



Autophagy Image Analysis

Junior Independent Work | Spring 2008

Adam Sanders and Hilary Coller

Adam Sanders
COS JIW
4/23/08

Autophagy: MATLAB Image Analysis

Abstract

As scientists and researchers continue to delve deeper and deeper into the intricate processes that modulate the cell cycle, increasingly advanced computational metrics will be needed to identify and quantify these cellular processes. I propose in this paper a computational means for the identification and quantification autophagy, a cellular process whose appearance is strongly related to entrance into and maintenance of cellular quiescence.

Introduction

The complex controls, switches, and balances that allow cells to make decisions as to whether or not division is in their best interest are some of the most fundamentally important pathways in all biology. These pathways are inescapably linked to the ultimate survival of an organism given times of resource shortages or difficult environmental circumstances and are especially significant for the development of new organisms. As a result of this importance, a vast amount of research and thought has been applied to furthering our understanding of the cell cycle. From this study, scientists have been able to develop a fairly accurate understanding of the *method* by which cells divide. Any reasonably prudent first-year biology student will be able to rattle off the different phases of the cell cycle and those actions that will likely occur at any given phase (a progression from G1 to S and so forth). For those out-of-touch with their days in freshman biology, a simplified cell cycle can be seen in Figure 1 below. This information is undoubtedly a vital aspect of our understanding and has informed our thought processes on an innumerable quantity of cellular diseases and other maladies. However, that is not to say that questions about the cell cycle do not remain. While a tremendous amount of effort and hard work has been applied to furthering our understanding of the processes by which cells elect to divide or not to divide, we are still plagued by diseases of the cell cycle. Cancer, a particularly relevant example, is among the leading causes of death in the United States as well as many other industrialized countries around the world (Boehm M 2003). Aside from strictly disease, research in this area holds tremendous promise for the study of stem cells and neurodegenerative disorders that cause paralysis. As a result of these important areas of research, interest in the field has undergone a noticeable shift from an emphasis on the *methods* by which cells divide, to their *reasons* for doing so. It is thought that through a more

complete understanding of this question, we will be able to better understand the peculiarities of cancer or even discover new ways of preventing tumor progression. The question of exactly why cells divide is a question fundamentally more difficult than the question of way in which they divide, and it's answer provides the ultimate goal in my research project. Specifically, I will examine methods by which researchers can measure a particular process that is associated with one phase in the cell cycle. The process is known as autophagy, and its occurrence is associated with the cell cycle state known as quiescence.

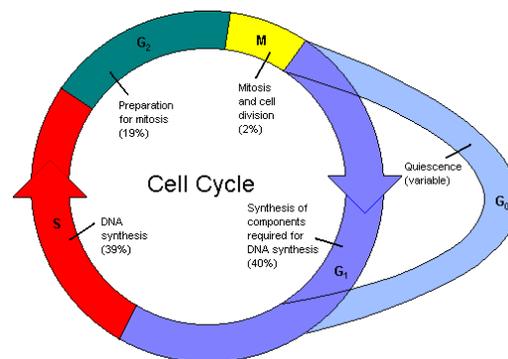


Figure 1: The Cell Cycle

Cellular Quiescence and Autophagy Background

While there exist many subcultures and topics within the broad class of research devoted to understanding the reasons for cell division, the Collier lab is primarily devoted to furthering our understanding of cells in a quiescent state. Quiescence for the purposes of biology is understood as the reversible arrest of cell growth and proliferation. It is the state in which you will find the majority of cells in your body, and the ability to enter quiescence is imperative for the survival of most organisms (Hippert, O'Toole and Thorburn 2006). Without this cellular state, organisms would consume and thus waste a tremendous amount of energy attempting to divide all of the time. Despite the importance and apparent ubiquity of this cellular state, serious questions remain. For example, some researchers are uncertain as to whether or not there exist subtypes of cellular quiescence that may be switched on with different cues and thus are used in response to different environmental circumstances (Juhász and Neufeld 2006). Others are more interested in examining the way in which autophagy, a process often found in quiescent cells, serves as an additional cell degradation pathway akin to apoptosis (Klionsky and Emr, Autophagy as a

Regulated Pathway of Cellular Degradation 2000). Regardless of their reasons for interest, however, all must rely on different assays to examine the processes active in a quiescent cell.

A particularly interesting cell process that often occurs during periods of cellular quiescence, and the topic of this analysis, is autophagy. Autophagy is a catabolic process by which nutrient-deprived cells are able to re-claim stored energy through the degradation of its own cellular components. The cell is literally able to eat itself in order to maintain homeostasis. Through this process, starved cells are able to reallocate nutrients from unnecessary cellular components to those that are found to be more essential. Though scientists have been aware of this process for over 50 years, it has only recently been recognized for its importance in cancer development and progression as well as in the determination of a cell's response to anti-cancer treatment (Hippert, O'Toole and Thorburn 2006). The process works as follows and is diagrammed in Figure 2 below. When a cell makes the determination that is resource-poor and requires additional energy, it encloses some cellular components within a membrane. These components could include cytoplasm, the gelatinous semi-transparent fluid that fills all cells, or even an entire organelle. Once this membrane has completely closed around the cellular machinery to be discarded, it becomes known as an autophagosome. This stage can be seen in the second phase of Figure 2 below. The third and final phase of autophagy occurs when a lysosome, a cellular membrane filled with highly acidic digestive enzymes used to digest useless organelles, food particles or viruses, fuses with the autophagosome in order to decompose the worn-out cellular material. This final lysosome-autophagosome fusion is known rather succinctly as the autophagolysosome. It is in this organelle that the cell components are finally broken down and returned to raw materials that the cell can then put to better use in those functions it finds most important.

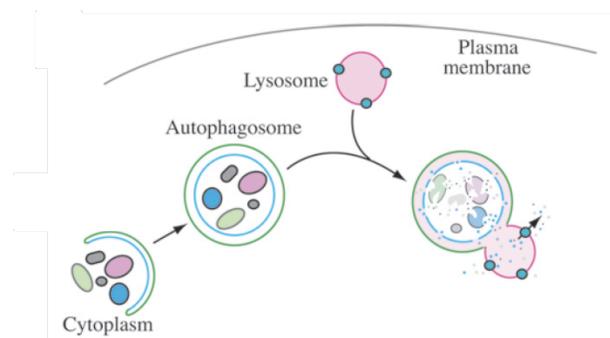


Figure 2: The Process of Autophagy

There have been a variety of methods tried in the past and used at current to visualize and quantify this process. One rather obvious and effective method has been to monitor the pH of the cell components. The goal of

this process was to locate the autophagolysosomes which would appear as large cellular spaces filled with highly acidic materials. However, the use of such pH-based assays can be problematic. There are many cellular components that are highly acidic in baseline cells (lysosomes, for example). The result of this is that researchers would be required to find some effective means of differentiating autophagolysosomes from any other cell component that happened to be acidic in any given cell. Figure 3 below demonstrates the use of just such an assay: Acrodine Orange. This dye interacts both with the autophagolysosomes as well as DNA. Fortunately, the dye fluoresces a different color when interacting with nucleic acids like DNA than when it interacts with autophagolysosomes. Below we can see the nucleus of a yeast cell fluorescing green while baseline levels of low-pH organelles fluoresce a more orange hue. At right one can see the results in the same cell type when autophagy is induced through the introduction of rapamycin.

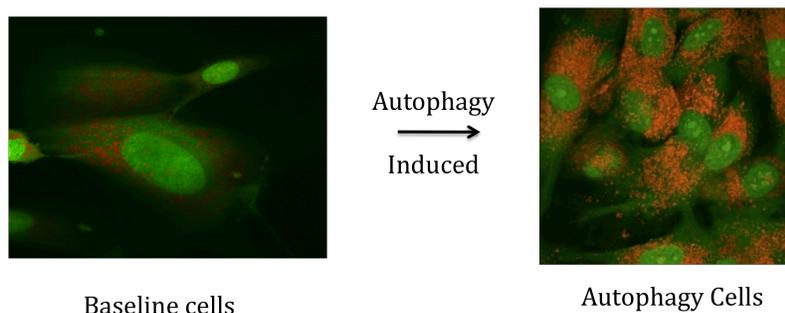


Figure 3: A pH assay for the determination of autophagy

More recently, however, a different method has been introduced that allows researchers to visualize the autophagolysosomes specifically. This method is known as GFP-LC3 fusion, and relies on the fusion of GFP, a fluorescent protein originally found in jellyfish, to LC3, a protein known to interact with autophagolysosomes. When cells devoid of any appreciable amount of autophagy are exposed to this protein and then subjected to blue light (the wavelength at which the GFP protein fluoresces), one can see the GFP protein well diffused throughout the cell, as there exist no autophagolysosomes with which this protein can interact. This phenomenon is shown in Figure 4 below at left. However, when the same cells are exposed to the starvation conditions created by researchers, the diffuse GFP-LC3 localizes to distinct points within the cell, presumably the autophagolysosomes that have been created in response to this cellular starvation. These points are known in the literature surrounding autophagy as puncta. One can clearly see how the autophagic cells at right might be distinguished from the baseline cells at left. In order to determine the amount of autophagy occurring in a given cell, one need only count the number of distinct

puncta that exist for that cell. These points are visually obvious to the human eye and getting a reasonably proximate count for any individual cell is a trivial step in research.

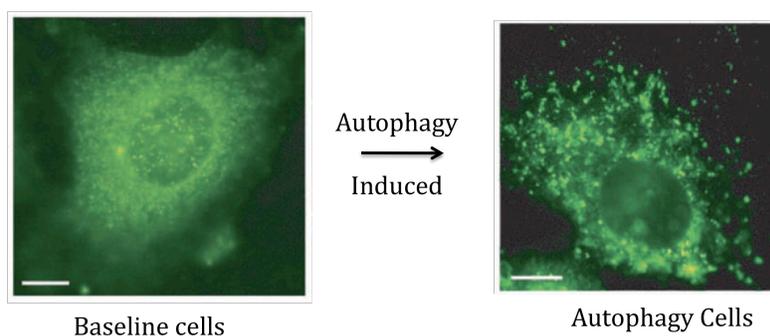


Figure 4: GFP-LC3 Assay showing puncta in autophagic cells

However, research in all areas of biology is rapidly moving in the direction of high-throughput data. Partially as a result of advances in computation, it is relatively easy and very desirable to take tens of hundreds of images for any given condition when examining autophagy. As one can imagine, this method would then become prohibitively time-consuming and highly prone to individual error. In one review of the processes by which autophagy can be quantified, it was said of GFP-LC3: “There are, however, practical issues with counting puncta manually and reliably, especially if there are large numbers per cell (although *this may be more accurate than relying on a software program*, in which case it is important to ensure that only appropriate dots are being counted)” (Klionsky, Abeliovich, et al. 2008, emphasis added). This quote neatly outlines the obvious necessity felt in the autophagy community for the development of a reliable computational method to quantify autophagy, and to do so on large-scale data sets. Shockingly enough, no current or widely available program exists for such quantification.

My Project

The preceding computational call-to-arms leaves many questions yet to be answered about my particular project. The first and perhaps most pressing concern regarded the exact type of image that was to be used. As mentioned previously, a great many methods exist for visually quantifying autophagy. It would be important to both learn as much as possible about the different types of images used in each analysis, as well as to examine each image type in order to determine what might best be suited for computational analysis. This decision, representing the first stage of the task at hand, would do a great deal in terms of guiding the rest of the project. Once the decision as to the particular type of image to be used was made, it would be important to closely analyze this image and attempt to find those factors within the image correlating to autophagy that would be least difficult to extract. The

difficulty in feature extraction would also have to be weighted against the information that could be extracted from those features. For example, previous studies had assumed that merely having a count of the number of puncta within any autophagic cell was sufficient to determine the exact amount of autophagy occurring in that cell. However, it might also be worthwhile to consider the average cell size, or the average size of the individual puncta, both tasks that would be difficult to do by hand. The third and penultimate stage of the project would be to develop an effective metric for the quantification of autophagy given both the image type and relative importance of the different types of information that could be gathered from these images as determined by stage two. This metric would have to be fast, reliable, and flexible enough to work on a variety of different data sets. In an ideal world, this metric could be applied to data sets garnered not only in experiments carried out by the Collier lab, but in all those labs concerned with the study of autophagy. The development of such a resource, though likely challenging, would represent a tremendous achievement in the field of autophagy.

The Original Data

The type of image selected for use in this project went through several iterations before a final workable set of images was reached. The original data set consisted of 24 TIF images, each about 300 kilobytes. The resolution on these images was 700x500 pixels. The images were taken in sets of three, one image for each type of assay that could be completed on the cell. The three different stains used were: DAPI (4'-6-Diamidino-2-phenylindole) – used to stain the nucleus of cells by attaching to DNA, GFP-LC3 – used to locate autophagosomes as mentioned previously, and phalloidin – a stain used to visualize the cell body. In addition, this set of images was also broken down into four different levels of quiescence (for example: deeply quiescence or entering quiescence). As each of these assays returned light to the microscope in only one wavelength (though at different intensities), each individual file could be considered a grayscale image. In fact, the microscope that took the pictures has the ability to recognize such a fact and automatically made the conversion. The results for the purposes of image processing were eight sets of images, where each set consisted of three matrices, each 700x500 cells with intensity values ranging from zero to 255. Figure 8 below is an example of one set of images from the original data set.

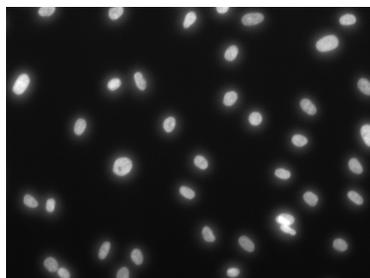


Figure 5: DAPI Staining

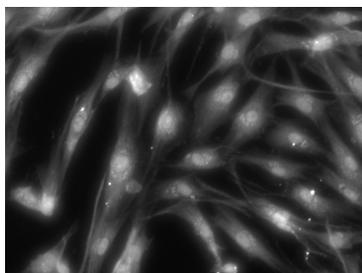


Figure 6: GFP-LC3 Probe

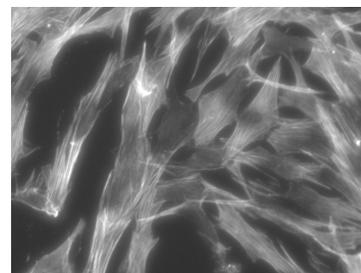


Figure 7: Phalloidin Stain

Figure 8: A set of images from the original data set

Image Histograms

One of the first ideas for these images was to simply examine the image histograms in the hope that these would reveal a passable threshold value for puncta. In ideal circumstances, each image would have two identifiable peaks. The first and lower intensity of these two peaks would be made entirely of the background of the image. Thus the first peak would contain both the cell body and background noise, while the second peak would contain the high-intensity peaks of interest – namely puncta. Figure 9 below is an example of just such a histogram. The picture at left corresponds to the histogram at right. One can visually identify the slightly bimodal arrangement of the histogram with peaks both at 25 and 125 intensity values. This result was initially exciting and provided hope for a very simple solution to the quantification problem. However, it was soon realized that these highly intense values were almost entirely in the nucleus. These high intensity points are problematic because autophagosomes are not found in the nucleus. Thus at some point during the creation of these images, the GFP-LC3 fusion must have found some affinity for being inside or very nearby the nucleus of cells. To make matters more clear, it was only necessary to subtract the corresponding DAPI image from the GFP-LC3 image. This effectively removed the nucleus from each cell and generated the image and corresponding histogram seen in Figure 10. From this new image and histogram, it becomes clear that the obvious bimodality in the original GFP-LC3 image has been lost.

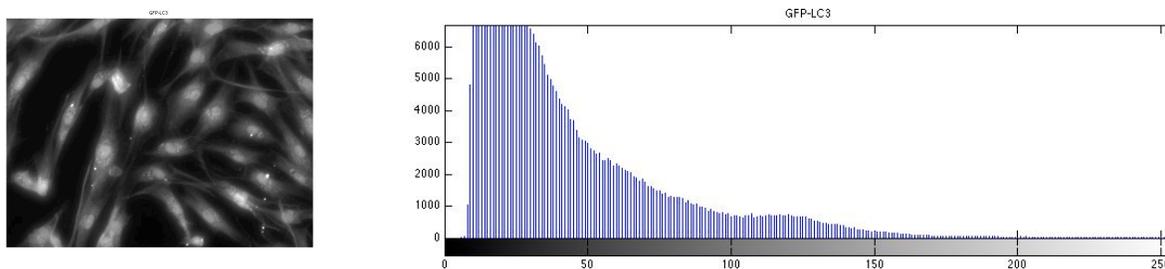


Figure 9: GFP-LC3 Fusion and intensity histogram

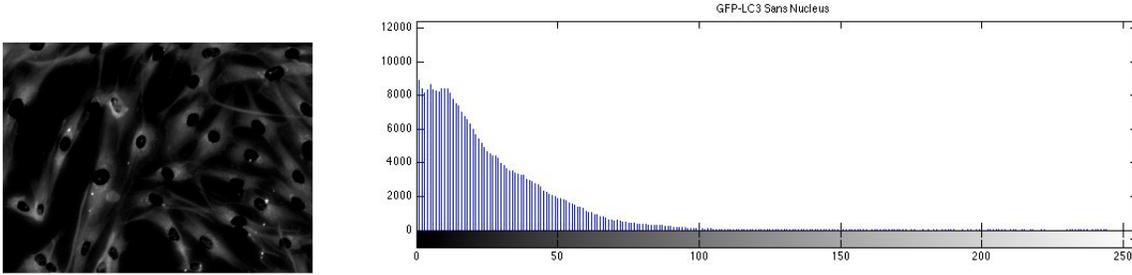


Figure 10: GFP-LC3 Fusion and Intensity Histogram with DAPI Subtraction

Further complicating analysis on these images was the unfortunate fact that the area immediately surrounding the nucleus, and especially the nucleus itself, were found to have intensity values in exactly the same range as did the autophagosomes. Even with the DAPI nucleus subtraction, the small halos left surrounding the nucleus were very similar in intensity to the autophagic puncta.

Image Blurring

While it seemed that simply finding some obvious threshold value above which puncta existed was likely out of the question, another direction in which to take this project was to blur the images multiple times and reexamine the histogram after each blurring. The thought process behind this analysis was that if there were points of high intensity surrounded by areas of lower intensity, we would be able to differentiate these two areas via blurring analysis. For areas of low intensity, the blurring was meant to make the cell background indistinguishable from the actual slide background. Meanwhile, the high intensity puncta resulting from the autophagolysosomes should be resilient to such blurring. However, when this analysis was actually attempted, the result was quite to the contrary. The blurring rapidly removed the high intensity puncta while exerting relatively little change on the lower-intensity cell bodies. The difficulty with this type of analysis could be traced back to the size of each puncta. The puncta in this set of images were largely found to have a diameter of less than 10 cells. On average, this value was about 7 cells across, while the center containing the highest intensity points would be found to be about one or two cells in diameter. One would have to find a filter that was smaller than this 2-pixel window in order to preserve the puncta. However, there is little sense in using a Gaussian filter for blurring that is 3 pixels across, as it would have almost no effect on any part of the image.

In addition to this attempted use of blurring, an attempt was also made to use blurring to remove the GFP halo found after DAPI subtraction was applied to the GFP-LC3 images. In order to carry out this transformation, one first subtracted the DAPI stain of an image from the corresponding GFP-LC3 stain. This results in the image

shown above in Figure 10. From this point, I applied a Gaussian blur to the original DAPI stain and then subtracted that image again from the GFP-LC3 image. The result was a dramatic and fairly uniform reduction in the GFP halo around the nucleus. Unfortunately, this also led to a severe reduction in the puncta surrounding the nucleus.

The New Data Set

While the results and analysis generated through work with the first data set were informative, both obviously left much to be desired. In regards to the actual data, the resolution of the photographs was relatively low. Additionally, while it was fairly easy to remove the nucleus of these images via thanks to the DAPI stained images, there were still many problems with the GFP-LC3 assay itself. That is not to say that working with such data would be impossible, merely very difficult. Happily, there was a second data set produced by the Collier lab during the same period as the first. These new images were taken on the much more powerful confocal microscope and were accordingly much higher resolution than those in the first dataset. Each TIF file had a resolution of 1024x1024 (~3MB). The images came in two varieties. The first of these were GFP-LC3 photographs taken under a variety of different conditions, while the second were Acridine Orange images. Figure 11 below shows both of these image types (GFP-LC3 at left, Acridine Orange at right).

Perhaps the most meaningful difference between the first and second datasets with regards to GFP-LC3 analysis was the absence of multiple images corresponding to the same cell. Rather than having three separate matrices that could be added, subtracted, or otherwise combined to extract useful information, this new data set only contained information in a single matrix of intensity values. However, further differences existed. In contrast with the older dataset, this new dataset was devoid of the problematic GFP-LC3 nucleus affinity. As can be seen below at left in Figure 11, the great majority of the high intensity peaks within the images are puncta. Additionally, the puncta in these images seem to be significantly brighter than any other points that would allow for relatively simple methods enumerating the puncta. Interestingly, the acridine orange stain would still have allowed for the existence of multiple matrices. As each of these images was in color, each was actually represented as a set of three matrices corresponding to the red, green, and blue intensity values of the image. Given the two different wavelengths of light produced by the acridine stain, it would be relatively simple to locate the nucleus of each of these cells by isolating the large gobs of green from the individual points of orange. While I did not deal directly with these images in the course of this project, they do represent an interesting direction for further work in this area.

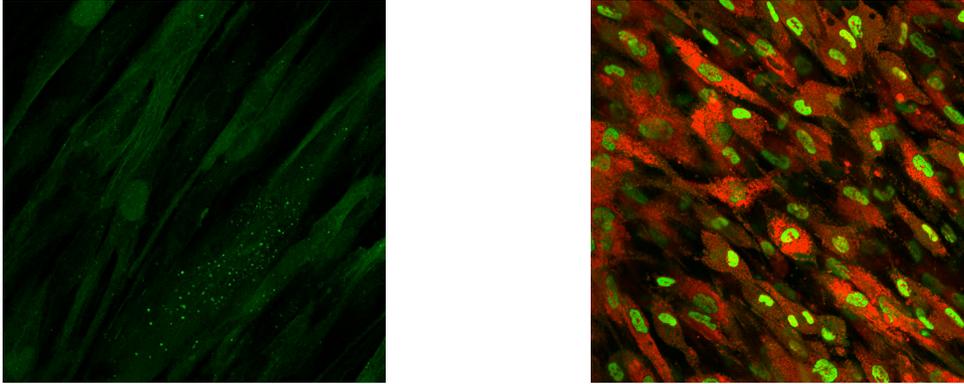


Figure 11: Example images from the new data set

Image Thresholds

My first and seemingly most obvious attempt to quantify autophagic cells was through the application of global thresholding. This technique, owing both to its ease of implementation and intuitive properties, is fairly prevalent in image processing applications and seemed a very reasonable direction for my initial foray into the new dataset (Gonzalez, Woods and Eddins 2004). In contrast to local thresholding, a point to which I will return near the end of this paper, global thresholding considers the histogram of an entire image in order to make a determination about which values are useful and which are not. The result of any thresholding experiment is a logical matrix of the same dimensions as the original image, where each cell represents either an up or down vote as to the relevance of the same cell in the original image. These values can be described by the following formula, where T represents some threshold value, g represents the original image matrix and f represents the transformed matrix of logical points:

$$\text{If } g(x, y) > T \rightarrow f(x, y) = 1$$

$$\text{If } g(x, y) \leq T \rightarrow f(x, y) = 0$$

While this formula is simple, an intelligent determination of the actual threshold value, T , is a challenging problem. Under ideal circumstances, one would be able to closely examine the image histogram and visibly identify two separate peaks. Taking the minimum point between these two peaks to be the threshold value, one could then separate these peaks via the formula listed above. Unfortunately the image histograms given in this new dataset made such an analysis impossible for a number of different reasons. One of the first and most obvious challenges in

the way of just such an analysis is the relatively small area represented by puncta. Manually selecting for an ideal threshold value as a test, I was able to locate only 86 total cells that could be identified as puncta in a possible 1048576 cells. In terms of percentages the area, on average, consumed by puncta is only $8.2016 \times 10^{-5}\%$ of the total possible area in a given image. This fact makes the visual identification of these puncta from a scan completely impossible. Further complicating the analysis is the fact that every individual puncta is actually a range of values resembling a 2D Gaussian distributed point. Though it would be far easier were these points flat little disks, the majority of the puncta in a given image actually come to a point making thresholding a difficult problem. Depending on the value eventually chosen, any given autophagic point could be a single pixel or perhaps something several times that size. After a great deal of trial and error, I developed a metric that seemed to work fairly well for discerning autophagic puncta from irrelevant background noise. Given the standard deviation and average intensity value of an entire image I , the ideal threshold for the images in this dataset appears to be about ten standard deviations above the average. This measure is obviously flawed, but it did allow me to rapidly begin probing deeper in to the best ways to quantify these puncta.

However, I quickly realized that simply counting the number of points above such a threshold value is completely irrelevant. The reason for this is the variability in the number and size of cells in any given image. This value can range anywhere from a single cell to more than thirty. Taking this example farther, were one to simply take the number of puncta in the case where there was a single cell and compare that number to the number of puncta in an image containing more than thirty cells, the results could very easily be off by thirty-fold. Given the need for precision mentioned in the review of autophagy near the beginning of this paper, an error of this magnitude would be intolerable. To make matters more complicated, there are additional cells in most images that are simply out of focus. Because cells are three-dimensional and most microscopes (including the one used for this analysis) can only capture images representing two dimensions, there are always out-of-focus cells in the lens. In fact, it is partially due to this cause that there is some variability in the intensity values of many puncta. The less intense puncta are likely the result of an out-of-focus cell whose puncta are similarly out of focus. These are both fascinating problems that lend themselves to some very exciting solutions in terms of generating new and three-dimensional images of the cells, however in consideration of the fact that I was only privy to the two-dimensional images on hand, they proved to be something of a problem.

The obvious result of this difficulty was that it was necessary to find some way of quantifying the number of cells in the visual pane. Under ideal circumstances, one would have access to the DAPI stained images of the same cells as was the case using the old image set. However, no DAPI stains were taken of these newer images so a new metric had to be evolved. Keeping in mind the thresholding done for the puncta, I felt that it might be relevant to attempt just such an analysis for the cell bodies. Thus, in addition to finding the areas of very high intensity and calling them puncta, I would also find those areas of slightly above-average intensity and define them as the cell body. Again through a process of trial and error I was able to derive a metric that was fairly efficient at finding those areas (not background) that contained cells. Unfortunately, however, this metric seemed less effective than the metric used to find autophagic points. This is a result of the relative similarity between the intensity values of many cells in the image and the background of that image.

Given both the binary matrix representing the position of purported cells as well as the binary matrix reporting the position of posited puncta, I developed a simple scoring mechanism to find an autophagy score for any given image. The metric simply finds the ratio of those two quantities as shown below:

$$\text{Autophagy Score} = \text{Total Area of Puncta} / \text{Total Area of Cell}$$

Figure 12 below shows the result of applying this formula to a single image. The left pane shows the original image representing what appear to be three distinct cells, at least two of which contain some autophagic points. The second pane shows the binary transformation of that image into the cell matrix. The areas of white are all those points greater than one standard deviation above the average. Finally, the third pane shows all those points the program decided were autophagic. As a reminder, those are points that were ten standard deviations above the average or greater. Taking the sum of all the points in the third matrix and dividing that number by the sum of all the points in the first matrix (and multiplying by some constant in order to improve the ease with which these images can be seen) we would arrive at a total autophagy score for this image of five.



Figure 12: Image thresholding example

Through the application of the formula mentioned above, I developed a program that could rapidly assign a score to a great number of images. This program requires only a directory and a matrix of general expressions in that directory that are to be explored, and it will rapidly generate an average score for each condition. An example of such a matrix would be [`*KRB*`, `*3MA*`, `*SS*`], from which the program would find and analyze all images containing KRB, 3MA and SS, and then compare the results. It is through this format that the meaningful differences between different image types can be determined. While it is interesting and necessary to ensure that the autophagy score for any given image is representative of what a human would generate, it is likely more informative and actually less challenging to determine whether or not this program can generate the results expected for a set of conditions. After all, it is likely in this stead that this program will find the majority of its use, especially considering the movement in biology towards large-scale datasets.

As mentioned previously in the Data Sets section, this new dataset contained a total of five different possible cell conditions. These conditions were: rapamycin (RAPA), KRB, Serum Starvation (SS), No Starvation (NS), and 3MA. As an additional method of breaking these images down, the images were also classified by cell-state. Cells could be subjected to any of the above conditions while growing (G), quiescent (Q), or deeply quiescent (DQ). The effects of each of these conditions can be found in Table 1 below. One important note about cellular state that might be worth mentioning in regards to G, Q and DQ is that we would not necessarily expect the cells in the deeply quiescent state to exhibit higher levels of autophagy when exposed to an environmental condition like KRB (which would induce autophagy). As these cells are already in a dormant state, it would be just as likely that they have already squirreled away the necessary resources before being subjected to the autophagy-inducing condition. It is for this reason that I have elected to follow primarily the effects of the environmental conditions as opposed to examining the relative effects of cell-state on the average autophagy score.

Table 1: Table of cellular conditions induced in the new dataset

Label	Condition Induced	Autophagy
KRB	Starvation	Induced
Serum Starved	Starvation	Induced
Rapamycin (RAPA)	Starvation	Induced
Not Starved (NS)	None	Baseline
3MA	Autophagy Inhibition	Reduced

The above table allows us to make some predictions about the outcome of the aggregate autophagy analysis. We obviously expect the conditions in which autophagy was induced to provide much higher average autophagy scores, while those in which starvation was not induced should be noticeably lower. However, further discussion with biologists and an examination of the literature allows us to make at least some distinctions even within the group of conditions that should induce autophagy. KRB is a particularly severe starvation medium. In contrast to serum starvation, where cells are only deprived of several important ingredients, cells grown or transferred into KRB medium are placed in a harsh environment consisting almost entirely of salts. As a result of this fact, we would expect KRB to be noticeably higher than either of the other two autophagy-inducing conditions. In terms of distinguishing RAPA from simple serum starved cells, the line becomes less finite. RAPA induces autophagy by inhibiting the signals sent by another protein, TOR, to divide. However, the activity of this inhibition is variable and depends largely on the condition of the cells in a given plate. Thus it is difficult to provide with any certainty the exact degree to which autophagy will be induced given a specific quantity of RAPA. Finally, we consider the anticipated results for both 3MA as well as the baseline non-starvation condition. As 3MA inhibits a protein in the pathway necessary for the activation of autophagy, we should anticipate an average of less autophagy in the 3MA condition than in the baseline condition. It is important to note at this point that the inhibition of autophagy does not mean a complete discontinuation of the process. Just as cells undergo some autophagy at a baseline level under normal circumstances, those cells exposed to 3MA will still undergo some decreased but still visible amount of the process.

The results of my analysis proved highly consistent with the predictions just mentioned. Shown in Figure 13 in order from the highest to the lowest autophagy scores are: KRB, SS, RAPA, NS, and finally 3MA. Perhaps the most striking feature of this figure is the obvious difference between the first three bars, those representing autophagy-inducing conditions, and the final two, those representing autophagy-neutral or autophagy-inhibiting conditions. In fact, every single autophagy-inducing condition received an average autophagy score of more than all

the autophagy-reducing conditions combined. Taking a closer look at Figure 13 we can also verify our predictions for the conditions with the highest autophagy scores within those conditions. As predicted, the cells in the KRB condition had a noticeably higher autophagy score than did any other condition. In addition, the 3MA condition was found to have a lower autophagy score than the baseline case.

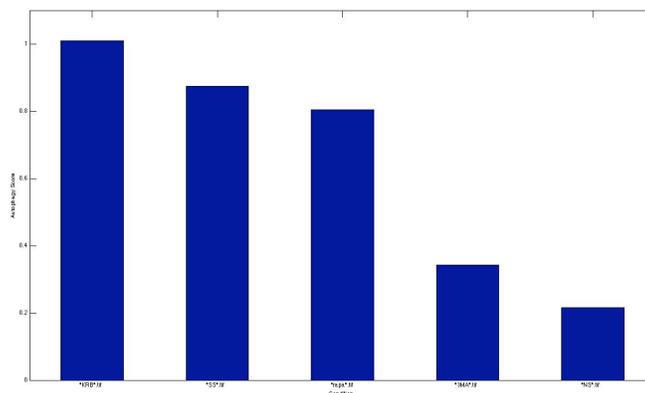


Figure 13: Autophagy score of each condition

This initial result was very encouraging and does much to prove the development of automatic autophagy-scoring programs is very possible. However, it becomes obvious when closely examining the results from individual images that this metric still leaves much to be desired.

Developing the Background

One of the main sources of difficulty in the previous analysis was the variability found in the background images. The main problem with generating this cell outline matrix was that many points in the original image were indistinguishable from background noise. The result of this difficulty would be a binary matrix that looked granulated and contained many gaps. This granulation can be seen in Figure 14 where we see both an original image (at left) as well as the binary transformation of that image (at right). In ideal circumstances the area delineated by this binary matrix would be completely solid and perhaps even recognizable as a cell. However, this is far from the case in images shown in Figure 14 below. Taking a closer look, it becomes apparent that this image suffers from some of the problems mentioned before. There are several out of focus cells and small amounts of background interference.



Figure 14: Original Thresholding Binary Matrix Creation

There are a number of ways to deal with such a problem, however the general idea behind all of these methods is finding points that deserve to be thinned or thickened that then performing the appropriate action. In order to find the best points for my analysis I abandoned completely the thresholding metric used in previous analysis in favor of a morphological approach as well as a number of image filters. The first stage of the conversion from the regular image to a binary background image was to morphologically open the image using a special filter. I chose to include only those portions of the image that contained nine pixels in a square formation. The aim of this process was to reduce the amount of background noise that was allowed in the image. Often the background interferes with the cell background by intruding single pixels or doubles and triples of pixels into the cell-background range. The requirement of fulfilling this square filter makes it impossible to include such pixels in the sum of the background image. While this transformation does allow for a dramatic reduction in the amount of useless background noise, it does to some extent thin out those pixels that do get through. Thus the next transformation is a dilation that is meant to expand on the selections that successfully pass the first test. This dilation finds pixels fulfilling the same 3x3 square pattern and expands the black areas around those images generating a fuller and more uniform image. The final transformation applied to these images is the closing of any of the small gaps remaining. This is done by running a 7x7 pixel mask over the image and looking for connected regions within which zero-value pixels exist. Whenever this mask finds such a region, all the pixels inside can be filled in with true values. The results of this background image processing can be seen in Figure 15 below. The image at left below is the original image, while the image at left is the final binary transformation of that image given the new set of analysis just outlined. It is very easy to identify the smoothness and effectiveness that this new metric allows in these images. A future direction cited for this analysis would almost certainly be attempting to locate and enumerate the cells on the plate rather than simply summing the binary matrix seen below. It is as yet unclear whether or not the use of cell size as opposed to cell number has a positive or negative effect on the outcome

of the autophagy analysis. In order to better answer this very question, one would have to develop a means to count the number of cells on each image, while avoiding both out-of-focus cells as well as background noise.



Figure 15: New Dilated Matrix Creation

Image Correlation

In a further move away from correlations and at the behest of Professor Szymon Rusinkiewics of Princeton University's Computer Vision course, it was determined that improvements could be made to the thresholding analysis mentioned earlier. Rather than attempting to locate high intensity points from a graph, it would likely be more worthwhile to define some standard puncta, and then find a matching score for all points in the matrix. This method would work by taking the filter and passing it over the matrix a cell at a time and calculating the Euclidean distance between every cell in the filter and the corresponding cell in the image itself. This immediately begged the question of just how to define the ideal puncta. The initial route tried for this analysis was an attempt to artificially create a standard puncta. Using a Gaussian function and a small filter matrix of roughly 12 pixels, one could develop a simple puncta. The advantages to such a process would be that this process could then easily be automated and would allow for a greater flexibility when incorporated into other data sets. An alternative to this process was to examine the current data set in question for a perfectly representative point. While the initial idea of self-generating points seemed an interesting challenge, it performed relatively poorly in competition with natively found points. Figure 16 below is a demonstration of just such a point. This image is a 12x12 clipping from an actual autophagic puncta. The image at right is the actual matrix representation of this 12x12 clipping.

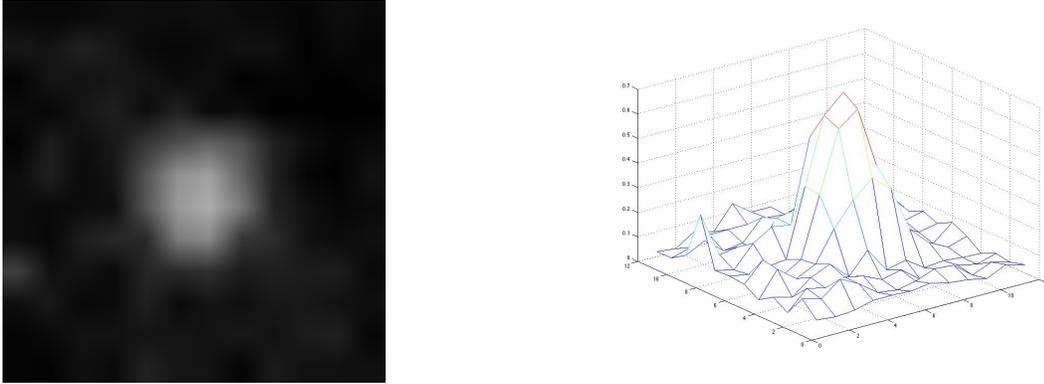


Figure 16: Actual Puncta and the matrix representation of that point

Perhaps most important to note in this image is the area of low intensity surrounding the puncta. It is really this area that generates the difference between thresholding and image correlation. Without this area (i.e. with a single pixel or small number of high-intensity pixels), this process would return exactly the same result as would any thresholding analysis. Given the sample puncta found in Figure 16, and the original image found below at left in Figure 17, the resulting correlation can be seen below at right. Points of low intensity are shown in blue while those of higher intensity are graphed in red. In addition, it is relatively simple to determine that there exist only a relatively small number of points that reach any appreciable height. This picture is drastically different than the original image where, rather than individual peaks, there are large areas of higher intensity. From these areas protrude a number of peaks, but the distance between the highest height and the average intensity within the actual cell body is significantly greater in the correlation condition than when using image thresholding.

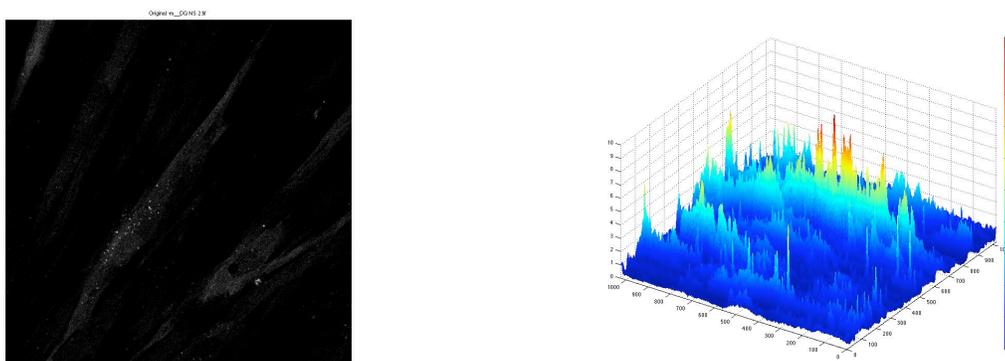


Figure 17: Puncta identified via correlation with simulated puncta

There are a great variety of advantages to using such a process. The first of these is in relation to the size of the points on the binary image. In thresholding analysis, there were often large blobs of points that fell above the

threshold value. In contrast to this type of analysis, correlation typically resulted in only one or two points that were above a certain correlation. This result is much more in keeping with the way autophagy has historically been classified. Previously, the only metric used for quantifying autophagy was the specific count of autophagic puncta in an image. With the correlation method, the size of puncta plays less of a role in the final autophagy score. In addition to autophagy size, this method also allows for the selection of puncta that are well below the highest intensity values in any given image.

However, the real test of this process would be to determine exactly how the results generated through image correlation compared with those originally generated by image thresholding. In answer to the question of how exactly the performance of this function differed from that of its predecessors one could only answer: not by much. Figure 18 below shows a comparison of the two methods in quantifying the amount of autophagy discovered in the G, Q, and DQ conditions. In viewing these two graphs it becomes fairly obvious that what changes have occurred are subtle. One can see a slight decrease in the G condition's autophagy mean with a corresponding increase in the average amount of autophagy occurring in the DQ condition. However, these changes are certainly nothing extraordinary. The bar graph corresponding to the KRB, SS, RAPA, NS, and 3MA conditions experienced a similarly minor degree of change.

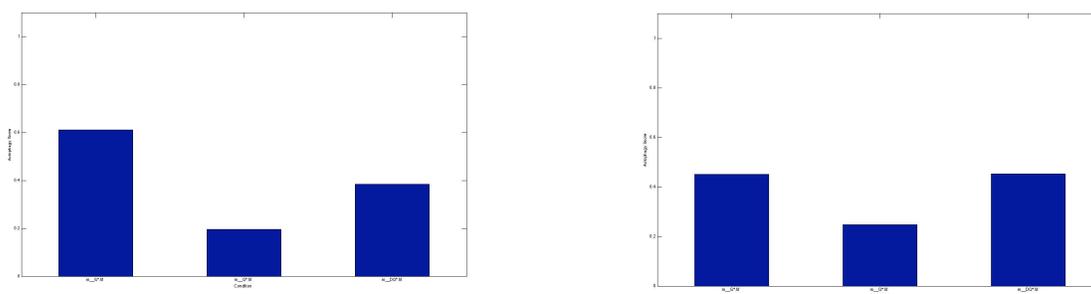


Figure 18: Comparison of correlation versus threshold for quiescence level

Perhaps more relevant than an examination of the average autophagy score across multiple images, would be an examination of the individual autophagy score given to a number of different images. Figure 19 represents just such a comparison. The figure shown at left is the result of the old image analysis program while the image at right demonstrates the newer correlation-based methods. The first thing one notices is the general reduction in the autophagy scores across many of the images. This is likely a result of the decreased size of the individual puncta, as well as the increased size and specificity of the background cell image. In addition, the new correlation-based

method does a great deal to fix one of the most obviously wrong classifications made by the thresholding method. The extremely large bar shown at bottom left seems an unrealistically high score for any individual image (though it does occur in a KRB condition). However, with the correlation based methods the unrealistic autophagy score has been reduced to a more reasonable range.

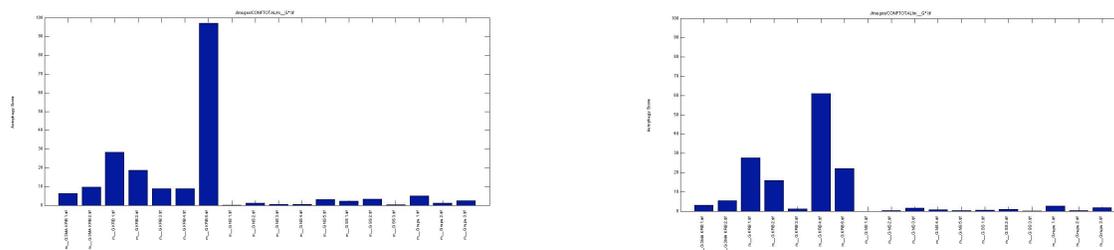


Figure 19: Comparison of correlation versus threshold for KRB

Conclusion

There are a number of things one can draw from this analysis, however the aspect that is most striking is that the creation of a computational metric to measure the amount of autophagy occurring in a given cell is very possible. This possibility is only improved by the increasing average size of datasets. While it is likely or even probable that any classifier will fail to produce adequate results given only a few unfortunate images, it is far more likely that the same classifier will, on average, be able to produce adequate results in merely a fraction of the time and effort it would take to garner those same results by hand. While there certainly exist barriers to the creation of any machine that works as well as do human eyes and brains for the image processing, with each passing day it seems that these barriers become less and less of a problem.

Future Directions

Future directions in this project are exciting as they are plentiful. One of the very first ideas that I would like to implement is a simple user interface. My program already allows one to specify regular expressions for the conditions as well as to examine the progress made at every stage in the process should you so choose. It would be relatively simple to aggregate these features into a simple menu through which one could control the process. As one of the future goals of this project was to make my program accessible, it would seem that this feature should be implemented first. A second and final direction would be to take the image analysis that has been accomplished in two dimensions and add a third dimension. There is little reason why one couldn't simply take a series of images of

any given plate of cells and layer them on top of one another. From this screen, one could develop a three-dimensional view of the cells. Rather than simply finding the autophagic density in two dimensions, one could actually find it in three. In speaking with the biologists in the Collier lab, this would actually be a far more appropriate means to quantify autophagy.

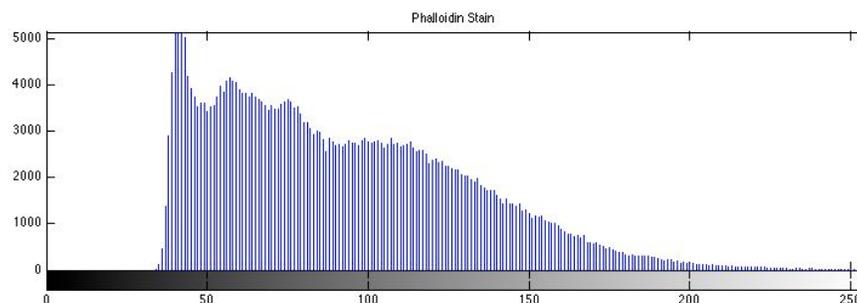
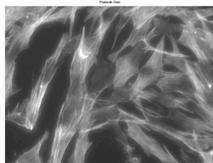
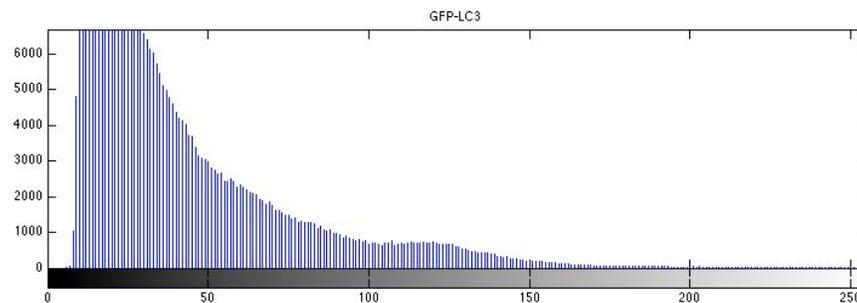
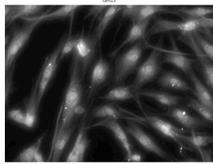
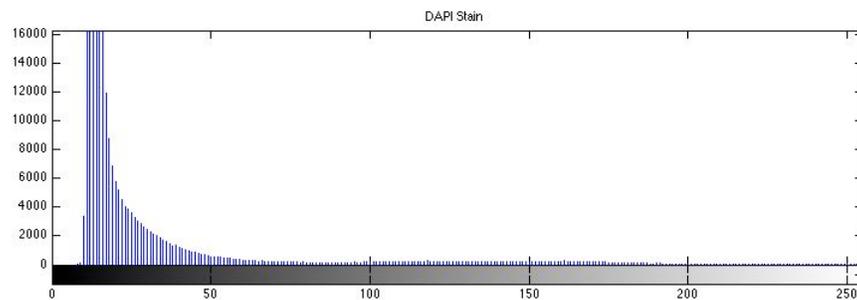
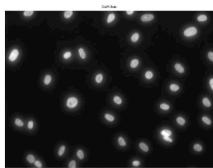
Works Cited

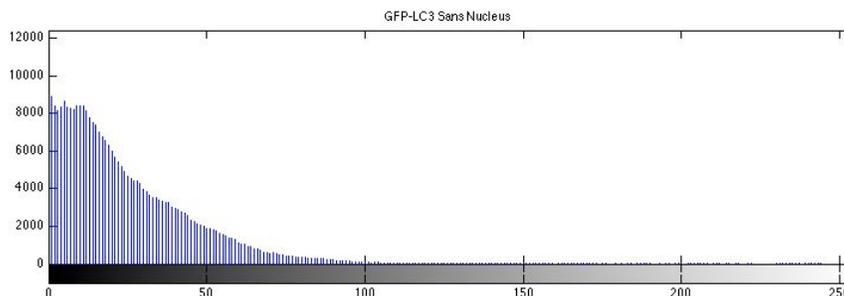
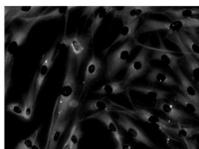
- Boehm M, Nabel EG. "PubMed." *The cell cycle and cardiovascular disease*. March 2003.
<http://www.ncbi.nlm.nih.gov/pubmed/14593697> (accessed April 2008).
- Gonzalez, Rafael C, Richard E Woods, and Steven L Eddins. *Digital Image Processing using MATLAB*. Upper Saddle River, NJ: Prentice Hall, 2004.
- Hippert, Melanie M., Patrick S. O'Toole, and Andrew Thorburn. "Autophagy in Cancer: Good, Bad, or Both?" *Cancer Research* 66 (2006): 9349-9351.
- Juhasz, Gabor, and Thomas P. Neufeld. "Autophagy: A Forty-Year Search for a Missing Membrane Source." *PLoS Biology* 4, no. 2 (2006): 161-164.
- Klionsky, Daniel J., and Scott D. Emr. "Autophagy as a Regulated Pathway of Cellular Degradation." *Science* *Biology* 290 (2000): 1717-1721.
- Klionsky, Daniel J., Hagai Abeliovich, Patrizia Agostinis, Devendra K. Agrawal, Gjumrakch Aliev, and David S. Askew. "Guidelines for the use and interpretation of assays for monitoring autophagy in higher eukaryotes." *Autophagy* 4, no. 2 (February 2008): 151-175.
- Shintani T, Klionsky DJ. "Autophagy in health and disease: a double-edged sword." *Science*, no. 306 (2004): 990-995.

Appendix

Image Histograms

Below are additional histograms for the several image types from the original data set with which I worked on this project. Of particular importance is the histogram of the GFP-LC3 stain found in pane two. Here we can see something of a bimodal peak (though it is very weak) suggesting two main sections of image intensity. While this result seemed promising, it completely disappeared after nucleus subtraction. This subtraction is shown in pane 4. Here we see the histogram has completely lost any bimodality. In addition it is important to note that the backgrounds of these images are not necessarily zero. In fact, the phalloidin stain found in pane three has a baseline value of nearly 50, or roughly 20% of the maximum allowable intensity. As a result it was necessary to convert each of these images to some standard baseline level through a subtraction of their minimum intensity value.



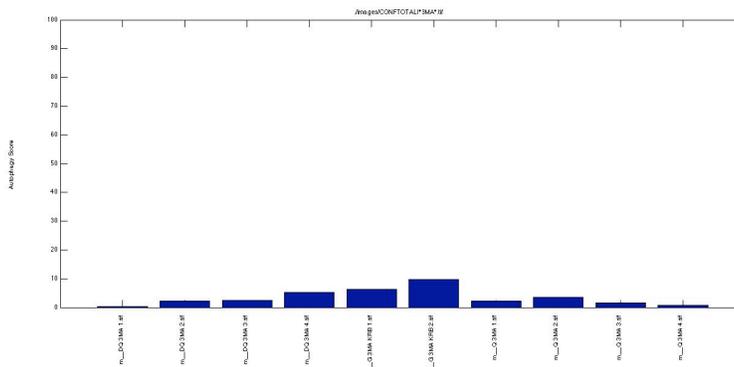


A Comparison of Conditions

The following are plots made of each of the different conditions found in the second data set from the initial program I developed. They seem to correlate well to the expected results. For example, KRB was expected to induce autophagy and has the highest average autophagy score. Each bar in the plot represents a single image's autophagy score as determined by the equation shown above in the paper. It is important to note when examining these plots that there is still some progress to be made on the autophagy scoring mechanism, as clearly there are some images that have scores well above what should be acceptable. However, when these images are combined in a high-throughput screen of autophagic cells containing hundreds of images, the effects of these outliers will likely be minimal.

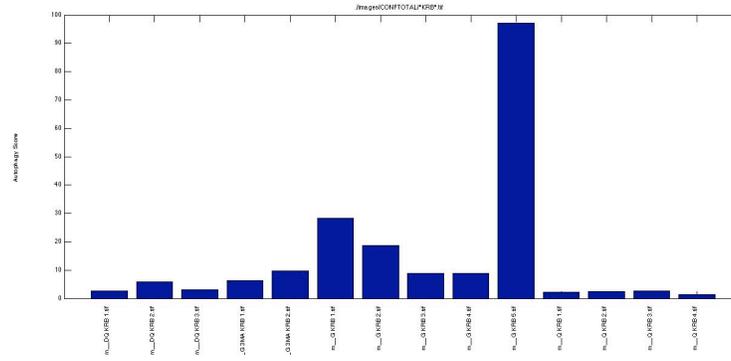
3MA

Inhibits Autophagy



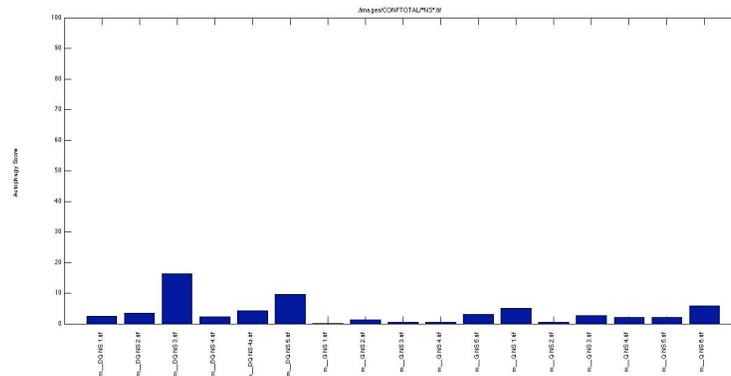
KRB

Induces Autophagy



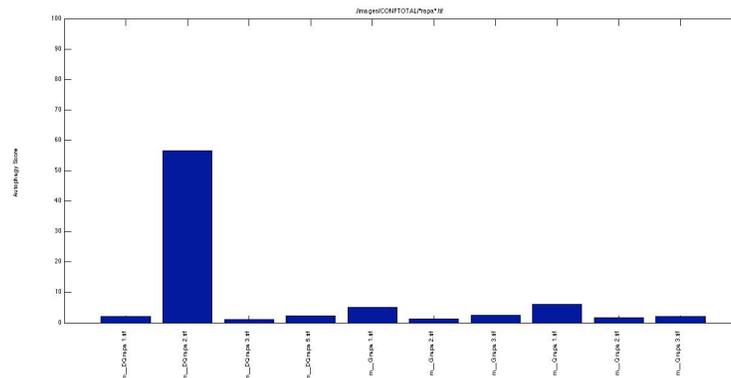
NS – Not Starved

Baseline Autophagy



RAPA - Rapamycin

Induces Autophagy



SS - Serum Starved

Induces Autophagy

