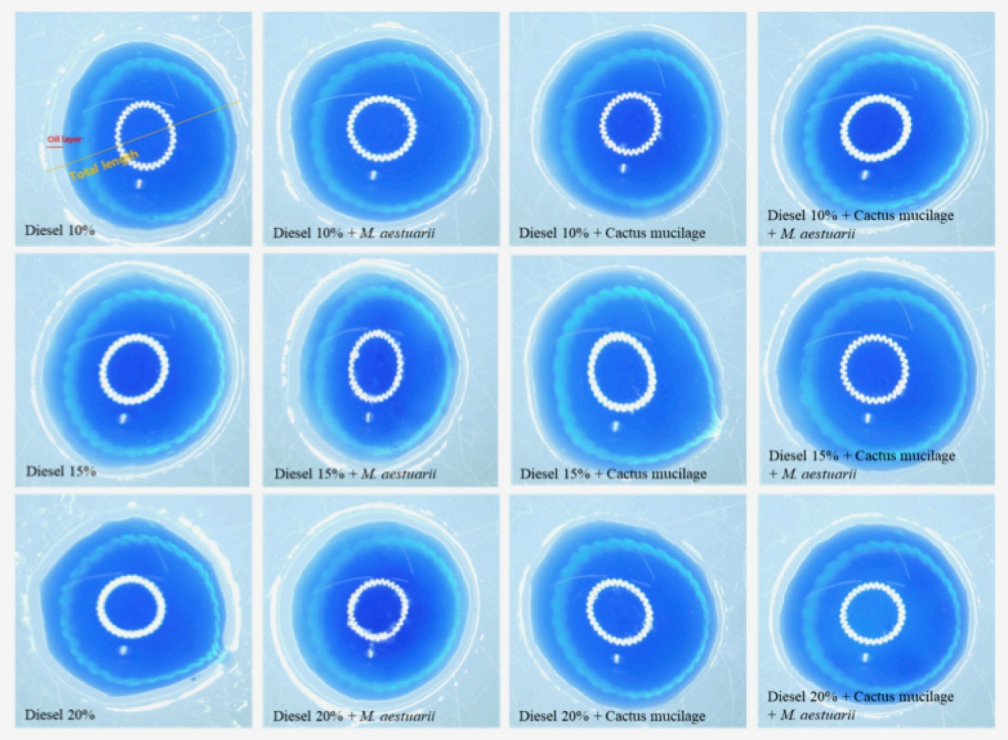
CJSJ

COLUMBIA JUNIOR SCIENCE JOURNAL



Development of Cactus Oil Dispersant Suitable for Oil-Degrading Marina Bacteria and Various Seawater Temperatures
Page 25

Volume 6 | 2020 - 2021 | cjsjournal.org

TABLE OF CONTENTS

ORIGINAL RESEARCH

REVIEW ARTICLES

Dear Readers,

I am so delighted to announce the publication of the sixth edition of the *Columbia Junior Science Journal*! We, as editors of this journal, have always strived to spotlight outstanding research done by high school students, and this issue features a collection of impressive and novel work in a wide range of scientific disciplines.

I would like to take this opportunity to thank everyone who was involved in the inception of this edition: the *Columbia Undergraduate Science Journal*, the *Columbia Junior Science Journal* editorial board, the wonderful students featured, their research mentors, their parents and peers, and anyone else who contributed. To the authors - you have made the best of a difficult year and we are all inspired by your drive and your unwavering commitment to scientific excellence.

During this tumultuous and confusing year of COVID-19, we recognize how difficult it has been to obtain the resources and laboratory access necessary for hands-on research. Due to these unfortunate circumstances and our dedication to making scientific publication as accessible as possible, we now accept both original research articles and review papers for CJSJ. We are proud to present both types of research in this issue. As we broaden both the scope and form of the scientific literature presented, we have received a record number of submissions this year, and we are ecstatic that the expansion of our journal has granted more students the opportunity to submit their research.

This unique edition of the *Columbia Junior Science Journal* explores a broad spectrum of scientific innovation ranging from microbial fuel-cell development to gamma-ray analysis of blazars, and it also includes timely work on COVID genetic analysis and diagnosis. The cover page features oil-water drops from a paper investigating the oil-degrading benefits of a natural dispersant from cactus mucilage. We hope that you read this set of articles as one cohesive publication and can be inspired, as we at CJSJ are, by the endless potential of the next generation's incredible scientific thinkers.

It has been a great pleasure serving as the 2020-2021 Editor-in-Chief of CJSJ and watching our editorial board, as well as the high school students, grow throughout this entire publication process. I look forward to seeing and reading about all of your continued scientific endeavors for many years to come!

Sincerely,

Deena Shefter
Editor-in-Chief
Columbia Junior Science Journal

Dear Readers,

It is with great pleasure that I announce the publication of the 2021 edition of the *Columbia Junior Science Journal*. This year's edition features innovative work from the next generation of scientists in a wide range of scientific disciplines. As President/Chief Editorial Officer of the *Columbia Undergraduate Science Journal*, I've had the privilege to work with these fine students, and I am impressed and encouraged by their curiosity and intellectual vigor.

This year, we received a record number of submissions to the *Columbia Junior Science Journal*. This is a testament not only to the vibrancy of the young, growing scientific community, but also to the tenacity of its members. Amidst a global crisis like no other, all who submitted to the *Columbia Junior Science Journal* shared in a great scientific endeavor that gives us hope for better days to come. I would like to thank each student, teacher, research mentor, parent, and everyone who made these submissions possible.

We must acknowledge that there are significant disparities in access to the resources necessary to produce work of the quality that is published in the *Columbia Junior Science Journal*. There are few high schools that are able to connect their students with laboratory work, and even fewer that have on-site research facilities and resources. By accepting review articles for the first time, the *Columbia Junior Science Journal* aims to provide all budding researchers with the opportunity to showcase their scientific skills. We will continue to work towards a more equitable scientific community for all students.

I am inspired by the prospect of a scientific future led by these fine young scientists. I encourage them to remember that the scientific process is constantly evolving, and that they are now part of that evolution. They should strive to maintain a scientific mind; experimentation, logical thinking, and curiosity should remain a fundamental principle.

The *Columbia Undergraduate Science Journal* editorial board is proud to present the 6th edition of the *Columbia Junior Science Journal*. I am grateful for the hard work of the *Columbia Junior Science Journal* team, and I specifically commend the outstanding effort of the *Columbia Junior Science Journal* Editor-in-Chief, Deena Shefter. In selecting, curating, and refining the articles here, she has overseen the production of our largest publication yet. I would also like to thank the *Columbia Undergraduate Science Journal* Faculty Advisory Board, a group of esteemed Columbia University professors whose support of the young scientific community ensures its growth and prosperity. Congratulations to all authors, and thank you to our readers!

Arya Rao
President, Chief Editorial Officer
Columbia Undergraduate Science Journal

2020-2021 Editorial Board

CJSJ Editor-in-Chief: Deena Shefter
President, Chief Editorial Officer: Arya Rao
CUSJ Editor-in-Chief: Isabella Leite
Director of Communications: Elin Hu
Director of Events: Hibah Vora
Director of Finance: Kanav Kalucha
Director of Internal Affairs: Kate Johnson
Director of Outreach: Shivali Verma

CJSJ Editors

Noah Krever
Mayeesa Rahman
Jennifer Lin
Teodora Bratu
Gloria Charite
Kazi Rabbe
Aswath Suryanarayanan

Editors

Jason Mohabir
Ilaria Simeone
Ashley Rosenberg
Katie Long
Neeraj Sakhrani
Ethan Chen
Maria Trifas
Caroline Haoud
Sophia Ladyzhets
Sahith Vadada
Lydia Wu
Arjun Kudinoor
Chloe Gong
Ellen Ren
Unal Yigit Ozulku
Shamara Yearwood

Associate Editors

Mehdi Elfilali
Nadiia Dubchak
Stan DeLaurentiis
Samantha Turk
Aaron Jackson
Anushka Gupta
Shloka Janapaty
Vanshika Sriram
Zoe Zager
Laurelle Lee Young
Natasha Sharma
Ana Beatriz Bonfim

Faculty Advisory Board

Hugh Ducklow.....Professor of Earth & Environmental Sciences
Ivana Hughes.....Director of Frontiers of Science, Senior Lecturer in Discipline – Chemistry
Marko Jovanovic.....Assistant Professor of Biological Sciences
Laura Kaufman.....Professor of Chemistry
Patricia Lindemann.....Lecturer in Discipline, Director of Undergraduate Studies – Psychology
Kyle Mandli.....Associate Professor of Applied Mathematics
Gerard Parkin.....Professor of Chemistry
Ron Prywes.....Professor of Biological Sciences
Rachel Rosen.....Associate Professor of Physics
Samuel Sternberg.....Assistant Professor of Biochemistry and Molecular Biophysics

PulmXNet: A Novel Deep Learning Architecture for the Diagnosis of COVID-19 Alongside Other Respiratory Illnesses (Bacteria and Viral Pneumonia, Tuberculosis) from Chest X-rays

Justin Shen[1] and Davesh Valagolam[1]

[1i, 1ii]Jericho Senior High School, 99 Cedar Swamp Rd, Jericho, NY 11753

***Abstract*—The novel coronavirus (COVID-19) has brought tremendous international detriment as it has forced the world into a global pandemic and shutdown. COVID-19 and other respiratory illnesses directly impact the pulmonary tract of the human body, leading to noticeable structural differences which can be identified via a chest x-ray (CXR). This study highlights the implementation of a novel convolutional neural network (CNN) in the diagnosis of 5-classes of CXRs (normal, COVID-19, tuberculosis, bacterial and viral pneumonia). This model was trained over 17 epochs and had 3,403,061 parameters. The overall 5-class accuracy for PulmXNet was 86.24%. The individual accuracy values for normal, COVID-19, tuberculosis, and bacterial and viral pneumonia were 85.59%, 85.45%, 84.16%, 89.33%, and 85.52%, respectively. This model was the first 5-Class CNN model to classify for normal, COVID-19, tuberculosis, bacterial and viral pneumonia CXRs and was able to maintain a relatively high accuracy and low parameters, demonstrating computational efficiency. Future investigations should look to further optimize this model with Bayesian or random search optimizations.**

I. Introduction

The novel coronavirus (COVID-19) first surfaced in Wuhan, China in December 2019, later escalating to global pandemic status prompting warranted concern from governmental and healthcare officials [2].

As COVID-19 directly impacts the respiratory tract, clear morphological abnormalities can be noticed in chest x-rays (CXRs) [3]. Similar morphological abnormalities are present in other respiratory illnesses such as pneumonia and tuberculosis allowing for their diagnosis through CXRs as well [3].

However, the main limitation with radiology-aided diagnosis is the high cost associated with the use of trained clinical radiologists. As a result, with the advancement of recent deep learning methodologies, prior research has identified the merit of using a hybrid radiology-deep learning methodology to diagnose respiratory illnesses (COVID-19, pneumonia, tuberculosis) [4]. These procedures can often be cheaper as well as more accurate than traditional radiological practices. Already, certain deep learning architectures have outperformed average radiologist performance [4].

Therefore, the purpose of this research was to construct a novel convolutional neural network architecture (PulmXNet), the first unified 5-class convolutional neural network (CNN) for the diagnosis of normal, COVID-19, tuberculosis, and bacterial and viral pneumonia. A 5-class model for respiratory illness diagnosis is crucial for its potential clinical applicability as the overlapping symptomologies of these illnesses can create difficulties for accurate diagnosis.

II. Methods

The data for COVID-19 and normal CXRs was sourced from Kaggle databases, the pneumonia CXRs were from the Mendeley dataset, and the tuberculosis CXRs were from the National Institutes of Health National Library of Medicine database [5,6,7]. The training to testing data ratio was 3:1 and the data was sheared, rotated, and scaled to increase noise and prevent training bias (Table 1). To limit PulmXNet's overemphasis on average accuracy metrics, only 300 images were tested from the bacterial pneumonia class.

TABLE I. DISTRIBUTION OF THE DATA USED IN PULMXNET

Data Type	Disease Class				
	Normal	*COVID-19*	*Bacterial Pneumonia*	*Viral Pneumonia*	*Tuberculosis*
Train	1006	164	2472	1120	293
Test	225	55	300	373	101

The images were scaled to 250px by 250px to reduce training parameters. The architecture consisted of 3 convolutional layers and 128 nodes for the neural network layer. The rectifier function was used between these layers and softmax was used to predict the output. The 5-class PulmXNet was trained over 17 epochs and had 3,403,061 parameters.

$$f(s)_i = \frac{e^{s_i}}{\sum_j^C e^{s_j}} \quad (1)$$

$$CE = -\sum_i^C t_i \log(f(s)_i) \quad (2)$$

The equation for softmax (1) predicts the probability of each type of disease based on the artificial neural network layer. Categorical cross entropy loss was calculated as a measure to determine the model's improvement on a finer scale. Loss is a more reliable measure for training as it determines more than simply the right and wrong about the prediction but details the degree to which the prediction was off (2).

The accuracy for the model was calculated to determine the effectiveness of the model. The precision is the straight average of the diseases removing any dataset biases. The recall is the true positive rate. The F1 score is a summary statistic that factors in both precision and recall.

III. Results and Discussion

TABLE II. ACCURACY METRICS FOR 5-CLASS PULMXNET

Disease Class	*Accuracy*	*Precision*	*Recall*	*F1 Score*
Normal	85.59%	85.59%	95.96%	90.48%
COVID-19	85.45%	85.45%	90.39%	87.85%
Bacterial Pneumonia	89.33%	89.33%	83.49%	86.31%
Viral Pneumonia	85.52%	85.52%	87.16%	86.33%
Tuberculosis	84.16%	84.16%	74.56%	79.07%

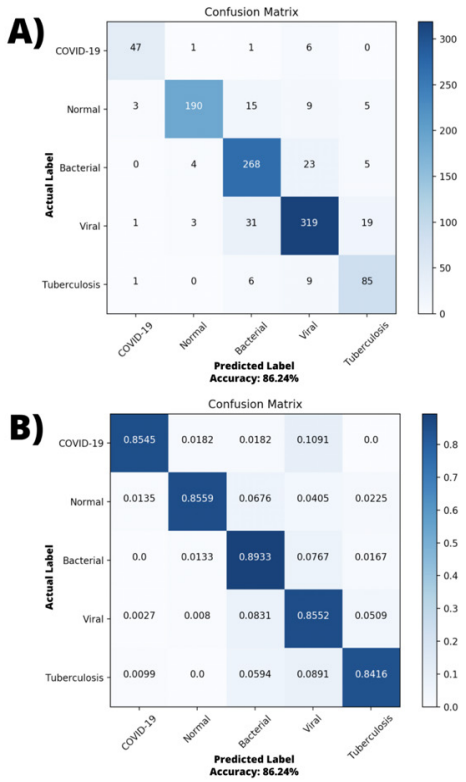


Figure 1. Confusion Matrices for PulmXNet. A) Non-normalized Confusion Matrix for 5-Class PulmXNet. B) Normalized Confusion Matrix for 5-Class PulmXNet.

The 5-class PulmXNet model distinguished between COVID-19, Normal, Bacterial, Viral, and Tuberculosis CXRs. Table 2 shows a summary of the performance metrics for the PulmXNet model. This model yielded an overall accuracy of 86.24% (Figure 1). The accuracy and precision values for normal, COVID-19, tuberculosis, and bacterial and viral pneumonia were 85.59%, 85.45%, 84.16%, 89.33%, and 85.52%, respectively (Table 2) . F1 scores were highest for the normal class (90.48%) and the lowest for the tuberculosis class (79.07%) (Table 2), which could be attributed to the low number of testing images compared to other classes.

This is the first proposed unified 5-class model for the classification of CXRs with the COVID-19 class. The PulmXNet model is also one of the first COVID-19 models to have a custom architecture without relying on pre-trained, generic convolutional layers, such as VGG-19 and ResNet. The model was manually trained to minimize the number of parameters for computational efficiency, while maximizing accuracy. With only 3,403,061 parameters, the 5-class PulmXNet model is far less than any of the current 4-class models which have a minimum of 9.4 million parameters [8]. In fact, PulmXNet has lower parameters for all 2-class and 3-class models, with the exception of the DarkNet model studied by Ozturk et. al [9].

IV. Conclusions

The widespread nature of x-ray imaging technology coupled with deep-learning models provides promise for cheaper radiological practices into the future. This research highlights the construction of the first unified CNN architecture that classifies CXRs for the following classes: healthy, COVID-19, tuberculosis, bacterial and viral pneumonia. By constructing PulmXNet's architecture from scratch, its use was optimized for chest x-ray diagnosis, unlike existing models in literature. Future investigations should optimize PulmXNet via Bayesian and random search optimization.

V. Acknowledgment

The authors of this paper would like to acknowledge the unconditional support from Dr. Serena McCalla and Dr. Jean-François Daneault. The authors would also like to thank the studies which compiled the data used in this experiment.

VI. References

[1] H. Li, S.-M. Liu, X.-H. Yu, S.-L. Tang, and C.-K. Tang, "Coronavirus disease 2019 (COVID-19): current status and future perspectives," International Journal of Antimicrobial Agents, vol. 55, no. 5, p. 105951, 2020.

[2] Y.-C. Liu, R.-L. Kuo, and S.-R. Shih, "COVID-19: The first documented coronavirus pandemic in history," Biomedical Journal, 2020.

[3] H. Chen, L. Ai, H. Lu, and H. Li, "Clinical and imaging features of COVID-19," Radiology of Infectious Diseases, Apr. 2020.

[4] A. I. Khan, J. L. Shah, and M. M. Bhat, "CoroNet: A deep neural network for detection and diagnosis of COVID-19 from chest x-ray images," Computer Methods and Programs in Biomedicine, vol. 196, p. 105581, 2020.

[5] M. E. H. Chowdhury, T. Rahman, A. Khandakar, R. Mazhar, M. A. Kadir, Z. B. Mahbub, K. R. Islam, M. S. Khan, A. Iqbal, N. Al-Emadi, M. B. I. Reaz, and M. T. Islam, "Can AI help in screening Viral and COVID-19 pneumonia?," IEEE Access, pp. 1–1, 2020.

[6] D. S. Kermany, M. Goldbaum, W. Cai, C. C. Valentim, H. Liang, S. L. Baxter, A. Mckeown, G. Yang, X. Wu, F. Yan, J. Dong, M. K. Prasadha, J. Pei, M. Y. Ting, J. Zhu, C. Li, S. Hewett, J. Dong, I. Ziyar, A. Shi, R. Zhang, L. Zheng, R. Hou, W. Shi, X. Fu, Y. Duan, V. A. Huu, C. Wen, E. D. Zhang, C. L. Zhang, O. Li, X. Wang, M. A. Singer, X. Sun, J. Xu, A. Tafreshi, M. A. Lewis, H. Xia, and K. Zhang, "Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning," Cell, vol. 172, no. 5, 2018.

[7] S. Jaeger, S. Candemir, S. Antani, Y.-X. J. Wáng, P.-X. Lu, and G. Thoma, "Two public chest X-ray datasets for computer-aided screening of pulmonary diseases," Quantitative Imaging in Medicine and Surgery, vol. 4, no. 6, pp. 475–477, Dec. 2014.

[8] Y. Oh, S. Park, and J. C. Ye, "Deep Learning COVID-19 Features on CXR using Limited Training Data Sets," IEEE Transactions on Medical Imaging, 2020.

[9] T. Ozturk, M. Talo, E. A. Yildirim, U. B. Baloglu, O. Yildirim, and U. R. Acharya, "Automated detection of COVID-19 cases using deep neural networks with X-ray images," Computers in Biology and Medicine, vol. 121, p. 103792, 2020.

Elucidating the Roles of Subtypes of B cells and Fibroblasts in the Immune Response of Rheumatoid Arthritis Joint Synovial Tissue

Barbara Timmerman[1] and Madhav Subramanian[2]

[1]Manhasset Secondary School, 200 Memorial Place, Manhasset, NY 11030

[2]Undergraduate student at Washington University of St. Louis, 1 Brookings Dr, St. Louis, MO 63130; Mentor through iResearch Institute, iResearch Foundation, P.O. Box 782, Glen Cove, NY 11542

***Abstract*- Rheumatoid Arthritis (RA) is an autoimmune disorder characterized by inflammation, while Osteoarthritis (OA) is developed through wear and tear on the body. B cells and Fibroblasts are activated by inflammation, however, the roles of subtypes of these cells in the progression of RA are unclear. Through single cell RNA sequencing on a dataset with RA and OA, the B4 subtype of B cells and the F2 subtype of fibroblasts were identified. The Immunoglobulin Heavy Constant Gamma 3 (IGHG3) gene, which codes for an immunoglobulin protein, was upregulated, suggesting its potential as a marker of RA. Through gene set enrichment analysis, upregulated pathways of the subtypes were identified. Immunoglobulin related pathways were upregulated in the F2 subtype suggesting that F2 cells are a crucial part of the immune response. Interferon related pathways were enriched in B4 cells suggesting that B4 cells regulate immunoglobulin class switching to Immunoglobulin G (IgG) molecules. The functions of the F2 and B4 subtypes are unique to RA and provide new information that is necessary to further elucidate the roles of B cells and Fibroblasts in RA.**

I. Introduction

Osteoarthritis (OA) and Rheumatoid Arthritis (RA) are two forms of arthritis. While OA occurs due to wear and tear on joints, RA is a chronic autoimmune disease characterized by inflammation and swelling of the synovial membrane in joints with eventual tissue and bone damage. Early detection and treatment are crucial for positive prognosis. [1]

Autoantibodies are produced by the body during RA against the body's antigens, which results in an aggressive immune response that destroys the cells, tissue, and eventually bones of the body. The inflammatory response is composed of immune cells, such as B cells and T cells, as well as fibroblasts and monocytes. [2] Currently researchers are focusing on key cells, such as B cells, T cells, Fibroblasts, and Monocytes that are involved in RA pathogenesis as potential targets. However, there are many subtypes for each of these cells, whose role in RA are unknown. The subtypes of B cells and fibroblasts were investigated using single cell RNA sequencing and gene set enrichment analysis to understand the dysfunctional immune response and the mechanisms that can be targeted to develop more efficient treatments and markers for RA. [1],[2]

II. Methods

Analysis of single cell RNA-sequencing was performed on a dataset of Rheumatoid Arthritis and Osteoarthritis, which included aligned and mapped data from the joint synovial tissue samples of n=30 human patients (20 RA and 10 OA). Seurat, dplry, ggplot2, and cowplot packages were downloaded into Rstudio to analyze and visualize the data. The data underwent quality control, reduction of dimensions, and then visualization of clusters with a Uniform Manifold Approximation and Projection (UMAP) graph. [3] The data was divided into four Seurat objects based on cell type: B cell, T cell, Fibroblasts, and Monocytes. The same pathway of analysis was used to identify clusters that represented the subpopulations of each cell type with a UMAP visualization. Each cluster was named to represent the cell type they were related to: B cell subtypes were referred to as B0-B5; Fibroblast subtypes were labeled F0-F5. Volcano plots and dot plots of the Seurat objects were also created in RStudio to identify differentially expressed genes in each group.

Gene Set Enrichment Analysis (GSEA) was utilized to define enriched pathways [3]. The first comparison was between RA and OA. The next two comparisons only included B cells and Fibroblasts, respectively, and was a comparison between RA and OA for each cell type. The last comparison was between a subpopulation and the rest of the subpopulations for the B4 subtype and the F2 subtype respectively. The c5 gene set database from the Molecular Signatures Database was used as the reference database for each GSEA. The GSEA software calculated statistical significance for each analysis and that data was used to determine which pathways were significantly unregulated. Enrichment analysis was considered statistically significant when the nominal p-value cutoff < 0.05 [3]. Enriched pathways from each of the GSEA results were combined into a bar graph with the normalized enriched scores (NES) to visualize the results.

III. Results and Discussion

Fibroblasts

The UMAP image shows the clusters of the subpopulations of fibroblasts split by condition (Fig 1A). F2 cells were only present in RA patients and made up 12.39% of the fibroblast population (Fig 1B). GSEA results indicated that immunoglobulin related pathways were enriched only in the F2 subtype (Fig 1C). F2 cells express IGHG1 and IGHG3 genes (Fig 1D).

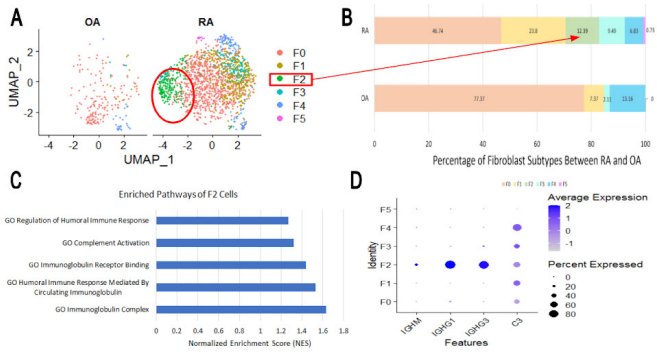


Figure 1: Differences of Fibroblast subtypes between OA and RA. **A.** The UMAP Visualization of the Fibroblast subtypes. **B.** The percentage of each subtype between conditions shows that F2 cells are unique to RA. **C.** Enriched Pathways in F2 Fibroblasts when compared to the rest of the fibroblast subtypes. **D.** Dotplot of the four core genes in Fibroblasts related to the GO Humoral Immune Response Pathway from C. (Figures by author)

The enriched complement activation pathway indicates that F2 fibroblasts directly promote inflammation through the creation of immunoglobulin immune complexes with fibroblasts (Fig 1C). F2 fibroblasts may be antigen presenting cells due to the gene expression levels of IGHG1 and IGHG3 (Fig 1D). [4]

B cells

The UMAP image shows the clusters of the subpopulations of B cells split by condition (Fig 2A). The bar graph shows that B4 cells were only present in the RA condition and made up 5.65% of the B cell population (Fig 2B). GSEA results indicate that interferon related pathways are enriched in B4 cells compared to the other B cell subpopulations (Fig 2C). B2 and B3 cells express IGHG2, IGHG3, and IGHG4 genes (Fig 2D).

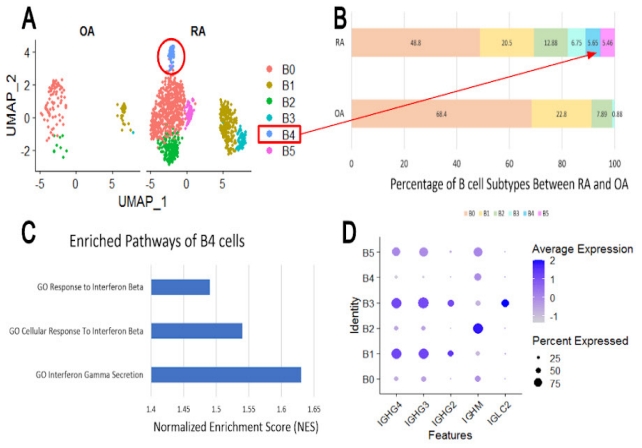


Figure 2: Differences of B cell subtypes between OA and RA. **A.** The UMAP Visualization of the B cell subtypes. **B**. The percentages of each subtype between conditions **C.** Enriched Pathways of B4 cells when compared to the rest of the B cell subtypes. **D.** Dot plot of the five core genes related to the Go Immunoglobulin Complex pathway. (Figures by author)

Interferon gamma secretion pathway was enriched in B4 cells. B cell subpopulations can produce IFN-γ in immune responses, such as the immune response in RA [5]. (Fig 2C) Interferon gamma (IFN-γ) has the ability to promote immunoglobulin class switching in B cells; IFN-y promotes IgG class switching to a more pathogenic subclass, IgG3, in RA [5]. This suggests that B4 cells are indicators of disease severity of RA and are essential to immunoglobulin class switching. Additionally, two pathways were enriched in B4 cells that indicated a role in responding to interferon beta. (Fig. 2C) Previous research has suggested that interferon betas have an anti-inflammatory role in the pathogenesis of RA [6].

IV. Conclusion

Immunoglobulin upregulated pathways and the presence of IGHG3 in both B cells and Fibroblasts indicates that the IGHG3 gene may be a potential biomarker for RA [7]. The F2 subtype provides evidence that there is a strong link between the immune response and fibroblasts through immunoglobulins [5]. The B4 subtype showed evidence of an important role in immunoglobulin class switching to igG and specific IgG1-4 molecules. This function has a direct link to disease severity as class switching of Igs indicates disease progression [1]. This relationship between B subtypes and immunoglobulins may uncover the possibility of dependent interactions between the functions of fibroblasts and B cells, and may lead to the development of a new treatment that can target these pathways of disease progression.

V. Acknowledgment

I would like to extend my gratitude to my mentor Madhav Subramanian and Dr. Serena McCalla of IResearch Institute, and my science research teacher Ms. Alison Huenger from the Manhasset Science Research Department.

VI. References

[1] H.Y. Yap, S.Z. Tee, M.M Wong, S.K. Chow, S.C. Peh, and S.Y. Teow., "Pathogenic role of immune cells in Rheumatoid Arthritis: Implications in clinical treatment and biomarker development," *Cells*, vol. 7, n.10, Oct. 9, 2018. [Online]. Available: https://doi.org/10.3390/cells7100161

[2] Zhang et al, "Defining inflammatory cell states in Rheumatoid Arthritis joint synovial tissues by integrating single-cell transcriptomics and mass cytometry," *Nature immunology*, vol. 20, n.7, pp. 928-942, May 6, 2019. [Online]. Available: https://doi.org/10.1038/s41590-019-0378-1

[3] Zhao et al, "Single-cell RNA sequencing reveals the heterogeneity of liver-resident immune cells in human," *Cell discovery*, vol. 6, n.22, April 28, 2020. [Online]. Available: https://doi.org/10.1038/s41421-020-0157-z

[4] A. Gierut, H. Perlman, and R.M. Pope, "Innate immunity and Rheumatoid Arthritis," *Rheumatic diseases clinics of North America*, vol. 36, n. 2, pp. 271–296, May, 2010. [Online]. Available: https://doi.org/10.1016/j.rdc.2010.03.004

[5] K.M. Pollar, D.M. Cauvi, C.B. Toomey, K.V. Morris, and D.H. Kono. "Interferon-γ and systemic autoimmunity", *Discovery medicine*, vol. 16, n. 87, pp. 123–131, Sept., 2013. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/23998448/

[6] T. Muskardin, and T.B. Niewold, "Type I interferon in rheumatic diseases", *Nature Reviews Rheumatology,* vol, 14, n. 4, pp. 214–228, Mar. 21, 2018. [Online]. Available: https://doi.org/10.1038/nrrheum.2018.31

[7] W.J. Falkenburg, D. van Schaardenburg, P. Ooijevaar-de Heer, G. Wolbink, and T. Rispens, "IgG Subclass Specificity Discriminates Restricted IgM Rheumatoid Factor Responses From More Mature Anti–Citrullinated Protein Antibody–Associated or Isotype-Switched IgA Responses," *Arthritis & Rheumatology,* vol. 67, n. 12, pp. 3124-3134, Aug. 5, 2015. [Online]. Available: https://doi.org/10.1002/art.39299

SafeBuild: The Risk-Based Utility Pole Design Software

Jennifer Lew[1] and Derek Fong[2]

[1] Palos Verdes Peninsula High School, 27118 Silver Spur Rd, Rolling Hills Estates, CA 90274
[2]California Public Utilities Commission, 505 Van Ness Ave, San Francisco, CA 94102

***Abstract*— On a yearly basis, California experiences fires, property damage, and prolonged outages due to the failure of electric utility infrastructure. In 2007, a Southern California Edison (SCE) pole broke, igniting the Malibu Fire. In 2011, in the San Gabriel Valley, 248 SCE poles broke, causing an outage to 440,000 customers for up to a week. In 2018, a component on a PG&E tower broke, igniting the *Camp Fire*, which destroyed 18,804 buildings and killed 86 civilians. These structural failures were due, in part, to the utility's failure to calculate the probability that its poles and towers could withstand known local wind speeds without breaking. Existing pole design software, SPIDAcalc and Osmose O-Calc, have major flaws. They provide inaccurate wind modeling analysis due to their failure to account for material strength variability. Additionally, the software calculates a strength factor using a Reference Wind Load, which is of little practical value as it cannot be compared against known local wind speeds. The current research corrects these errors by introducing a program, SafeBuild, that, given a wind gust, provides the probability that a structure can withstand the wind gust without breaking.**

I. INTRODUCTION

In California, failures of overhead electric distribution facilities have caused numerous wildfires, outages, and property losses. In 2007, an overloaded Southern California Edison (SCE) pole broke, igniting a fire that destroyed 14 structures. In 2011, in the San Gabriel Valley, 248 SCE poles broke, causing an outage to 440,000 customers, some for as long as a week. In 2018, a deteriorated component on a Pacific Gas & Electric Company tower broke, igniting the *Camp Fire*, which destroyed 18,804 buildings and killed 86 civilians. In these cases and others, all of which occurred during windy conditions, the cause of the incident was attributed to the utility's failure to design its facilities to withstand known local wind speeds. In the case of poles, the design software used by utilities (e.g., SPIDAcalc and Osmose O-Calc) have a major flaw: the software calculates the pole's safety factor by dividing the pole's bending moment resistance by the bending moment due to a Reference Wind Load. This method treats the safety factor as a strength factor, which underestimates the bending moment due to lateral loads, i.e. the P-Δ effect. In other words, the current method underestimates the true bending of the pole due to wind. Indeed, General Order 95 – the California statute governing the design of overhead facilities – requires the safety factor to be treated as a load factor (instead of a strength factor), which will properly account for the P-Δ effect. Furthermore, the method is misleading because the bending moment resistance of the pole is calculated using the median wood fiber strength; as a result, there is a 50% chance that the pole is weaker than the utilities assume. Finally, the method is misleading because safety factors obfuscate the actual windspeed that a pole should be able to withstand. For example, a very clear result such as "this wood pole can withstand a 112 mile per hour wind gust with a 95% probability of success" is not achievable via the existing software. Instead, the software provides that at some Reference Wind Load (such as 8 pounds per square foot), the pole has some strength factor (a value of 4, for example). This result has limited practical value and cannot be compared to known local wind speeds. The current research developed a structural design software, SafeBuild, that improves upon existing software: for any wind gust inputted by the user, SafeBuild calculates the probability that the pole can successfully withstand the wind gust based on the inherent material strength variability of wood. Note that, while SI units are the staple in most fields, SafeBuild uses US Customary Units, as this is the industry standard. SPIDAcalc and Osmose O-Calc also use US Customary Units.

II. METHODS

SafeBuild's methodology can be broken into several steps: (1) the engineer designs a utility pole using a spreadsheet; the design includes the pole class (from which the pole's groundline circumference C_G and pole top circumference C_{TOP} can be obtained) and wood species; the design also includes data for all supported conductors and equipment, e.g. a transformer; (2) the engineer enters a wind gust in miles per hour (mph) and then initiates the calculation; (3) SafeBuild converts the wind gust into a wind load, in pounds per square foot, and applies the wind load to all exposed surfaces; (4) using Finite Element Analysis, SafeBuild calculates the total bending moment BM_{TOTAL} at the pole's groundline due to the applied wind load; (5) SafeBuild calculates the probability that the pole can withstand the wind gust by determining the z-score z related to the bending moment resistance of the pole, taking into account the median wood fiber strength $F_{b,\ median}$ and its coefficient of variation CV (≈ 0.2 for wood).

While SafeBuild's Finite Element Analysis method is standard within the industry, its probability calculation is uniquely used. The calculation for finding z for the bending moment resistance of the pole is as follows:

$$BM_{TOTAL} = 0.000264\ (F_{b,\ median})(1 - z*CV)(C_G) \qquad (1)$$

The value of z is iterated until (1) becomes true. Once z is known, a table of z-scores can then provide the probability that the pole can withstand the wind gust.

III. EXPERIMENTAL DESIGN AND RESULTS

Using SafeBuild, a hypothetical utility pole was wind loaded in three trials. The pole had a length of 50 feet (ft), with 6 ft underground. The groundline circumference C_G and pole top circumference C_{TOP} was 35.72 inches and 21.59 inches, respectively, with a linear taper. The pole supported

the facilities shown in Table I. For Trial 1, the pole is made of Red Cedar ($F_{b,\ median}$ = 5800 lb_f/in^2) and the wind gust is assumed to be 112 mph. The results, shown in Fig. 1, show that the pole has an 8.69% chance of withstanding the wind gust. For Trial 2, the pole is made of Douglas Fir ($F_{b,\ median}$ = 8000 lb_f/in^2) and the wind gust remains 112 mph. The results, shown in Fig. 2, show that the pole has a 67.36% chance of withstanding the wind gust. For Trial 3, the pole is made of Douglas Fir and the wind gust is lowered to 92 mph. The results, shown in Fig. 3, show that the pole has a 97.13% chance of withstanding the wind gust.

TABLE I. CONDUCTORS AND EQUIPMENT SUPPORTED ON THE EXAMPLE POLE

Conductor Attachments						
Left Span Length[a]	***Right Span Length***	***Diameter***	***Height***[b]	***Weight or Density***	***No.***	***Type***
400 ft	300 ft	0.25 in	44 ft	0.2 lb/ft	3	Wire
400 ft	300 ft	0.50 in	33 ft	0.5 lb/ft	1	Wire
400 ft	300 ft	0.75 in	24.5 ft	1 lb/ft	1	TV cable
400 ft	300 ft	0.50 in	20.3 ft	0.5 lb/ft	1	Phone cable
Equipment Attachments						
0	0	0	37.5 ft	600 lb	1	XMFR

a. Span Length represents the length of the conductor span to the left or right of the pole.

b. Height represents the attachment height of the conductor or equipment on the pole.

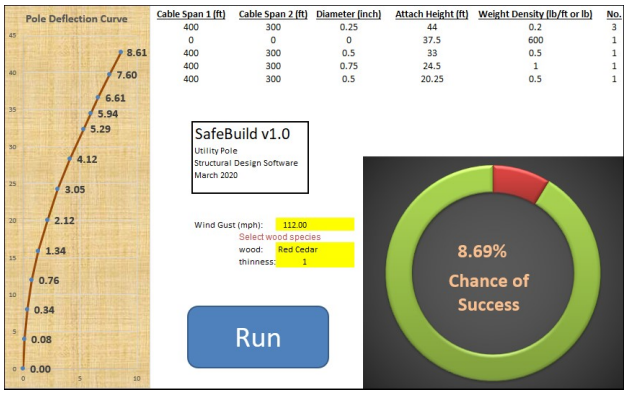


Figure 1. First trial: the pole has an 8.69% chance of not breaking.

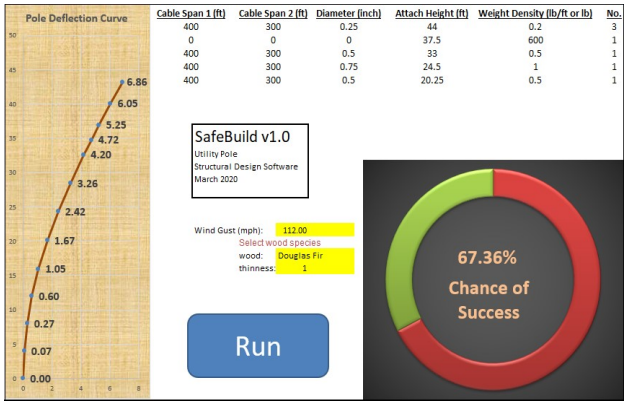


Figure 2. Second trial: the pole has a 67.36% chance of not breaking.

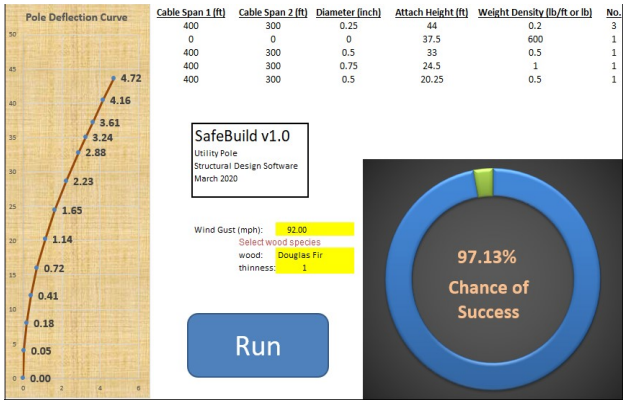


Figure 3. Third trial: the pole has an 97.13% chance of not breaking.

IV. DISCUSSION

Current pole design tools calculate a pole strength factor based on a Reference Wind Load (which is typically a blanket value applied indiscriminately). However, strength factors have limited practical value because they cannot be compared against known local wind speeds; instead, an engineer can only discern that a pole with a higher strength factor is stronger than a pole with a lower strength factor. SafeBuild takes a completely different approach. Instead of spitting out a pole strength factor, SafeBuild calculates the probability that a pole can withstand a given wind gust. This means that SafeBuild can be used to design an overhead electric system using a risk-based methodology. For example, an engineer can use SafeBuild to design any pole to have a 99% probability of withstanding the greatest 50-year wind gust in a given area. SafeBuild can prevent utility poles from being underbuilt, thereby enhancing public safety; at the same time, SafeBuild can prevent utility poles in low wind areas from being overbuilt, thereby saving ratepayer money. Therefore, it is worth further exploring SafeBuild.

ACKNOWLEDGEMENT

I would like to thank my mentor, Mr. Derek Fong, PE, Senior Utilities Engineer Supervisor at the California Public Utilities Commission, for his many detailed correspondences during this undertaking.

REFERENCES

[1] M. M. Alam, B.E. Tokgoz, and S. Hwang, "Framework for Measuring the Resilience of Utility Poles of an Electric Power Distribution Network", *International Journal of Disaster Risk Science*, vol. 10, 2019, pp. 270-281.

[2] C. Paio, T. Shupe, R.C. Tang, and C.Y. Hse, "Finite element analysis of wood laminated composite poles", *Wood and Fiber Science*, vol. 37, 2005, pp. 535-541.

[3] P. Eavis, I. Penn, California Says PG&E Power Lines Caused Camp Fire That Killed 85 [Internet]. The New York Times. [cited 2020 Aug 19]. Available from: https://www.nytimes.com/2019/05/15/business/pge-fire.html

[4] Edison taken to task in devastating 2011 San Gabriel Valley wind storm [Internet]. The Los Angeles Times. [cited 2020 Aug 19]. Available from: https://latimesblogs.latimes.com/lanow/2013/01/edison-taken-to-task-in-devastating-2011-wind-storm-.html

[5] N. Kim and B.V. Sankar, Introduction to Finite Element Analysis and Design. New York, NY: John Wiley & Sons, 2009.

The Effect of Ginkgo Biloba Extract on Beta-Amyloid Aggregation in C. elegans

Benjamin Punzalan[1], Virginia Flanagan[2], Alison Huenger[3]

[1, 2, 3] Manhasset Secondary School, 200 Memorial Place, Manhasset NY, 11030

***Abstract*— In 2019, the Alzheimer's Association estimated that approxima tely 5.6 million people in the United States alone are affected by Alzheimer's disease. Alzheimer's is a neurodegenerative disorder associated with the accumulation of beta-amyloid proteins, resulting in inflammation that disrupts synaptic functioning. The purpose of this experiment was to determine the extent to which *Ginkgo biloba* extract could be used to remediate beta-amyloid-induced paralysis in C. elegans. *Ginkgo biloba* is an antioxidant containing ginkgolides, which reduce inflammation by regulating cytokines. Additionally, *Ginkgo biloba* contains flavonoids, which reduce oxidative stress by limiting the buildup of free radicals, another characteristic of Alzheimer's. Transgenic strain CL2120 worms, which express beta-amyloid paralysis, were exposed to either 50µg/mL, 100µg/mL, 150µg/mL, or 0µg/mL (control) of *Ginkgo biloba* extract. After 48 hours, paralysis was determined by prodding individual worms with a platinum worm pick, with full body movement indicating the worm was not paralyzed, and no movement or only head movement indicating paralysis. Statistical analysis of the data using IBM SPSS v. 25 ANOVA followed by a Post Hoc Scheffe with $p<0.05$ showed that there was a statistically significant reduction in paralysis in worms treated with *Ginkgo biloba* extract when compared to the control. Overall, there is a correlation between the addition of *Ginkgo biloba* extract and a reduction in beta-amyloid-induced paralysis. This data may lend itself in the future to studies observing not only the use of *Ginkgo biloba* extract as a treatment for Alzheimer's symptoms, but also as a method of Alzheimer's prevention.**

I. Introduction

Alzheimer's disease (AD) is a progressive neurodegenerative disorder that causes cognitive impairment, including memory loss and confusion. On a cellular level, AD develops by the enzymatic activities of β-secretase and γ-secretase, rather than α-secretase and γ-secretase, on Amyloid Precursor Protein (APP). This amyloidogenic pathway leads to the formation of amyloid-beta (Aβ) peptides. Aβ aggregation causes inflammation, which can disrupt synaptic functioning, resulting in neurological damage and disruption of cell communication. (Bali et al, 2010). This describes the widely accepted amyloid cascade hypothesis that states that the aggregation of Aβ peptides results in neurotoxicity, senile plaque development, and Alzheimer's-related dementia [1].

Another common pathology of AD is oxidative stress, caused by Aβ aggregation. Oxidative stress is characterized by the formation of free radicals, namely reactive oxygen species (ROS), which are highly reactive molecules that lack a valence electron. Due to this, they are unstable, and oxidize many compounds in the brain, including DNA and proteins, causing significant neurodegeneration [2].

AD is also associated with the formation of neurofibrillary tangles (NFT) of tau proteins in neurons. Aβ releases kinases into neurons, which are enzymes that allow for the hyperphosphorylation of tau, which is when all phosphorylation sites are saturated. This causes the tau proteins to change structure, thus resulting in their loss of function, which is necessary for maintaining the structural integrity of microtubules in neurons [3].

Ginkgo biloba is an antioxidative chemical with many properties contributing to its ability to combat neurodegenerative disorders. One component of *Ginkgo biloba* extract is ginkgolides, which is a platelet-activating factor (PAF) receptor antagonist that is able to limit the binding of the platelet-activating factor (PAF) to the platelet activating factor receptor (PAFR), a G-protein-coupled-receptor (GPCR). Through the PAF signaling cascade, inflammation is caused by PAF binding to PAFR, which results in the synthesis of mediators such as leukotrienes, pro-inflammatory chemicals. By preventing PAF binding to PAFR, ginkgolides are able to reduce inflammation that is characteristic of AD [5]. The flavonoid component of *Ginkgo biloba* extract is also responsible for limiting the production of free radicals by donating electrons to ROS, inhibiting them from destabilizing compounds in the brain. Additionally, *Ginkgo biloba* extract is associated with the upregulation of protein phosphatase, which is able to reduce the hyperphosphorylation of tau [1].

II. Methods

C. elegans strain CL2120 (donated by the CGC) was separated into 4 groups, one control and three experimental groups (0, 50, 100, and 150 µg/mL of *Ginkgo biloba* extract). *C. elegans* were cultured in petri dishes containing nematode growth medium (NGM) agar and *E. coli* OP50 as a food source, and prior to conducting trials, all *C. elegans* were age synchronized through centrifugation [6].

150 mL of *Ginkgo biloba* extract were added to 3mL of agar in 35 mm petri dishes, with concentrations varying per experimental group. After 48 hours of exposure to the respective concentrations, the paralysis assay was

conducted. In *C. elegans,* paralysis is a common symptom of beta-amyloid aggregation, suggesting the presence of AD, and thus testing for paralysis indicates whether beta-amyloid plaques are present in the worms. The paralysis assay involved using a platinum wire pick to transfer individual *C. elegans* onto a new petri dish containing only NGM agar. Following this, paralysis was tested using a platinum wire worm pick, where individual worms were stimulated lightly on the head. Following stimulation, worms were determined to be paralyzed or not paralyzed, based on whether they exhibited any body movement. Three trials of the paralysis assay were performed, and 15 worms per group were tested per trial.

Following paralysis assay completion, averages were compiled per group in Microsoft Excel. Statistical analysis was performed using a One-Way Anova followed by a Post-Hoc Scheffe test with $p<0.05$ in IBM SPSS v. 25 to determine significance between groups.

III. Results and Discussion

After performing statistical analysis, it was determined that all three concentrations of *Ginkgo biloba* extract (50, 100, 150 μg/mL) significantly reduced beta-amyloid induced paralysis. 75% of the control group exhibited paralysis, compared to only 10% of the 50 μg/mL group, 3.33% of the 100 μg/mL group, and 20% of the 150 μg/mL group, as shown in Figure 1. While the three experimental groups were statistically significant from the control, they were not statistically significant from each other.

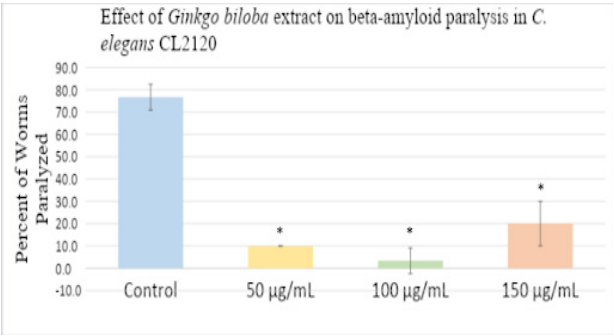


Figure 1: Percentage of CL2120 C. elegans paralyzed after 48 hours, n=15. Statistical significance determined by a One-Way ANOVA with a post-hoc Scheffe test ($p<0.05$). Error bars represent a 95% confidence interval. Significance is indicated by asterisks. (Graph by Author)

The probable causes for this significant reduction in paralysis are the ginkgolide and flavonoid components of the *Ginkgo biloba* extract. The ginkgolide component, which regulates platelet activation via limiting PAF binding to PAFR, allows for reduction of inflammation caused by the aggregation of platelets and the release of inflammatory chemicals. The limitation of beta-amyloid aggregation results in reduced paralysis, as evidenced by the results and data collected. Additionally, the flavonoid component is able to limit the production of ROS through its ability to donate a valence electron, thus negating the threat of ROS oxidizing easily-oxidizable compounds in the brain. This reduces neurodegeneration, and thus prevents paralysis from occurring in *C. elegans* [1].

IV. Conclusion

To conclude, the purpose of this experiment was to analyze the effects of *Ginkgo biloba* extract on reducing beta-amyloid paralysis in transgenic *C. elegans* strain CL2120. After performing the paralysis assay, it was determined that all experimental concentrations of *Ginkgo biloba* extract significantly reduced the percentage of *C. elegans* paralyzed. In addition, this study suggests that, pending further research, *Ginkgo biloba* extract, along with other similar antioxidative chemicals, may be applied in a human medical context as treatment for AD. In the future, *Ginkgo biloba* extract could be studied as a potential treatment contributing to the prevention of AD onset, rather than a remedy for AD symptoms. Additionally, further testing may show that *Ginkgo biloba* extract could be combined with other anti-inflammatory or antioxidative chemicals such as epigallocatechin gallate (EGCG) to test the effect of multiple treatments on AD symptom remediation.

V. Acknowledgments

I would like to thank Mrs. Flanagan, Mrs. Zhu, Ms. Cheryl, and Ms. Huenger for their support throughout the commencement of my research, along with the Manhasset Secondary School and the Manhasset Secondary School Science Research Program for allowing me to conduct my research.

VI. References

[1] Luo, Yuan, Yanjue Wu, Marishka Brown & Christopher Link (2009) Caenorhabditis elegans Model for Initial Screening and Mechanistic Evaluation of Potential New Drugs for Aging and Alzheimer's Disease, Methods of Behavioral Analysis in Neuroscience, 2nd Edition

[2] Shi, Chun, Jun Liu, Fengming Wu & David Yew (2010) Ginkgo biloba Extract in Alzheimer's Disease: From Action Mechanisms to Medical Practice, International Journal of Molecular Sciences, 11(1), 107-123

[3] Osmosis (2016) Alzheimer's disease - plaques, tangles, causes, symptoms and pathologies https://www.youtube.com/watch?v=v5gdH_Hydes, retrieved 10/4/19

[4] Link, Christopher (2019) *C. elegans* models of age-associated neurodegenerative diseases: Lessons from transgenic worm models of Alzheimer's disease, Experimental Gerontology, 41(10), 1007-1013

[5] Yost, Christian, Andrew Weyrich & Guy Zimmerman (2010) The Platelet Activating Factor (PAF) Signaling Cascade in Systemic Inflammatory Responses, Biochimie, 92(6), 692-697

[6] Sulston, J. & Hodgkin, J. (1988) Method. *In:* Wood, W.B. (ed.) The Nematode *Caenorhabditis elegans*. Cold Spring Harbor Laboratory, Cold Spring Harbor, 587-606

An Automated Census of Globular Cluster Systems in Virgo Cluster Dwarf Galaxies

Broderick Nies[1], Nicole Tian[2], Thendral Kamal[3], Eric Peng (PhD)[4i], Youkyung Ko (PhD)[4ii], Kaixiang Wang (PhD candidate)[4iii], and Puragra Guhathakurta (PhD)[5]

[1]Ralston Valley High School, 13355 W 80th Ave, Arvada, CO 80005, USA
[2]The Harker School, 500 Saratoga Ave., San Jose, CA 95129, USA
[3]Sharjah English School, Mleiha Road, Opposite Sharjah Planetarium, Sharjah, United Arab Emirates
[4i,4iii]Department of Astronomy, Peking University, Beijing, China
[4ii]Korea Astronomy and Space Science Institute, Daedeokdae-ro 776, Yuseong-gu Daejeon 34055, Republic of Korea
[5]Astronomy & Astrophysics Department, University of California Santa Cruz, 1156 High Street, Santa Cruz, CA 95064, USA

***Abstract*- This study outlines a method to detect globular cluster (GC) systems within Virgo cluster dwarf galaxies with the aim of creating a robust new census of GCs. Compact objects are selected in each image with the Source Extractor software package, and color-selection is performed to pick out possible GC candidates — objects within each dwarf galaxy that match the color and brightness profile of known GCs. Additionally, artificial GC sources are added to the image, and the percentage recovered by Source Extractor is used as a means to quantify the degree of completeness of their detection algorithm. Combined with an analysis of the GC background and the known shape of the GC luminosity function, this study obtains a statistical estimate of the total number of GCs associated with several dwarf galaxies. Further analysis of the new clusters discovered using this method could provide insight into the hierarchical formation of galaxies and the tidal events that modify them within the Virgo Cluster.**

I. Introduction

Residing over 65 million light years away from the Milky Way is the Virgo Cluster, the closest galaxy cluster to the Local Group. The Virgo Cluster is home to a particular abundance of dwarf galaxies, the most common being dwarf ellipticals, smaller galaxies with an ovular, rather than spiral or irregular, shape. In and around these dwarf galaxies are agglomerations of some of the oldest stars in the universe. These star clusters, known as globular clusters (GCs), are tightly bound by self-gravity and contain a range of 10,000 to 10 million stars. GCs have played key roles in the understanding of stellar evolution and dynamics and have served as fossil records of their host galaxies. One of the most significant impacts of GC research to date has been the accurate estimation of the minimum age of the universe from the study of GCs within the Milky Way [1]. By studying the metallicity, or amount of non-hydrogen elements in a star cluster, it is possible to determine the approximate age of a GC. It is also believed that galaxies are formed hierarchically, growing larger by accreting smaller star systems; thus, the study of GC systems could shed light onto the evolution of galaxies.

The dataset analyzed in this paper is part of the Next Generation Virgo Cluster Survey (NGVS), a 104 deg^2 optical survey of the Virgo Cluster in the *u*g'i'z'* bands using the MegaCam instrument on the Canada-France-Hawaii Telescope [4]. For the purposes of this project, the selection process relies on the *u*, g,* and *i* filters, or absorption colors of ultraviolet, blue light, and near infrared wavelengths, as well as making use of the point source depth of $g \approx 25.9$ mag to obtain a complete census of GCs of lower brightness than has been recorded so far. To develop the GC selection method, two images of Virgo dwarf galaxies that have been processed and calibrated by the NGVS collaboration are analyzed. Specifically, this process uses images in which the galaxy light is subtracted using smooth model images with elliptical symmetry and ring median filtering.

Our main work involved running Source Extractor (SExtractor), a light-source detection software that uses aperture photometry to record the brightness of each light source in an image, producing a catalog of several thousand sources of interest.

II. Methods

The process of GC selection begins by running SExtractor on the two model-subtracted dwarf galaxy images. The flux of each source detected by SExtractor was estimated using 4- and 8-pixel diameter apertures. As a result of SExtractor's aperture photometry, the light not captured by each aperture must be corrected for, as well as extinction, or dimming of light by interstellar dust; the average of these corrections was applied to all of the sources. In this process, the Source Extractor photometry data is compared to a pre-calibrated NGVS catalog with *some* of the sources in the dwarf galaxy image. With such a narrow field of view, it is safe to employ constant extinction and aperture corrections across the image; applying the average of these corrections across every source gives accurate photometric data for the dimmer GCs of interest.

Sources with a g-band magnitude brighter than 24.5 and a concentration index between $-0.1 < i_4 - i_8 < 0.15$ were selected, as these are likely objects with the same compactness that is typical of Virgo Cluster GCs. Based on a previous study of the Virgo Cluster [7] conducted by astronomy researchers using imaging from the Hubble Space Telescope (HST), the GC luminosity function peaks at around a g-magnitude of 24. Thus, a cut at 24.5 will detect roughly 70% of all GCs, as seen in Figure 1. On the other hand, this range in concentration index was determined empirically by looking at the distribution of concentration indices for known GCs. Exact point sources should have a concentration index of zero, but GCs are slightly extended and there is some scatter that must be accounted for. After plotting these compact sources on a color-color diagram, all sources that fit inside the color-space typical of GCs in M87 were taken to be GC candidates [5].

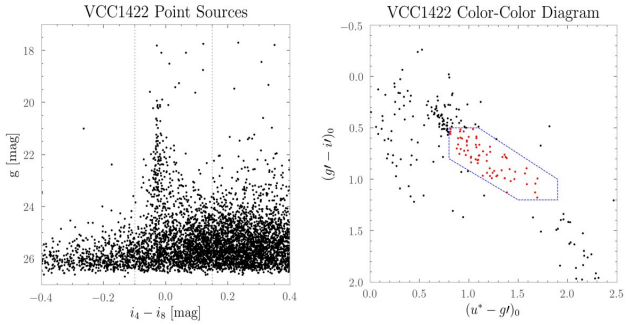


Figure 1 (left), g-band magnitude as a function of concentration index for the galaxy VCC 1422. The dashed lines show the range of values that we selected point source candidates from.
Figure 2 (right). Color-color diagram *(u*-g') vs. (g'-i')* of point sources in the dwarf galaxy VCC 1422. GC candidates are highlighted in red, and the blue polygon constrains the color-space for our GC candidates.

However, SExtractor was not able to detect all sources in an image, as some objects are obscured by image artifacts, bright foreground stars, etc. To correct for these 'missed' GCs, a checking procedure must be prepared to calculate a completeness ratio, or the percentage of sources in the image that are detected by SExtractor.

The first step in this procedure was to create a controlled number of artificial GCs. After light from a point source travels through interstellar dust and the earth's atmosphere, it reaches the telescope as a two-dimensional flux distribution, which can be modeled by a point spread function (PSF) [3]. Depending on the galaxy's location in the sky, a corresponding 31 by 31 pixel PSF is retrieved from the NGVS PSF server, which is added instead of a clean point source. Gaussian noise was also added to each pixel of the PSF to make them more realistic. Artificial GCs were positioned in a uniform square grid across the image, with their spacing and quantity set as adjustable parameters. In order to generate more data, four more output images are produced by shifting the PSF grid a half step to the right, a half step down, and a half step diagonally down and to the right for a fuller coverage of the pixels in the image. Our Python code iterated through all galaxies and color bands for magnitudes in the range 23.0 to 28.0 (four times each to accommodate the dithering of our GC grid). Then, the completeness ratio is computed by dividing the number of SExtractor-detected artificial stars by the total number of stars added.

Two details must be considered: the relationship between the completeness and the magnitude, and the position of the stars in the image. The completeness ratio was found to decrease as the magnitudes increased, since larger magnitudes equate to fainter images. This completeness testing is split into five distinct annuli based on their radius from the center, as shown in Figure 3. Finally, the completeness test is run in rings of range 10 arcsec, beginning from 0-10 arcsec and ending at 40-50 arcsec.

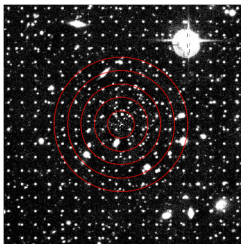


Figure 3. Galaxy VCC 940 has been divided into five annuli as shown by the red rings, ranging from a radius of 0-50'' from the galaxy center.

A sample of the completeness graph from 0-10 arcsec is shown in Figure 4, with each of the red crosses representing the completeness ratio at that specific magnitude. The resultant completeness ratios are fit to the Pritchet function algorithm [6]. A particularly useful characteristic of the Pritchet Function is the "g_limit." It returns the magnitude when source detection is at 50% accuracy, providing a uniform measure for completeness across the magnitude bins.

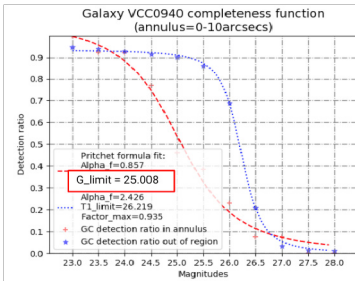


Figure 4. A completeness graph of the region of the image with a radius of 0-10'' from the center fit to the Pritchet Function. Completeness within the 10'' annulus is shown in red, and completeness outside 10'' is shown in blue.

III. Data Analysis

As seen in Figure 5, the g_limit generally increased as the artificial stars moved farther away from the center of the galaxy, indicating a higher completeness ratio for the outer annuli. This effect occurs because the patch of sky behind a light-subtracted galaxy still retains its original, higher noise floor. This leads to a low signal-to-noise ratio, crowding out GCs and making source detection more difficult. An increase in the magnitude corresponded to a lower completeness ratio and g_limit since larger magnitudes produce a fainter image, making source detection more difficult.

	0-10 arcseconds	10-20 arcseconds	20-30 arcseconds	30-40 arcseconds	40-50 arcseconds
g_limit	25.008	26.07	26.167	26.229	26.086

Table I. The g_limit, or g-band magnitude for which Source Extractor detects GCs with 50% accuracy, for each annulus in galaxy VCC 940.

The detection process was refined on galaxies VCC 940 and VCC 1422, selected due to good light subtraction quality, clear elliptical shape and the presence of confirmed GCs. Photometric analysis was also run on seven other galaxies of varying light subtraction quality. In the case of VCC 940, SExtractor detected 3,055 objects, 41 of which passed the criteria for concentration index, magnitude, and color. Out of these, 22 GC candidates were contained visually within the core of the galaxy with a radius of ~50''. Accounting for the Globular Cluster Luminosity Function, correcting for completeness, and calculating a calculated GC background of ~0.612 GCs per arcmin^2 gives approximately 47 GCs in VCC 940 – a slightly larger number than was found by our research group using Hubble Space Telescope images without accounting for completion or implementing model subtraction.

IV. Results and Further Implications

In the future, we hope to run this method on the full dataset of 1,145 galaxies, and we are currently constructing a computerized pipeline to automatically detect the total number of GCs within each dwarf galaxy's gravitational

influence. To enhance the accuracy of our data, we also plan to develop a program to determine the completeness of the color selection process by placing artificial GCs with brightnesses set in *multiple* filter bands in each image and feeding them through the GC color selection program.

By comparing the colors of our GCs across galaxies of varying surface brightness, size, and composition, we believe that this GC catalog will be useful to pin down correlations between galaxy GC properties. By analyzing this data across the entire Virgo Cluster, it will be possible to draw conclusions about the formation of galaxies through the merging of globular clusters with a higher resolution than previously possible.

V. Acknowledgments

Broderick Nies, Nicole Tian, and Thendral Kamal conducted their research under the auspices of the Science Internship Program at the University of California in Santa Cruz. All members wish to acknowledge the very significant cultural role and reverence that the summit of Mauna Kea has always had within the indigenous Hawaiian community. We are most fortunate to have the opportunity to work with observations from this mountain.

VI. References

[1] Ashman, Keith M., and Zepf, Stephen E. “Globular Cluster Systems” (Cambridge Astrophysics) *Cambridge University Press* (1998)

[2] Bertin, Emmanuel, and Stephane Arnouts. "SExtractor: Software for source extraction." *Astronomy and Astrophysics Supplement Series* 117.2 (1996): 393-404.

[3] De Gennaro, Geoff G. A postprocessing method for spatially variant point spread function compensation. *The University of Utah* (2010) [4] Ferrarese, Laura, et al. "The next generation Virgo Cluster survey (NGVS). I. Introduction to the survey." *The Astrophysical Journal Supplement Series* 200.1 (2012): 4.

[5] Lim, Sungsoon, et al. "Globular Clusters as Tracers of Fine Structure in the Dramatic Shell Galaxy NGC 474." *The Astrophysical Journal* 835.2 (2017): 123.

[6] McLaughlin, Dean E., et al. "Washington photometry of the globular cluster system around NGC 3311. 2: Spatial structure and mass spectrum." *The Astronomical Journal* 109 (1995): 1033-1043.

[7] Peng, Eric W., et al. "The color-magnitude relation for metal-poor globular clusters in M87: confirmation from deep HST/ACS imaging." *The Astrophysical Journal* 703.1 (2009): 42.

CJSJ

Investigating the therapeutic potential of capsaicin and curcumin: a comparative study on neuroblastoma and hypothalamic cells

Lucy Zha[1], Grace Lawson[2], Rolf Karlstrom Ph.D.[3], and Wei Zhu Ph.D.[4]

[1]The Wheatley School, 11 Bacon Rd, Old Westbury, NY 11568
[2,3]University of Massachusetts Amherst, University of Massachusetts Amherst, MA 01003
[4]SUNY College at Old Westbury, 223 Store Hill Road Campus Center H-310, Old Westbury, NY 11568

***Abstract*—Neuroblastoma is one of the most common malignant pediatric tumors. However, current cancer therapy usually involves surgical removal or chemotherapy, which are both damaging to normal cells. Phytochemicals are proposed as cancer treatment alternatives. Curcumin and capsaicin, two polyphenolic compounds from turmeric and pepper, possess anti-inflammatory and antioxidant properties, which suggest their neuroprotective benefits. This study investigates anti-cancer effects of curcumin and capsaicin on neuroblastoma cells in vitro while also probing potential effects of curcumin on hypothalamic neurogenesis in vivo. The experiments demonstrated that curcumin and capsaicin synergistically induced cell death, inhibited metastasis through decreasing MMP9 levels, and disrupted tumor growth of neuroblastoma cells. Both compounds also induced cell death of hypothalamic cells in vitro, and curcumin inhibited hypothalamic neurogenesis, possibly through the Notch signaling pathway. These results illustrate that while curcumin and capsaicin are effective at treating neuroblastoma cells, their toxicity on hypothalamic cells in vitro and potential impacts on neurogenesis in vivo must also be considered in conjunction with their anti-cancer effects in future research.**

I. INTRODUCTION

Neuroblastoma, a malignant pediatric tumor in the sympathetic nervous system, is the most common cancer in infants in their first year[1]. The 5-year survival rate is only 40% for children with stage-4 metastatic neuroblastoma [2]. Currently, the most predominant cancer treatment involves surgical or radiation-based tumor removal and chemotherapy, which can have serious toxic effects on non-targeted tissues. Hence, scientists started exploring phytochemicals' anti-cancer potential, which are generally non-toxic, inexpensive, and can be found in common household spices.

Curcumin, a polyphenolic compound extracted from turmeric, and capsaicin, a primary capsinoid extracted from peppers of the Capscum genus, are known for their antioxidant and anti-inflammatory properties. On normal stem cells, curcumin can increase hippocampal neurogenesis in adult mice [3]. In colorectal cancer cells, curcumin decreased the survival rate of the cancer cells and inhibited metastasis through a downregulation of the matrix metalloproteinases gene group (MMP) [4]. Capsaicin acts as a cancer suppressor or promoter in different cancer cell types [5]. However, the combined effect of curcumin and capsaicin on neuroblastoma cells remains unknown.

II. METHODS

MTT cell viability test was done on human neuroblastoma cells SK-N-SH and rat hypothalamus cell line R9 to detect reduction of MTT by mitochondrial dehydrogenase to a yellow formazan product after 24 hours of treatment. Cell viability was calculated as follows:

Cell viability (%) = experimental OD value/control OD value × 100.

The neuroblastoma cells' metastatic ability after treatment was analyzed through a sandwich enzyme-linked immunosorbent assay (ELISA) that detects the level of matrix metallopeptidase (MMP) -9 and a scratch-healing assay. Wound healing was assessed after 24 h. The percent decrease in migration was calculated as follows:

Decrease in migration (%) = 1 − migration distance (pixels)/control migration distance (pixels) x100

Neuroblastoma cell colony formation was also examined by measuring the percent area decrease of surviving colonies using ImageJ.

To investigate hypothalamic neurogenesis, transgenic larval zebrafish Tg(Tp1bglob:GFP), which expresses Notch-responsive cells, at 5 days post fertilization (dpf) were treated with curcumin. EdU immunohistochemical staining incorporates 5-ethynyl-2'-deoxyuri-dine (EdU) into the cells during the S phase of the cell cycle to express proliferating cells. The brains were dissected and imaged for the fluorescent EdU+ and TP1+ cells. Only cells in the posterior recess were counted.

III. RESULTS AND DISCUSSION

Curcumin and capsaicin decreases neuroblastoma and hypothalamic cell viability

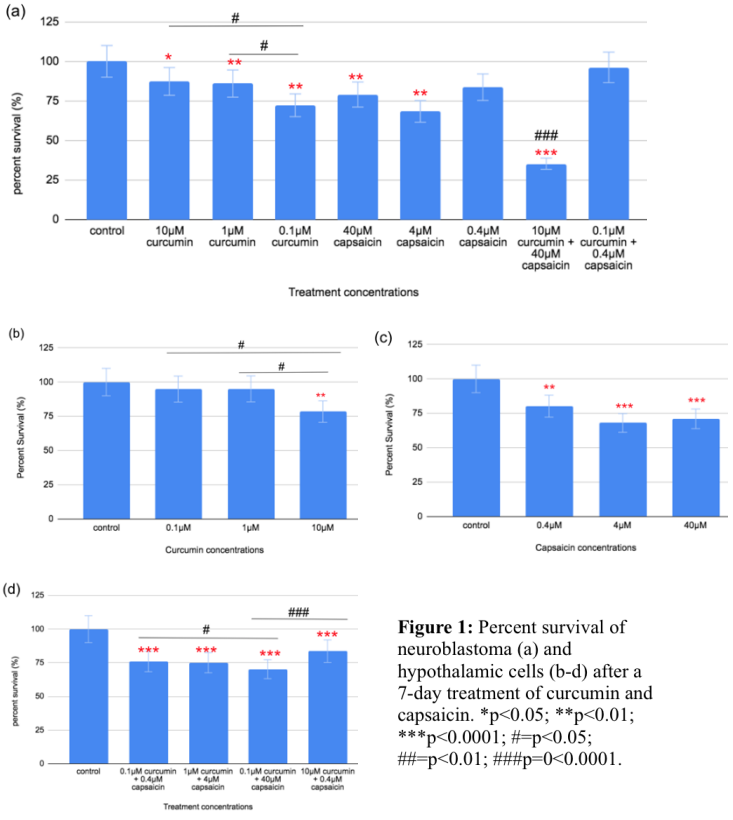
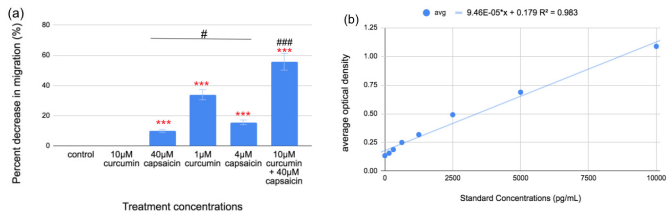


Figure 1: Percent survival of neuroblastoma (a) and hypothalamic cells (b-d) after a 7-day treatment of curcumin and capsaicin. *p<0.05; **p<0.01; ***p<0.0001; #=p<0.05; ##=p<0.01; ###p=0<0.0001.

Curcumin and capsaicin decreases neuroblastoma cell migration and MMP9 expression



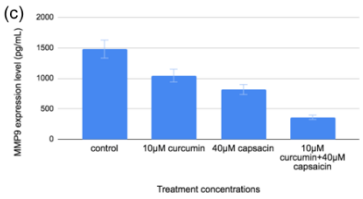


Figure 2: Effects of curcumin and capsaicin on neuroblastoma cell migration (a-b) and ELISA MMP9 level (c). *p<0.05; ***p<0.0001; #=p<0.05; ##=p<0.01; ###p=0<0.0001.

Curcumin and capsaicin decreases neuroblastoma colony formation

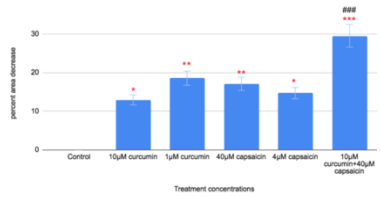


Figure 3: Effects of curcumin and capsaicin on the area of colonies. *p<0.05; **p<0.01; ***p<0.0001; #=p<0.05; ##=p<0.01; ###p=0<0.0001.

In this study, both curcumin and capsaicin decreased neuroblastoma cell viability and inhibited neuroblastoma cell metastasis, the invasion of cancer cells to neighboring cells which increases the tumor's risk level. An overexpression of MMP9, a protein involved in tumor invasion, metastasis and angiogenesis, has been examined in multiple malignant tumors such lung, pancreatic, breast cancers, etc. [7]. The results from the ELISA assay shows that curcumin and capsaicin reduced MMP9 expression level, indicating both chemicals inhibit metastasis through suppressing MMP9 expression. Curcumin and capsaicin also disrupted the tumor formation by reducing the area of neuroblastoma colonies. Normal cells accumulate in masses until they reach a finite density; however, cancer cells proliferate abnormally, resulting in high-density cell masses [8]. Prior studies of curcumin and capsaicin on other cancer cells are consistent with this study's findings that curcumin and capsaicin can interfere with tumor formation by disrupting the cell proliferation cycle, inhibit metastasis through regulating MMP9 level, and induce apoptosis [9] [10]. The most prominent anti-cancer effect occurred at the curcumin-capsaicin combination at 10μM and 40μM at all assays on neuroblastoma cells, indicating the two chemicals can synergistically suppress neuroblastoma growth (Fig 1, 2, 3).

Curcumin decreases hypothalamic proliferating cells and Notch responsive cells

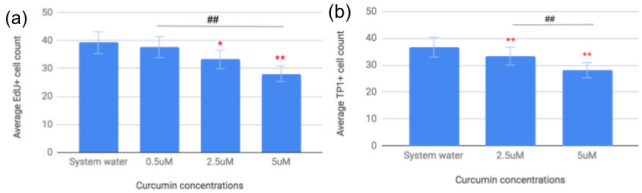


Figure 4: Effects of curcumin on (a) EdU+ (proliferating) cells and (b) TP1+ (Notch-responsive) cells. Boxed area indicates the posterior recess of the hypothalamus. *p<0.10; **p<0.05; ##p<0.05.

Curcumin and capsaicin decreased viability of hypothalamic cells, and that curcumin decreased cell proliferation in the postembryonic hypothalamus in zebrafish and reduced the number of Notch-signaling-responsive neural precursors in the hypothalamic stem cell niche (Fig 4). Notch is one of the three cell signaling pathways in zebrafish that regulates the neural stem cell cycle by inhibiting the differentiation of neural progenitor cells from oligodendrocyte precursors to oligodendrocytes and promoting the differentiation of glial precursors to astrocytes (Fig 5) [11]. According to a study done by Li et. al, blocked Notch signaling by DAPT, a γ-secretase inhibitor that indirectly inhibits Notch, leads to a significant decrease of proliferative cells from the control and the suppression of cell proliferation effects of curcumin in mice's hippocampus [12]. If considered together with Li et al's study, the results from this research may suggest that curcumin can block Notch signaling, which consequently triggers a positive feedback loop by decreasing proliferative cells and further inhibition of Notch signaling from curcumin.

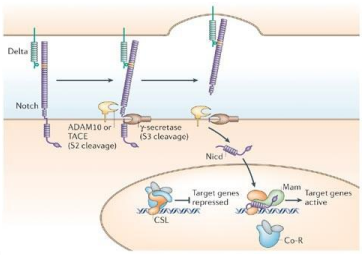


Figure 5: An Overview of Notch Signaling Pathway

IV. CONCLUSION

While curcumin and capsaicin can be effective at treating neuroblastoma cells, they can decrease hypothalamic cell viability and postembryonic neurogenesis through the Notch signaling pathway, both of which are crucial for maintaining the body's homeostasis. By comparing the results of curcumin and capsaicin on neuroblastoma and hypothalamic cells holistically, this study provides insights on the clinical feasibility of curcumin and capsaicin on neuroblastoma treatment.

V. ACKNOWLEDGEMENT

I would like to thank the Summer College Program at the University of Massachusetts Amherst and Dr. Wei Zhu of SUNY Old Westbury for providing me the opportunity to conduct this research.

VI. REFERENCES

[1]Tonini, G. Growth, progression and chromosome instability of Neuroblastoma: a new scenario of tumorigenesis? BMC *Cancer* 17, 20 (2017).
[2]Neuroblastoma stages and prognostic markers. (2018, March). Cancer.
[3]Hewlings, S. J., & Kalman, D. S. (2017). Curcumin: A Review of Its Effects on Human Health. *Foods* (Basel, Switzerland), 6(10), 92.
[4]Kim, S. J., Son, T. G., & Park, H. R. (2008). Curcumin stimulates proliferation of embryonic neural progenitor cells and neurogenesis in the adult hippocampus. *Journal of Biological Chemistry*, 283.
[5]Klinger, N. V., & Mittal, S. (2016). Therapeutic potential of curcumin for the treatment of brain tumors. *Oxidative Medicine and Cellular Longevity*.
[6]Yang, J., Li, T. Z., Xu, G. H., Luo, B. B., Chen, Y. X., & Zhang, T. (2013). Low-concentration capsaicin promotes colorectal cancer metastasis by triggering ROS production and modulating Akt/mTOR and STAT-3 pathways. *Neoplasma*, 60(4), 364–372.
[7]Huang H. (2018). Matrix Metalloproteinase-9 (MMP-9) as a Cancer Biomarker and MMP-9 Biosensors: Recent Advances. *Sensors* (Basel, Switzerland), 18(10), 3249.
[8]Cooper GM. The Cell: A Molecular Approach. 2nd edition. Sunderland (MA): Sinauer Associates; 2000. The Development and Causes of Cancer.
[9]Clark, R., & Lee, S. H. (2016). Anticancer Properties of Capsaicin Against Human Cancer. *Anticancer research*, 36(3), 837–843.
[10]Wang, L., Chen, X., Du, Z., Li, G., Chen, M., Chen, X., Liang, G., & Chen, T. (2017). Curcumin suppresses gastric tumor cell growth via ROS-mediated DNA polymerase γ depletion disrupting cellular bioenergetics. *Journal of experimental & clinical cancer research*: CR, 36(1), 47.
[11] Bray, S. Notch signalling: a simple pathway becomes complex. *Nat Rev Mol Cell Biol* **7,** 678–689 (2006). https://doi.org/10.1038/nrm2009
[12]Li, J., Han, Y., & Li, M. (2019). Curcumin promotes proliferation of adult neural stem cells and the birth of neurons in Alzheimer's disease mice via notch signaling pathway. *Cellular Reprogramming*, 21(3).

The Remedial Effect of *Mucuna pruriens* extract on Copper-induced Parkinsonian Behaviors in *Daphnia magna*

Ericka Lai[1], Alison Huenger[2], and Virginia Flanagan[2]

[1, 2i, 2ii] Manhasset High School, 200 Memorial Place, Manhasset, NY 11030

***Abstract*—In 2018, the Parkinson's Foundation estimated healthcare for Parkinson's Disease (PD) (treatment, social security payments, lost income) to be a $52 billion industry. PD is a neurodegenerative disorder that induces difficulty in movement and reduced motor skills. The aim of this study was to determine the effects of exposure to *Mucuna pruriens* extract (MPE) on reduced locomotion as a result of copper-induced Parkinsonism in *Daphnia magna*. Further study of *M. pruriens* on *D. magna* served to validate its efficacy on multiple model organisms. Parkinsonian behavior was induced by exposing *D. magna* to copper sulfate (15.9 mg/L) for 96 hours. It was hypothesized that exposure to MPE would increase locomotion in *D. magna* that had previously exhibited Parkinsonian behaviors. *D. magna* were treated with MPE (20, 40 µg/mL) for 96 h. Their movement was then recorded microscopically for 1 min after each exposure. Movement was quantified with the use of wrmTrck, an ImageJ plugin. The distance each daphnid traveled and the frequency at which they spun in terms of bends was measured. Analysis of the data using IBM SPSS v. 25 ANOVA followed by a post hoc Scheffe ($p < .05$) showed that both groups treated with MPE significantly increased movement from the group treated with only copper sulfate. Frequency of spinning significantly decreased in groups exposed to MPE compared to daphnids only exposed to $CuSO_4$. It was concluded that *M. pruriens* shows efficacy as a treatment for impaired behaviors found in Parkinson's disease in *D. magna*.**

I. Introduction

It is estimated that 930,000 Americans will be diagnosed with Parkinson's disease (PD) by 2020. [1] This number is predicted to further increase to 1.2 million in 2030. PD is a neurodegenerative disorder that develops when the neurons of the substantia nigra in the brain die or become damaged, resulting in the loss of ability to produce dopamine. [2] The most widely prescribed medication for PD is levodopa (L-DOPA) and is converted by the brain into dopamine. [3] However, it causes unwanted effects, such as nausea and vomiting, and even dyskinesia or motor fluctuations. An alternative to L-DOPA that does not have these consequences would be beneficial.

Mucuna pruriens (MP), also known as the velvet bean, is a subtropical to tropical legume that has been used to treat neurodegeneration, like Parkinsonism and tremors. Its seeds are a source of L-DOPA, but previous studies have presented that pure *M. pruriens* is a safe alternative to L-DOPA. [4,5] To confirm the efficacy of *M. pruriens* as a possible remedy for PD across multiple model organisms, *D. magna* were utilized for their sensitivity to changes in water chemistry, as well as them being cost-effective organisms.

wrmTrck is an ImageJ plugin used for C. elegans motility analysis. [6] However, it can also be used on other experimental organisms, including *D. magna*. A dopamine receptor signaling pathway, mediated by putative D2-like receptors, was found to be involved in the control of *Daphnia* swimming behavior. [7] Using wrmTrck, the movement of *D. magna* was monitored in terms of distance traveled and number of body bends, in which their circular movement was quantified. Improved movement in *D. magna* would suggest the efficacy of *M. pruriens* as treatment for copper-induced Parkinsonism.

II. Materials and Methods

Copper sulfate (CuSO4) was used to induce *D. magna* with Parkinsonism. *D. magna* were obtained from Carolina Biological. A concentration of 15.9 mg Cu/L of copper sulfate was used to induce each experimental group of Daphnia magna with Parkinsonian-like symptoms. [8] Each group of *D. magna* was exposed to 500 µL of copper sulfate in a 100 mL container for development of PD for 96 h. *M. pruriens* in its form of *M. pruriens extract* (MPE) was used to remediate the symptoms of PD from Daphnia magna. 20 ug/mL and 40 ug/mL concentrations of MPE were utilized in this study. [9] Each experimental copper sulfate-exposed group was exposed to 120 mL water with dissolved pure MPE from BulkSupplements. All experimental groups were exposed to MPE for 96 h in a controlled environment, and examined after exposure.

Figure 1. D. magna exposed to different concentrations of copper sulfate and/or MPE were observed under a microscope. (Photo by Author)

Daphnia movement was examined through microscopy with the Motic SMZ-171, an industrial stereo microscope. *D. magna* were placed into a small petri dish filled halfway with water. The petri dish was then placed on the stage of the microscope, where the light source illuminated the specimens. Using an iPhone XS, the movement of daphnids per group was recorded for 1 min. After a video was recorded, wrmTrck was used to quantify the data. First, a video of daphnid motility was opened in ImageJ. The background was then subtracted using the rolling ball method in order to remove the uneven background, and the "Rolling ball radius" was adjusted. Next, the movie was converted to binary, where adjusting the threshold indicated what pixels

were daphnids and what pixels were background. Afterwards, ImageJ was informed of the real-world scale of the animals in the movie. The "wrmTrck" plugin was then be run, and factors such as minimum (minSize) and maximum (maxSize) worm area in pixels, maxVelocity, "maxAreaChange," "minTrackLength," "FPS," and "bendThreshold," were adjusted. A Results window popped up, containing the tracking analysis of all detected specimens. Movement was quantified based on distance traveled in a span of 1 min by each daphnid, and the frequency of spins in terms of bends.

Data was analyzed using IBM SPSS v. 25 One-way ANOVA followed by a post hoc-Scheffe test, where $p<.05$ and was graphed using Excel.

III. Results and Discussion

The purpose of this study was to investigate the effects of MPE on movement in *D. magna* as a Parkinson's model. It was hypothesized that *M. pruriens* would be effective in treating Parkinson's symptoms induced by copper sulfate in *D. magna*.

	Average Distance Traveled by *D. magna* (mm)			
	Control	**$CuSO_4$**	**MPE (20 ug/mL)**	**MPE (40 ug/mL)**
Distance	18.4 ± 2.40	4.63 ± 2.82	16.5 ± 4.34	18.5 ± 9.59

Figure 2. D. magna exposed to MPE exhibited significantly greater movement than daphnids exposed solely to copper sulfate. (Table by Author)

Statistical analysis using IBM SPSS one-way ANOVA followed by a post-hoc Scheffe test where $p<0.05$ showed that both groups treated with MPE significantly increased movement from the group treated with only copper sulfate. The groups exposed to MPE demonstrated similar movement to that of the untreated control. The frequency of spinning was greatest with exposure to only $CuSO_4$, and decreased from exposure to MPE. However, spinning increased as the concentration of MPE increased, which could have been due to a higher amount of natural levodopa. These results suggest similar findings to the climbing assay in Johnson et al (2018), where *D. melanogaster* treated with MPE exhibited significantly increased climbing distance compared to flies exposed only to toxins. In the current study, daphnids treated with MPE after $CuSO_4$ exposure exhibited a 256% (20 ug/mL) and a 300% increase in movement compared to those exposed solely to $CuSO_4$. Thus, the alternate hypothesis was supported.

IV. Conclusion

Significant improvement in movement was exhibited by a novel model organism, *D. magna*, after MPE treatment. With exposure to increased concentrations of MPE (20, 40 ug/mL), daphnids demonstrated the ability to travel greater distances, signifying remediation of PD-like symptoms. However, decreased spinning exhibited at the lower MPE concentration indicated that higher concentrations of MPE exhibited higher concentrations of L-DOPA, which may have resulted in dyskinesia-like symptoms. The overall success of MPE in remediating PD-like symptoms in *D. magna* suggests that it may exhibit similar success in other model organisms for which MPE treatment has not been studied, and more importantly, humans. A cheaper alternative to the current treatment process of PD that could be provided by MPE would enable a greater population of PD patients to have access to treatment. *M. pruriens* is a legume, which can be easily grown in a variety of climates and regions. Its use in the future as a treatment for PD would be to act as a natural, cost-effective alternative to current treatments. Due to the COVID-19 pandemic, full data collection could not be completed. Future study of *M. pruriens* on PD model organisms would provide further support for its efficacy as an alternative treatment for PD.

Acknowledgment

I would like to thank my mentors, Mrs. Flanagan and Ms. Huenger, for giving me the opportunity to participate in this research.

References

[1] Parkinson's Foundation (2018), Parkinson's Foundation, retrieved 9/18/19, https://www.parkinson.org/Understanding-Parkinsons/Statistics

[2] Michael J. Fox Foundation (2019) Dyskinesia, retrieved 9/25/19, https://www.michaeljfox.org/news/dyskinesia

[3] National Institute of Aging (2017), National Institute of Aging, retrieved 9/18/19, https://www.nia.nih.gov/health/parkinsons-disease

[4] Cilia, Robert, Janeth Laguna, Erica Cassani, Emanuele Cereda, Nicolò G. Pozzi, Ioannis U. Isaias, Manuela Contin, Michela Barichella, and Gianni Pezzoli (2017) *Mucuna pruriens* in Parkinson disease: A double-blind, randomized, controlled, crossover study, Neurology, Volume 89, Issue 5, P. 432-438

[5] Katzenschlager, R., A. Evans, A. Manson, P.N. Patsalos, N. Ratnaraj, H. Watt, L. Timmermann, R. Van der Giessen, and A.J. Lees (2004) Mucuna pruriens in Parkinson's disease: a double blind clinical and pharmacological study, Journal of Neurology, Neurosurgery, and Psychiatry, Volume 75, Issue 12, P. 1672-1677

[6] Nussbaum-Krammer, Carmen I., Mário F. Neto, Renée M. Brielmann, Jesper S. Pedersen, and Richard I. Morimoto (2015) Investigating the Spreading and Toxicity of Prion-like Proteins Using the Metazoan Model Organism *C. elegans*, JoVe Journal, 10.3791/52321

[7] Barrozo, Enrico R., David A. Fowler, and Matthew L. Beckman (2015) Exposure to D2-like dopamine receptor agonists inhibits swimming in Daphnia magna, Volume 137, P. 101-109

[8] Mashock, Michael J., Tyler Zanon, Anthony D. Kappell, Lisa N. Petrella, Erik C. Andersen, and Krassimira R. Hristova (2016) Copper Oxide Nanoparticles Impact Several Toxicological Endpoints and Cause Neurodegeneration in *Caenorhabditis elegans*, PloS ONE, Volume 11, Issue 12, P. e0167613

[9] Johnson, Shelby L., Hyun Y. Park, Nicholas A. DaSilva, Dhiraj A. Vattem, Hang Ma, and Navindra P. Seeram (2018) Levodopa- Reduced *Mucuna pruriens* Seed Extract Shows Neuroprotective Effects against Parkinson's Disease in Murine Microglia and Human Neuroblastoma Cells, *Caenorhabditis elegans*, and *Drosophila melanogaster*, Nutrients, Volume 10, Issue 9, P. 1139-1153

Interpretable Convolutional Neural Networks for Building Damage Assessment in Satellite Imagery

Thomas Y. Chen[1] and Ethan Weber[2]

[1]Academy for Mathematics, Science, and Engineering, Morris Hills High School, 520 W Main St, Rockaway, NJ 07866
[2]Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, 32 Vassar St, Cambridge, MA 02139

***Abstract*— Natural disasters ravage the world's cities, valleys, and shores constantly. Having precise and efficient mechanisms for assessing infrastructure damage is essential to channel resources and minimize the loss of life. Using a dataset that includes labeled pre- and post- disaster satellite imagery, we train multiple convolutional neural networks to assess building damage on a per-building basis. We present a highly interpretable deep-learning methodology that seeks to explicitly convey the most useful information required to train an accurate classification model. Our findings include that ordinal-cross entropy loss is the most optimal loss function to train on and that including the type of disaster that caused the damage in combination with a pre- and post-disaster image best predicts the level of damage caused. Our research seeks to computationally aid in this ongoing humanitarian crisis.**

I. Introduction

Natural disasters devastate countless vulnerable communities and countries annually, killing on average 60,000 people each year worldwide [5]. The timely allocation of resources in the event of these tragedies is crucial to saving lives. The frequency and severity of these disasters will only continue to increase, exacerbated by climate change [6]. The catastrophic impact of natural disasters and their increasing prevalence motivates the problem addressed in this work. Deep neural networks (DNNs) have been used to locate and classify building damage within satellite imagery [2]. However, the current literature is limited in interpreting what exactly these neural networks are learning and identifying key predictors for assessing building damage. Thus, we present a novel analysis of the most important information that a deep learning model needs to assess building damage. We use a convolutional neural network (CNN) architecture called residual neural network (ResNet), pre-trained on ImageNet data. In our approach, we train multiple CNNs on xBD satellite imagery data [3] with different modalities of input, as well as with different loss functions, and compare accuracy on the validation set. We aim to explicitly provide insight into the most effective ways to train models to classify levels of building damage, maximizing the efficiency of the emergency response after a natural disaster, which has the potential to save lives and reduce economic strain.

II. Methods

For this work, we utilize the xBD dataset [3], which covers a wide range of disasters in fifteen countries. One of xBD's main purposes is to demonstrate changes between pre- and post-disaster satellite imagery to aid in detecting the damage caused. Therefore, each post-disaster building is labeled as one of the following: "unclassified," "no damage," "minor damage," "major damage," or "destroyed." The classification benchmark utilized is called the Joint Damage Scale (JDS) [3]. We use the xBD dataset because it incorporates a variety of disaster and building types, as well as geographical locations (for cross-region generalization), allowing for diversity in training the model. Additionally, the high-resolution imagery allows for detailed change detection between pre- and post-disaster images. These factors currently make xBD the leading dataset for building damage detection using labeled satellite imagery [3]. Using the segmentation ground truth masks (sets of coordinates constituting building polygons) provided, we extract individual building polygons to train on.

We train a baseline classification model to classify buildings by damage level, as defined by the JDS. Post-disaster images were the only model input. Notably, our baseline model does not use change detection because pre-disaster imagery is not taken as input. The model architecture is ResNet18, an 18-layer CNN [4]. This baseline model uses the cross-entropy loss function, which is defined as:

$$\sum_{c=1}^{4} y_{o,c} \log(p_{o,c}) \qquad (1)$$

where $y_{o,c}$ is a 0/1-binary indicator of whether c, as a label, correctly classifies observation o, and $p_{o,c}$ is the predicted probability that observation o is of the class c. The network is trained on 12,800 building crops with a batch size of 32. The crops are divided in a 0.8:0.2 ratio for training and validation data, respectively. The Adam optimizer, which is for adapted learning, is set at a learning rate of 0.001. The model trained for 100 epochs on NVIDIA Tesla K80 GPUs.

We train models that improve upon the performance of the baseline model. To do this, we introduce other model inputs, namely the pre-disaster image (in combination with the post-disaster image) and the type of disaster (e.g. volcano, wind) that caused the building damage. To train a model that takes in both pre-disaster images and their corresponding post-disaster images, we concatenate the RGB channels of the two and use that as input. To train a model that takes in the pre-disaster image, post-disaster image, and disaster type, we do the same, but also concatenate a one-hot encoded representation of the disaster type in one of the later layers of the CNN. Furthermore, we experiment with other loss functions, namely mean squared error loss (MSE) and ordinal cross-entropy loss to train these models. We define mean squared error as:

$$\frac{1}{b}\sum_{i=1}^{b}(y-\hat{y})^2 \qquad (2)$$

where b is the batch size, y is the ground truth (a class from 0 to 3 representing each damage level), and $\hat{y}$ is the prediction. Ordinal cross-entropy loss differs from cross-entropy loss in that it considers the distance between the ground truth and the predicted class (hence "ordinal"). This function is useful because the building damage classification problem involves different and increasing levels of damage from “no damage” to “destruction.” To implement ordinal cross-entropy loss as the loss function, we treat it as generic multi-class classification and encode the classes “no damage,” “minor damage,” “major damage,” and “destroyed” as [0, 0, 0], [1, 0, 0], [1, 1, 0], and [1, 1, 1], respectively [1].

III. Results and Discussion

In Table 1, we present model accuracy on the validation set across nine different models, which are differentiated by three different input combinations and three loss functions. The baseline model, which is trained with post-disaster data only and the cross-entropy loss function, has an accuracy of 59.5%, as shown, while the most accurate model has an accuracy of 74.6%. It is important to note that all models were trained and validated on data that is evenly split between building crops of each class (no damage, minor damage, major damage, and destroyed), so a purely blindly guessing model would achieve approximately 25% accuracy.

TABLE 1. COMPARISON OF VALIDATION ACCURACY ON 9 DIFFERENT MODELS

Model Input(s)	**Model Accuracy on Validation Set**		
	Mean Squared Error	***Cross-Entropy Loss***	***Ordinal Cross-Entropy Loss***
Post-Disaster Image	45.3%	59.5%	64.2%
Post-Disaster Image + Pre-Disaster Image	50.2%	68.3%	71.2%
Post-Disaster Image + Pre-Disaster Image + Disaster Type	49.7%	72.7%	74.6%

Percentages represent model’s validation accuracy given loss function and model input.

Much of our results confirm our hypotheses. Accuracy on the validation set improves when more modes of useful information are inputted into the model (accuracy generally increases moving down the rows of Table 1). This is justifiable given the intuitive assumption that the more information the model has to work with, the more accurate predictions it should make. A large part of our research addressed which types of input aid the convolutional neural networks in making accurate predictions. From the generated results, it seems that having the aspect of change detection (when the pre-disaster image is concatenated with the post-disaster image and inputted) is useful, along with the type of disaster. We also note that models using ordinal cross-entropy loss as their criterion for optimization perform the most accurately. As previously mentioned, ordinal cross-entropy loss is most specifically applicable for a classification problem that involves an ordinal scale (in this case, the JDS), as opposed to categories with no intrinsic ordering. MSE, not surprisingly, showed itself to be the least effective loss function to use for training. This result is justifiable because MSE is primarily used in regression problems, not classification problems. We find that cross-entropy loss models fall somewhere in between.

However, we note that none of the accuracy numbers are necessarily optimal. This can be explained by the fact that the differences between categories, particularly between minor-damage and major-damage, are largely difficult to discern for both humans and computers. This is a challenge that comes with non-binary classification tasks with building damage, and it has been acknowledged by many, including [3]. In addition, there is some noisy data in the dataset and cleaning it more thoroughly would most likely yield marginally more accurate predictions.

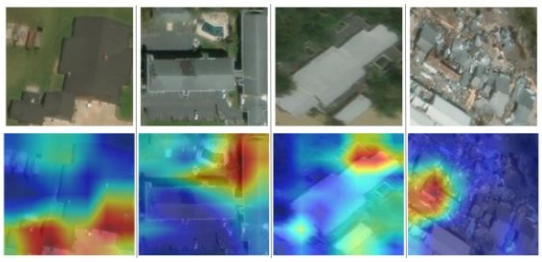


Figure 1: Gradient class activation maps depict which parts of the building crop lead the baseline model to predict a certain classification. On the top are the original images (crops) and on the bottom are the corresponding gradient class activation maps. The images included are only post-disaster images.

IV. Conclusion

The main insights that can be drawn from our work include using individualized building crops instead of semantic segmentation to train models and performing experiments with various combinations of model inputs and loss functions to explicitly examine their differences. Practically, our work (a novel, more interpretable approach) and others in the field advance methods for more robust emergency responses and more efficient allocation of resources, which saves lives.

Acknowledgment

I would like to thank my mentor, Mr. Ethan Weber, without whom this research would not have been possible.

References

[1] Cheng, Jianlin, Zheng Wang, and Gianluca Pollastri. "A neural network approach to ordinal regression." *2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence)*. IEEE, 2008.

[2] Gueguen, Lionel, and Raffay Hamid. "Large-scale damage detection using satellite imagery." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2015.

[3] Gupta, Ritwik, et al. "Creating xBD: A dataset for assessing building damage from satellite imagery." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*. 2019.

[4] He, Kaiming, et al. "Deep residual learning for image recognition." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.

[5] Ritchie, Hannah, and Max Roser. "Natural disasters." *Our World in Data* (2014).

[6] Van Aalst, Maarten K. "The impacts of climate change on the risk of natural disasters." *Disasters* 30.1 (2006): 5-18.

Protective effect of Gallic acid against Aβ neuronal toxicity and neuroinflammation for the treatment of Alzheimer's Disease

Jierui Wang[1], Wei Zhu[2]

[1]Jericho Senior High School, 99 Cedar Swamp Rd, Jericho, NY 11753
[2]223 Store Hill Road Campus Center H-310, Old Westbury, NY 11568

***Abstract*- Alzheimer's Disease (AD) is one of the leading causes of death and dementia in elderly people. Previous research indicates that AD may take its roots from Amyloid Beta (AB) aggregation and neuroinflammation. As 3,4,5-trihydroxy benzoic acid, gallic acid, demonstrates its potential to alleviate AD induced cytotoxicity and oxidative stress, this study researched the effect of gallic acid on pathological formation of AD. Gallic acid was tested against AB induced cytotoxicity, neurons apoptosis, and lipopolysaccharide (LPS) induced neuroinflammation, via a variety of biological assays, including cell viability assays, cell mitigation assays, and ELISA assays. According to the data and statistical analysis, gallic acid attennued the neuronal apoptosis by reversing the cytotoxic effects induced by AB. Moreover, gallic acid also reduced the production of AB by significantly down regulating the protein expression of Amyloid precursor protein. Gallic acid indicated its effects on neuroinflammation since it significantly inhibited LPS induced pro-inflammatory cytokines released from immune cells. All the results demonstrated that gallic acid has a positive impact on AD by inhibition of AB aggregation and AD induced cytotoxic effects, as well as anti-inflammatory effects. Overall, gallic acid gives new insight into the treatment of AD.**

I. Introduction

Alzheimer's Disease (AD) is the most predominant cause of neurodegenerative dementia and one of the leading sources of morbidity and mortality in the aging population [1].

Two hallmarks for the pathological formation of AD are known: (1) the extracellular plaques formed by amyloid-beta (Aβ) aggregates and (2) the intracellular tangles formed by hyperphosphorylated tau proteins[2]. Aβ, an amino acid commonly found in the amyloid plaques with AD, forms tangles in between the synapses of neurons in the hippocampus, blocking the communications between neurons and resulting in memory and learning impairments[3]. Since Expression of Amyloid precursor protein (APP) determines the production of Aβ, overexpression of APP results in excess Aβ production and thus aggregation [3]. In addition toAβ aggregation, bacterias induce microglia to secret pro-inflammatory cytokines, resulting in neuron inflammations, which eventually leads to neuron apoptosis [4]. Apoptosis has a significant negative impact on the developments of AD symptoms [2].

Gallic acid (GA) has been researched for its anti-inflammatory effects and effects on alleviating cell apoptosis in allergic research, obesity, and cancer. *In vivo* and *in vitro* studies on attenuating effects of GA on cytotoxicity and inflammation suggest a possible therapeutic application of this agent in other cytotoxicity-induced inflammatory diseases. Since AD is associated with neuron loss, as a result, researching GA 's effects on cytotoxicity, neuron survival rates, inflammatory responses will greatly reflex GA's effects on healing AD [5].

Thus, the goal of this study is to identify possible mechanisms by which GA may alleviate and target AD pathogenesis and symptoms and to study the possibility of this chemical to both decrease plaque aggregation and neuroinflammation, as well as reduce cell death and inflammatory cytokines induced by Aβ and bacteria.

II. Methods

Gallic acid, 3,4,5-trihydroxy benzoic acid, the powder was obtained from Sigma-Aldrich. It was then diluted with distilled water to create 10uM and 1uM concentrations.

HTB-11 human neuroblastoma cells (ATCC, USA), 3-day transfer, inoculum 3×10^5 cells (ATCC), and Raw 264.7 murine macrophage-like cells (ATCC) were used as neuron cell models to study the effects of 3,4,5-trihydroxy benzoic acid on the effects on neurons apoptosis and amyloid-beta induced cytotoxicity.

To assess the alleviative effects of GA on neuron apoptosis, cell proliferation assays were used to determine neuronal survival rates. Active cells reduced the yellow tetrazolium MTT into purple formazan. Spectrometer measured the absorbance and cells survival rates under the wavelength of 595nm after 24 hours incubation.

To analyze the effects of GA on Aβ production and neuroinflammation, Enzyme Linked Immunosorbent Assay (ELISA)assay was performed on cytokines and protein, including APP, IL-1β, TNF-α, IL-4, and IL-10 following the manufacturer's protocol (BosterBio, Pleasanton, California). A microplate absorbance reader is used for results analysis.

P-values, calculated from two-tailed T-tests, were employed to determine the statistical significance of the results. $P<0.05$ indicates statistically significant results.

III. Results

GA was found to have a positive effect in treating Alzheimer's diseases as it alleviates neuro apoptosis, attenuates inflammation, and reduces the overproduction of Aβ.

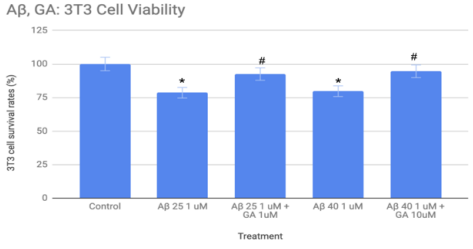


Figure 1. Fragment Aβ25 and full-length Aβ40 toxicity treated with GA; 10uM GA is the most effective in treating Aβ40 with 94.61 ± 0.096% for 24 hours. Error bars denote standard deviation. * indicates statistical significance, $p<0.05$, when data is compared to control, and # indicates statistical significance, $p<0.05$, when the data is compared to the group only treated with Aβ25 1uM or Aβ40 1uM.

Cells have the lowest survivability when treated with Aβ25 1uM, with a survival rate of 78.57%. Both treatments of Aβ25 1uM and GA 10uM and Aβ40 1uM and GA 10uM increase cell survivability. Moreover, treatment of Aβ40 1uMand GA 10uM has the highest survival rate of 94.96%, while the treatment of Aβ25 1uM and GA 10uM is slightly lower in survival rate with 92.50%. This suggests that GA may be more effective at targeting full-length Aβ cytotoxicity (figure 1).

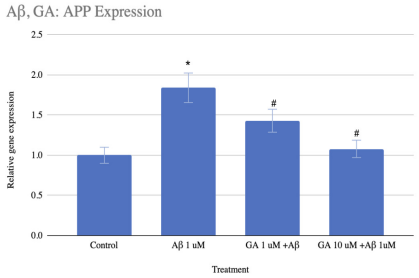
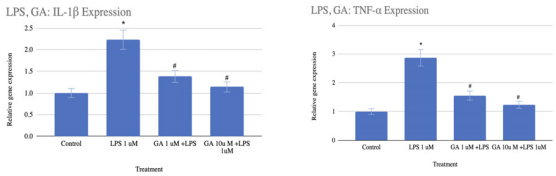


Figure 2. ELISA for APP gene expression under concentration treatment of GA. The most effective treatment is GA 10 uM + Aβ 1uM. The error bar represents standard error. * indicates statistical significance, p<0.05, when data is compared to control, and # indicates statistical significance, p<0.05, when the data is compared to the group only treated with Aβ 1uM.

Aβ 1 uM increased APP expression to 1.84 times the normal APP expression level. Both concentrations of GA significantly reduced APP expression level, while GA 10uM is the most effective by reducing APP expression to only 1.08 times the normal APP expression level. This suggests that GA is prominent in reducing Aβ stimulative effect on APP expression, and the higher concentration may be more effective (figure 2).



Figures 3a and 3b. ELISA for IL-1β and TNF-α gene expression under GA concentrations treatment; most effective treatment for most IL-1β expression is GA 10uM +LPS 1uM, with relative gene expression of 1.14±0.15, and the most effective treatment for TNF-α expression is GA 10uM + LPS 1uM, with relative expression of 1.24±0.13. Error bars represent standard errors. * indicates statistical significance, p<0.05, when data is compared to control, and # indicates statistical significance, p<0.05, when the data is compared to the group only treated with LPS 1uM.

IL-1β and TNF-α are two pro-inflammatory cytokines. Bacterial LPS increases both of their expressions to 2.23 times and 2.87 times their normal expression levels. Treatment with both concentrations of GA has a significant impact on decreasing the cytokines expression levels. GA 10um + LPS 1uM is the most effective, bringing both IL-1β and TNF expression to 1.14 times and 1.24 times their normal expression levels. This suggests GA can potentially reduce the inflammation triggered by toxins, such as LPS, and higher concentration may be a more effective option (Fig 3a & 3b).

IV. Discussion

Data from the experiment indicates that GA can be a potential treatment due to its effects to target the molecular mechanisms of AD. Aβ aggregation and resultant neuron apoptosis has been understood as the main cause of AD, and in the experiment, GA significantly alleviates cell death that was originally induced by Aβ. This can be translated into that GA has an alleviative and protective effect against Aβ-induced cytotoxicity.

APP is known to be the protein that makes excess Aβ and responsible for the Aβ aggregation in AD patients. The overexpression of APP will cause the excess production of Aβ, and a large amount of Aβ will induce APP expression again. Data from the experiment shows that when Aβ is used alone, the APP expression becomes significantly greater than its normal expression. However, GA shows its potential effect to decrease and reverse the overexpression of APP and thus reduce the production of Aβ since the expression level of APP is greatly reduced after GA is used with Aβ. Moreover, a greater concentration of GA reduces more APP expression in the experiment, implying that a high concentration of GA may be a more effective option for APP overexpression.

Pro-inflammatory cytokines, such as IL-1β and TNF-α, are overexpressing and causing the neuron's death in neuroinflammation. The data from the experiment indicates that LPS, when added alone, will greatly increase the expression of IL-1β and TNF-α, indicating that LPS very likely causes serious inflammation in the cells. However, treatment with GA reduced the overexpression of IL-1β and TNF-α to close-to-normal levels, suggesting that GA may reverse and alleviate the inflammatory responses that LPS induced. As a result, GA will be a possibility to target bacterial LPS induced inflammatory effects and neuron apoptosis.

V. Conclusion

Overall, this research investigated the effects of GA on AD. The result indicates that GA can be a prominent option to alleviate neuron death, prevent overproduction of Aβ, attenuate inflammation. It ultimately indicated that GA has the potential to target and prevent Aβ tangles. As a result, GA can be considered as a potential treatment based on the ways to attack the symptoms and pathologies of AD.

Future studies need to test GA effects in vivo AD models, such as C.elegans or experimental rats. It will be crucial to test the GA effect in vivo too since the impacts and performances of the chemicals may be different when used in the human body.

VII. Acknowledgments

I would like to thank Dr. Wei Zhu for all his help and the amazing opportunity to conduct research.

VI. References

[1]J. Weller and A. Budson, "Current understanding of Alzheimer's disease diagnosis and treatment", *F1000Research*, vol. 7, p. 1161, 2018. Available: 10.12688/f1000research.14506.1 [Accessed 10 December 2020].

[2]V. Mantzavinos and A. Alexiou, "Biomarkers for Alzheimer's Disease Diagnosis", *Current Alzheimer Research*, vol. 14, no. 11, 2017. Available: 10.2174/1567205014666170203125942 [Accessed 10 December 2020].

[3]R. O'Brien and P. Wong, "Amyloid Precursor Protein Processing and Alzheimer's Disease", *Annual Review of Neuroscience*, vol. 34, no. 1, pp. 185-204, 2011. Available: 10.1146/annurev-neuro-061010-113613 [Accessed 10 December 2020].

[4]R. Moir, R. Lathe and R. Tanzi, "The antimicrobial protection hypothesis of Alzheimer's disease", *Alzheimer's & Dementia*, vol. 14, no. 12, pp. 1602-1614, 2018. Available: 10.1016/j.jalz.2018.06.3040 [Accessed 10 December 2020].

[5]S. Kim et al., "Gallic Acid Inhibits Histamine Release and Pro-inflammatory Cytokine Production in Mast Cells", *Toxicological Sciences*, vol. 91, no. 1, pp. 123-131, 2005. Available: 10.1093/toxsci/kfj063 [Accessed 10 December 2020].

Nanohole Arrays in Ultra-Low Concentration Biosensing

Tomosuke Yamaguchi[1], Alson Chan[2], Mustafa Mutlu[3]

[1]Santa Rosa High School, 1235 Mendocino Ave, Santa Rosa, CA 95401, USA
[2]Gunn High School, 780 Arastradero Rd, Palo Alto, CA 94306, USA
[3]University of California, Santa Cruz, Nanoengineering Group, Baskin Engineering Room 230, 1156 High St, Santa Cruz, CA 95064, USA

***Abstract*— Millions of medical tests are performed annually in the United States. However, current medical testing infrastructure and methods are prone to delays and errors. Biosensing using plasmonic nanohole arrays (NHA) can meet the demand for accurate and accessible testing methods. NHAs are metallic nanoscale structures with voids that allow nano-sized particles to pass through while accumulating biomarkers on its surface. Using finite-difference time-domain (FDTD) simulations, a method to quantify electrodynamic interactions, researchers can observe the presence and concentration of accumulated biomarkers accurately and efficiently. This technique enables the detection of a variety of biomarkers, such as immunoglobulins and antigens, at ultralow concentrations. These detection results can be interpreted by medical professionals to diagnose diseases at early stages.**

I. Introduction

Medical diagnoses are often the first step in treating or preventing medical complications. Yet, up to 12 million (~5%) medical cases are estimated to be misdiagnosed annually, ~13.5% of them being infectious diseases [1]. Extant research suggests that biosensing, the detection of biological particles with an analytical device, is more rapid and accurate using plasmonic nanohole arrays (NHAs) compared to existing diagnostic technologies such as polymerase chain reaction (PCR) and enzyme-linked immunosorbent assays (ELISA) [2,3].

The structure of NHAs [Fig. 1] consists of a silicon nitride substrate coated in a gold metal film with an evenly spaced array of cylindrical voids [Fig. 2][2,3]. Yanik et al. detected a variety of DNA and RNA viruses with minimal processing of samples using this structure, displaying the simplicity and versatility of biosensing with NHAs [2]. Their experiments showed that NHAs are capable of detecting a large range of particles at ultralow concentrations without complex purification processes [2]. More recent experiments have further optimized the silicon nitride layer by creating a silicon wafer patterned with UV light and coated in a light-sensitive material [3]. These new innovations in silicon technology suggest NHAs are a viable solution to medical testing [2,3].

Zhu et al. further observed the behavior of electromagnetic phenomena existing on the NHAs [3]. They affirmed that the conductive coating on the NHA interacts with electrons in biomolecules to create propagating high-energy electromagnetic waves called surface plasmon polaritons (SPP) [3]. The frequency of these waves varies based on the refractive index (RI) of the accumulated particles [3]. When photons interact with SPPs propagating on the underside of the NHA, the transmission of light through the NHA is enhanced, resulting in a phenomenon known as extraordinary optical transmission (EOT) [2-4]. The EOT effect allows researchers to observe and measure light transmission [2,3]. Unlike traditional photonic biosensors, these plasmonic devices can couple incident light directly, enabling researchers to observe EOT as the light exits the NHA [2].

Additionally, Zhu et al. utilized the finite-difference time-domain technique (FDTD) to observe the levels of light transmission that result from bioaccumulation of varying RI [3]. Novel FDTD simulations enabled the measurement of light transmission without fabrication and the use of an electron microscope for experiments [3]. In this method, variables are isolated and time is discretized, allowing for the direct measurement of light transmission without effect from other variables [3]. Furthermore, spectral shifts of light transmission at differing RIs were credited to an effect called red shifting [2,3,5]. As the RI of the bioaccumulation increases, the velocity of incident light decreases, causing the light to stretch and increase in wavelength [5]. Based on these shifts, they found that specific RIs could be matched to various bioparticles and further matched to the wavelength of transmitted light [2,3]. The higher the concentration of such particles, the more red shifting will occur [2,3].

There remains a limited understanding of these optical properties in experimental settings and possible limitations to wider applications of NHAs in biosensing since SPPs may only penetrate the depth of an accumulated particle [2,3]. In our research, we sought to apply novel FDTD simulation techniques to observe the transmission spectra and red shifts through an NHA in relation to different RIs. Our results affirming the optical behaviors of NHAs coupled with bioaccumulation of varying RIs support broader applications to medical testing and diagnostics.

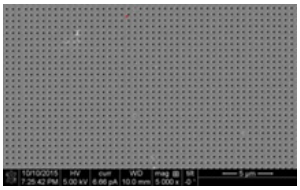


Figure 1: The surface of an NHA with a gold film and silicon nitride substrate is captured by a Scanning Electron Microscope (SEM).

II. METHODS

To predict the relationship between bioaccumulation, RIs, and light transmission in an experimental setting, we modeled and simulated an NHA in Lumerical Inc. FDTD Simulator. Our design [Fig. 2], based on designs of Zhu et al., consisted of a 125 nm thick conductive gold film placed above a 100nm thick substrate. Like previously completed simulations [3], the RI of the environment surrounding the NHA was set to 1.33, that of water. Water suspension reflects all escaped light internally, preventing any light from escaping the system [3].

A 3-by-3 nanohole array was evenly spaced across the plane and situated so its height would match that of the sensor. The light source was set along the negative z-axis to model the injection of light on a sensor when studied under an electron microscope. An FDTD calculator was applied normal to the y-axis to detect the light transmissions within our tests. Finally, an 8nm thick protein layer was added to the surface of the gold layer and nanoholes. This simulates the accumulation of particles taken from a biological sample. Light transmission was recorded from a series of simulations with protein layers of different RIs. FDTD analyses were utilized to compute light transmission within each simulation scenario.

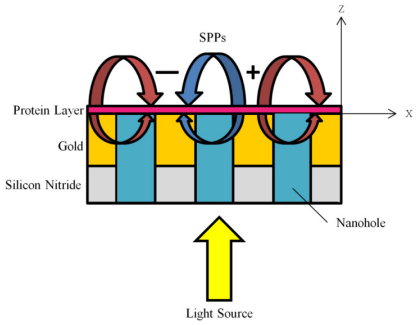


Figure 2: SPPs are illustrated at the metal-dielectric interface. A light source is injected from the negative z direction. Direction of light does not affect the behavior of SPPs or EOT, but it must be normal to the y-axis.

III. Results

After adjusting RIs in multiple trials, we observed differing transmission peaks, confirming that we can diagnose a biological irregularity using NHA. Figure 3 displays the wavelength results of our FDTD simulations, measured in nanometers (nm). Without protein accumulation (green), the maximum light transmission was observed at 685 nm. This means the wavelength of transmitted light without bioaccumulation measured 685 nm. When the refractive index of the simulated protein layer measured 1.53 (blue), 1.6 (red), and 1.7 (teal), the transmission peaks were observed at 692nm, 694nm, and 698nm, respectively. Thus, when the bioaccumulation refractive index rose, the transmission peak moved further to the right.

We reason that red shifting causes transmitted light to increase in wavelength as the RI increases. We further deduce that as the incident light slows and refracts at greater angles due to red shifting, the EOT will decrease. This effect can be observed by the decreasing peak levels of transmission (measured in a.u.) of each respective curve. This means researchers can identify the RI of any biomarker with the wavelength of light transmitted through it. For example, if an immunoglobulin of RI 1.6 accumulates on a biosensor, it can be detected based on the wavelength of light it transmits. By comparing different levels of light transmission, biosensing with NHAs can inform the presence and concentrations of specific molecules. These detection results can be further utilized by medical professionals to diagnose diseases.

IV. Conclusions

Using the FDTD technique, NHAs are capable of discerning differences in transmission peaks between molecules and thus allow researchers to detect the presence of biological irregularities using known RIs. Our findings suggest that using FDTD to measure the EOT that originates from bioaccumulation on an NHA effectively distinguishes between samples with or without proteins present. This is supported by our FDTD simulations, which show the usability of NHAs in accurately differentiating between biological particles by comparing transmission peaks. Therefore, this method will allow for the detection of biological differences at a broad spectrum of wavelengths and a wide variety of particles. Additionally, published research [2] shows binding sites or receptors can be mobilized on NHAs to detect a variety of biological particles under ultra-low concentrations and within the same sample. This enables the detection of a wide variety of proteins and particles, informing the diagnosis of diseases before they become more serious. Once further optimization is reached, plasmonic biosensing could provide an effective and efficient way for doctors to determine the biological composition of samples, diagnosing and thus preventing future diseases.

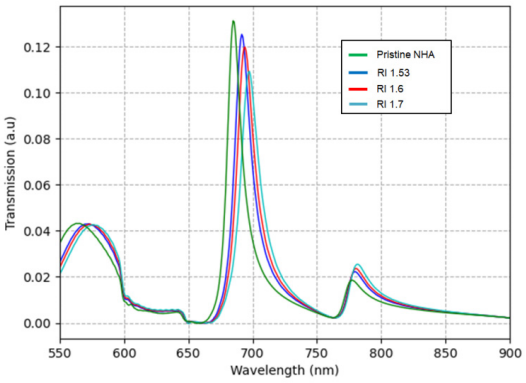


Figure 3: EOT peaks varied based on the RI of protein layers and wavelengths of light.

Acknowledgments

We extend our deepest gratitude to our mentors Professor Ali Yanik, Mustafa Mutlu, and Ahsan Habib for their expert guidance and allowing us to participate in their research. We thank Alexa Lowe, Safaa Hussain, and Grace Xia for their contributions. We also appreciate the UCSC Science Internship Program and the UCSC Nanoengineering Group for providing us the opportunity to participate in research.

References

[1] Newman-Toker D.E.; Wang Z.; Zhu Y.; Nassery N.; Tehrani A.S.; Schaffer A.C.; Yu-Moe C.W.; Clemens G.D.; Fanai M.; Siegal D. “Rate of diagnostic errors and serious misdiagnosis-related harms for major vascular events, infections, and cancers: toward a national incidence estimate using the ‘Big Three’” Diagnosis (2020). https://doi.org/10.1515/dx-2019-0104.
[2] Yanik, A.A.; Huang, M.; Kamohara, O.; Artar, A.; Geisbert, T.W.; Connor, J.H.; Altug, H. “An optofluidic nanoplasmonic biosensor for direct detection of live viruses from biological media.” Nano Letters, 10.12 (2010), pp. 4962-4969. https://doi.org/10.1021/nl103025u
[3] Zhu, X.; Cao N.; Thibeault, B.J.; Pinsky, B.; Yanik, A.A. “Mechanisms of Fano-resonant biosensing: Mechanical loading of plasmonic oscillators.” Optics Communications (2020), pp. 125780. https://doi.org/10.1016/j.optcom.2020.125780.
[4] Ebbesen, T.W.; Lezec, H.J.; Ghaemi, H.F.; Thio, T.; Wolff, P.A. “Extraordinary optical transmission through subwavelength hole arrays.” Nature, 391.6668 (1998), pp. 667-669. https://doi.org/10.1038/35570.
[5] Kevin A. Tetz, Lin Pang, and Yeshaiahu Fainman, "High-resolution surface plasmon resonance sensor based on linewidth-optimized nanohole array transmittance," Opt. Lett. 31, 1528-1530 (2006)

Development Of Cactus Oil Dispersant Suitable For Oil-Degrading Marine Bacteria And Various Seawater Temperatures

Evelyn Hur[1], Sunny Kim[2]

[1]Seoul International School, Seoul, South Korea
[2]Genuine Research, Seoul, South Korea

***Abstract*— Following an oil spill, dispersants are sprayed in large volumes onto the ocean surface and work by emulsifying oil particles. However, most chemical dispersants contain toxic properties and inhibit natural oil-degradation by bacteria. This study investigates the effects of a natural dispersant from *Opuntia ficus-indica* cactus mucilage on oil-degrading *Marinobacter* growth and oil emulsification by testing marine broth absorbance and qualitatively rating dispersion respectively. The *Opuntia ficus-indica* was selected due to its unique surface-active properties, allowing it to naturally enhance dispersion. The results indicate that the mucilage enhances oil-degrading *Marinobacter* sp, *M. aestuarii, M. antarcticus,* and *M. maritimus* populations. Microscopic inspection suggests that mucilage is an effective dispersant both alone and in combination with the *Marinobacter*. The mucilage was also equally effective at various temperatures and without certain standard pre-treatment from previous studies. Lastly, a cell viability test indicated that both acidic and neutralized pH cactus mucilage did not damage cell growth, and is thus safe to use in oceanic waters. In short, cactus mucilage is a natural nontoxic dispersant that promotes oil-degrading *Marinobacter* and works well at various oceanic conditions.**

I. Introduction

When oil spreads across the ocean surface following an oil spill, less sunlight is able to penetrate the water, reducing phytoplankton population and harming marine wildlife [1]. To prevent oil from clumping on the surface, dispersants are applied to emulsify oil through amphipathic properties [2]. However, there has been much controversy surrounding the use of dispersants due to the toxicity of early generation of chemical dispersants. For instance, chemical dispersants may suppress natural oil degradation by marine bacteria [3].

This study tests the oil-dispersing abilities of the cactus *Opuntia ficus-indica*, a natural nontoxic alternative to chemical dispersants, at various ocean temperatures and conditions. In this study, a slimy fluid visible in the cross section of the pad was extracted and studied [4]. In the past, powder cactus mucilage was found to be an effective and eco-friendly dispersant at various salinities [5]. Although insoluble in water, the mucilage swells in water, allowing it to potentially precipitate ions, bacteria, and particles from aqueous solutions. The mucilage also holds unique surface-active properties, allowing it to enhance dispersion and reduce high polarity liquid surface tension [5]. Since there are currently no studies on the effect of cactus mucilage on oil-degrading *Marinobacter* growth, this study investigates whether cactus mucilage can coexist with *Marinobacter* oil-degradation and the conditions under which this would be a possibility.

II. Methods

Cactus mucilage was extracted from the pads of *Opuntia ficus-indica* by blending, heating, and neutralization [4].

The effect of cactus mucilage on oil-degrading *Marinobacter* sp., *Marinobacter aestuarii, Marinobacter antarcticus,* and *Marinobacter maritimus* population growth was first tested by incubating each species in marine broth with 10% diluted mucilage, measuring absorbance at 600nm after 48 hours, then calculating growth rate.

The efficacy of oil dispersion by the mucilage on diesel oil was then observed via visual microscopic inspection of a drop of oil-water-dispersant solution once, mixed with methylene blue and second, under fluorescent light. Images were rated 1-5 based on number and size of oil particles and oil-water drop diameter ratio was calculated. Emulsification was not only analyzed for the mucilage alone, but also in combination with the various *Marinobacter* species.

The same two tests were repeated for mucilage after various pre-treatment including autoclave heating, alcohol-drying, etc. to find optimal conditions. The effect of seawater temperatures 4°C and 25°C were also tested.

Lastly, a cell cytotoxicity test on SK-N-MC cells was conducted to ensure that the cactus mucilage is non-toxic to marine wildlife. SK-N-MC, neuroblastoma cancer cells from the brain, were chosen to simply test if the mucilage had any adverse effects on general cell growth. The cells were seeded on a 12-well plate with DMEM, fetal bovine serum and penicillin streptomycin. After 48 hours, the cells were treated with diluted mucilage. Cell viability was calculated using a CCK-8 kit and measuring absorbance at 450nm.

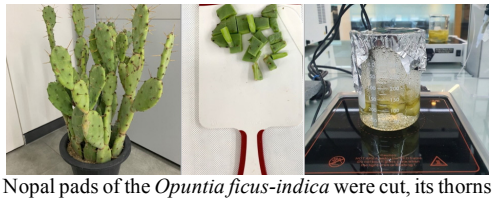


Figure 1. Nopal pads of the *Opuntia ficus-indica* were cut, its thorns removed, heated, and neutralized with 0.1M NaOH.

Table 1. *Marinobacter* sp. characteristics

	Marinobacter sp.	*Marinobacter aestuarii*	*Marinobacter antarcticus*	*Maribacter maritimus*
Catalogue number	KCTC 22332	KCTC 52913	KCTC 23684	KCTC 62399
Culture media	Marine agar	Marine agar	Marine agar	Marine agar

III. Results

A comparison of average growth rate for each *Marinobacter* species showed that *M.* sp., *M. aestuarii,* and *M. antarcticus* respectively all increased after the addition of cactus mucilage in marine broth. *M. maritimus* did not show significant growth, however. Since *M. aestuarii* showed the highest growth rate among the *Marinobacter* species, this bacteria was used to test oil emulsification in the next step.

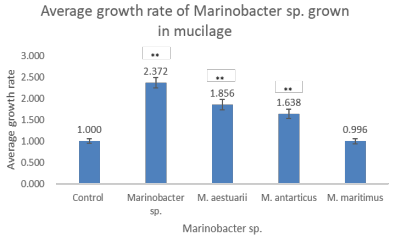


Figure 2. Average growth rate of *Marinobacter* sp. grown in mucilage. The data are shown as mean ± S.E.M. (n=3). **P<0.01, vs control.

To measure oil emulsification efficacy, oil layer ratio was calculated as seen below. A lower ratio means more effective oil emulsification. For all diesel concentrations, cactus mucilage alone reduced oil layer ratio. Combined mucilage and *M. aestuarii* further reduced the ratio at 10% and 20% diesel. Overall, the greatest emulsification was at 20% diesel for the mucilage and *M. aestuarii* combined.

Comparison of the various pre-treated mucilage from the cacti indicated no significant differences in *Marinobacter* growth and oil emulsification at various conditions. Although neutralized mucilage at pH 7.8 showed higher marine broth absorbance than at pH4.5, both mucilage pH groups still showed higher absorbance than the control (no mucilage). The only outlier was *M. aestuarii.* This meant that regardless of whether the cactus mucilage was left acidic or neutralized, the mere presence of mucilage increased *Marinobacter* growth compared to the control. Thus, mucilage acidity and other pre-treatment factors had little effect on the *Marinobacter* and oil emulsification.

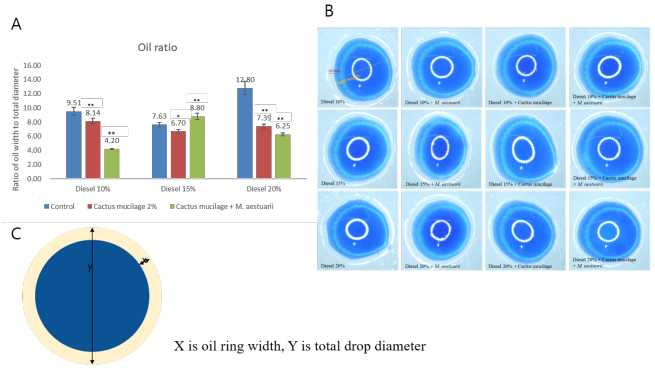


Figure 3. Oil dispersion efficacy of mucilage with *M.* aestuarii, A. Ratio of width to total diameter for methylene blue drop test, B. Microscopic image of methylene blue oil-water drops, C. Calculation method of ratio. The data are shown as mean ± S.E.M. (n=3). $^{**}P<0.01$, $^{*}P<0.05$, vs control.

Additionally, differences in seawater temperature did not show any adverse effects on *Marinobacter* and oil dispersion. Although all four species exhibited slightly higher absorbance at 25°C than 4°C, given the nature of these bacteria, such an effect is reasonable. Moreover, the correlation between emulsification score and temperature was insignificant as seen by overlapping error bars. The only outliers were Mucilage + *M. aestuarii* and with *M. maritimus*. This indicates that the cactus mucilage can maintain *Marinobacter* populations while enhancing emulsification at both arctic and tropical oceanic temperatures.

Lastly, in the cell cytotoxicity test, both pH4.5 and pH7.8 mucilage showed similar absorbance to those of both controls as seen in the overlapping error bars. Thus, the results indicate that this acidity does not pose a significant threat to cell viability, meaning that the mucilage will likely be non-toxic to marine wildlife. As such, cacti neutralization may not be necessary as seen in previous studies [5].

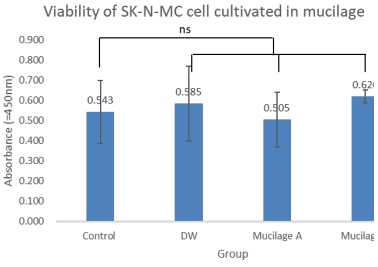


Figure 4. Viability of cells as calculated by absorbance. Control= no mucilage; DW= distilled water; Mucilage A= pH4.5; Mucilage B= pH7.8. The data are shown as mean ± S.E.M. (n=3). ns: Not significant vs control.

IV. Discussion & Conclusion

Overall, cactus mucilage did not harm but instead promoted the growth of *Marinobacter*, and oil emulsification was effective both alone and in combination with *Marinobacter* species at low and high oceanic temperatures. Furthermore, pH, heating temperature, and alcohol drying did not have a significant effect on *Marinobacter* growth and oil dispersion, meaning that certain standard procedures may be eliminated from previous studies [5]. Lastly, mucilage at both acidic and neutral pH did not harm cell viability, ensuring that the cactus will be safe to use in oceanic waters.

Still, further research needs to be done on the interaction between oil-degrading bacteria and natural dispersants by observing the effect of dispersants on not only the bacteria's population, but it's oil-degradation rate. Other methods of preparation and application other than using mucilage in powdered and liquid form may also be studied such as long-term freezing, mixing with other dispersants, etc. For instance, an alternative oat-derived dispersant can be prepared through a variety of drying techniques such as fluid bed drying, subjecting the powder to a vacuum with or without the addition of heat, or simply air drying [6]. In short, this study provides an eco-friendly alternative to chemical dispersants that works well with oil-degrading marine bacteria at various oceanic conditions.

Acknowledgements

I would like to thank my mentor Dr. Sunny S. Kim for giving me the opportunity to conduct this research.

References

[1] Doerffer, J. W. (1992). *Oil spill response in the marine environment*. Pergamon Press.

[2] Guo, F., Stebbins, D. D., Toomey, R. G., & Alcantar, N. A. (2019). Interfacial Phenomena of Natural Dispersants for Crude Oil Spills.*Langmuir,35*(48),15904-15913. doi:10.1021/acs.langmuir.9b02036

[3] Kleindienst, S., Seidel, M., Ziervogel, K., Grim, S., Loftis, K., Harrison, S.,Joye, S. B. (2015). Chemical dispersants can suppress the activity of natural oil-degrading microorganisms. *Proceedings of the National Academy of Sciences, 112*(48), 14900-14905. doi:10.1073/pnas.1507380112

[4] Cardenas, Adriana & Goycoolea, Francisco & Rinaudo, Marguerite. (2008). On the gelling behaviour of 'nopal' (Opuntia ficus indica) low methoxyl pectin. *Carbohydrate Polymers*. 73. 212-222. 10.1016/j.carbpol.2007.11.017.

[4] Alcantar, N. A., Fox, D. I., Thomas, S., & Toomey, R. G. (2015). *Use of cactus mucilage as a dispersant and absorbant for oil in oil-water mixtures*. US9163374B2

[5] Potter, R. (1995). Proteinaceous oil spill dispersant. US5399350A

Genome Analysis of a Novel Photoarsenotroph, *Rhodobacter sp.* str. ORIO

Pranav Kirti[1], Isaac Chang[1], Madeline Day[1], Sanjin Mehić[2], and Chad Saltikov[2]

[1i]Lynbrook High School, 1280 Johnson Ave, San Jose, CA 95129
[1ii]Saratoga High School, 20300 Herriman Ave, Saratoga, CA 95070
[1iii]Amador Valley High School, 1155 Santa Rita Rd, Pleasanton, CA 94566
2i, 2ii Department of Microbiology and Environmental Toxicology, University of California, Santa Cruz, 1156 High St, Santa Cruz, CA 95064

***Abstract*— Photosynthetic arsenite oxidation (photoarsenotrophy) is an anoxygenic process where arsenite is used as an electron donor for growth. Through this process, bacteria oxidize arsenite (As(III)) to arsenate (As(V)), a less toxic substance. Extremophiles from hypersaline environments, which are difficult to grow and genetically manipulate, were previously used to determine that photoarsenotrophy is encoded by the *arxB2AB1CD* gene cluster. Here, we analyzed a novel freshwater bacterium, *Rhodobacter sp.* str. ORIO, for its potential to serve as a model organism in the study of photoarsenotrophy. We utilized NCBI BLAST to gather *arx* sequences, EBI Clustal Omega to construct phylogenetic trees, and InterPro, Pfam, Phobius, and SAPS to analyze ORIO's *arx* gene products individually. Lastly, we searched for the *arx* pathway in over 35,000 metagenomes from JGI IMG/MER and mapped the presence of photoarsenotrophs worldwide. Our findings indicate that ORIO contains an *arx* gene cluster similar to those previously studied, and that *arxA-like* genes are ubiquitous in nature, with concentrations in the U.S. and Southeast Asia. In conclusion, these results suggest that ORIO can serve as a model organism for photoarsenotrophy – providing the first step toward bioremediation by detoxifying arsenic-contaminated water and improving water quality.**

I. Introduction

Arsenic contamination affects up to 200 million people worldwide; exposure can have carcinogenic effects and lead to arsenicosis, arsenic poisoning that affects multiple organs [1]. Arsenic-contaminated groundwater is a detrimental problem in lower- and middle-income countries with large geologic arsenic deposits such as Bangladesh, where an estimated 27% of wells exceed the 10 µg/L contaminant limit [2]. The need for clean water will only increase as the human population grows.

Bacteria play a crucial role in cycling arsenic through the environment, as they utilize arsenite as an electron donor and energy source during cellular respiration (oxidizing it into arsenate). Therefore, understanding their function will help researchers remove arsenic from water and mitigate contamination. Investigations on bacteria have led to a hypothesized model of the *arx* gene cluster, which encodes for the anaerobic arsenite oxidase complex and is responsible for photoarsenotrophy [3]. However, photoarsenotrophy has only been researched in extremophiles, which are difficult to grow and genetically manipulate in the laboratory. To circumvent these difficulties and simultaneously fill this research gap, we analyzed the genome of *Rhodobacter sp.* strain ORIO, a freshwater photoarsenotroph, and we surveyed numerous metagenomes for *arx*-like gene clusters. Our findings suggest that ORIO is a suitable model organism for photoarsenotrophy and the *arx* gene cluster is globally ubiquitous in a diverse group of environments (Fig 2B).

II. Methods

Organism: Rhodobacter sp. str. ORIO (Owens River Isolate Oxidizer) is the first non-extremophile photoarsenotroph discovered and was isolated from Owens River (CA) sediments in 2014 by the Saltikov Lab (University of California, Santa Cruz). Sanjin Mehić (PhDc) cultured *Rhodobacter sp.* str. ORIO with PNS medium [4].

Genome Assembly and Annotation: We used raw 150 bp paired-end Illumina MiSeq sequencing reads to assemble the ORIO genome. Forward and reverse reads were uploaded to PATRIC [5], a database for bacterial genome analysis, and its Unicycler pipeline was used to assemble the genome. Next, we used RAST-tk via PATRIC's annotation service [6] to annotate the genome using the default setting.

Phylogenetic Analysis: After generating our genome draft, we created phylogenetic trees for each gene in the *arxXSR* and *arxB1AB2CD* gene clusters. We used the ORIO *arx* sequences obtained from our draft genome with NCBI BLAST [7] to search for the individual *arx* genes, collect the top 25 matches, and create an alignment file using EBI Clustal Omega [8]. Then, we inputted the alignment into EBI Simple Phylogeny [8] to create a phylogenetic tree and used iTOL [9] to customize each tree.

Protein Analyses: To investigate the biochemical properties of the *arx* genes, we analyzed each gene from the *arxXSR* and *arxB1AB2CD* gene clusters with EBI's bioinformatics toolkits: InterPro [10], Pfam [12], Phobius [12], and SAPS [8]. These programs predicted each gene product's molecular weight, protein family/domains, hydrophobicity, and transmembrane topology.

Metagenomic Analysis and Mapping: We blasted ORIO's *arxA* gene in over 35,000 metagenomes on JGI IMG/MER [13, 14] to determine the *arx* genes' environmental relevancy. For each metagenome hit, we searched for the hypothesized "GRGWG" active site in [15] the protein translation of *arxA*. We created a map that illustrates where *arxA* and other *arx* genes can be detected (Fig. 2).

III. Data Analysis

Genome Assembly and Annotation: PATRIC's genome assembly feature provided us with the summary statistics found in Table 1. ORIO's genome size and GC content falls into the range commonly seen in *Rhodobacter* species. We found the entire *arx* gene cluster on a 27.5 kB contig.

Table 1. ORIO GENOME CHARACTERISTICS

GC Content	*Contigs*	*Base Pairs*	*N50*
68.2%	160	4,601,101	217,804

a. The assembled and annotated genome had a course consistency of 98.0%, a fine consistency of 96.8%, 3089 protein-encoding genes with functional assignment, and 1623 protein-encoding genes without functional assignment.

Phylogenetic Analysis: We created protein phylogenetic trees for each of the *arx* genes. The phylogenetic trees indicate that the genera *Rhodobacterales*, *Rhodobacteraceae*, and *Rhodospirillales* contain closely related *arx* sequences to ORIO. Furthermore, Table 2 includes bacteria closely related to ORIO based on the *arx* genes' phylogeny.

Table 2. CLOSELY RELATED BACTERIA TO ORIO

ORIO Gene	*Closest Relative*
arxA	*Thiocapsa rosea*
arxB1	*Rhodovulum steppense*
arxB2	*Rhodovulum steppense*
arxC	*Ectothiorhodospira sp.* PHS-1
arxD	*Azoarcus sp.* CIB

a. Selected bacteria are sister taxa of ORIO based on each gene's phylogeny.

Protein Domain Analyses: InterPro predicted protein domains and conserved sequences present in each *arx* gene translation; Pfam generated predictions for the protein encoded by each *arx* gene; Phobius created a graph conveying predictions for transmembrane topology and signal peptides; and SAPS produced a list of protein sequence properties, including molecular weight. This data was used to reinforce the current hypothesized photoarsenotrophy model (Fig. 1). For example, it was hypothesized that *ArxC* is a transmembrane protein, which our Phobius results demonstrated. Specifically, we verified the location and function of *ArxA*, *ArxB*, *ArxB2*, and *ArxC*.

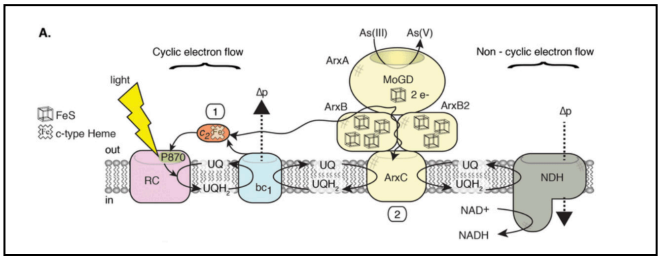


Figure 1. The current hypothesized model of photoarsenotrophy. This model was developed by the Saltikov Lab using extremophiles [16].

Metagenomic Analysis and Mapping: We searched *arxA*'s protein sequence in BLAST, which yielded 371 hits for bacteria with *arx*-related genes within their metagenomes (Fig. 2A). Specifically, we filtered our results by searching for the *ArxA* active site sequence (GRGWG). Our map exemplifies the ubiquity of *arx* genes and indicates that *arx* genes can be found on all seven continents. Also, we found that bacteria containing the *ArxA* active site were distributed among 11 primary environments, which included hot springs, wastewater, and freshwater areas (Fig. 2B).

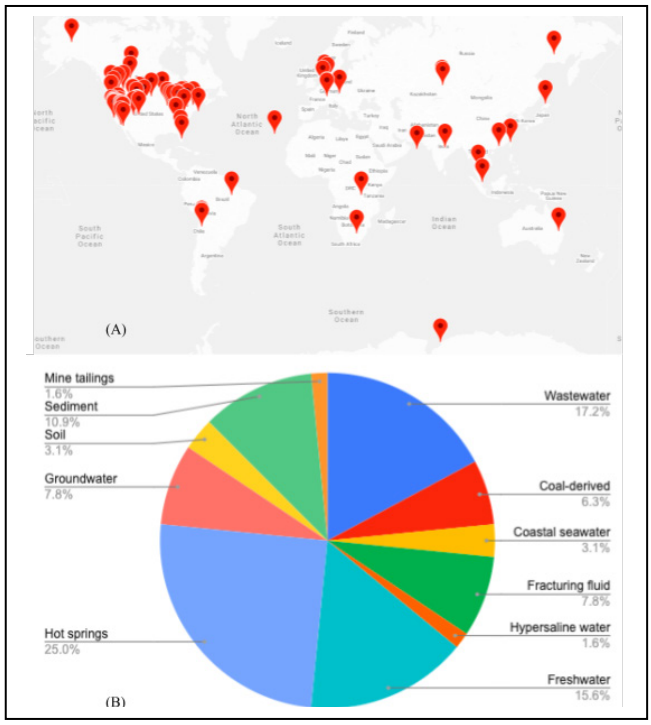


Figure 2. (A) Locations with the *ArxA* active site present. (B) A pie chart depicting the environments where *arx* genes were found.

IV. RESULTS/DISCUSSION

Overall, our findings are fundamental to understanding photoarsenotrophy and have established *Rhodobacter sp.* str. ORIO as an excellent model organism for photoarsenotrophy. This is because its *arx* genes are closely related to previously studied photoarsenotrophs. Protein domain analyses validated that each ORIO *arx* gene product contains the same features that were previously described in photoarsenotroph studies. For example, we detected hypothesized molybdenum-containing domains in ArxA, iron-sulfur clusters in the ArxB2 and ArxB1 putative proteins, and several transmembrane domains in ArxC (Fig. 1). Our discovery will allow researchers to more easily and efficiently genetically manipulate *arx* genes. This will expedite the overall study of photoarsenotrophy and drive novel experimentation.

Lastly, we are the first group to detect *arxA*-like sequences in metagenomes on all seven continents (Fig. 2A). Freshwater, groundwater, and wastewater environments accounted for 40.2% of *arxA* hits, whereas hot spring and saline environments accounted for 29.7% (Fig. 2B). This is interesting since nearly all previous photoarsenotrophy studies have been conducted in hot spring and saline environments. When taken into consideration with ORIO's genetic malleability [17], ORIO emerges as the leading model organism for studying photoarsenotrophy.

Future studies should aim to complete the genome sequence of *Rhodobacter sp.* str. ORIO and perform gene knockouts to further enhance our knowledge of the mechanisms behind photoarsenotrophy. In the long term,

the goal is to harness the unique oxidative ability of these bacteria to develop compact technology that will improve water quality through bioremediation.

V. Acknowledgements

This project was made with the support of our primary mentor, Sanjin Mehić (PhDc), and faculty advisor, Dr. Chad Saltikov (PhD). We would also like to thank the Science Internship Program and the University of California, Santa Cruz for providing us with an opportunity to conduct open-ended research.

VI. References

[1] Wang Q, et al., Environ Microbiol, 2018, 20: 1782-1793. doi:10.1111/1462-2920.14108
[2] Islam MS, Islam F, Accessed 2020, IWA Publishing. https://www.iwapublishing.com/news/arsenic-contamination-groundwater-bangladesh-environmental-and-social-disaster
[3] Hernandez-Maldonado J, et al., Environ Microbiol, 2017 Jan;19(1):130-141. doi: 10.1111/1462-2920.13509.
[4] Lindquist, J. 2014 July;21. https://www.splammo.net bact102/102pnsb.html
[5] Davis JJ, et al., Nucleic Acids Res. 2020 Jan 8;48(D1):D606-D612. doi: 10.1093/nar/gkz943.
[6] Brettin, T., et al., Sci Rep. 2015. 5: 8365.
[7] BlastP [internet]. Bethesda (MD): National Library of Medicine (US), National Center for Biotechnology Information; 2004 – [cited 2020 October 6]. Available from: https://www.ncbi.nlm.nih.gov/blast/
[8] Madeira F, Park YM, Lee J, et al., Nucleic Acids Research, 2019 Jul;47(W1):W636-W641. DOI: 10.1093/nar/gkz268.
[9] Ivica Letunic, et al., Nucleic Acids Research, Volume 47, Issue W1, 02 July 2019, Pages W256–W259, https://doi.org/10.1093/nar/gkz239
[10] Alex L Mitchell, et al., Nucleic Acids Research, Jan 2019, (doi: 10.1093/nar/gky1100)
[11] Sara El-Gebali, et al., Nucleic Acids Research, Volume 47, Issue D1, 2019, Pages D427–D432, https://doi.org/10.1093/nar/gky995.
[12] Käll L, Krogh A, Sonnhammer EL. J Mol Biol. 2004 May 14;338(5):1027-36. doi: 10.1016/j.jmb.2004.03.016.
[13] Chen IA, et al., Nucleic Acids Res. 2019 Jan 8;47(D1):D666-D677. doi: 10.1093/nar/gky901.
[14] Supratim Mukherjee, et al., Nucleic Acids Research, Volume 47, Issue D1, 08 January 2019, Pages D649–D659, DOI:10.1093/nar/gky977
[15] Glasser, Nathaniel R., et al. Proceedings of the National Academy of Sciences, vol. 115, no. 37, National Academy of Sciences, Sept. 2018, pp. E8614–23. doi:10.1073/pnas.1807984115.
[16] Hernandez-Maldonado J., 2017. Dissertation. University of California Santa Cruz
[17] Mehic S and Saltikov CW. ASM Microbe. June 20-24, 2019. Poster. San Francisco, CA. DOI: 10.13140/RG.2.2.13409.66403/1

CJSJ

DeepWaste: Instantaneous and Ubiquitous Waste Classification using Artificial Intelligence for Combating Climate Change

Yash Narayan[1]

[1]The Nueva School, 6565 CA-35, Hillsborough, CA 94010

***Abstract*—Accurate waste disposal, at the point of disposal, is crucial to fighting climate change. When materials that could be recycled or composted get diverted into landfills, they cause the emission of potent greenhouse gases. Current attempts to reduce erroneous waste disposal are expensive, inaccurate, and confusing. In this work, we propose *DeepWaste*, an easy-to-use mobile app, that utilizes highly optimized deep learning techniques to provide users instantaneous waste classification into trash, recycling, and compost. We experiment with several convolution neural network architectures to detect and classify waste items. Our best model, a deep learning residual neural network with 50 layers, achieves an average precision of 0.881 on the test set. We demonstrate the performance and efficiency of our app on a set of real-world images.**

I. Introduction

Every year, the world generates over 2 billion tons of solid waste [1]. In the U.S., even though 75% of this waste is capable of being recycled, only 34% is actually recycled [2]. Further, 91% of plastic isn't recycled [3] and only about 5% of food and other organic waste is composted [4]. This waste generates over 1.5 billion metric tons of CO2 equivalent greenhouse gases [1], contributing nearly as much to climate change as all the cars on the U.S. roads.

Despite massive investment to educate the public about accurate waste disposal, efforts so far have been only moderately successful. People are often confused by what they can recycle or compost. Signs and boards found ubiquitously near waste bins are difficult to understand and often incomplete. Furthermore, disposal of waste varies based on the local recycling facilities' capabilities, and therefore rules for disposal are subject to change on a county-by-county basis [5].

Errors in waste disposal constitute not only missed opportunities to recycle or compost, but also lead to the contamination of recycling and compost bins. Often, an entire bin can end up at a landfill due to a single error, leading to contamination. Data from the National Waste and Recycling Association show that human confusion in the identification and correct disposal of waste into our waste bins results in nearly 25% of recyclables getting contaminated [6], diverting materials that could be recycled into our landfill. When a recyclable or compostable material ends up in the landfill, it releases methane, which is several times more potent than CO2 in contributing to global warming [7].

In this work, we leverage the recent improvements of convolution neural networks (CNNs) for image-recognition tasks [8] and the availability of increased computational power on modern-day cell phones, to provide a novel approach for waste identification that is fast, low-cost, and accurate for anyone, anywhere. Our application, ***DeepWaste***, is the first mobile app targeted at the problem of erroneous waste disposal, at the point of disposal through deep learning.

II. Previous Work

The topic of waste classification has recently begun to garner some research interest, but the attempts reported across literature to solve this problem have suffered from low accuracy (ranging from 22% to low 70%) [9], [10], or a network size that is too big for real-time application [11]. Further, most of the previous attempts to mitigate the aforementioned problem of erroneous disposal envisaged deployment within a "smart bin" or the use of a commercial and industrial grade binning system within a recycling plant [12], [13], requiring expensive hardware that costs thousands of dollars. The high cost of these solutions has so far been a deterrent to their large-scale adoption. Our approach is novel, as it allows the use of widely available mobile phones, and therefore has the potential of large-scale adoption at little or no cost.

Finally, none of the previous approaches have targeted compost classification. This is a significant problem because when compostable material such as food scraps and green waste gets into a landfill, it is generally compacted down and covered. This removes the oxygen and causes it to break down anaerobically. Eventually, this releases methane, a greenhouse gas that is 25 times more potent in warming the earth than carbon dioxide over a 100-year timescale. Our work is the first work in literature, to the best of our knowledge, that considers not only recycling, but also compost as a new category for classification.

III. Methods

Classifying waste using machine learning is challenging for three reasons. First, whether a waste is recyclable or compostable depends on the properties of the material, which can be hard to detect simply from the image. Second, the material can come in any shape, such as a broken bottle, or a crumpled can, or deformed plastic; any machine learning technique needs to deal with this variation. Third, material that is recyclable depends on the capabilities of the local recycling center, so the app needs to take this geographical variable into account.

Since there was no public dataset available to accomplish this task of waste classification, a dataset was constructed

from scratch by contacting various recycling centers and collecting images from the local neighborhood. Towards this goal, we implemented user-based model training capabilities so that users can easily take a picture, label it, and upload it to the cloud for further training. In total, we manually collected 1218 images items at various lightings and angles, with 396 images containing compostable item(s), 427 images containing recyclable item(s), and 395 images containing landfill item(s). Utilizing this data, we experiment with several state-of-the-art convolution neural network methods, including InceptionV3, Inception ResnetV2, Resnet 50, Mobile Net, and PNAS Net. All CNNs used were initialized with weights pre-trained on ImageNet. During training, each input image was rotated with an angle randomly selected and also randomly flipped, cropped, and blurred for data augmentation. Each method outputs a confidence for an inputted image. Hyper parameters specific to each method are set to the best values described in the original work.

IV. Results

Out of the various CNNs benchmarked on the dataset, Resnet50 showed the best accuracy and convergence on the test set in terms of average precision and thus was optimized and subsequently deployed inside of a mobile app using Apple CoreML. Core ML optimizes on-device performance by leveraging the CPU, GPU, and Neural Engine while minimizing its memory footprint and power consumption. The DeepWaste model is running strictly on the user's mobile device, therefore removing the need for internet connection and sharing data. The benchmarked models and their respective average precision scores on the test set can be found in Table 1.

Accuracy	InceptionV3	Inception ResnetV2	Resnet 50	MobileNet	PNAS Net
Trash	0.771	0.773	0.761	0.751	0.722
Recycle	0.891	0.783	0.924	0.949	0.864
Compost	0.806	0.806	0.882	0.873	0.841
Overall	0.84	0.82	0.881	0.842	0.852

Table 1: CNN performance on test set

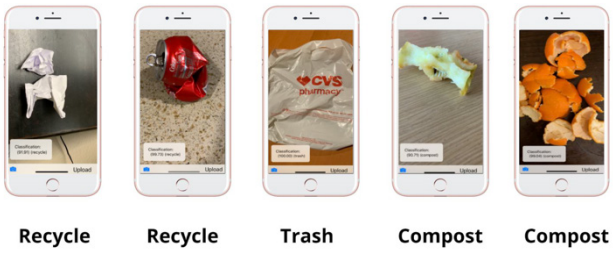


Figure 1: *DeepWaste app classification output*

Fig. 1 shows the DeepWaste app classifying items commonly confused in real life. A user can simply point their phone camera to any piece of waste and get instantaneous feedback, with an average prediction time of around 100ms. DeepWaste is able to correctly identify items with high accuracy, even when the shape has been deformed, such as a crushed soda can, an apple core, crumpled paper, and a plastic bag. Note that the plastic bag in *Fig 1* is classified as trash because plastic bags, films, and wraps cannot be recycled in a curbside recycling bin; they must be dropped-off at a special retail stores that can collect plastic grocery bags for recycling [14]. Throwing this plastic bag into the recycling bin has the potential of contaminating the entire bin.

V. Conclusion and Future Work

In this work, we present ***DeepWaste***, a mobile application that uses highly optimized deep learning techniques to provide users instantaneous waste classification, enabling users to accurately dispose of waste into recycling, compost, or trash. Currently, DeepWaste is available to beta users for testing. In future work, we aim to construct a larger data set by releasing the app to the general public so that more users can add to the growing DeepWaste database. We have started a conversation with local recycling companies to explore if they would be interested in promoting DeepWaste to their customers. The initial feedback is encouraging, and we hope to continue this conversation and start a trial soon. We hope that our work can reduce the amount of incorrect waste disposal and over time, raise more awareness around the environmental impacts of waste on our climate. If DeepWaste can even reduce erroneous waste disposal by 1%, it will be equivalent to removing over 6.5 million gasoline-burning passenger vehicles from the road, demonstrating the potential for machine learning techniques to tackle challenging problems related to climate change.

References

[1] Kazam, Yao, Bhada-Tata, F. Van Woerden, "What a Waste 2.0: A Global Snapshot of Solid Waste Management to 2050," World Bank, 2018, pp. 3-5.
[2] United States EPA, "National Overview: Facts and Figures on Materials, Wastes and Recycling," 2016.
[3] L. Parker, "A Whopping 91 Percent of Plastic Isn't Recycled," National Geographic, 2018.
[4] United States EPA, "Food: Material-Specific Data," 2016..
[5] Jula, Megan "How Recycling Varies by Where You Live," 2019
[6] M. Koerth, "The Era Of Easy Recycling May Be Coming To An End," FiveThirtyEight, 2019.
[7] UNECE, "Methane Management," 2020
[8] United States EPA, "Overview of Greenhouse Gases," 2018.
[9] G. Thung, M. Yang, "Classification of Trash for Recyclability Status," 2016, Stanford CS229 Project Report.
[10] O. Awe, R. Mengitsu, V. Sreedhar, "Final Report: Smart Trash Net: Waste Localization and Classification," 2016, Stanford CS229 Project Report.
[11] C. Bircanoğlu, M. Atay, F. Beşer, Ö. Genç and M. A. Kızrak, "RecycleNet: Intelligent Waste Sorting Using Deep Neural Networks," Innovations in Intelligent Systems and Applications (INISTA), Thessaloniki, pp. 1-7, 2018, doi: 10.1109/INISTA.2018.8466276
[12] D. Vinodha, J. Sangeetha, B. Cynthia Sherin, M. Renukadevi, "Smart Garbage System with Garbage Separation Using Object Detection," International Journal of Research in Engineering, Science and Management 2020.
[13] D. Ziouzios, M. Dasygenis, "A Smart Recycling Bin for Waste Classification," Panhellenic Conference on Electronics & Telecommunications (PACET).pp. 1-4, 201doi: 10.1109/PACET48583.2019.8956270.
[14] EPA, "Frequent Questions on Recycling," 2018

CJSJ

Gamma-Ray Analysis of the Most Energetic Blazars to Probe the Cosmos

Yashika Batra[1], Je-Won Im[1], Nathan Nguyen[1], and Olivier Hervet[2]

[1i] Evergreen Valley High School, 3300 Quimby Rd, San Jose, CA 95148
[1ii] Choate Rosemary Hall, 333 Christian St, Wallingford, CT 06492
[1iii] Atholton High School, 6520 Freetown Rd, Columbia, MD 21044
[2] University of California Santa Cruz, 1156 High St, Santa Cruz, CA 95064

***Abstract*— Direct measurement of the Extragalactic Background Light (EBL) is difficult due to foreground emissions. An alternative method is to indirectly probe the EBL from its interaction with blazar gamma (γ) rays. The Fermi Large Area Telescope (Fermi-LAT) and H.E.S.S. collaborations proposed using a scaling factor α to normalize EBL density based on previously existing models. However, initial normalizations analyzing 10 years of data from the Fermi-LAT Fourth Source Catalog Data Release 2 (4FGL-DR2) resulted in numerous outliers, whose values differed more than 3σ from an existing EBL model. We performed a new spectral analysis on 12 years of Fermi-LAT observations, focusing on outlier and bright sources. The changes of α derived from our analysis resolve the issue for most of the outlier sources, while creating a new outlier from our "bright sources" sample. By estimating the factor α for a large number of blazars, this study will contribute to the creation of a density map of the EBL.**

I. INTRODUCTION

Proper measurements of the Extragalactic Background Light (EBL), the sum of all infrared to optical light emitted since reionization, are key to understanding the universe's makeup and evolution. Our study uses 12 years of data from the NASA space Fermi Large Area Telescope (LAT) which observes γ-rays from ~100 MeV to TeV energies. Using blazar sources from the Fermi-LAT Fourth Source Catalog Release 2 (4FGL-DR2) [1], and Fermipy, a Fermi analysis python package (Wood et al. 2017), we created optimal observed spectral energy distributions (SEDs), or plots of flux over energy. The relationship between observed and intrinsic spectra is shown in equation 1:

$$\Phi_{obs} = e^{-\alpha\tau(E,z)}\Phi_{intr}$$

Equation 1. Φ_{obs} and Φ_{intr} are the observed and intrinsic spectra, α is the normalization value, and τ(E,z) [3] is EBL optical depth.

An optimal α value is considered after performing a likelihood test with a given intrinsic spectrum [2].

II. METHODS

Fermipy analysis produces counts maps (Fig. 1), residual maps (Fig. 2), likelihood fits (Fig. 3), and SEDs (Fig. 4).

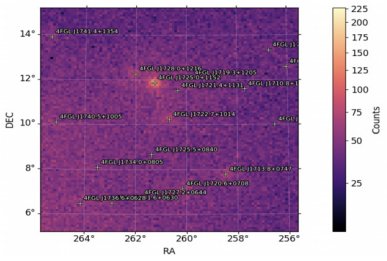


Figure 1. Photon data counts map of TXS 1720+102. The source in the center is surrounded by the region of interest, or the total area from which photon and background data is considered.

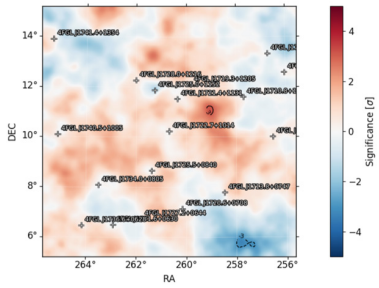


Figure 2. Residual plot of TXS 1720+102, illustrating the difference between values of the counts map and spectral models of all sources and background in the region of interest. Near zero σ represents a good fit.

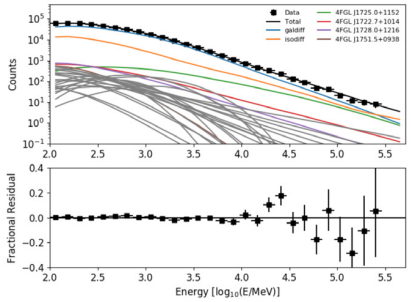


Figure 3. Likelihood fit of all sources in the region of interest. The red line shows the best spectral fit of TXS 1720+102.The top portion is a map of counts vs. energy. The bottom is a graphical residual plot.

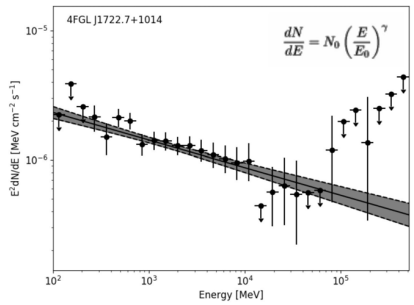


Figure 4. SED of TXS 1720+102.

One way to optimize SEDs for analysis is through bin reduction. Upper limits lack error bars, indicating a 95% chance the flux value is below the horizontal line. Having upper limits in the middle energies of an SED signifies an oversampling of flux emission, indicative of a non-optimized spectrum for EBL normalization (Fig. 4). Thus, we perform iterative bin reductions by merging coincident energy bins to remove these upper limits. Another way to optimize SEDs is through spectral model change. With 12 years of data, the default 4FGL-DR2 spectral models used for 10 years of data may not accurately describe the SEDs produced in our study and can be changed. The models mostly used were Power Law, Log Parabola, and Power Law with exponential cutoff (Eq. 2, 3, 4). Here, $\frac{dN}{dE}$ is flux density, and E is energy.

$$\frac{dN}{dE} = N_0 \left(\frac{E_0}{E}\right)^{\gamma}$$

Equation 2. Power Law. Set factors: E_0(scaling factor). Free factors: N_0 (prefactor), γ(index)

$$\frac{dN}{dE} = N_0 \left(\frac{E}{E_0}\right)^{-(\gamma+\beta \log(\frac{E}{E_b}))}$$

Equation 3. Log Parabola. Set factors: E_b(scaling factor). Free factors: N_0 (norm), γ(index 1), β(index 2).

$$\frac{dN}{dE} = N_0 \left(\frac{E_0}{E}\right)^{\gamma} \exp(\frac{E}{E_{cut}})^b$$

Equation 4. Power Law with Exponential Cutoff. Set factors: E_0(scaling factor). Free factors: N_0 (prefactor), γ(index 1), b (index 2), E_{cut}(cutoff)

With two more years of data than 4FGL-DR2, we were able to significantly improve the underlying spectral model of certain sources (Fig. 5).

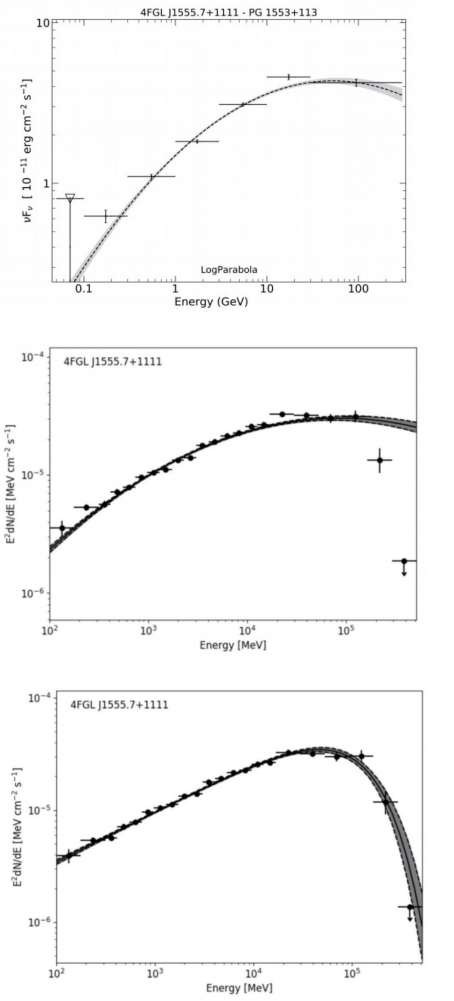


Figure 5. Changes between the 4FGL-DR2 catalog (top) and analysis using 12 years of data (bottom) for source 4FGL J1555.7+1111: bins are smaller, and a model change from Log Parabola to Power Law with Exponential Cutoff shows a much better fit of the Fermi-LAT data.

III. RESULTS AND DISCUSSION

Our analysis using 12 years of data created 54 new SEDs for 18 outlier and 36 additional bright sources.

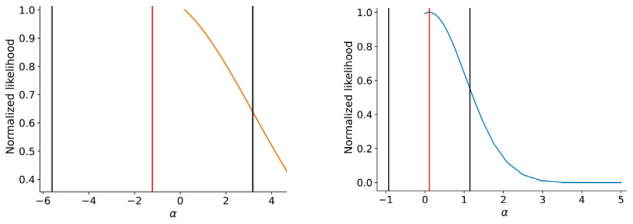


Figure 6. Likelihood profiles of α using 10 years of data (left) and 12 years of data (right) for source 4FGL J1555.7+1111. These plots are produced from our analyzed spectra by the UCSC Particle Physics team hosting our research. α deviations from normalization using 10 years of data are resolved with our analysis using 12 years of data [2].

While we were able to resolve 17 outlier sources, analysis of bright sources significantly changed the normalization value of the EBL, creating a new outlier source: one of the bright sources studied differed more than 3σ from the newly calculated normalization value. As analysis of the corrected outlier 4FGL sources reduced discrepancies between the source value and a nominal EBL model [3], it can be surmised that issues in the initial catalog caused σ-disagreement rather than the intrinsic spectra of our sources. However, causes of discrepancy for the new outlier source are unknown; problems could be in the nature of the source, or in the analysis itself. For instance, sources with specifically high variability can have a reconstructed 12-years spectra unable to show any given real state of activity, inducing bias in α values. Additionally, for certain sources, we had to change the spectral model from the original catalog description. For 4FGL J1555.7+1111, the large energy bin at 100 GeV in the 4FGL-DR2 spectrum (See fig. 5 top) minimized the real spectral curvature at high energies. This can lead to an underestimation of EBL absorption, creating an outlier source. Future work will include checking which of the mentioned possibilities may be creating the new outlier source. The results of this study will contribute to the first EBL skymap ever created from normalization and to the probing of possible anisotropies.

ACKNOWLEDGEMENTS

We thank our mentor, Dr. Olivier Hervet, and the UCSC VERITAS team for their assistance. We also give gratitude to the UCSC Science Internship Program for guiding us throughout this project, and the editors of CJSJ.

REFERENCES

[1] Ballet, J., Burnett, T.H., Digel, S.W., & Lott, B., "Fermi Large Area Telescope Fourth Source Catalog Data Release 2," *The Astrophysical Journal Supplement Series*, Volume 247, Issue 1, id.33, 37 pp., 2020.

[2] Biasuzzi, B., Hervet, O., Williams, D. A., & Biteau, J., "Normalization of the extragalactic background light from high-energy gamma-ray observations," *Astronomy & Astrophysics*, Volume 627, id.A110, 12 pp., 2019.

[3] Franceschini, A., & Rodighiero, G., "The extragalactic background light revisited and the cosmic photon-photon opacity," *Astronomy & Astrophysics*, Volume 614, id.C1, 5 pp., 2017.

[4] Wood et al., "Fermipy: An open-source Python package for analysis of Fermi-LAT Data," *Proceedings of Science*, Volume 301, 2017.

Observing the Random Dispersion of Compartmentalized Droplets in an Oil-Water Emulsion in Microgravity

William Tang[1], Jennifer Song[1], Julia Kim[2]

[1i,1ii,2]Valley Christian High School International Space Station Research Lab, 100 Skyway Dr #130, San Jose, CA 95111

***Abstract*— By injecting a droplet of oil into water, a contact surface forms between the two liquids, known as a liquid-fluid interface. The jamming of nanoparticle-ligands assemblies at this interface forms a permeable membrane. In microgravity, multiple droplets of a liquid can be suspended in another exterior liquid phase, forming a stable emulsion without coalescing of the interior droplets. The negligible force of gravity on the International Space Station (ISS) allows for the random distribution of these encapsulated droplets, whereas on Earth buoyant forces immediately affect the structures due to differences in density. On earth and the ISS, liquid-fluid interfaces were created by injecting a solution of POSS nanoparticles and silicon oil into aqueous polyacrylic acid, and their shapes were compared. In microgravity, the formation of fluid-liquid-fluid and fluid-liquid-gas structures further proved the stability of these systems in the absence of significant buoyant forces. The formation of stable, dispersed structures in liquid-fluid emulsions allows for a wide variety of applications including drug delivery, reagent encapsulation for on-demand reactive systems, and all-liquid batteries.**

I. Introduction

The contact surface between two different liquids constitutes a liquid-fluid interface. Differences in the interaction forces between the two liquids changes the interfacial potential energy as molecules are moved toward the border. Work must be done against the unbalanced interaction forces to bring the molecule to the interface. To reduce potential, liquid structures tend to minimize surface area by forming spherical domains, coalescing, or absorbing particles to the interface. In addition, surfactants, which reduce interfacial energy, are often used to stabilize smaller droplets, since the energy cost of the droplets is reduced.

Polyacrylic acid (PAA), which exhibits low binding energies and a slight negative charge, is dissolved in water and subsequently injected in a solution of silicone oil. The inherent negative charge of the water-oil interface hinders the polymer from binding to the interface; furthermore, the low-binding energy of the nanoparticles prevents the formation of an irreversible bond.

To anchor the charged PAA to the interface, positively charged POSS nanoparticles (NP) are mixed into the oil, forming NP-ligand assemblies at the interface by electrostatic attraction. The absorption of the nanoparticle-ligands at the interface, in addition to the droplet's tendency of lowering interfacial surface area, irreversibly jams the particles in place, forming a flexible, permeable structure with special chemical properties. This structure can then be modified by injecting or reducing the volume of the droplet, shining a light of a specific wavelength to weaken the membrane, or adjusting the pH level of the liquid to alter the characteristics of the structure.

*Research supported by The Quest Institute.

W. Tang and J. Song are researchers with the International Space Station Research Lab at Valley Christian High School, San Jose, CA 95111 USA. (e-mails: william.tang@warriorlife.net and jennifer.song@warriorlife.net)

II. Methods

Experimental Setup:

A solution of PAA in water is initially injected into a sealed chamber. Silica capillary tubing connected to a bi-directional pump protrudes into the chamber, functioning as an injection needle. By turning on the pump, a solution of NP and silicone oil is injected into the PAA solution, forming a spherical droplet 2 millimeters in diameter. Pictures are taken between intervals.

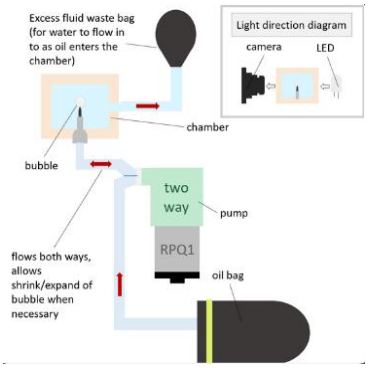


Figure 1: Experiment Block Diagram

Flight Test:

The experiment module was placed in a microgravity environment on the ISS and ran autonomously. The experiment loop consists of 7 one second injection intervals followed by 7 one second retraction intervals. Each injection interval increases by 30 milliseconds per loop, while retraction times stay constant. Photos are taken between intervals.

III. Results

For the first fourteen cycles, the injected droplet can be seen oscillating on the needle, slowly increasing in size. However, after several cycles, in which we were able to observe a slight wrinkling effect of the membrane, the droplet volume continued increasing such that it adhered to the chamber by surface tension, and the unexpected formation of different structures was observed. When the needle detached from the oil droplet, it retracted PAA solution from the surrounding chamber and injected it into the larger droplet of oil. These smaller droplets dispersed into the surrounding oil, eventually stabilizing in individual locations. Coalescence was not observed, and the droplets of PAA with diameter around 0.08 mm remain in place inside the 3-layer emulsion, encapsulated inside the oil droplet even during subsequent cycles of injection. Thus, we were able to observe the formation of water-oil-water emulsions, and the stable, undisturbed distribution of all-aqueous structures separated only by a NP-surfactant membrane. Similarly, gaseous bubbles are

injected into the oil droplet in a later phase, and the chambers of air remain encapsulated inside the oil. Some of the air bubbles do coalesce, but at a rate much slower than that on Earth, and the droplet grows as it envelopes gas. The formation of these fluid-liquid-fluid and fluid-liquid-gas structures proves the concept of stable encapsulation of both gaseous and liquid phases, which do not tend to coalesce in the surrounding liquid. The experiment demonstrates significant evidence that randomly dispersed, all-liquid structures can be stably formed microgravity due to the lack of significant

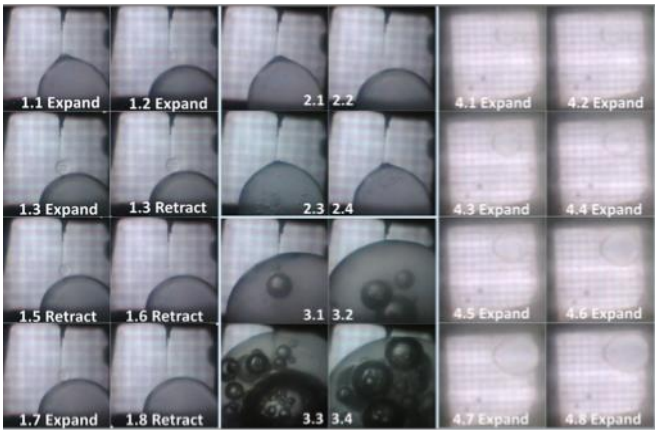


Figure 2: Pictures of the chamber from the camera.

buoyant forces arising from liquid density differences.

On the ground, we were able to obtain valuable data through testing as well. Due to gravity and density differences, buoyant force causes the water droplets to accumulate at the top of the larger host droplet and individual droplets to coalesce before a membrane is stabilized from the assembly of POSS nanoparticles and PAA at the interface. As seen from our ground tests, the silicone oil droplet, denser than the aqueous PAA solution, sagged downward, falling off the injection needle regardless of chamber orientation. The formed droplet could not reach the size of the droplet in the flight experiment, and all droplets that fell coalesced into a single oil bubble, adhering to the bottom chamber panel. Thus, due to gravity, it is extremely difficult to form the compartmentalized, all-liquid systems on the ground, or observe the characteristics of interior liquid-fluid interfaces.

We were able to evaluate our hypothesis with the data collected, and used ImageJ, an open source image processing package, to identify trends, take precise measurements, compile animations, and calculate area of the droplet, in order to convert the 1271 photos collected into a visual form.

IV. Conclusion

Originally, we wanted to observe the effect of the reduction in surface area of the droplet on the buckling of the nanoparticle surfactant assemblies. As observed in the injection/reduction phase, there were no significant wrinkles on the droplet. However, a permeable, irregular membrane can be seen in Figure 2 images 2.-2.4, suggesting that the membrane undergoes wrinkling, yet gravity causes sagging and vertical wrinkling due to the constant downward pull. As the liquid-fluid interface minimizes its free energy by decreasing total surface area and coalescing into spherical shapes, the presence of an uneven interface indicates the existence of a semi-permeable membrane.

While we had originally expected the formation of a singular liquid-fluid-liquid emulsion, we instead observed the formation of multiple layers of encapsulated liquid-fluid interfaces (figure 2, images 3.1-3.4). On Earth, the interior droplets would concentrate at the top of the host liquid, forming gradient structures based on fluid densities. On the International Space Station, the droplets stay randomly dispersed, generating a randomly mixed droplet configuration that can only be achieved in microgravity.

Nanoparticle systems at liquid-fluid interfaces cover a vast field of technological applications, including the stabilization of dispersion devices, flotation systems, encapsulation, and pharmaceutical formulations. The stability and random, undisturbed emulsion of gas and liquid droplets encased by a nanoparticle surfactant membrane demonstrates the potential of reagent encapsulation, opening a wide window of potential applications in biomedical engineering, drug delivery, and selective diffusion and reactions. The all-liquid structures could be configured to be sensitive to a certain stimulant, which breaks down the nanoparticle monolayer weakening it to allow for the selective or partial diffusion of the enclosed liquid, or the coalescing of multiple droplets into one structure. For example, the physical and chemical properties of the liquid membrane could be altered as the surrounding solution reaches a certain pH threshold, or after being exposed to certain wavelengths of light, changing the properties of interaction between structures as well as the surrounding environment. We would recommend future microgravity experiments to explore these characteristics.

The modifiable, permeable characteristics of liquid-fluid interfaces and all-liquid structures can be applied in the cosmetics industry, with compartmentalized reagents interacting only after a certain amount of time, or in the medical industry, as a chemical transport vessel breaking down under certain conditions. Studying liquid-fluid interfaces in microgravity, where the distribution of liquids is unaffected by density, allows for the observation of new characteristics of these unique structures.

Acknowledgment

We are thankful to our mentors, Marcel Marc, Howell Ivy, and Emeka Okekeocha, as well as Prof. Thomas Russell and Dr. Paul Kim at Lawrence Berkeley National Laboratories.

References

[1] Forth, J., Kim, P. Y., Xie, G., Liu, X., Helms, B. A., and Russell, T. P. (2019). Building Reconfigurable Devices Using Complex Liquid–Fluid Interfaces. *Advanced Materials*, *31*(18), 1806370.

[2] Shi, S., Liu, X., Li, Y., Wu, X., Wang, D., Forth, J., and Russell, T. P. (2018). Liquid Letters. *Advanced Materials*, *30*(9), 1705800.

[3] Lichtenhan, J. D., Pielichowski, K., and Blanco, I. (2019, October 22). POSS-Based Polymers. Retrieved April 20, 2020, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6835917/

[4] Toor, A., Forth, J., Araujo, S. B. D., Merola, M. C., Jiang, Y., Liu, X., … Russell, T. P. (2019). Mechanical Properties of Solidifying Assemblies of Nanoparticle Surfactants at the Oil– Water Interface. *Langmuir*, *35*(41), 13340–13350

A Novel Weighted Approach to Predict Protein Fold Type

Eeshwar Krishnan[1]

[1]Pennsylvania Leadership Charter School, 1585 Paoli Pike, Westchester PA 19380, USA

***Abstract*— Prediction of protein fold type is the first step in determining protein folding of amino acids. Predicting fold type is a difficult multi-class machine learning problem. Past work in predicting fold types has shown poor overall accuracy, although the methods worked well for determining some protein fold types. In this paper, we describe a novel framework to predict protein folding using a weighted approach, combining different machine learning approaches in a principled manner. This approach uses a weighted voting method to combine results from different machine learning methods to improve the accuracy of predicting fold type over individual methods. Results show an increase in the accuracy of protein fold measurements. Furthermore, the framework can be expanded to include new and emerging deep learning methods, and can serve to enable protein folding prediction.**

I. Introduction

Protein folding is a difficult and intricate problem in molecular biology. Proteins are created in unfolded, random shapes of amino acids. Through interactions with hydrogen bonding and other Intermolecular Attractive Forces (IMAFs), the protein folds into a different shape, known as the native shape. The shape of the protein helps define its function and how it interacts with the rest of the cell. Changing the folding pattern of the protein can make it become toxic or malfunction in the cell. This could lead to different conditions including Alzheimer's, Parkinson's, cystic fibrosis, and many others [1]. Understanding how these proteins fold can advance treatment of these diseases by helping the development of new medicines that have particular impacts on the cell.

However, determining folds of a protein is computationally intensive. Each protein can consist of 20 different amino acids that combine and adopt one of several trillion shapes [2]. To determine how a particular sequence of amino acids will fold, one can compare the sequence of proteins to a known set, and if the sequences are similar, the protein folds will likely be similar as well. There are many sequences that are not similar to others, and as a result the folds cannot be determined by a search. In these cases, one proposed method to determine protein folds is to first predict the fold type, such as Globin-like, Long alpha-hairpin, and Cytochrome c. After the fold type is determined, the specific fold shape can then be found through a local search [3].

The key to the success of this approach is predicting the fold type. The general problem for predicting fold type is a multi-class machine learning problem, where the goal is to label a sample into one of several defined types of fold. Several studies have attempted to use machine learning to predict the fold type of a protein molecule [4,5,6,7,8]. The results of these approaches are shown in Table 2. Each of these studies used the same set of data for training and testing, to allow for comparison of methods. While none of these approaches are accurate enough to be used in many general cases, the results show that each method worked for some subset of the cases. For example, the hierarchical classifier proposed by Lin *et al* [6] did fairly well with small protein fold types (average accuracy 70.7%) but much more poorly on membrane and cell surface protein types (41.9%). This suggested that a combination of these methods could be successful in accurately predicting many different protein folds. In this paper, we propose a weighted framework that combines different machine learning algorithms to predict fold type, and demonstrate how this approach can significantly improve accuracy over existing methods.

II. Methods

In this paper, a weighted approach is proposed to combine the output of different machine learning approaches to optimize classification performance. Based on the work of previous researchers, different machine learning approaches work well for different fold types. The idea is to weight the output of different algorithms based on their accuracy for different fold types, so that algorithms that perform well for identifying a particular fold type are weighted higher for those fold types. This framework is extendible to new classification approaches, thus allowing for continued improvement from new researchers while leveraging the strengths of existing approaches.

First, 20 different machine learning algorithms were implemented to predict fold type. These algorithms were selected to represent a wide cross-section of machine learning algorithm types. Each algorithm was developed using Weka [9], with a custom Java wrapper around the algorithm.

J48 Decision Tree	Random Committee	Weighted Instance	Bagging
Voted Perception	Replacement-Based Decision	Sequential Minimal Optimization	Additive Regression
Logistic Model Tree	Bayes Net	Partial Decision Tree	Random Subspace
Linear Regression	Multinomial Naïve Bayes	Apriori with Subgroups	Neural Network

Table 1: List of Algorithms Used

Training and test data to develop and test the model was taken from data on-line, and made freely available to researchers [12]. The 20 fold types correspond to the fold types in [3]. The training and test sets had 238 and 290 samples respectively, across the different folds. Each sample had 125 features from 6 different parameter sets: amino acid composition, predicted secondary structure,

hydrophobicity, normalized Van der Wahl volume, polarity, and polarizability. This was the same training and test data used by other researchers, and so the methods can be directly compared. Full details of the training and test set data, as well as the definitions of the folds, are available at [3]. Each algorithm was optimized using the training set. Feature selection was used to minimize the number of features per model and reduce the impact of overfitting. The final accuracy of each algorithm, defined as the number of times the algorithm selected the correct fold divided by the total number of samples, was computed on the training set.

The results from the training set confirmed the results of previous studies that different algorithms work very well for certain folds, and not others. For example, the J48 decision tree works very well to identify fold type 12 (95.9% accuracy) but very poorly for fold type 15 (7.4% accuracy). Conversely, local weighted learning works well for fold type 15 (95.8% accuracy) but poorly on fold type 12 (34.5% accuracy). In selecting the algorithms for this study, we were careful to ensure that each fold type worked well (greater than 80% accuracy) on the training set for at least one algorithm.

The results from the training set were used to develop a weighted voting system. Each fold in the test set was run against all of the algorithms. For each algorithm *i*, it computes the fold *j* that has the maximum likelihood. It then casts a vote for that particular fold. The value of the vote is weighted to the accuracy for that algorithm for that fold, $\mathbf{x}_{ij}$, as measured in the training set. After all the algorithms are run, the votes for each fold are added up, and the fold with the highest number of votes is declared the "winner" and that fold is selected.

III. Results and Discussion

The weighted voting method resulted in an overall accuracy of 83.2% on the unseen test set, where accuracy is defined as the percentage of the test set folds that were correctly classified in the appropriate fold type. The results of the weighted voting method, compared to other methods in literature using the same training and test set, are shown in Table 2 below.

Reference	Algorithm used	Accuracy
Ding *et al* [4]	Support Vector Machines and Neural Networks	20.5%
Chinnasamy *et al* [5]	Naïve Bayes	58.8%
Lin *et al* [6]	Hierarchical Classifier	60.1%
Jo *et al* [7]	Random Forest	40.8%
Gromiha *et al* [8]	Linear Regression	57.1%
Weighted Voting (this paper)	**Multiple**	**83.2%**

Table 2: Accuracy of Weighted Voting Versus Other Methods on Protein Fold Prediction Using Same Test Set

The results clearly demonstrate that a weighted voting method can substantially improve the accuracy of predicting a fold given parameters of that fold over current published methods. The accuracy of this method suggests that it can be used as a first step to determining the specific protein fold for a set of amino acids. The approach leverages the ability of certain approaches to work well for certain fold types, and combines them in a principled manner to optimize classification performance.

Furthermore, this approach can be expanded to leverage new methods to classify protein folds. Recently, Google has recently published results on using a deep learning approach, AlphaFold, to predict protein folds [11,12]. Applying these newer deep learning method methods, and combining them with current approaches, could further enhance the accuracy of the predictions. By measuring the accuracy of new approaches, such as Google's AlphaFold, on each fold type, these newer methods can be incorporated into this voting scheme. The approach described in this paper thus provides a framework for incorporating new predictive models to further improve performance.

Determining fold type is just a first step in predicting the overall protein fold. If a reliable method can be developed to address this problem, a new generation of treatments can be developed for a host of diseases, and bring hope to many people.

References

[1] Selkoe, Dennis J. "Cell Biology of Protein Misfolding: The Examples of Alzheimers and Parkinsons Diseases." *Nature Cell Biology*, vol. 6, no. 11, 2004, pp. 1054–1061.

[2] Everts, Sarah. "Protein Folding: Much More Intricate than We Thought." CEN RSS, cen.acs.org/articles/95/i31/Protein-folding-Much-intricate-thought.html.

[3] Dubchak, Inna, et al. "Recognition of a Protein Fold in the Context of the SCOP Classification." *Proteins: Structure, Function, and Genetics*, vol. 35, no. 4, Jan. 1999, pp. 401–407.

[4] Ding, C H, and I Dubchak. "Multi-Class Protein Fold Recognition Using Support Vector Machines and Neural Networks." Bioinformatics (Oxford, England), U.S. National Library of Medicine, Apr. 2001.

[5] Chinnasamy, A, et al. "Protein Structure and Fold Prediction Using Tree-Augmented Naive Bayesian Classifier." Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing, U.S. National Library of Medicine, 2004.

[6] Lin, Chen, et al. "Hierarchical Classification of Protein Folds Using a Novel Ensemble Classifier." PloS One, Public Library of Science, 2013.

[7] Jo, Taeho, and Jianlin Cheng. "Improving Protein Fold Recognition by Random Forest." BMC Bioinformatics, BioMed Central, 2014.

[8] Gromiha, M Michael, et al. "FOLD-RATE: Prediction of Protein Folding Rates from Amino Acid Sequence." Nucleic Acids Research, Oxford University Press, 1 July 2006.

[9] Holmes, Geoffrey; Donkin, Andrew; Witten, Ian H. (1994). "Weka: A machine learning workbench". Proceedings of the Second Australia and New Zealand Conference on Intelligent Information Systems, Brisbane, Australia.

[10] http://www.nersc.gov/~cding/protein

[11] "AlphaFold: Using AI for Scientific Discovery." Deepmind, deepmind.com/blog/article/AlphaFold-Using-AI-for-scientific-discovery

[12] https://www.bloomberg.com/news/articls/2020-11-30/deepmind-s-alphafold-corssesthreshold-in-solving-protein-riddle

Predicting Migraines with Machine Learning and Feature Selection

Rebecca Zhu[1], Yan Luo[2]

[1]Nashua High School South, 36 Riverside Street, Nashua NH, 03062, USA
[2]University of Massachusetts Lowell, 220 Pawtucket Street, Lowell MA, 01854, USA

***Abstract*—Often seen as severe headaches, migraines lower the quality of life for patients. To combat ineffective monitoring applications that can only provide information about previous attacks and not future occurrences, multiple types of machine learning models such as the Random Forest (RF), Logistic Regression (LR), and Support Vector Machine (SVM) models were trained to predict migraines occurrences. The data used was augmented by employing Synthetic Minority Oversampling Technique (SMOTE). By using feature selection methods, more discriminative features were selected to train the models. The RF model outperformed existing models with a classification accuracy of 0.9924 on the testing data.**

I. Introduction

Migraines are the third most prevalent and sixth most disabling illness in the world [1]. Frequently, migraines are associated with sensitivity to light and loud sounds, nausea, and vomiting [1]. Currently, many migraine tracking diaries only allow patients to track retrospective information and provide no information to prevent future migraines [2]. However, machine learning models such as the Random Forest (RF), Logistic Regression (LR), and Support Vector Machine (SVM) models can be utilized to predict future migraine occurrences, allowing patients to take preventative measures. These models were created using Python and the scikit-learn library [3].

II. Methods

The dataset used was called "Analysis of Trigger Factors in Episodic Migraineurs Using a Smartphone Headache Diary Applications," and was published on Plos One in 2016 [4]. There were numerous migraine triggers or features that could increase a chance of a migraine, such as "stress," "excess_sleep," "smoking," etc. The dataset contained 4,679 entries including both migraine and no-migraine days.

After preprocessing the data, several feature selection methods such as the Random Forest Feature Importance attribute (RFFI), Principal Component Analysis (PCA), Recursive Feature Elimination (RFE), and the novel Jenson-Shannon Divergence ranking values (JSD) were applied to the models to select more discriminative data to train on by dropping features with lower importances [5]. Additionally, the dataset was severely imbalanced, with 4,243 no-migraine occurrences and 336 migraine occurrences. To mitigate the bias in the dataset, Synthetic Minority Oversampling Technique (SMOTE) was utilized, where the number of under sampled instances of a class is increased using the k-nearest neighbors algorithm [6].

The models were evaluated on 20% of the dataset and the other 80% of the dataset was used to train the models. The metrics used to evaluate the results of the research and their formulae are shown below in Equation 1:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN},\ F1 = \frac{2*precision*recall}{precision+recall} \quad (1)$$

where *TP*=true positive, *TN*=true negative, *FP*=false positive, and *FN*=false negative migraine occurrences. F1 scores were utilized because accuracy by itself does not provide detailed insight about the performances especially if the dataset is imbalanced. Thus, F1 scores are calculated based on precision and recall metrics. Additionally, macro-average scores were used because they do not take class imbalance into consideration, which allowed the results to be compared before and after applying SMOTE. *Figure 1* depicts the mechanism used for all three ML models.

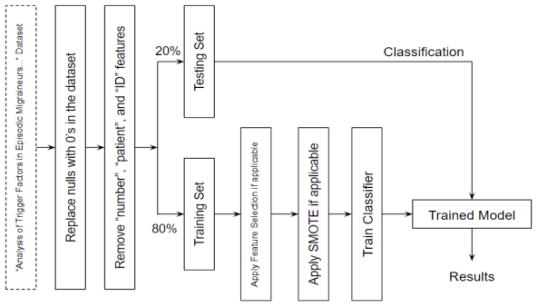


Figure 1. The proposed approach for the ML models to classify migraines.

III. Results and Discussion

The hyperparameters for each ML model were found through grid search on the training data. The classification metrics were found on the testing data. Specifically, the RF model contained 30 trees and the SVM model used a polynomial kernel function. In Table 1 without SMOTE, it was evident that the RF and SVM models outperformed the LR model, as the RF and SVM models both achieved an accuracy of 0.99017. However, the RF model achieved a higher macro-average F1 score than the SVM model because it predicted fewer false negatives.

TABLE I. Classification Metrics for ML Models

ML Model SMOTE?	***RF Model*** No	***RF Model*** Yes	***LR Model*** No	***LR Model*** Yes	***SVM Model*** No	***SVM Model*** Yes
Accuracy:	0.9902	**0.9913**	0.9571	0.9673	0.9902	0.9913
Macro F1 Score:	0.9707	**0.9741**	0.8918	0.8884	0.9704	0.9740

While the RF and SVM models performed well before applying SMOTE, there was an inherent bias because the number of no-migraine occurrences was more than 12 times the number of migraine occurrences. To see if the imbalance in the classes caused false negatives, SMOTE was applied. Without SMOTE, it seemed obvious that the model could have kept predicting no-migraine and getting a high accuracy. Thus, other metrics such as F1 scores were considered in addition to accuracy. After applying SMOTE, there were 3,409 instances of the migraine and no-migraine class each. However, SMOTE was only applied on the training data so that the testing data is still the original data. As shown in Table 1 with SMOTE, all three models had an improvement

in accuracy and F1-score. Because macro-average scores do not take class imbalance into consideration, the improvement in F1 scores means the RF and SVM classifiers improved in the ability to predict fewer false negatives. Additionally, the RF model performed the best out of the three models, achieving an accuracy and macro-average F1 score of 0.9913 and 0.9741 respectively with SMOTE.

Since the RF model achieved the highest result, it was used to evaluate the four feature selection methods. Running feature selection methods on the RF classifier without SMOTE was not performed because the RF classifier with SMOTE already produced better accuracy and F1 metrics, which were evaluated before feature selection.

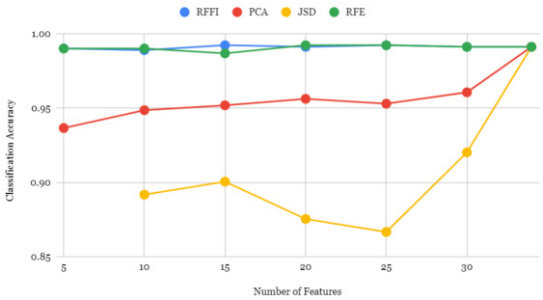


Figure 2. Comparing different feature selection methods using the RF model with SMOTE.

Figure 2 shows the classification accuracy of each method according to different numbers of features, namely 5, 10, 15, 20, 25, 30, and 34 or all features. It is observed that feature selection with PCA and JSD degrade in performance when fewer features are selected. The best performance with the RF model using SMOTE and feature selection was when the top 15 features were selected using the RFFI method, which achieved an accuracy of 0.9924, an increase from before applying feature selection. When 15 of the most discriminative features were selected by the RFFI method, the RF model achieves the best results. Therefore, around half of the features in the dataset are crucial in training the RF model for a more accurate migraine prediction.

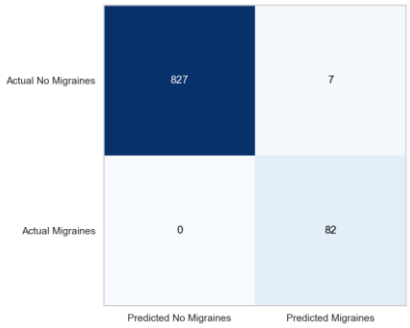


Figure 3. The confusion matrix for the RF model with SMOTE and top 15 features selected by RFFI.

Figure 3 shows the confusion matrix (CM) or the distribution of TN, TP, FP, and FN for migraine prediction. The top left and bottom right corner depict the number of correctly predicted migraine and no migraine occurrences. After applying SMOTE and the RFFI feature selection method, however, the model predicted zero migraines as no-migraines, which meant the false-negative rate fell to zero. Additionally, the model only inaccurately predicted seven no-migraines as migraines, which was similar to the metric before applying SMOTE and feature selection. Thus, SMOTE and RFFI were able to decrease the false-negative rate while maintaining the false-positive rate, which meant the overall classification performance increased.

Figure 4 shows the importance of the top 5 features selected by RFFI, which included "nausea_vomiting," "helping_factors," "sound_sensitivity," "light_sensitivity," and "rest." The most important factor or "nausea_vomiting" had an importance of 0.4 evaluated by the RFFI method, with the second most important feature with an importance of 0.19. These importances measure the contribution of the features to the classification result.

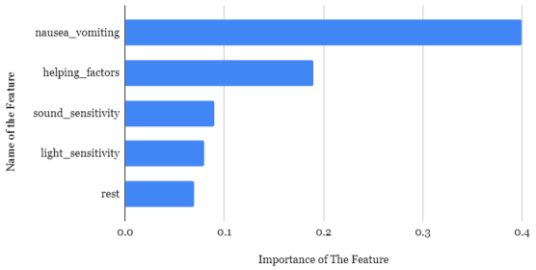


Figure 4. The top 5 features given by RFFI attribute feature selection.

IV. Conclusion

A demand exists for accurate migraine prediction through ML. All models achieved high results, but the best model was the RF model with SMOTE and the top 15 features from RFFI. This model achieved an accuracy of 0.9924 and a macro-average F1 score of 0.9774, which outperformed existing models on the standard dataset. The research showed that SMOTE improved the results of all classifiers and the RFFI feature selection method worked best, which decided the most important feature in predicting migraines was nausea or vomiting. In the future, an application where users can plug their symptoms, health factors, and biometrics could be created where the ML algorithm could predict in real-time whether or not a patient would get a migraine.

V. Acknowledgements

I would like to thank Professor Yan Luo from UMass Lowell's ECE Department for his valuable feedback and help.

References

[1] Migraine Research Foundation. (n.d.). Migraine Facts. Retrieved from https://migraineresearchfoundation.org/about-migraine/migraine-facts

[2] Migraine Buddy. (n.d.). Migraine Buddy. Retrieved from https://migrainebuddy.com/

[3] Scikit-learn: Machine Learning in Python, Pedregosa *et al.*, JMLR 12, pp. 2825-2830, 2011.

[4] Min Kyung Chu, J.-M. K. (n.d.). Analysis of Trigger Factors in Episodic Migraineurs Using a Smartphone Headache Diary Applications. Retrieved from https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0149577

[5] Karl Pearson F.R.S. . (1901). LIII. On lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, *2*(11), 559–572. https://doi.org/10.1080/14786440109462720

[6] Nitesh V. Chawla, Kevin W. Bowyer, Lawrence O. Hall, and W. Philip Kegelmeyer. 2002. SMOTE: synthetic minority over-sampling technique. J. Artif. Int. Res. 16, 1 (January 2002), 321–357.

[7] Google Research. (2017, April 6). Federated Learning: Collaborative Machine Learning Without Centralized Training Data. Retrieved from https://ai.googleblog.com/2017/04/federated-learning-collaborative.html

The Remediation of Wastewater Using a Novel Microbial Fuel Cell with Optimized Electricity Generation and an Algae Bioreactor

Sanjna Kedia[1], Emily Ma[2], and Alison Huenger[3]

[1,2,3]Manhasset High School, 200 Memorial Pl, Manhasset, NY, 11030

Abstract- The USA generates the maximum amount of wastewater per capita across the world, and according to the EPA, approximately $25 billion is annually spent on its treatment. In addition to the high costs, high residues of nitrogen and phosphorus are found in the remediated clean water. The purpose of this study was to create a scalable, novel microbial fuel cell (MFC)/algae bioreactor that would be more efficient in pollutant removal as well as energy consumption than current aeration technologies. The MFC consists of two chambers (cathode and anode) separated by a Nafion membrane. The control treatment, aeration, and MFC took 11 days, 3 days, and 0.9 days, respectively to remediate the water (90% dissolved oxygen increase). Additionally, the MFC was able to generate electricity at a sustainable voltage (0.62 V max). Anabaena biomass increase in the algae bioreactor effectively reduced nitrate levels. As shown in this study, MFC treatment holds promise for a more electrically efficient, time efficient and cost-efficient method for treating wastewater.

I. Introduction

3% of the United States' electrical load is due to wastewater treatment, which is equivalent to the electricity use of 9.6 million households [1]. The cost and electrical load of treating wastewater is largely due to the aeration technology as it consumes 45-75% of wastewater treatment plant costs [2]. Aeration reactors treat wastewater by bringing the pollutants in contact with air, causing biodegradation of the contaminants. To solve the problem of energy efficiency, microbial fuel cells (MFCs) were used to treat wastewater, as MFCs are devices that use bacteria as catalysts to oxidize organic and inorganic matter and generate current through the transfer of electrons from anode to cathode [3]. Although MFCs have been studied for many years, they have not been put into practical use due to hindering points such as cost and durability, with materials like the forward osmosis membrane requiring very specific storage conditions and carbon paper electrodes being too flimsy.

This study created an optimized design of the most commercially viable materials to construct the MFC. However, MFCs are not completely efficient in removing nitrates and phosphates from wastewater. These pollutants are linked to zones of eutrophication and hypoxia, or "dead zones" where life cannot be sustained due to reduced oxygen levels because of algal blooms and excess nutrient deposits[4]. Dissolved oxygen is an important factor to measure water quality as it directly indicates an aquatic resource's ability to support aquatic life. To further remove nitrates from the water, algae treatment of the water was undertaken as algae contain a high affinity transport system NrtABCD permease, which allows nitrate to enter the algae and be converted into ammonium, which reduces pollutant levels in the cleaned wastewater [5].

So, the purpose of this study was to engineer a novel, scalable, cost-efficient, and time efficient MFC with optimized electricity generation and enhanced pollutant removal with an algae bioreactor, whose efficiency should be higher than current aeration and control treatments.

II. Methods

In phase one, the control, aeration, and MFC were run until a 90% increase in dissolved oxygen was observed. The control treatment consisted of still artificial wastewater, made according to the OPEC formula, and the aeration reactor consisted of artificial wastewater circulated by an aquarium diffuser. *E. coli* K-12 (Carolina Biological) served as the exoelectrogenic microbial agent for the MFC (figure 1), which was constructed by joining two, 300mL plastic containers together with a Nafion membrane. The anode compartment consisted of carbon felt electrode and was fed wastewater solution inoculated with 15 colonies of *E. coli* K-12. The cathode compartment consisted of carbon cloth electrode with platinum as catalysts and 5% sodium chloride as a draw solution. Dissolved oxygen (DO) and electricity generation were measured using a Vernier dissolved oxygen probe and Vernier energy sensors, respectively.

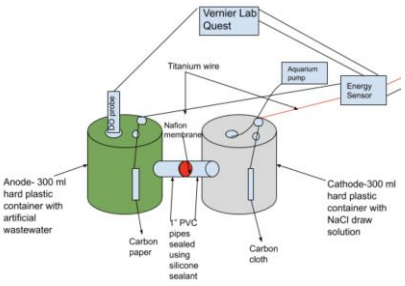


Figure 1: Schematic of MFC *(by authors)*

Figure 2: Algae bioreactor setup *(by authors)*

In phase 2, an algae bioreactor was constructed using plastic water bottles containing *Anabaena* algae solution (Carolina Biological), plus the remediated wastewater from phase one (figure 2) to further remediate nitrates via biomass increase over one week.

III. Results/Discussion

The purpose of this study was to determine the efficiency of a novel MFC to treat wastewater in terms of electricity generation and pollutant removal. The control, aeration, and MFC reactor took 11 days, 3 days, and 0.9 days, respectively, to observe a 90% increase in dissolved oxygen levels (figure 3). This increase indicated that enough oxygen had reentered the water to combat the effects of hypoxia and eutrophication. The MFC remediated the wastewater significantly faster than both the control and aeration, due to the Nafion membrane, which filtered out particulates. This membrane, in addition to the *E. coli* K-12, transferred the hydrogen ions to the cathode to combine with the oxygen molecules, facilitating the transfer of oxygen from the air to the water, increasing the DO in the anode chamber.

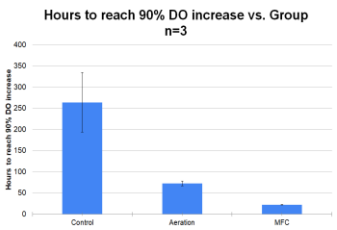


Figure 3: The amount of time, in hours, it took the Control reactor, Aeration reactor, and Microbial Fuel Cell to reach a 90% increase in DO from the initial readings. *(graph by authors)*

The MFC group showed the capability of producing an increased voltage of 0.126V on average (figure 4), proving *E. coli* K-12 to be a valid exoelectrogenic microbial agent which concurs with Sugnaux et. al (2013) [6] as it created an electrical current between the anode and cathode, by oxidizing the organic substrate of the wastewater. The wastewater in the anode chamber of the MFC is oxidized by the *E. coli* K-12 into carbon dioxide and hydrogen ions. The hydrogen ions pass through the membrane to the cathode chamber, where it combines with oxygen molecules from the air to produce pure water molecules. The overall reaction ends up converting the organic substrate and oxygen into carbon dioxide, pure water, and electricity.

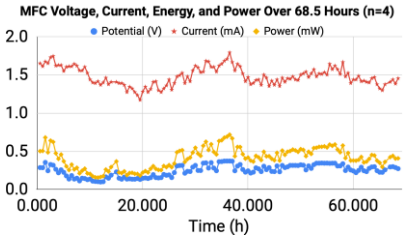


Figure 4: Novel MFC testing over 68.5 hours for variables including voltage (blue), current (red), and power (yellow). All variables experienced fluctuation, reaching a maximum at around hour 35 of operation. *(graph by authors)*

The average efficiency of the MFC was 75% which is comparably larger than the standard 40% efficiency of aeration reactors, as aeration requires inputted power from an aquarium pump, which is lost as thermal energy to the environment. Additionally, it does not produce any power of its own. However, the MFC used the energy from the aquarium pump, while also generating its own power. At hour 35 of (figure 5), a 135% efficiency was observed, meaning the MFC was producing more power than it consumed.

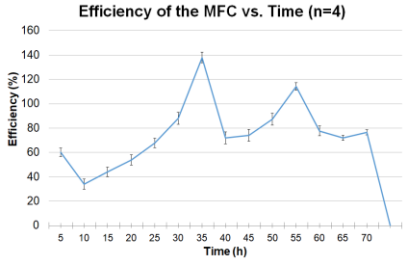


Figure 5: The efficiency of the MFC system as time increased from 0 to 70 hours. Energy efficiency was calculated using the equation: P_{output}/P_{input}. *(graph by authors)*

The microbial fuel cell is 42.72% cost efficient while the aeration reactor is 29.48% cost efficient (figure 6). The MFC cost $10.68 to construct while the aeration reactor cost $17.69. The increased cost efficiency of the MFC is attributed to carbon felt electrode and commercially viable Nafion membrane used in this device. The MFC is a more practical instrument to use in wastewater treatment, as it can help reduce electricity costs, and save money spent in the water sector, while also producing a greater power efficiency.

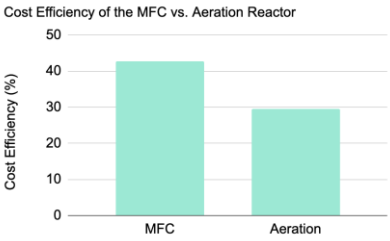


Figure 6: Cost efficiency (%) of the Microbial Fuel Cell compared to that of the Aeration Reactor. *(graph by authors)*

On day one of testing the algae bioreactor, nitrate levels began at 75mg/L, and decreased to 23.5mg/L after the full seven-day period (figure 7). This decrease was due to the growth of algal biomass in the water as algae used the enzyme nitrogenase to convert nitrates into atmospheric nitrogen by consuming the pollutant. The reduced nitrate level achieved by the bioreactor may be suitable for other aquatic species, such as fish, to live in, as the nitrate safe limit for such species is 25mg/L [7].

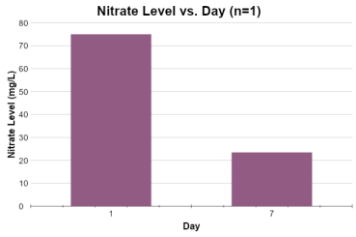


Figure 7: Nitrate removal by the algae bioreactor over a span of 7 days. 10 mg/L of nitrate is considered a safe drinking water standard [7]. *(graph by authors)*

IV. Conclusions

The alternative hypothesis was accepted as the rate of the remediation of wastewater (90% DO increase) was significant in the MFC as compared to the control group and aeration group. The MFC design was able to produce electricity indicating a preservation of energy through wastewater treatment. As a result, the MFC indicated a higher energy efficiency in treating wastewater as compared to current aeration technologies. In the future, a larger scale model of the MFC can be tested to observe the efficacy of this prototype design at the industrial level. This novel MFC has a patent pending (application number: 62987455). Moreover, the electrodes and Nafion membrane used were determined to be the most cost and energy efficient as well as readily available. Hence, this design is scalable, commercially viable and capable of being used in industry which reduces the burden on fossil fuels, reduces the time taken to remediate wastewater, and mitigates the increased electricity consumption projection over the following decades.

Acknowledgments

We would like to thank Ms. Alison Huenger, our science research teacher, for all her help and guidance throughout the research process.

References

[1] McCarty, Perry L., et al. "Domestic Wastewater Treatment as a Net Energy Producer–Can This Be Achieved?" *Environmental Science & Technology*, vol. 45, no. 17, 2011, pp. 7100–7106.doi:10.1021/es2014264

[2] Huggins, Tyler, et al. "Energy and Performance Comparison of Microbial Fuel Cell and Conventional Aeration Treating of Wastewater." *Journal of Microbial & Biochemical Technology*, 2013, doi:10.4172/1948-5948.s6-002

[3] Logan, Bruce E, et al. "Microbial Fuel Cells: Methodology and Technology†." *EnvironmentalScience and Technology*, vol. 40, no. 17, 7 June 2006, pp. 5181–5192., doi: 10.1002/9781118255728.ch7

[4] "Clean Watersheds Needs Survey." *EPA*, Environmental Protection Agency, 1 Apr. 2019, www.epa.gov/cwns.

[5] Hu, Qiang, et al. "Removal of Nitrate from Groundwater by Cyanobacteria: Quantitative Assessment of Factors Influencing Nitrate Uptake." *Applied and Environmental Microbiology*, vol. 66, no. 1, 2000, pp. 133–139., doi:10.1128/aem.66.1.133-139.2000.

[6] Sugnaux, Marc, et al. "Probing Electron Transfer with Escherichia Coli: A Method to Examine Exoelectronics in Microbial Fuel Cell Type Systems." *Bioresource Technology,* vol. 148, 2013, pp. 567-573., doi: 10.1016/j.biortech.2013.09.004

[7] "Ambient Water Quality Value for Protection of Sources of Potable Water." *NEW YORK STATE - HUMAN HEALTH FACT SHEET* -, EPA , 1991,www.epa.gov/sites/production/files/2015-06/documents/ny_hh_644_w_03121998.pdf

SARS-CoV-2 Variant Analysis

Henrik Torres[1], Amith Vasantha[2], Helen Chow[3], Gepoliano Chaves PhD[4]

[1]Choate Rosemary Hall, 333 Christrian Street, Wallingford, CT 06492
[2]Basis Independent Silicon Valley, 1290 Parkmoor Ave, San Jose, CA 95126
[3]Lowell High School, 1101 Eucalyptus Dr, San Francisco, CA 94132
[4]University of Chicago, 5801 S Ellis Ave, Chicago, IL 60637

***Abstract*— The SARS-CoV-2 virus started the novel coronavirus pandemic. SARS-CoV-2 is an RNA virus that causes infection through the binding of the virion's spike protein to a cell's ACE2 receptor. The SARS-CoV-2 virion cleaves its way into the cell and deposits its RNA genome that hijacks the cell's RNA replication system to produce more virions. During replication, genetic variance arises through single nucleotide polymorphisms (SNPs) that can enable a zoonotic jump or affect the transmissibility or lethality of a virus [3]. Our research focused on studying these SNPs from collected FASTQ and FASTA files of human, pangolin, and bat SARS-CoV-2 genomes on online databases such as the NCBI SRA Browser and GISAID and running files through variant call pipelines. Our results confirmed SNP frequencies at locations in the genome that matched those of Yin [1]. Genomic comparison of SARS-CoV-2 between the humans, bats, and pangolins showed a distinct mutation at location 23403 on the spike protein that was only present in humans, possibly a mutation that facilitated a zoonotic jump. Regional frequency analysis of collected samples showed regional clustering and similarities. The results from our research further existing knowledge of the SARS-CoV-2 virus and can be further expanded upon to create regional vaccines specifically tailored to mutations that affect certain protein mechanisms.**

I. INTRODUCTION

The coronavirus pandemic has caused an unprecedented public health crisis with the potential to worsen if mutations were to arise that increased transmissibility or infectivity. This novel coronavirus is named SARS-CoV-2, and it comes from a family of coronaviruses that cause a respiratory infection. Past coronavirus outbreaks have been SARS in 2003 and MERS in 2012 [1]. However, the spread of SARS-CoV-2 has caused a pandemic of proportions last seen only in 1918 with the influenza. Proliferated spread of SARS-CoV-2 has led to more than 22 million confirmed cases and 777,000 deaths as of August 18, 2020 [2]. SARS-CoV-2 is an RNA virus with a single-stranded positive RNA genome. Structural proteins of SARS-CoV-2 virions include the spike protein, envelope protein, membrane protein, and nucleoprotein. To infect a cell, it uses its spike protein, located on the outer membrane protein, to bind to a cell's ACE2 receptor. SARS-CoV-2 cleaves its way into a cell with this binding mechanism and deposits its genetic material to hijack the cell's RNA replication mechanisms [3].

Within replication mechanisms needed to produce more SARS-CoV-2 virions occur mutations known as single nucleotide polymorphisms (SNPs). Our research focused on identifying these frequent SNPs, understanding the regional differences of SARS-CoV-2 variants, and searching for possible variants indicating a zoonotic jump (the transmission of a virus between species).

II. METHODS

Data was collected from the NCBI SRA Database and the GISAID Database in the form of FASTQ files and FASTA files, respectively. FASTQ files and FASTA files both function to store genomic information, with the important distinction being that FASTA files store a nucleotide sequence, while FASTQ files add quality scores to each nucleotide. Several FASTQ and FASTA files were collected from NCBI and GISAID with each continent except Antarctica being equally represented. To analyze the FASTQ and FASTA files, we used a variant calling pipeline (a shell script) on the University of California, Santa Cruz Hummingbird Server. Our pipeline aligned each genome file with the established SARS-CoV-2 reference genome from the UCSC Genome Browser, NC_045512, and stored the results in a Binary Alignment Map (BAM) file using Picard; then, we used GATK (a software supplied by Broad Institute along with Picard) to find SNPs and store them in a Variant Calling Format (VCF) file using a tool named Picard; then Picard marked duplicates and constructed a BAM Index. From the BAM Index, GATK (a genomic analysis tool) realigns the indels, calls and filters variants, and produces a Base Quality Score Recalibration. This pipeline was used for analyzing Sequence Read Archive (SRA) files from the NCBI SRA Database.

For the FASTA files from the GISAID Database, a simpler version of the aforementioned pipeline was used that realigned reads through Burrows-Wheeler Aligner (BWA), converted BAM alignment of the files to Sequence Alignment Map (SAM) alignment through Samtools, created .vcf (variant call format) files with Samtools, and marked SNPs and indels with BCFTools.

III. DATA ANALYSIS

After running the pipelines on the UCSC Hummingbird Server, we gathered the results in a spreadsheet and calculated the frequencies and regional frequencies (Figure 1). We constructed a Meta-Analysis Table (data not shown) containing information on the region of origin of the virus, size of the sequencing file analyzed, when the database was the Gene expression Omnibus and sequencing strategy adopted. The purpose of this table is to keep stable information about the virus for future reference. We also constructed a Frequency Table which we used to compute frequencies of different viral genotypes per different geographic region. This table also allowed execution of the pipeline for extraction of the different genotypes present in the files analyzed.

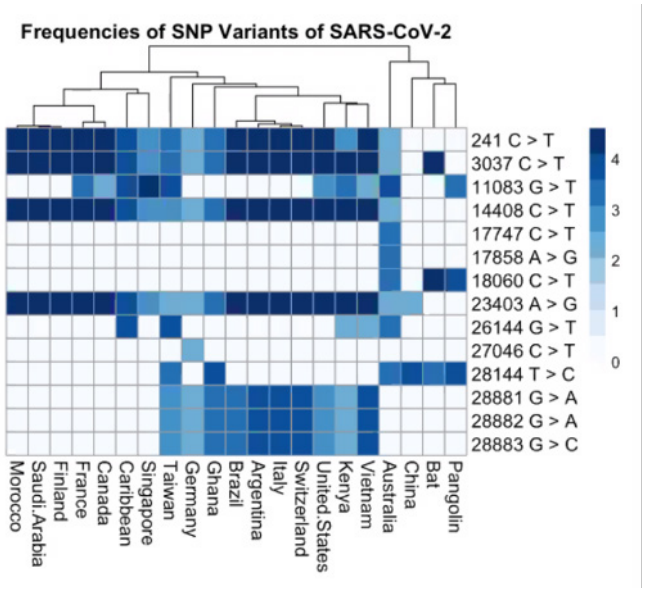


Figure 1. **Comparison of FASTA sequences from human samples in different regions around the world, bat samples, and pangolin samples.** The FASTA sequences were gathered from the GISAID Coronavirus Database and run through the variant call pipeline. The figure shows regional clustering based on the frequencies of SNP mutations. Samples taken from different countries become more genetically dissimilar as the virus traveled further from the epicenter.

IV. RESULTS AND DISCUSSION

The present COVID-19 pandemic highlighted the need for investing in understanding viral diseases, their spread, and the need for common hygiene practices across countries. Other measures that might become more important from here on is characterization of the genomic variants associated with severity of COVID-19 and incidence in certain geographic locations. Here, we compared FASTA sequences isolated from SARS-CoV-2 from across the globe. We found several overlaps between our own calculated frequencies and the frequencies found by Yin (2020). With these similar results, we were able to confirm their findings and take the research one step further by comparing the SARS-CoV-2 genomes of the samples from intermediate hosts to samples from humans. In Figure 1, we mapped the regional clustering of the different variants of SARS-CoV-2. As expected, they clustered regionally. As the virus travelled further from the epicenter, genetic variation increased and began to cluster regionally (Figure 1).

Additionally, we found that a mutation in position 23403 was present only in SARS-CoV-2 sequences infecting humans and not in bats or pangolins (Figure 1). This finding agrees with the notion that a different strain of the virus underwent mutation in order to infect humans. Furthermore, mutation in position 14408, which seems to be related to mutation 23403, was also identified in our study, present in all human sequences found infecting humans, except the Chinese samples (Figure 1). Additionally, we found that mutations in SARS-CoV-2's helicase (from SNPs in positions 17747 and 17858 of the genome) only occur in samples collected from Australia, and mutations in SARS-CoV-2's membrane glycoprotein (from SNPs in position 27046) only occur in samples from Germany (Figure 1).

Much is still unknown about SARS-CoV-2 and the disease it causes, COVID-19, but the research grows each day. Moving forward, further research can be conducted on how exactly SNPs in different regions of the genome alter protein structure and function. We have already found that the mutation at position 23404 may have facilitated a zoonotic jump, but more research must be done on this topic to confirm the claim. Additionally, more research must be done to determine why our heat map showed a closer clustering between bat sequences and human sequences because this may be an indication that they were not the original hosts of the virus. It is only a matter of time before all the SNPs and new arising SNPs are understood concretely. Given that there are more mutations in the genome of SARS-CoV-2 that fundamentally change how the virus operates in different regions, in the future scientists would be able to administer the most effective vaccines or antibody therapies for each region to prevent the spread of COVID-19.

V. ACKNOWLEDGEMENTS

We would like to acknowledge our mentor Dr. Gepoliano Chaves, PhD, for his endless guidance on this project and commitment to work. We would also like to thank the Science Internship Program (SIP) at the University of California, Santa Cruz for facilitating this research opportunity and providing a dedicated team for all participating interns. Additionally, we would like to acknowledge the NCBI SRA Database, the GISAID Database, and the submitting labs and hospitals around the world; this research would not have been possible without their work.

VI. REFERENCES

[1] Yin, C. (2020). Genotyping coronavirus SARS-CoV-2: Methods and implications. Genomics, 112(5), 3588-3596. https://doi.org/10.1016/j.ygeno.2020.04.016

[2] WHO Coronavirus Disease (COVID-19) Dashboard. (n.d.). Retrieved August 31, 2020, from https://covid19.who.int/

[3] Walls, A. C., Park, Y.-J., Tortorici, M. A., Wall, A., McGuire, A. T., & Veesler, D. (2020). Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. Cell, 181(2), 281-292.e6. https://doi.org/10.1016/j.cell.2020.02.058

Deep Brain Stimulation in the Cortico-Striato-Thalamo-Cortical Pathway and its Effect on Obsessive-Compulsive Disorder

Arya Khokhar[1]

[1]Homestead High School, 21370 Homestead Rd, Cupertino, CA 95014, USA

***Abstract*- Obsessive-compulsive disorder (OCD) is a neuropsychiatric disorder in which repetitive behaviors are done to relieve anxiety caused by repeated and intrusive thoughts. About 20% of OCD patients remain resistant to therapies and medications and are linked to suicidal behavior and lack of social functioning. Deep brain stimulation (DBS) has been considered as a last-resort solution for these patients. Recently, neuroimaging techniques have shown significant differences in the activity of the cortico-striato-thalamo-cortical (CSTC) pathway in OCD patients, supporting the first studies of DBS in the anterior limb of the internal capsule (ALIC), which is a part of this pathway. Since then, studies have expanded DBS into other locations of the CSTC pathway. With all these different regions being studied, many patterns have been found. However, as each location has a different degree of efficiency in each trial, the final goal should be to be able to determine which location will be most beneficial for patients. The purpose of this paper is to compare the studies and effects of DBS on OCD patients in varying parts of the CSTC pathway and discuss the goals and experimental setups of future studies to determine the best combination of stimulation parameters and DBS locations for patients.**

I. INTRODUCTION

Obsessive-compulsive disorder (OCD) is a disabling neuropsychiatric disorder that has a lifetime prevalence of approximately 2.3% within our population.[7] It is characterized by the presence of obsessions, which are persistent and uncontrollable thoughts or impulses, and compulsions, repetitive behaviors done to diminish the anxiety and discomfort associated with the obsessions. Effective treatments for OCD include cognitive behavioral therapy (primarily exposure and ritual prevention) and medications (mainly antidepressants such as selective serotonin reuptake inhibitors (SSRIs) and clomipramine, as well as antipsychotics that manipulate the transmission of monoaminergic neurotransmitters). Unfortunately, despite this assortment of treatment options, 40-60% of patients do not respond to SSRIs and 10-20% of patients remain resistant to all therapies.[3] These patients are diagnosed with treatment-resistant OCD and it is in these cases that deep brain stimulation (DBS) can be considered as a last resort treatment option.

DBS is a neurosurgical technique that uses electrical current delivered to specific locations of the brain through the means of implanted electrodes to regulate abnormal neural activity.[22] While the precise mechanism of DBS is unknown, there is evidence that shows DBS exerts its effect through both the activation and inhibition of brain areas by stimulating positive and negative feedback loops.[7] Via a pulse generator embedded in the chest and a connecting cable that runs under the skin, the electrodes are supplied power. Clinicians then set parameters that determine how strong of an impulse is generated, how long it lasts, and how many times per second it is delivered. The adjustability and reversibility of DBS allows neuromodulation to be done without serious side effects. DBS has been used since the mid-1980s to treat movement disorders such as Parkinson's disease and was first used for the treatment of OCD in 1999.[13] DBS has so far only been tested in those with severe OCD which is characterized by a 32-40 on the Yale-Brown Obsessive-Compulsive Scale (Y-BOCS). The Y-BOCS scale is a test that is used to determine the severity of OCD by asking the patient a series of questions related to the impact of the symptoms on their lives. Since then, following multiple studies, DBS has been approved by the Food and Drug Administration in the United States.[5]

This review covers the regions of the cortico-striato-thalamo-cortical (CSTC) pathway that have been tested for OCD, their overall effectiveness, and their shortcomings. By discussing the differences and results of studies, a better action plan is formulated to determine the best treatment parameters for OCD patients using DBS.

II. CSTC PATHWAY

The regions targeted in most DBS studies are a part of the cortico-striato-thalamo-cortical (CSTC) pathway which is a brain circuit that controls movement execution, habit formation, and reward.[19] Recent studies and analysis using neuroimaging show abnormal activity and anatomical differences in the CSTC circuits of patients with OCD. The affected regions include the orbitofrontal cortex (OFC), the anterior cingulate gyrus (ACC), the prefrontal cortex (PFC), and the ventral striatum.[7] The current leading hypothesis is that OCD is associated with hyperactivity of the CSTC loop, and while the precise mechanism of DBS is unknown, its effects on patients with OCD could be explained by an inhibition of the CSTC network.[15] The CSTC network projects from the frontal cortex to specific targets in the striatum. It then goes through the basal ganglia, through direct and indirect pathways, to the thalamus and back to where it started in the frontal cortex. Neuroimaging findings that relate the involvement of the CSTC pathway in the pathophysiology of OCD show elevated activity in the nodes of this circuit in OCD patients at rest, which are accentuated during symptom provocation and attenuated toward normal with successful treatment.[15, 18] There are two widely studied stimulation areas of the CSTC pathway for OCD patients.[7] The first is the striatal region, which includes the anterior limb of the internal capsule (ALIC), the ventral striatum/ventral capsule (VS/VC), the ventral caudate nucleus, the nucleus accumbens (NAc), the bed nucleus of the striata terminalis (BST), and the medial forebrain. The second is the subthalamic nucleus (STN).[19]

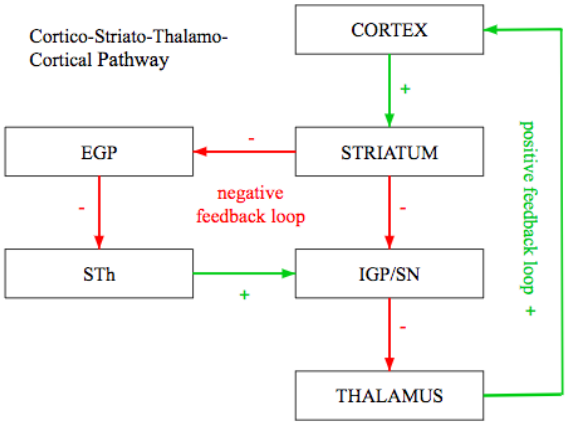


Figure 1 - Feedback loops within the CSTC pathway
The CSTC pathway uses both positive and negative feedback loops in regions of the brain to influence habit formation and reward.

III. STRIATAL REGION

A. Anterior limb of the internal capsule (ALIC)

The ALIC is a bundle of fibers that connects the thalamus to the frontal lobe, caudate nucleus and putamen. It is known to be linked with the processing of emotion, cognition, decision making, and motivation, and abnormalities in its white matter are often found in psychiatric illnesses such as OCD.[17]

The first case study of DBS for the treatment of OCD that was done in 1999, targeting the ALIC. In this study, four patients were implanted with bilateral DBS, and three were reported to have beneficial effects (Nuttin et al. 1999).[13] Since a definition for improvement is not given, it cannot be said to which degree the patients got better. However, more detailed data was given for one patient who reportedly had a 90% Y-BOCS reduction. Similar to this study, there have been many more that propose beneficial effects of DBS in the ALIC for OCD. In another study done with 6 patients, a blinded double random crossover design was used to choose four of those patients who were given continuous stimulation. All four of the patients were shown to have a 61.3% decrease in the Y-BOCS scale from postoperative off to stimulation-on.[16] Unfortunately, during the stimulation-off period, all four patients' symptoms severity approached the baseline levels.[16] An open-label study with 20 patients showed a 33% reduction on the Y-BOCS scale in the 40% responders one year after surgery.[20] In an open label study, both the patients and the researchers know which treatment is being administered. Therefore, the fact that it is an open label study introduces positive bias, which is the overestimation of the positive effects, into the results. It also enhances the placebo effect among patients especially considering that the results are based on self-reports, making the data from this study subjective. According to the data, DBS did not show an improvement in anxiety and depression. In fact, 35% of patients reported a sudden increase in anxiety after stimulation was stopped. Adverse events were limited to one case due to hardware infection.[20] Effects due to changes in DBS settings, such as hypomania, disinhibition, lack of concentration, transient loss of energy, sleep disturbances, and >20% weight gain also occurred. These side-effects were temporary. All of the studies above, however, had many varying factors among patients such as stimulation parameters and follow-up periods which makes it difficult to compare the data directly. Since there are so many variable factors, it is hard to determine whether it was stimulation parameters, the location, or the follow-up time that caused the improvement of the patients. For future studies, researchers should decide on only one varying factor, whether it be stimulation time, parameters, or DBS location.

The inconsistent conclusions in YBOCS reduction and response rates between studies in individual clinical trials show that they cannot predict the response towards DBS in a large group. A study done by Hartman et al. used neuroimaging techniques to follow the progress of ALIC-DBS.[25] It showed that the two best responder target locations within the ALIC had stronger connectivity between the right middle frontal gyrus (MFG), which is known to be associated with executive functions and adapting in response to changing task requests. On the other hand, the two non-responder targets had stronger connectivity with the right thalamus and the orbital part of the right inferior frontal gyrus, which is implicated in task-switching and maintenance of compulsive behavior. This shows that beneficial ALIC-DBS relies on the adaptation of the brain and particular executive functions and that the effectiveness of ALIC-DBS is largely based on which fiber pathways it influences within the active DBS site. However, since each patient's body reacts in a different way, in order for the stimulation parameters to be optimized, each patient had a different set of stimulation parameters. This means that while the location of DBS stimulation and the follow-up period were the same, the parameters among patients being tested were not. Therefore studies from different locations cannot be compared because then there are two changing variables, the location and the stimulation parameters. While ALIC-DBS results may not be consistent now, further investigation into influencing specific fiber pathways may yield patterns and responses that are more applicable to the general public of OCD patients.

B. Ventral striatum/ventral capsule (VS/VC)

The junction of the ventral capsule/ventral striatum is referred to as the 'VC/VS' and is now one of the most popular targets for OCD. However, one of the first randomized studies done for VC/VS failed to establish the difference between "on" and "off" stimulation patients. The result of the study could be explained by the short follow up time of only two months for active stimulation.[4] After one year though, out of those same patients, 66.7% were responders showing a more than 35% decrease on the Y-BOCS scale. However, the reduction in symptoms and criteria for responders is unknown so the true efficiency cannot be determined. Most studies show an average response rate of 50% within a year after the surgery.[4] Greenberg et al. did a double-blind crossover study with 26 patients and implanted them with electrodes in the VS/VC. After a follow-up period of 3-36 months, 10 patients (38.5%) were responders and had a Y-BOCS reduction from 0-62.5%. Since the time before the follow-up influences the effectiveness of DBS, the varying time of

follow-ups between patients may explain the drastic range of Y-BOCS reductions found in the trial.[10] However, almost all of the patients had a secondary anxiety disorder; 82% of patients had major depressive disorder. These disorders often have similar symptoms as OCD since it is also a mental anxiety disorder that can cause anxiety and frustration. It is possible that the big reduction in Y-BOCS scores could be due to a response the DBS exerted on the comorbidity, and not on OCD itself. While this does not mean that DBS had no effect on OCD as there still were patients with no comorbidities who improved, it does mean that the extent to which it helps is uncertain and may not be as positive as it appears. Still, compared to other forms, VS/VC-DBS, shows a higher response rate.

Similar to ALIC-DBS, it was also associated with transient cognitive side-effects that were mainly related to changes within links to the reward and motivational system. Certain patients experienced stimulation-induced symptoms and temporary hypomania.[10] A study by Tyagi et al. compared the effectiveness of VS/VC and anteromedial subthalamic nucleus (amSTN) DBS in the same patients and looked for differences in mood, cognitive flexibility and associated neural circuitry.[1] They found that the patients with VS/VC stimulation had a greater improvement in mood.[1] Their implant site was within the VC and connected mainly to the medial orbitofrontal cortex (medial OFC). This improvement in mood can be explained by previous abnormal functional connectivity found by functional MRI (fMRI) in OCD patients. The MRI scans showed the medial OFC to be hyperresponsive to threat stimuli.

C. *Nucleus accumbens (NAcc)*

While many case reports have been documented investigating NAcc-DBS, there are only a small handful of larger studies. In the open-label design study conducted by Kohl et al., 18 patients were implanted with electrodes with the bilateral NAcc. In a double-blind study, neither the researchers nor the patients know which treatment is being administered. Hence, compared to a double-blind study, open-label studies added positive bias into their conclusions. After one year, there were around 50% responders and 16.7% partial responders.[12] In a second study, 16 patients also received bilateral NAcc stimulation and after a follow-up period of eight months, 9 (56.25%) patients were responders. Responders showed a mean decrease of 46% on the Y-BOCS scale. After the eight months and a double-blind crossover study with a two-week period, the difference in Y-BOCS score between active and sham stimulation was a notable 25%.[8] Sham stimulation is used in research for a placebo effect and indicates an inactive or weak form of stimulation. This is different from off-stimulation in which there is no stimulation at all. The time for off-stimulation was different for each patient and so were the stimulation parameters. The varying off-stimulation period may result in different effects in patients as the patients with lower off-stimulation may experience a higher Y-BOCS reduction. There was also no control group who had sham stimulation for the whole period of the study. This means that the effectiveness and numbers from the study are not fully reliable because the placebo group is altered throughout the experiment and no longer provides an accurate baseline for comparison. Aside from mild forgetfulness and word-finding problems, depression and anxiety were decreased significantly within patients.

Resting-state fMRI scans in 16 NAcc-DBS responders showed reduced frontostriatal connectivity between NAcc, lateral prefrontal cortex (LPFC), and the medial PFC before the treatment when compared to afterwards.[6] NAcc-DBS also reduced excessive frontal low-frequency oscillations caused by symptom-provoking events. These oscillations are known to be linked to the severity of OCD symptoms.[11] By decreasing the excessive frontostriatal connectivity and allowing the natural processing of stimuli to return, targeting regions of the NAcc with the strongest connectivity to the LPFC and medial PFC might further enhance the beneficial effects of NAcc-DBS.

D. *Bed nucleus of the striata terminalis (BST)*

Although the BST is not typically cited in OCD pathology, it is known to have a role in emotional learning and there has been evidence relating it to anxiety responses.[14] While not many DBS studies have been done in the BST for OCD, the few that have been done look promising. One study conducted by Luyten et al. monitored patients with electrodes planted in BST and patients with electrodes in the bilateral ALIC simultaneously.[14] After a follow-up period of 48 months, 80% of the BTS stimulated patients were responders and had an average Y-BOCS reduction of 50%. The patients with ALIC implanted electrodes had a 16.7% response rate and demonstrated a 22% Y-BOCS reduction. Since they were a part of the same experiment, many varying factors were the same. ALIC-DBS patients seem to show a notably lower response rate and Y-BOCS reduction than BST-DBS. Compared to past ALIC-DBS experiments however, the success and improvement shown are rather low. This may be because this study has patients with implants in both the BST and ALIC and therefore may not have many factors that directly boost the performance of the ALIC-DBS like the solely ALIC-DBS focused studies. It is also possible that the exact region within the ALIC that the electrodes were implanted was not very beneficial as the precise location of the DBS activity has been proven to influence the benefits of DBS. Regardless, BST-DBS shows great promise for the future as its responder rate and Y-BOCS reduction are high and while not enough studies have been done in order to predict its effects on a larger population, hopefully future studies focus more on this region.

IV. SUBTHALAMIC NUCLEUS

The anteromedial subthalamic nucleus (amSTN) is another targeted region for DBS in OCD patients. Since stimulation of the subthalamic nucleus with DBS was previously used for Parkinson's patients, the concept of stimulating this region was known for improving motor symptoms.[2] However, after studies were conducted, it was found that the stimulation of the dorsolateral STN affects primarily motor networks while stimulation of the anteromedial STN (amSTN-DBS) affects limbic and

associative functions.[9] The first double-blind crossover study of amSTN-DBS with 16 OCD patients had a 75% responders after a mere 3 month follow up period. The patients had a decrease in OCD severity by 39%.[1] A meta-analysis reported that 44% of patients could be considered responders to STN-DBS.[3] While the responding rate and severity decrease are not extraordinary compared to the other DBS techniques, the small follow-up period to obtain those similar results is. Most other locations used for DBS require at least one-year post-surgery to obtain the same results, as shown in Goodman et al.

Authors	Target	Patients	Design	Stimulation Parameters	Follow up (months)	Y-BOCS Reduction
Luyten et al, 2015[14]	Bilateral ALIC and BST	6 - ALIC; 15 - BST;	Double-blind, random crossover	F: 85-130 Hz, PW: 90-450 µs, A: 3-10.5 V	48	22% - ALIC 50% - BST
Nuttin et al, 1999[13]	ALIC	1	Case study	F: 100 Hz, PW: 210 µs, A: 5-7 V	--	90%
Nuttin et al, 2003[16]	ALIC	4	Double-blinded random crossover	F: 100 Hz, PW: 210-450 µs, A: 4-10 V	21	61.3%
Huys et al, 2019[20]	ALIC	20	Open-label	F: 100-135 Hz, PW: 210 µs, A: 4-7 V	12	33%
Greenburg et al, 2010[10]	VS/VC	26	Double-blind crossover trial	F: 100-130 Hz, PW: 90-450 µs, A: 2-8 V	3-36	0 - 62.1%
Goodman et al, 2010[4]	VS/VC	6	Double-blind crossover	F: 130-135 Hz, PW: 90-210 µs, A: 2-5 V	12	>35%
Kohl et al, (2015)[12]	NAcc	18	Open-label	F: 120-130 Hz, PW: 90-150 µs, A: 3.8-7 V	12	--
Denys et al, 2010[6]	Bilateral NAcc	16	Double-blind crossover	F: 130 Hz, PW: 90 µs, A: <= 5 V	8	46%
Mallet et al, 2002[2]	amSTN	16	Double-blind crossover	F: 130 Hz, PW: 90 µs, A: 3 V	3	39%

Figure 2 - **Summary of DBS Studies for OCD treatment**
This table lists the researchers, target region, design, stimulation parameters, follow-up period, and Y-BOCS reduction, of different studies that have tested DBS for the treatment of OCD.

V. CONCLUSION

DBS in the CSTC pathway improves OCD symptoms through a variety of different locations, each having its own benefits. Each DBS target influences through its own specific pathway and electrodes implanted in different areas within the same region have been shown to have varying outcomes. Future studies should focus more on optimizing the targets chosen as it has already been shown that certain fiber pathways are more effective than others. By experimenting more in specific regions, more precise patterns and pathways that are more beneficial can be found. Plus, since stimulation parameters are easy to adjust after surgery, it is highly critical that the location of the implant is best suited towards each patient to decrease their symptoms to the highest degree. In order to find this, studies should make sure that there is only one varying factor, and that factor should be the location. By keeping all other variables, such as stimulation parameters, the same, the location can be best optimized. Since placing the electrodes is the most invasive and difficult part of DBS, it will be easier to optimize other variables afterwards and create a better understanding and treatment plan for patients with OCD using DBS if the location is ideal.

VI. ACKNOWLEDGMENTS

I would like to thank Ms. Annie Phan for reviewing my paper and providing valuable feedback.

REFERENCES

1. Tyagi H., Apergis-Schoute A., Akram H., Zrinzo L., Hariz M., Joyce E. (2019 May 01). A Randomized Trial Directly Comparing Ventral Capsule and Anteromedial Subthalamic Nucleus Stimulation in Obsessive-Compulsive Disorder: Clinical and Imaging Evidence for Dissociable Effects. Biological Psychiatry. https://www.biologicalpsychiatryjournal.com/article/S0006-3223(19)30063-0/fulltext
2. Mallet L., Mesnage V., Houeto J., Pelissolo A., Yelnik J., Behar C., et al. (2002, October 26). Compulsions, Parkinson's disease, and stimulation. *The Lancet.* https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(02)11339-0/fulltext
3. Alonso P., Cuadras D., Gabriëls L., Denys D., Goodman W., Greenberg B., Jimenez-Ponce F., Kuhn J., Lenartz D., Mallet L., Nuttin B., Real E., Segalas C., Menchon J. (2015, July 24). Deep Brain Stimulation for Obsessive-Compulsive Disorder: A Meta-Analysis of Treatment Outcome and Predictors of Response. *Plos One.* https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0133591
4. Goodman W., Foote K., Greenberg B., Hill C., Rasmussen S., Okun M. (2010, February 01). Deep Brain Stimulation for Intractable Obsessive Compulsive Disorder: Pilot Study Using a Blinding, Staggered-Onset Design. *Biological Psychiatry.* https://www.biologicalpsychiatryjournal.com/article/S0006-3223(09)01426-7/fulltext
5. Karas P., Lee S., Jimenez-Shahed J., Goodman W., Wiswanathan A., Sheth S. (2019, January 08). Deep Brain Stimulation for Obsessive-Compulsive Disorder: Evolution of Surgical Stimulation Targets Parallels Changing Model of Dysfunctional Brain Circuits. *Frontiers Neuroscience*. https://www.frontiersin.org/articles/10.3389/fnins.2018.00998/full
6. Senova S., Clair A., Palfi S., Yelnik J., Domenech P., Mallet L. (2019, December 13). Deep Brain Stimulation for Refractory Obsessive-Compulsive Disorder: Towards an Individualized Approach. *Frontiers Psychiatry.* https://www.frontiersin.org/articles/10.3389/fpsyt.2019.00905/full
7. Tastevin M., Spatola G., Régis J., Lançon C., Richieri R. (2019, May 15). Deep brain stimulation in the treatment of obsessive-compulsive disorder: current perspectives. *Neuropsychiatric disease and treatment.* https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6526924/
8. Denys D., Mantione M., Figee M, et al. (2010, October 4). Deep Brain Stimulation of the Nucleus Accumbens for Treatment-Refractory Obsessive-Compulsive Disorder. *JAMA Psychiatry.* https://jamanetwork.com/journals/jamapsychiatry/fullarticle/210896
9. Mulders A., Plantinga B., Schruers K., Duits A., Jassen M., Ackermans L., Leentjens A., Jahanshahi A., Temel Y. (2016, November 10). Deep brain stimulation of the subthalamic nucleus in obsessive-compulsive disorder: Neuroanatomical and pathophysiological consideration. *European neuropsychopharmacology: the journal of the European College of Neuropsychopharmacology.* https://pubmed.ncbi.nlm.nih.gov/27838106/
10. Greenberg B., Gabriels L., Malone D., Rezai A., Friehs G., Okun M., Shapira N., Foote K., Cosyns P., Kubu C., Malloy P., Salloway S., Giftakis J., Rise M.,

Machado A., Baker K., Stypulkowski P., Goodman W., Rasmussen S., Nuttin B. (2010, January). Deep brain stimulation of the ventral internal capsule/ventral striatum of obsessive-compulsive disorder: worldwide experience. *Molecular psychiatry*. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3790898/

11. Knyazev G. (2011, October 02). EEG delta oscillations as a correlate of basic homeostatic and motivational processes. *Neuroscience & Biobehavioral Reviews*. https://www.sciencedirect.com/science/article/abs/pii/S0149763411001849?via%3Dihub

12. Kohl, S., Gruendler, T. O., Huys, D., Sildatke, E., Dembek, T. A., Hellmich, M., Vorderwulbecke, M., Timmermann, L., Ahmari, S. E., Klosterkoetter, J., Jessen, F., Sturm, V., Visser-Vandewalle, V., & Kuhn, J.(2015). Effects of deep brain stimulation on prepulse inhibition in obsessive-compulsive disorder. *Translational psychiatry*. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5068764/

13. Nuttin B., Cosyns P., Demeulemeester H., Gybels J., Meyerson B. (1999, October 30). Electrical stimulation in anterior limbs of internal capsules in patients with obsessive-compulsive disorder. *Lancet*. https://pubmed.ncbi.nlm.nih.gov/10551504/

14. Luyten L., Hendrickx S., Raymaekers S., Gabriëls L., Nuttin B. (2015, August 25). Electrical stimulation in the bed nucleus of the stria terminalis alleviates severe obsessive-compulsive disorder. *Molecular Psychiatry*. https://www.nature.com/articles/mp2015124

15. Saxena S., Rauch S. (2000 September). Functional neuroimaging and the neuroanatomy of obsessive-compulsive disorder. *The Psychiatric Clinics of North America*. https://pubmed.ncbi.nlm.nih.gov/10986728/

16. Nuttin B., Gabriëls L., Cosyns P., Meyerson B., Andréewitch S., Sunaert G., Maes A., Dupont P., Gybels J., Gielen F., Demeulemeester H. (2003). Long-term electrical capsular stimulation in patients with obsessive-compulsive disorder. *Neurosurgery*. https://pubmed.ncbi.nlm.nih.gov/12762871/

17. Emos M., Agarwal S. (2020, August 10). Neuroanatomy, Internal Capsule. *StatPearls*. https://www.ncbi.nlm.nih.gov/books/NBK542181/

18. Saxena S., Brody A., Schwartz J., Baxter L. (1998). Neuroimaging and frontal-subcortical circuitry in obsessive-compulsive disorder. *The British Journal of Psychiatry*. https://pubmed.ncbi.nlm.nih.gov/9829024/

19. Milad M., Rauch S. (2011, December 2). Obsessive-compulsive disorder: Beyond segregated cortical-striatal pathways. *Trends in cognitive sciences*. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4955838/

20. Huys D., Kohl S., Baldermann J., Timmermann L., Volker S., Visser-Vandewalle V., Kuhn J. (2019). Open-label trial of anterior limb of internal capsule-nucleus accumbens deep brain stimulation for obsessive-compulsive disorder: insights gained. *Journal of Neurology, Neurosurgery & Psychiatry*. https://jnnp.bmj.com/content/90/7/805

21. Claire A., Haynes W., Mallet L. (2018, June 05). Recent advances in deep brain stimulation in psychiatric disorders. *F1000Research*. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5989145/

22. Zhang, C., Li, D., Jin, H., Zeljic, K., & Sun, B. (2017). Target-specific deep brain stimulation of the ventral capsule/ventral striatum for the treatment of neuropsychiatric disease. *Annals of translational medicine*. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5673773/

23. Luyten L. (2020). The Bed Nucleus of the Stria Terminalis: Translational Deep Brain Stimulation to Reduce Anxiety. *The Neuroscientist: a review journal bringing neurobiology, neurology and psychiatry*. https://pubmed.ncbi.nlm.nih.gov/31971488/

24. Schrues K., Baldi S., Huevel T., Goossens L., Luyten L., Leentjens A., Ackermans L., Temel Y., Viechtbauer W. (2019, August 05). The effects of deep brain non-stimulation in severe obsessive-compulsive disorder: an individual meta-analysis. *Translational Psychiatry*. https://www.nature.com/articles/s41398-019-0522-6

25. Hartmann C., Lujan J., Chaturvedi A., Goodman W., Okun M., McIntyre C., Haq I. (2016, January 19). Tractography Activation Patterns in Dorsolateral Prefrontal Cortex Suggest Better Clinical Responses in OCD DBS. *Frontiers Neuroscience*. https://www.frontiersin.org/articles/10.3389/fnins.2015.00519/full

Small Molecule Inhibition of Oncogenic KRAS and Downstream Signaling Pathways in Pancreatic Ductal Adenocarcinoma

Pranavi Garlapati[1]

[1]Texas Academy of Maths and Sciences, 1705 W Sycamore St, Denton, TX 76201

***Abstract*— Common recurring mutations in certain cell signaling pathways continue to be found in the progression and metastasis of Pancreatic ductal Adenocarcinoma (PDAC), with Ki-ras2 Kirsten rat sarcoma viral oncogene (KRAS), a family of genes controlling control cell growth, being the most frequent one. Understanding the nuances of this mutation and its downstream effectors can provide a blueprint for effective treatment methods using small molecule inhibitors. This review will elucidate how activation of the mutated KRAS protein leads to the eventual activation of various intracellular pathways, leading to cell proliferation. In addition, new targeted therapeutic treatments in the form of small molecule inhibitors will also be discussed.**

I. Introduction

With late stage diagnosis and unsuccessful treatment methods, Pancreatic ductal adenocarcinoma (PDAC) is on the road to become the second most common cause of cancer-related deaths by 2030 [1]. Over the past 10 years, new therapeutic approaches have emerged to increase the survival rate of patients with PDAC, but lack of early diagnosis causes prognosis to remain disappointing. Identifying the type of mutated cell signaling pathway in the patient and using targeted therapy may solve this problem. The KRAS isoform of Ras is a commonly mutated oncogene, a gene that due to stressors can turn into a tumor cell. KRAS plays a major role in the progression and invasion of PDAC tumors, and its constitutive activation is found in over 90% of PDAC cases [2]. The Mutational KRAS protein is permanently attached to GTP (Guanosine triphosphate) and continually activates downstream signaling pathways such as nuclear factor κB (NF-κB), phosphoinositide 3-kinase (PI3K), JUN N-terminal kinase (JNK), and the Mitogen-activated protein kinase (MAPK) pathway, which all lead to tumorigenesis, the formation of tumors [3].

II. Inhibition of KRAS

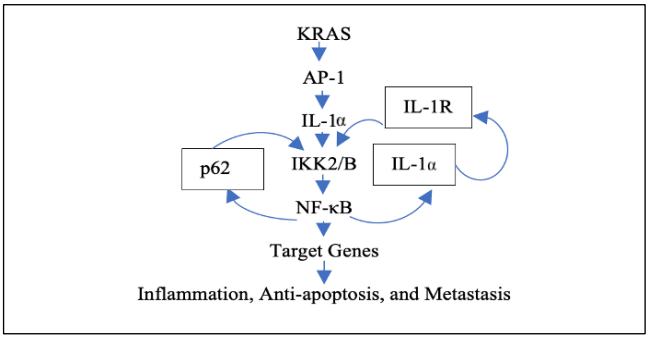
The role of KRAS has been studied in detail over the past few decades as it is a key marker for PDAC prognosis. However, due to its constant activation of GTP and its reoccurrence and resistance after treatment, it has earned itself the reputation of being an undruggable molecule. Several small molecule inhibitors are being developed to target this seemingly undruggable mutant. Recently, Sotorasib (AMG 510) has been identified as a potential small molecule inhibitor of the KRAS G12C mutation. In a recent phase 1 trial monitoring the dosage and safety of the treatment, 129 patients diagnosed with various cancer types were used to detect the effectiveness of the drug. The study had a majority of 59 patients with Non-small-cell lung carcinoma (NSCLC). The drug was administered orally with monotherapy- meaning only the single drug and nothing else- once daily. In the subgroup with NSCLC, 32.2% of patients showed an objective response to the drug (percentage of patients whose disease showed a positive response to the drug), and 88.1% had disease control (percentage of patients whose disease regressed or remained stable) [4]. Further testing on PDAC may give us valuable insight on the effectiveness of this small molecule inhibitor. Another drug, Adagrasib (MRTX849), was recently recognized as a potential inhibitor of the $KRAS^{G12C}$, the most common mutation of KRAS. It is the first of all $KRAS^{G12C}$ mutation inhibitors to move to clinical trials. *In vitro*, tumor regression was observed in 17 of 26 (65%) patient-derived xenograft models (tissue taken from humans and placed into mice) that had $KRAS^{G12C}$ mutations at dose levels of 30 -100 mg/kg/day of Adagrasib. Adagrasib caused significant tumor reduction in lung and colon adenocarcinoma patients, indicating the need for further testing of this inhibitor in PDAC [5].

III. NF-κB Pathway

NF-κB is an especially important transcription factor that plays diverse roles in PDAC, including inflammation, angiogenesis, control of apoptosis, and proliferation through transcription of various cytokines, leading to cancer-related inflammation. Constitutive NF-κB activation is present in 90% of cancers, including PDAC [6]. Studies show that mutant KRAS and several cytokines activate NF-κB resulting in the expression of proinflammatory signals. The mechanism by which NF-κB is activated by KRAS is shown in Figure 1. First, KRAS activates AP-1. In response to the activation of AP-1, an upregulated transcription factor in mutated KRAS cells, Interleukin-1α (IL-1α), a cytokine, is overexpressed. IL-1α then activates the IKK2/B complex, which causes the consequent expression of NF-κB. This is the short term activation of NF-κB. In order to sustain this activation long term, NF-κB then continually results in the expression of p62 and IL-1α, which act in a feedforward loop in order to keep NF-κB activated long term. Thus, proinflammatory and

antiapoptotic responses are achieved in KRAS mutated and NF-κB activated cells through the auto-regulatory loop caused by IL-1α and p62 [7].

Figure 1: NF-κB activation in KRAS mutated tumor cells through p62 and IL-1α feedforward loop

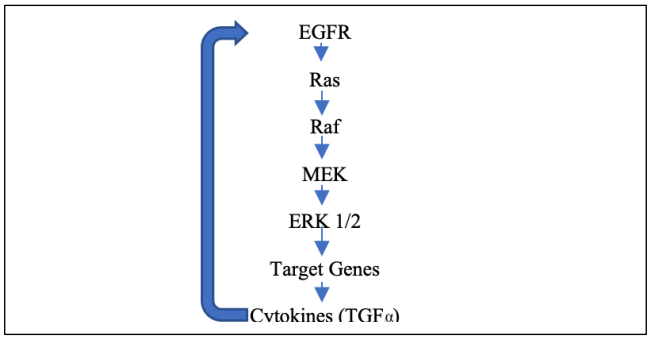


This pathway illustrates that inhibition of IL-1α overexpression can hinder PDAC development by suppressing NF-κB activation because NF-κB is a downstream effector of IL-1α. The mechanism by which IL-1α induces the NF-κB pathway is through the IL-1R receptor. The IL-1 acts as a link between the KRAS mutation and the NF-κB pathway needed for mutant KRAS induced PDAC, making it an important therapeutic target. A recent study showed that Anakinra, an FDA approved treatment, inhibits the IL-1 receptor and can decrease tumor progression by inhibiting IL-1α induced NF-κB activity. The study used various human PDAC lines as well as testing *in vivo* on orthotopic nude mouse models, genetically engineered mice implanted with tumors, (n=20, 5 per group) to demonstrate the increased effectiveness of Anakinra in decreasing tumor growth along with the potential benefits of using it alongside chemotherapy (Gemcitabine, the most common and effective chemotherapeutic treatment currently) [8]. TPCA-1 (GW683965) is also found to be a potential treatment due to its direct inhibition of the IKK2/B complex upstream of NF-κB. The study, using tumor xenograft models, also found that through the dual inhibition of NF-κB and another simultaneous tumorigenic pathway (STAT3), the inhibitor represses the transcription of IL-6. This can have profound effects on the autocrine mechanism in which the NF-κB acts [9]. BAY 11-7082 (BAY) is another small molecule inhibitor that acts similar to TPCA-1 by inhibiting the IKK complex upstream of NF-κB. It has profound effects on tumor development by acting as an anti-inflammatory protein, which is key to tumor regression. Moreover, BAY demonstrated broad-spectrum inhibition of other oncogenic cellular pathways including PI3K/Akt/IKK/NF-κB, ERK/JNK/AP-1, TBK1/IRF-3, and JAK-2/STAT-1 [10]. Further testing may reveal the extensive power BAY 11-7082 has in fighting PDAC.

IV. MAPK/ERK Pathway

The MAPK pathway plays a key role in cell differentiation and proliferation. Mutational Ras, activated by the binding of a signaling ligand on the Epidermal growth factor receptor (EGFR), causes the activation of downstream effectors such as Raf and eventually MAPK, also called MEK in mammals. MEK can also activate Extracellular signal-regulated kinase (ERK), which moves to the nucleus to activate a variety of transcription factors, ultimately leading to the transcription of various cytokines and cell survival. Furthermore, ERK activation can result in the upregulation of Transforming growth factor alpha (TGFα), and EGFR ligand, creating an autocrine feedback loop [11].

Figure 2: MAPK/ERK activation in KRAS Mutated tumor cells through Raf



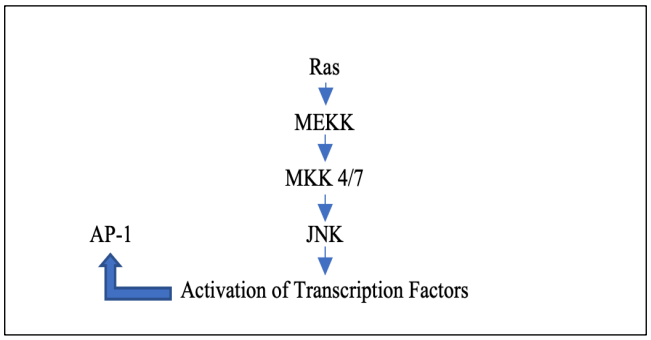
Potential small molecular inhibitors of this pathway are still in pre-clinical evaluations, but some drugs show promise. The Raf inhibitor, LY3009120, prevents the downstream phosphorylation and activation of MEK/ERK in Ras mutated melanoma. In nude mice with Patient-derived xenografts (PDX) tumors, a dosage of 15 to 30mg/kg of LY300912 showed almost complete tumor regression with no toxicity observed [12]. Downstream of Raf, CI-1040 (PD 184352) is the first MEK inhibitor shown to inhibit tumor growth *in vivo*. The drug was administered once daily with a dosage of 100 mg/day up to 1600 mg/day in a Phase 1 trial with a sample size of 66 patients. 1 patient showed partial response for 12 months, and 19 patients showed a stable disease lasting a median of 5.5 months [13]. Unfortunately, the Phase II trial showed no significant response worth noting [14]. A variation of CI-1040 (PD 184352) known as PD 0325901 was developed after the failure of the first drug. Compared to CI-1040, PD 0325901 had a 70-fold increase in potency in pre-clinical evaluations. Patients received doses of up to 20 mg and in a sample size of 27 patients, 2 patients were reported to have partial response and 8 patients had stable disease [15]. A third MEK inhibitor, ARRY-142886 (AZD6244) shows promise as it shows potent activity in human tumor xenograft models [16]. Overall, MEK Inhibition shows promise as stable disease is observed in patients with various types of cancer and it is warranted for further research. In some cases, MEK inhibition can be toxic, and it might be important to consider the ablation (removal) of the upstream receptor, EGFR, instead. Studies show that mutated KRAS upregulates EGFR expression, and the inhibition and ablation of EGFR significantly decreases tumorigenesis *in vivo* because EGFR is necessary for MEK/ERK activation [17]. That being said, a recent study showed that patients with KRAS mutations (181 patients) showed less improvement with targeted EGFR treatment than patients without (86 patients), pointing to the need for more understanding of this approach [18]. In contrast, targeted deletion of both EGFR and Raf prevented PDAC development in KRAS driven tumor models as well as

human PDX models for 2 years with no precursors to tumor progression, while also not affecting downstream MEK/ERK activation, eliminating the need to worry about treatment-related toxicity [19].

V. JNK Pathway

The c-Jun NH_2-terminal kinase (JNK) Pathway has a dualistic role in cancer development because it has both pro and anti-tumor functions. Ras acts as a switch converting JNK from having an anti-tumor role to a pro-tumor role. JNK is usually activated by stress cytokines such as IL-1α and tumor necrosis factor α (TNF-α) [20]. Ras activates downstream effectors, mitogen-activated extracellular signal-regulated kinase (MEK), Mitogen-activated protein kinase (MAPK) kinase 4 or 7, and eventually JNK. JNK causes the transcription of Activator Protein-1 (AP-1) commonly implicated in cell proliferation and differentiation [21]. This pathway is clearly shown in Figure 3. Dysregulation of the JNK pathway may increase cell proliferation [22]. Several loss of function studies have shown that JNK activation is necessary for Ras-induced transformation of tumor cells [23].

Figure 3: JNK activation pathway in KRAS Mutated tumor cells

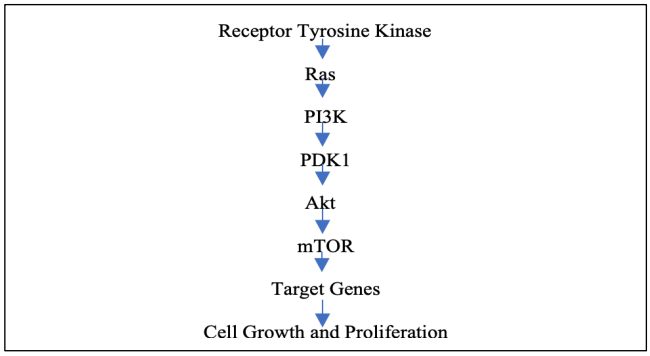


Because JNK is implicated in many different cancers, therapeutically targeting JNK is in the best interest for PDAC patients [24]. SP600125, a selective inhibitor, is shown to significantly inhibit JNK and prevent the activation of many downstream cytokines. SP600125 was administered at dosages up to 15 mg/kg in female CD-1 mice. SP600125 showed significant results with high specificity to the JNK pathway [25]. An IC50 of a drug helps indicate the dosage of the drug necessary to reduce the response by half. The IC50 of SP600125 for JNK2 is equal to 40 nM and for JNK3, it is 90 nM [26]. A recent study investigated the efficiency of another small molecule inhibitor, Bentamapimod (AS602801), of the JNK pathway. Bentamapimod has been shown to be effective in the treatment of inflammatory endometriosis. The study explored its effects on human cancer cells both *in vitro* and *in vivo*. In cancer stem cells derived from human pancreatic cancer, Bentamapimod demonstrated cytotoxicity [27]. AS602801 has an IC50 for JNK1 = 80 nM, for JNK2 = 90 nM, and for JNK3 = 230 nM.

VI. PI3K Pathway

The PI3K pathway plays a role in cell metabolism, growth, and proliferation. PI3K is bound to Receptor Tyrosine Kinases (RTKs) in the cell membrane and activation of the receptor by an extracellular signal activates PI3K to convert Phosphatidylinositol (3,4)-bisphosphate (PIP_2) lipids into Phosphatidylinositol (3,4,5)-trisphosphate (PIP_3). Mutated Ras facilitates the binding of phosphoinositide-dependent kinase-1(PDK1) to PIP_3 at the cell membrane causing the downstream activation of Protein kinase B (Akt) and Mammalian target of rapamycin (mTOR) as shown in Figure 4 [28]. The PI3K/AKT signaling pathway is one of the most deregulated pathways in cancer and is commonly found in KRAS driven-pancreatic cancer both in human PDAC and mouse models [29][30].

Figure 4: PI3K activation in KRAS Mutated tumor cells



PI3K therapeutic treatments have been unsuccessful thus far as the specificity of the molecules and the toxicity at which they have worked has limited their potential [31]. However, several inhibitors are in Phase III of clinical trials. Gedatolisib (PF04691502), a dual PI3K/mTOR, has shown potent activity and decreased tumor density in head and neck squamous cell carcinomas by inhibiting Akt proliferation. The comparison between control tumor cell lines and those treated with Gedatolisib using an XTT assay after 48 hours showed statistical significance with P value ≤ 0.01 [32]. BKM120 (Buparlisib) is an orally administered reversible PI3K inhibitor. Both *in vitro* and *in vivo*, Buparlisib shows significant antiproliferative activity [33]. A maximum tolerated dose (MTD) of 100 mg/day showed efficiency as well as safety in a study with sample size of 64 patients [34]. This study proved an earlier study testing Buparlisib on 15 Japanese patients with advanced solid tumors and another study with 35 Western patients that also settled on 100 mg/day as the most efficient dose [35][36].

Conclusion

From an improved understanding of the dysregulated cell-signaling pathways mentioned (NF-κB, MAPK/ERK, JNK, and PI3K), many targeted therapeutic treatments have emerged in the past decade. The high mutation rate and inherent drug resistance of PDAC serves as a barrier to effective treatment. This paper elucidates some of the most novel treatments that hold potential to target parts of these mutated pathways as a framework to combat tumor progression, but much remains to be learned on how to effectively utilize these treatments to target KRAS mutated pancreatic tumors in patients. Note that most pathways act

synergistically through cross talk, so targeting many parts of different pathways may offer better treatment plan clinically. For example, NF-κB and JNK share common upstream activators and may act synergistically to regulate cancer cell survival. Looking ahead, it may be valuable to research the effects of a combination of the small molecule inhibitors given for more effective treatment. The small molecule inhibitors discussed can also be combined with dosages of the current effective chemotherapeutic drugs for PDAC, Gemcitabine and 5-fluorouracil (5-FU). Gemcitabine has already shown promise when combined with Anakinra as mentioned above. Applying targeted therapy to the tumor microenvironment is another approach that can be researched. Often the tumor microenvironment with its dense stroma and inflammatory cells makes drug delivery ineffective, so combining small molecule inhibitors with immunotherapy might be a more effective course of treatment. As of now, there is only one FDA approved immunotherapeutic drug for PDAC, and it is Pembrolizumab (Keytruda). Utilizing it along with targeted therapy in patients with PDAC is worthy of consideration. Combination therapy holds a lot of promise in the treatment of PDAC, but it is also important to consider the toxicity that occurs when multiple drugs are administered. Further studies using small molecule inhibitors with other treatments must be conducted before deeming this as an effective course of treatment.

ACKNOWLEDGMENT

I would like to thank Dr. Jie Fu and Dr. Paul Chiao at the MD Anderson Cancer Center for their guidance and feedback when writing this paper.

REFERENCES

[1] Rahib L, Smith BD, Aizenberg R, Rosenzweig AB, Fleshman JM, Matrisian LM., et al. “Projecting Cancer Incidence and Deaths to 2030: The Unexpected Burden of Thyroid, Liver, and Pancreas Cancers in the United States.” *Cancer Research*, vol. 74, no. 11, 2014, pp. 2913–2921., doi:10.1158/0008-5472.can-14-0155.

[2] Hruban H, Goggins M, Parsons J, Kern E. Progression model for pancreatic cancer. Clin Cancer Res 2000;6:2969–72.

[3] Buscail, L., Bournet, B., & Cordelier, P. (2020). Role of oncogenic KRAS in the diagnosis, prognosis and treatment of pancreatic cancer. Nature Reviews Gastroenterology & Hepatology, 17(3), 153-168. doi:10.1038/s41575-019-0245-4

[4] Hong, D. S., Fakih, M. G., Strickler, J. H., Desai, J., Durm, G. A., Shapiro, G. I., . . . Li, B. T. (2020). KRASG12C Inhibition with Sotorasib in Advanced Solid Tumors. *New England Journal of Medicine, 383*(13), 1207-1217. doi:10.1056/nejmoa1917239

[5] Hallin, J., Engstrom, L. D., Hargis, L., Calinisan, A., Aranda, R., Briere, D. M., Sudhakar, N., Bowcut, V., Baer, B. R., Ballard, J. A., Burkard, M. R., Fell, J. B., Fischer, J. P., Vigers, G. P., Xue, Y., Gatto, S., Fernandez-Banet, J., Pavlicek, A., Velastagui, K., Chao, R. C., … Christensen, J. G. (2020). The KRAS^{G12C} Inhibitor MRTX849 Provides Insight toward Therapeutic Susceptibility of KRAS-Mutant Cancers in Mouse Models and Patients. *Cancer discovery*, *10*(1), 54–71. https://doi.org/10.1158/2159-8290.CD-19-1167

[6] Prabhu, L., Mundade, R., Korc, M., Loehrer, P. J., & Lu, T. (2014). Critical role of NF-κB in pancreatic cancer. Oncotarget, 5(22), 10969–10975. https://doi.org/10.18632/oncotarget.2624

[7] Ling, J., Kang, Y., Zhao, R., Xia, Q., Lee, D., Chang, Z., . . . Chiao, P. (2012). KrasG12D-Induced IKK2/β/NF-κB Activation by IL-1α and p62 Feedforward Loops Is Required for Development of Pancreatic Ductal Adenocarcinoma. Cancer Cell, 21(1), 105-120. doi:10.1016/j.ccr.2011.12.006

[8] Zhuang, Z., Ju, H. Q., Aguilar, M., Gocho, T., Li, H., Iida, T., Lee, H., Fan, X., Zhou, H., Ling, J., Li, Z., Fu, J., Wu, M., Li, M., Melisi, D., Iwakura, Y., Xu, K., Fleming, J. B., & Chiao, P. J. (2016). IL1 Receptor Antagonist Inhibits Pancreatic Cancer Growth by Abrogating NF-κB Activation. *Clinical cancer research : an official journal of the American Association for Cancer Research*, *22*(6), 1432–1444. https://doi.org/10.1158/1078-0432.CCR-14-3382

[9] Nan, J., Du, Y., Chen, X., Bai, Q., Wang, Y., Zhang, X., Zhu, N., Zhang, J., Hou, J., Wang, Q., & Yang, J. (2014). TPCA-1 is a direct dual inhibitor of STAT3 and NF-κB and regresses mutant EGFR-associated human non-small cell lung cancers. *Molecular cancer therapeutics*, *13*(3), 617–629. https://doi.org/10.1158/1535-7163.MCT-13-0464

[10] Lee, J., Rhee, M. H., Kim, E., & Cho, J. Y. (2012). BAY 11-7082 is a broad-spectrum inhibitor with anti-inflammatory activity against multiple targets. *Mediators of inflammation*, *2012*, 416036. https://doi.org/10.1155/2012/416036

[11] Molina, J. R., & Adjei, A. A. (2006). The Ras/Raf/MAPK Pathway. *Journal of Thoracic Oncology, 1*(1), 7-9. doi:10.1016/s1556-0864(15)31506-9

[12] Peng, S., Henry, J., Kaufman, M., Lu, W., Smith, B., Vogeti, S., . . . Flynn, D. (2015). Inhibition of RAF Isoforms and Active Dimers by LY3009120 Leads to Anti-tumor Activities in RAS or BRAF Mutant Cancers. *Cancer Cell, 28*(3), 384-398. doi:10.1016/j.ccell.2015.08.002

[13] Rinehart, J., Adjei, A. A., Lorusso, P. M., Waterhouse, D., Hecht, J. R., Natale, R. B., . . . Meyer, M. B. (2004). Multicenter Phase II Study of the Oral MEK Inhibitor, CI-1040, in Patients With Advanced Non-Small-Cell Lung, Breast, Colon, and Pancreatic Cancer. *Journal of Clinical Oncology, 22*(22), 4456-4462. doi:10.1200/jco.2004.01.185

[14] Rinehart, J., Adjei, A. A., Lorusso, P. M., Waterhouse, D., Hecht, J. R., Natale, R. B., . . . Meyer, M. B. (2004). Multicenter Phase II Study of the Oral MEK Inhibitor, CI-1040, in Patients With Advanced Non-Small-Cell Lung, Breast, Colon, and Pancreatic Cancer. *Journal of Clinical Oncology, 22*(22), 4456-4462. doi:10.1200/jco.2004.01.185

[15] Lorusso, P., Krishnamurthi, S., Rinehart, J. R., Nabell, L., Croghan, G., Varterasian, M., . . . Meyer, M. B. (2005). A phase 1–2 clinical study of a second generation oral MEK inhibitor, PD 0325901 in patients with advanced cancer. *Journal of Clinical Oncology, 23*(16_suppl), 3011-3011. doi:10.1200/jco.2005.23.16_suppl.3011

[16] Lee, P., Wallace, E., Yeh, T., Poch, G., Litwiler, K., Pheneger, T., ... & Winkler, J. (2004). ARRY-142886, a potent and selective MEK inhibitor: III) Efficacy in murine xenograft models correlates with decreased ERK phosphorylation.

[17] Ardito, C., Grüner, B., Takeuchi, K., Lubeseder-Martellato, C., Teichmann, N., Mazur, P., . . . Siveke, J. (2012). EGF Receptor Is Required for KRAS-Induced Pancreatic Tumorigenesis. *Cancer Cell, 22*(3), 304-317. doi:10.1016/j.ccr.2012.07.024

[18] Chiramel, J., Backen, A., Pihlak, R., Lamarca, A., Frizziero, M., Tariq, N., . . . Mcnamara, M. (2017). Targeting the Epidermal Growth Factor Receptor in Addition to Chemotherapy in Patients with Advanced Pancreatic Cancer: A Systematic Review and Meta-Analysis. *International Journal of Molecular Sciences, 18*(5), 909. doi:10.3390/ijms18050909

[19] Blasco, M. T., Navas, C., Martín-Serrano, G., Graña-Castro, O., Lechuga, C. G., Martín-Díaz, L., . . . Barbacid, M. (2019). Complete Regression of Advanced Pancreatic Ductal Adenocarcinomas upon

Combined Inhibition of EGFR and C-RAF. *Cancer Cell, 35*(4). doi:10.1016/j.ccell.2019.03.002

[20] Tournier C. (2013). The 2 Faces of JNK Signaling in Cancer. *Genes & cancer*, *4*(9-10), 397–400. https://doi.org/10.1177/1947601913486349

[21] Garces de Los Fayos Alonso, I., Liang, H. C., Turner, S. D., Lagger, S., Merkel, O., & Kenner, L. (2018). The Role of Activator Protein-1 (AP-1) Family Members in CD30-Positive Lymphomas. *Cancers*, *10*(4), 93. https://doi.org/10.3390/cancers10040093

[22] Davis, R. J. (2000). Signal Transduction by the JNK Group of MAP Kinases. *Cell, 103*(2), 239-252. doi:10.1016/s0092-8674(00)00116-1

[23] Johnson, R., Spiegelman, B., Hanahan, D., & Wisdom, R. (1996). Cellular transformation and malignancy induced by ras require c-jun. *Molecular and cellular biology*, *16*(8), 4504–4511. https://doi.org/10.1128/mcb.16.8.4504

[24] Suzuki, S., Okada, M., Shibuya, K., Seino, M., Sato, A., Takeda, H., Seino, S., Yoshioka, T., & Kitanaka, C. (2015). JNK suppression of chemotherapeutic agents-induced ROS confers chemoresistance on pancreatic cancer stem cells. *Oncotarget*, *6*(1), 458–470. https://doi.org/10.18632/oncotarget.2693

[25] Bennett, B. L., Sasaki, D. T., Murray, B. W., O'Leary, E. C., Sakata, S. T., Xu, W., Leisten, J. C., Motiwala, A., Pierce, S., Satoh, Y., Bhagwat, S. S., Manning, A. M., & Anderson, D. W. (2001). SP600125, an anthrapyrazolone inhibitor of Jun N-terminal kinase. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(24), 13681–13686. https://doi.org/10.1073/pnas.251194298

[26] Cicenas, J., Zalyte, E., Rimkus, A., Dapkus, D., Noreika, R., & Urbonavicius, S. (2017). JNK, p38, ERK, and SGK1 Inhibitors in Cancer. *Cancers*, *10*(1), 1. https://doi.org/10.3390/cancers10010001

[27] Okada, M., Kuramoto, K., Takeda, H., Watarai, H., Sakaki, H., Seino, S., Seino, M., Suzuki, S., & Kitanaka, C. (2016). The novel JNK inhibitor AS602801 inhibits cancer stem cells in vitro and in vivo. *Oncotarget*, *7*(19), 27021–27032. https://doi.org/10.18632/oncotarget.8395

[28] Hemmings, B. A., & Restuccia, D. F. (2012). PI3K-PKB/Akt pathway. *Cold Spring Harbor perspectives in biology*, *4*(9), a011189. https://doi.org/10.1101/cshperspect.a011189

[29] Baer, R., Cintas, C., Therville, N., & Guillermet-Guibert, J. (2015). Implication of PI3K/Akt pathway in pancreatic cancer: When PI3K isoforms matter? *Advances in Biological Regulation, 59*, 19-35. doi:10.1016/j.jbior.2015.05.001

[30] Kennedy, A. L., Morton, J. P., Manoharan, I., Nelson, D. M., Jamieson, N. B., Pawlikowski, J. S., McBryan, T., Doyle, B., McKay, C., Oien, K. A., Enders, G. H., Zhang, R., Sansom, O. J., & Adams, P. D. (2011). Activation of the PIK3CA/AKT pathway suppresses senescence induced by an activated RAS oncogene to promote tumorigenesis. *Molecular cell*, *42*(1), 36–49. https://doi.org/10.1016/j.molcel.2011.02.020

[31] Hanker, A. B., Kaklamani, V., & Arteaga, C. L. (2019). Challenges for the Clinical Development of PI3K Inhibitors: Strategies to Improve Their Impact in Solid Tumors. *Cancer Discovery, 9*(4), 482-491. doi:10.1158/2159-8290.cd-18-1175

[32] Herzog, A., Bian, Y., Vander Broek, R., Hall, B., Coupar, J., Cheng, H., Sowers, A. L., Cook, J. D., Mitchell, J. B., Chen, Z., Kulkarni, A. B., & Van Waes, C. (2013). PI3K/mTOR inhibitor PF-04691502 antitumor activity is enhanced with induction of wild-type TP53 in human xenograft and murine knockout models of head and neck cancer. Clinical cancer research : an official journal of the American Association for Cancer Research, 19(14), 3808–3819. https://doi.org/10.1158/1078-0432.CCR-12-2716

[33] Maira, S. M., Pecchi, S., Huang, A., Burger, M., Knapp, M., Sterker, D., Schnell, C., Guthy, D., Nagel, T., Wiesmann, M., Brachmann, S., Fritsch, C., Dorsch, M., Chène, P., Shoemaker, K., De Pover, A., Menezes, D., Martiny-Baron, G., Fabbro, D., Wilson, C. J., … Voliva, C. F. (2012). Identification and characterization of NVP-BKM120, an orally available pan-class I PI3-kinase inhibitor. Molecular cancer therapeutics, 11(2), 317–328. https://doi.org/10.1158/1535-7163.MCT-11-0474

[34] Wu, Y. L., Zhang, L. I., Trandafir, L., Dong, T., Duval, V., Hazell, K., & Xu, B. (2016). Phase I Study of the Pan-PI3K Inhibitor Buparlisib in Adult Chinese Patients with Advanced Solid Tumors. Anticancer research, 36(11), 6185–6194 https://doi.org/10.21873/anticanres.11212

[35] Ando, Y., Inada-Inoue, M., Mitsuma, A., Yoshino, T., Ohtsu, A., Suenaga, N., Sato, M., Kakizume, T., Robson, M., Quadt, C., & Doi, T. (2014). Phase I dose-escalation study of buparlisib (BKM120), an oral pan-class I PI3K inhibitor, in Japanese patients with advanced solid tumors. Cancer science, 105(3), 347–353. https://doi.org/10.1111/cas.12350

[36] Bendell, J. C., Rodon, J., Burris, H. A., de Jonge, M., Verweij, J., Birle, D., Demanse, D., De Buck, S. S., Ru, Q. C., Peters, M., Goldbrunner, M., & Baselga, J. (2012). Phase I, dose-escalation study of BKM120, an oral pan-Class I PI3K inhibitor, in patients with advanced solid tumors. Journal of clinical oncology : official journal of the American Society of Clinical Oncology, 30(3), 282–290. https://doi.org/10.1200/JCO.2011.36.1360