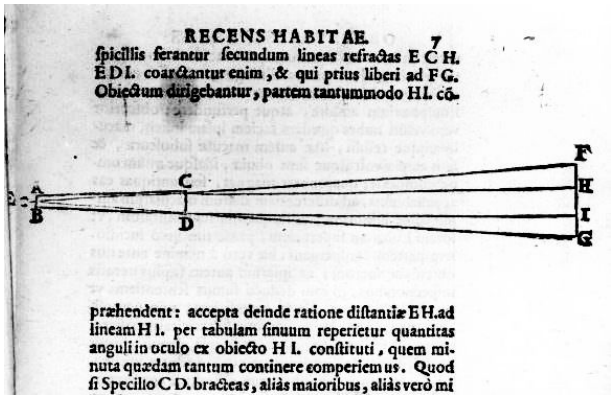
A picture is worth 1000 hypotheses: using computers to analyze biological images

Kyle Karhohs

Broad Institute of Harvard & MIT

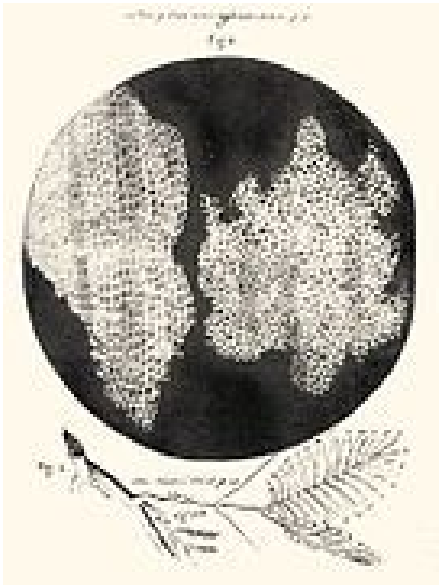
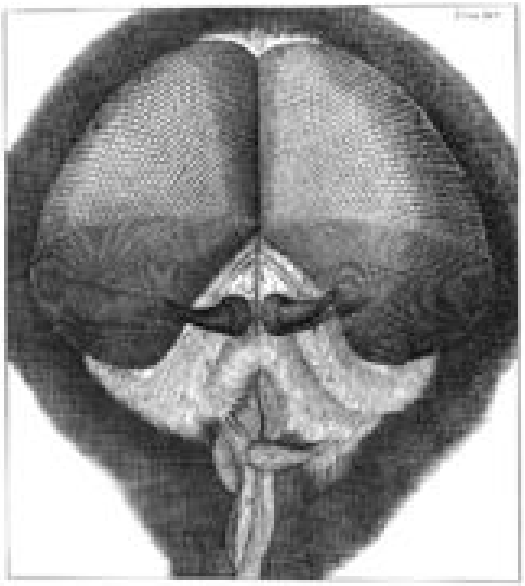
The origins of the microscope

- A microscope sees objects that are too small to see by the human eye.
- Optical microscopes were invented almost simultaneously with the telescope.
  - Galileo invented versions of both circa 1600 shortly after the original was developed in the Netherlands
  - The original compound microscopes can be described as “two lenses in a tube”



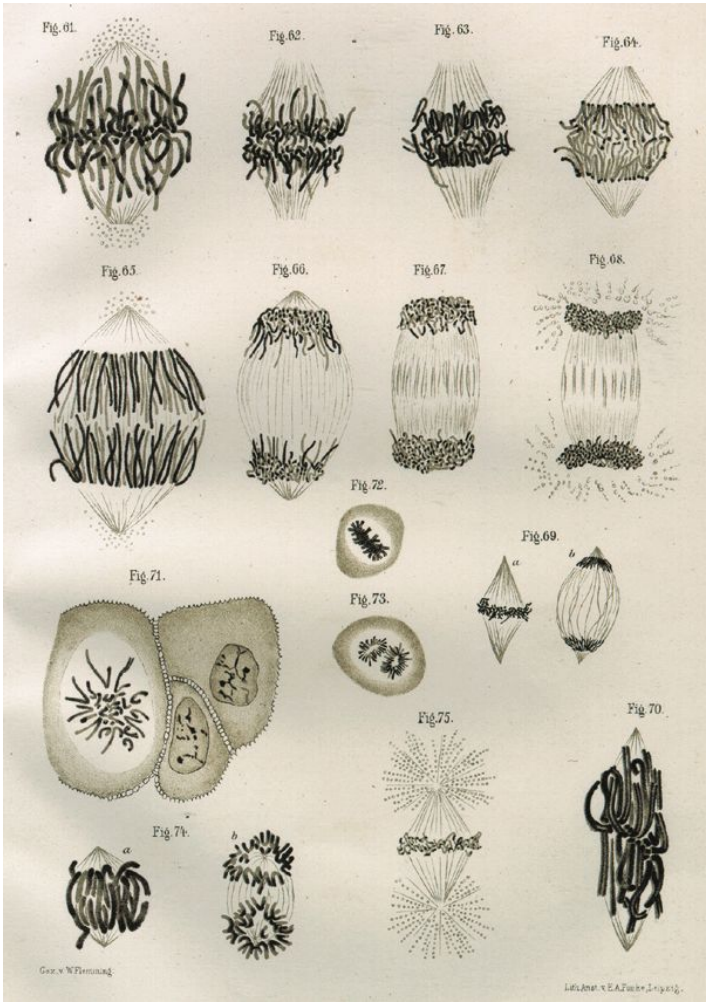
“Images” were originally hand drawings

- Robert Hooke’s Micrographia is a landmark publication documenting what microscopy reveals about the natural world (1665)
- Coined the word “cell” to describe cork

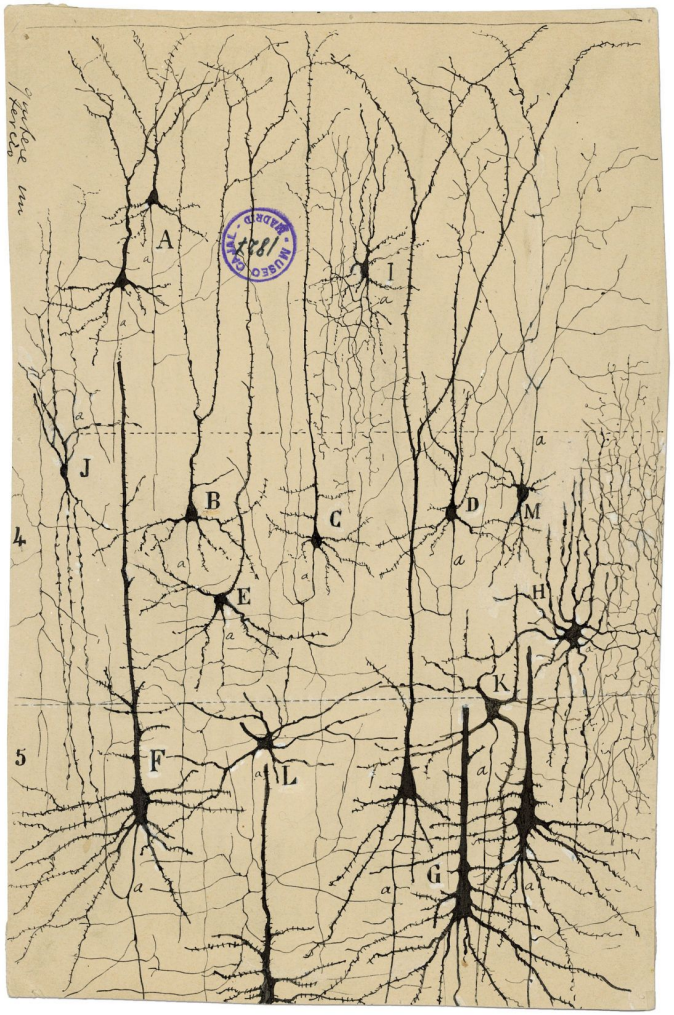

“Cell Theory” was driven by microscopy.

- Cell Theory states that all living organisms are made up of cells.
  - Introducted by Matthia Schleiden after observing hundreds of plants.
- Walther Flemming’s observations in 1882 of mitosis revealed details about the mechanics of how cells divide.
  - Key to explaining how organisms made of cells can reproduce.



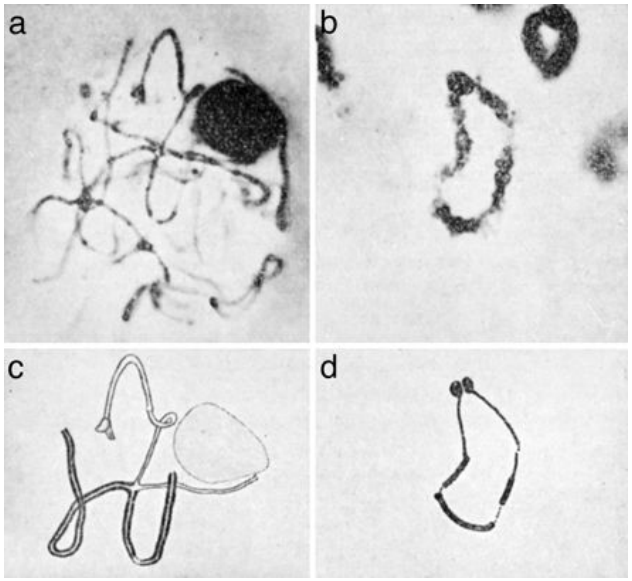
Many discoveries were made by treating biology samples with chemicals

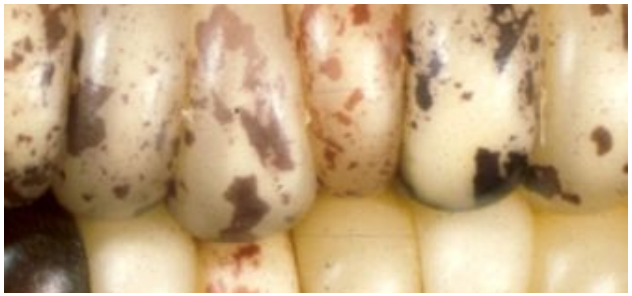
- Innovations in staining biological samples revealed previously unseen aspects of biology
- In 1906, Santiago Cajal won the Nobel Prize for his work investigating neurons using a silver based stain



Microscopy contributes to many fields of biology

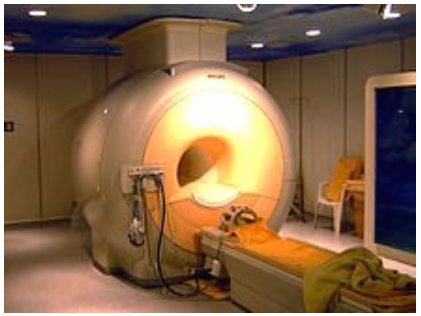
- Barbara McClintock received the Noble Prize in 1983 for her discovery of mobile genetic elements.
  - This discovery gave insight into why spots appear in certain corn species. It is a genetic mechanism also responsible for the dysregulation of genes in cancer.

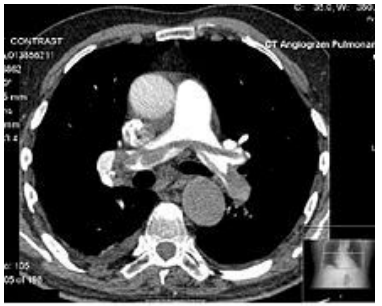


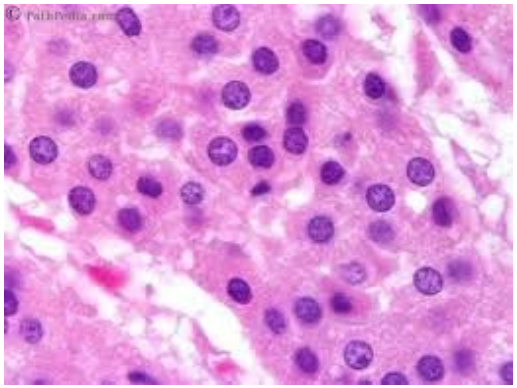


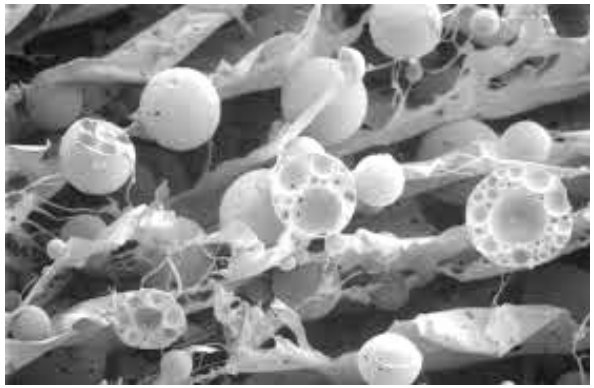
The computer revolution led to the diversification of microscopy and images

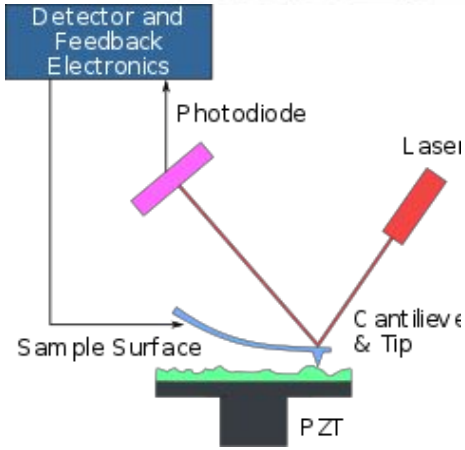
- Images come from many different sources and modalities
  - MRI and CT Scans
  - Atomic Force Microscopy
  - Electron Microscopy
  - *Epifluorescence and Light Microscopy*





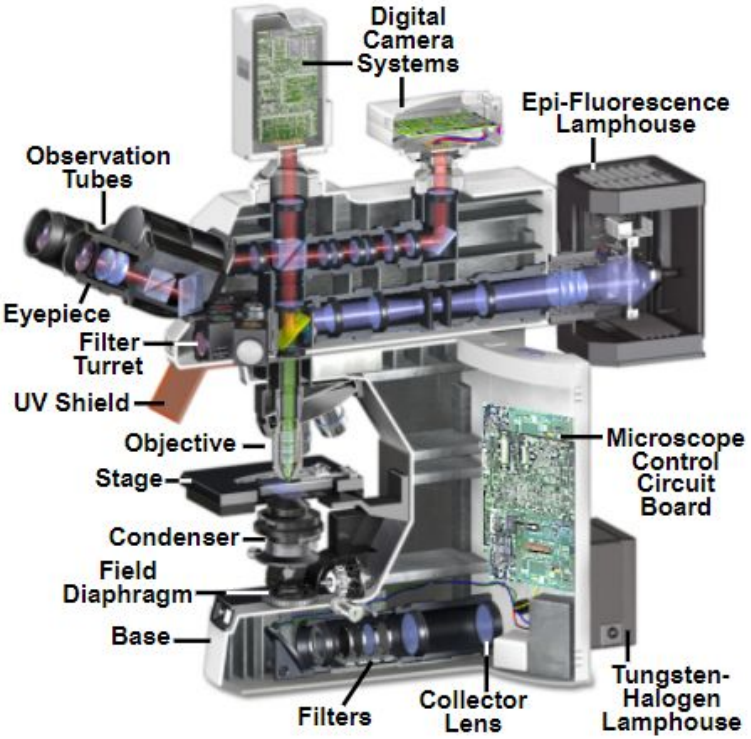


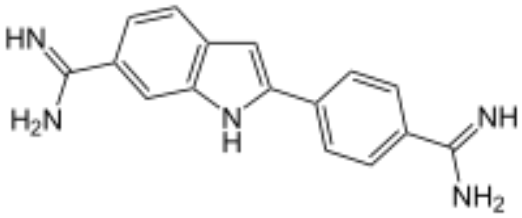


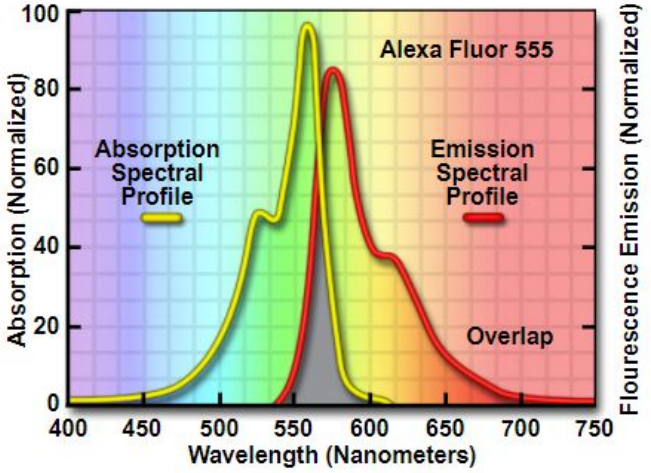


Fluorescence microscopy is essential to modern biology research

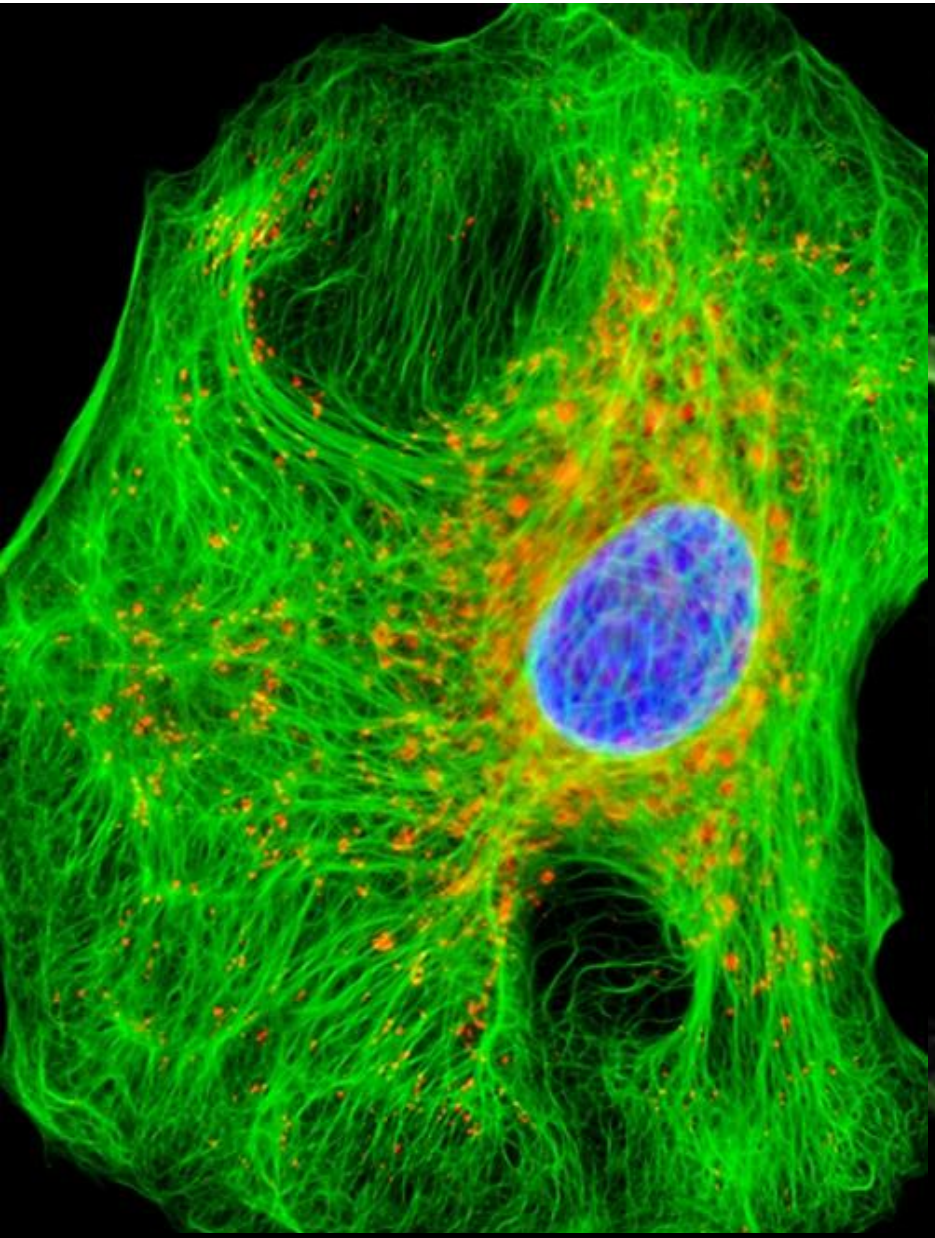
- Microscopes today are a sophisticated combination of lenses, mirrors, and electronics.
- Fluorescence: when a fluorescent compound is excited by light, it then emits light at a slightly longer wavelength after it “relaxes”.
- Fluorescent antibodies, protein-tags, and DNA oligos illuminate details within cells with great specificity.







Microscopes produce images that contain a wealth of information



http://www.microscopyu.com

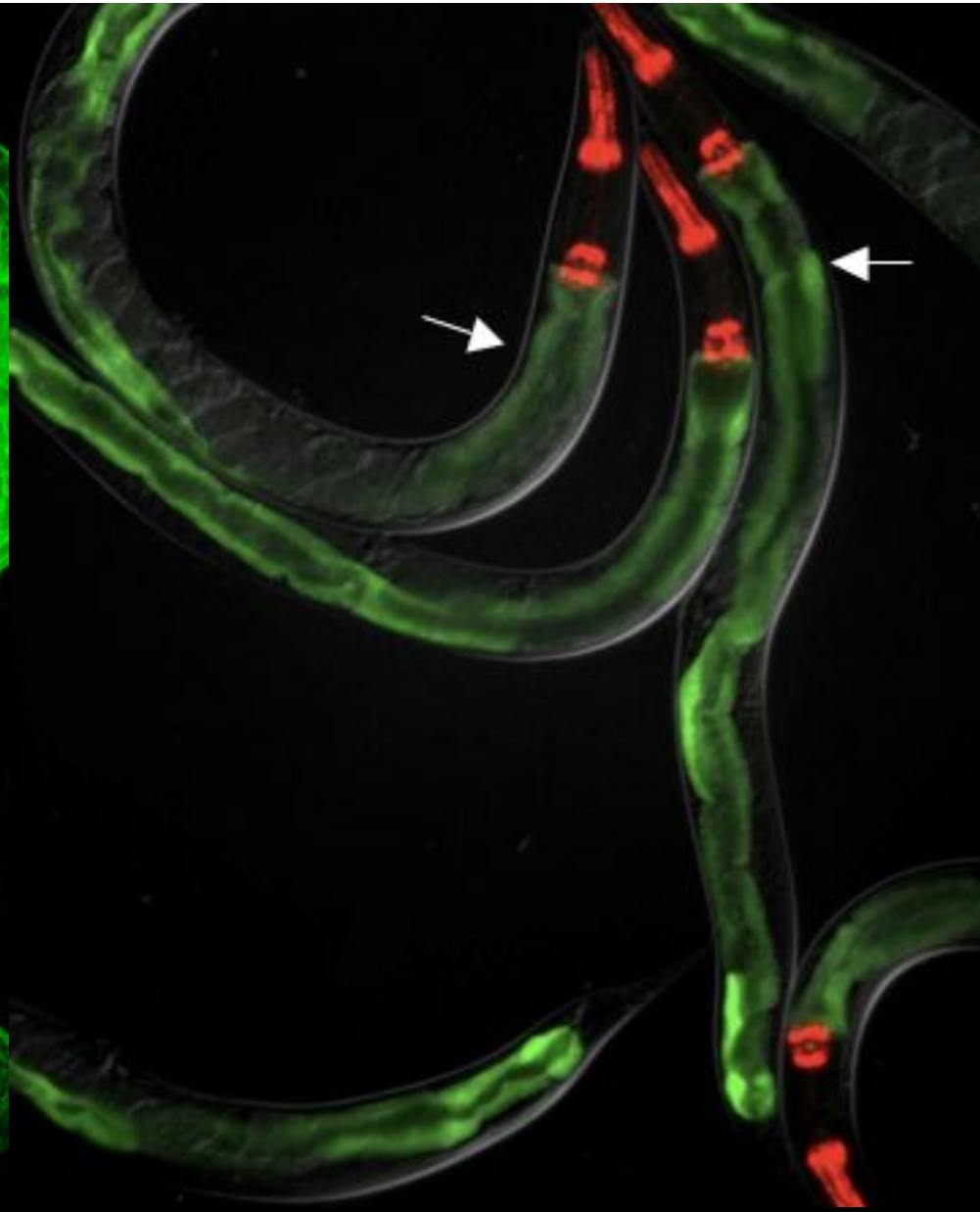
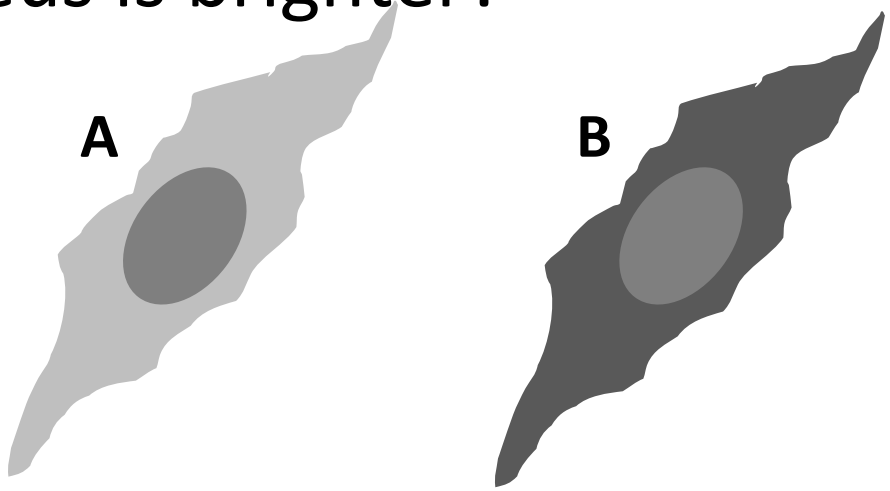


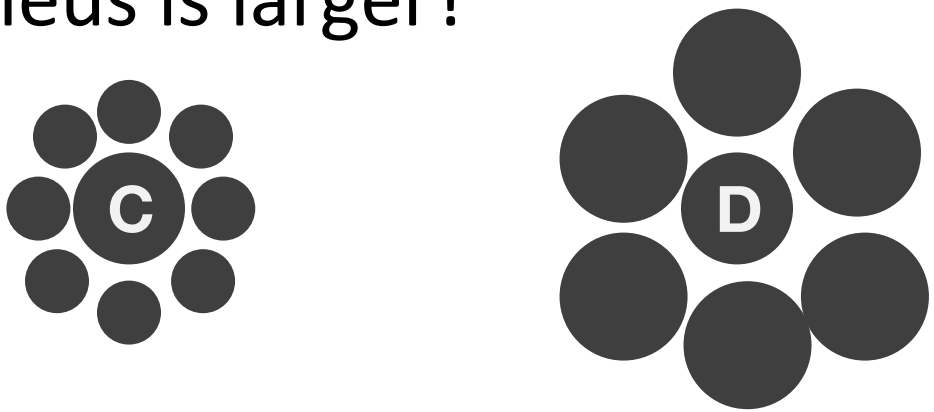
Image: Javier Irazoqui

Quantifying images makes analysis reproducible, automated, and statistical

Which nucleus is brighter?

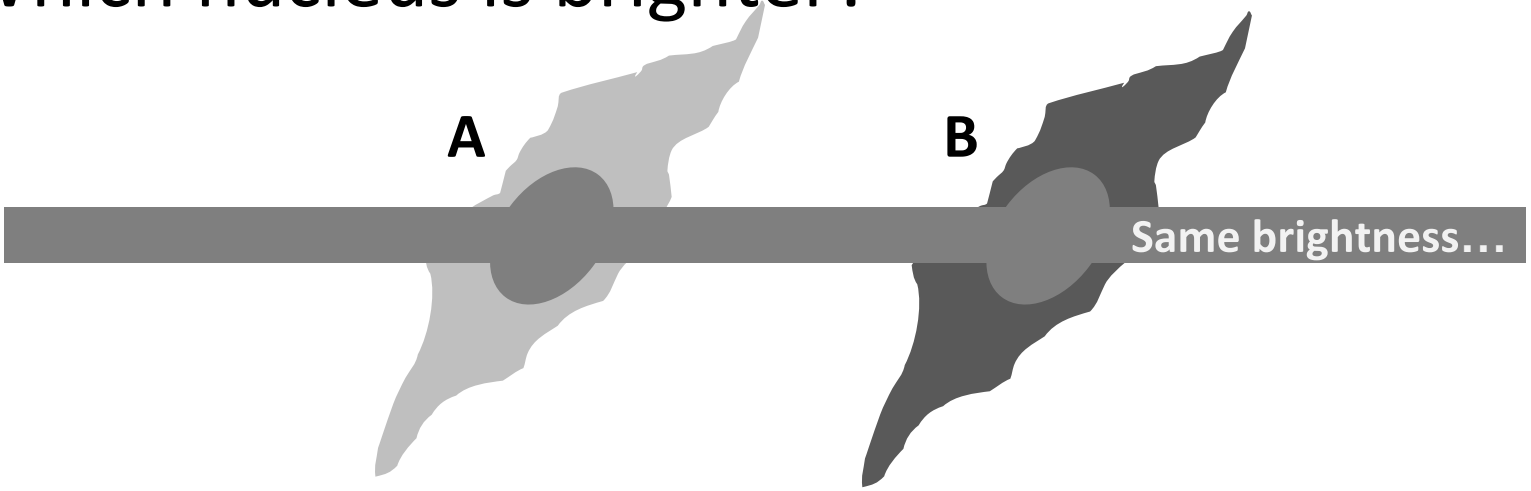


Which nucleus is larger?

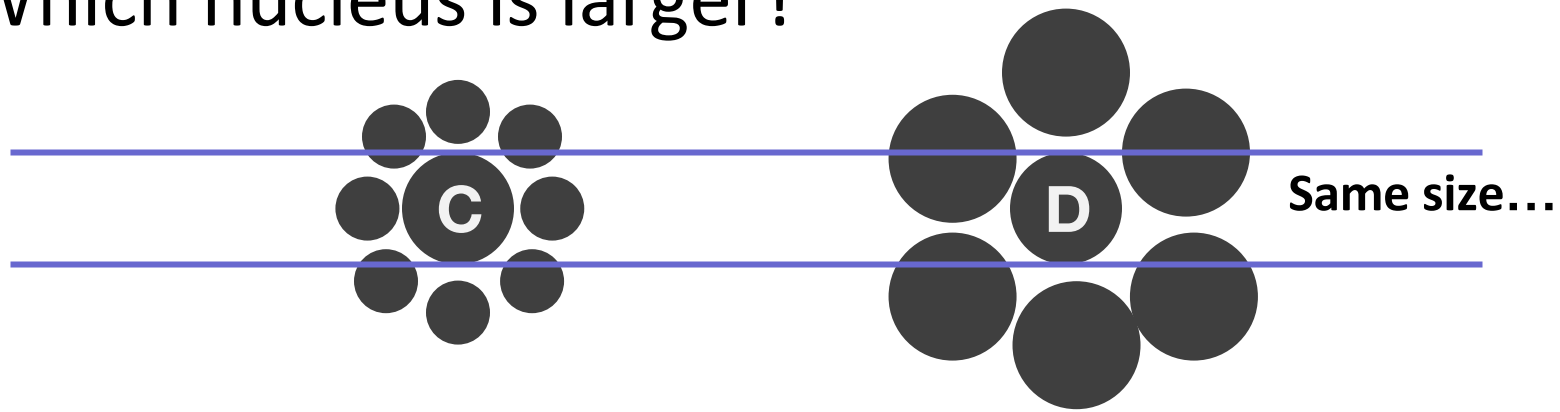


Quantifying images makes analysis reproducible, automated, and statistical

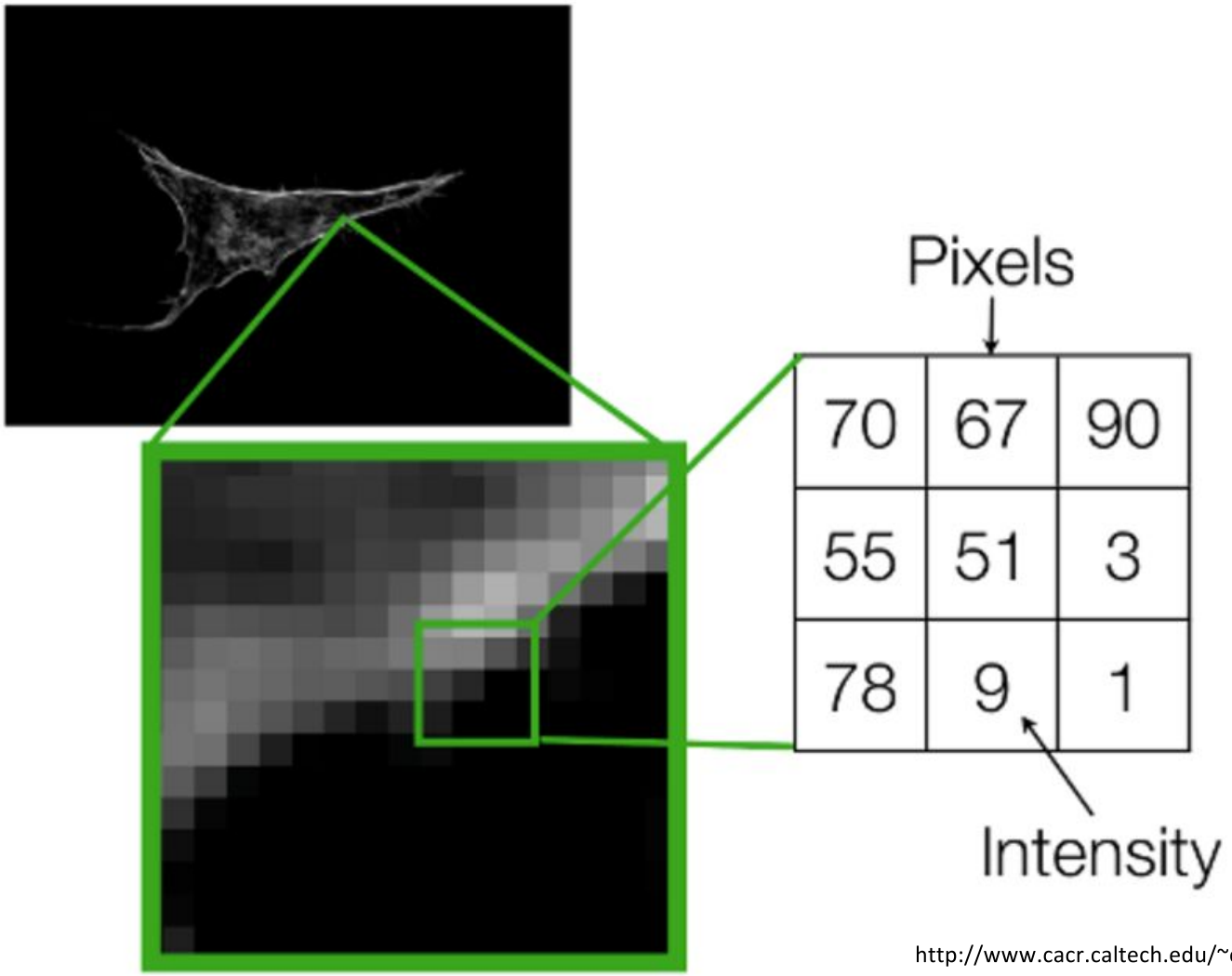
Which nucleus is brighter?



Which nucleus is larger?



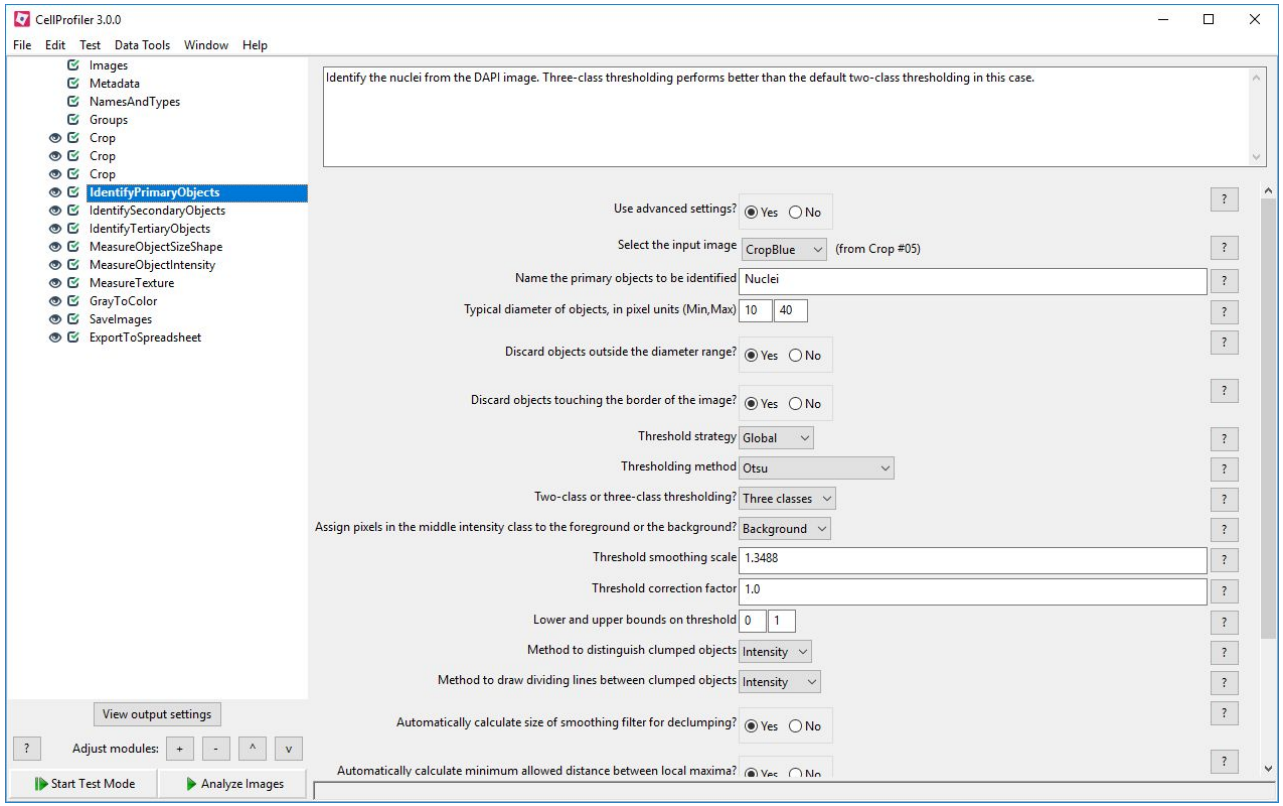
Images are numbers



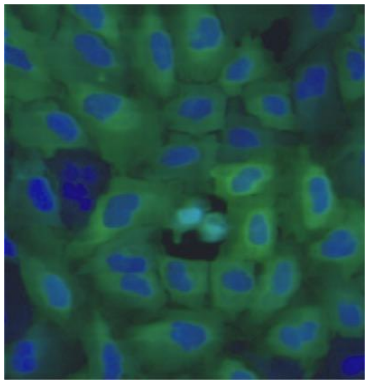
http://www.cacr.caltech.edu/~cunha/bi199/one.html

CellProfiler is a software for analyzing biological images

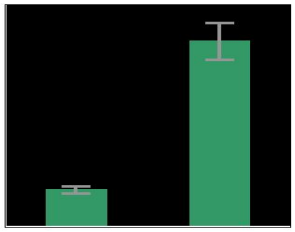
- User friendly interface provides powerful image analysis methods to non-programmers



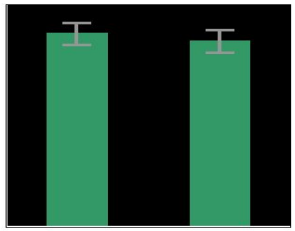
Visual appearance can be quantified



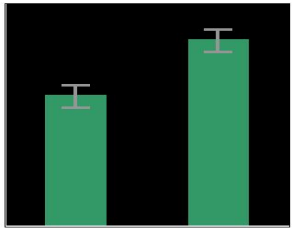
Localization



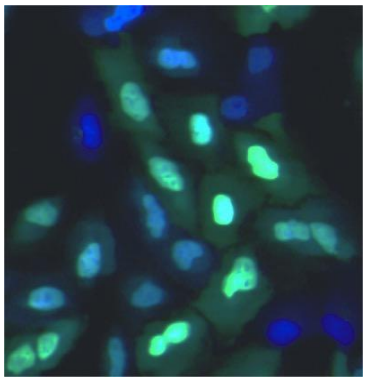
mRNA or protein levels



morphology



… + hundreds of other features



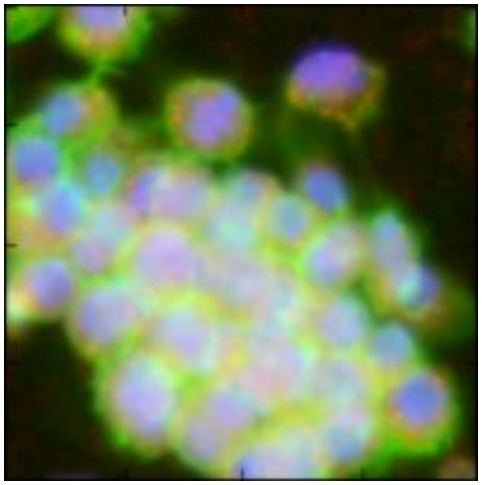
Automatic image analysis:

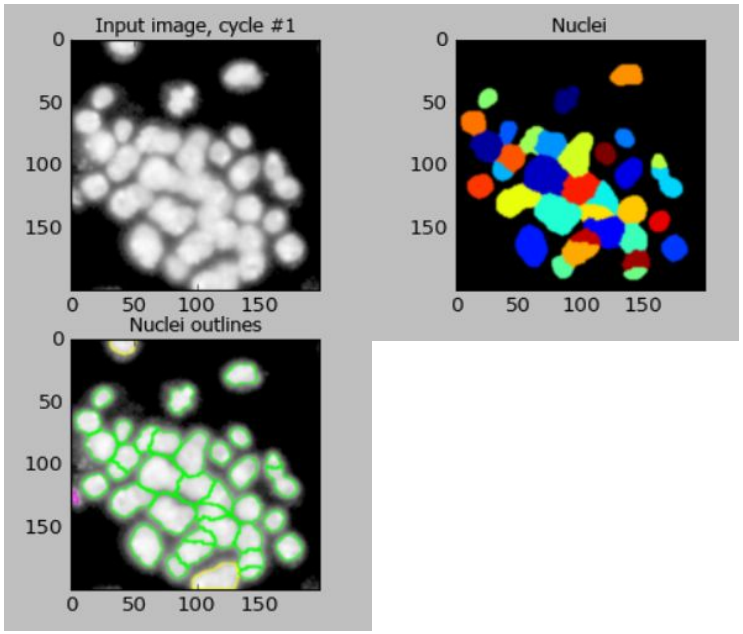
- Objective
- Quantitative, with statistics
- Can measure multiple properties at once
- Can quantify heterogeneity (single-cell measurements)
- Distinguishes subtle changes, even those undetectable by eye
- Faster, less tedious

A pipeline in CellProfiler quantifies the biology within an image

- Measure everything, ask questions later
- A typical pipeline does 3 tasks:

Process Images → **Segment Biology** → **Measurements & Features**





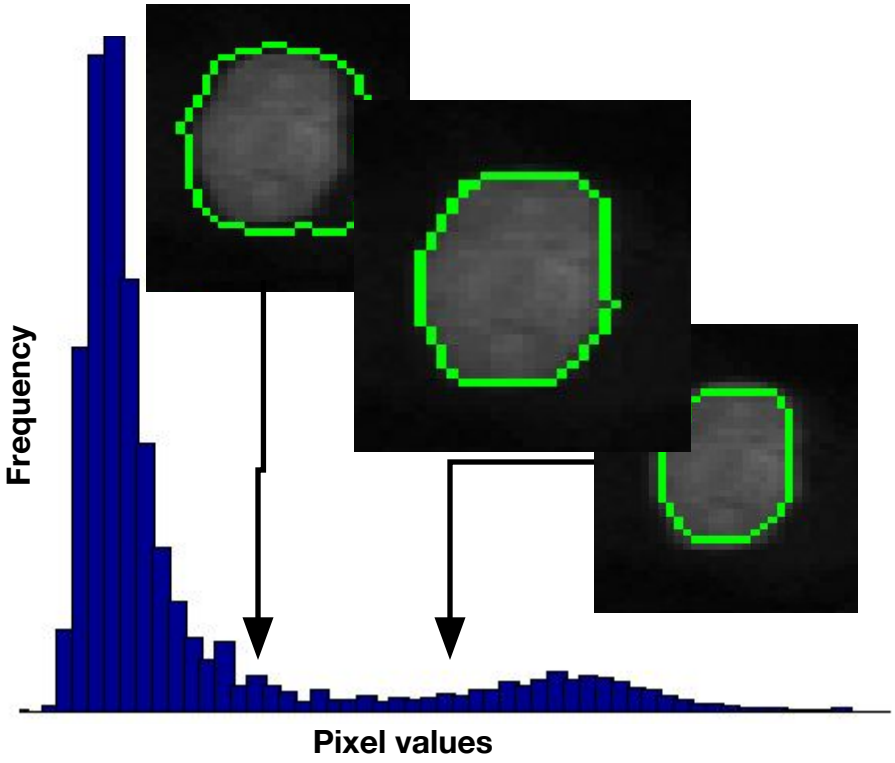
Meas. & Features

Cells

Segmentation - Image thresholding

Definition: Division of the image into background and foreground

What is the best threshold value for dividing the intensity into foreground and background pixels?

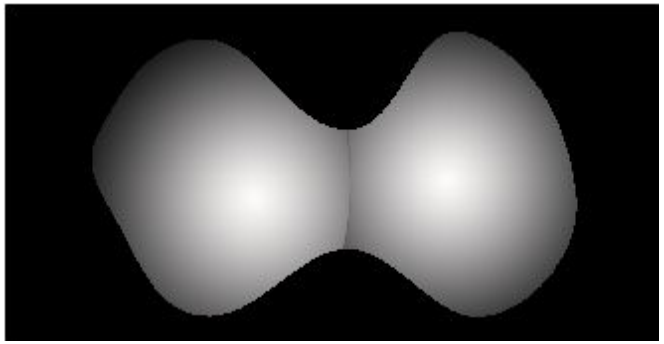


Methods:

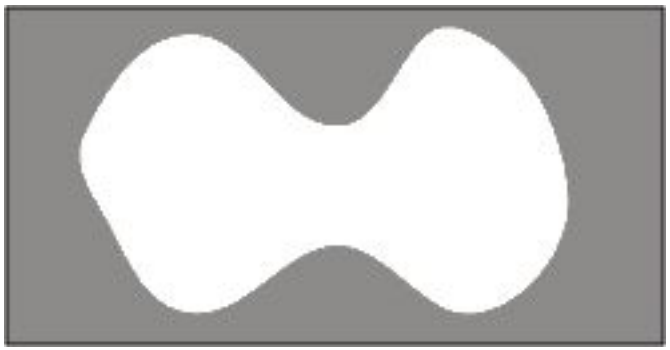
- Automatic: Good for readily identifiable foreground / background
- Otsu: Choose between 2- or 3-class if mid-level intensities present
- Background, RobustBackground: Good for images in which most of the image is comprised of background

Segmentation - Splits and Merges

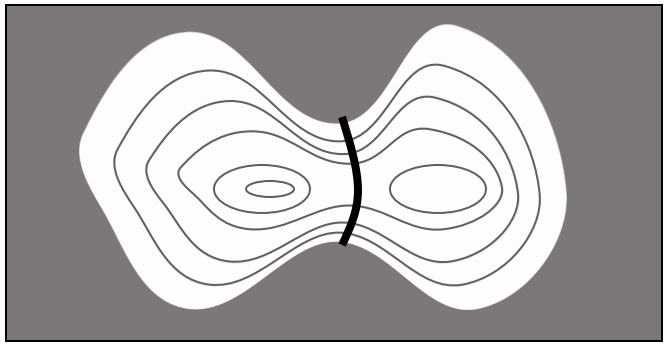
Once the images are loaded, how do you find objects of interest?



- **Step 1:** Distinguish the foreground from the background (thresholding)



- **Step 2:** Split/merge “objects” properly

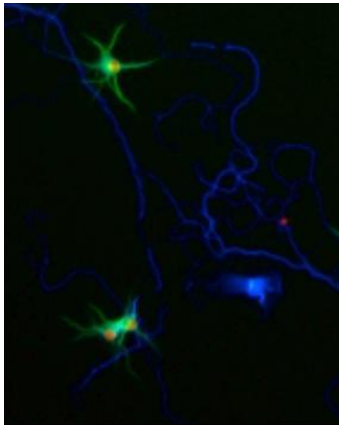


CellProfiler has quantified a wide variety of biology within images

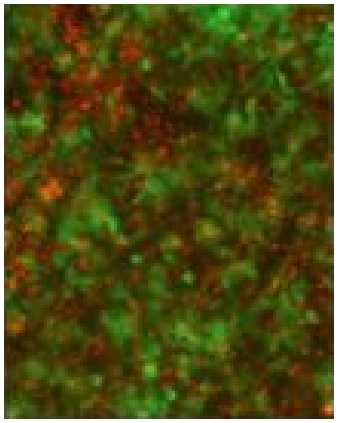
http://cellprofiler.org/examples/

Yeast

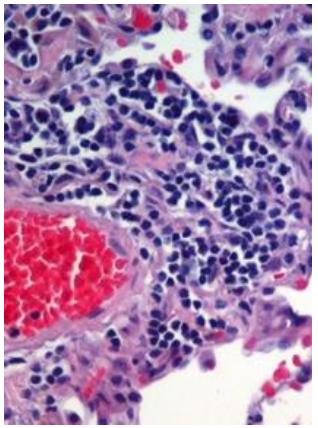
Neurons



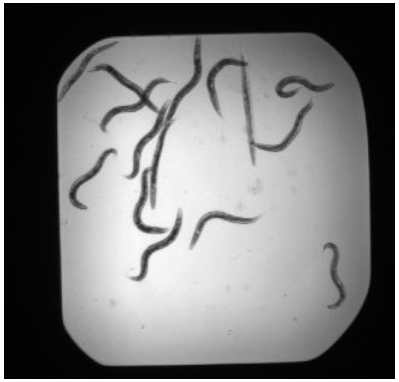
Co-cultures



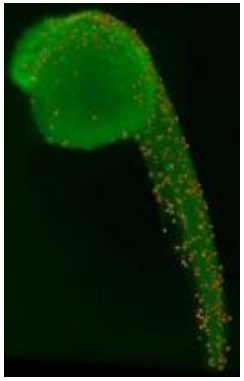
Tissue



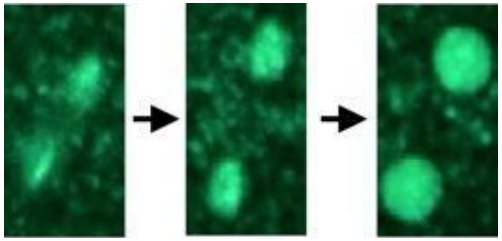
C. elegans



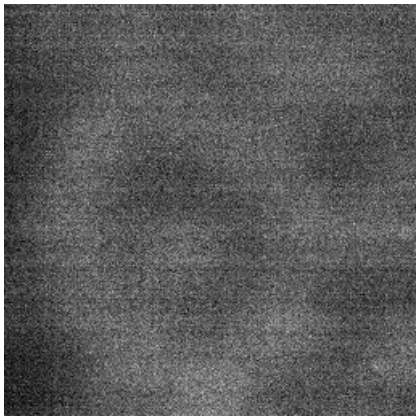
Zebrafish



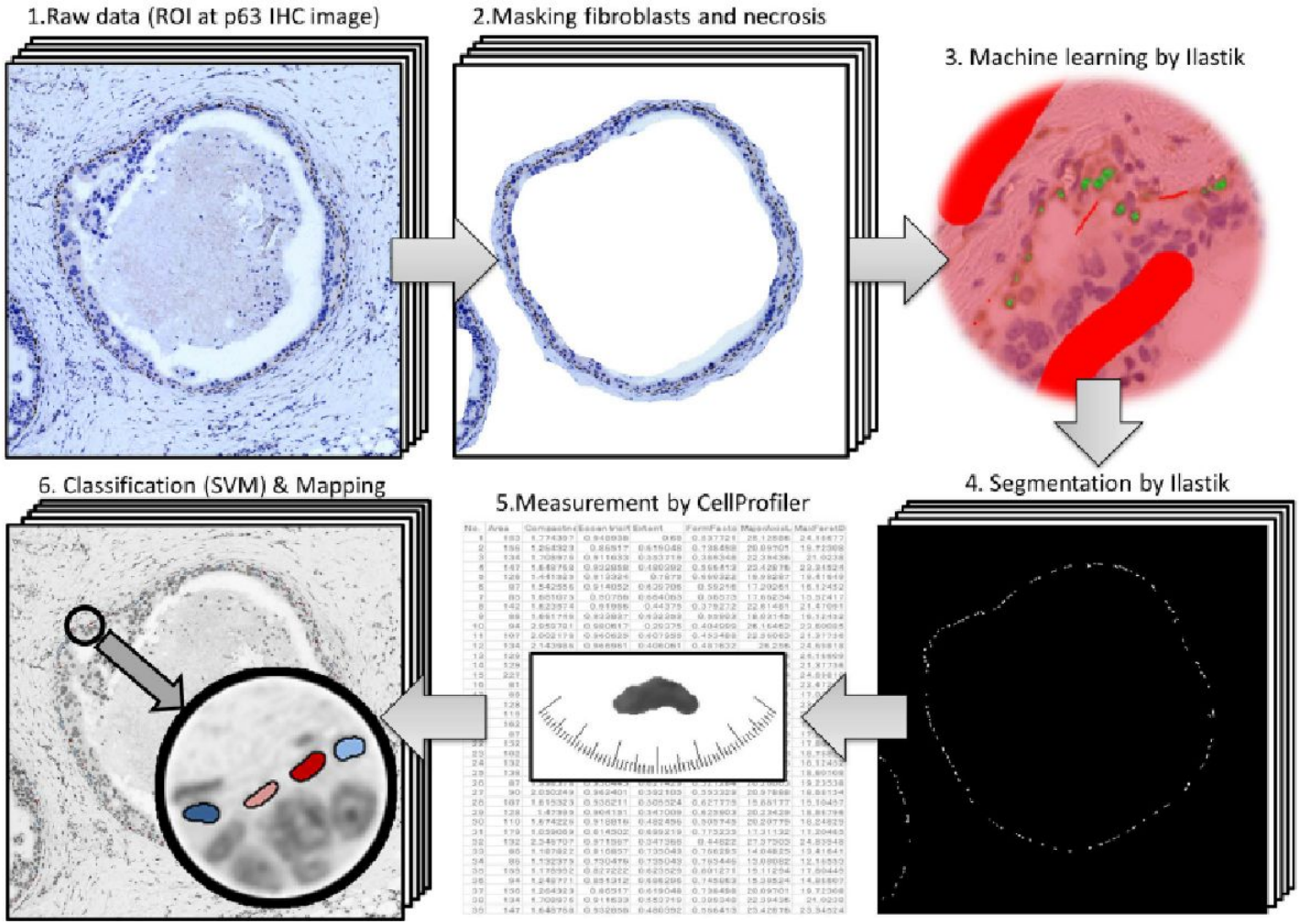
Movies



3D

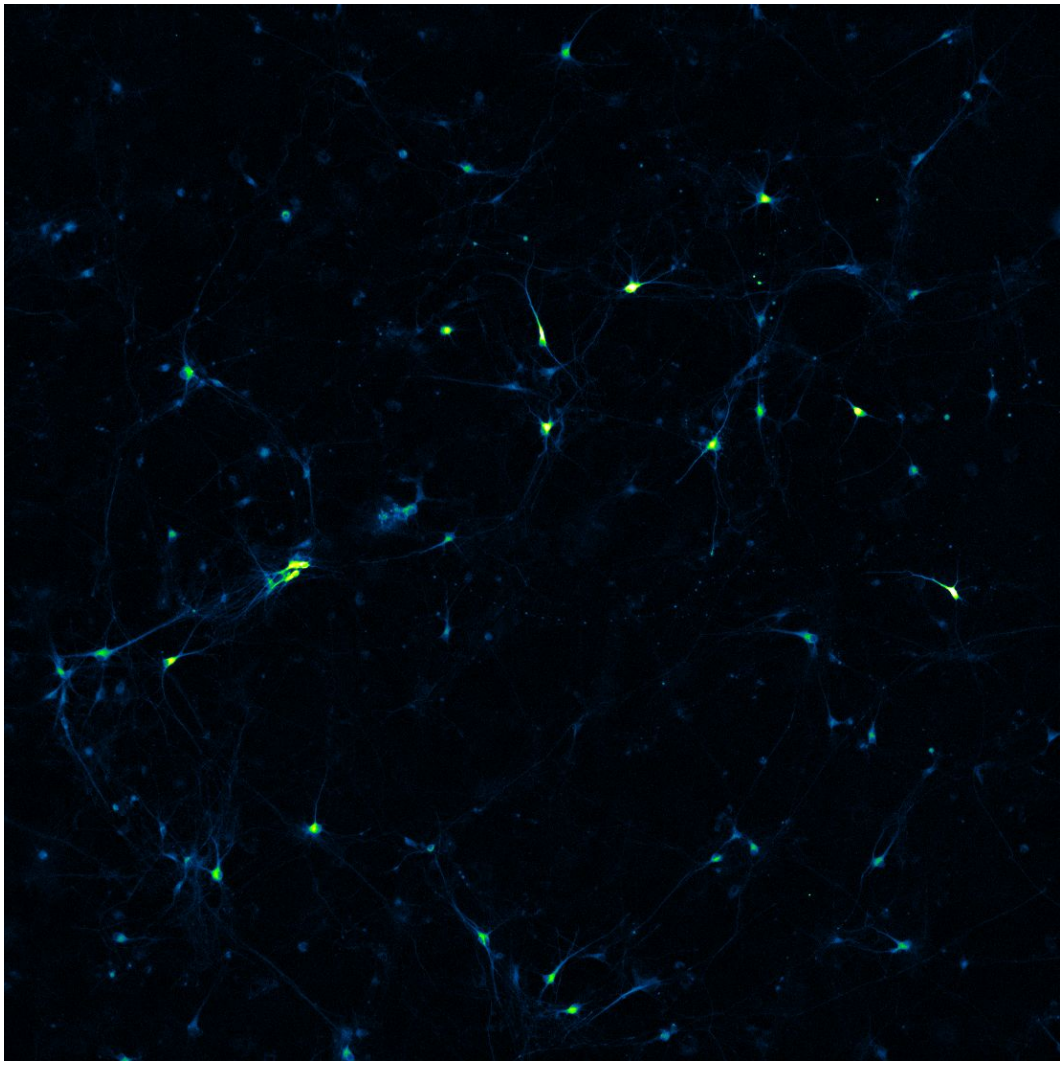


Example: Tissue analysis



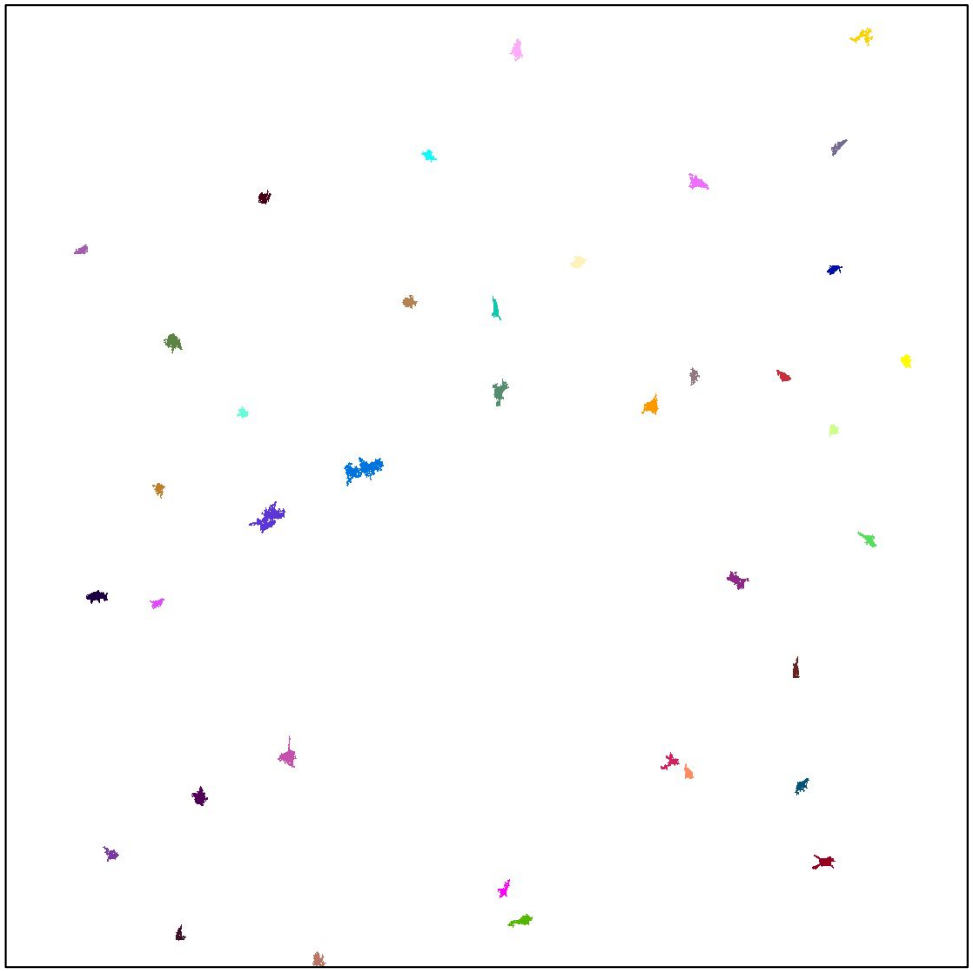
Yamamoto et al, Sci Rep 2017

Example: GCAMP fluorescence of motor neurons



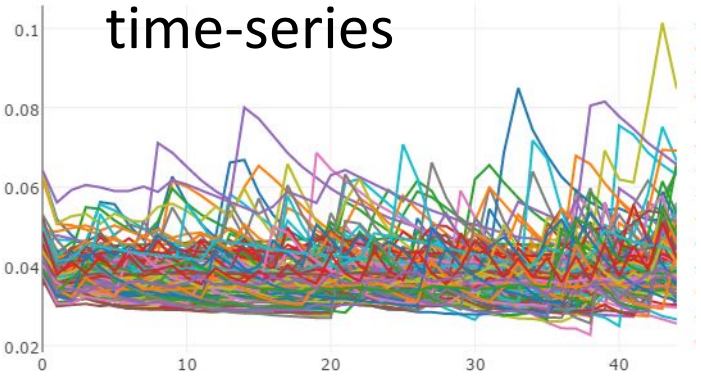
CellProfiler segments the max projection

- A CellProfiler pipeline finds regions of high intensity and segments these into

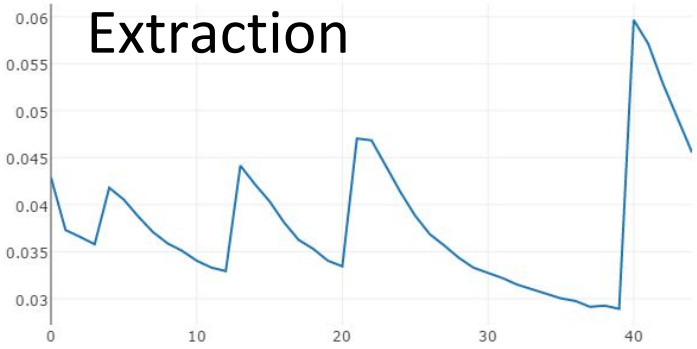


Quantify neuron spiking time-series

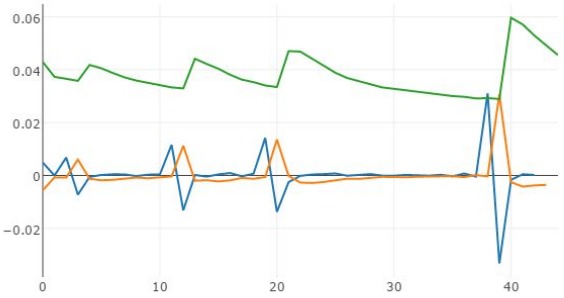
3. Correct time-series



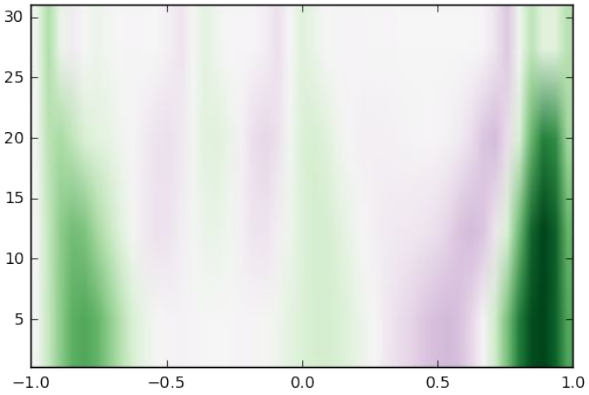
4. Feature Extraction



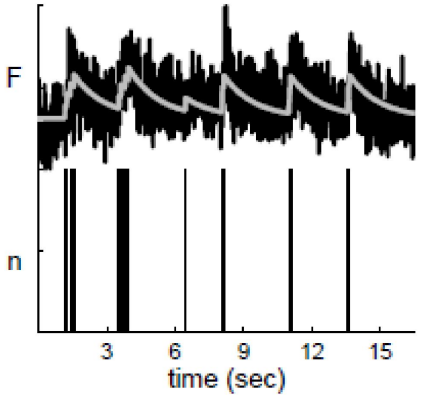
4.a Roet Peaks



4.b Wavelet Peaks

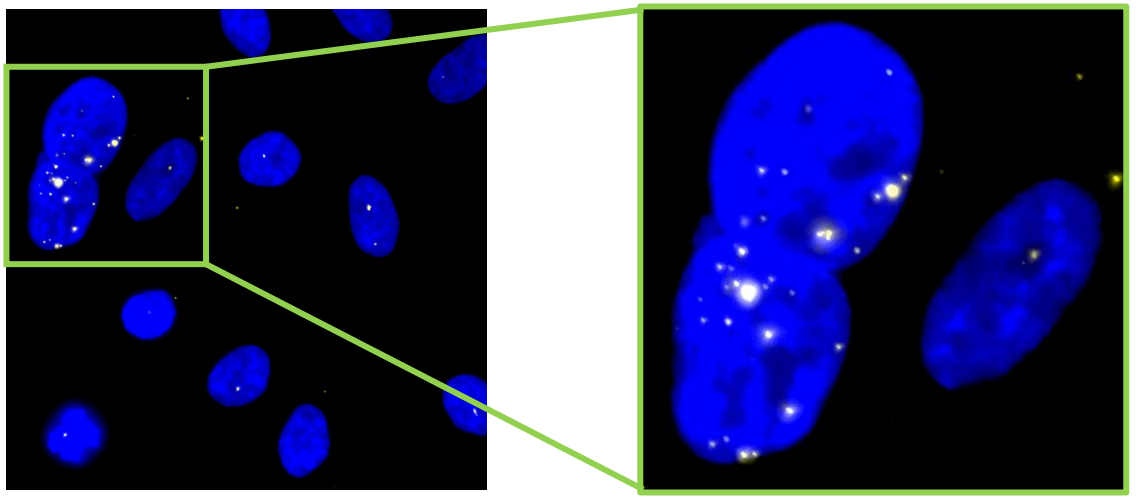


4.c oopsi spikes



mRNA ISH with cultured cells

- U2OS cells

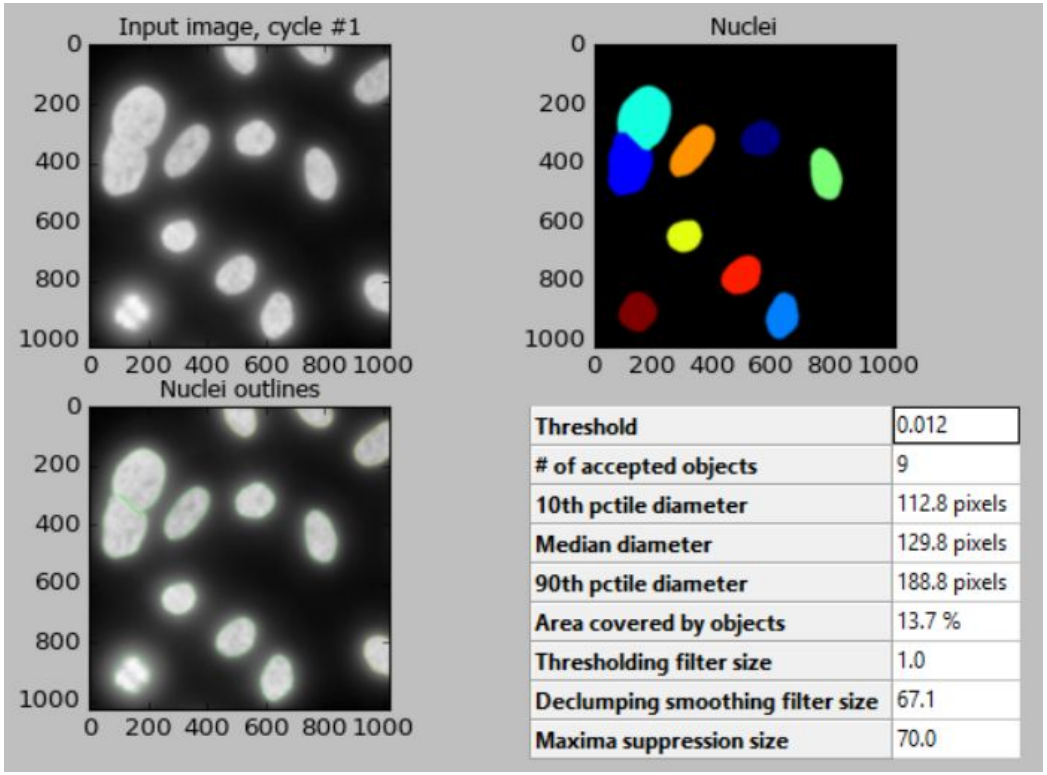


- Widefield microscope
- Z-stacks, 2-channels: nuclei and mRNA
- ISH Assay

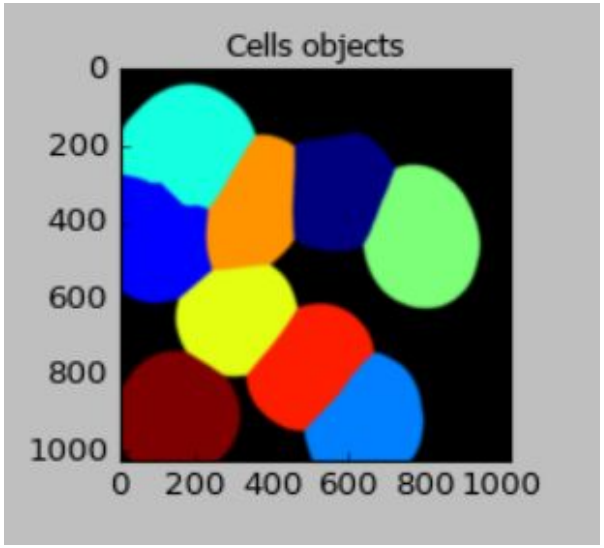
Leah Bury,
Cheeseman Lab,
Whitehead
Institute

mRNA ISH with cultured cells

- Max projections of nuclei and mRNA are inputs into CellProfiler
- Nuclei are segmented. The region around the nuclei are expanded to find mRNA near each nuclei

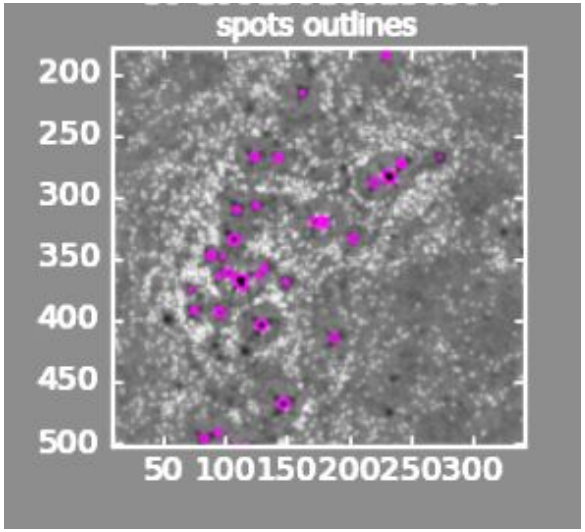


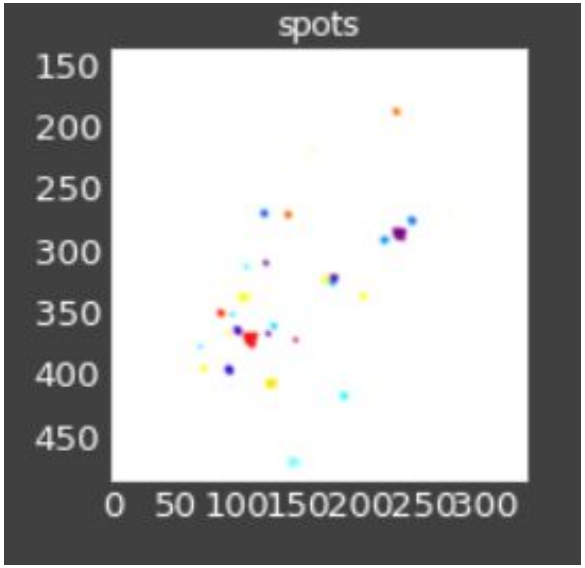
Threshold	0.012
# of accepted objects	9
10th pctile diameter	112.8 pixels
Median diameter	129.8 pixels
90th pctile diameter	188.8 pixels
Area covered by objects	13.7 %
Thresholding filter size	1.0
Declumping smoothing filter size	67.1
Maxima suppression size	70.0

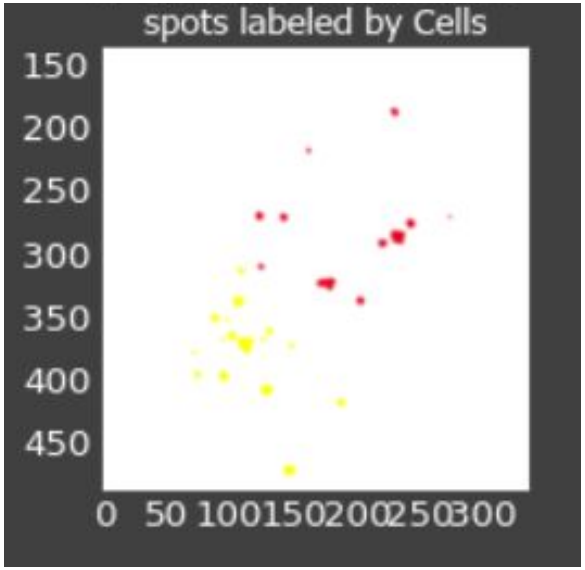


mRNA ISH with cultured cells

- Background signal is subtracted from mRNA channel and mRNA “spots” are enhanced
- Each mRNA is segmented
- mRNA is assigned to each cell by proximity to the nucleus

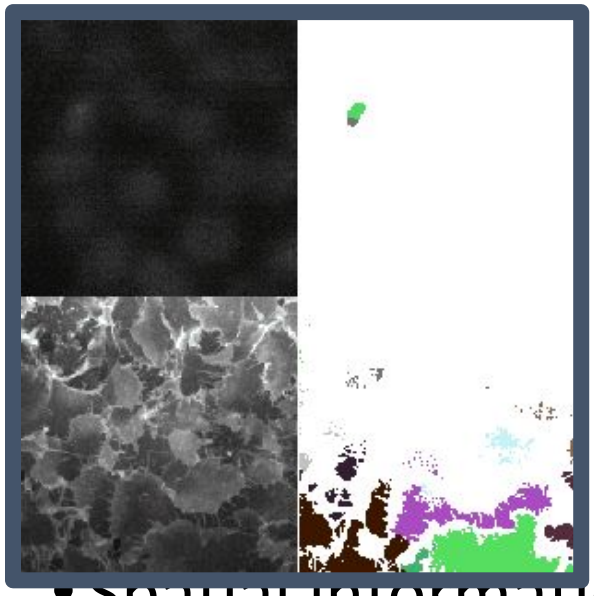


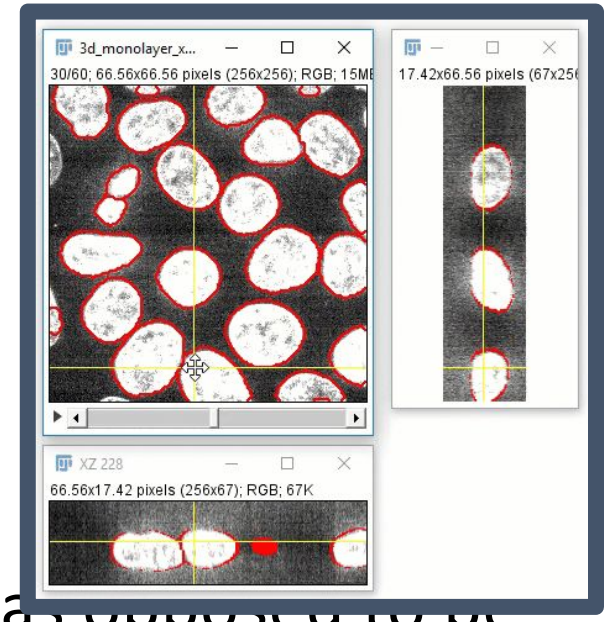




CellProfiler and 3d images

- The *Summer 2017* update of CellProfiler will include expanded 3D functionality



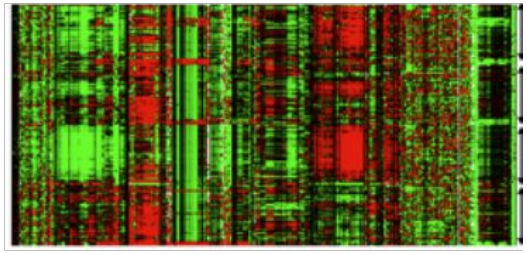


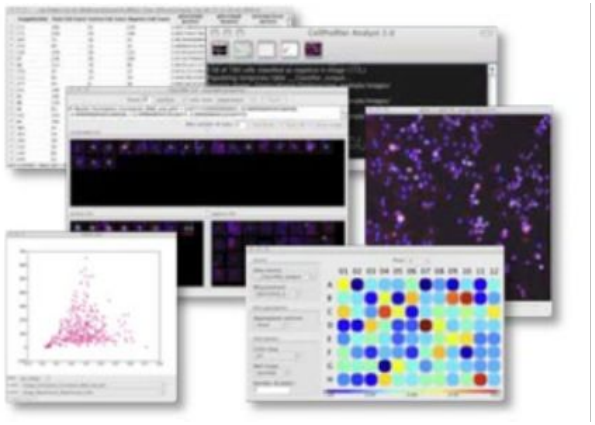
- Spatial information is preserved, as opposed to be removed through a max projection

Liya Ding,
Allen Institute

CellProfiler™ Analyst
data exploration software

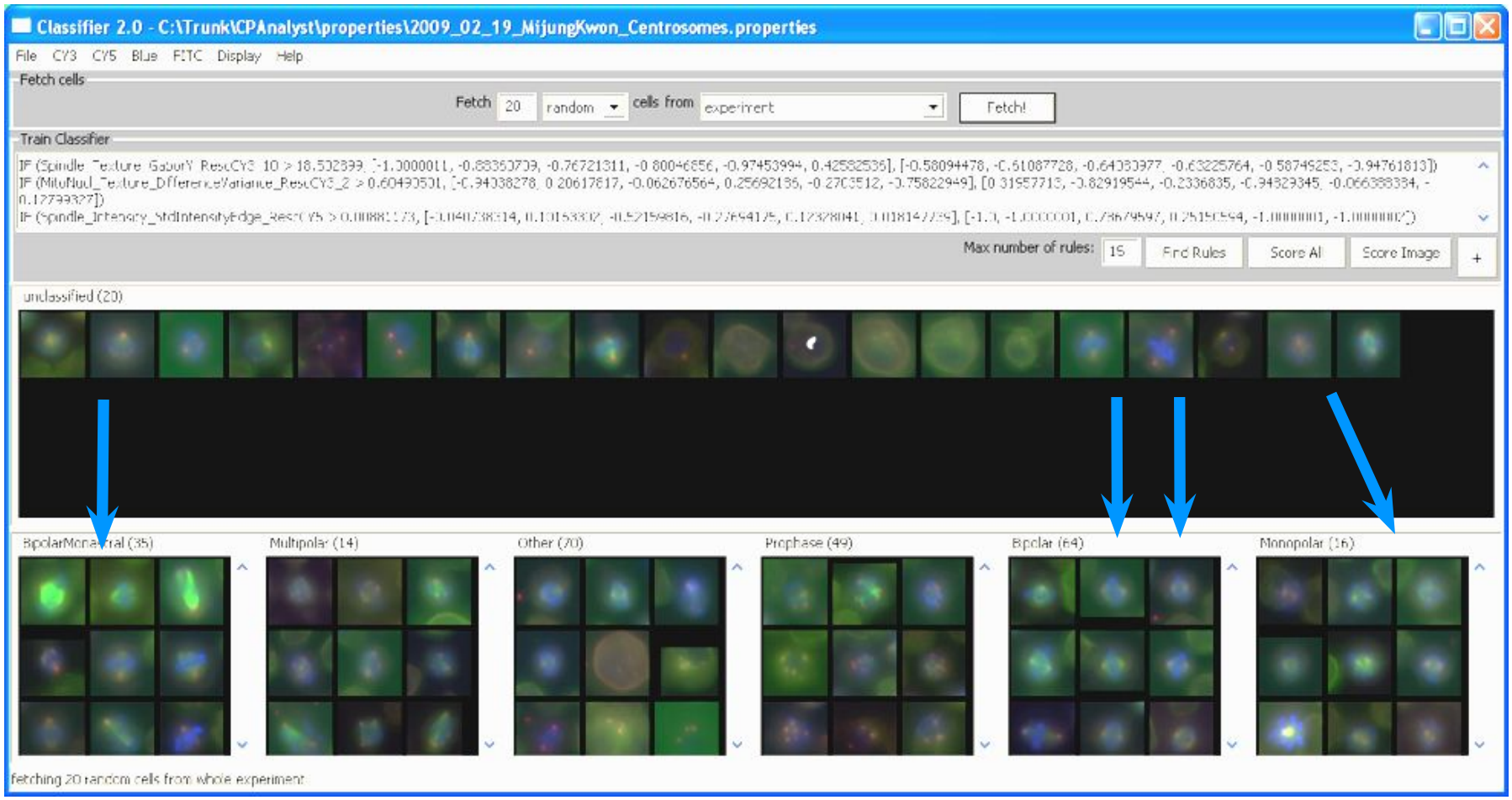
- **Explore** large, image-derived data sets
- **Score phenotypes** with machine learning





- **Goal:** Support data analysis and exploration based on individual cells

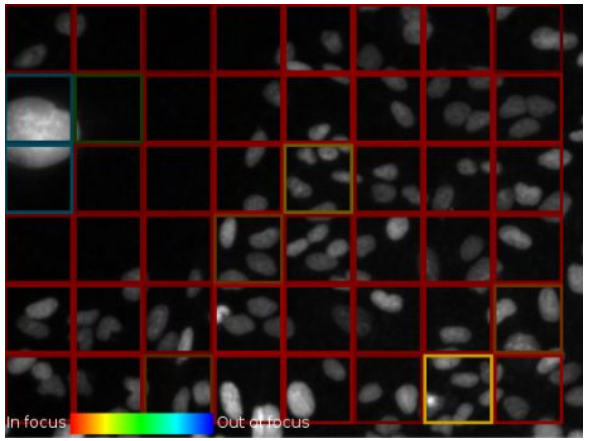
Phenotyping / Classification / Machine learning

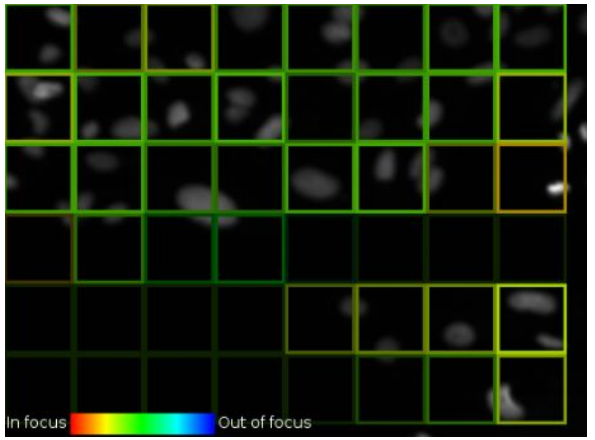


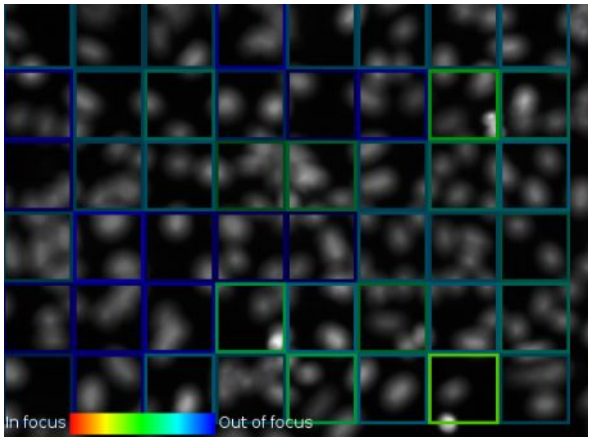
CellProfiler™ Analyst
data exploration software

Future Directions

- Distributed CellProfiler: a project for processing images with CellProfiler on cloud services such as AWS or Google Cloud.
- Deep Learning Modules: pre-trained DNN models can be added to CellProfiler pipelines
  - **MeasureImageQuailty** will determine what parts of the image are in-focus.





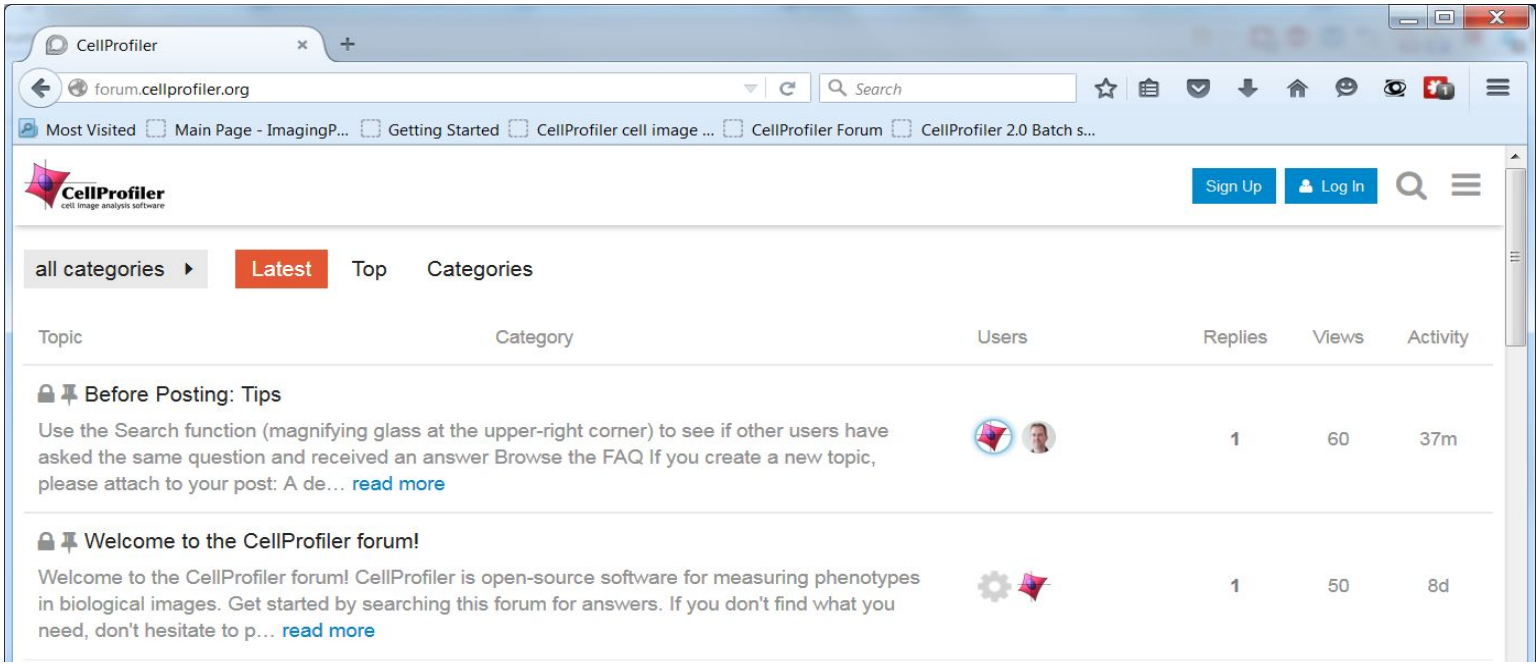


Conclusion

1. CellProfiler can quantify images from **diverse** biological contexts in an automated workflow
2. CellProfiler can analyze complex image sets including **3D**, **time-series**, **tissue** and **high-content screening**.
3. CellProfiler is **open-source** and actively developed, so it will grow with the data as image sets become larger and more complex and incorporate the latest methods

CellProfiler help is found online

- Documentation and examples are at cellprofiler.org
- Direct help from the Broad Institute and other users can be found at the forum.cellprofiler.org



Thank you!

The Anne Carpenter Lab and Imaging Platform