An ImageJ Macros for the Automated Analysis of *S. Cerevisiae* Populations using Microscopy Data

Abstract

Microscopy data has previously been used to analyze the growth of a population of yeast cells (Ginovart et al 2018), as this information can reveal how a population is responding to its environmental conditions. Yeast cell populations move through different growth phases over time as individual cells grow buds and then divide (Gregorio and Duennwald 2018), and so the growth of a yeast population can be analyzed using microscopy data by counting the number of cells seen at a specific time point in a time course experiment that do or do not have buds. However, this process is often time consuming, as it requires researchers to both gather a large volume of images and then analyze each one, making computational tools allowing for the automated analysis of groups of images important. This project aimed to construct a program written using the ImageJ Macros programming language that can take folders of images as input and output the number of budded cells and the number of unbudded cells in each image in a csv format that would allow the user to then export the data for further analysis.

Introduction

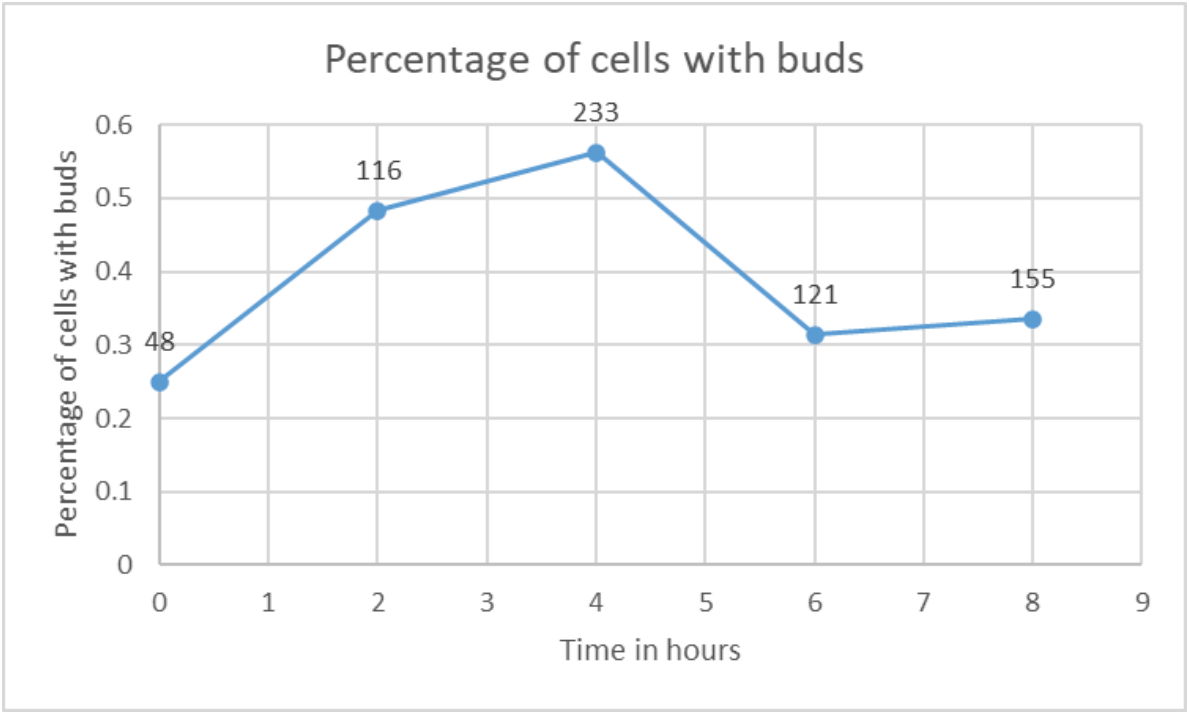
The percentage of a population of *S. cerevisiae* cells with buds varies over time as individual cells grow and the population responds to its environmental conditions (figure 1a), with this percentage typically peaking when cells in the population are in the mid-log phase of growth. The percentage of cells in a population with buds can be measured through placing part of a liquid cell growth culture under a microscope and taking microscopy images for later analysis. The software ImageJ is often used by researchers to analyze microscopy images, but manually analyzing large numbers of images and large numbers of cells can be time consuming, while analyzing less data can harm the accuracy of the data collected. Therefore, this project aimed to develop an ImageJ Macros program that can take an entire directory of images as an input and return the number of budded cells and the number of unbudded cells in each image to the user as an output. Images of *S. cerevisiae* cells taken over an eight hour time course under two different environmental conditions, standard YPD media and YPD media adjusted to a pH of 8, were used to construct this program. All images were 2048 by 2048 pixels large and represented an area of 133.12 by 133.12 um.

Results

The program takes each image in a directory and processes it before analyzing the cells it finds in the image and classifying said cells as either budded or unbudded. The first step of processing is, if necessary, converting the image to an 8-bit grayscale format before applying the threshold function in ImageJ to generate a binary image. This image is then converted to a binary mask, a process that involves inverting all of the pixels in the image, turning white pixels into black pixels and black pixels into white ones (figure 1b). This is necessary because the built-in functions in ImageJ that can identify objects in an image identify white space as an object, not black space, and so having the cells be black and the background be black can cause issues, particularly with the software measuring the inner parts of hollow cells. This is part

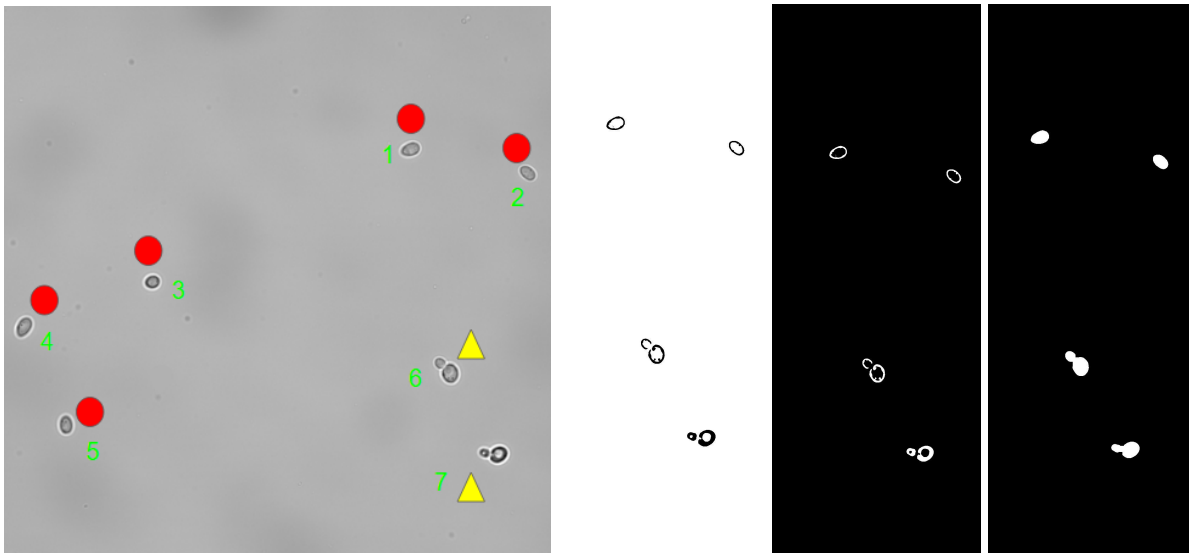
of why cells are measured most accurately when a function to fill in the outlines of the cells detected is used, but this function often completely erases cells depicted in black pixels on a white background. Because the borders of cells are often broken up during the thresholding process, several rounds of dilation are used to connect the borders, followed by several rounds of erosion to return the cells to their original size. When the cells are most dilated, the fill function is used to ensure that the cells will remain intact during the erosion process. Unfortunately, sometimes during dilation cells collide with each other or with other noise in the image. For each image used as data to construct this program, the minimum number of rounds of dilation needed to connect all cells and the maximum number of rounds of dilation able to be tolerated by each image before collisions began were noted. The result of this analysis was that 96.7% of images needed 6 rounds of dilation or less, while 96.99% of images could tolerate at least 6 rounds of dilation, meaning that 6 rounds of dilation was the number that best maximized the number of cells with properly reconstructed borders while minimizing the number of cells that collided with other objects in the image. From there, 6 rounds of erosion returned the cells to their original sizes, creating the image seen in figure 1b.

The last step prior to cell classification was to identify objects within a certain set of parameters that indicate that they are cells and not random noise transformed into objects by processing, and gather various measurements from each of these cells for use in the classification process. In order to avoid counting noise as cells, a specific range of measurements for both the size and circularity of cells needed to be designated. For this, the minimum cell size, maximum cell size, minimum noise object size, maximum noise object size, and minimum and maximum circularity for the cells in the image were documented. Analysis of this data revealed that cells ranged in size from 3.004 um^2 to 35.785 um^2 with a circularity range from 0.321 to 0.97, with the maximum possible measurement being 1. Because the smallest noise particle too large for this range was 67.096 um^2, the size 67.095 um^2 was chosen as the upper threshold. Most instances of noise fell outside of the size range, but instances that did not were further analyzed in order to ensure that the combination of size and circularity as thresholds prevented most if not all instances of noise were not counted as cells. Of the 38 images with instances of noise that were not eliminated by the size range, 13 had all instances of noise removed by the circularity threshold. This means that out of the 226 images in which some noise was detected, all noise was removed from the pool of data points analyzed by the program as cells for 201 images, for a noise removal success rate of 89.94%. Out of concern that lowering the size or circularity thresholds further would result in large numbers of cells not being counted and analyzed, the ranges 3.004 um^2 to 67.095 um^2 and 0.321 to 1 were used in the program as thresholds. It was found that 99% of budded cells had a circularity of 0.785 or less, while 99% of unbudded cells had a circularity measurement greater than 0.785, so 0.785 was chosen as the initial threshold used to separate budded cells from unbudded cells (figure 1d). Additionally, for cells with a circularity measurement greater than 0.785, it was found that 93% of unbudded cells had a perimeter (measured in um) of less than 13.784, while 90% of budded cells had a perimeter of 13.784 or more. For cells with a circularity of 0.785 or less, 91% of budded cells had a perimeter greater than 13.042, while 86% of unbudded cells had a perimeter of 13.042 or less. As a result, these thresholds were used in addition to the original circularity measurement to classify cells (figure 1d).



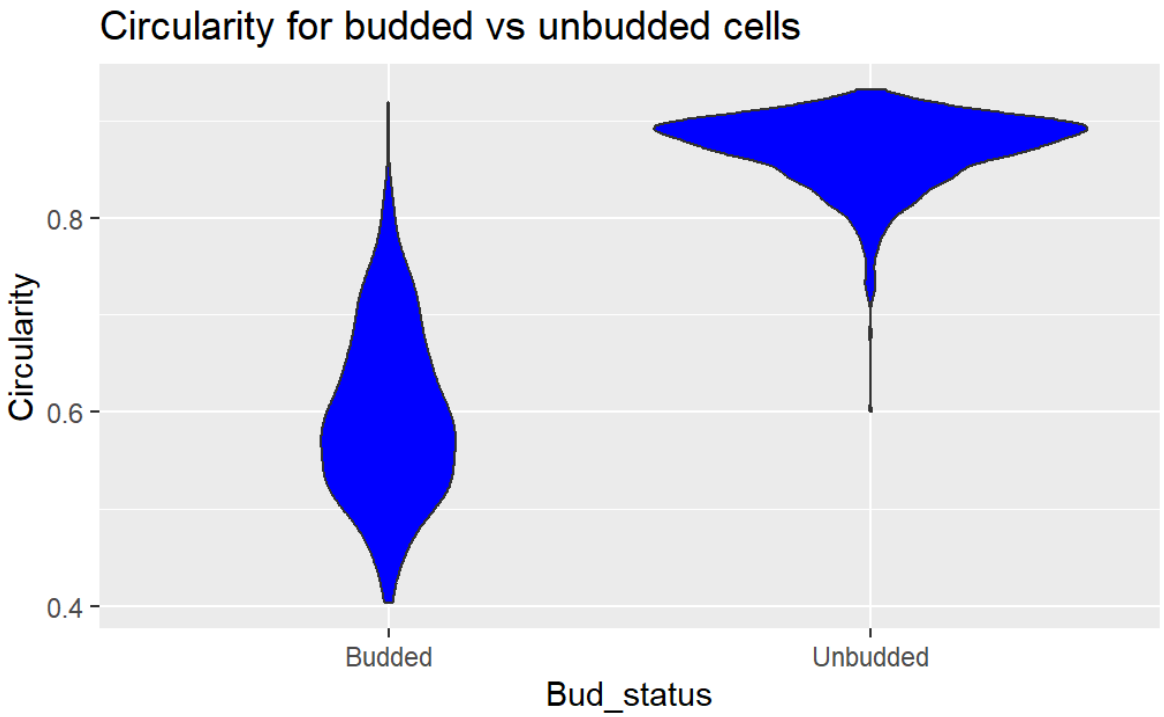
a).

b).

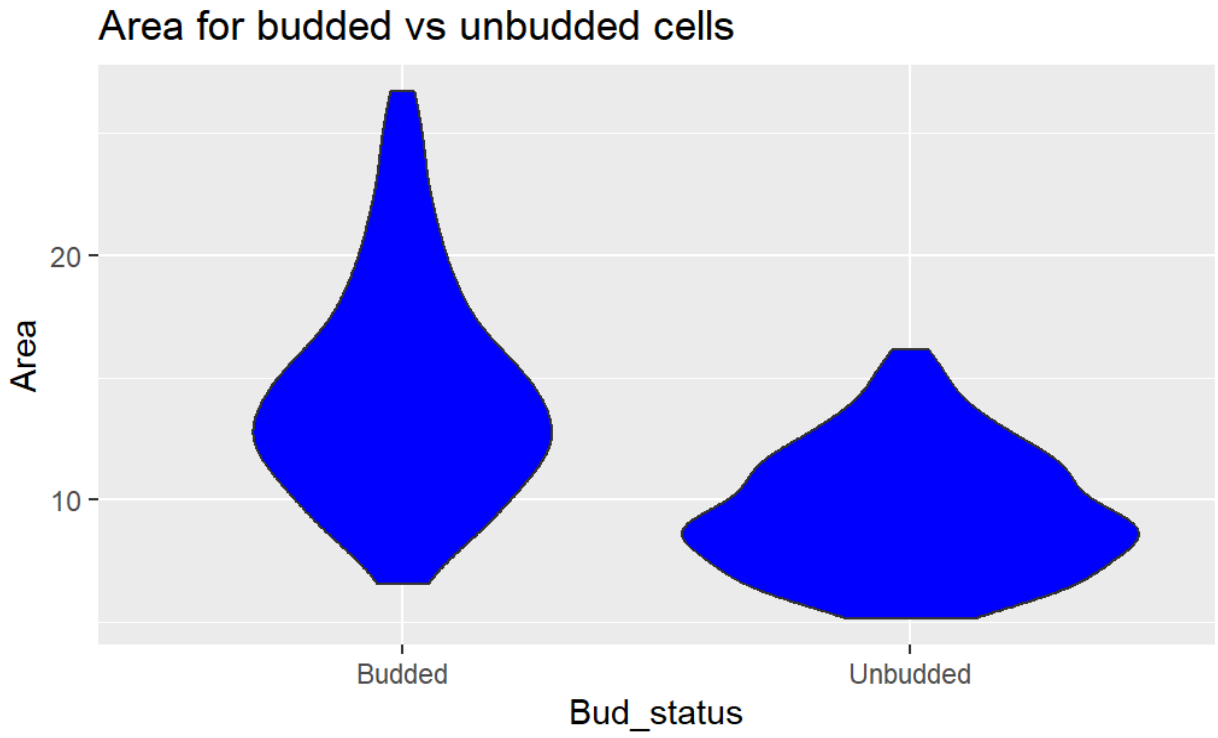


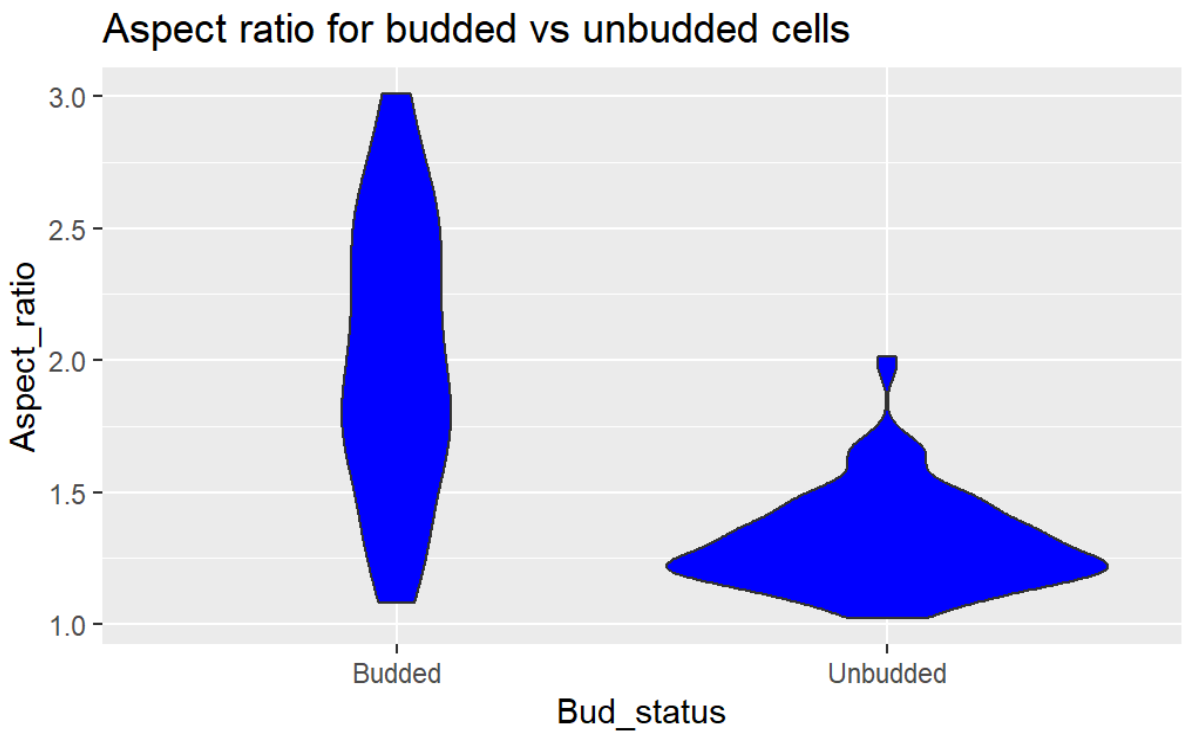
Cell number	Circularity measurement	Area measurement	Aspect ratio	Perimeter
1	0.84	10.152	1.187	10.401
2	0.855	8.657	1.519	11.283
3	0.9	7.753	1.478	11.771
4	0.834	11.255	1.489	12.323
5	0.842	9.286	1.973	13.02

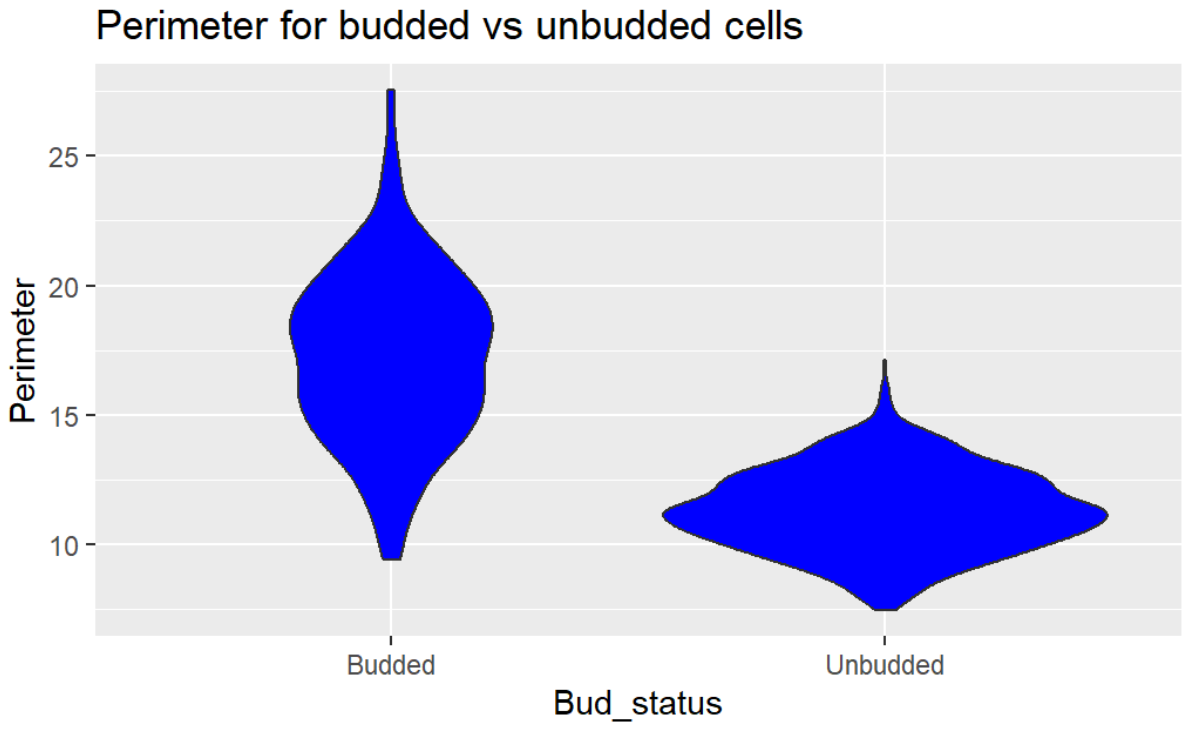
6	0.621	18.083	1.584	19.586
7	0.608	16.262	1.898	19.066



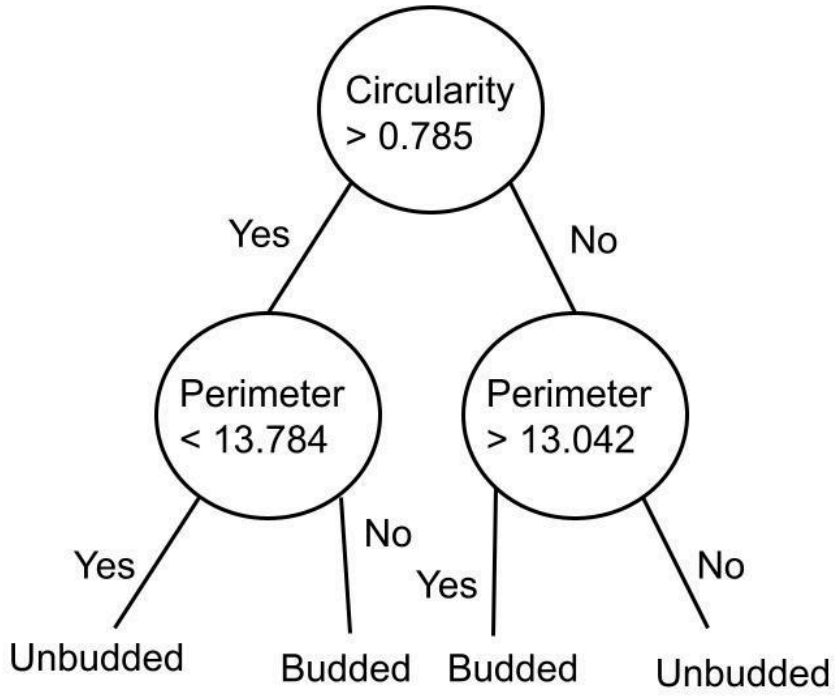
c).




Aspect ratio for budded vs unbudded cells
3.0
2.5
2.0
1.5
1.0
Aspect_ratio
Budded
Unbudded
Bud_status


Perimeter for budded vs unbudded cells
25
20
15
10
Perimeter
Budded
Unbudded
Bud_status

d).



e).

Perimeter	Objects counted as budded cells	Objects counted as unbudded cells
Accurately classified	421	844
Inaccurately classified	100	91
Total	521	935
Percent accurate	81%	90%

a). Graph showing the percentage of cells photographed at a given time point to be budded. Each data point shows a number corresponding to the number of cells photographed at the corresponding time point.

b). Example of an image used in constructing the program, with each cell numbered, unbudded cells indicated with red circles, and budded cells indicated with yellow triangles. A table containing morphological measurements for each cell can be found below the image. To the right of the image, a series of three panels demonstrating the image processing pipeline used by the program to the detection and measurement of cells.

c). A series of violin plots showing the distribution of four different measurements between budded and unbudded cells. Circularity and perimeter were chosen for use in the decision tree

because they possessed the greatest difference in distribution and the clearest thresholds between the two categories.

d). A visual representation of the decision tree used by the program.

e). A classification table showing the accuracy of the program.

Discussion

This project resulted in an ImageJ Macros program that can analyze directories of microscopy images of *S. cerevisiae* cells and return the number of budded and unbudded cells in each image to the user. This project also outlined the morphological parameters that vary the most between budded and unbudded yeast cells (figure 1c). While the exact thresholds most effective for the classification of other species of yeast may vary, this project was able to construct a general decision tree that can be altered for similar studies of different types of yeast cells in the future.

Materials and methods

S. cerevisiae cells were streaked out on plates of YPD and allowed to grow at room temperature for 72 hours. Then, the cell plates were used to inoculate liquid cultures made with either 3 mL of standard YPD liquid media or YPD liquid media adjusted to pH 8. The cell cultures were then incubated at 30 degrees Celsius for 24 hours. 50 uL of each solution was then removed and placed into a new solution containing an additional 3 mL of liquid growth media. 10 uL of each culture of these new cultures was removed for the purposes of imaging before all cultures were placed back in a 30 degree Celsius incubator. The 10 uL previously removed were diluted with 1 mL of either standard YPD or pH 8 YPD depending on which media the cells had previously been grown in. These solutions were sonicated with 23% amplitude at 10 seconds each to break up cells and provide single cells to image before 15 uL of each solution was removed and placed on a microscope slide for imaging. This process was repeated every 2 hours with some variations in the amount of cell culture removed for dilution and imaging to account for the growth of the cell population and the need for single cells to image. At 2 hours, 80 uL were used, at 4 hours, 40 uL were used, at 6 hours, 20 uL were used, and at 8 hours, 10 uL were used. Images were taken at 100x magnification and each image represented 133.12 um by 133.12 um of area on the microscope slide. The images taken, ImageJ script created, and additional materials are available at the GitHub repository https://github.com/FrannieMurphy903/Summer2025.

References

Ginovart, Marta, Rosa Carbó, Mónica Blanco, and Xavier Portell. “Digital Image Analysis of Yeast Single Cells Growing in Two Different Oxygen Concentrations to Analyze the Population Growth and to Assist Individual- Based Modeling”. Frontiers in Microbiology, January 4, 2018, 2628. https://doi.org/10.3389/fmicb.2017.02628

Di Gregorio, Sonja & Duennwald, Martin. “Yeast as a model to study protein misfolding in aged cells”. FEMS yeast research, May 13, 2018, foy054.
http://dx.doi.org/10.1093/femsyr/foy054