum indicum*), moringa (*Moringa oleifera*), turmeric (*Curcuma*), and cinnamon (*Cinnamomum verum*) are natural antioxidants and potential options for cancer therapy (7-11). The harmful effects of chemotherapeutic agents, which can be toxic to non-cancerous and healthy cells, can perhaps be avoided with the use of natural products (12-13).

Garlic is usually used in cooking, not only as a flavor component but also for its medicinal properties. These properties are due to sulfur compounds which alter tumor growth, have cancer preventative potential, help improve immune function, and have many other therapeutic effects in treating infection among other properties to cover the list (14). Recently, Dong *et al* (2014) studied the ability of garlic to inhibit colon cancer growth via the inhibition of PI3K/Akt signaling pathway (15).

Sesame is one of the oldest and most important oilseeds (16). The medicinal properties of the leaves and the seed oil were identified four decades ago and hence sesame products are used as ingredients in several cooked dishes to leverage these benefits (17). The medicinal properties include anticancer, antioxidative, anti-immunoregulation and anti-hypersensitivity effects. Sesame consists of many macronutrients such as proteins, carbohydrates, antioxidants, lignans, and tocopherols. (18). Many sesame compounds like

sesamol, sesamin exert inhibitory effects on NF-κB and ERK/p38 MAPK signaling pathways and thereby suppress the inflammatory responses (19).

Moringa oleifera's leaves and bark have been a very popular food item since the 18th century (20). Several different studies suggest moringa leaves have antiviral, anti-heart disease, pro-immune, and anti-inflammatory properties (21-26). Moringa oleifera extract decreases the inflammatory mediators (NF-κB) but increases anti-inflammatory cytokines, IL-10 and IκB-α (27).

Turmeric root, which belongs to the ginger family, is widely used to add flavor to food in Asia and India. The main part of turmeric is curcumin, a yellow compound which purportedly has anti-cancer properties because of its ability to inhibit microsomal PGE2 synthase 1 (28). It has been shown to have several beneficial medicinal properties including anticarcinogenic, antioxidant, anti-inflammatory, antifertility, antidiabetic, antimicrobial, and antifibrotic effects. Due to turmeric's great natural antiseptic, disinfectant, anti-inflammatory, and analgesic properties it is also used in Ayurvedics, an alternative medicine involving a holistic ("whole-body") healing system (29). Administering the right dosages of curcumin to rodents showed great promise in preventing many types of cancer (30).

Cinnamon is a spice with a sweet and spicy fragrance that is used in cooking and aromatherapy. Cinnamon primarily contains vital oils and other derivatives, such as cinnamaldehyde, cinnamic acid, and cinnamate. Cinnamon has several medicinal properties such as antioxidant, anti-inflammatory, antidiabetic, antibacterial, and anticancer effects (31). Recent studies have shown that cinnamon also has activities that aid in the treatment of neurological disorders, such as Parkinson's and Alzheimer's diseases (32). Ongoing research aims to study the anti-cancer properties of cuminaldehyde as this compound activates the Nrf2 pathway, which is involved in strengthening the cellular response against stressors such as carcinogen exposure (33-34).

Drosophila melanogaster has been used as a model system for several decades to study the effects of various drugs (35-36). *Drosophila* have been used extensively in animal studies because they are invertebrates and can be easily studied. In addition, Drosophila are useful for studying human disease since nearly 75% of human disease-causing genes are present in the *Drosophila* genome (37).

An ideal system in which cancer can be studied is in the *Drosophila* eye, since tumors in this location do not impact the viability of the flies and can easily be detected and measured via use of simple microscopes. The oncogene RET in *Drosophila* eyes is linked to thyroid and lung cancers in humans. There are many ways in which tumors can be induced in the fly eye, the easiest being the overexpression of an oncogene using the GAL4/UAS (upstream activation sequence) flies which are heterozygous for the given driver and responder. GAL4 is a transcriptional activator derived from yeast, GAL4 binds UAS and turns on expression of the downstream gene. This system requires a cross between male flies carrying a GAL4 only expressed in the eyes (such as GMR-GAL4) and virgin females that carry the UAS-"oncogene" (for example UAS-RasV12 or UAS-RetMEN) (38-39). In this study we used the GMR-GAL4 cross with RetMEN system. The eyes of the offspring express tumors in the eye photoreceptors and have what is called a "rough eye" (an irregular and uneven, usually smaller eye).

Our work addresses several unanswered questions in the field, especially since there are currently few studies on the anti-cancer properties of moringa oleifera. While there are several studies using natural products to treat cancer, there is no systematic study comparing the effects of these specific natural products in a cancer model. We hypothesize that the common natural products sesame, cinnamon, garlic, moringa and turmeric will have an anti-cancer impact on *Drosophila*. Our data showed that these natural products cannot be used to reduce tumors once it has completely formed. However, our data confirmed that these natural products can be used to reduce cancer cell growth when treated early.

RESULTS

Testing the effects of 100% concentrated natural products on tumor growth and fly viability

Our goal was to compare the effects of each of the natural products at 100% concentration (i.e. garlic, sesame, moringa, turmeric and cinnamon) on tumor growth using the GMR-GAL4 cross with RetMEN system that expresses oncogenic RET in the drosophila eye, causing cancer. We placed 12-15 flies with eye tumors in one of five test vials, each containing one of the five natural products. We also set up two control vials with no natural product, containing either flies with eye tumors or wild type healthy flies. Four to five days after the experiment was set up, all the flies in the test vials died whereas control and reference samples were alive.

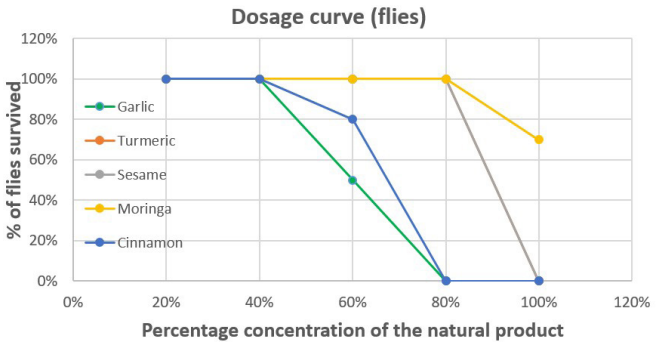


Figure 1. The number of flies survived in each of the vials vs the percentage concentration of each of the natural products (garlic, sesame, moringa, turmeric, and cinnamon)**.** All flies died in the garlic and cinnamon vials at concentrations above 40%. With turmeric and sesame, all flies died in the vials with concentrations greater than 80%. With moringa, flies in all vials survived except for those exposed to moringa at a 100% concentration. Note that the data for turmeric and sesame are overlapping. The plot for turmeric is hidden behind sesame plot in the graphs.

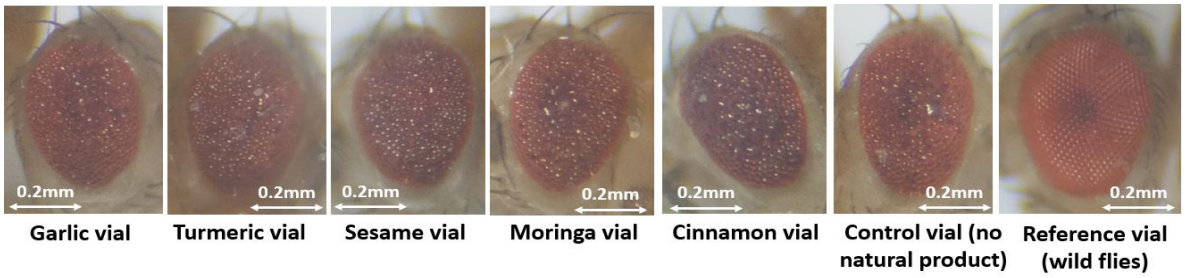


Figure 2. Representative images of the eyes of the flies from each of the natural products. From left to right: eyes of flies in vials with garlic, turmeric, sesame, moringa, cinnamon, Control sample (no natural products), and wild flies (healthy flies eyes).

This showed that a 100% concentration is lethal to the flies. Then we decided to determine the concentrations at which the flies would survive.

Dose-response curves can be used to determine the concentrations for each natural product at which flies remain viable. We sought to generate dose-response curves and use these concentrations for the experiment using the cancer flies. The number of flies that survived in each of the vials was plotted against the percentage concentration of each of the natural products (**Figure 1**). All the flies died in the garlic and cinnamon vials at concentrations above 40%. With turmeric and sesame, all the flies died in the vials with concentrations greater than 80%. With moringa, flies in all vials survived except for those exposed to moringa at a 100% concentration. Based on our results, we selected concentrations of 10%, 25%, and 40% for garlic and cinnamon, 25%, 50%, and 75% for turmeric and sesame, and 30%, 60%, and 90% for moringa.

Testing if we can treat cancer using dosage curve concentrations (on adult flies)

Next, we aimed to compare the effects of each of the various concentrations of the natural products (i.e. garlic, sesame, moringa, turmeric and cinnamon) on tumor characteristics. Vials with different concentrations (from the dosage curve) were prepared for each of the natural products (added 12-15 cancer flies) along with control sample vials, one with GAL4/UAS flies (without any natural product) and other healthy flies (without any natural product). Four to five days after the experiment, we examined the eyes of the flies in each vial under the microscope. Unfortunately, all the flies in all the test sample vials and concentrations showed blurry eyes with no improvement. Representative images of the eyes from each of the natural products are shown (**Figure 2**). This implies that the natural products may not be able to cure cancer once it is formed in the adult stages.

Imaginal discs of *Drosophila* eyes are formed during the initial larval stage (40). Most of the food that the flies eat during their life is consumed during the larvae stage (41**)**. So, we hypothesized that these natural products can help prevent cancer if the larvae feed on the experimental food.

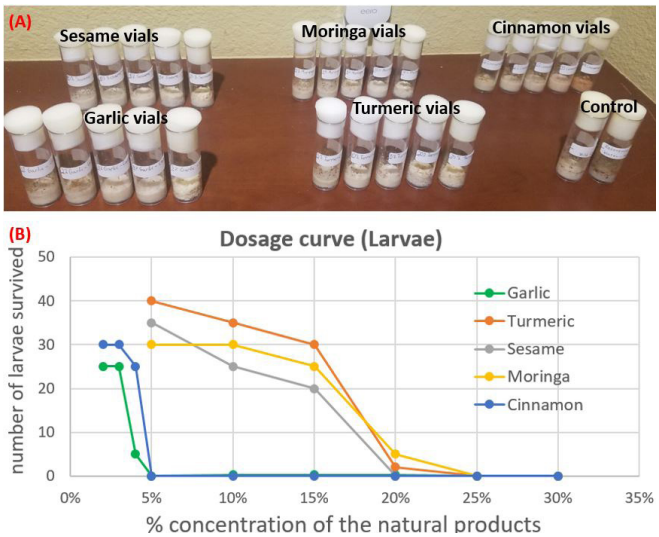


Figure 3. A) Experimental setup with flies in vials containing garlic, turmeric, sesame, moringa, cinnamon, and no product (control). **B)** Number of surviving flies in each vial plotted against the percentage concentration of each of the natural product.

Testing if we can reduce the tumors using dosage curve concentrations (on larvae)

We set up a cross between UAS-RetMEN males and GMR-GAL4 females and the larvae but after two weeks none of the experimental vials with natural products had any larvae in them. The control and reference vials had a normal number of larvae. This shows that the dosage curve for adult flies is different from that of the larvae. Hence, we sought to generate a dosage curve for the larvae.

One week after the larvae were exposed to various dosages of natural products, most of the vials had viable larvae except for vials with high concentrations of each of the natural products (**Figure 3A)**. We counted the larvae in each of the vials, and we plotted the survival rate of the larvae in each vial against the natural product concentration, generating dose-response curves (**Figure 3B**). Based on the results of the experiment, we selected concentrations of 1%, 2%, 3%, 4% and 5% for garlic and cinnamon, and 5%, 10%, 15%, 20% and 25% for sesame, moringa and turmeric.

The phenotype of rough versus smooth eyes can be visibly seen under the microscope (**Figure 4**). The *Drosophila* compound eye contains approximately 700-750-unit eyes known as 'ommatidia' (42). Ommatidia can be seen under the

microscope and can be evaluated for uniformity qualitatively.

A decrease in the size of the eye is one of the symptoms of cancer. The size of the eye is compared to that of the control and healthy flies' eyes based on the measurements of length and width of the eyes. The average length and width of the flies' eyes in each of the vials of the natural products was plotted compared to each other and the control and healthy flies' sample (**Figure 5A-B**). Around 10 flies for each Turmeric and sesame treatment groups were analyzed and the average length and width for each sample was calculated. Flies treated with turmeric and sesame showed smaller lengths compared to the control sample. Garlic- and cinnamon-treated samples showed the same lengths compared to the control sample. mMringa-treated samples showed longer lengths than the control sample and close to healthy wild flies. Turmeric- and sesame-treated samples showed smaller widths compared to the control sample. Garlic- and cinnamon-treated samples showed the same widths compared to the control sample. moringa-treated samples showed widths longer than the control sample and close to healthy wild flies. When compared the size among the different natural treatments, sesame- and turmeric-treated flies showed the smallest sizes (both length and width), moringa-treated flies showed the largest size and cinnamon- and garlic-treated samples had intermediate sizes.

We also performed a simple t-test to compare each of the natural treatment conditions to the control sample to determine if the treatment had a significant effect on the length and width of the flies' eyes. For the length of the eyes, 10%, 15%, 20% moringa samples were significantly higher than the control samples with the *p*-values <0.05. For the samples 10%, 15%, 20% turmeric, 5%, 10%, 15%, 20% sesame are significantly lower than the control samples with the *p*-values <0.05. Rest of all the samples are not significantly different to the control samples with the *p*-values >0.05.

For the width of the eyes, 15%, 20% moringa samples showed significantly higher than the control samples with the *p*-values <0.05. For the samples 5%, 10%, 15%, 20% turmeric, 10%, 20% sesame are significantly lower than the control samples with the *p*-values <0.05. Rest of all the samples are not significantly different to the control samples with the *p*-values >0.05.

We plotted the lifespan (or survival rate) of flies for each of the samples including the control sample (**Figure 6A-E**). As the number of days increased, the percent of flies that survived decreased after around 30 days. The natural products fall under three distributions as shown in the plots. Garlic falls in one distribution where all the flies survived until 30 days and all of them died around 80 days. Flies treated with garlic distribution matches to that of the control sample distribution. The average lifespan for garlic and control sample was around 50 days. sesame, cinnamon and moringa fall in a 2nd distribution where all the flies survived until 35-40 days and all of them died around 95-100 days. The average lifespan for sesame, cinnamon and moringa samples was around 65-70 days. Turmeric falls in the 3rd distribution where

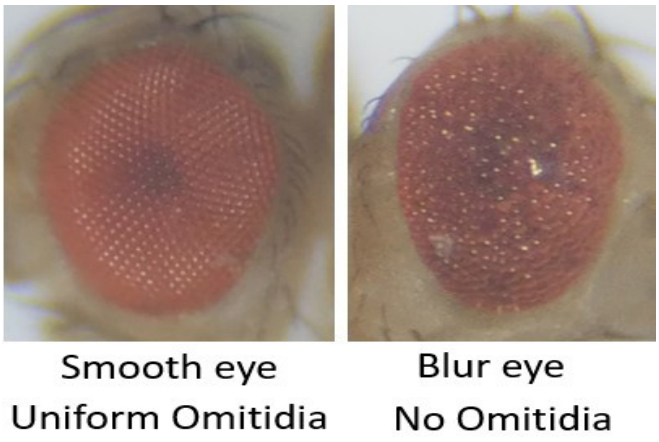


Figure 4. Comparison of Drosophila eye phenotypes visible under the microscope. Left image shows smooth eye and uniform ommatidia in healthy wild fly eye and right image shows rough eye and non-uniform ommatidia in cancer fly eye.

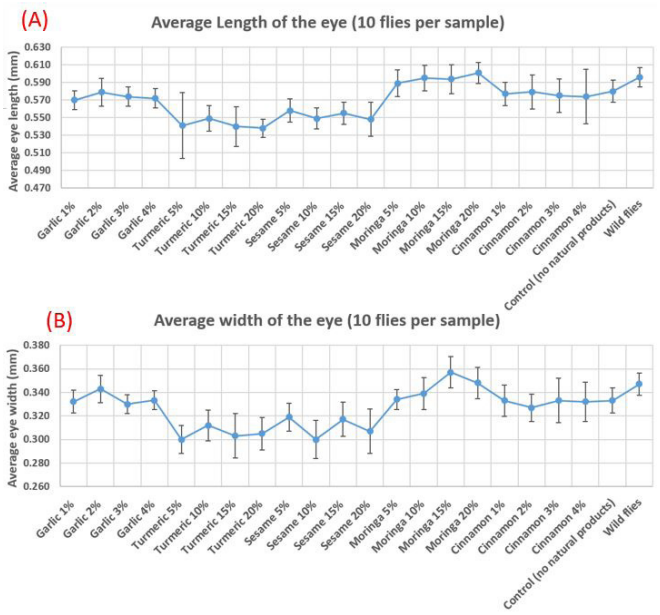


Figure 5. The average length and width of the flies' eyes in each of the vials of the natural products compared to each other and the control sample. A) For the length of the eyes, 10%, 15%, 20% moringa samples were significantly higher than the control samples with the *p*-values 0.034, 0.04, 0.014 respectively. For the samples 10%, 15%, 20% turmeric, 5%, 10%, 15%, 20% sesame are significantly lower than the control samples with the *p*-values 0.015, 0.038, <0.001, 0.049, 0.01, 0.022, 0.044 respectively. For the samples 5% moringa, 5% turmeric, 1%, 2%, 3%, 4% garlic, 1%, 2%, 3%, 4% cinnamon are not significantly different from the control samples with the *p*-values 0.23, 0.82, 0.68. 0.56, 0.62, 0.59, 0.33, 0.38, 0.45, 0.66 respectively.
B) For the width of the eyes, 15%, 20% moringa samples showed significantly higher than the control samples with the p-values 0.009, 0.039 respectively. For the samples 5%, 10%, 15%, 20% turmeric, 10%, 20% sesame are significantly lower than the control samples with the p-values 0.003, 0.038, 0.021, 0.028, 0.01, 0.041 respectively. For the samples 5%, 15% sesame, 5%, 10% moringa, 1%, 2%, 3%, 4% garlic, 1%, 2%, 3%, 4% cinnamon are not significantly different to the control samples with the p-values 0.69, 0.72, 0.27, 0.31, 0.36. 0.24, 0.33, 0.35, 0.33, 0.52, 0.39, 0.34 respectively.

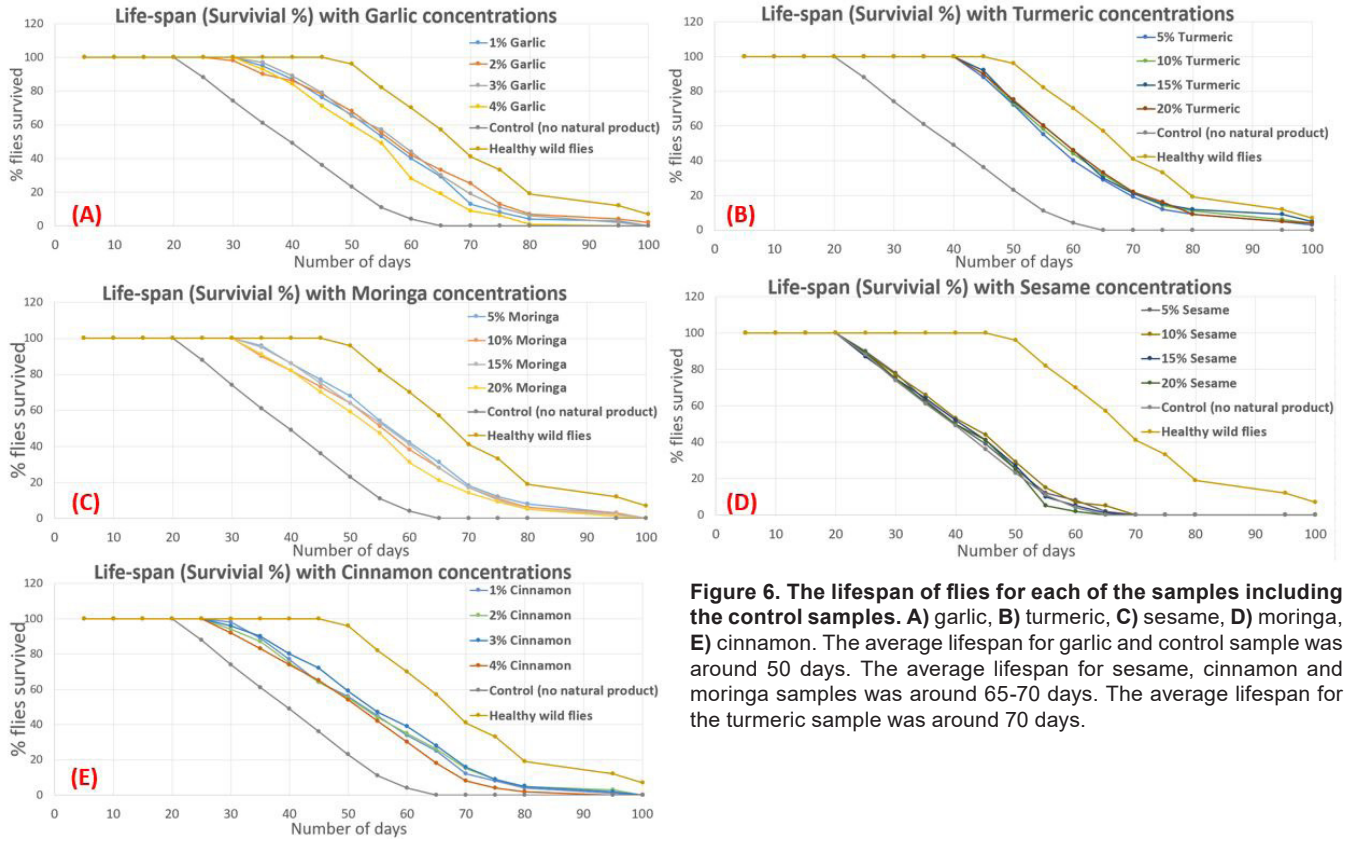


Figure 6. The lifespan of flies for each of the samples including the control samples. A) garlic, **B)** turmeric, **C)** sesame, **D)** moringa, **E)** cinnamon. The average lifespan for garlic and control sample was around 50 days. The average lifespan for sesame, cinnamon and moringa samples was around 65-70 days. The average lifespan for the turmeric sample was around 70 days.

all the flies survived until 50 days and all of them died around 100 days. The average lifespan for the turmeric sample was around 70 days. All the natural products showed increased lifespan (except for garlic which was similar to the control sample) compared to the control sample. Turmeric showed the longer lifespan compared to all the other natural products. The lifespan for the healthy wild flies is longer than all the natural products including the control sample, as expected.

No concentrations of garlic, turmeric, sesame, and cinnamon led to uniform ommatidia for any of the flies, and so did not prevent cancer (**Figure 7A-E**). All concentrations of moringa showed some flies with uniform ommatidia. For moringa, the number of flies with uniform ommatidia increased with natural product concentration until 15% and at 20% it decreased because the number of flies that survived were less. So, moringa showed better performance for tumor characteristics. Our study revealed that 15% moringa is the best candidate, among those tested in our study, to prevent tumors of the eye in this fly model of cancer. The number of flies with smooth eyes and the number of flies with uniform ommatidia for each of the vials including the control sample were plotted (**Figure 8A-B**). 5% and 10% sesame showed some flies with smooth eyes but all the concentrations of moringa showed flies with smooth eyes. For moringa the number of flies with smooth eyes increased with natural product concentration until 15% and at 20% it decreased because the number of flies that survived at 20% concentration were less. So, moringa showed better performance for tumor inhibitory characteristics.

DISCUSSION

We hypothesize that the common natural products sesame, cinnamon, garlic, moringa and turmeric will have an anti-cancer impact on *Drosophila*. This study showed that natural products (sesame, garlic, moringa, turmeric, cinnamon) cannot be used to treat cancer once it is completely formed during the adult stage of a fly's life cycle. During the larval stages, imaginal discs of the *Drosophila* fly's eyes are formed and during the pupal stages, eyes are developed (40). As a result, by the time we treated the adult flies with the experimental food, the tumors in their eyes were already formed. We therefore concluded that these natural products cannot cause tumor regression in adult flies. Since most of the food consumption in the flies is done during the larval stages, we hypothesized that treating flies with natural compounds before the larval stages could prevent cancer (41).

The larval lethal dose for each of the natural products, as defined by the dosage response curves, was lower than for adult flies. Since the larvae generally consume a large dose of food, they may not be able to survive at high concentrations as the adult flies do. Garlic and cinnamon lethal doses needed were considerably lower than other natural products for both larvae and the adult flies. It may be due to either the color or the odor of the two natural products. Cinnamon and turmeric

were noted to have a strong odor and garlic was observed to have a strong odor that is persistent.

Though these natural products may not slow down or reverse tumor growth, they could potentially be used for preventing cancer. Sesame showed some improvement in the tumor incidence in *Drosophila* eyes at 5% and 10% concentrations. At 10% sesame concentration, about 30% of the flies that were examined showed smooth eyes with no visible tumors. Although the flies in the sesame test samples demonstrated decreased tumor incidence, they did not show uniform ommatidia in the eyes. Moreover, the average length and width of the flies' eyes in the sesame samples were smaller compared to the control sample and most of the other natural products. Smaller eyes generally are a symptom of disease. This showed that sesame alone cannot prevent the cancer activity completely in these flies. This showed that sesame treated flies had decreased tumor incidence, they still had diseased eyes, meaning that sesame alone cannot prevent tumor formation.

Moringa showed better results in most of the concentrations. Interestingly, as the concentration of the moringa increased, the anti-cancer performance (tumor in the eyes and size of the eyes) improved, showing the best performance at 15% and decreased at 20%. At 20% concentration, the sample size was less than at 15% concentration, because only eight larvae survived at the higher concentration. Moringa showed bigger eye sizes, both lengths and widths, compared to the control sample and were very close to healthy wild flies. One thing to note is even in the moringa-treated flies, none of the vials showed 100% of the flies free of tumors. Some of the larvae may not have consumed or absorbed the moringa fully and hence when those larvae turned into adult flies still showed tumor symptoms.

Though sesame has better absorbance samples treated with sesame did not show the better performance compared to the samples treated with moringa, since moringa has an anti-tumor compound called niazimicin which helps prevent

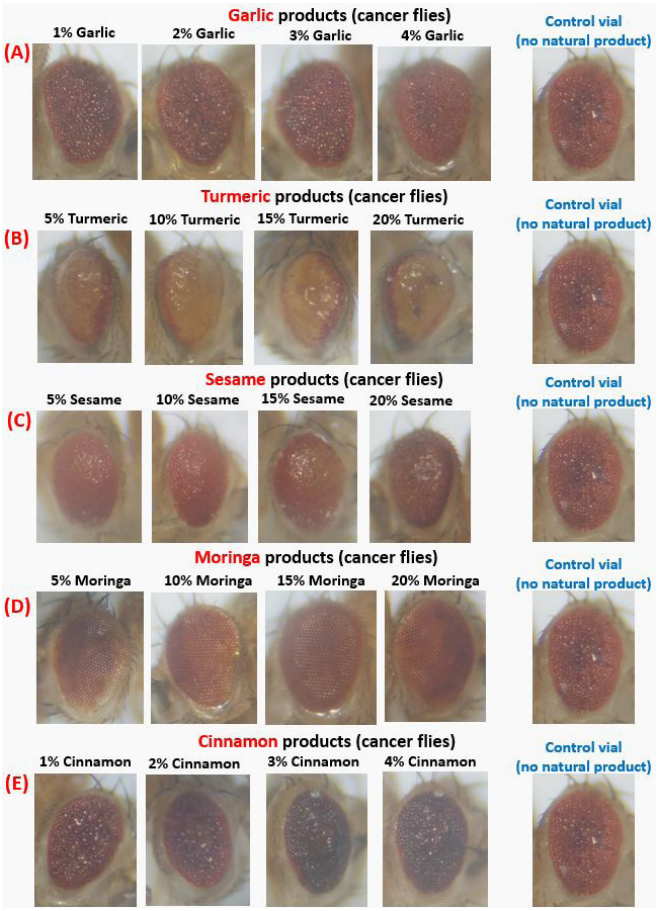


Figure 7. Representative pictures of the eye from each of the concentrations of the various extract vials compared to the control sample. A) garlic, **B)** turmeric, **C)** sesame, **D)** moringa, **E)** cinnamon.

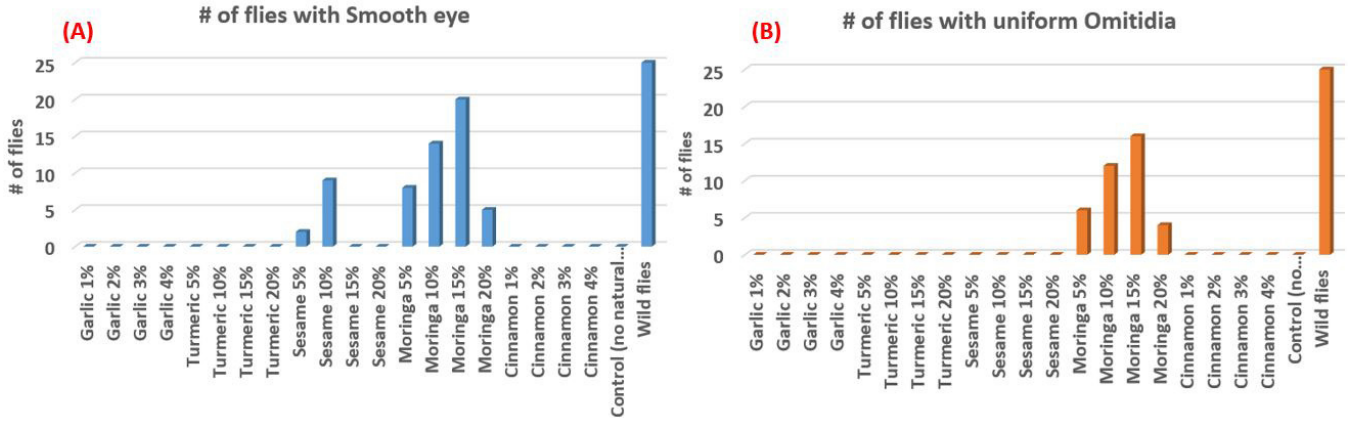


Figure 8. Flies with smooth eye and ommatidia. A) The number of flies with smooth eye for each of the test vials (out of 25 total flies taken for analysis), across five concentrations of each natural product, and the control sample. **B)** The number of flies with uniform ommatidia for each of the test vials (out of 25 total flies taken for analysis), across four concentrations of each natural product, and the control sample.

tumors, it showed better performance compared to other natural products (43-44). It is interesting to note that niazimicin is not present in any of these other natural products.

Though turmeric products did not prevent cancer, it helped increase the lifespan compared to control samples and other natural products. Curcumin impacts ageing and age-related diseases at the organismal and cellular level. On an organismal level, curcumin helps postpone the ageing symptoms and delays the on-set of diseases related to old-age (45). On cellular level, Curcumin enhances the activity of some of the anti-aging proteins and decreases the activity of some of the age-causing proteins (e.g., NF-κB, mTOR).

However, there are some limitations in this study. The microscope used in this study had low resolution, and as such, limited the accuracy of the eye measurements. Although the flies were anesthetized for image acquisition, they still moved slightly, making it difficult to capture good quality images. The methodology used to estimate the size of the flies' eyes is not robust as it assumed the size of the wild flies and compared the test samples to that. Since we are comparing the size of the eyes among each of the samples the actual numerical values are not as important for tumor performance study.

In this study, we did not test the performance of a combination of different natural products. Since both sesame and moringa showed promising results in terms of decreasing tumor activity, we plan to combine these natural products in various proportions and test the various endpoints in future experiments. Also, for future study we would like to combine these natural products with medicines that are already FDA-approved for the treatment of cancer, such as chemotherapy.

MATERIALS AND METHODS

Summary of the procedure

To test the effectiveness of 100% concentration of natural compounds, we took five test vials with each of the natural products and 12-15 cancer flies (eye tumors) were added into all of them. A 6th vial was added without any natural products (just water) to compare with the test samples. We also added a 7th vial with wild flies (healthy flies without cancer) for reference to compare regular healthy flies' eyes. Natural Products that were used include commercially available pure garlic, turmeric, sesame, moringa, and cinnamon. It was noted that the pure cinnamon mixed poorly with the food. Flies were allowed to feed for 4-5 days. Pure natural products are denoted as 100% concentration.

In order to test the effectiveness of lower concentrations of natural compounds obtained using dosage curve (on adult flies), we took 25 vials (5 for each natural product) with varying concentration and 12-15 cancer flies (eye tumors) were added to the vials. All the natural products (except cinnamon) were diluted using water. Cinnamon, which was not soluble in water, was diluted using PEG-400 (poly-ethylene glycol). A 16th vial was added without any natural products (just water) to compare with the test samples. We also added a 17th vial with wild flies (healthy flies without cancer) for reference to compare regular healthy flies' eyes and waited for 4-5 days till the flies fed on the food.

To test the effectiveness of lower concentrations of natural compounds obtained using dosage curve (on larvae), we set up the 17 vials including control and reference vials. We then set up a cross between UAS-RetMEN males and GMR-GAL4 females by adding 6-8 flies each type. The flies were allowed to mate for 4-5 days.

All the parent flies were separated from the vials to prevent confusion with the progeny. The larvae were grown into flies and checked their eyes for the "rough eye" tumor/overgrowth. After 2 weeks all the larvae turned into adult flies. The eyes of all the flies were examined in the microscope and measurable endpoints such as length and width of the eye, number of flies with no blur, number of flies with uniform ommatidia, lifespan of the eyes were used to compare the performance of each of the natural products for cancer prevention.

Drosophila melanogaster care

Drosophila wildtype flies were purchased from Carolina Biological Supply Company. The mutant strains (RetMEN2B mutant gene and UAS-GAL4 driver gene) were purchased from Bloomington *Drosophila* Stock Centre. The *Drosophila melanogaster* were cared for using the guidelines of the Carolina ® *Drosophila* manual (46). The flies in all the vials were fed with Formula 4-24® Instant *Drosophila* Medium food (15ml) and water(15ml) 6-8 grains of yeast were added to each of the vials. The stock of cancer flies was prepared by taking several vials and setting up a cross between flies having RetMEN2B mutant gene and UAS-GAL4 driver gene.

Dosage curve

For the adult flies' dose response, we took 20 vials (4 vials for each natural product) with varying concentrations from 20-100% and added wild flies (assuming the survival conditions of the wild flies were similar to the cancer flies). Then 20%, 40%, 60% and 80% concentrations of each product were prepared and added to the vial. The total volume of the natural products and water was always maintained at 15ml. For example, to prepare 20% moringa concentration we added 3ml of moringa to 12ml of water and mixed it well. After 4-5 days, the total number of flies that survived in each vial were counted and plotted against the concentrations.

For the larvae dosage curve, twenty vials (four for each natural product) were prepared with varying concentrations. The 21st vial without any natural products (just water) and 22nd vial with wild flies was also included. Then in each of the 21 vials (including the control sample), 6-8 UAS-RetMEN and GMR-GAL4 flies were added.

Measuring size of the eyes, lifespan of the flies

The microscope used in this study was AmScope SE306R-PZ-LED Stereo Microscope, 20X/40X/80X Magnification. The size of the eyes was measured by comparing them with the known size of wild flies' eyes. The size of the wild flies' eyes

was assumed as 0.6mm long and 0.35mm wide (47). Pictures of all the flies were taken under the microscope with the same zoom and lens. The eye sizes of the test samples are calculated by comparing the eye sizes of both the samples on the pictures. Around 10 flies were taken and averaged to get the length and width of eyes of flies in each vial. The lifespan of the flies was measured by counting the number of flies that died every 5 days (initially when few of them died) and eventually counted the number of flies that survived (when most of the flies died).

Statistical analysis

A simple two-sample t-test was used to test the difference between the lengths and width of the eyes of the various test samples against the control or healthy flies. We then obtained the *p*-value (the probability that the sample means are different) for each of the comparisons. If the p value was <0.05 we reject the null hypothesis that there is no difference between the means and conclude that a significant difference does exist.

ACKNOWLEDGEMENTS

We would like to acknowledge Dr. Ainhoa Perez and Ms. Chana Hecht for providing initial samples of the Drosophila flies and clarifying questions about crossing the flies and the experiment which improved our understanding of the overall study. Furthermore, we wish to thank the reviewers and editors for giving us valuable constructive feedback so we could improve our report.

Received: February 21, 2020
Accepted: May 25, 2020
Published: May 31, 2020

REFERENCES

1. Stewart *et al*. "Cancer prevention as part of precision medicine: 'plenty to be done'." *Carcinogenesis,* vol. 37, no. 1, 2015, pp. 2–9.
2. Freddie Bray *et al*. "Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries." *American Cancer Society,* CA: *A Cancer Journal for Clinicians*, vol. 68, no. 6, November/December 2018, pp. 394-424
3. Kulmira *et al*. "Adverse Effects of Cancer Chemotherapy: Anything New to Improve Tolerance and Reduce Sequelae?" *Front Pharmacol.,* Vol 9, 2018, pp. 245.
4. Teles *et al*. "Cyclophosphamide administration routine in autoimmune rheumatic diseases: a review" *Brazilian Journal of Rheumatology* (English Edition), vol 57, no. 6, 2017, pp. 596-604
5. Rebecca *et al*. "Apoptosis in cancer: from pathogenesis to treatment" *Journal of Experimental & Clinical Cancer Research,* vol. 30, no .1, 2011, pp. 87.
6. Gonzalez-Perez *et al*. "IntOGen-mutations identifies cancer drivers across tumor types." *Nature Methods.,* vol. 10, no.11, 2013, pp.1081-2, doi: 10.1038/nmeth.2642. Epub 2013 Sep 15.
7. Schäfer *et al*. "The immunomodulation and anti-inflammatory effects of garlic organosulfur compounds in cancer chemoprevention." *Anti-Cancer Agents in Medicinal Chemistry,* vol. 14, no .2, 2014, pp. 233-40.
8. Abou-Gharbia *et al*. "Effects of processing on oxidative stability of sesame oil extracted from intact and dehulled seeds." *Journal of the American Oil Chemists' Society*, vol 74, no. 3, 1997, pp. 215-221.
9. Mukunzi *et al*. "Comparison of volatile profile of moringa oleifera leaves from Rwanda and China using HS-SPME." *Pakistan Journal of Nutrition,* vol. 10, 2011, pp. 602–608.
10. Wilken *et al*. "Curcumin: a review of anti-cancer properties and therapeutic activity in head and neck squamous cell carcinoma." *Mol. Cancer,* vol. 10, no. 1, 2011, pp.12.
11. Ho-Keun *et al*. "Cinnamon extract induces tumor cell death through inhibition of NFκB and AP1" *BMC Cancer*, vol. 10, no. 392, 2010.
12. Khamphio *et al*. "Sesamol induces mitochondrial apoptosis pathway in HCT116 human colon cancer cells via pro-oxidant effect." *Life Sci.,* vol. 158, 2016, pp. 46–56.
13. Nayak *et al*. "Sesamol prevents doxorubicin-induced oxidative damage and toxicity on H9c2 cardiomyoblasts." *J. Pharm. Pharmacol.,* vol. 65, no. 7, 2013, pp.1083–1093.
14. Singh *et al*. "Medicinal values of garlic (Allium sativum L.) in Human Life: An Overview" *Greener Journal of Agricultural Sciences*., vol. 4, no. 6, 2014, pp. 265-280
15. Dong *et al*. "Aged black garlic extract inhibits HT29 colon cancer cell growth via the PI3K/Akt signaling pathway." *Biomedical Reports* 2, 2014, pp. 250-254
16. Komivi *et al*. "The Emerging Oilseed Crop Sesamum indicum Enters the "Omics" Era" *Front Plant Sci*., vol. 8, 2017, pp. 1154
17. Geetha *et al*. "Sesamol: an efficient antioxidant with potential therapeutic benefits." *Medicinal chemistry,* vol. 5, no .4, 2009, pp. 367–371.
18. Prasad *et al*. "A Review on Nutritional and Nutraceutical Properties of sesame." *Journal of Nutrition & Food Sciences*, vol. 2, no. 2, 2012, pp. 1-6
19. Majdalawieh *et al*. "Sesamol, a major lignan in sesame seeds (Sesamum indicum): Anti-cancer properties and mechanisms of action." *European Journal of Pharmacology,* vol. 855, 2019, pp. 75-89.
20. Chumark *et al*. "The in vitro and ex vivo antioxidant properties, hypolipidaemic and antiatherosclerotic activities of water extract of moringa oleifera Lam leaves." *Journal of Ethnopharmacology,* vol. 116, 2008, pp. 439–446.
21. Miyachi *et al*. "Benzyl isothiocyanate inhibits excessive superoxide generation in inflammatory leukocytes: implication for prevention against inflammation-related carcinogenesis." *Carcinogenesis,* vol. 25, 2004, pp. 567–

575.
22. Faizi *et al*. "Isolation and structure elucidation of new nitrile and mustard oil glycosides from moringa oleifera and their effect on blood pressure." *Journal of Natural Products,* vol. 57, 1994, pp. 1256–1261.
23. Waiyaput *et al*. "Inhibitory effects of crude extracts from some edible Thai plants against replication of hepatitis B virus and human liver cancer cells." *BMC Complementary and Alternative Medicine,* vol. 12, 2012, 246–252.
24. Sultana *et al*. "Effect of extraction solvent/technique on the antioxidant activity of selected medicinal plant extracts." *Molecules,* vol. 14, 2009, pp. 2167– 2180.
25. Kumar *et al*. "Antibacterial and antioxidant activity of extract of moringa oleifera leaves-An in vitro study." *International Journal of Pharmaceutical Sciences Review and Research,* vol. 12, 2012, pp. 89–94.
26. Kumar *et al*. "Retinoprotective Effects of moringa oleifera Via Antioxidant, Anti-Inflammatory, and Anti-Angiogenic Mechanisms in Streptozotocin-Induced Diabetic Rats." *Journal of ocular pharmacology and therapeutics,* vol. 29, no .4, 2012, pp. 420-426
27. Tan *et al*. "moringa oleifera Flower Extract Suppresses the Activation of Inflammatory Mediators in Lipopolysaccharide-Stimulated RAW 264.7 Macrophages via NF-κB Pathway." *Mediators of Inflammation,* 2015, pp. 1-11
28. Koeberle *et al*. "Curcumin blocks prostaglandin E2 biosynthesis through direct inhibition of the microsomal prostaglandin E2 synthase-1." *Mol Cancer Ther.,* vol. 8, 2009, pp. 2348-2355.
29. Verma *et al*. "Medicinal properties of turmeric (Curcuma longa L.): A review" *International Journal of Chemical Studies,* vol. 6, no. 4, 2018, pp. 1354-1357
30. Sharma RA *et al*. "Curcumin: the story so far." *European Journal of Cancer,* vol. 41, 2005, pp. 1955-1968
31. Dhanushka *et al*. "Anti-inflammatory activity of cinnamon (C. zeylanicum and C. cassia) extracts - Identification of E-cinnamaldehyde and o-methoxy cinnamaldehyde as the most potent bioactive compounds" *Food & Function,* vol. 6, no. 3, 2015
32. Kuen-daw *et al*. "Cuminaldehyde from Cinnamomum verum Induces Cell Death through Targeting Topoisomerase 1 and 2 in Human Colorectal Adenocarcinoma COLO205 Cells" *Nutrients,* vol. 8, no. 6, 2016, pp. 318
33. Rao *et al*. "cinnamon: A Multifaceted Medicinal Plant" *Evidence-Based Complementary and Alternative Medicine,* 2014, pp. 1-12
34. Wondrak *et al*. "The Cinnamon-Derived Dietary Factor Cinnamic Aldehyde Activates the Nrf2-Dependent Antioxidant Response in Human Epithelial Colon Cells." *Molecules,* vol. 15, 2010, pp. 3338-3355
35. Pandey *et al*. "Human disease models in *Drosophila melanogaster* and the role of the fly in therapeutic drug discovery." *Pharmacol. Reviews,* vol. 63, no. 2, 2011, pp. 411–436.
36. Huo *et al*. "The applications and advantages of *Drosophila melanogaster* in cancer research." *Hereditas,* vol. 36, no. 1, 2014, pp. 30–40
37. Fortini *et al*. "A Survey of Human Disease Gene Counterparts in the Drosophila Genome" The *Journal of Cell Biology,* vol. 150, no. 2, 2000, pp. F23-F29
38. Zoranovic *et al*. "A genome-wide Drosophila epithelial tumorigenesis screen identifies Tetraspanin 29Fb as an evolutionarily conserved suppressor of Ras-driven cancer" *PLOS Genetics*, 2018, pp. 1-36
39. Das *et al*. "*Drosophila* as a Novel Therapeutic Discovery Tool for Thyroid Cancer" *THYROID*, vol. 20, no. 7, 2010, pp. 689-95
40. Dübendorfer *et al*. "Development and differentiation in vitro of *Drosophila* imaginal disc cells from dissociated early embryos" *Journal of embryology and experimental morphology,* vol. 33, no. 2, 1975, pp. 487-498
41. Zera *et al*. "The physiology of life history trade-offs in animals." *Annu. Rev. Ecol. Syst.,* vol. 32, 2001, pp. 95-126
42. Ross *et al*. "Principles of Drosophila Eye Differentiation" *Curr Top Dev Biol.,* vol. 89, 2009, pp.115–135
43. Abdul Wakil *et al*. "Evaluation of rice bran, sesame and moringa oils as feasible sources of biodiesel and the effect of blending on their physicochemical properties" *RSC Advances*, vol. 4, no. 100, 2014, pp. 56984-56991
44. Natural bioactive compound from moringa oleifera against cancer based on in silico screening" *Jurnal Teknologi,* vol. 78, no. 5, 2016, pp. 315-318
45. Anna Bielak-Zmijewska *et al*. "The Role of Curcumin in the Modulation of Ageing." *International Journal of Molecular Sciences,* vol. 20, no. 5, 2019, pp. 1239
46. Flagg, Raymond O., Ph.D. Carolina ® *Drosophila* Manual. N.p.: *Carolina Biological Supply*, 2005.
47. Characterization of the Genetic Architecture Underlying Eye Size Variation Within *Drosophila melanogaster* and *Drosophila* simulans, *G3: GENES, GENOMES, GENETICS*, vol. 10, no. 3, 2020, 1005-1018.

Sponsorship

Editor's Circle $10,000+

ALLOY THERAPEUTICS

Patron $5,000+

NEW ENGLAND BioLabs® Inc.
enabling technologies in the life sciences

PORTFOLIOS WITH PURPOSE®

Institutional Supporters

HARVARD UNIVERSITY

HARVARD MEDICAL SCHOOL

Tufts UNIVERSITY

Charitable Contributions

We need your help to provide mentorship to young scientists everywhere.

JEI is supported by an entirely volunteer staff, and over 90% of our funds go towards providing educational experiences for students. Our costs include manuscript management fees, web hosting, creation of STEM education resources for teachers, and local outreach programs at our affiliate universities. We provide these services to students and teachers entirely free of any cost, and rely on generous benefactors to support our programs.

A donation of $30 will sponsor one student's scientific mentorship, peer review and publication, a six month scientific experience that in one student's words, 're-energized my curiosity towards science', and 'gave me confidence that I could take an idea I had and turn it into something that I could put out into the world'. **If you would like to donate to JEI, please visit https://emerginginvestigators.org/support, or contact us at questions@emerginginvestigators.org.** Thank you for supporting the next generation of scientists!

'Journal of Emerging Investigators, Inc. is a Section 501(c)(3) public charity organization (EIN: 45-2206379). Your donation to JEI is tax-deductible. '

emerginginvestigators.org