A Novel Approach to Tracking Mitochondrial and Peroxisomal Proteins

Proposed Start Date

July 4, 2022

Introduction

The use of artificial intelligence in biological research has increased considerably for its ability to analyze data efficiently and economically. Currently, it is primarily used in bioinformatics for genome studies. Historically, AI has also been used in image analysis to complete tasks such as facial recognition, object classification, and as a preliminary diagnostic tool in medical offices for detecting malignant tumours. The branch of AI that specifically deals with such image analysis is called Computer Vision (CV). I propose a novel approach to protein translocation analysis between organelles in live cells using CV.

During my lab placement at Kim Lab in Toronto, ON over the summer, I will be studying the translocation of protein between the mitochondria and peroxisome. The mitochondria is an organelle that is involved in metabolic pathways that lead to the production of ATP, a molecule that provides energy for most cellular processes. The peroxisome performs tasks such as the oxidation of very long and branched-chain fatty acids, which are essential to the body. For example, the myelin sheath, the protective outer membrane for neurons, is composed of fatty acids. Problematic peroxisome function can damage the myelin sheath, leading to adrenoleukodystrophy, a severe brain disorder. Peroxisomes also contain reactive oxidative species such as hydrogen peroxide which oxidizes and destroys cell waste.

A relatively recent discovery showed that many of proteins appearing in mitochondria can also appear in the peroxisome, and vice-versa via translocation. (Koyano, 2019)

An example of this interaction is between PEX2 and USP30. PEX2 is a peroxisomal signalling protein responsible for adding ubiquitin, a protein marker, to the peroxisome, which signals it for pexophagy, destruction of the organelle. Conversely, USP30 is a deubiquitinase protein that was originally thought to exist only on the mitochondria to remove ubiquitin. It was found that USP30 also exists in small numbers in peroxisomes, showing that on peroxisomes, USP30 and PEX2 function in equilibrium so to maintain basal peroxisome count and quality. (Riccio, 2019)

Problem, Hypothesis, and Aims

Findings like this illustrate how much more remains to be discovered in the field of proteomics, and I believe that an effective and efficient way of potentially elucidating new protein to protein interaction mechanisms and colocalization is to use CV to analyze these proteins.

I will be using fluorescent microscopy, a technique where fluorescently tagged proteins are visualized *in vivo* by energizing the protein at specific light wavelengths. To create fluorescent protein, DNA encoding for protein domains that fluoresce under certain light wavelengths is spliced into existing protein-encoding genes in bacteria to produce fluorescent proteins for use in experiments. Images can then be taken for analysis.

I would develop my CV software that would be able to track fluorescently tagged proteins in the cell over time. My software aims to specifically:

- Use the fluorescent signal of proteins to determine cell borders
- Detect the location of peroxisomes and mitochondria in cells via fluorescent protein expression using fluorescent domains such as GFP and RFP

- Detect proteins in each organelle and investigate how this may be affected by cellular conditions
- Suggest that a certain protein is involved in both organelles from past datasets

My hope is that my project can set a standard for CV-based molecule tracking in cells, and if all goes well, to potentially release it as an open-source platform for other scientists to use. The proposed platform does not have to be exclusive to mitochondrial and peroxisomal proteins.

Methodology

Preparation Work – Pre-Start Date

I will be working at the Kim Lab throughout the summer, researching the nature of the protein, BCL-Rambo. The protein induces apoptosis in cells through a unique C-terminal expression (Kataoka, 2001). The protein shows no interaction with the rest of the BCL family, both pro and anti-apoptotic. Interestingly, the method of inducing apoptosis is not controlled by the mitochondrial pro-apoptotic pathway, Bcl-x_2, or the upstream death receptors, FLIP and FADD-DN. It is however controlled by caspase inhibitors, IAPs. (Kataoka, 2001) In mammalian cells, the protein is localized to the mitochondria, but at the lab, we are curious to see where BCL-Rambo localizes to in the case where: (a) it is overexpressed in the cell and (b) if the cell's mitochondria are damaged or absent.

To identify where BCL-Rambo is going in the cell under these conditions, co-localization assays will be used. In this assay, the protein of interest and a control protein with known localization patterns are transformed with fluorescent tags and are imaged. These proteins' movement can be recorded over time, and have their fluorescent images merged together to provide insight into the protein of interest's localization patterns.

During this time, data will collected and manually analyzed, measuring key factors such as fluorescence intensity, and the peroxisomes' size, morphology, and density.

I also plan on following the Computer Vision lectures series by Stanford University's School of Engineering to develop some intuition with tools used in CV, and looking at the libraries the OpenCV offers for semantic image segmentation.

Cell Recognition – 1-2 Weeks

The first component of the program will be to create the cell-border recognition portion of my software. It is often difficult, especially to the untrained eye, to identify cell membranes from fluorescent images.

To do this, an algorithm called canny edge detection (CED) will be used. The steps of the algorithm are:

1. Noise reduction
2. Gradient calculation
3. Non-maximum suppression
4. Double threshold
5. Edge Tracking by Hysteresis

In lay terms, an image that has its border analyzed through CED will be blurred to remove 'noise', have a rough outline constructed, and then be fine-tuned until an acceptable border outline is produced.

I expect that modifications will be made to completed border images to extraneous border lines inside the circled cells.

The dataset for this program will come both from the fluorescent images that I will be taking earlier in the summer and from the bank of images that the lab has amassed over time.

In the case that the COVID-19 pandemic prevents me from going to the lab earlier in the summer, I will exclusively use the lab's dataset to train my model.

Co-Localization Assay Analysis – 1-2 Weeks

The next portion of the program would be performing the analysis of the co-localization assay. The aim of this portion of the program is to create a robust system that is able to correctly identify whether or not a protein co-localizes with a control protein to a specific area in the cell.

There are a variety of factors that can be used to identify co-localization.

Firstly, I could look at the pixel intensity of the coloured proteins compare the values. Statistical analyses could then be applied to it such as the Pearson's correlation coefficient and Mander's Colocalization Coefficients (Dunn, 2011) (Khushi, 2018). Then, the training dataset can be used to determine the constraints and sensitivity of this data.

Next, I could also attempt to locate the peroxisomes in the cells and detect the presence of the pixel colour itself within the organelle, as well as comparing it to the control protein. This could potentially reduce error due to endogenous protein levels and chance pixel overlap in merged images. This could be done using semantic image segmentation, where a pixel map corresponding to the location of the peroxisome is created (Jordan, 2018).

Buffer Time – 1 Week

A week of buffer time is left in the case that either portion of the project takes longer than expected, and to allow for time developing things like an intuitive GUI for use in the lab and any additional testing.

Research Advisor

For this project, I will be under the supervision of Dr. Peter Kim at SickKids hospital in Toronto, Canada and the University of Toronto's Department of Biochemistry, along with Victoria Riccio, a PhD candidate at the University of Toronto at Dr. Kim's lab.

Dr. Kim will be providing me with weekly progress reports, suggest pertinent literature, and provide me with existing datasets.

Victoria will be working with me in analyzing the microscopy data and tagging the datasets.

Outcomes

The final product of my project would be a robust CV-based model that could accurately detect protein localization patterns in the peroxisome, and potentially other organelles. While this program is being developed in the context of analyzing mitochondrial proteins going to the peroxisome, I believe that more training with other molecules and organelles can turn my project into a general tool that can be used at all labs for a variety of use cases.

Ethics

This project does not involve working with humans and/or animals.

Sources

Dunn, K. W., Kamocka, M. M., & McDonald, J. H. (2011). A practical guide to evaluating colocalization in biological microscopy. *American Journal of Physiology-Cell Physiology*, *300*(4), C723–C742. https://doi.org/10.1152/ajpcell.00462.2010

Jordan, J. (2018, May 22). *An overview of semantic image segmentation.* Jeremy Jordan. https://www.jeremyjordan.me/semantic-segmentation/

Kataoka, T., Holler, N., Micheau, O., Martinon, F., Tinel, A., Hofmann, K., & Tschopp, J. (2001). Bcl-rambo, a Novel Bcl-2 Homologue That Induces Apoptosis via Its Unique C-terminal Extension. *Journal of Biological Chemistry*, *276*(22), 19548–19554. https://doi.org/10.1074/jbc.m010520200

Koyano, F., Yamano, K., Kosako, H., Kimura, Y., Kimura, M., Fujiki, Y., Tanaka, K., & Matsuda, N. (2019). Parkin-mediated ubiquitylation redistributes MITOL/March5 from mitochondria to peroxisomes. EMBO reports, 20(12), e47728.

Khushi, M., Napier, C. E., Smyth, C. M., Reddel, R. R., & Arthur, J. W. (2017). MatCol: a tool to measure fluorescence signal colocalisation in biological systems. *Scientific Reports*, *7*(1), 8879. https://doi.org/10.1038/s41598-017-08786-1

Riccio, V., McQuibban, A., & Kim, P. (2019). USP30: protector of peroxisomes and mitochondria. Molecular & Cellular Oncology, 6, 1-2.