# Time-lapse Microscopy Exercise

*NOTE: All the data for this exercise are in a separate ~500MB zip file. This exercise does not come with a separate hints document. For hints or guidance, refer to the solution file* `TimelapseMicroscopySolution.m`.

In this exercise, you will analyze time-lapse imaging data to see whether there is a growth cost to yeast for responding to galactose. In other words, you will look at an image sequence of yeast cells growing under a microscope. In the middle of the experiment, galactose is added to the culture media. Your job is to determine if the cells temporarily grow more slowly due to the addition of galactose.
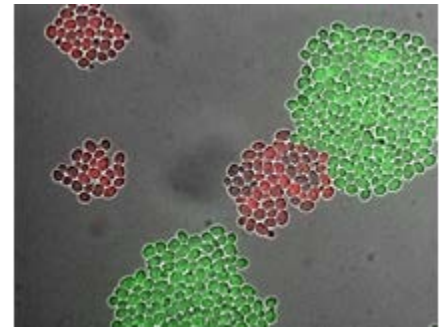
This requires writing code for multiple tasks, so let's go through it step by step. As usual, we start by exploring the data, then we prototype our analysis code on a small subset of the data, and finally we scale up to analyzing all the data, and plot our results.

**Exercise 1. Explore the data.** For this experiment, this means watch a movie of the cells.

a. **Load the file `overlay_sequence.mat` into MATLAB.** This contains a pre-processed image sequence which shows cells growing. (This is a 300MB file and may take a minute.) What variable is loaded? How many dimensions does it have and what does each dimension represent?

b. **Play the image sequence as a movie**. Use MATLAB's `implay` command (type `doc implay` or `help implay` to figure out how to use `implay`). What do you see?

   You'll notice that some cells are green and some cells are red. These are actually two different genotypes. The green cells are wildtype yeast that have been tagged with YFP (i.e. they express Yellow Fluorescent Protein constitutively). The red cells are ΔGAL2 and tagged with mCherry. Wildtype yeast (green) respond to galactose by inducing the GAL pathway, a set of genes that allow the cells to metabolize galactose and grow. ΔGAL2 cells (red) lack the galactose transporter and should not induce GAL genes when galactose is introduced.

c. **Can you tell by eye which strain is growing faster?** This data consists of 97 images, or *frames*, that were taken by the microscope every 10 minutes. The cells were imaged in a chamber that had a constant flow of nutrient media containing raffinose, a relatively poor carbon source. Starting in frame 31 (t=5 hours), galactose was also added to the flow. Figure out how to use `implay` to watch the movie at a slow frame rate (8 frames per second). You can also play the movie slowly with the "next frame" and "previous frame" buttons. The bottom right corner of the window tells you which frame you are on. **Does anything different happen after galactose induction (frame 31)?**

**2. Segment the images**. You can't tell by eye which strain is growing more quickly, let alone by how much and if the growth rate changes over time. We need to quantify cell growth computationally. This means *segmenting*, or identifying the cells, and then extracting some metric of growth from the segmentation.

a. **Plan image analysis strategy.** One way we can quantify growth is to count how many cells there are in each frame, and plot cell counts versus time. While conceptually simple, counting cells can be hard to do accurately without some advanced image analysis operations (see BONUS). Instead of trying to distinguish cells from each other (which is hard if cells are close together or out of focus), we can just segment cells from background using their fluorescence (which is easy to do using thresholding). Then, we can use the total area (i.e. number of pixels) occupied by cells, rather than number of cells, as a proxy for cell mass. **Before writing any code, think about what commands or matrix operations will you need to do the segmentation this way.**

b. **Test out some segmentation code on a single frame.** Load the first frame of the YFP channel (whose filename ends with '**_t1.TIF**' in the folder **images/YFP**) and save it to a variable **img**. Pick a threshold and use it to segment cells from background. How does your segmentation compare to the original image? Where does thresholding tend to make mistakes? Calculate the area of the cells in this frame, in pixels.

   Try your code on other frames, and on the mCherry images. Should you use the same threshold for each image, or calculate different thresholds for different images?

c. **Put your code for segmenting a single frame into a function.** What inputs and outputs should this function have? For example, you can pass a matrix representing a frame and have it return the pixel area of the cells.

d. **Apply your segmentation to all frames.** In other words, loop through all 97 frames in both YFP and mCherry channels, load the corresponding image file, and call your function from part c. Should you use a different threshold for each frame, or a single threshold for all frames?

   Before writing any code, use Windows explorer or Finder to simply look through the images you need to load. What are their file names? How will you load them in the correct order?

   **HINT**: The following code will give you the name of the 10th frame in the YFP channel (and store it in the variable **frame_filename**):

```
YFP_filename_root = 'images/YFP/run1_w2Wheel YFP_s11_t';
frame_filename = [YFP_filename_root num2str(10) '.TIF'];
```

After you run your code, you should have arrays containing the pixel area of cells in each frame of both channels.

**3. Compare the growth of the two strains.** We want to see how adding galactose to the cells affects their growth rate. How do we define growth rate? Let's play with the data and come up with a definition.

    a. **Plot the pixel area of cells versus time (or frame number).** You should plot two lines: one for the YFP cells and another for the mCherry cells. Make sure to label your axes and create a legend for the two lines.

       Try taking the log base 2 of the pixel areas and plotting that instead. Which axis scaling is better?

       At frame 37, galactose is added to the media. Plot a vertical line at the moment of galactose induction. To do this, first realize that `plot([x1 x2], [y1 y2])` will plot a line between the coordinates (x1,y1) and (x2,y2). Now type `ylim` at the command line after you plot something. What does `ylim` return, and where do those values lie on your y-axis? Now how do you use this to plot a vertical line?

    b. **Do some sanity checks on your data.** Do the plots look how you expected them to? Are the cells growing exponentially? What are possible caveats or limitations of this data? For example, before the media switch, we don't expect a difference in growth because ΔGAL2 shouldn't affect metabolism of raffinose. Are your plots consistent with this expectation? If not, what might be the reason?

    c. **Consider the effects of galactose induction.** We are interested in whether the cells grow differently after galactose induction (t=5 hours, or frame 31). Is there a difference between the growth of cells in raffinose versus in raffinose and galactose? How long does it take before there is a difference in growth between the wildtype and ΔGAL2 cells? Is there a cost to GAL pathway induction (i.e. does growth of the wildtype strain slow immediately after galactose induction)? To compare the two strains, you may want to normalize pixel areas to the time of media switch.

    d. **Consider more data.** It's hard to make conclusions from the small number of cells in a single microscope field. To get better statistics, you can image and analyze more fields. In fact, 20 microscope fields (about 15 GB) of data were acquired for this experiment. To save you some time, the results of a threshold segmentation on all 20 fields are provided

for you. **Load the file `allfields.mat`**. This contains data of pixel areas. What variables were loaded and what do their dimensions mean?

**Combine the data from all the fields and plot the results**. How should you combine the data? Note that all the fields are from a single flow chamber under identical culture conditions. How does this affect your conclusion about the effect of galactose induction?

**4 (BONUS). Segment by counting cells.** For the above analysis we used the pixel area of all cells as a proxy for cell mass. An alternative measure of growth is the number of cells over time. Revise your segmentation code to determine the number of cells in each frame, and then reconsider the question of how galactose induction affects growth rate. What are the differences, qualitative or quantitative, between using cell number and total cell area as measures of growth?

You will need the commands **`bwlabel`** and **`regionprops`** to count segmented regions—look these up in the MATLAB documentation (or in Google) to see details. You will also need a variety of filters and morphological operations, which you can learn about in the image processing packet from Day 4. As a start, you may want to consider using **`imtophat`** as a way of evening out fluorescence halo around cells.

*Created 6/1/2012 by JW.*