Automatic Cell Segmentation with *Zymomonas mobilis*

Ana-Karina Potcoava

Defense Date: October 31, 2023

Department of Biochemistry
University of Colorado at
Boulder

Committee Members

Jeffery Cameron, PhD: Department of Biochemistry – Thesis Advisor
Halil Aydin, PhD: Department of Biochemistry – Honors Council Representative

Jason Zietz, PhD: Department of Information Sciences – Outside Reader
John Rinn, PhD: Department of Biochemistry – Additional Committee Member

Abstract

Zymomonas Mobilis ZM4 is a well-known industrial gram-negative bacterium known for its high sugar uptake bioethanol production. With single-cell imaging and machine learning algorithms advancing rapidly, it will be much easier to understand the metabolic system of the bacterium. A pre-cursor to metabolic manipulation of the bacterium is to evaluate the cell growth of various populations by using an inducible CRISPR-interference (CRISPRi) toolset on fluorescent and metabolic targets to determine the construction of strains capable of manipulating cell metabolism to increase bioethanol production. To evaluate cell growth, we implemented single-cell time-lapse microscopy, which revealed population heterogeneity in knockdown rate. These observations led to using a cell segmentation program (CyAn) to analyze biomass and growth rates generated from metabolic manipulations. This program was coupled with the development of cell tracking software for *Z. Mobilis* using dilated Convolutional Neural Networks (dCNNs) to simplify the cell tracking pipeline and to open new avenues of tracking for the future.

Key Words: CRISPRi, Machine Learning, dilated Convolutional Neural Networks

Acknowledgements

I want to thank Dr. Jeffery Cameron for entrusting me with such an invodlved project that taught me how to think about experimentation and allowed me to apply my interdisciplinary skills in computation and biochemistry. It was a formative experience that taught me many lessons I will take into future projects.

Many have helped within the Cameron Lab, and the guidance everyone has provided with this research project. Thank you especially to Margaret Habib for her mentorship and support throughout the last year. This project was a series of failures and small progress that taught me more about microbiology and lab protocol than I could have imagined. I also want to thank my fellow undergraduates who worked with this bacterium, Nasreen Marikar and Delwin Mayben.

Thanks to Dr. Jian Wei Tay for his one-on-one consultations on microscopy, data acquisition, and CyAn troubleshooting. His support sparked the start of my cell tracking exploration and understanding of microscopy.

I would also like to thank Dr. John Rinn and Dr. Jason Zietz for their courses and instruction. They played a significant role in my understanding of bioinformatics and the various programming languages I know.

Lastly, a thank you to my friends and family. I want to thank my mother for listening to my concerns with this project and always helping me work through them. I also want to thank my father for trying to understand the world of undergraduate research and listening to my microbiology discussions that were precursors to my future talks. Lastly a thank you to everyone involved, because without you I would not have been able to make headway on this project.

Introduction

Zymomonas mobilis is a Gram-negative non-model organism ethanologenic bacterium that can ferment sugars to produce ethanol with high efficiency. It is commonly found in sugary plant materials such as fruits and popular fruit drinks (4). *Z. mobilis* has been studied for use in industrial fermentation processes like bioethanol production and use in fuels (4). Despite its biofuel potential, there are still challenges in scaling up the use of Z. mobilis in industrial settings.

The pathway that *Z. mobilis* utilizes for its high ethanol production is the Entner-Doudoroff pathway (E-D), which is a glycolytic pathway and leads to *Z. mobilis's* ability to ferment sugars, which has high ethanol production as a result (4). The ZM4 AG4628 *Zymomonas mobilis* strain will be used in all experiments (7). Ag4628 is a strain developed in the Guss lab that is used in conjunction with protocols the lab created for gene editing.

The Eckert Lab utilizes CRISPRi to divert its metabolism towards the production of other compounds of interest using protocols previously used on *Psuedomonis Putida* (2), another Gram- negative bacterium that can degrade organic compounds. The novel aspect of this previous research was the exploration of single-cell level behavior of cell growth by using a CRISPR interference toolset, which is what the project seeks to do with *Z. mobilis*. The same CRISPR interference tool set was implemented from this previous study as both bacteria utilize the E-D pathway and have similar metabolic functions (2, 7).

The project utilizes single-cell time-lapse microscopy to image *Zymomonas* cells as they grow in different media conditions and gene modifications with CRISPRi. CRISPRi is like CRISPR in that it utilizes a guide RNA (gRNA) to identify a gene

sequence within the genome. The difference lies in the Cas9 protein that it uses, which is a mutant form of Cas9 (dCas9), whose endonuclease activity is removed by mutating the endonuclease domains and positions itself in the target region to stop transcription (2).

To test that the gRNAs from the prior experiment work with the new bacteria, GFP (green fluorescent protein) will be added into the bacteria on a plasmid. The gRNAs will target the transcription of GFP so that the GFP protein will not express in daughter cells of the population.

Another point of interest for Zymomonas's metabolism is that it has an anaerobic function that allows it to fix dinitrogen (N_2) and produce ethanol as well (8). This is an area of research that is newer; we will just be analyzing single time-lapse images from a prior experiment.

Preliminary imaging shows population heterogeneity, which is where a population will showcase separate cell expression, such as cells that fluoresce and cells that do not. The heterogeneity within various populations requires further analysis to understand this bacterium's ability to circumvent the CRISPR machinery that allows escapees to dominate the population and cell growth rate effects. To track these cells, the analysis toolbox CyAn will be employed, which was developed for cyanobacteria cell tracking by the Cameron Lab (5). CyAn utilizes different segmentation algorithms to track cells in the films taken through single-cell timelapse microscopy. The segmentation program will use the watershed algorithm (1), which treats the image as a topographic map, and each pixel is assigned a value that creates "watershed lines" based on programmed threshold intensity to define the boundaries of the cells within the image (3).

To track a cell population with a disappearing fluorophore, another fluorophore must be included that remains constant. Therefore, mKATE2 will

be implemented as it is a red fluorophore, so it can be easily identified when compared with GFP. It also had a significant difference in excitation and emission which creates images with high contrast.

Implementation of machine learning through dCNNs (dilated Convolutional Neural Networks) has been explored as an alternative to the traditional segmentation mentioned. This is because the double fluorophore system has proven to be challenging to engineer within the cells. A dCNN has many components to it, but to start, a Neural Network is a program that uses algorithms to make connections between cases like that of a human brain; they can receive multiple inputs and generate an output that can get sent to other neurons (14). Neural networks are fully connected to their inputs. Therefore, a CNN is a Convolutional Neural Network; the difference is that this program uses convolutional layers that allow the inputs to be semi-connected, which creates something called a feature map (15). As the program runs through its different layers of filters, the feature map gets populated with information. Feature extraction is the convoluted part of this (13).

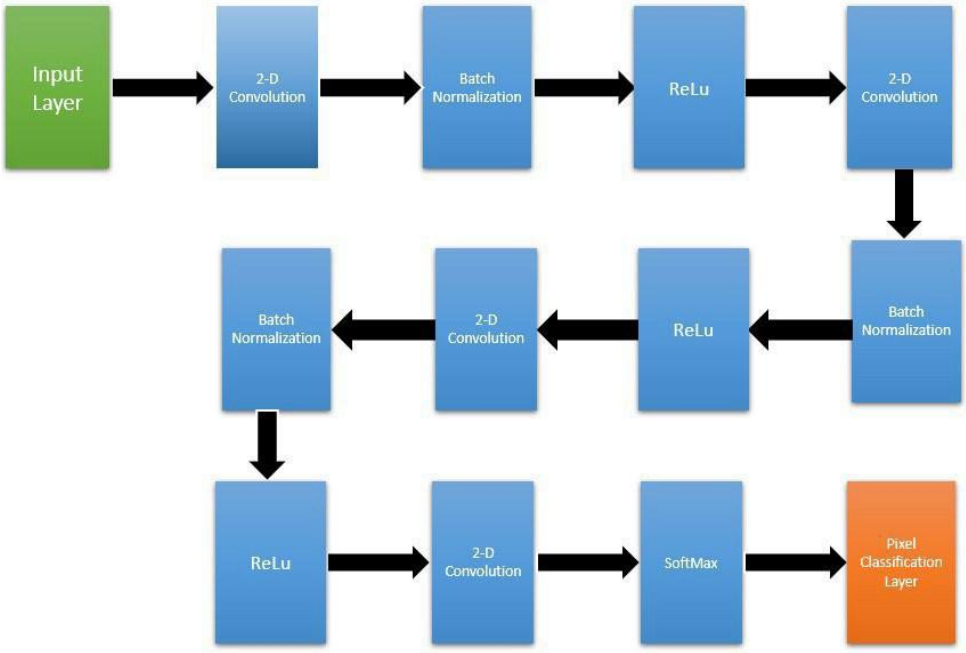

Figure 1: A map of all the convolution layers that the DCNN developed for *Z. mobilis*. The input layer accepts an image that is then sent through the 2-D convolution layer to get a small amount of data to create a feature map. Batch normalization exists to compress the size of the feature map by decreasing connections between layers. Rectified Linear Unit (ReLU) and makes any value below zero as zero (a thresholding function). SoftMax is used to predict the class of the input image (sorts image as class 'cell' or class 'background'). Pixel classification layer sorts every pixel into a category or label.

Dilation is the inflation of the input. It does this expansion by inserting holes between its consecutive elements. This allows for pixels to be skipped in a way so that a larger area of the input is scanned. Hence, a dCNN uses multiple filters to convolute and, therefore, extract more features over a larger area via dilation. Since this entire process involves nearly no human input for the selection of features within the cells, it has been termed Automatic Segmentation (1).

This research seeks to create a tracking and analysis software for *Zymomonas mobilis* that can identify characteristics of cell populations that lead to better solutions in metabolic engineering.

Methods and Materials

There are three separate sections for simplicity: the Microbiology methods for CRISPRi and Fluorescent Protein methods, live cell imaging, and how cell segmentation is done.

Microbiology

For the CRISPRi GFP knockdown, a strain from Adam Guss (Oak Ridge National Laborotories, 7), AG4826, was used. A dCas9, promoter, RBS (ribosome binding site), and gRNA insertion site developed by Jason Peters and optimized for Zymomonas, were cloned out of plasmid pJMP2480 and used as the cargo to be integrated into the *Zymomonas* genome. They were cloned from pJMP2480 plasmid with Spectinomycin antibiotic resistance (12). Competent cells were acquired from New England Biolabs. Once the CRISPR system was integrated into the *Z. mobilis* genome guide, two of the gRNAs targeting GFP from Fenster et al (2) were used. For the case of *Z. mobilis*, Vectors 2 and 4 were selected along with a non-targeting vector.

A strain of AG4826 expressing both GFP and mKATE2 was used. Plasmids containing the CRISPRi system and the two gRNAs plus a non-targeting control were transformed in, and 10ml RMG overnight cultures were started from a single colony from each plate. They were diluted down to OD 0.05 the day of imaging. Once cell culture plates grew, 15 mL falcon tubes were prepared with 10mL RMG media, antibiotic- mKATE2 ChlorR 120mg/L, and for mKATE, 1mM of IPTG was added to the test tube.

Microscopy

Brightfield and fluorescent timelapse microscopy were performed using a Nikon TiE inverted wide-field microscopy with a near-infrared–based Perfect Focus System. A 100x oil immersion objective was used for all images acquired (1.45 numerical aperture). The NIS Elements AR software was used to control the microscope, and a Hamamatsu sCMOS camera was used to capture the images. For cells with mKATE2 protein, the RFP channel from the NIS Elements AR software was used (Ex: 555nm, Em: 595nm, Filter Range: 571-628nm). For cells with GFP protein, the GFP channel was used (Ex: 470nm, Em: 515nm, Filter Range: 503-542nm). The BF (bright field) channel functioned as a reference channel. Images that make up the movies are taken at 5-minute intervals over 6-12 hours.

Slides preparation for the samples: plates were poured with 40mL RMG media, and 0.40g of Agarose was added. The mixture was then microwaved for 2 minutes in an Erlenmeyer flask. After the plate was left to cool for 15 minutes, small cubes were cut out and flipped over to dry for 3 minutes. Upon drying 1μL of sample was pipetted onto the RMG-Agar cube. Depending on the variance of samples, 3-4 spots of sample would be pipetted onto the cube. After drying for 5- 8 minutes, the agar pad was flipped into a well plate. The slide top was sealed with parafilm to ensure the agar pad did not dry out over longer imaging stints. For cultures that contained mKATE2 1mM IPTG was added to the agar pad to induce the CRISPR system.

Cell segmentation and dCNNs

The CyAn package automates cyanobacteria identification in microscope images and uses the following functions: cell segmentation tool, cell data measurement, tracking individual cells and cell division, and image registration. The cell segmentation uses an adaptive threshold algorithm and various morphological filters that creates a challenge when segmenting with two fluorophores. Therefore, we are working on developing a cell segmentation software package based on using deep learning neural networks. The core of the software has been made, and tested on data presented within the paper, but needs to be tailored to function with larger data sets.

The time-lapse movies acquired were in the .nd2 file format, so they first needed to be converted into TIFF files with image J. Bright Field images were separated from fluorescent channels to segment on as these were the largest amount of images that were simpler to acquire. Once these images are acquired, they are put through a pre-processing MATLAB program to cut up each individual from the movie into 'patches' of 128px * 128px. This is to create more images for the program to train on and just large enough to avoid bisecting cells (14).

A program is then set up to scan through the images, eliminate blank images, and check edge conditions to avoid bisected cells (cells cut in half or only remnants). After this, manually, one must select what percentage of images will go towards the training data set and what percentage will be part of the validation data set. These are usually separate folders.

How to use a trained network: There are many convolutional layers to work through the first though is always the input layer, which accepts an argument of inputSize (the 128px * 128px patches) which uses 8-bit encoding.

Next is creating two classes, "Cells" and "Background," for later use. The most important arguments are then the ds, layers, and options arguments. The argument: ds is for the training and validation data, where the folders that contain this will be input. The layers argument is for the different CNN layers available and can be customized to suit various data sets. The options argument then asks for the type of training the data set should be taken through. This data set was made with the maxEpochs option.

For traditional segmentation, the program developed in the Cameron lab, CyAn, was implemented. This program is what is used to generate lineage trees and growth plots. It is also used to generate the binary masks it creates by using thresholding algorithms (5). These binary masks serve as the ground truths (binary masks) the DCNN uses as its training set.

Results and Discussion

Dinitrogen Fixation

Z. mobilis has presented as a bacterium of interest due to its potential use in the biofuels industry. To better understand the metabolic processes of the organism, both different growth conditions, CRISPRi vectors, fluorescence microscopy, and cell tracking were implemented to find which routes of metabolic manipulation may be most helpful and how cell tracking can be implemented to understand the observations that arose about *Z. mobilis* colonies from timelapse microscopy videos.

Since the relatively recent finding of *Z. mobilis's* ability to fix dinitrogen as its sole source, there have been many studies made to understand better how this affects *Z. mobilis's* metabolic pathways. While there have been many studies proving that N2 can be fixed by the bacterium, which does not reduce ethanol production and could potentially improve it, imaging and studies of the physical characteristics of the cells have not been prevalent.

The WT strain findings in Figures 3.1A and 3.1B show that nitrogen fixation occurs with the cells as cell growth is registered from individual cell lineages within the populations. Some cells do not grow or grow at a significantly slower rate throughout the 20 hours the cells were imaged. This is reflected in the cell divisions in figure 3.1A, where the first cell division appears after the 6-hour mark. This is of note, as *Z. mobilis* has shown in previous preliminary imaging sessions to take between 30-40 minutes to double. However, Figures 3A and 3B have some revelations; as can be seen, there are cells (orange arrows) that seem to have rapidly grown into colonies in 3B, but other cells that were not around those colonies appear to have grown larger, but no cell division has occurred. In a way, it is as if some cells (3B) are still in a form of anaphase. When the videos of these cells growing are played, individual cells can be observed doubling but having difficulty splitting.

The WT strain findings in Figures 3.1A and 3.1B show that nitrogen fixation occurs with the cells as cell growth is registered from individual cell lineages within the populations. Some cells do not grow or grow at a significantly slower rate throughout the 20 hours the cells were imaged. This is reflected in the cell divisions in figure 3.1A, where the first daughter cells appear to come about after the 6-hour mark. This is of note, as *Z. mobilis* has shown in previous preliminary imaging sessions to take between 30-40 minutes to double. However, Figures 3A and 3B have some revelations; as can be seen, there are cells (orange arrows) that seem to have rapidly grown into colonies in 3B, but other cells that were not around those colonies appear to have grown larger, but no cell division has occurred. In a way, it is as if some cells (3B) are still in a form of anaphase. When the videos of these cells growing are played, individual cells can be observed doubling but having difficulty splitting.

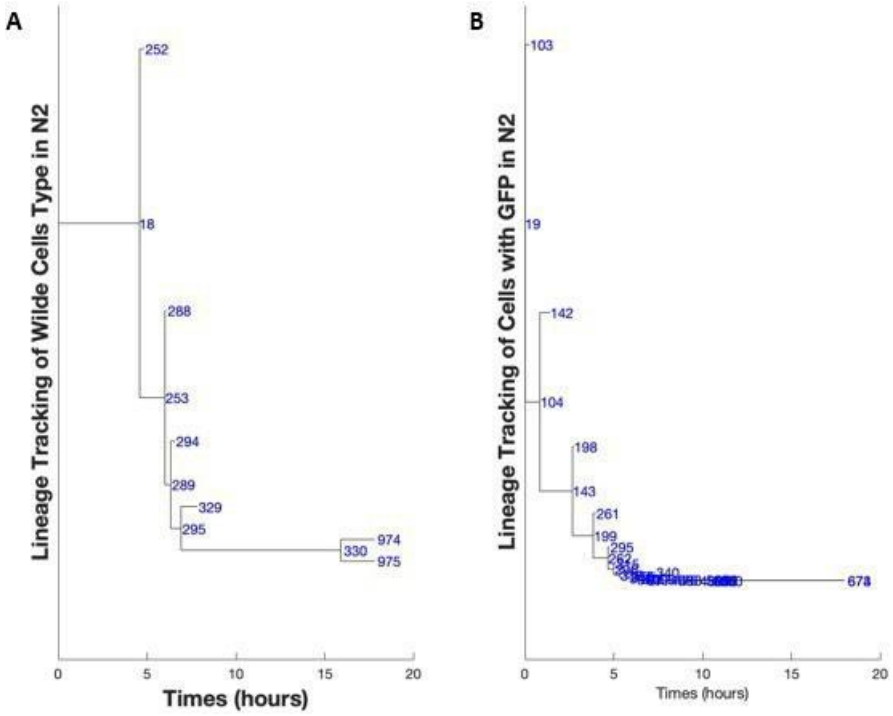

Figure 2: Biomass generated from two AG4826 cells grown in a chamber of N2. A) Lineage of control cell culture. B) Lineage of a cell that has a GFP fluorescent marker.

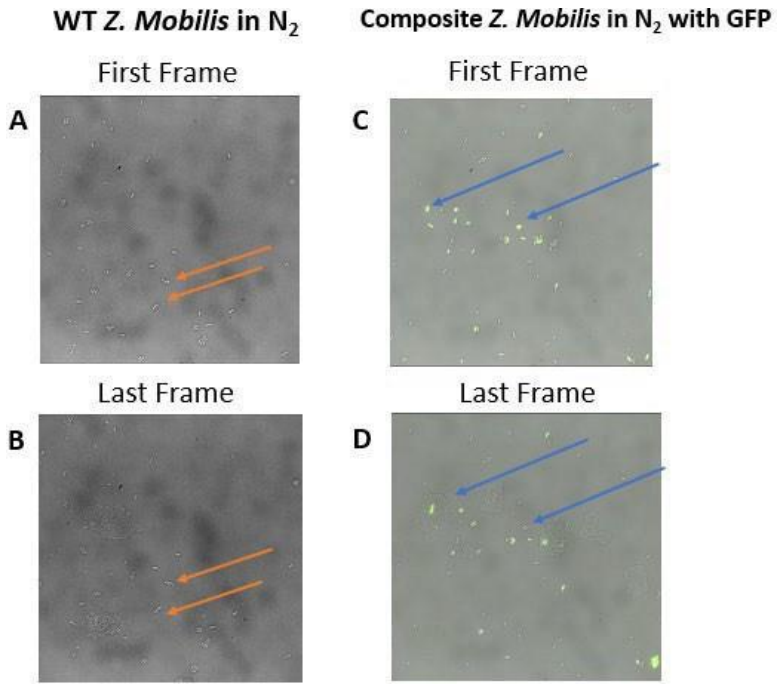

Figure 3: Still images of WT and cells with GFP incorporation grown in N2. A) First frame of WT cells in brightfield, B) Last frame of the timelapse movie in brightfield low growth is observed, C) First frame of BF (brightfield) and fluorescent channel time-lapse composited, D) Last frame of timelapse of with brightfield and fluorescent time-lapse composite.

While there seems to be cell growth observed to prove that N2 fixation is happening, there are two separate cell growth characteristics within the cells. As Figure 2B shows, the GFP stain biomass grows substantially, but there is a slow step of cell growth that is happening for the cell to have this behavior. More film studies and data images must be acquired to understand better if cell heterogeneity exists for these two differently characterized cell populations and to understand the effects of dinitrogen fixation on cell growth and morphology.

CRISPRi Knockdowns

Cell tracking was used to observe CRISPRi knockdowns of GFP in real time using two gRNAs that had been shown in previous experiments to reduce fluorescence, plus a non-targeting control. Figures 4A and 4B serve as the control, and it is shown that nearly all of the

cells within the Non-targeting population fluoresce to an extent. The fluorescence intensity fluctuation within the cells depends on the growth pattern and potential adaptation of cells to eliminate fluorescence. Figures 4C and 4D and 4E and 4F show two of the most promising GFP knockdown vectors. As can be seen throughout the imaging, the cell population seems to lose the intensity of the fluorescence, but there is still fluorescence present. More films need to be done so that the fluorescence decay time can be found to determine how many divisions it takes for fluorescence to be 'lost' or if there is an error with the CRISPRi system.

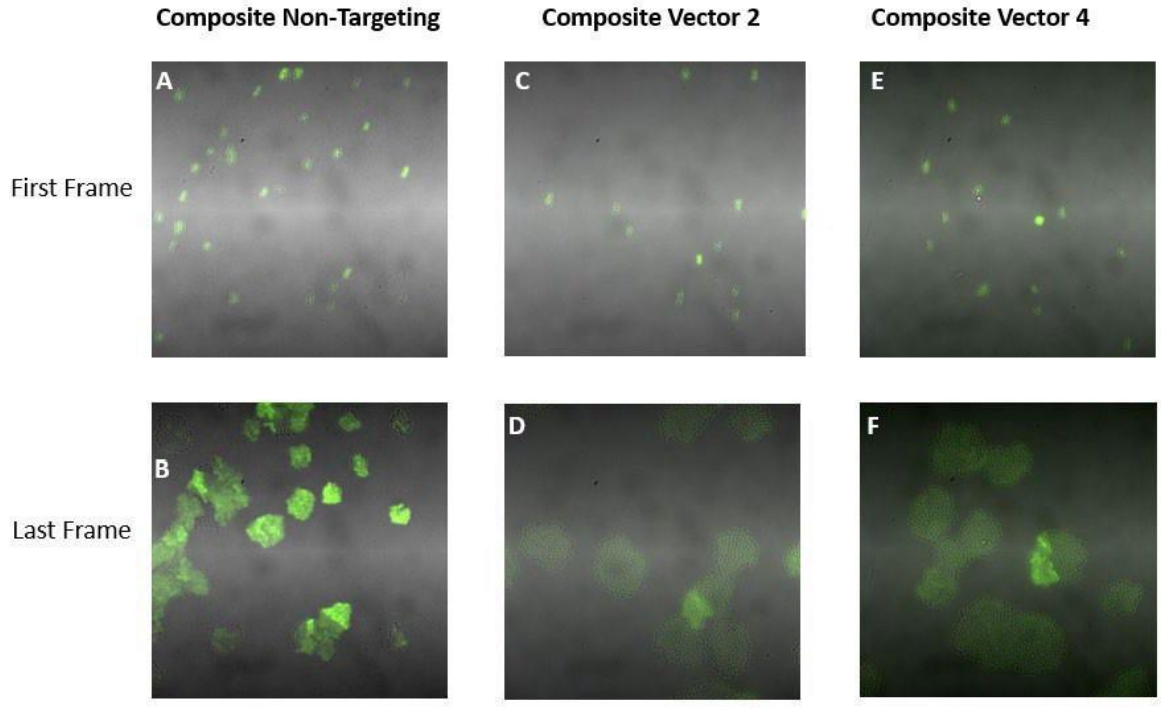

Figure 4: Lineage tracking of cells with CRISPRi knockdown to determine vectors for future use. A) First frame of the composite image of brightfield and fluorescent channel for the control sample. B) Last frame of the composite image of brightfield and fluorescent channel for the control sample. C) Cells with a CRISPRi knockdown of GFP, first frame with cells that have vector 2. D) In Cells with a CRISPRi knockdown of GFP using Vector 2, there is population heterogeneity, with many cells showing decreased fluorescence. E) Cells with a CRISPRi knockdown of GFP, first frame with cells that have vector 4, F) Cells with a CRISPRi knockdown of GFP using Vector 4, there is population heterogeneity, with many cells showing a decrease in fluorescence.

Of note, within the cell lineage pathways of these samples, it appears that the Vector 2 population has the most cell divisions happening within cells compared to Vector 4 and the non-targeting population. This is despite these cells having the least number of colonies from start to end. It is of note that as the cell division time is increased (5A, 5B, 5C) within these samples, it takes a substantial amount of time for the cells to start proper division. This is potentially due to the incorporation of GFP and the CRISPR system. These long periods of division also lead to the question of population health, which will be discussed later.

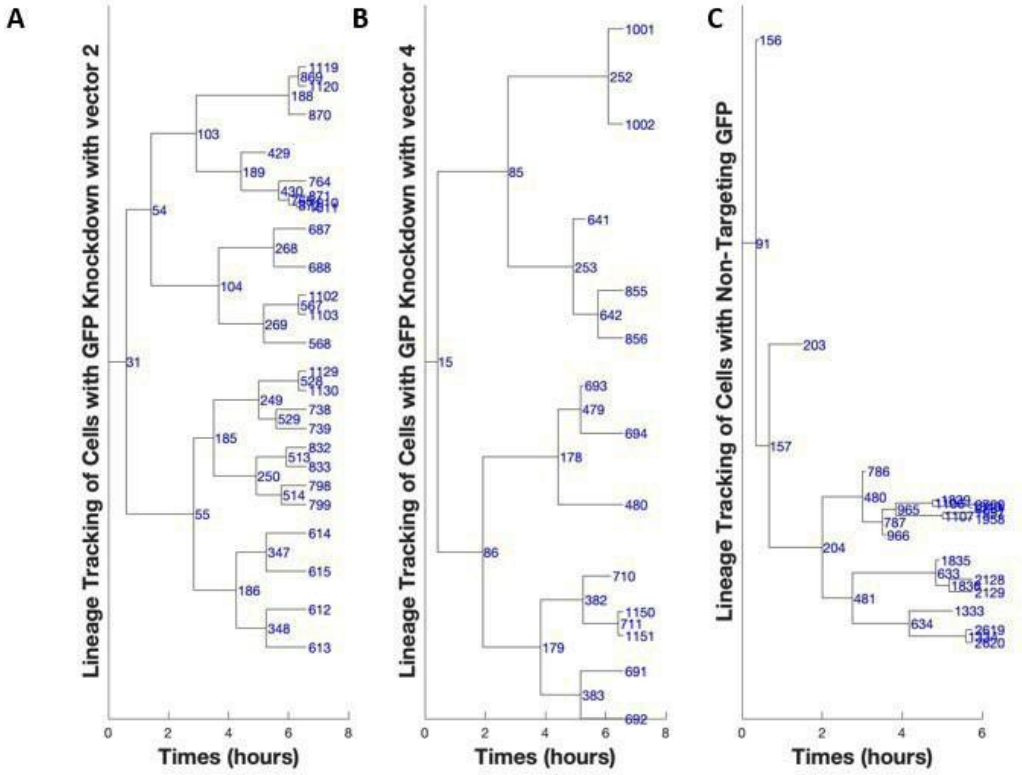

Figure 5: Comparison of biomass generated from cells engineered with CRISPRi knockdown vectors and a non-targeting cell population. A) It is observed that there is a high cell division rate within the cells that grow in the population treated with vector 2. B) It is observed that there is a low cell division rate within the cells that grow in the population treated with vector 4. C) The NT cells do not have a consistent growth plot but still have high cell division rates.

Figure 6 shows the behavior of the relationship between the CRISPR system and the population's growth rate. Figure 6C shows an exponential increase in the amount of cells and

cell divisions compared to 6A and 6B, which show a near halving of cell divisions and population size throughout the video recording. More movies must be made using wild-type *Z. mobilis* cells without engineering to set up a baseline. There also needs to be implemented a software solution where fluorescence is tracked over time to observe fluorescence dilution better.

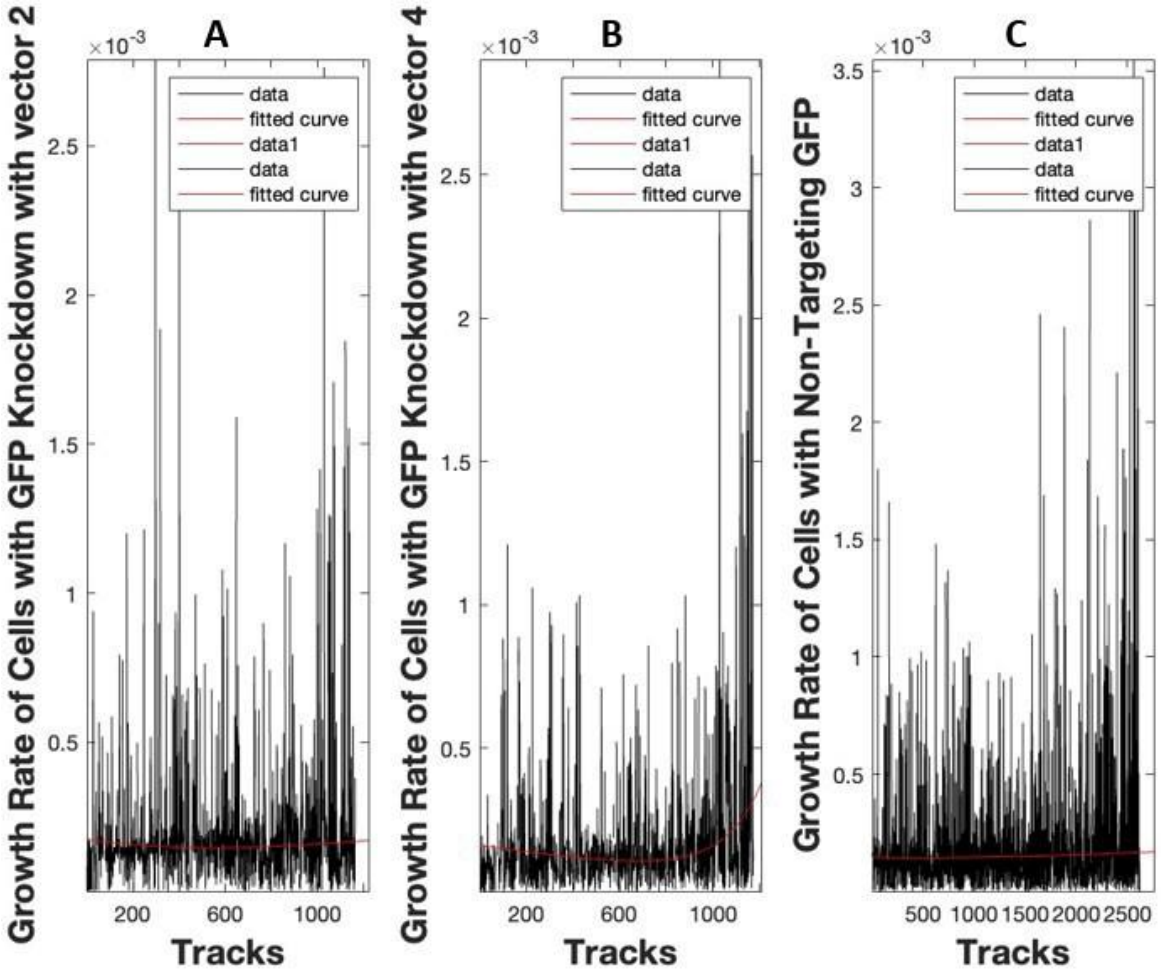

Figure 6: 'Tracks' are the IDs for every new cell, and their growth rate is the size of the major axis length of each cell. This is the measurement of the length of the largest cell axis in relation to the cell ID as cell division occurs. A) This is the curve-fitted plot for cells with the vector 2 CRISPRi mechanism, which shows a relatively steady axis growth rate. B) The curve fitted plot for cells with vector 4 CRISPRi mechanism that shows an increase in growth rate (axis size) for the cells. C) This is the control sample that does not have a CRISPRi mechanism implemented for GFP knockdown. The growth rate is steady and shows that there are more cell divisions and more cells in general within the population.

Double Fluorophore Approach

To track fluorescence over time, with disappearing GFP, in a cell population, a separate fluorophore that remains constant must be implemented to ensure properly transformed cells are tracked. mKATE2 was the most promising second fluorophore to grow within the AG4826 strain used and was then studied to understand its effects on the bacterium better. IPTG was added induce protein expression. Imaging was done to ensure that this was effective for the overall fluorescence of the culture.

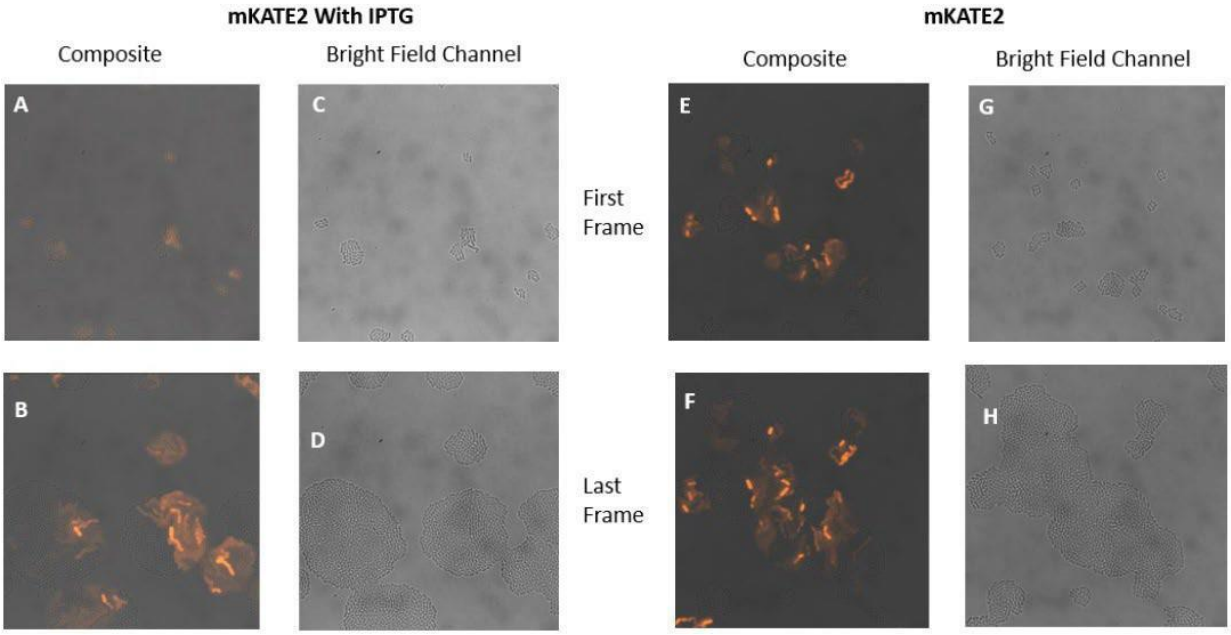

Figure 7: Figures A-D are cells with an inserted fluorescent protein mKATE2 plasmid grown on an agar pad with the inducer IPTG during the time-lapse imaging. Figures E-H are cells with a mKATE2 fluorescent protein plasmid grown without IPTG in the agar. Figures A, B, E, and F are composite images of the RFP fluorescent channel and the Bright Field channel to compare how well the protein's fluorescence relative to the population of cells imaged. Figures C, D, G, and H are brightfield still images at the same time points as the composite images to demonstrate the population heterogeneity.

From visual analysis alone, it can be concluded that IPTG does affect mKATE2 protein expression. In Figures 7B and 7F, while not all the cells are induced with IPTG fluorescence,

those that are not induced have less overall fluorescence within the population, which proves that IPTG does improve the overall protein expression.

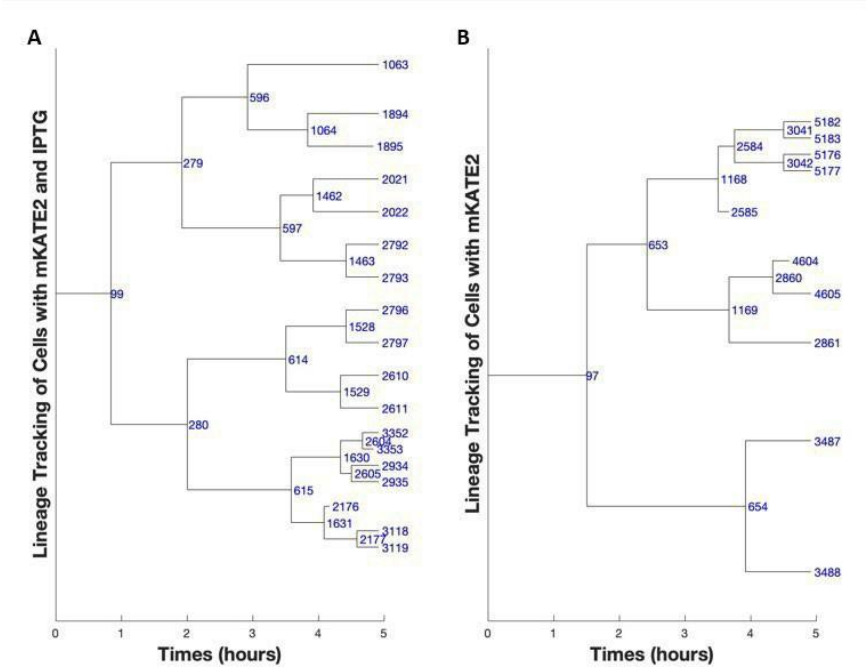

Figure 8: Comparison lineage maps for one cell from each cell population to better represent the effects of IPTG influence on the cell populations. A) A lineage map for one cell within the cell population treated with IPTG demonstrating many cell divisions. B) The lineage tacking for a cell within the population not treated with IPTG shows fewer cell divisions and general biomass in comparison.

From the lineage maps of cells (Figure 8A), it can be observed that IPTG can have a role in cell division as well since more divisions occurred for the induced population, but this is an inconclusive result. This is not the main scope of the experiment, but rather an observation of note that came up within the experiment.

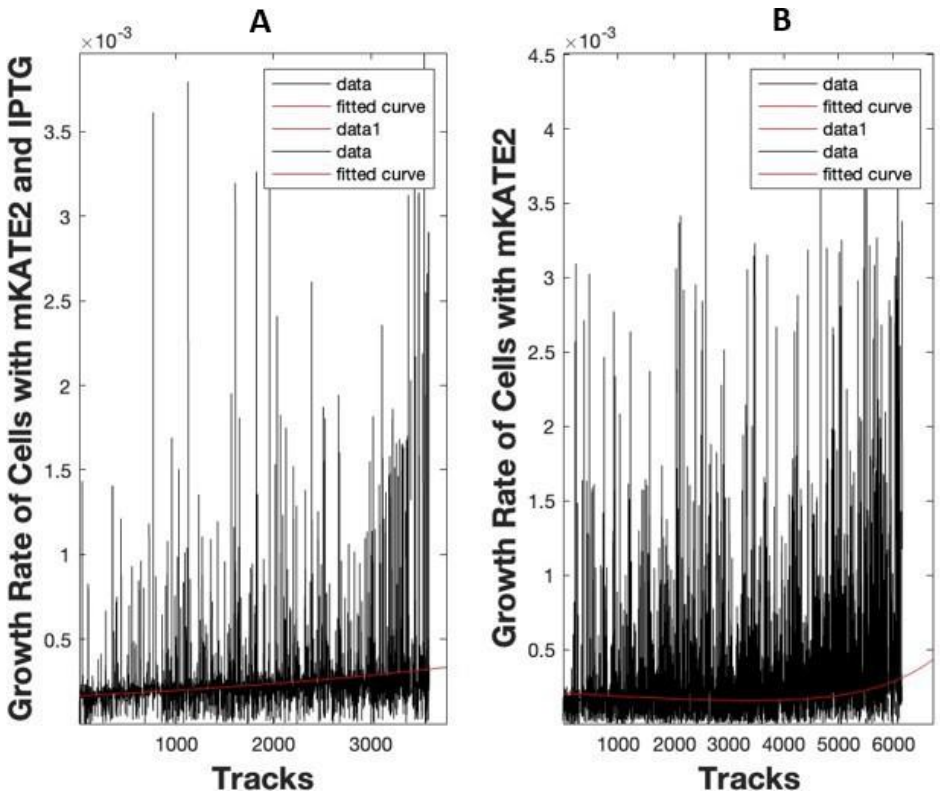

Figure 9: A comparison of growth rate (major axis length of cells) and Tracks (cells and cell divisions) that shows there are more cells divided in B that are untreated than that of A.

Furthermore, the conversation of population health needs to be addressed (Figure 9A) would suggest that the induced population health was lower than that of the untreated due to the higher amount of cell divisions and larger cell growth rate. Further data must be gathered to understand better if the cells that grow within the untreated population (9B) manage to circumvent the mKATE2 folding and why there is such heterogeneity in such a population (7F). The same can be said for the induced population; while there is better overall fluorescence, most of the population still does not correctly fluoresce.

Two Fluorescent Proteins, One Cell

While it is possible to include two fluorescent proteins within the *Z. mobilis* bacterium, this is no easy feat, and the results show that the cells do not grow as well with the integration of two fluorophores.

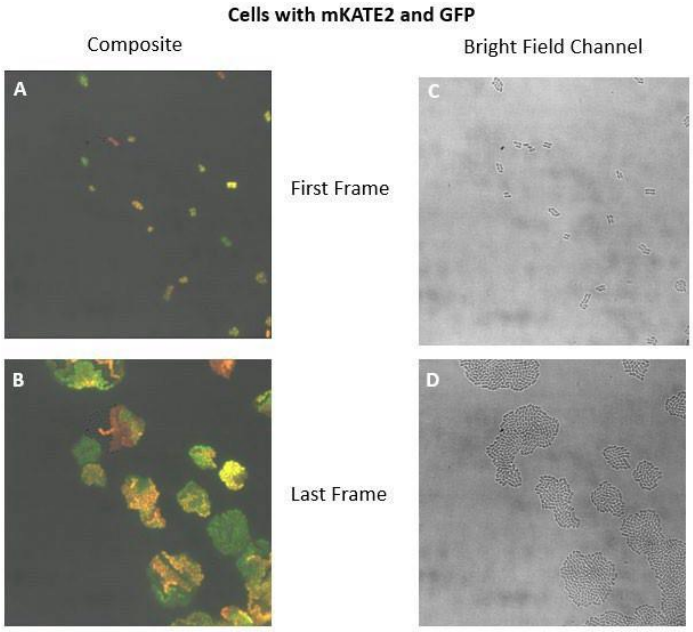

Figure 10: Figures A and B are composites of the GFP fluorescent channel, RFP fluorescent channel, and BF channel to show how well the population fluoresces over time with both fluorescent proteins. Figures C and D are still images of the Bright field channel at the same time points as figures A and B to compare how much of the cell population fluoresces. There is much heterogeneity in the population as most cells fluoresce at least one fluorophore type, half incorporated both, while some do not fluoresce at all.

In these still images of the double fluorophore system from Figure 10B, one can see that while there are yellow/ orange cells (mKATE2 and GFP fluorescence), most only fluoresce with one fluorophore. From Figures 10A and 10C, it can observed that the population is nearly all fluorescent. However, by the end, it is apparent that (10B v 10D) many cells still circumvent the fluorophore system or do not incorporate both fluorophores.

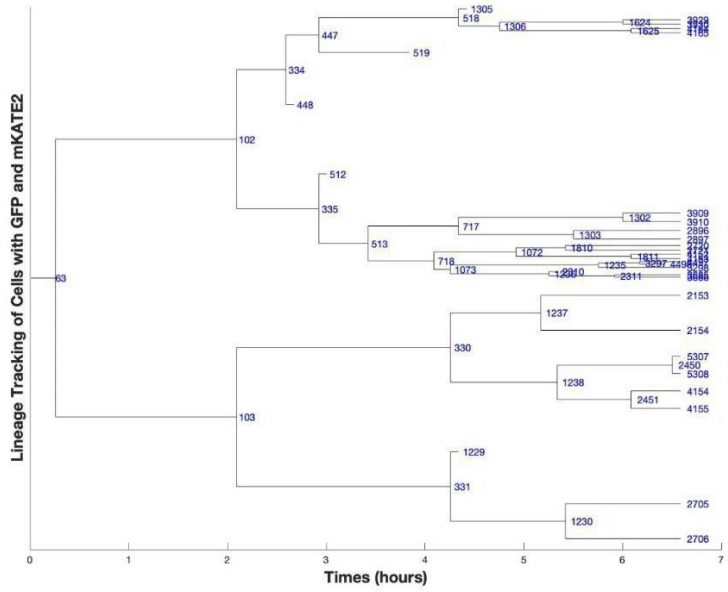

Figure 11: A lineage tree of a cell that contains fluoresces GFP and mKATE2. There are many cell divisions, indicating that once both proteins are incorporated, the cells can grow and go on to create many daughter cells, which indicates a relatively healthy population.

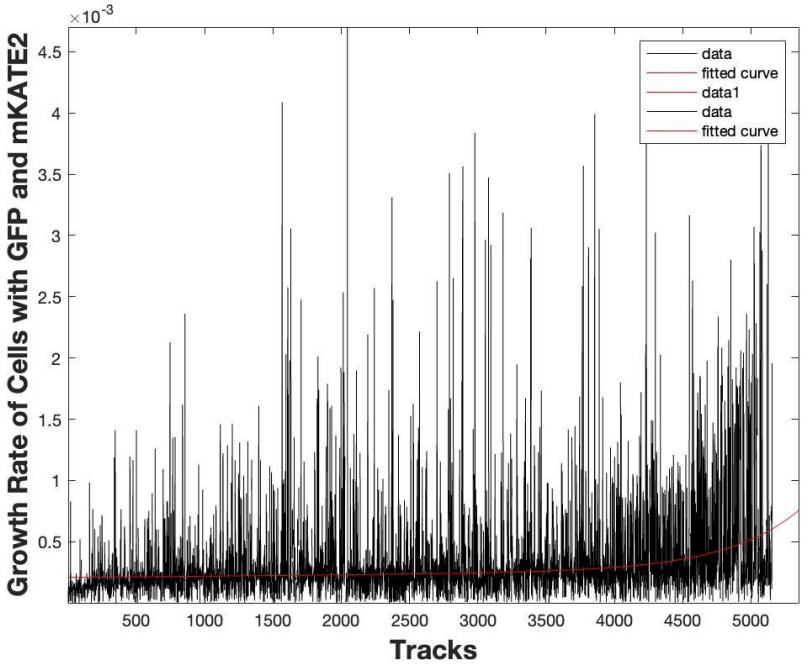

Figure 12: A graph of cells dividing and the growth rate of these cells. There is a large amount of tracks indicating population health.

Some solutions can implemented to improve this system, but considering a CRISPR system will need to be added later, it has proven a path that may need alternative fixes via software.

Cell Tracking with dCNNs

A way to circumvent using a double fluorophore system within cell tracking, which is the project's overall goal, is by using machine learning, specifically dCNNs (de-convolutional Neural Networks). Since these networks can be trained on thousands of captured films, such as only to recognize fluorescence, this makes it a more favorable method. Due to the ability of a dCNN to increase the depth of special features, it can identify and track, which allows for the tracking of cells using just fluorescence or other markers.

Traditional segmentation is still used to a certain extent, as the CyAn algorithm still generates the ground truth (binary mask) for training the system to use and train the dCNN system developed for *Z. mobilis*. Unfortunately, this system requires much more development as it has been trained on only around 1,500 images.

In the methods, it was described that a certain amount of data (movies) cleaning needs to happen before any form of network training can begin. However, the results are so far promising. Figure 13 shows that there is promise in using this method as it can provide more of a computational challenge. However, it would allow the focus to shift from fluorescent protein incorporation to CRISPR engineering, as that is the primary goal.

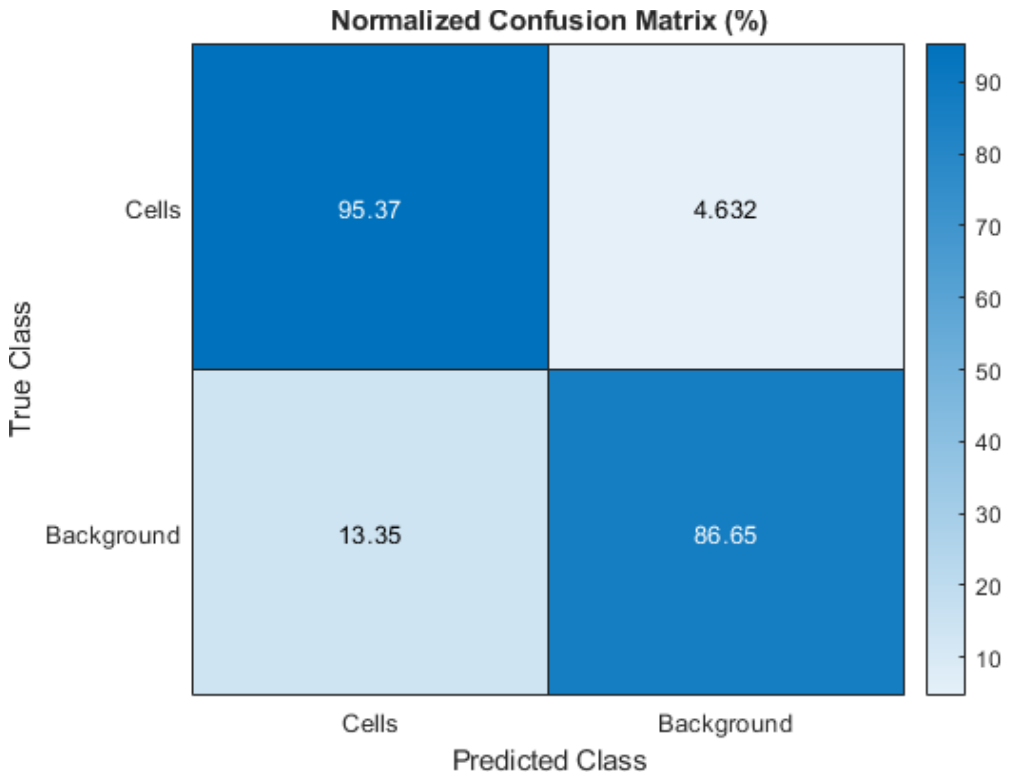

Figure 13: This shows how well the trained dCNN can properly classify the cell and background of an image. As can be seen, it can be accurate most of the time, but there are still things that could be improved, such as 13.35 % of backgrounds getting classified as cells instead.

While Figure 3.13 is impressive in values more data needs to be acquired, cleaned, and collected to train the network to avoid overtraining. Overtraining is an occurrence that is best described as when someone memorizes a story but does not actually learn how to read, so if the pages of a story were presented out of the memorized order, the network (person) would not know to adapt as it had not learned to identify specific cell types, but the pages. In the same way, the network, if there is too small of a data pool, will not learn to identify only cells but small quirks/artifacts in the images of the data training set.

Conclusion

The findings suggest that there is promise in studying manipulations of *Z. mobilis*'s metabolism using single-cell segmentation. By studying cell division and growth rates of

populations, there may be answers about cell growth and behavior concerning the incorporation of plasmids (GFP, mKATE2, GFP CRISPRi knockdown) on population health and cell division. The study of cell division under the condition of dinitrogen fixation is also fascinating as the cells can fix N2, but there appear to be two types of cells growing or two types of reactions from the cells. This is an experiment worth repeating to understand how these cells behave.

Upon looking at the results, there may be more effective methods to understand the CRISPRi GFP knockdown system than a two-fluorophore system. Testing to determine whether a CRISPRi vector system would allow the cell population to develop with two fluorophores is of interest to make sure the method does/ does not hold promise. At the time of writing, this system appears to be the wrong path to head down, given how challenging it is to get consistent expressions of fluorophores within the cells. Specifically, mKATE2 would need to stay at a consistent signal while GFP is depleted, but since not all the cells even express mKATE2, it would be a challenge to pursue a tracking system using this method.

Cell tracking by using dCNNs is a viable path that would simplify some of the microbiology aspects of the project and allow the focus to be on understanding what these various cellular and environmental manipulations do to the cells. Further, a more robust fluorescence measurement method and population tracking would be interesting to develop. Developing more advanced lineage maps and diagrams would provide answers and give a deeper understanding of cellular metabolism.

References

1. Battenberg, E.F., & Bischofs-Pfeifer, I. (2006). A System for Automatic Cell Segmentation of Bacterial Microscopy Images.
2. **Fenster JA,** Werner AZ, **Tay JW, Gillen M,** Schirokauer L, **Hill NC**, Watson A, Ramirez KJ, Johnson CW, Beckham GT, **Cameron JC**, Eckert CA. "Dynamic and single cell characterization of a CRISPR-interference toolset in Pseudomonas putida KT2440 for β-ketoadipate production from p-coumarate" *Metabolic Engineering Communications* (2022), https://doi.org/10.1016/j.mec.2022.e00204.
3. Keraudren, K., Spitaler, M., Braga, V.M.M., Rueckert, D., & Pizarro, L. (2011). Two-step watershed segmentation of epithelial cells. Journal of Microscopy, 271(1), 48-58. http://www.miaab.org/miaab-2011-heidelberg-papers.html.
4. Martien, J. I., Trujillo, E. A., Jacobson, T. B., Tatli, M., Hebert, A. S., Stevenson, D. M., Coon, J. J., & Amador-Noguez, D. (2021). Metabolic Remodeling during Nitrogen Fixation in Zymomonas mobilis. *mSystems*, *6*(6), e0098721. https://doi.org/10.1128/mSystems.00987-21.
5. Tay, J. W., & Cameron, J. C. (2021). CyAn: A MATLAB toolbox for image and data analysis of cyanobacteria. bioRxiv. https://doi.org/10.1101/2021.09.09.459068.
6. Selinummi, J., Ruusuvuori, P., Podolsky, I., Ozinsky, A. O., & Gold, E. S. (2009). Bright field microscopy as an alternative to whole cell fluorescence in automated analysis of macrophage images. Journal of Microscopy, 235(1), 84-93. https://doi.org/10.1371/journal.pone.0007497.
7. Lal, P. B., Wells, F., Myers, K. S., Banerjee, R., Guss, A. M., & Kiley, P. J. (2021). Improving Mobilization of Foreign DNA into Zymomonas mobilis Strain ZM4 by Removal of Multiple Restriction Systems. *Applied and environmental microbiology*, *87*(19), e0080821. https://doi.org/10.1128/AEM.00808-21
8. Palamae, S., Choorit, W., Chatsungnoen, T., & Chisti, Y. (2020). Simultaneous nitrogen fixation and ethanol production by Zymomonas mobilis. Journal of Biotechnology, 313, 38-43. doi: 10.1016/j.jbiotec.2020.03.016
9. Yang, S., Fei, Q., Zhang, Y., Contreras, L. M., Utturkar, S. M., Brown, S. D., Himmel, M. E., & Zhang, M. (2016). Zymomonas mobilis as a model system for production of biofuels and biochemicals. Microbial Biotechnology, 9(6), 699-717. doi: 10.1111/1751-7915.12408
10. Ioffe, S. and Szegedy, C., 2015. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv preprint arXiv:1502.03167*.
11. Xu, B., Wang, N., Chen, T. and Li, M., 2015. Empirical evaluation of rectified activations in convolutional network. *arXiv preprint arXiv:1505.00853*.
12. Banta, A. B., Enright, A. L., Siletti, C., & Peters, J. M. (2020). A High-Efficacy CRISPR Interference System for Gene Function Discovery in Zymomonas mobilis. *Applied and environmental microbiology*, *86*(23), e01621-20. https://doi.org/10.1128/AEM.01621-20
13. Mehta, S., Koncel-Kedziorski, R., Rastegari, M. and Hajishirzi, H., 2018. Pyramidal Recurrent Unit for Language Modeling. *arXiv preprint arXiv:1808.09029*.
14. Andrei-Alexandru Tulbure, Adrian-Alexandru Tulbure, Eva-Henrietta Dulf, A review on modern defect detection models using DCNNs – Deep convolutional neural networks, Journal of Advanced Research,Volume 35, 2022, Pages 33-48, ISSN 2090-1232, https://doi.org/10.1016/j.jare.2021.03.015.
15. E. Shelhamer, J. Long and T. Darrell, "Fully Convolutional Networks for Semantic Segmentation," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 4, pp. 640-651, 1 April 2017, doi: 10.1109/TPAMI.2016.2572683.

16. Banta, A. B., Enright, A. L., Siletti, C., & Peters, J. M. (2020). A High-Efficacy CRISPR Interference System for Gene Function Discovery in Zymomonas mobilis. *Applied and environmental microbiology*, *86*(23), e01621-20. https://doi.org/10.1128/AEM.01621-20