Rapid Multiplexing of Proteins for Deciphering Spatially Resolved Biophysical Organelle Networks

Undergraduate Thesis
Presented to
The Academic Faculty

by

Yukina Yajima

In Partial Fulfillment
of the Requirements for the
Undergraduate Research Option in
Biomedical Engineering

Georgia Institute of Technology
December 2021

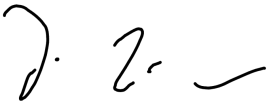
Rapid Multiplexing of Proteins for Deciphering Spatially Resolved Biophysical Organelle Networks

Approved by:

Dr. Ahmet Coskun, Advisor
Department of Biomedical Engineering
Georgia Institute of Technology

_____

Dr. Johannes Leisen
Department of Chemistry
Georgia Institute of Technology



_____

Date Approved: 12/17/2021

Table of Contents

1. Chapter 1 - Introduction 4
2. Chapter 2 - Materials and Methods 9
3. Chapter 3 - Results 17
4. Chapter 4 - Conclusion 25
5. References 26

Chapter 1

Introduction

Image analysis technology for investigating cell features in biological research has been advancing over the past several decades. Finding protein associations within cells through image analysis has the potential to deepen the understanding of disease microenvironment and how the structural organization changes as cells transition from normal to abnormal cell behaviors. Our study aims to establish a robust, rapid protein analysis pipeline for deciphering spatial organelle networks within a single cell using the CellProfiler software and deep learning techniques.

To observe and analyze cell features, visual inspection by scientists has been a powerful and common image analysis method, but this method is often more time-consuming, subjective, and nonquantitative [1]. The application of analysis software as well as deep learning techniques have shown to help improve the segmentation and analysis processes for identifying and visualizing cellular and subcellular features [2]. We're specifically interested in observing key organelles in stem cells such as the nucleus, mitochondria, Golgi and endoplasmic reticulum. Identifying these organelles through segmentation and quantifying biophysical features such as cell size, symmetry, intensity, and shape across multiple cell images can lead to discoveries of trends and associations that exist between and within organelles. The result can then be used for statistical analyses such as t-SNE analysis and ratiometric analysis, as well as for visualization of the structures in 2D and 3D for further spatial analysis. Our ultimate goal is to propose an accurate and reproducible multiplexed protein imaging and analysis method that could aid in better visualizing and understanding various disease states and microenvironments, which could then be used to find or design drugs and therapy methods that work best depending on patient needs.

The use of mesenchymal stem cells have become a promising therapy method for treating a variety of diseases including myocardial infarction and diabetes, as they have the ability to repair damaged cells by differentiating into replacement cells and by modulating immune responses [3, 4]. Therefore, analyzing the spatial organelle networks within mesenchymal stem cells can lead to a better understanding of cell functions in order to design appropriate treatment methods. Although there have been recent studies focusing on spatial organelle analysis, these studies have used spatial data of each organelle from different cells, which reduces the accuracy of the resulting spatial information. A highly multiplexed protein imaging and analysis method can be used to obtain spatial information of organelles within the same cell, which can then be used to compare and understand differences in spatial organization between normal and abnormal cells.

Literature Review:

Our study aims to develop a highly multiplex protein imaging and analysis method for stem cell organelle networks within a single cell. Previous studies with similar focus have explored how intracellular structures exhibit cell-to-cell variation in biophysical features such as the location and size of organelles. Gut et al. developed a data-driven computer vision approach to generate multiplexed protein maps (MPMs), which was then used to compare subcellular spatial protein distribution between single cells based on the knowledge that the phenotype of a cell is influenced by aspects such as their functional state, abundance, and morphology of organelles [5]. Similarly, Viana et al. developed a reproducible microscope imaging pipeline and investigated the impact of cell and nuclear size on the variation in the size of organelles using 25 isogenic human induced pluripotent stem cell (hiPSC) lines [6]. Their results showed that the

impact of cell and nuclear size metrics on the variation in cytoplasmic structure volumes is structure-dependent, and that this correlation is consistent with the wide range of cell functions that these structures regulate [6]. Although our study uses mesenchymal stem cells instead of the hiPSC, the findings and approach from this study can be taken to the next level by including other key biophysical characteristics into our focus, such the location, symmetry, shape, and intensity of organelles and associated proteins, which can then be compared to find trends and interactions between organelles.

Among several automated imaging software packages that have been developed in recent years is a software called CellProfiler [7]. CellProfiler is a free, open-source system that was designed for flexible, high-throughput cell image analysis [7]. The main purpose of this software is to identify and measure a variety of biological objects in images [1]. Studies that utilize CellProfiler for image analysis have shown that the reliability and applicability of this software is promising. Weiss et al. compared the results for the quantification of Schwann cell (SC) purity and proliferation between the automated CellProfiler program and manual counting [8]. No significant difference was found between the two methods for the total cell count, SC purity, and SC proliferation rate, which suggests the accuracy of the established CellProfiler pipelines for the analysis of these features [8]. One limitation with this study is that the analysis was done on a 2D level, so maintaining accuracy in 3D spatial analysis for organelles in a single cell can contribute to a better understanding of the structures of disease microenvironments. This is one of the goals of our study, as proposing a more detailed and precise 3D analysis method would be a notable step forward in better understanding how stem cells can be utilized for disease treatment. The location and morphology of the proteins would be observed after segmentation, where information such as the size of organelles, intensity of each marker and the colocalization

of multiple markers can be extracted. This information can then be used for 3D visualization of organelle networks within a single cell without the use of spinning-disk confocal microscopes.

It has also been found that CellProfiler in combination with machine learning techniques is suitable for the analysis of proteome dynamics on the single-cell level [9, 10]. Chong et al. performed an evaluation on protein abundance and localization on a proteome scale of a single yeast cell, then developed binary classifiers to generate localization data [9]. This field of study is prominent because the quantitative approach to analyze protein abundance and localization can be applied to various types of studies such as comparative studies, single-cell analyses, modeling, and prediction [9]. The possibility of using machine learning for pattern recognition in addition to a single-cell analysis using CellProfiler is worth noting as this increases the flexibility in the types of cells and features of analysis.

Other studies have focused on the applicability of machine learning to identify and classify various features within a single cell with higher accuracy and speed [11, 12]. Kriston-Vizi et al. showed that an automated high-content screening image analysis can be performed to identify selected compounds, which had results similar to that of threshold-based method [12]. Similarly, Hamilton et al. found that classification of subcellular protein localization using machine learning and threshold adjacency statistics was able to provide high accuracy with a calculation speed faster by an order of magnitude compared to other commonly used methods [11]. These findings demonstrate a remarkable step forward in the application of machine learning techniques in improving the speed and accuracy of the automated analysis methods compared to conventional visual analysis methods as well as previously released analysis software.

The above methods demonstrate a significant advancement towards the application and automation of image analysis methods for various features on a single-cell level. These studies provide insights on how open-source software packages such as CellProfiler can facilitate the automation of image analysis as well as how deep learning technology could be added for a more flexible and accurate pipeline for analyzing a variety of cell types and features of focus. A previous study has trained their segmentation algorithm using a subset of 600 images through two training assays, which were validated manually with an in-house scoring interface in Python using Napari [6]. A similar training and validation method can be used to develop and validate a robust algorithm using deep learning to segment and analyze organelle networks. Our study builds upon these past findings to apply these automated image analysis methods for investigating spatial organelle networks within a single stem cell. This information about the associations found in stem cells on a single-cell level can eventually aid researchers in better understanding microenvironments of normal and abnormal cells and to explore treatment methods using mesenchymal stem cells.

Chapter 2

Materials and Methods

2.1 Preparing Images

Images of mesenchymal stem cells to be processed were chosen based on the clarity and consistency among all markers. Mesenchymal stem cells extracted from two sources were used and compared: umbilical cord (UC) and bone marrow (BM) stem cells. Over 100 file formats are currently readable by CellProfiler, including BMP, GIF, JPG, PNG, TIF, DIB, LSM, and FLEX. For our study, all images were kept in TIF format because it has the ability to store color and transparency information using an alpha channel, which allows for fast compression and decompression while retaining high image quality. Fluorescent imaging was performed to obtain all the images in our laboratory using reagents and the microscope listed in Table 2.1. 11 markers were used to investigate 7 organelles: DAPI, HSP60, TOM20, Nucleolin, β-tubulin, ATF6, GOLPH4, Sortilin, Concanavalin A, Phalloidin, and WGA. The full names of the abbreviated markers are included in **Table 2.1**. The target organelles of each of the markers are shown in**Table 2.2**, some of which show overlaps in their target. These markers were selected in order to capture a variety of organelles pictured in **Figure 1**, which allows us to compare across cells and determine if any of the organelles show trends or unique characteristics [13]. Specifically, we focused on biophysical characteristics such as major/minor axis lengths, perimeter, and intensity.. Both masked images with already selected cells and unmasked images can be used as the input of the pipeline.

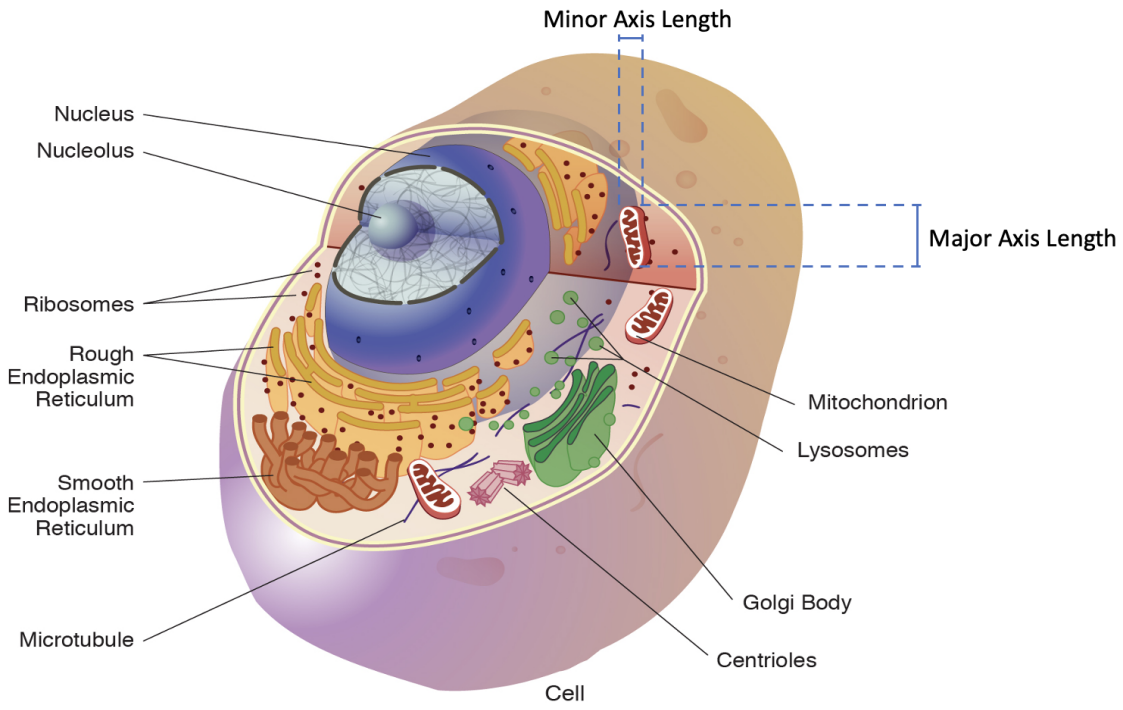


Figure 1: Organelles labeled in a single cell. Dashed lines show an example of how minor/major axis lengths, two of the parameters of interest, are measured for a segmented mitochondria.

REAGENT OR RESOURCE	**SOURCE**
Antibodies	
ATF6: Activating Transcription Factor 6 (Rabbit monoclonal)	Abcam
B Tubulin (Mouse monoclonal)	Santa Cruz Biotechnology, Inc
Concanavalin A (Alexa Fluor®488 conjugate)	Invitrogen
GOLPH4: Golgi Integral Membrane Protein 4 (Rabbit monoclonal)	Abcam
HSP60: Chaperonins (Rabbit monoclonal)	Abcam

Nucleolin (Rabbit monoclonal)	Abcam
Phalloidin (Alexa Fluor®555 conjugate)	Invitrogen
Sortilin (Rabbit monoclonal)	Abcam
Tom20: Translocase of outer mitochondrial membrane 20 (Mouse monoclonal)	Santa Cruz Biotechnology, Inc
Wheat Germ Agglutinin (Alexa Fluor®647 conjugate)	Invitrogen
Chemicals	
10x Phosphate-buffered saline	Sigma-Aldrich
Alexa Fluor®647 Conjugation Kit (Fast) - Lightning-Link®	Abcam
CSM Blocking solution	1% BSA, 0.1% Triton X-100 in PBS
DAPI (4',6-Diamidino-2-Phenylindole, Dihydrochloride)	Invitrogen
Hydrogen peroxide solution (H2O2 30% (wt/vol))	Fisher Chemical
Sodium hydroxide (NaOH) solution 1M	HiMedia
Triton X-100	Fisher BioReagents
Software and Algorithms	
NIS Imaging Software	https://www.nikonmetrology.com/en-us/industrial-microscopes/nis-software-nis-elements-microscope-imaging-software
RNAscope® HiPlex Image Registration Software	https://acdbio.com/image-registration-software
ImageJ	https://imagej.nih.gov/ij/
Colocalization Colormap plugin	https://sites.google.com/site/colocalizationcolormap/home

Coloc2 plugin	https://imagej.net/Coloc2
Python	https://www.python.org/
PARC	https://github.com/ShobiStassen/PARC
Microscope	
Eclipse TE2000-U Inverted Research microscope with motorized stage and filter wheel	Nikon

Table 2.1: List of all antibodies, chemicals, software/algorithms, and microscope used to obtain images of stem cells for this study.

Marker	Target	Function
HSP60	Mitochondria	Heat shock protein/chaperonin - molecular chaperone which assists protein folding; essential for the protein import into the mitochondrial matrix
TOM20	Mitochondria	Translocase of outer membrane of mitochondria - facilitates the import of nucleus-encoded precursor proteins across the mitochondrial outer membrane
Nucleolin	Nucleoli	Induces chromatin decondensation by binding to histone H1. Also involved in rDNA transcription, rRNA maturation, ribosome assembly and nucleo-cytoplasmic transport.
β-tubulin	Microtubules (β-tubulin)	Tubulin is the major constituent of microtubules. Heterodimer of α- and β-subunits. Microtubules affect cell shape, cell transport, cell motility, and cell division by interacting with microtubule associated proteins (MAPs). β isotypes are capable of forming the mitotic spindle and the interphase network.
ATF6	Endoplasmic Reticulum	Activating transcription factor 6 - Acts during ER stress by activating unfolding protein response (UPR) target genes
GOLPH4	Golgi apparatus	Golgi integral membrane protein - Plays a role in endosome to Golgi protein trafficking; mediates protein transport along the late endosome-bypass pathway from the early endosome to the Golgi
Sortilin	Localized to membranes of the Golgi apparatus, endoplasmic reticulum, endosomes, lysosomes and nucleus.	Sorting receptor in the Golgi compartment and a clearance receptor on the cell surface. Regulate intracellular protein transport through secretory or endocytic pathways.
Concanavalin A	Endoplasmic reticulum	Con A selectively binds to a-mannopyranosyl and a-glucopyranosyl residues. These are commonly found in the cell membrane of mammalian cells and tissues.
Phalloidin	F-Actin	Phalloidin is a peptide belonging to a family of toxins originating from the deadly Amanita phalloides mushroom. It binds to filamentous actin (F-actin). Actin takes part in many cellular functions - ranging from cell motility and the maintenance of cell shape and polarity to the regulation of transcription.
WGA	Golgi apparatus and plasma membrane	WGA is a carbohydrate-binding lectin that binds to sialic acid and N-acetylglucosamine residues of glycoproteins. WGA conjugates label the cell membrane of mammalian cells and tissues. Also stains various proteins and lipids in the golgi apparatus that are glycosylated.

Table 2.2: List of markers of interest used in this study, their target organelles, and functions.

2.2 Installing and Setting Up CellProfiler

We used the CellProfiler software as the main analysis tool for our study. CellProfiler addresses a number of challenges involved in previously published automated systems. First, this single open-source program has the ability to perform analysis for several purposes such as investigation of cell size, intensity, and quantification [1]. This is a useful improvement as it eliminates the need for using multiple software packages that function separately for different purposes. In addition, CellProfiler does not require any accompanying image acquisition hardware that previously developed programs require and are often costly [1]. Another remarkable difference between CellProfiler and previous programs is its modular design and graphical user interface. The user-friendly design of the program makes it suitable for analysis of various biological objects and assays without requiring programming [1]. This is a notable benefit because there are other flexible analysis programs that allow for interactive analysis, but the need for creating custom algorithms becomes a challenge without knowledge and experience in programming. Due to these benefits we chose to use this software to automate the analysis process and explore how this tool can be used in combination with deep learning techniques for an accurate multiplex analysis of stem cell organelles.

Installation instructions from the web page to install CellProfiler were followed to download the software. All test images were imported and classified based on the marker names under the “NameAndType” module. The output location was manually selected for the data spreadsheet containing the measurement results.

2.3 Measuring Size/Shape Features

To first identify and segment the nuclei of cells, “IdentifyPrimaryObject” was first used on DAPI, a marker often used for visualization of cell nuclei in immunofluorescent histological

studies because of its high affinity for DNA [14, 15]. The global two-class Otsu thresholding method was used, which is an approach that calculates the threshold separating the two classes of pixels, foreground and background, by minimizing the variance within each class [16]. Using the segmented nuclei, the “IdentifySecondaryObject” module was used to segment the general shapes of the organelles marked by each marker using the same thresholding method. After images of every marker were segmented, all the size/shape features were measured using the “MeasureObjectSizeShape" module for all the markers. Finally, the shape/size data were exported into an excel file in the output location specified in the earlier step.

2.3 Measuring Intensity Features

With the segmented images from the previous step, the “MeasureObjectIntensity” module was used to make intensity measurements including the mean and total intensity of the selected area. The units of intensity from this module are described as “intensity units” or “arbitrary intensity units” because microscopes are not calibrated to an absolute scale. Intensity corresponds to the amount of light emitted, which represents the amount of stained protein in the cell [17]. This information can be applied in further analysis such as colocalization, which is a useful method in determining if a protein is localizing to an organelle [18].

2.4 Visualization and Statistical Analyses

The data stored in an excel spreadsheet was manually sorted to extract and analyze only the features of interest, which are area, intensity, major/minor axes, and perimeter of each organelle. The measurements of interest were converted into CSV files for visualization. Heatmaps were generated using Python with dendrograms to compare the biophysical

measurements and to determine any close relationships or associations between markers and between the two sources (UC and BM). The z-score was calculated to evaluate the distribution of the output data. Ratiometric analysis is performed to detect changes to the local environment, and t-distributed stochastic neighbor embedding (t-SNE) analysis to reduce dimensionality for visualizing the data in two dimensions data [19].

Chapter 3

Results

Segmentation was performed to first identify the nuclei, then to identify the organelles. Example images that show the segmentation of the nuclei and the organelles are shown in **Figure 1-2**.

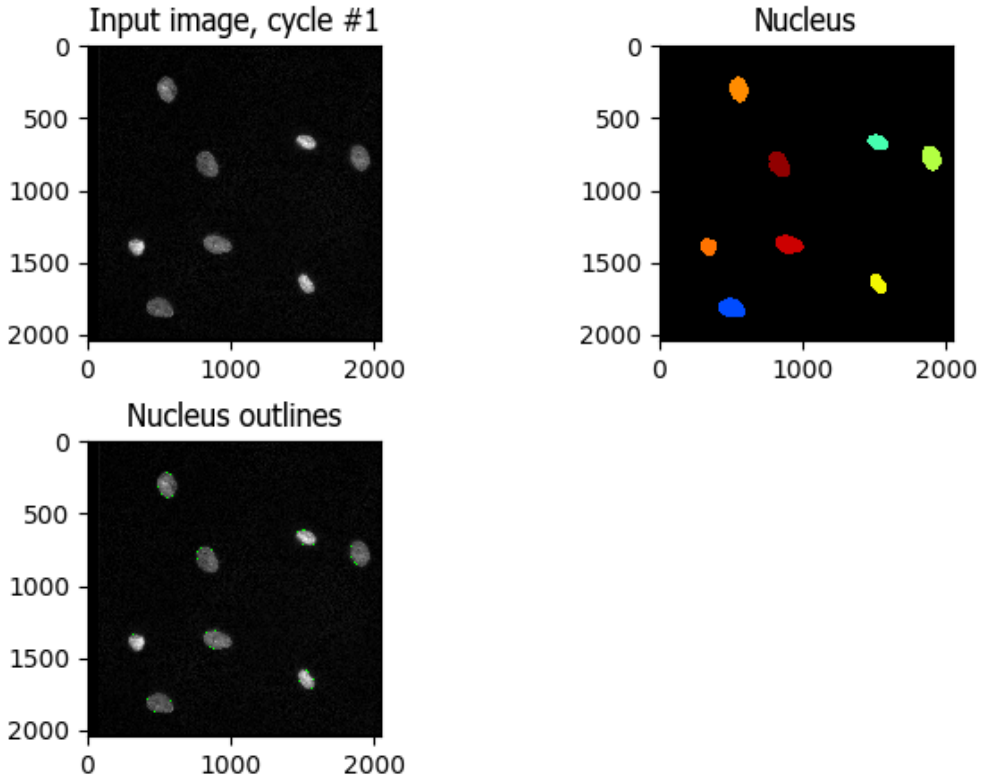


Figure 1: Segmentation steps of the nuclei, consisting of the original image (top left), outlines detected using global Otsu thresholding method with two classes (bottom left), and the segmented nuclei (top right). The segmented nuclei in the output are presented in different colors for clear visualization, which will be used as masks to obtain data for each cell separately.

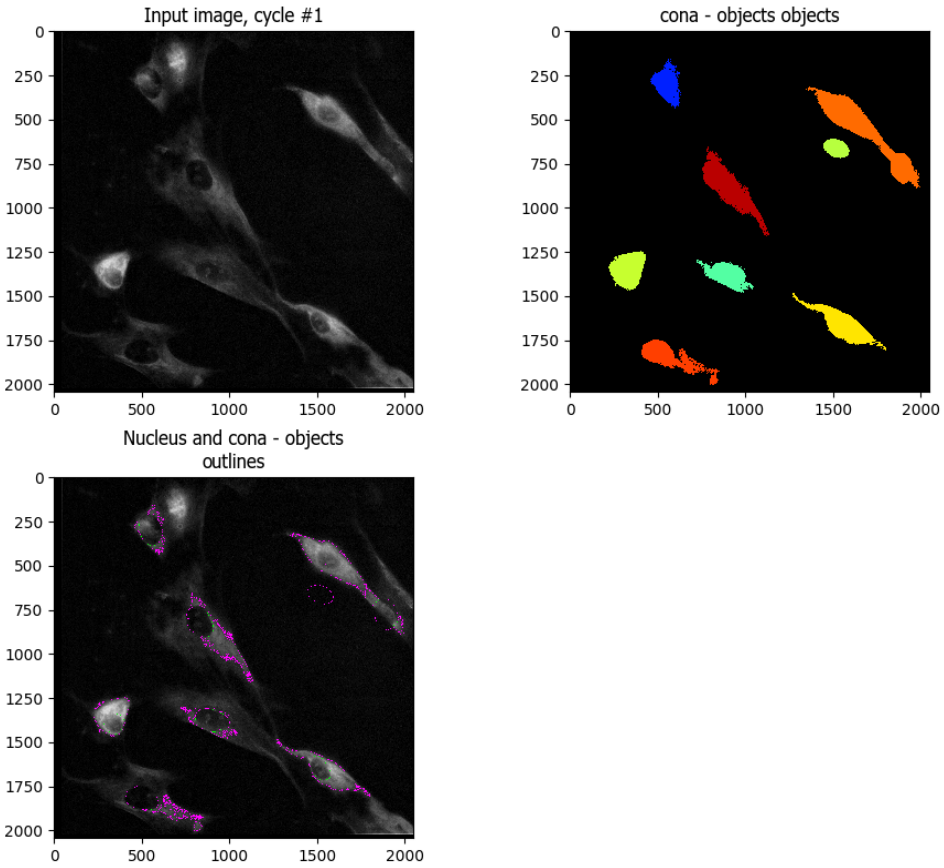


Figure 2: Segmentation steps of the nuclei, consisting of the original image (top left), outlines detected using global Otsu thresholding method with two classes (bottom left), and the segmented nuclei (top right). The segmented organelles in the output are presented in different colors for clear visualization, which will be used as masks to obtain data for each cell separately.

Figures 3-8 are heat maps that quantify morphological features found in images of 14 selected cells (7 cells from umbilical cord and 7 cells from bone marrow) using 10 selected markers. Specifically, the features of focus were major/minor axes, perimeter, and integrated intensity. We plan to continue investigating more features and categorizing the features into the following categories: area/shape features, intensity based features, image quality (blurriness and

granularity), and texture features. Once the data are visualized, trends and associations will be examined across markers as well as across the two cell types, UC and BM.

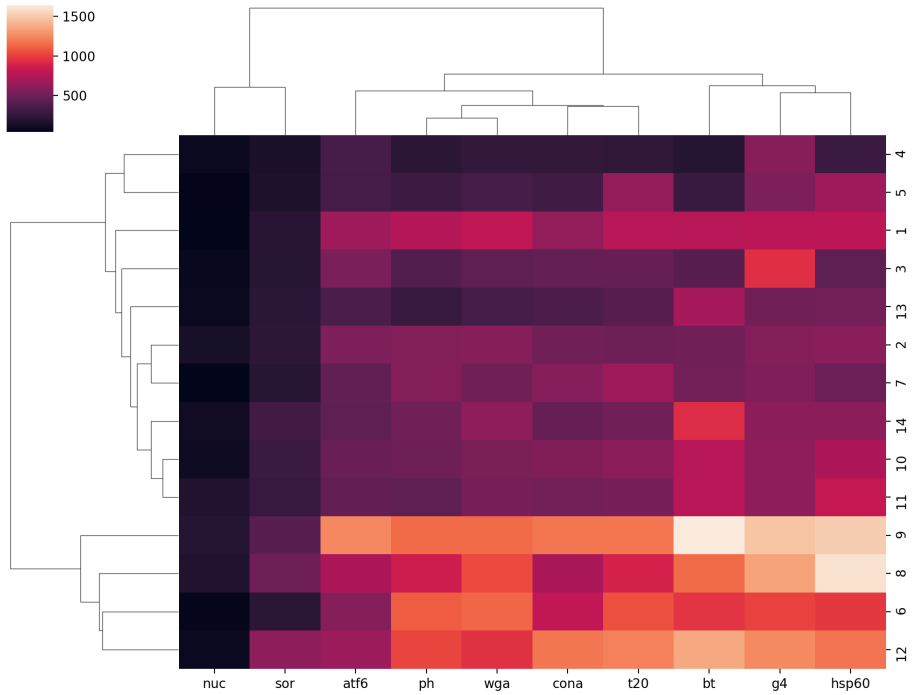


Figure 3: Heatmap comparing the raw major axis length across 10 markers and 14 selected cells
X-axis: 10 selected markers
Y-axis: Cell type and assigned cell number
Color bar: Major axis length, raw (pixels)

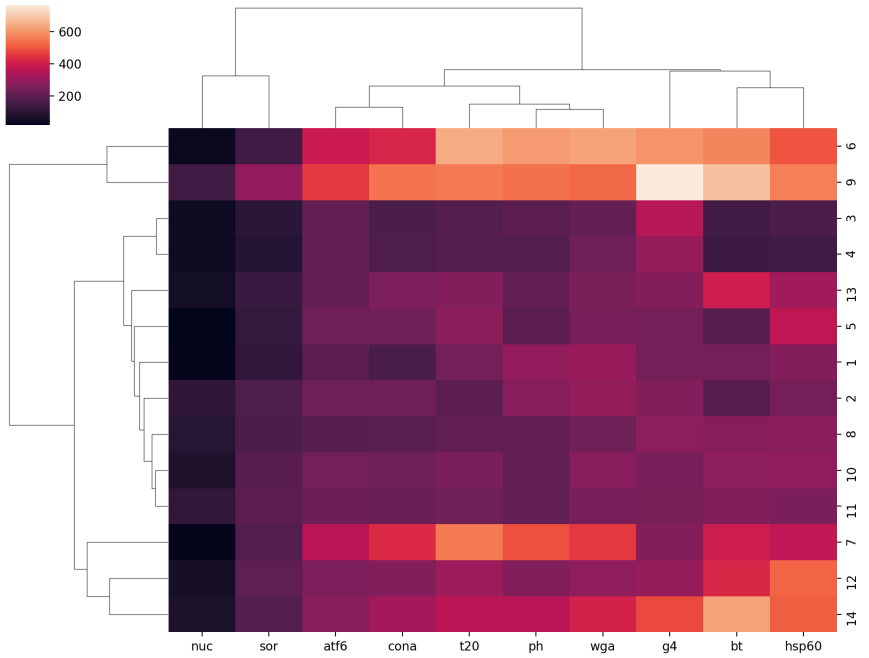


Figure 4: Heatmap comparing the raw minor axis length across 10 markers and 14 selected cells
X-axis: 10 selected markers
Y-axis: Cell type and assigned cell number
Color bar: Minor axis length, raw (pixels)

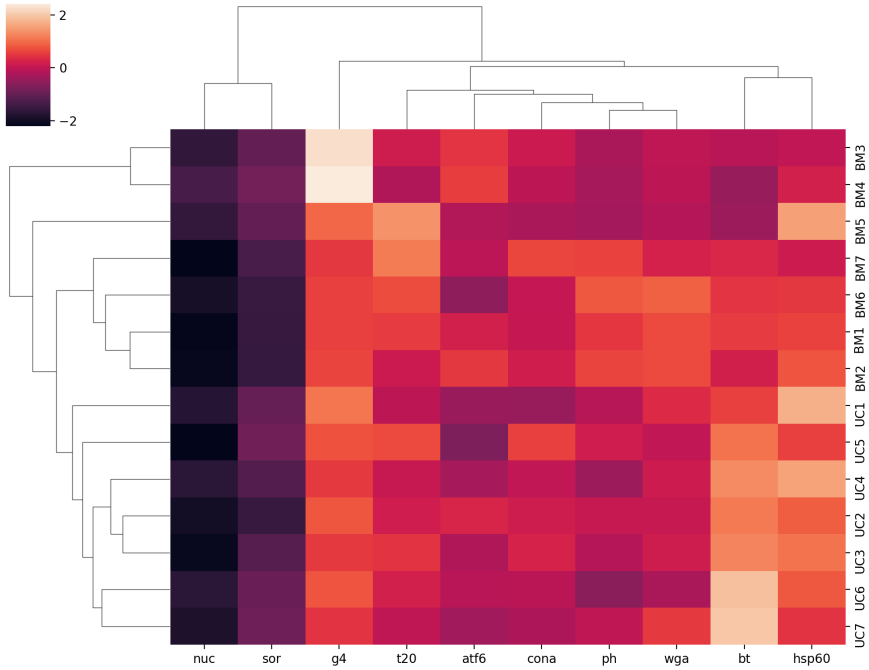


Figure 5: Heatmap comparing the normalized major axis length across 10 markers and 14 selected cells
X-axis: 10 selected markers
Y-axis: Cell type and assigned cell number

Color bar: Minor axis length normalized using z-score (unitless)

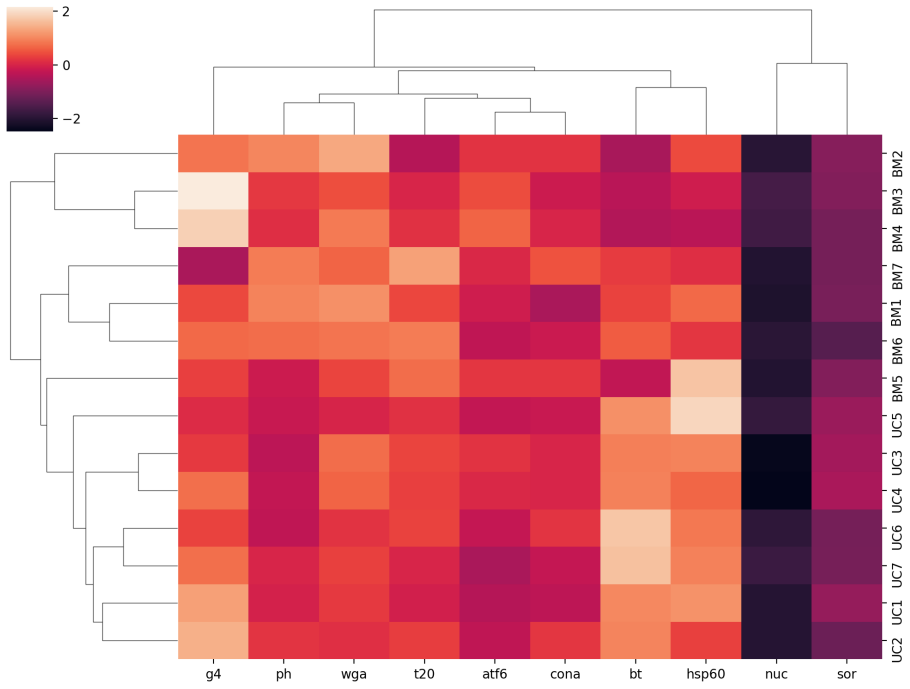


Figure 6: Heatmap comparing the normalized minor axis length across 10 markers and 14 selected cells
X-axis: 10 selected markers
Y-axis: Cell type and assigned cell number
Color bar: Minor axis length normalized using z-score (unitless)

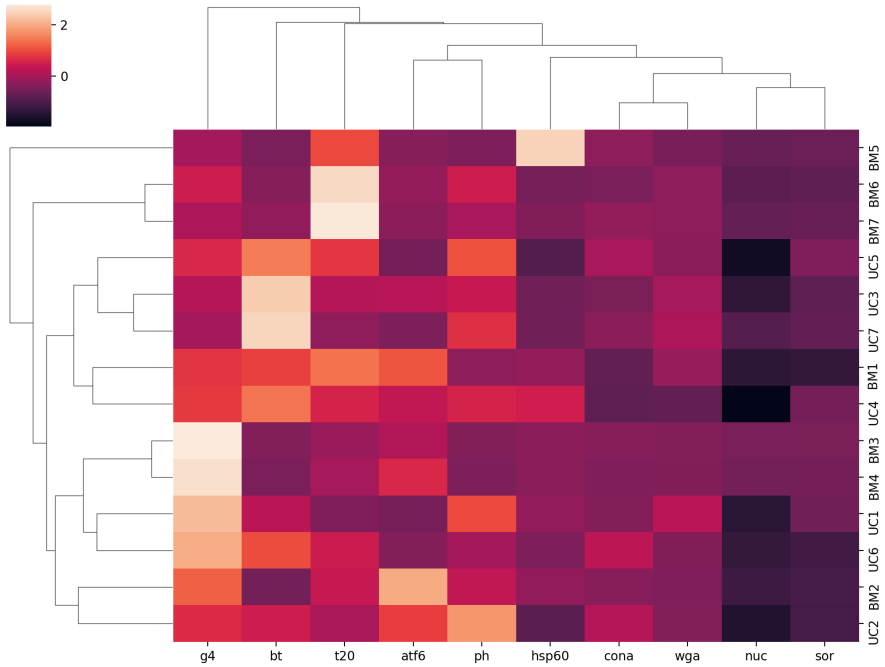


Figure 7: Heatmap comparing the normalized perimeter across 10 markers and 14 selected cells
X-axis: 10 selected markers
Y-axis: Cell type and assigned cell number
Color bar: Perimeter normalized using z-score (unitless)

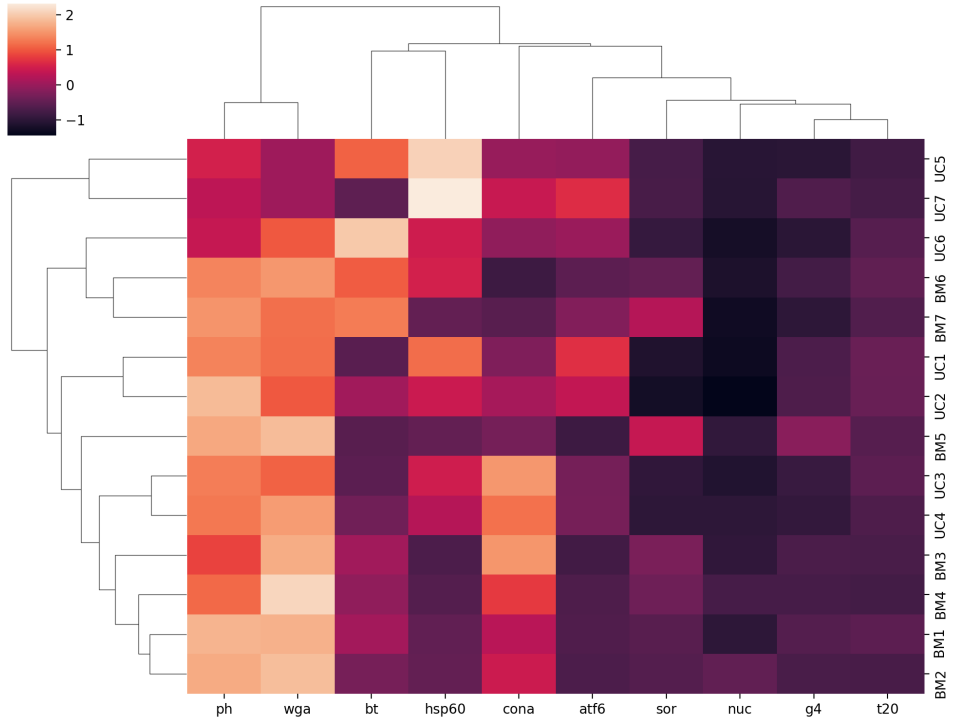


Figure 8: Heatmap comparing the normalized integrated intensity across 10 markers and 14 selected cells
X-axis: 10 selected markers
Y-axis: Cell type and assigned cell number
Color bar: Integrated intensity normalized using z-score (unitless)

Figures 3-4 are examples of raw data represented in the heatmap format. In order to focus on the variation between markers and better visualize the differences, z-score normalization was performed for all features using **Equation 1**:

$$z = \frac{x - \mu}{\sigma} \qquad (1)$$

In **Equation 1**:

z is the $z-$score
x is a random member
μ is the mean of population
σ is the standard deviation of population

Z-score was calculated for each cell, with each population consisting of 10 values for the 10 markers. The x-axis is labeled with the selected 10 protein markers to target different organelles, and the y-axis is labeled with the assigned cell number.

In **Figure 5**, the nucleolin marker generally has the lowest major axis lengths across all cells followed by the sortilin marker. Other markers have varying major axis lengths when compared across cells, with Beta-tubulin and HSP60 generally having high values.

In **Figure 6**, the nucleolin marker generally has the lowest minor axis lengths across all cells followed by the sortilin marker. Other markers have varying minor axis lengths when compared across cells, with Beta-tubulin and HSP60 generally having high values, which are the same characteristics as the major axis.

In **Figure 7**, the nucleolin and sortilin markers generally have the lowest perimeter across all cells. Other markers have varying perimeter when compared across cells, with GOLPH4, Beta-tubulin, TOM60, ATF6, and Phalloidin generally having high values.

In **Figure 8**, the nucleolin, GOLPH4, and t20 markers generally have the lowest integrated intensity across all cells followed by the sortilin marker. Other markers have varying perimeter when compared across cells, with WGA and Phalloidin having high values.

From all figures, the low values for the nucleolin marker were expected, because the nucleolin is a protein that is distributed inside the nuclei while other markers target organelles that are larger than the nuclei. However, data for the other markers vary more with a few trends that may be interesting to look at. For example, looking at normalized major/minor axis data, UC

cells generally have larger values for Beta-tubulin and HSP 60, which target the microtubules and mitochondria, respectively. This could be a key difference between UC and BM cells, although a larger data set may help find and confirm such a trend.

Chapter 4

Conclusion

Although processing and analysis of the images of more features are still in progress, there are a few points to note from the data we were able to obtain. Heatmap is a useful method to display large sets of data in a clear, visual manner, allowing us to compare values across markers and across cells. Generally, the nucleolin marker showed the smallest values among the 10 markers, while the values for the other markers varied depending on the feature. The dendrograms from the major/minor axis length show a split between UC and BM cells, indicating that there is a difference between the two groups. This may be a key finding, as our goal is to find trends and associations between markers or organelles as well as between the two cell types. The next step in the analysis would be to investigate more features and create one combined heat map that shows all of the features (averaged among UC cells and BM cells) on one axis against the ten markers on the other axis. After finalizing which features to measure and include in the visualization, statistical tests such as the t-distributed stochastic neighbor embedding (t-SNE) can be performed to reduce dimensionality for visualizing the data in two dimensions data.

References

[1] Lamprecht, M. R., Sabatini, D. M., & Carpenter, A. E. (2007). CellProfiler: free, versatile software for automated biological image analysis. *Biotechniques, 42*(1), 71-75. doi:10.2144/000112257

[2] Ozaki, Y., Yamada, H., Kikuchi, H., Hirotsu, A., Murakami, T., Matsumoto, T., ... Konno, H. (2019). Label-free classification of cells based on supervised machine learning of subcellular structures. *PLoS One, 14*(1), e0211347. doi:10.1371/journal.pone.0211347

[3] Wei, X., Yang, X., Han, Z. P., Qu, F. F., Shao, L., & Shi, Y. F. (2013). Mesenchymal stem cells: a new trend for cell therapy. *Acta pharmacologica Sinica*, *34*(6), 747–754. doi:10.1038/aps.2013.50

[4] Fan, X. L., Zhang, Y., Li, X., & Fu, Q. L. (2020). Mechanisms underlying the protective effects of mesenchymal stem cell-based therapy. *Cellular and molecular life sciences : CMLS*, *77*(14), 2771–2794. doi:10.1007/s00018-020-03454-6

[5] Gut, G., Herrmann, M. D., & Pelkmans, L. (2018). Multiplexed protein maps link subcellular organization to cellular states. *Science (New York, N.Y.)*, *361*(6401), eaar7042. doi:10.1126/science.aar7042

[6] Viana, M. P., Chen, J., Knijnenburg, T. A., Vasan, R., Yan, C., Arakaki, J. E., ... & Rafelski, S. M. (2020). Robust integrated intracellular organization of the human iPS cell: where, how much, and how variable?. *bioRxiv*. doi:10.1101/2020.12.08.415562.

[7] Carpenter, A. E., Jones, T. R., Lamprecht, M. R., Clarke, C., Kang, I. H., Friman, O., . . . Sabatini, D. M. (2006). CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biol, 7*(10), R100. doi:10.1186/gb-2006-7-10-r100

[8] Weiss, T., Semmler, L., Millesi, F., Mann, A., Haertinger, M., Salzmann, M., & Radtke, C. (2020). Automated image analysis of stained cytospins to quantify Schwann cell purity and proliferation. *PLoS One, 15*(5), e0233647. doi:10.1371/journal.pone.0233647

[9] Chong, Yolanda T., Koh, Judice L. Y., Friesen, H., Kaluarachchi Duffy, S., Cox, Michael J., Moses, A., . . . Andrews, Brenda J. (2015). Yeast Proteome Dynamics from Single Cell Imaging and Automated Analysis. *Cell, 161*(6), 1413-1424. doi:10.1016/j.cell.2015.04.051

[10] Dürr, O., & Sick, B. (2016). Single-Cell Phenotype Classification Using Deep Convolutional Neural Networks. *Journal of Biomolecular Screening*, *21*(9), 998–1003. doi:10.1177/1087057116631284

[11] Hamilton, N. A., Pantelic, R. S., Hanson, K., & Teasdale, R. D. (2007). Fast automated cell phenotype image classification. *BMC Bioinformatics, 8*(1), 110. doi:10.1186/1471-2105-8-110

[12] Kriston-Vizi, J., Lim, C. A., Condron, P., Chua, K., Wasser, M., & Flotow, H. (2010). An Automated High-Content Screening Image Analysis Pipeline for the Identification of Selective Autophagic Inducers in Human Cancer Cell Lines. *Journal of Biomolecular Screening*, *15*(7), 869–881. doi:10.1177/1087057110373393

[13] *Organelle*. Genome.gov. (n.d.). Retrieved December 17, 2021, from https://www.genome.gov/genetics-glossary/Organelle

[14] Lin, C. S., Xin, Z. C., Dai, J., & Lue, T. F. (2013). Commonly used mesenchymal stem cell markers and tracking labels: Limitations and challenges. *Histology and histopathology*, *28*(9), 1109–1116. https://doi.org/10.14670/HH-28.1109

[15] *DAPI (4',6-diamidino-2-phenylindole): Thermo Fisher Scientific - US*. DAPI (4',6-diamidino-2-phenylindole) | Thermo Fisher Scientific - US. (n.d.). Retrieved December 17, 2021, from https://www.thermofisher.com/jp/ja/home/life-science/cell-analysis/fluorophores/dapi-stain.html

[16] Goh, T. Y., Basah, S. N., Yazid, H., Aziz Safar, M. J., & Ahmad Saad, F. S. (2018). Performance analysis of image thresholding: Otsu technique. *Measurement*, *114*, 298–307. https://doi.org/10.1016/j.measurement.2017.09.052

[17]*Fluorescence intensity*. BMGLabtech.com. (n.d.). Retrieved December 17, 2021, from https://www.bmglabtech.com/fluorescence-intensity/

[18] *Colocalization analysis*. ImageJ Wiki. (n.d.). Retrieved December 17, 2021, from https://imagej.net/imaging/colocalization-analysis

[19] Kobak, D., & Berens, P. (2019). The art of using t-SNE for single-cell transcriptomics. *Nature Communications*, *10*(1), 5416. doi:10.1038/s41467-019-13056-x