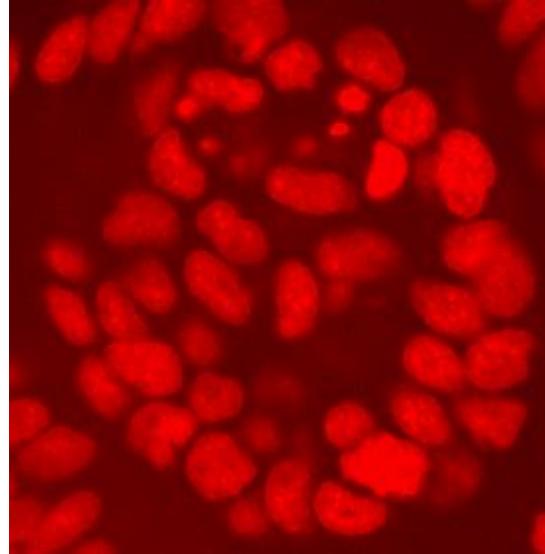


# Bio-image analysis, bio-statistics, programming and machine learning for computational biology

Robert Haase & Anna Poetsch

# Lecture overview: Bio-image analysis

- Quantitative bio-image analysis
  - Transformations, Projections and Filtering
  - Image segmentation
  - Tracking objects
  - Measuring in images
- Goal: Allow you to quantify observations.
- Software
  - ImageJ / Fiji for manual image analysis
  - Python for automation

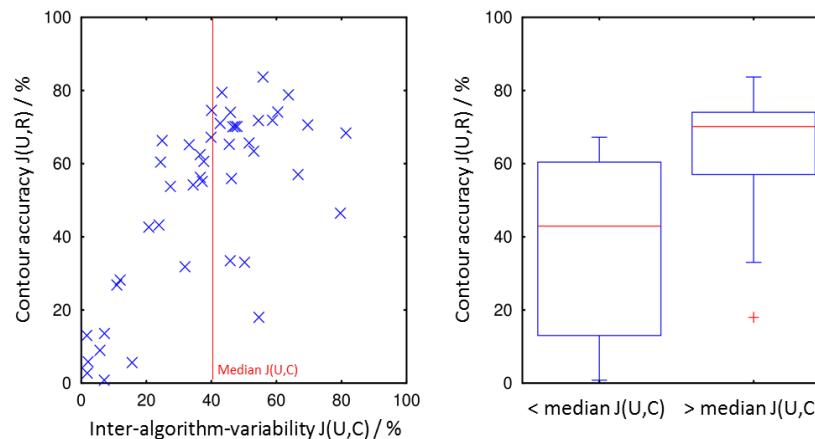
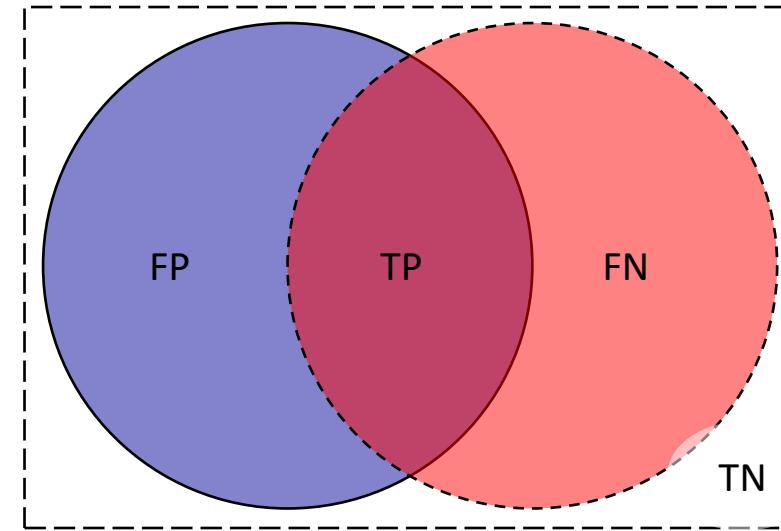


Results							
	Area	Perim.	Circ.	AR	Round	Solidity	
1	2644	182.21	1.00	1.07	0.93	1.00	
2	2680	183.78	1.00	1.09	0.92	1.00	
3	3176	199.49	1.00	1.02	0.98	1.00	
4	2733	185.35	1.00	1.07	0.93	1.00	
5	3804	218.34	1.00	1.05	0.96	1.00	
6	4188	229.34	1.00	1.09	0.92	1.00	
7	2832	188.50	1.00	1.06	0.94	1.00	
8	2970	193.21	1.00	1.12	0.89	1.00	

# Lecture overview: Bio-statistics



- Bio-statistics
  - Inferential statistics
  - Descriptive statistics
- Inter-/intra-observer-variability
- Method comparison: human versus algorithm
- Accuracy, precision, recall
- **Goal: Allow you to draw conclusions from quantified observations.**
- Dr. Anna Poetsch will teach you this part.

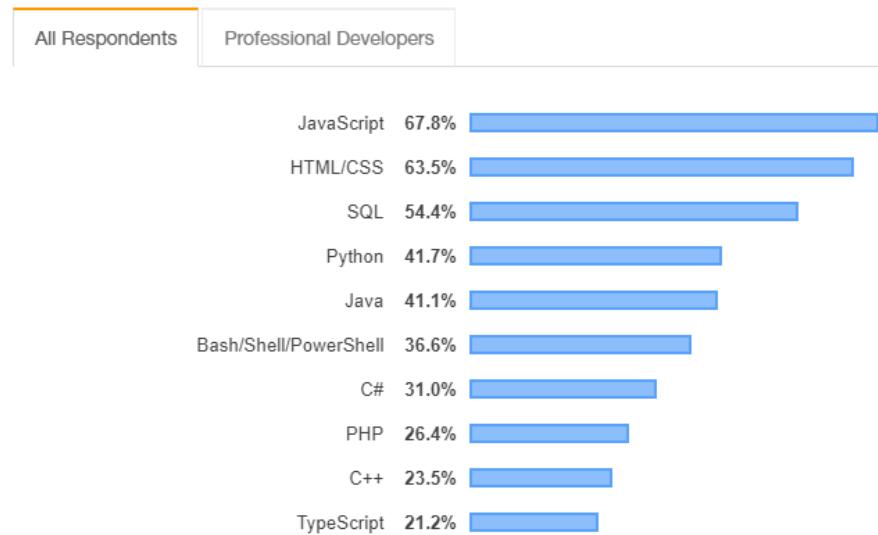


# Lecture overview: Programming

- We will focus on Python programming.
- “Python is the most wanted language for the third year in a row, meaning that developers who do not yet use it say they want to learn it.”
- Goal: **Allow you to do things automatically instead of suffering long time when doing it manually.**

## Most Popular Technologies

### Programming, Scripting, and Markup Languages



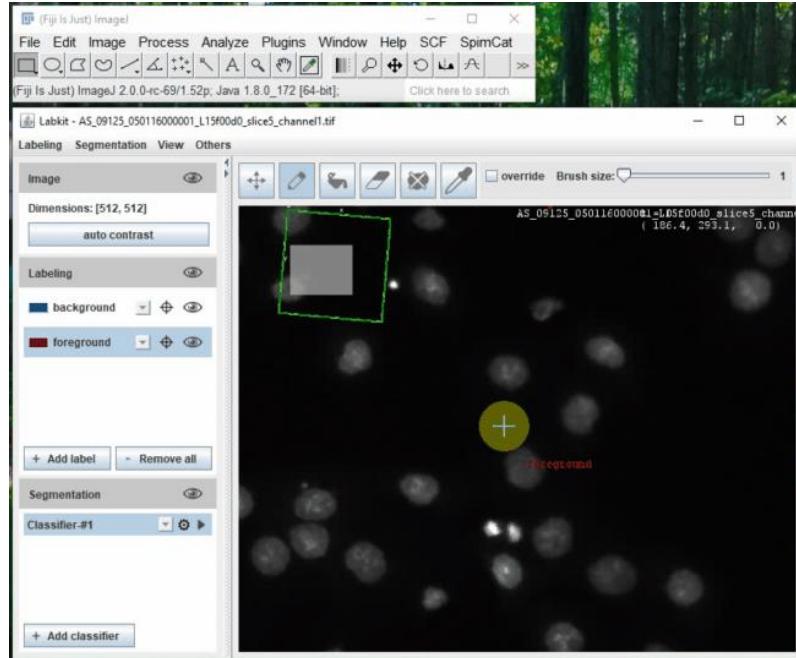
## Most Loved, Dreaded, and Wanted

### Most Loved, Dreaded, and Wanted Languages

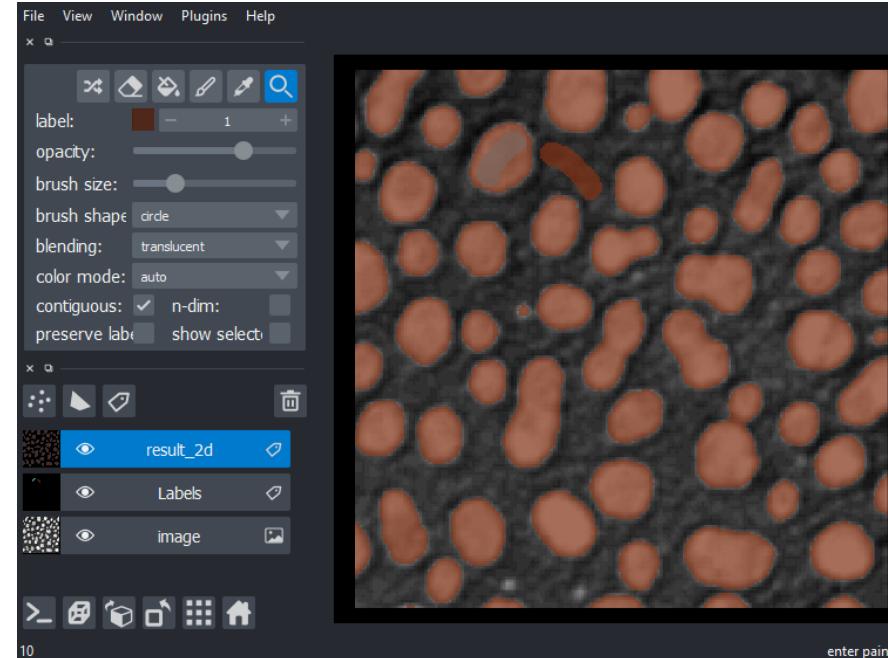


# Lecture overview: Machine learning

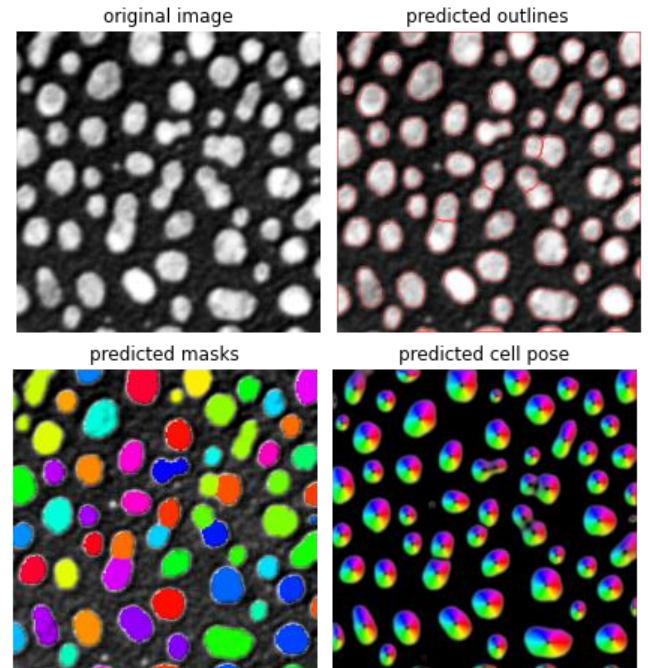
- Machine learning
  - in the context of image analysis
- Computers can *learn* where, what objects in images are...
- Goal: **Give you an insight into state-of-the-art methods and future developments**



Pixel classification (Labkit)



Pixel / object classification (Python / napari)



Nuclei segmentation with  
StarDist and CellPose

- 13.4.2021 - General introduction, setting up Fiji and Python
- 20.4.2021 - Introduction to python programming and data structures
- 27.4.2021 - Introduction to python algorithms: loops, conditions and functions
- 4.5.2021 - Image processing and filtering
- 11.5.2021 - Image segmentation + 3D Image Processing
- 18.5.2021 - Quantitative measurements, Processing folders
- 1.6.2021 - Introduction to Biostatistics
- 8.6.2021 - Descriptive statistics
- 15.6.2021 - Method Comparison - Bland-Altman analysis
- 22.6.2021 - Hypothesis testing
- 29.6.2021 - Big data and data visualization
- 6.7.2021 - Machine Learning I
- 13.7.2021 - Machine Learning II
- 20.7.2021 - Summary

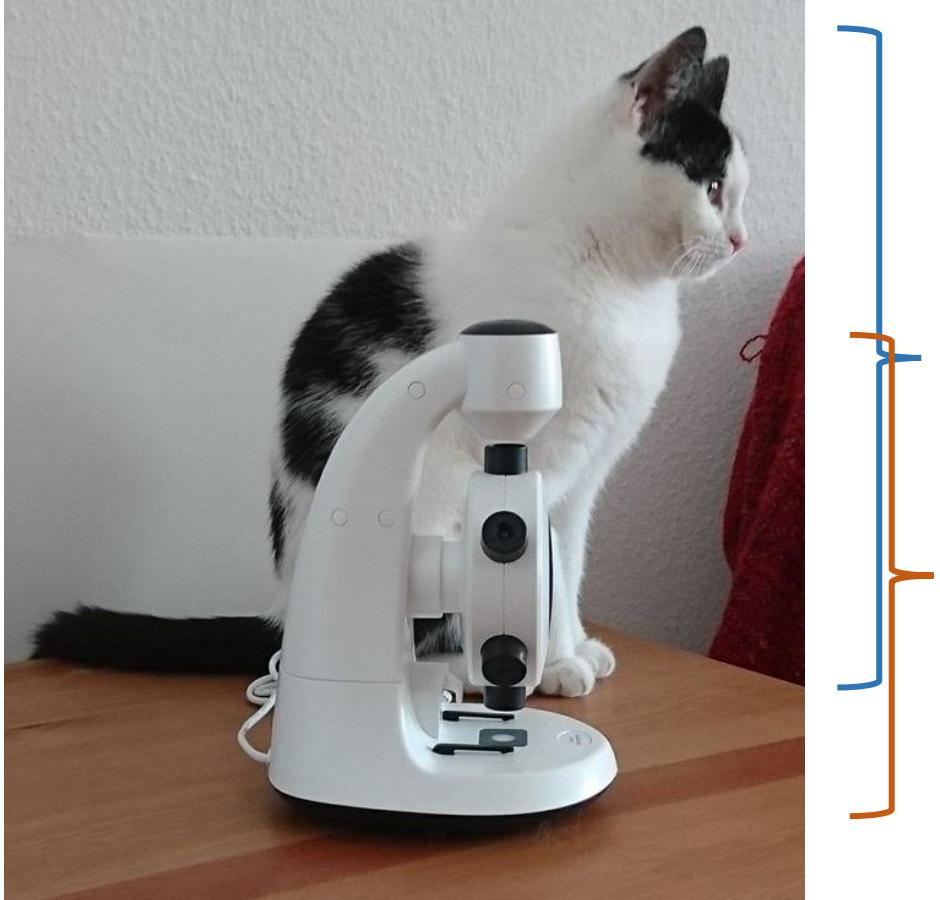
Anna Poetzsch  
Robert Haase

- Every week will follow the same rough scheme
  - < 1 h lecture
  - < 1h demonstration
  - < 1h homework
- Questions can be asked in the zoom meetings any time by speaking up. Or you post them anonymously: (see link to etherpad in email)
- Exam will cover the semester content accordingly
  - Theory of
    - image analysis,
    - statistics and
    - machine learning
  - Basics of programming
    - write simple < 10 line programs and
    - read code and describe what it does

# Bio-Image Analysis

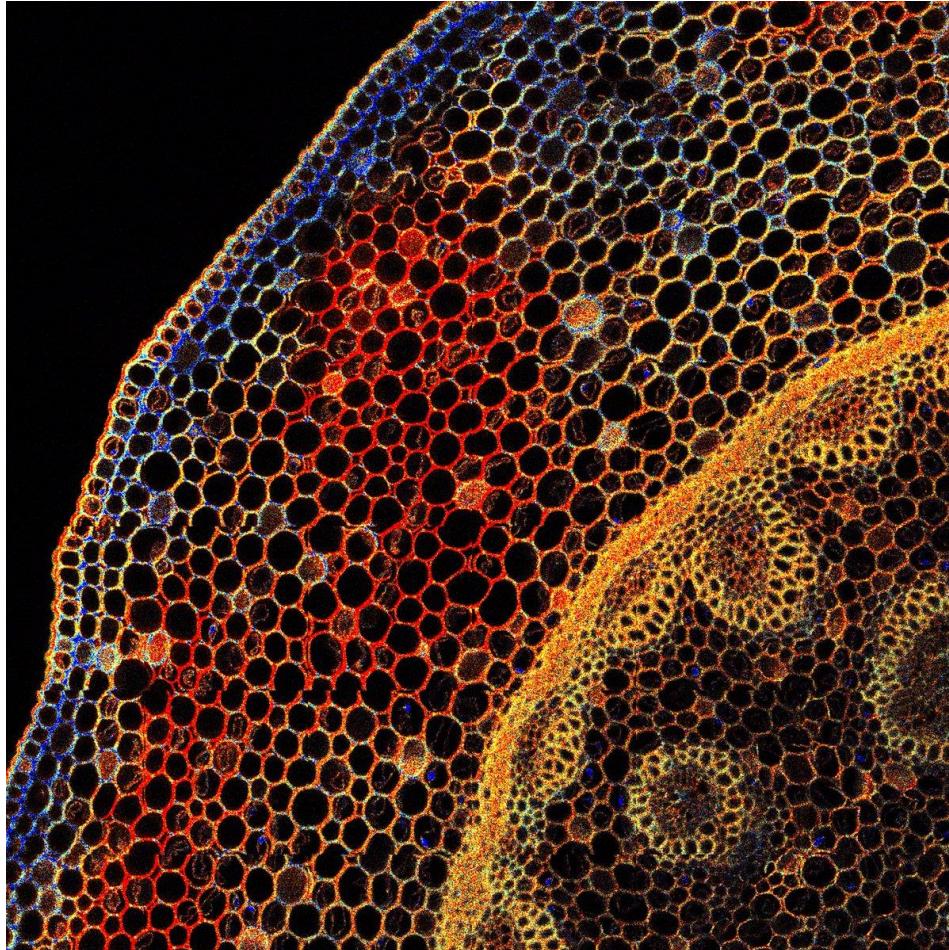
Robert Haase

- Deriving quantitative information from images of biological samples taken with microscopes



cat height = 1.5 x microscope height

- Deriving quantitative information from images of biological samples taken with microscopes



How many cells are there in the outer ring?

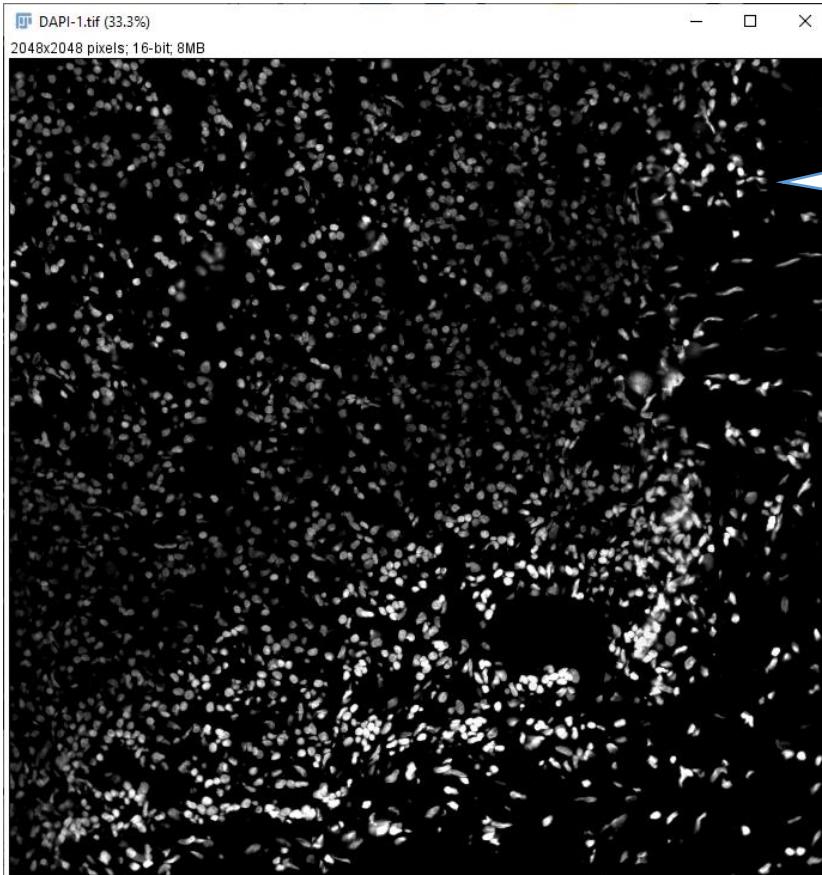
How high is the intensity in the inner ring?

How is the signal in the blue channel distributed?

Image data source: Lorenzo Scipioni, U. California / Irvine

# Objective bio-image analysis

- Measurements should be objective, not influenced by human interpretation



Nuclei in this image are ...

... more dense than in this image.

Use automation for less subjective analysis.

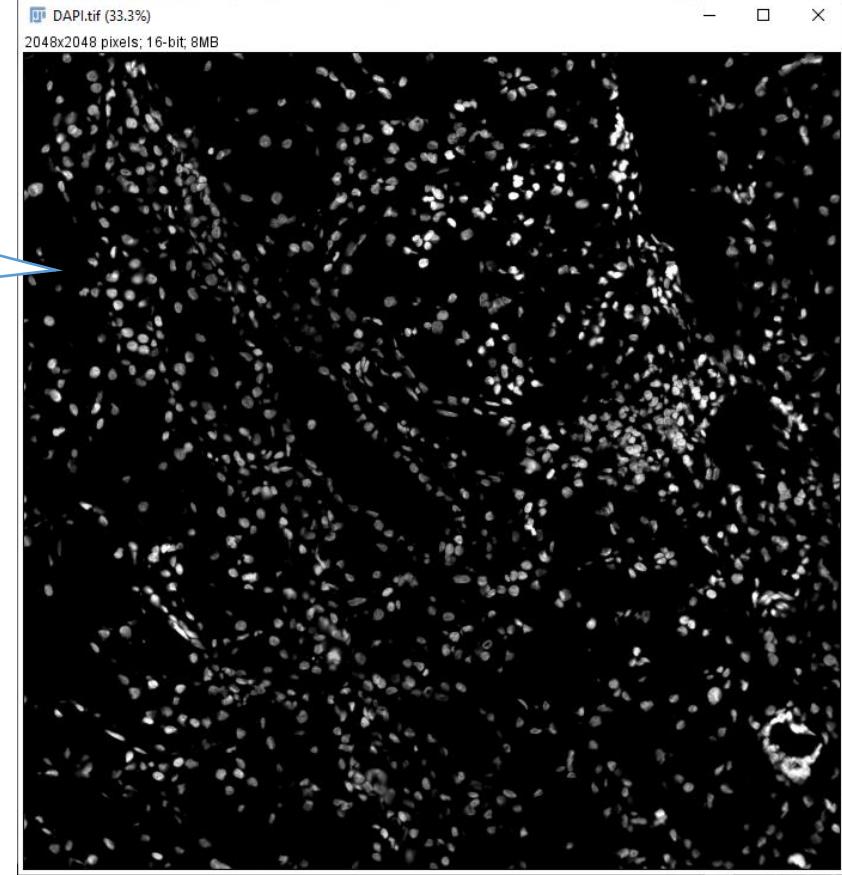
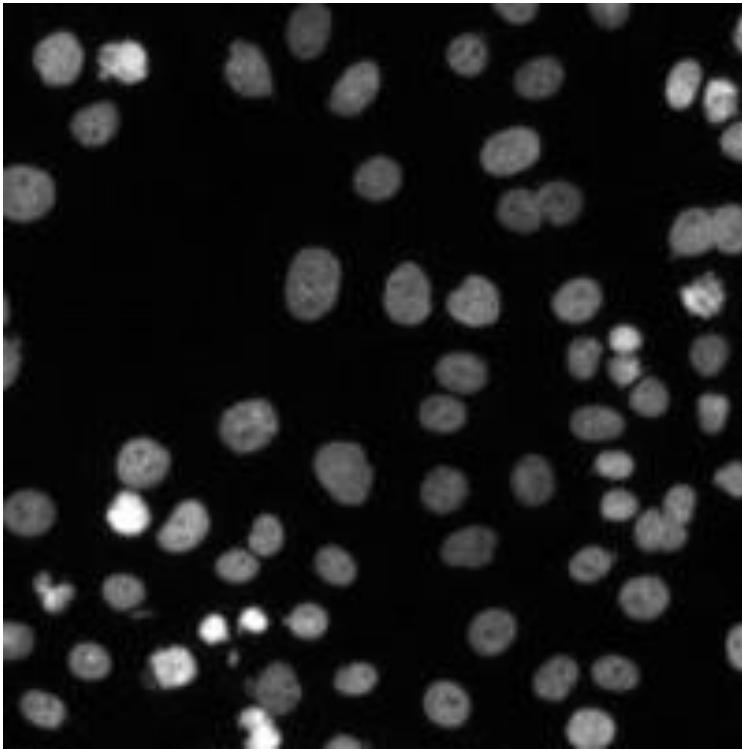


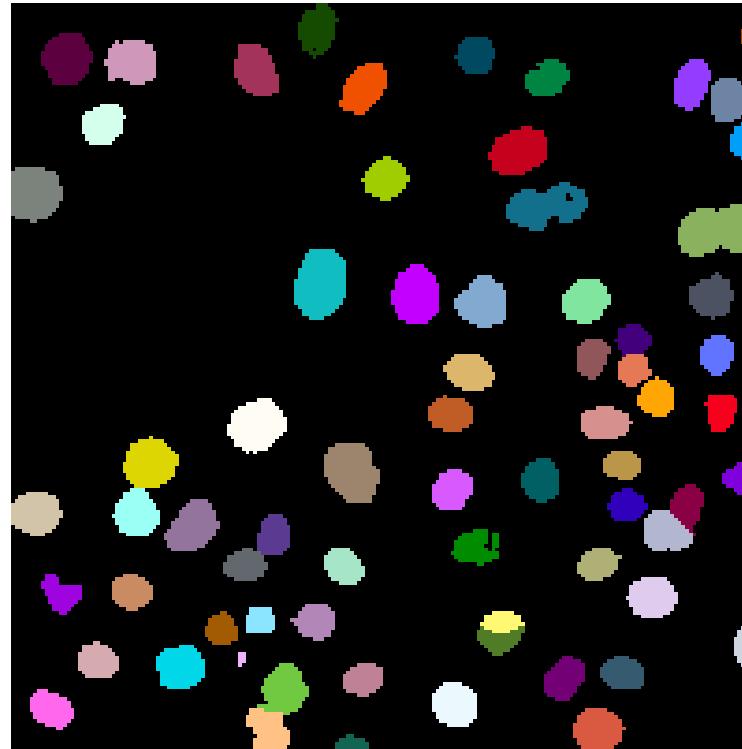
Image data source: Pascual-Reguant, Anna. (2021). Immunofluorescence staining of a human kidney (#2, peri-tumor area) obtained by MELC [Data set]. Zenodo. <http://doi.org/10.5281/zenodo.4434462> licensed CC-BY 4.0

- Algorithms must be reliable (trustworth). Visualization helps gaining trust in automated methods.

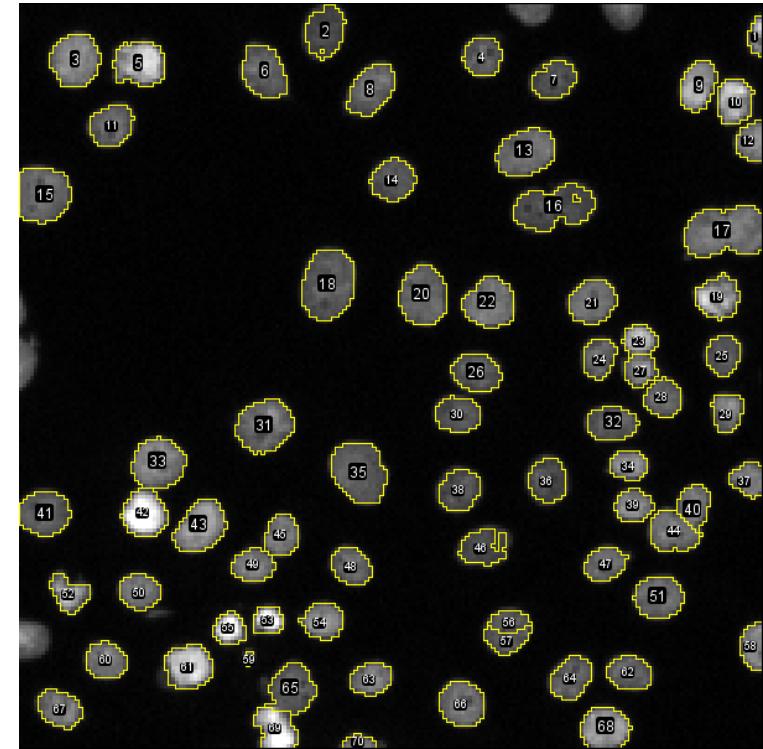
Original image



Label image



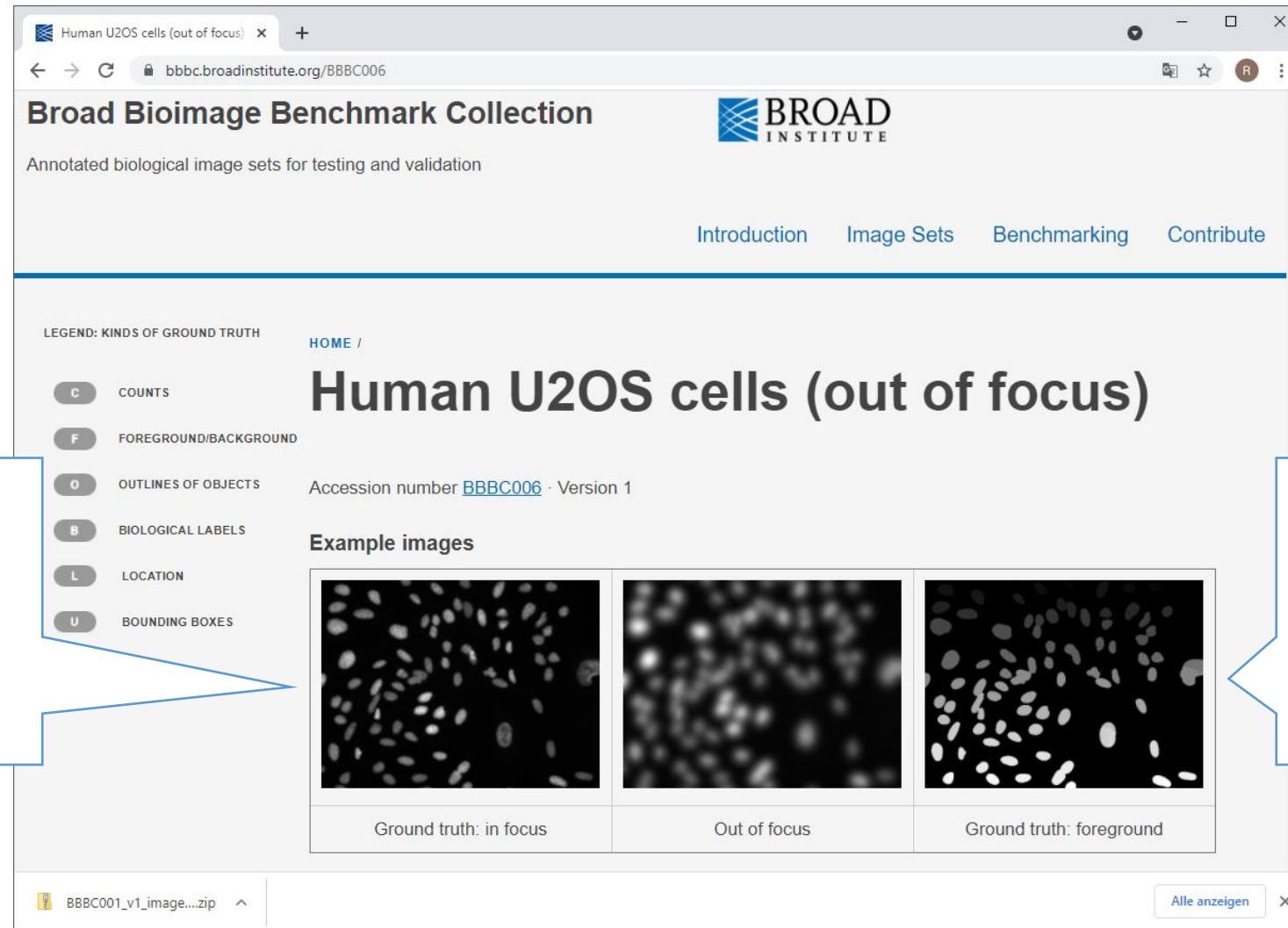
Overlay



There are 70 nuclei  
in this image.

# Reliable bio-image analysis

- Algorithms must be reliable (validate methods). Publicly available benchmark data sets allow to compare algorithms on common data.



Original image data

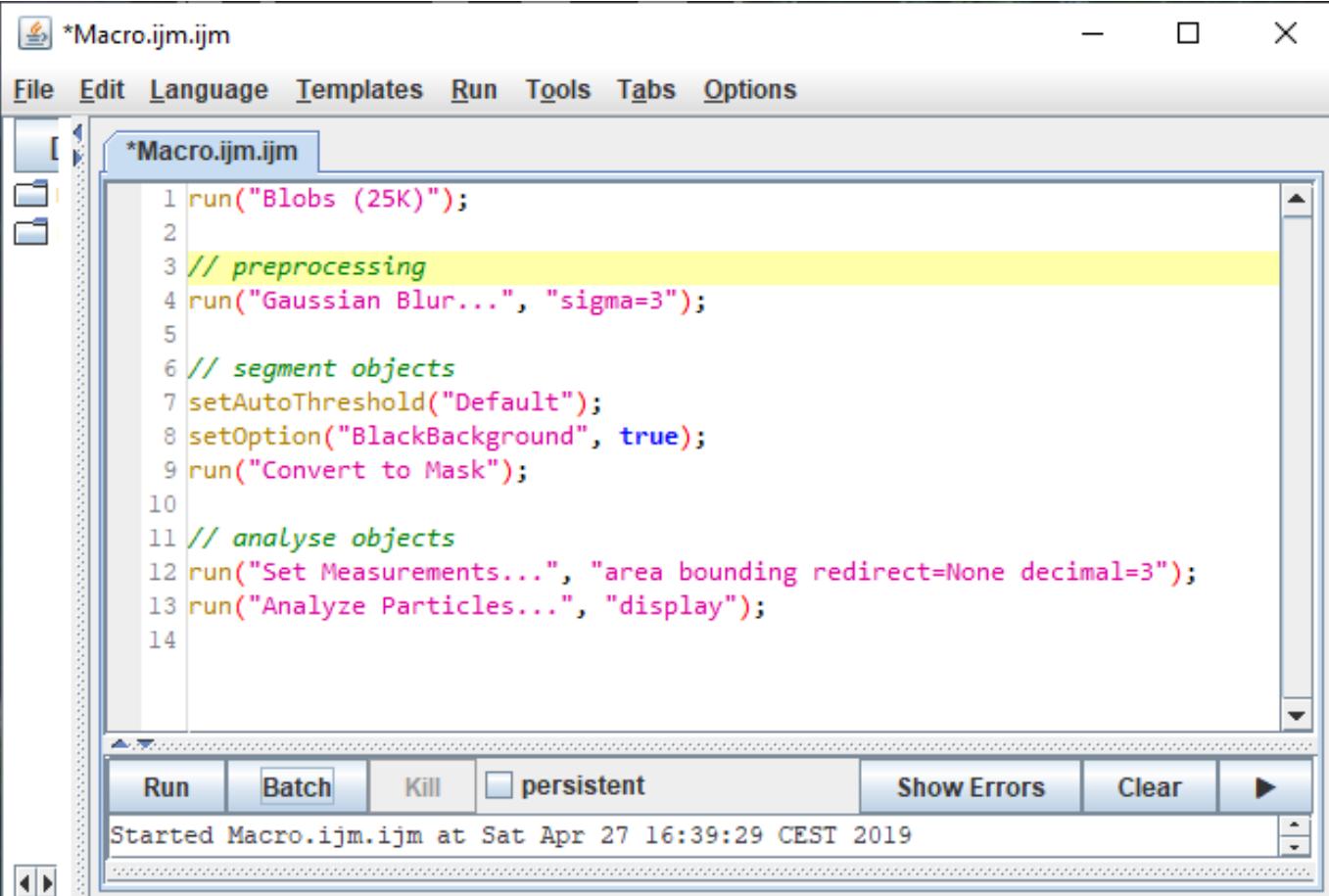
"Ground truth" label images

# Reproducible bio-image analysis

- “The image data was analyzed with ImageJ.”

Can you reproduce what they did?

Can you reproduce what they did?



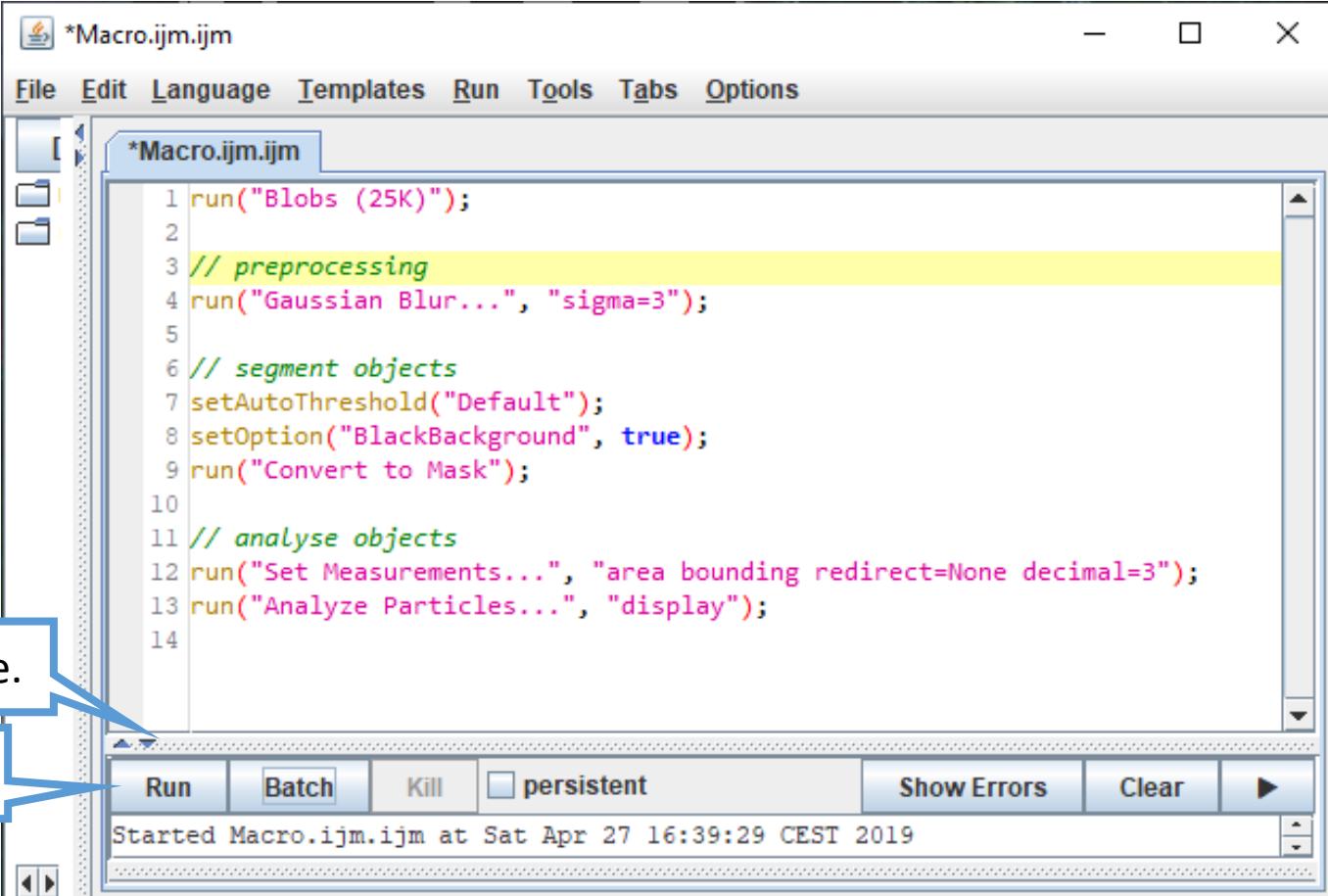
```
*Macro.ijm.ijm
File Edit Language Templates Run Tools Tabs Options
I *Macro.ijm.ijm
1 run("Blobs (25K)");
2
3 // preprocessing
4 run("Gaussian Blur...", "sigma=3");
5
6 // segment objects
7 setAutoThreshold("Default");
8 setOption("BlackBackground", true);
9 run("Convert to Mask");
10
11 // analyse objects
12 run("Set Measurements...", "area bounding redirect=None decimal=3");
13 run("Analyze Particles...", "display");
14

Run Batch Kill persistent Show Errors Clear ►
Started Macro.ijm.ijm at Sat Apr 27 16:39:29 CEST 2019
```

# Repeatable Bio-image analysis

- Compared to wet-lab experiments, image analysis (in-silico) experiments are typically repeatable.
- In wet-lab experiments, samples may get destroyed while executing the experiment.
- Repeatability is a property of the experiment. You cannot improve repeatability by better documentation.

- However, you need to pay some extra attention to have repeatable image processing
  - Scripts need to be written in a way that their execution is repeatable.
  - Test it!



```
*Macro.ijm.ijm
File Edit Language Templates Run Tools Tabs Options
I *Macro.ijm.ijm
1 run("Blobs (25K)");
2
3 // preprocessing
4 run("Gaussian Blur...", "sigma=3");
5
6 // segment objects
7 setAutoThreshold("Default");
8 setOption("BlackBackground", true);
9 run("Convert to Mask");
10
11 // analyse objects
12 run("Set Measurements...", "area bounding redirect=None decimal=3");
13 run("Analyze Particles...", "display");
14
```

Run it once.

Run it twice.

Are results identical?

# Introduction to bio-image analysis

- Bio-image analysis is supposed to be
  - **Quantitative**
    - We derive numbers from images which describe physical properties of the observed sample.
  - **Objective**
    - The derived measurement does not depend on who did the measurement. The measurement is free of interpretation.
  - **Reliable (trustworthy / validated)**
    - We are confident that the measurement is describing what it is supposed to describe.
  - **Reproducible**
    - Somebody else can do the experiment under *different conditions* and gets similar measurements. For this, documentation is decisive!
  - **Repeatable**
    - We can do the same experiment twice under the *same conditions* and get similar measurements.

# Image analysis is part of the experiment



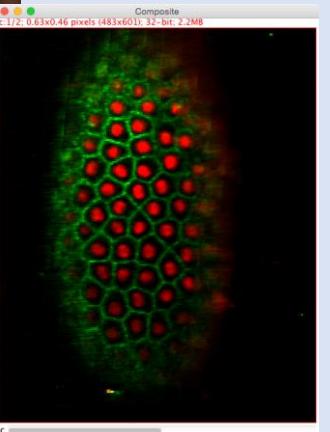
**PoL**  
Physics of Life  
TU Dresden



Observation

$$p_{ij}(t) = \frac{[\tau_{ij}(t)]^\alpha \cdot [\eta_{ij}]^\beta}{\sum_{j=1}^n [\tau_{ij}(t)]^\alpha \cdot [\eta_{ij}]^\beta}$$

Modeling



Imaging

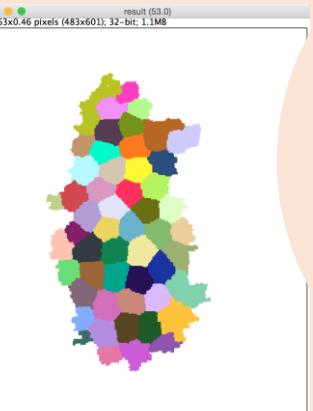


Image processing

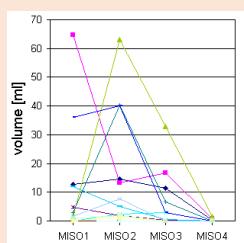
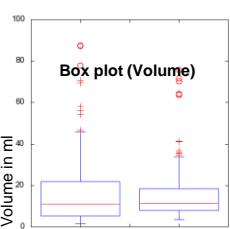


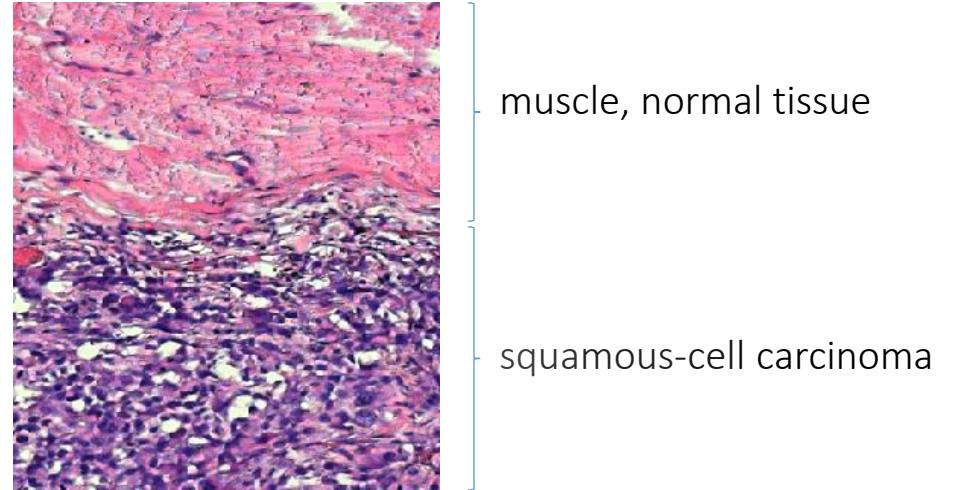
Image analysis  
Bio-statistics



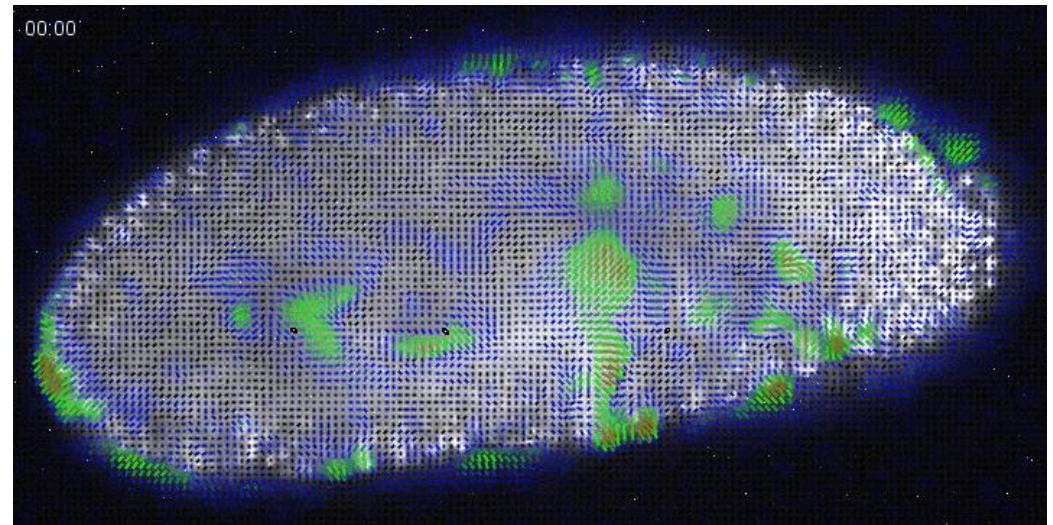
April 2021

# Common questions

- Typical questions bio-image analysts deal with
  - Is signal intensity different under varying conditions?
  - How many cells are in my image?
  - How high is cell density?
    - Bio-statistics / medicine
  - How are different tissues characterized?
    - Machine learning



- Typical questions bio-image analysts struggle with
  - What force drives the observed processes?
  - What is the lineage tree of one particular cell?
  - Are observation A and observation B related?
  - Are structures observed in different color channels colocalized?



# Image analysis is part of the experiment

- Think about how to analyze your images before starting the experiment.
  - Consider adapting your experiment so that quantitative image analysis can be performed easily.
- Think about controls, counter-proves, an easy to falsify null-hypothesis.
  - Be a lazy scientist. Do simple experiments.
- How can you exclude yourself from the experiment?
  - Think of blinding yourself or fully automate analysis.
- One experiment usually answers just one or less questions.

# Introduction to Bio-image analysis: Terminology

Robert Haase

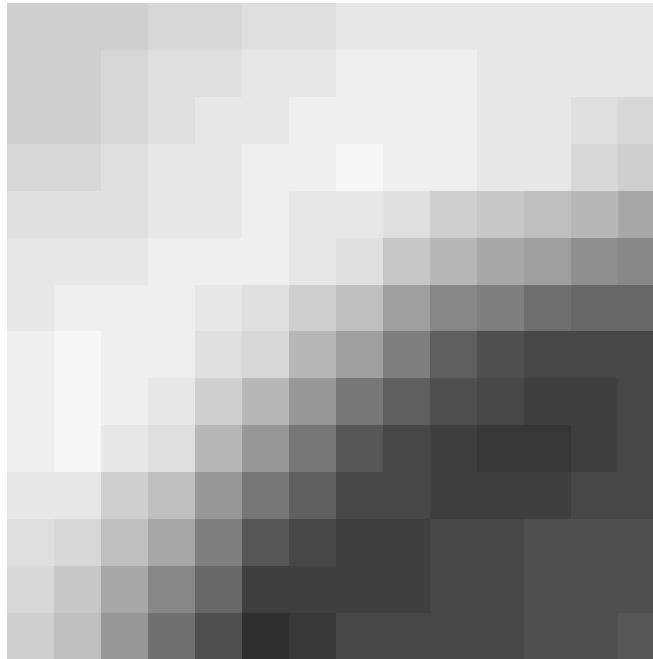
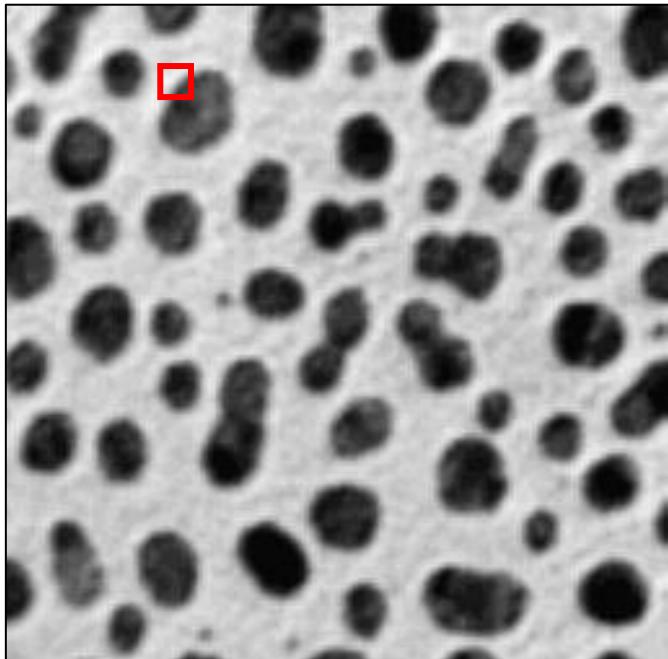
With material from

Benoit Lombardot, Scientific Computing Facility, MPI CBG

April 2021

# Images and pixels

- An image is just a matrix of numbers
- Pixel: “picture element”
- The edges between pixels are an artefact. In reality, they don’t exist!

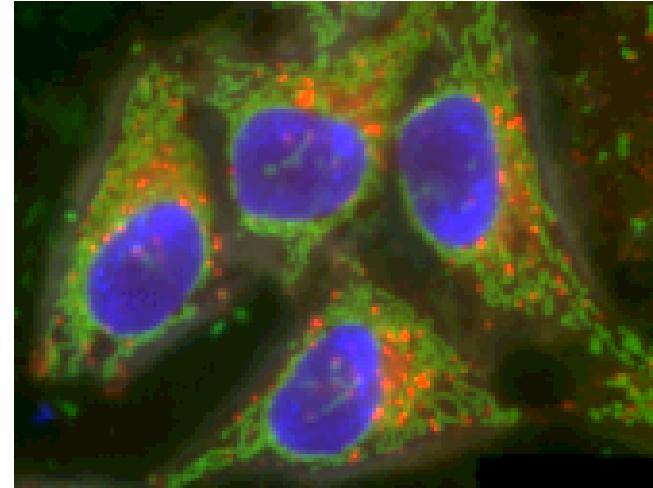
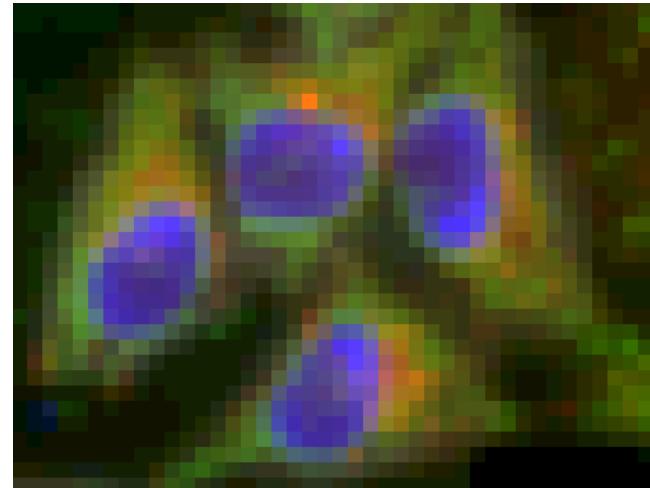
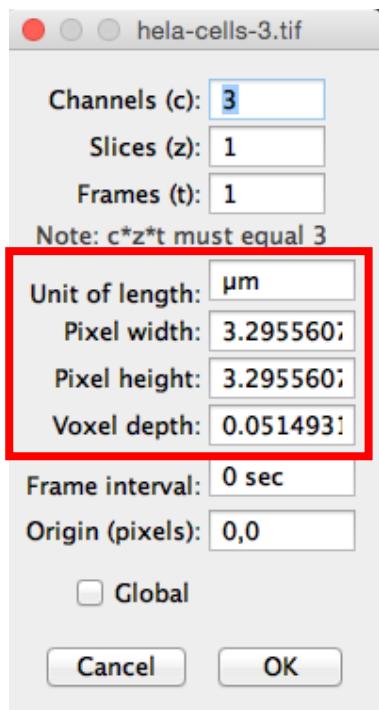


48	48	48	40	40	32	32	24	24	24	24	24	24	24	24
48	48	40	32	32	24	24	16	16	16	16	24	24	24	24
48	48	40	32	24	24	16	16	16	16	16	24	24	32	40
40	40	32	24	24	16	16	16	8	16	16	24	24	40	48
32	32	32	24	24	16	24	24	32	48	56	64	72	88	
24	24	24	16	16	16	24	32	56	72	88	96	112	120	
24	16	16	16	24	32	48	64	96	120	128	144	152	152	
16	8	16	16	32	40	72	96	128	160	176	184	184	184	
16	8	16	24	48	72	104	136	160	176	184	192	192	192	184
24	24	48	64	104	136	160	184	184	192	192	192	184	184	184
32	40	64	88	128	168	184	192	192	184	184	176	176	176	176
40	56	88	120	152	192	192	192	184	184	184	176	176	176	176
48	64	104	144	176	208	200	184	184	184	184	176	176	176	168



# Pixel size versus resolution

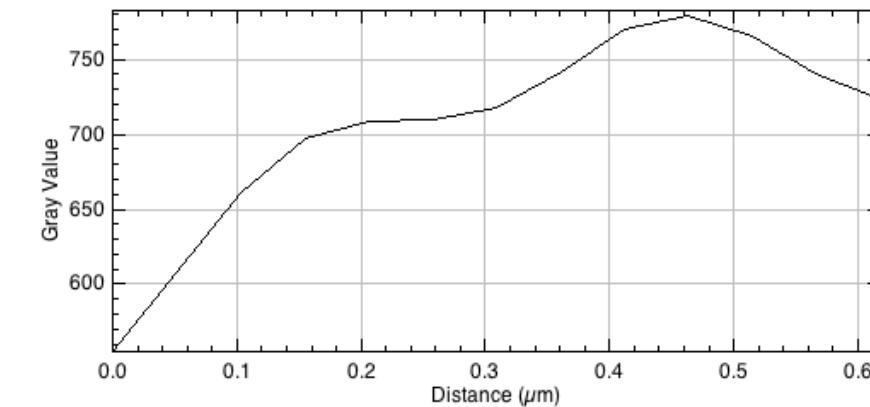
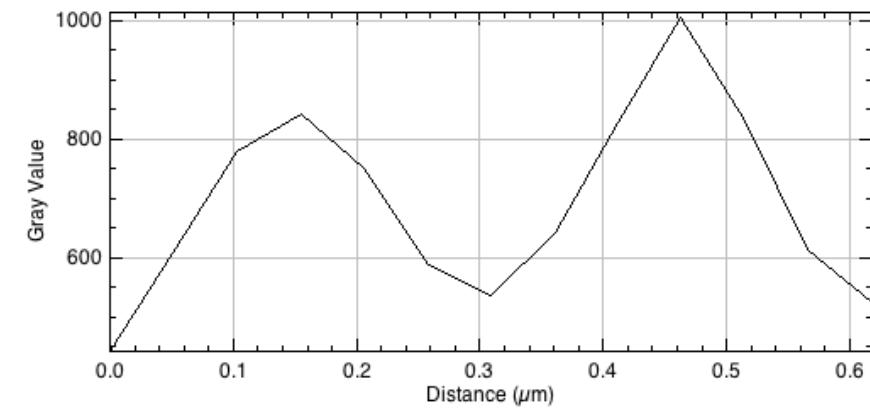
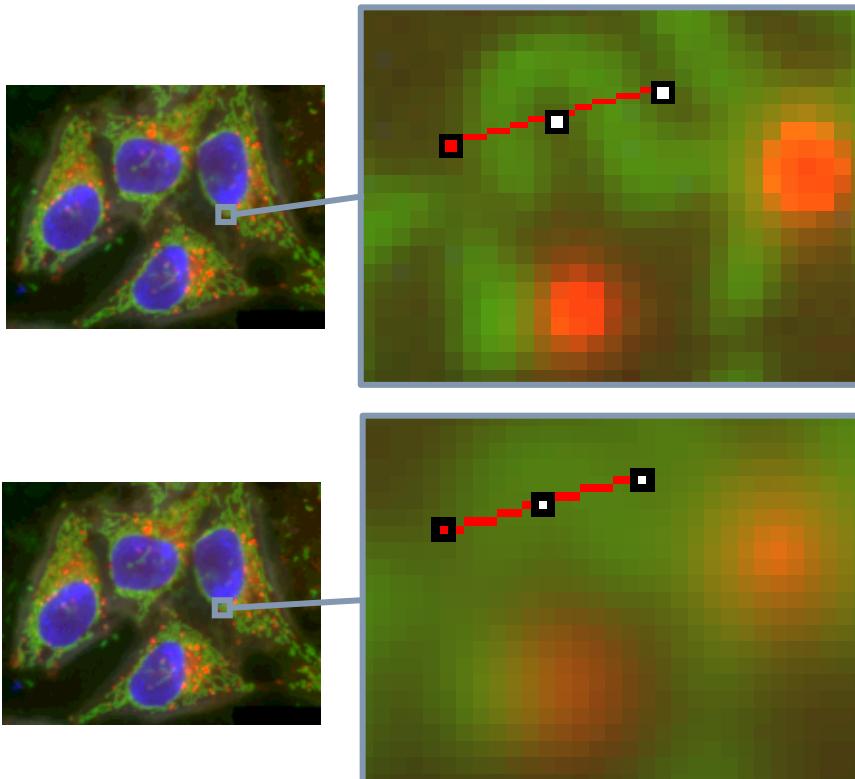
- Pixel size is a digital property of an image.
- You configure it during the imaging session at the microscope.



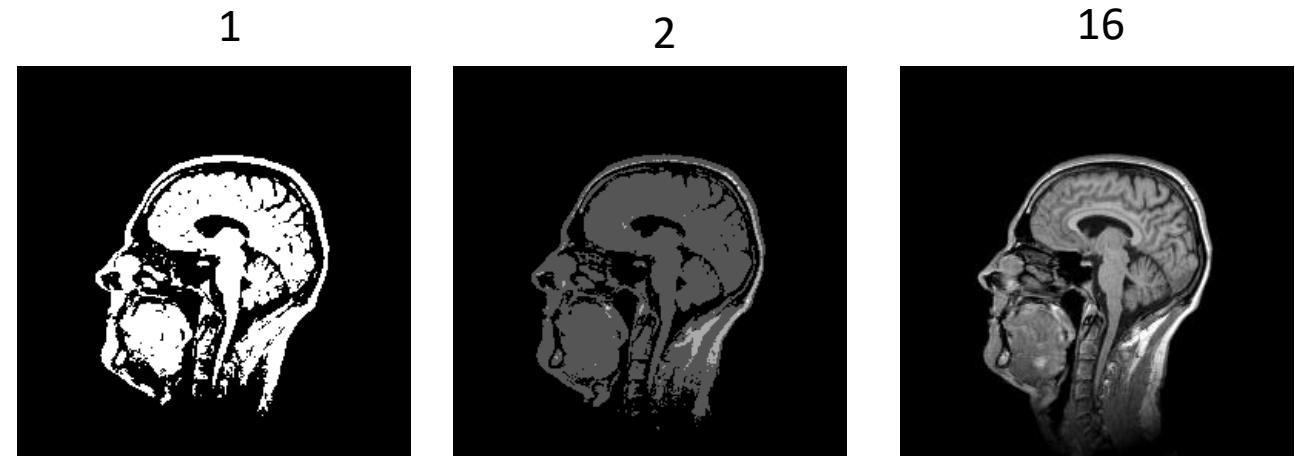
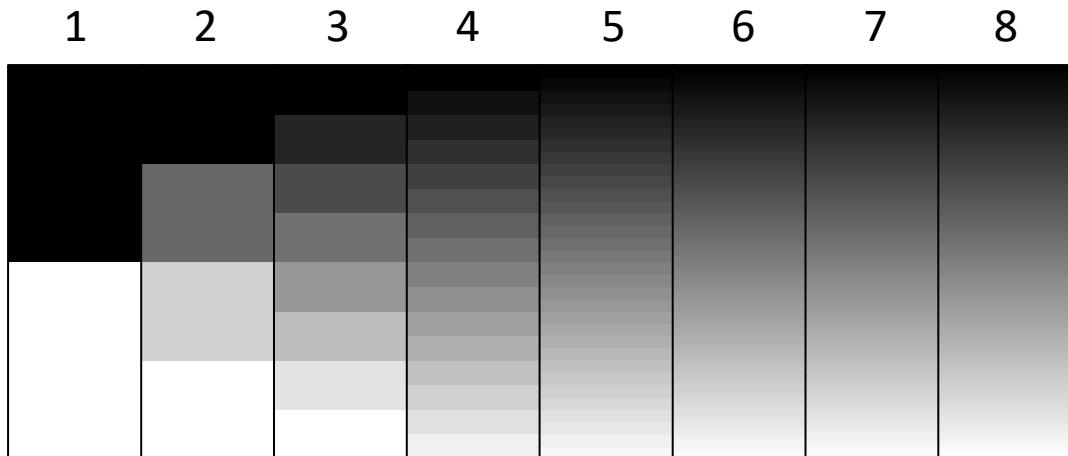
- We are not talking about resolution!

# Pixel size versus resolution

- Resolution is a property of your imaging system.
- The measure of how close object can be in an image while still being differentiable, is called spatial resolution.



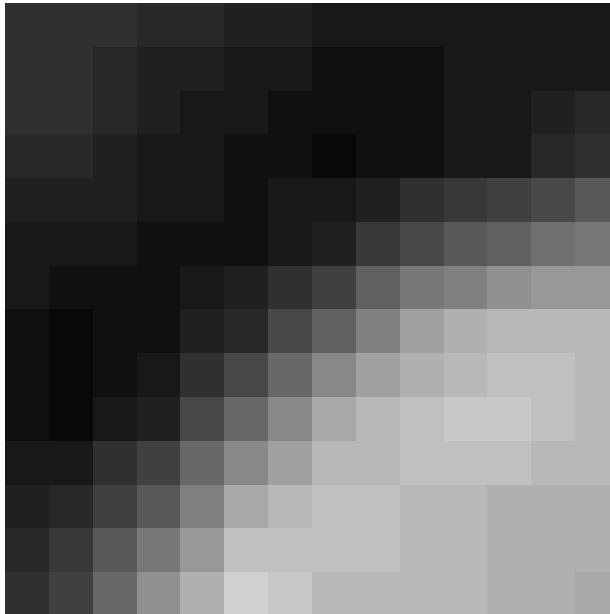
- A bit is the smallest memory unit in computers, *atomic data*.
- The bit-depth  $n$  enumerates how many different intensity values are present in an image:
  - $2^n$  grey values
- In microscopy, images are usually stored as 8, 12 or 16-bit images.



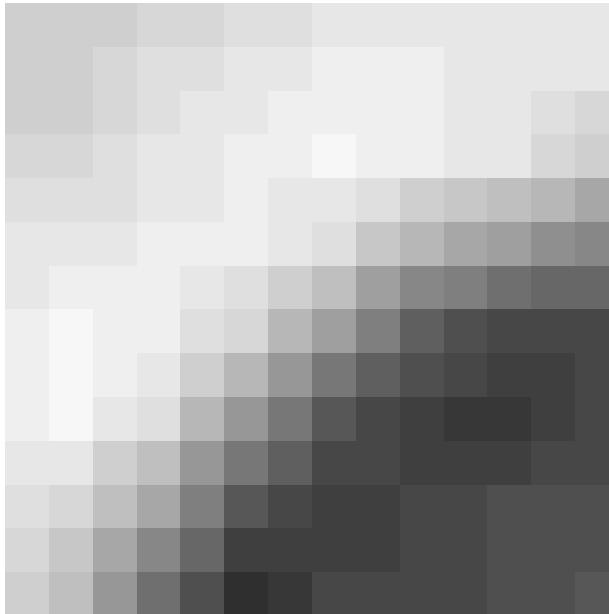
# Lookup tables

- The lookup table decides how the image is displayed on screen.
- Applying a different lookup table doesn't change the image. All pixel values stay the same, they just appear differently

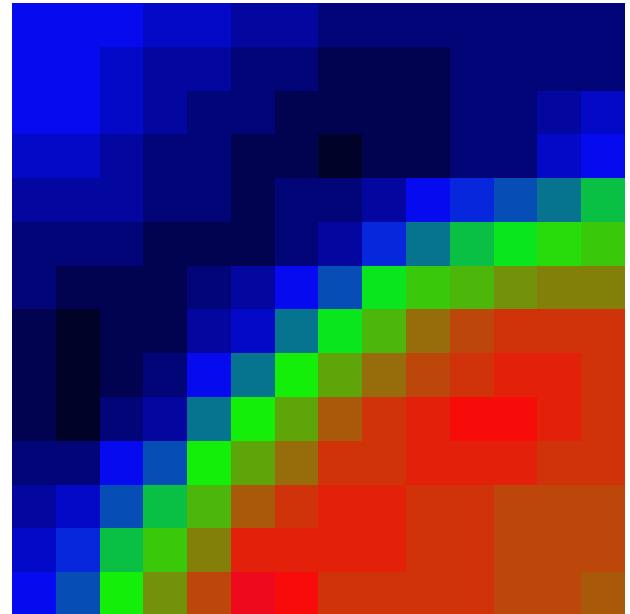
Pixel value	Display color
0	Black
1	Dark gray
2	Medium gray
...	
255	White



Pixel value	Display color
0	Black
1	Dark gray
2	Medium gray
...	
255	Black



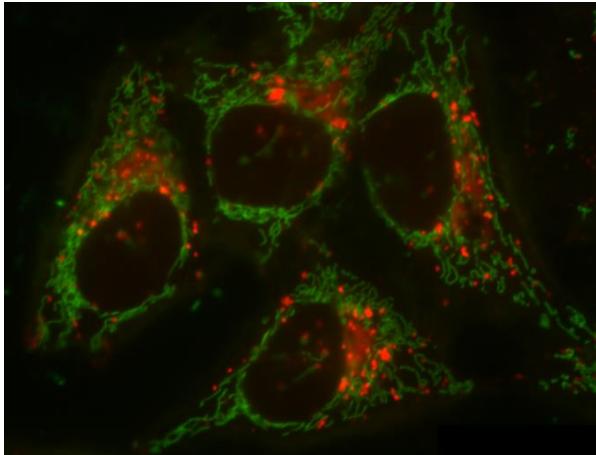
Pixel value	Display color
0	Red
1	Orange
2	Yellow
...	
255	Blue



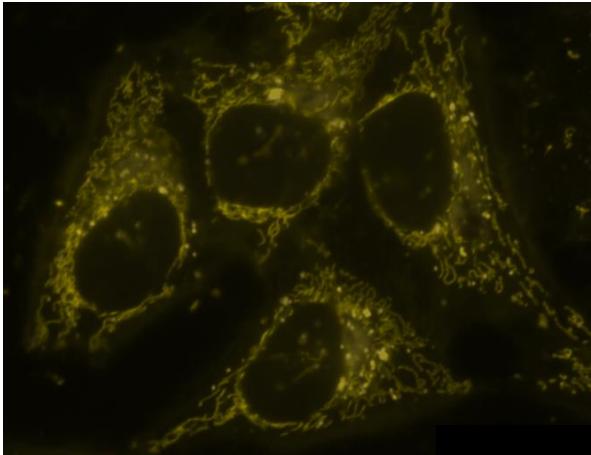
# Lookup tables

- Choose visualization of your color tables wisely!
- Think of people with red/green blindness!

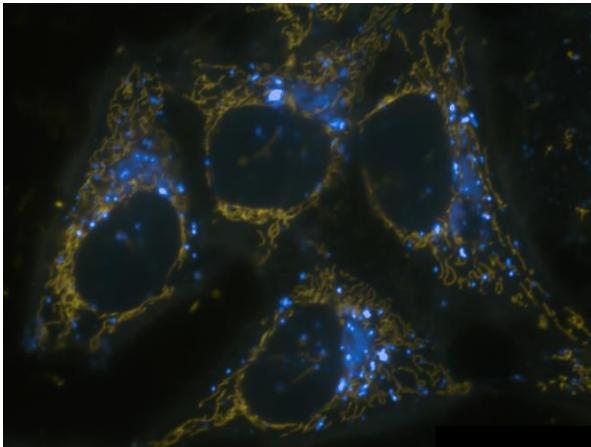
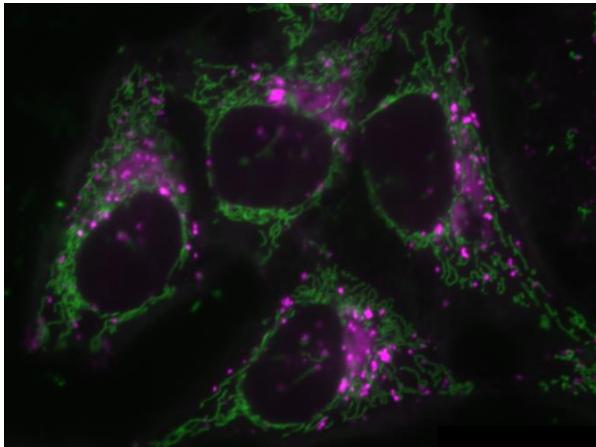
Default view



Red/green blind people see it like this



Replace red with  
magenta!

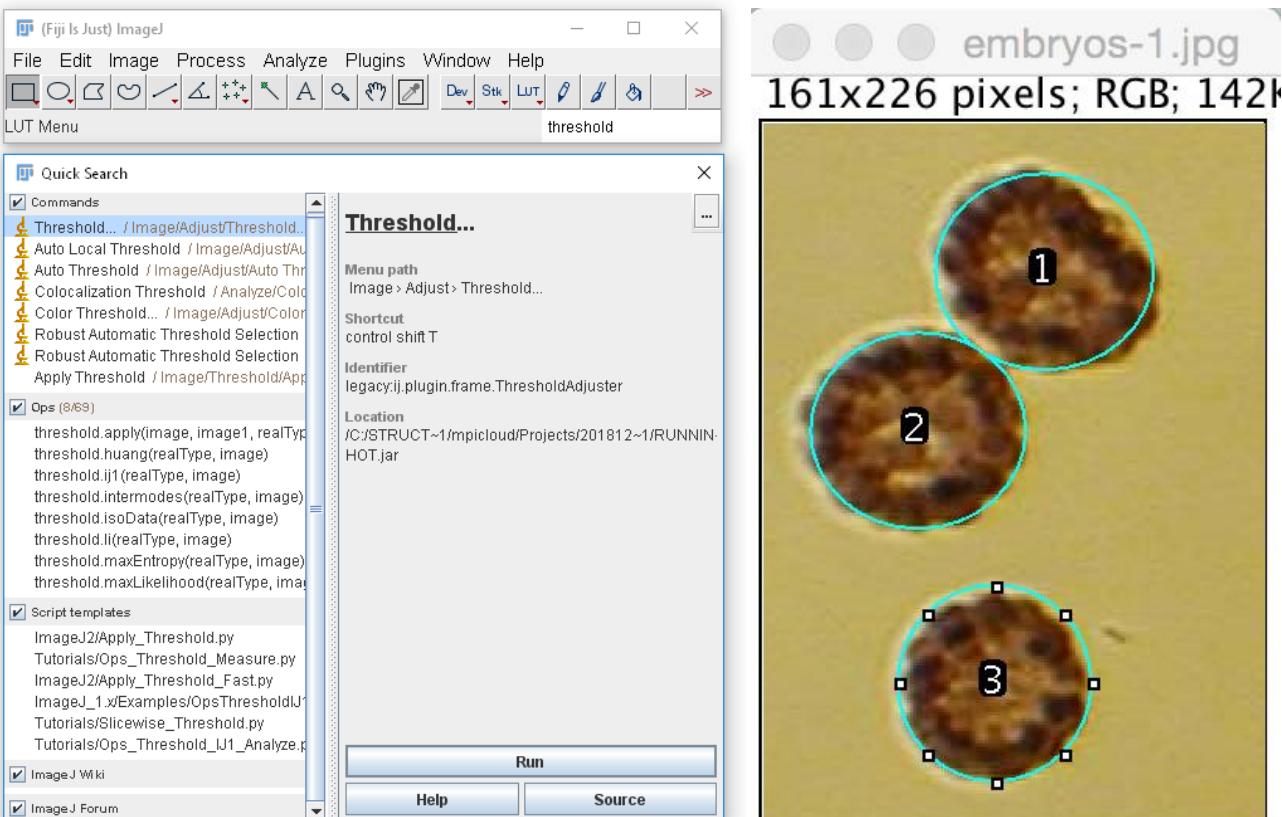


In this lecture you learned

- *Bio-image analysis*
  - Quantitative
  - Objective
  - Reproducible
  - Repeatable
  - Reliable
- What are bits, pixels, voxels, images
- Resolution versus pixel size
- Lookup tables
- Histograms

Coming up next

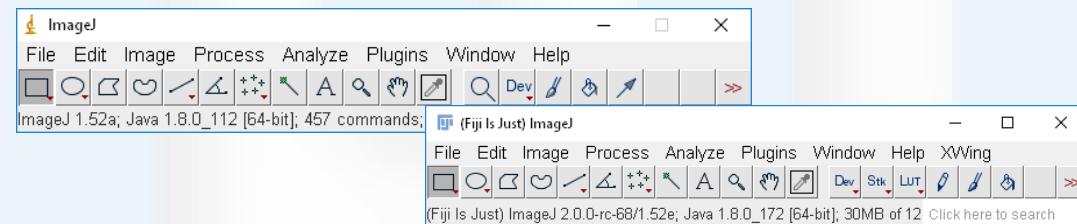
- Introduction to Fiji
- Regions of interest
- Basic measurements



# Biolimage Analysis with Fiji

Robert Haase, Myers lab, MPI CBG

With material from Benoit Lombardot, SCF MPI CBG

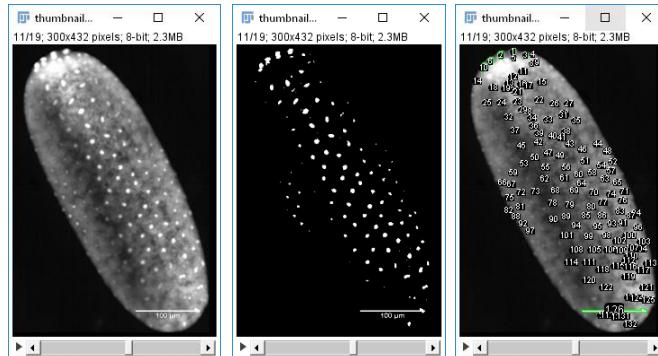
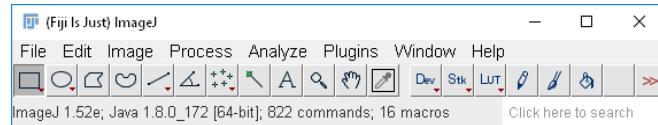


April 2021

# Fiji: Image analysis with the head in clouds

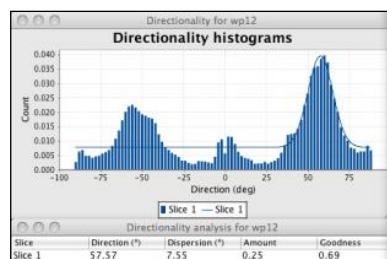
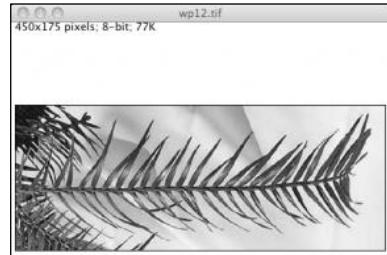
- Fiji is just ImageJ - with batteries included.

## Core ImageJ

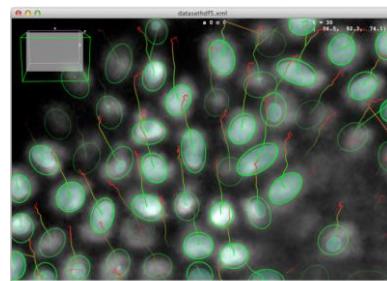


Results						
	Area	Mean	StdDev	Mode	Min	Max
1	53	218.245	41.615	255	104	255
2	73	243.466	27.442	255	135	255
3	17	215.941	23.360	177	177	249
4	1	181.000	0.000	181	181	181
5	10	249.000	8.602	255	229	255
6	64	243.984	19.667	255	178	255
7	2	225.500	0.707	225	225	226

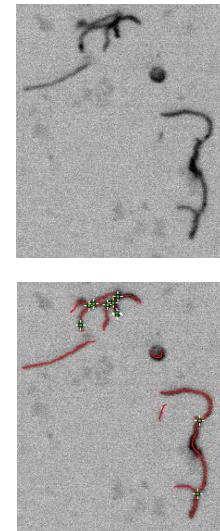
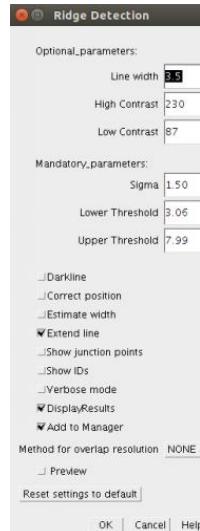
## Directionality



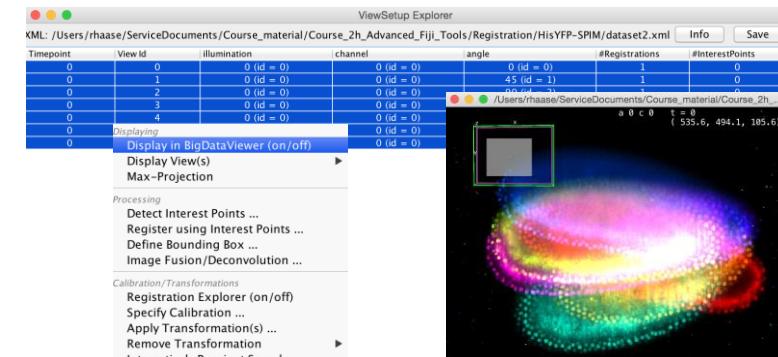
## BigDataViewer



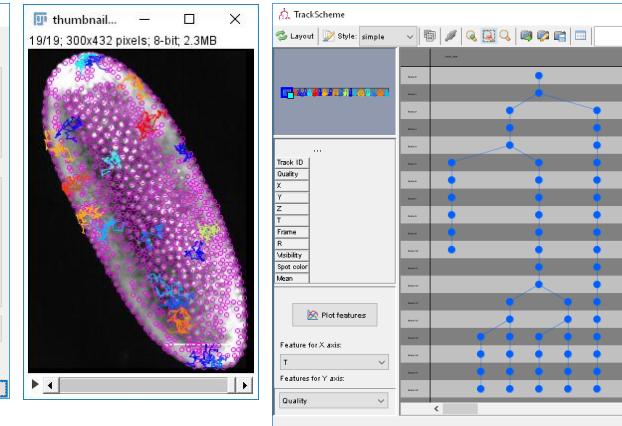
## Ridge detection



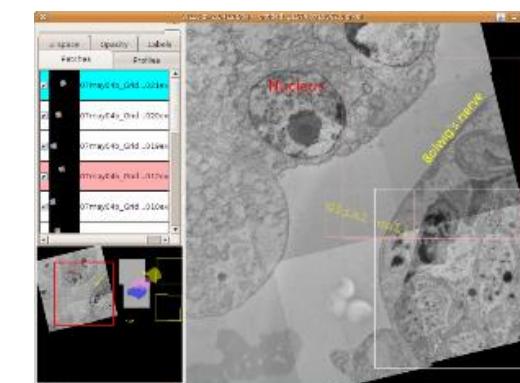
## Multiview fusion



## Trackmate

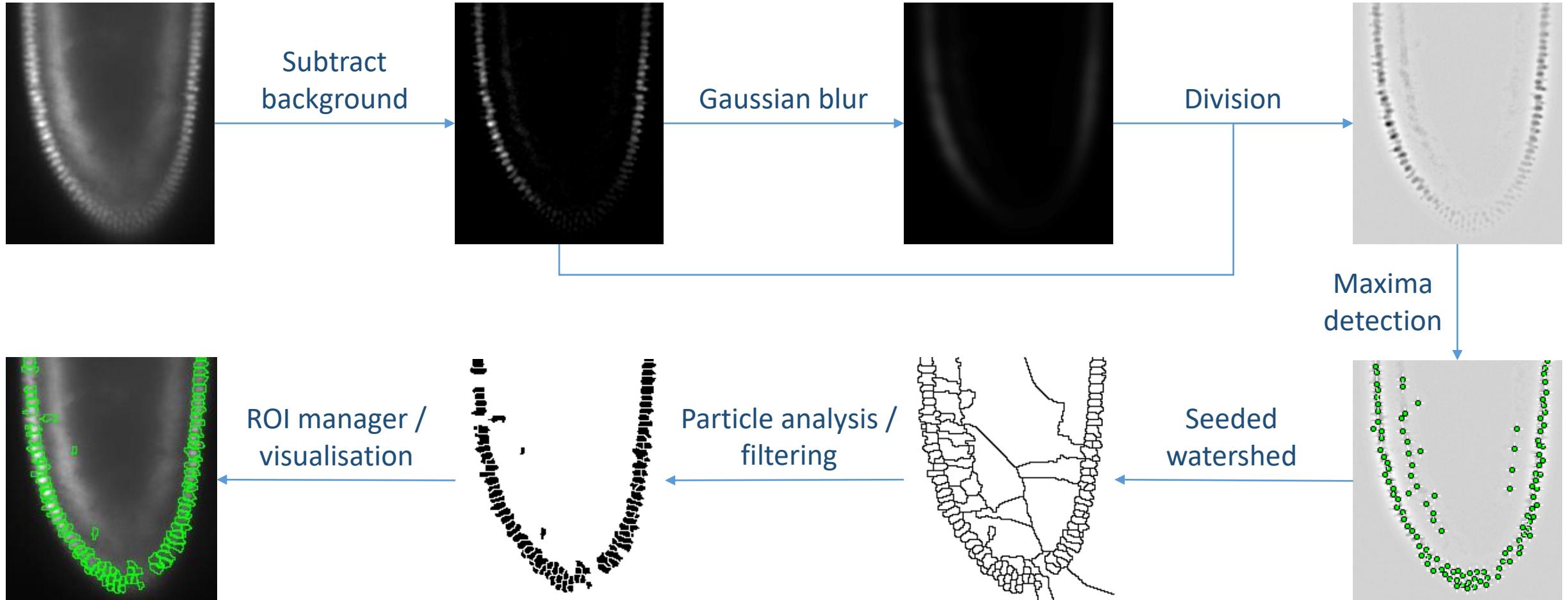


## TrakEM



# Fiji: Image analysis with feed on the ground

- Low level image processing in all shades of grey

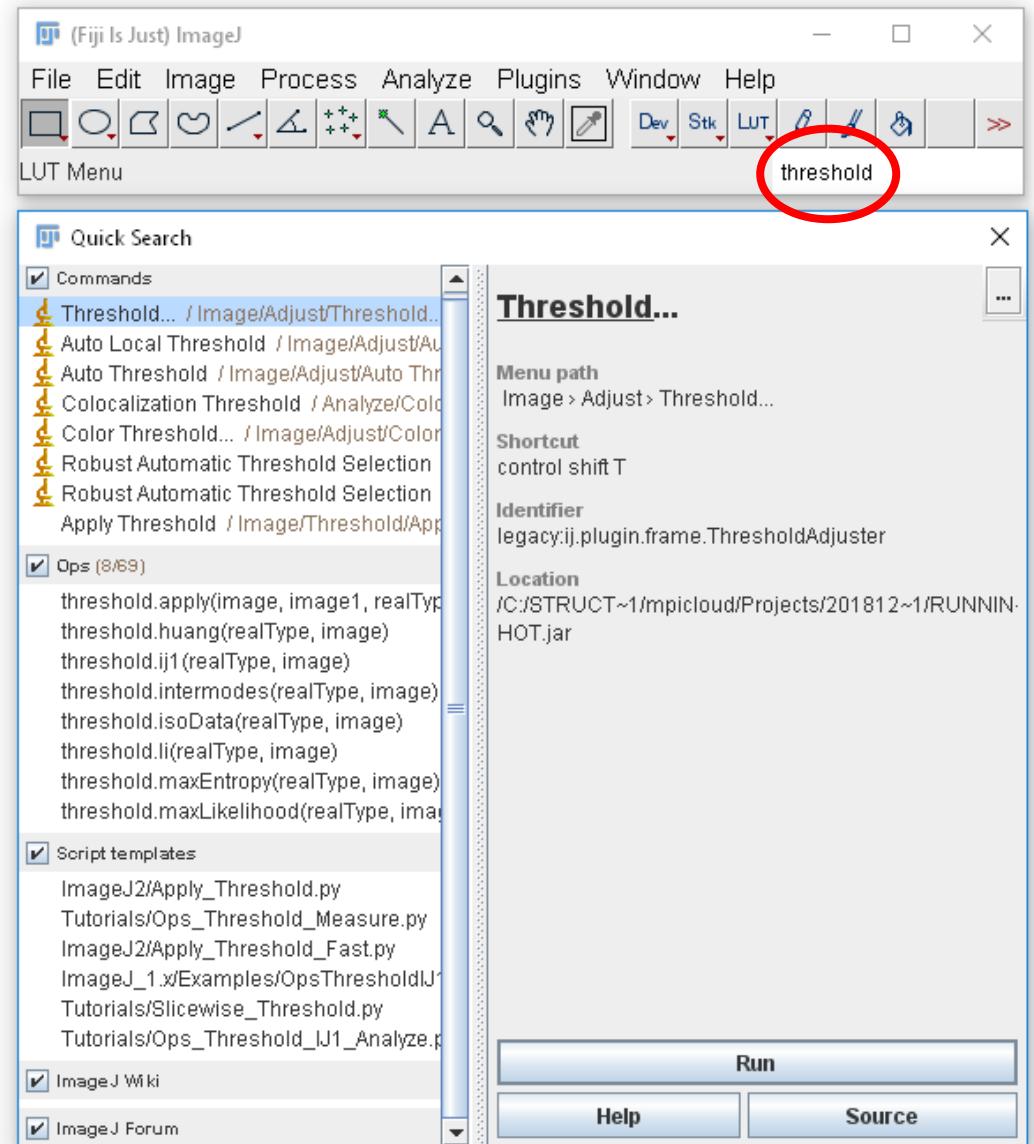


# Fiji user experience

- The menus are confusing and not very well organized

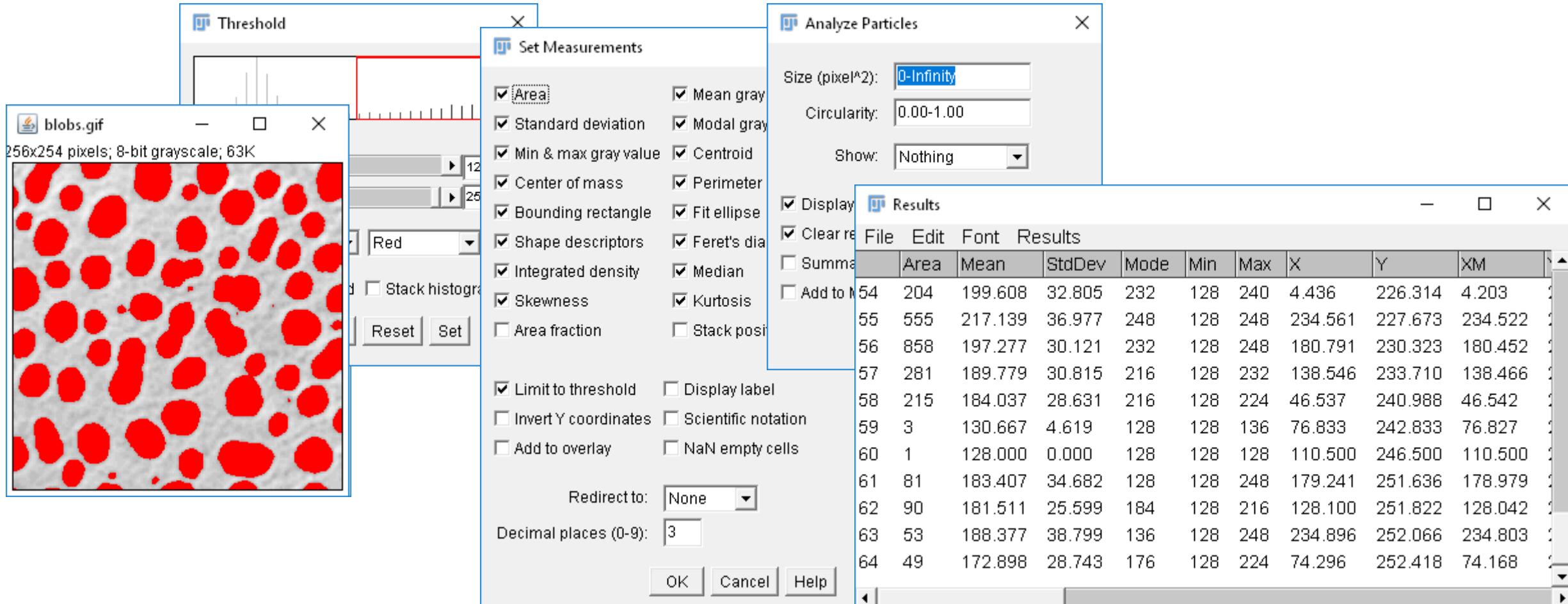
- Use the search field!

- It shows you where the plugin is located
- You can run it from here (press enter)
- Allows you searching
  - for plugins / menus
  - Operations (ops)
  - ImageJ wiki
  - Forum



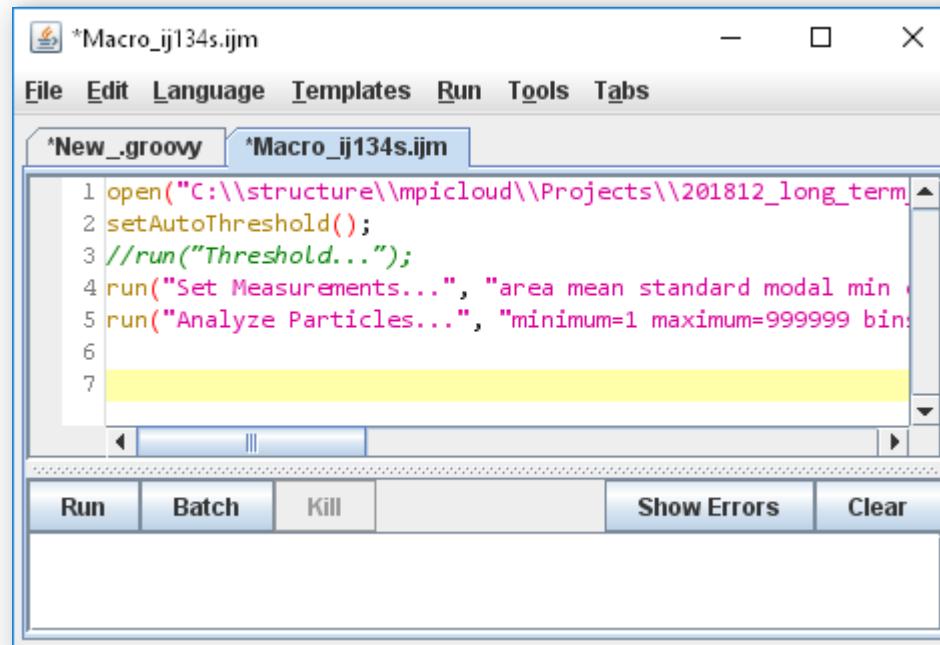
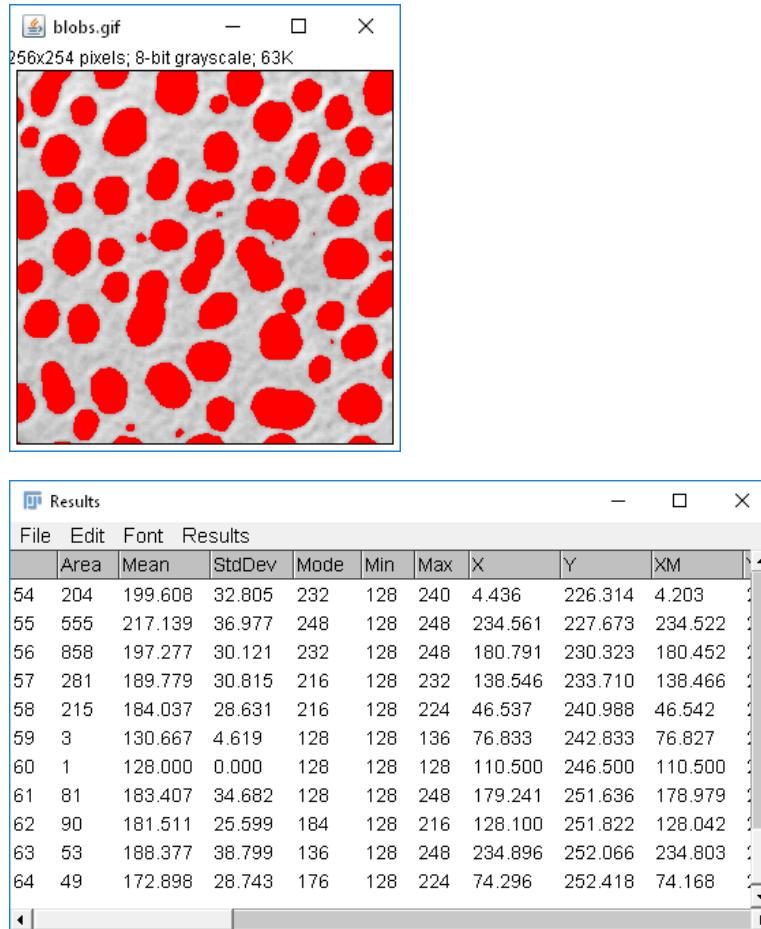
# Fiji example workflow

- Popular basic analysis pipeline: thresholding + particle analysis
- From images to raw data for statistics and \*omics post-processing



# Fiji & reproducible research: macros

- The macro recorder and the script editor allow scripting Fiji and thus, reproducible workflows.
- Run workflows on hundreds or thousands of images!



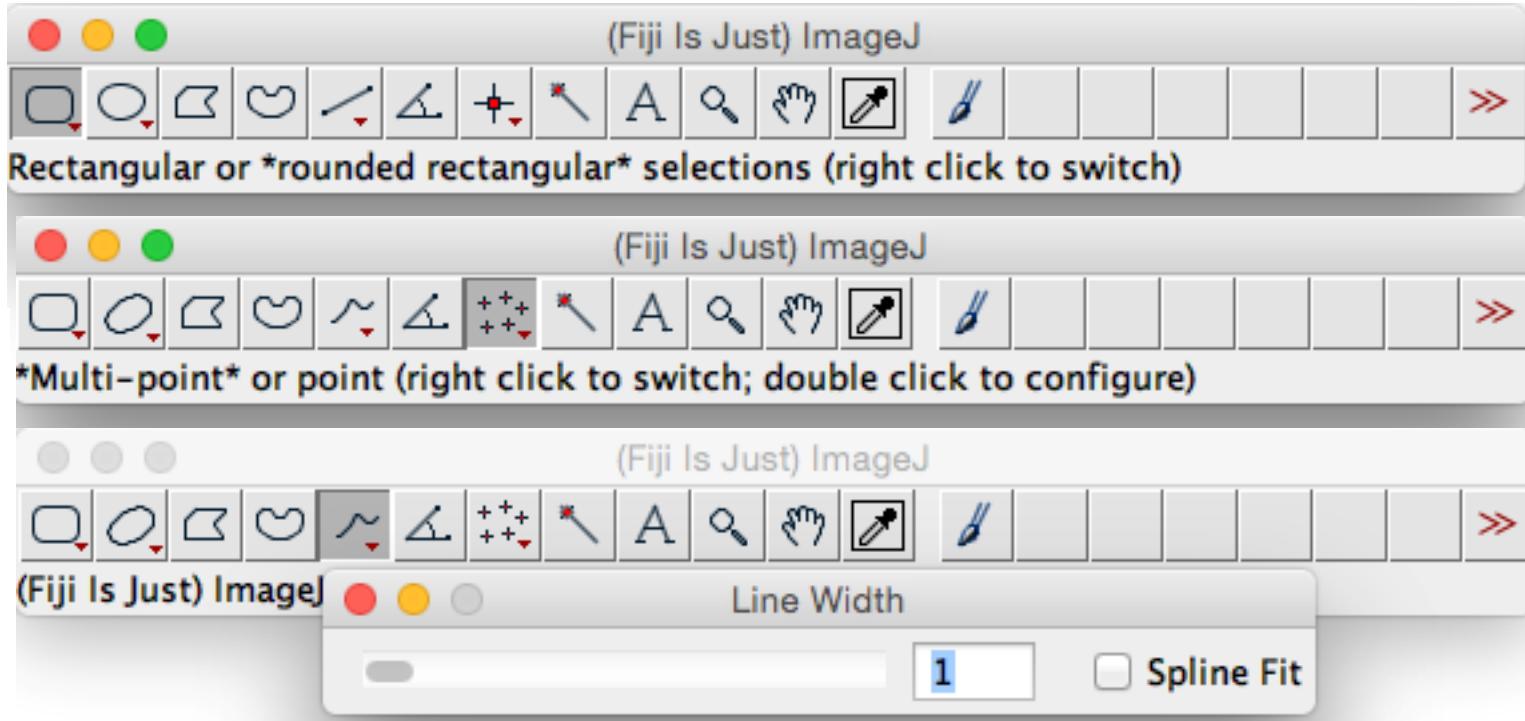
The figure shows the Fiji Script Editor window titled "\*Macro\_ij134s.ijm". The menu bar includes File, Edit, Language, Templates, Run, Tools, and Tabs. The tabs section shows two tabs: "New\_groovy" and "Macro\_ij134s.ijm", with "Macro\_ij134s.ijm" selected. The code area contains the following Groovy script:

```
1 open("C:\\\\structure\\\\mpicloud\\\\Projects\\\\201812_long_term");
2 setAutoThreshold();
3 //run("Threshold...");
4 run("Set Measurements...", "area mean standard modal min");
5 run("Analyze Particles...", "minimum=1 maximum=999999 bin:
```

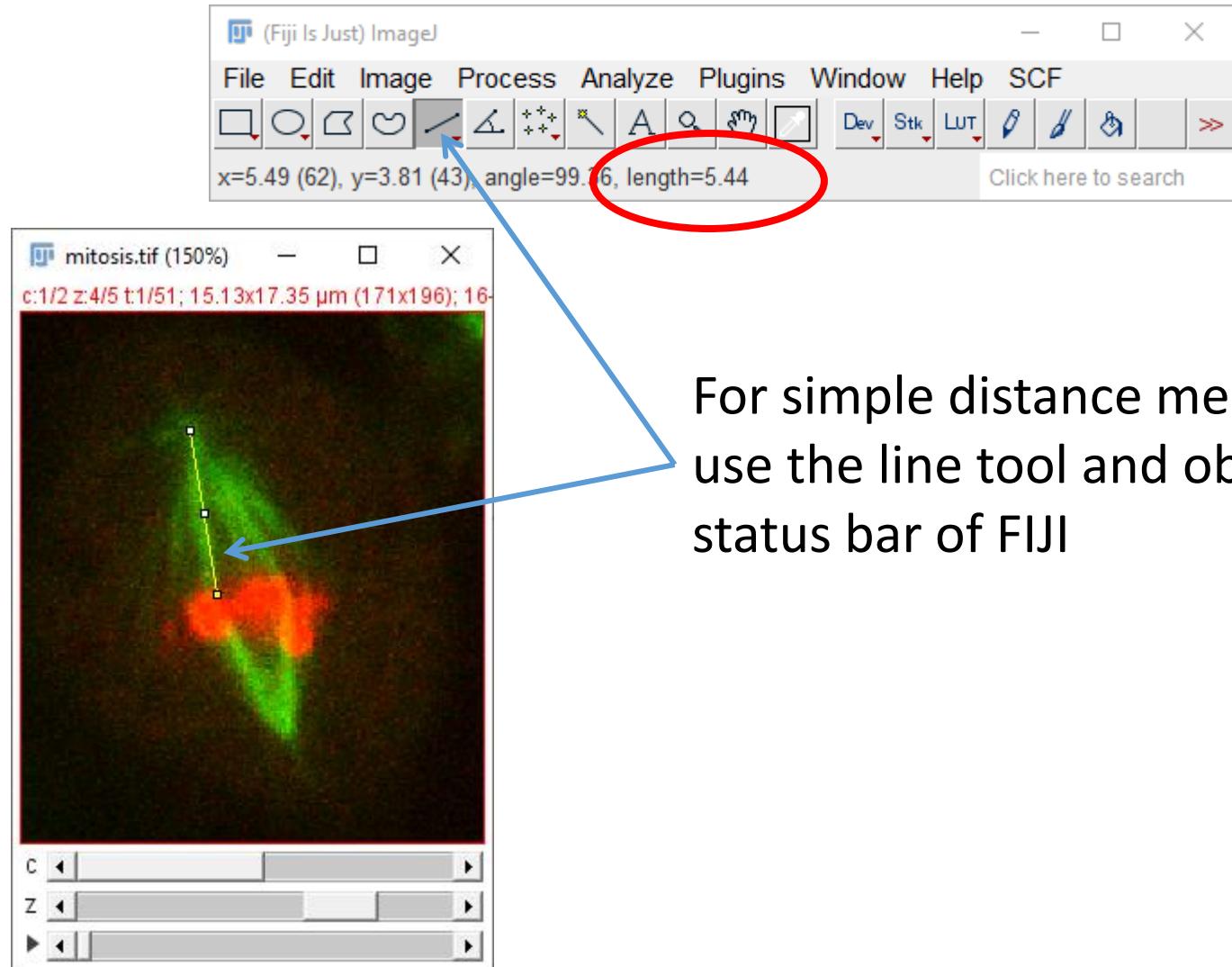
Below the code area is a toolbar with buttons for Run, Batch, Kill, Show Errors, and Clear.

# Fijis user interface

- There are more tools in the toolbar than expected...
- Use the right click to discover them!



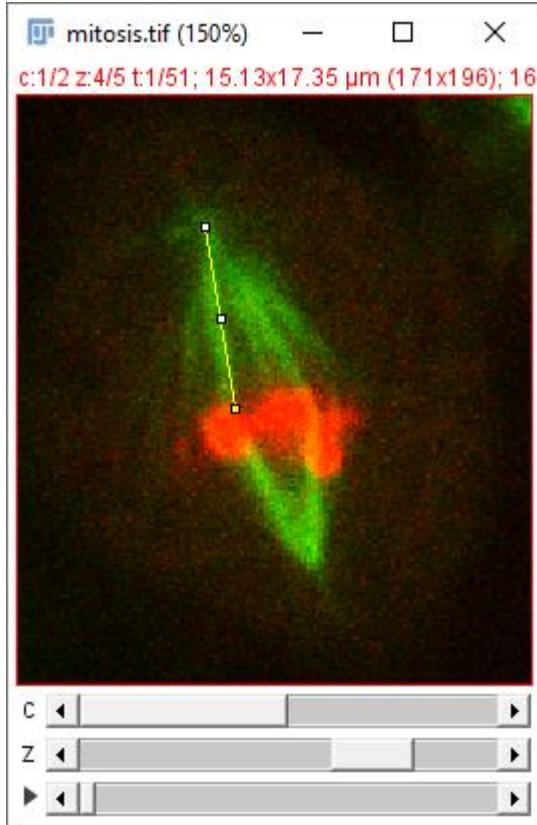
# Distance measurements



For simple distance measurements,  
use the line tool and observe the  
status bar of FIJI

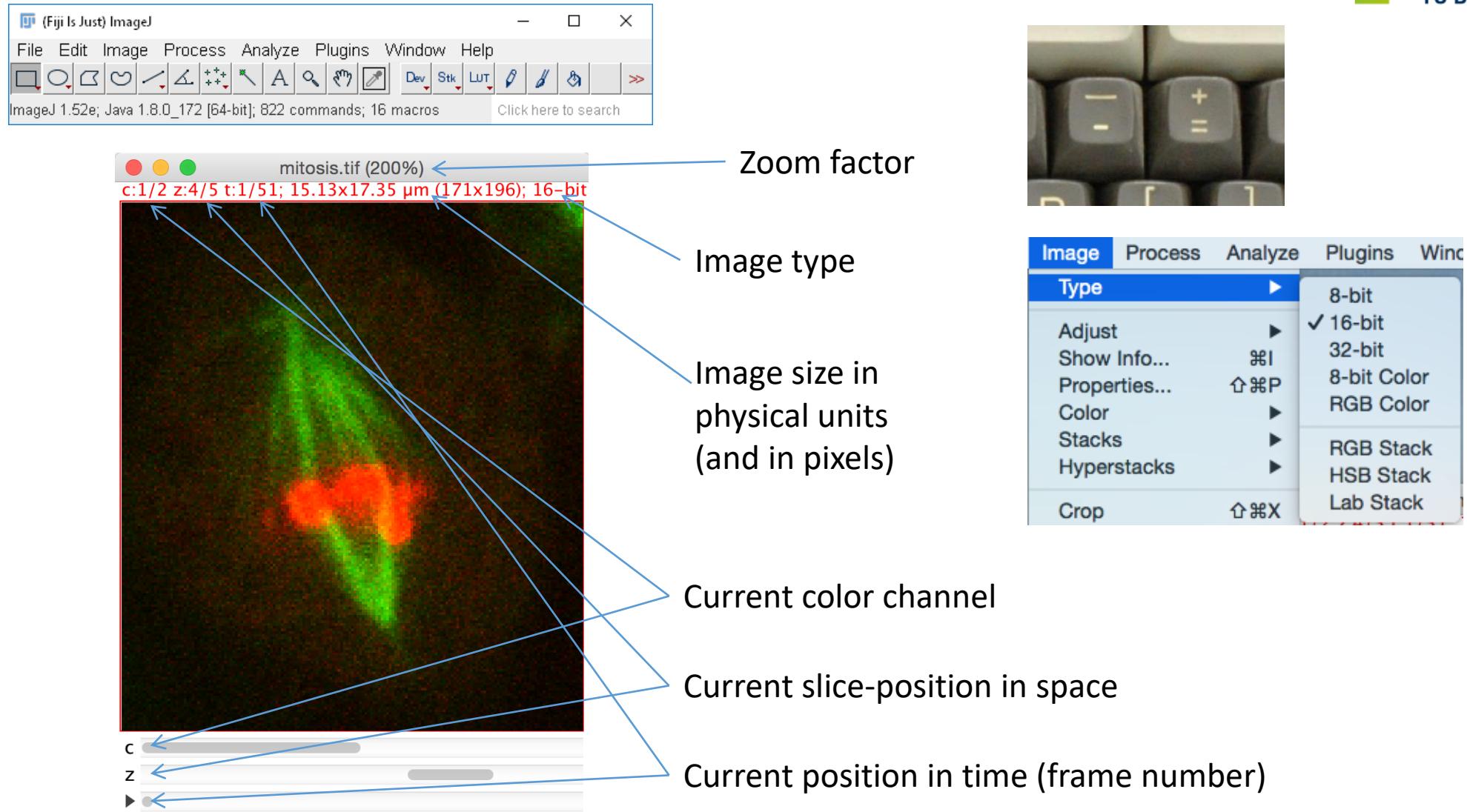
# Distance measurements

- Use the M key to store a length-measurement to the results table.



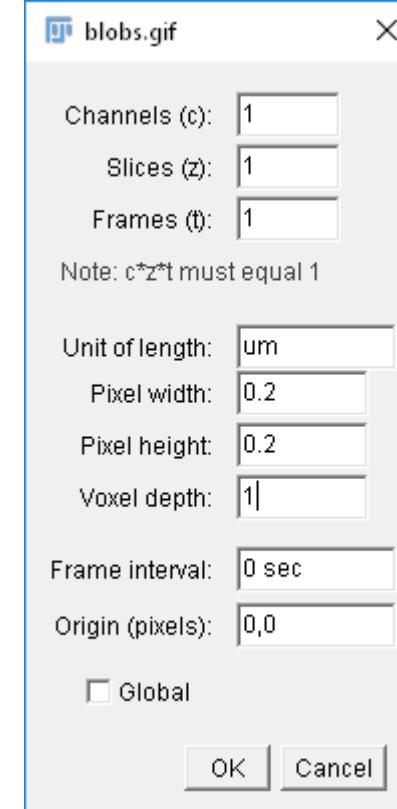
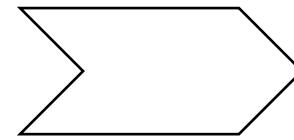
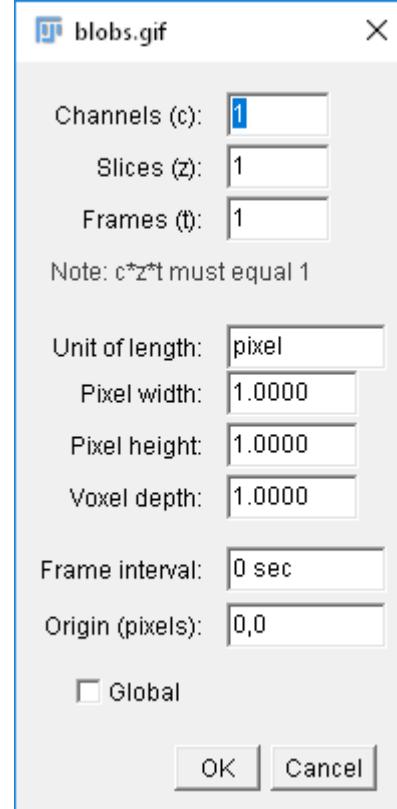
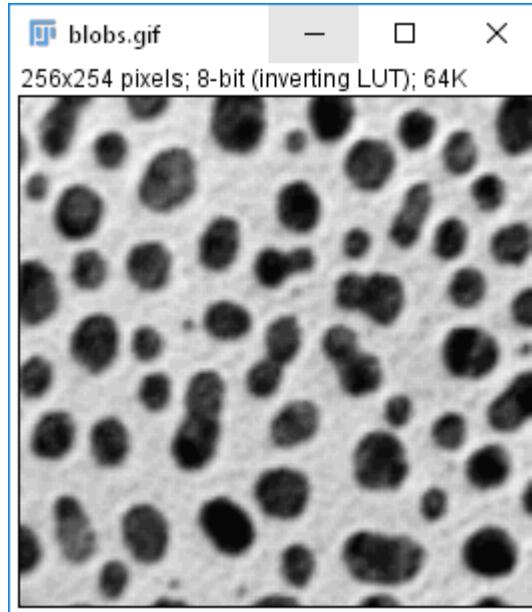
	Angle	Length
1	99.310	5.441

# Fijis user interface



# Correct pixel size

- Set pixel size by using *Image > Properties*

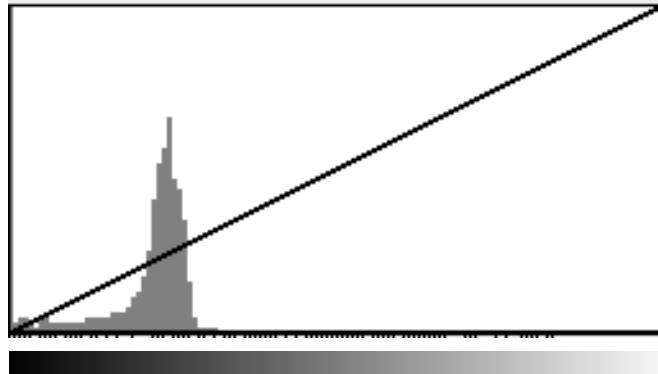
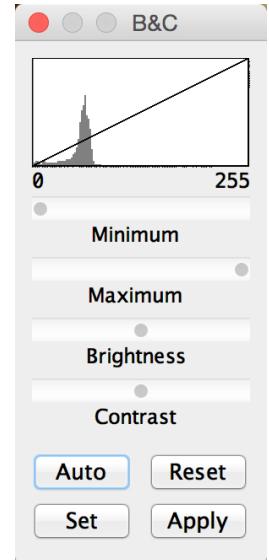
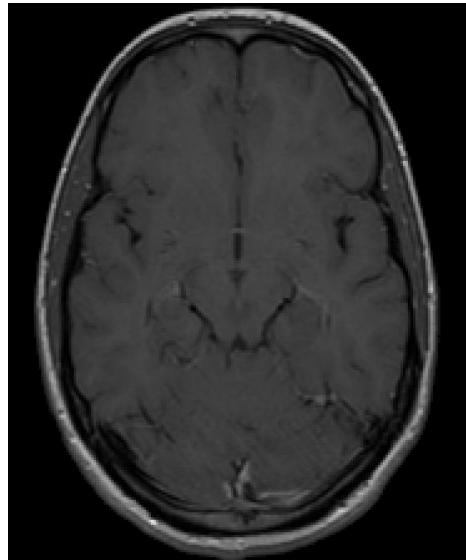


Here you can enter “um” instead of “ $\mu\text{m}$ ”. It will be interpreted as micrometer

- *Analyze > Set scale* is a workaround; it's not precise!

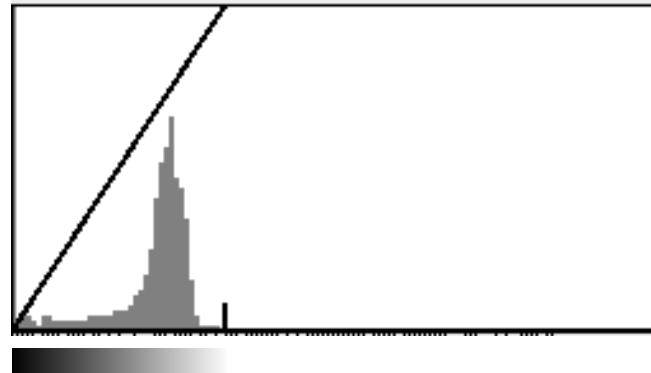
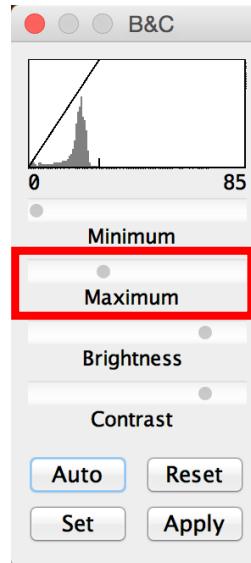
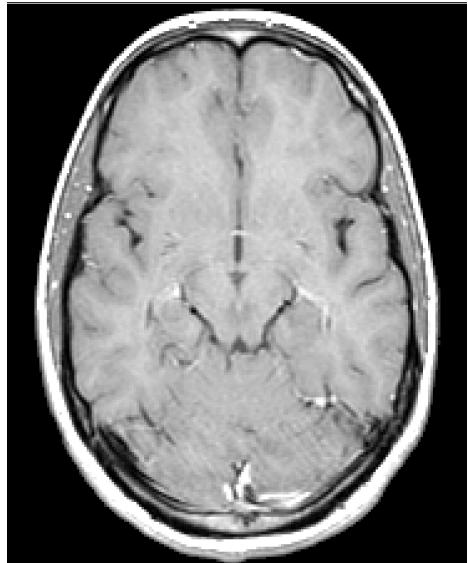
# Visualisation: brightness / contrast

- After opening an image, adjust Brightness & Contrast in order to see it properly
- Use the menu **Image > Adjust > Brightness & Contrast (CTRL+SHIFT+C)**



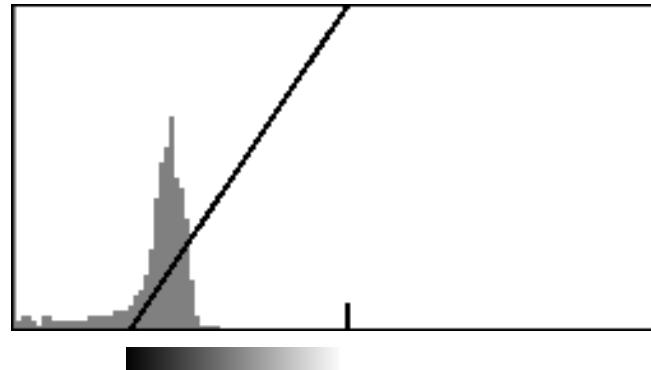
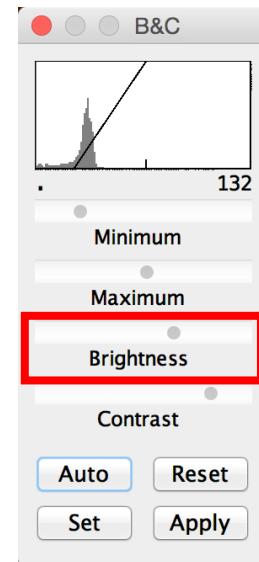
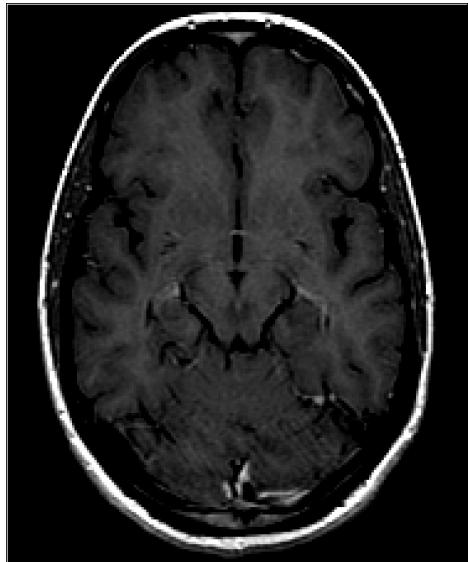
# Visualisation: brightness / contrast

- After opening an image, adjust Brightness & Contrast in order to see it properly
- Use the menu Image > Adjust > Brightness & Contrast (CTRL+SHIFT+C)



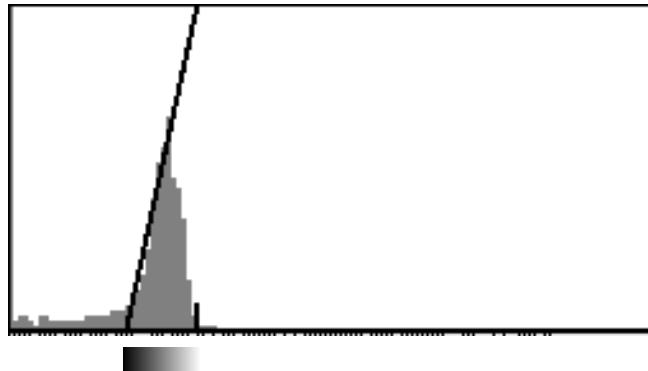
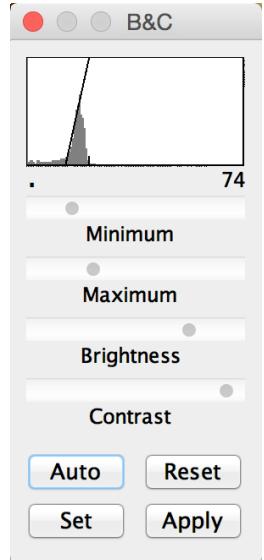
# Visualisation: brightness / contrast

- After opening an image, adjust Brightness & Contrast in order to see it properly
- Use the menu **Image > Adjust > Brightness & Contrast (CTRL+SHIFT+C)**



# Visualisation: brightness / contrast

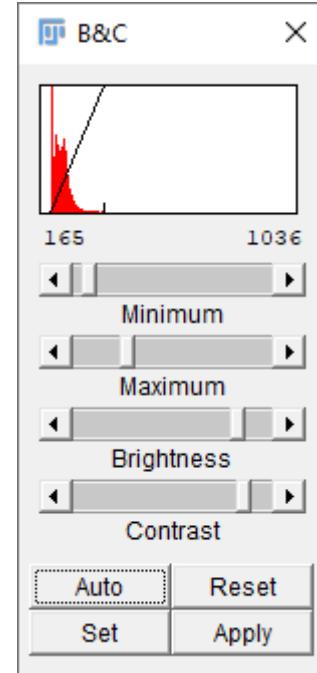
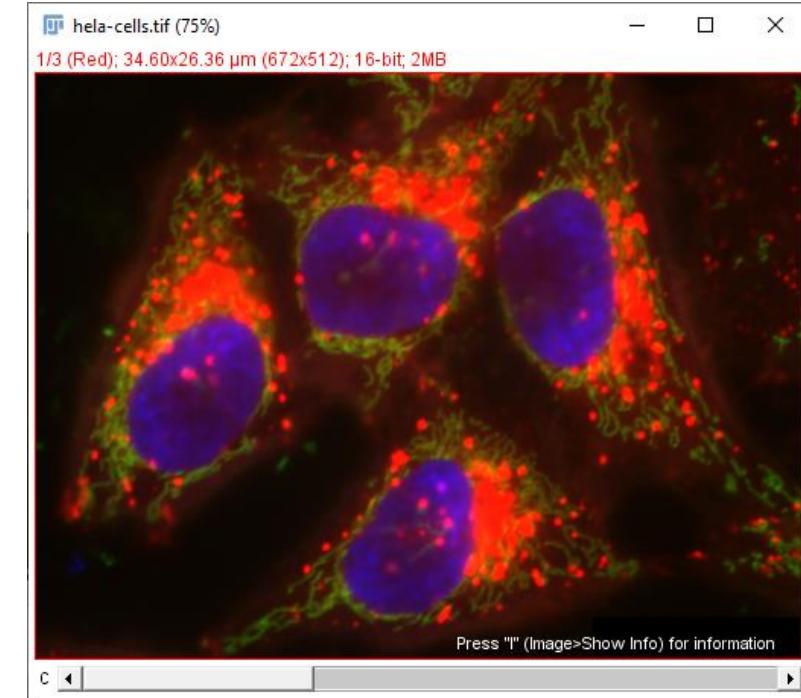
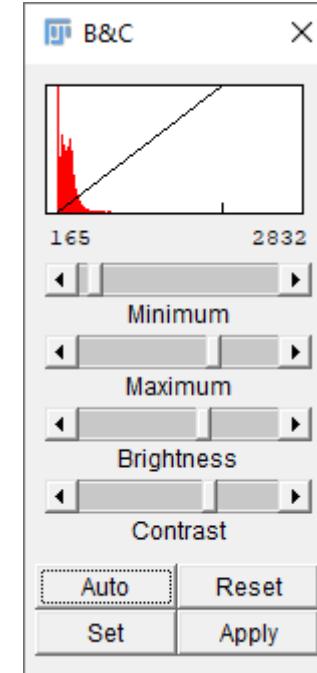
- Define the grey value range to visualize
- Also known as “windowing”
- Menu *Image > Adjust > Brightness/Contrast* or SHIFT+CMD+C



- Try to “window” the object you would like to see best.
- This may be the “hill” in the histogram

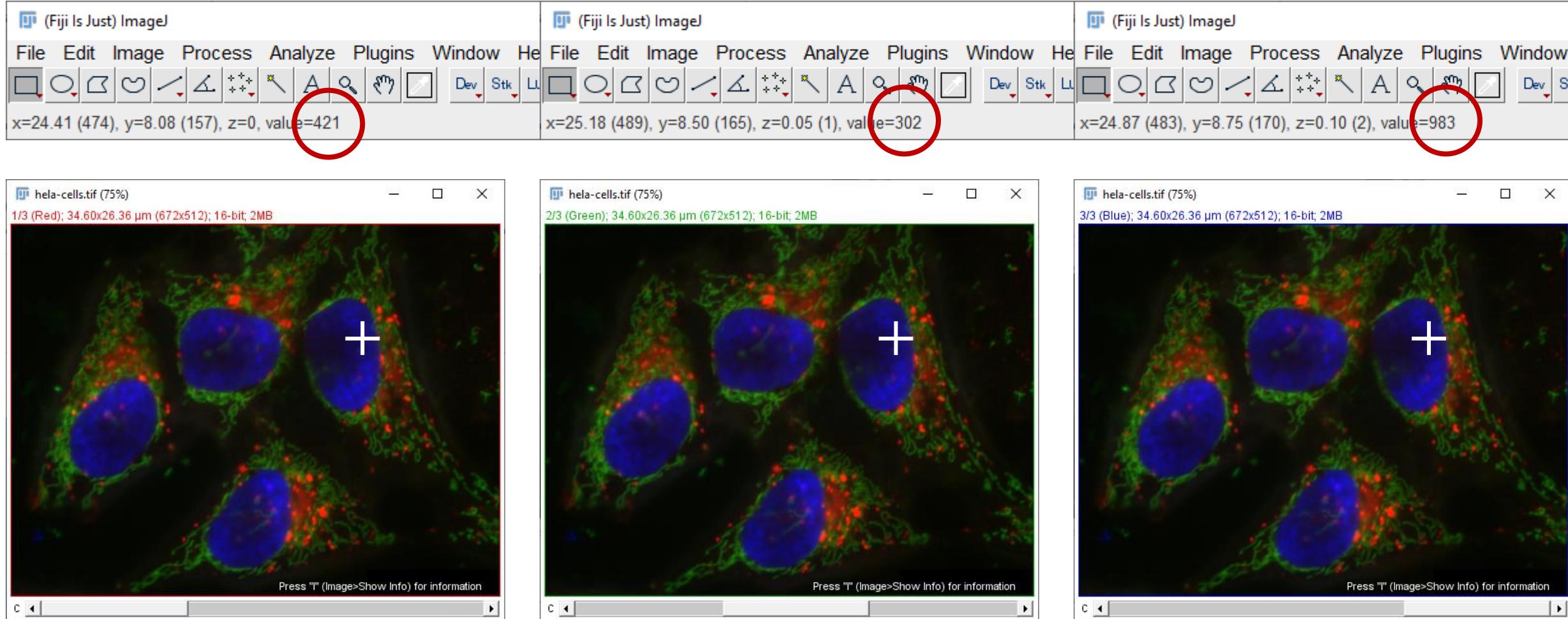
# Channel selection

- In multi-channel images, brightness and contrast must be configured for channels individually



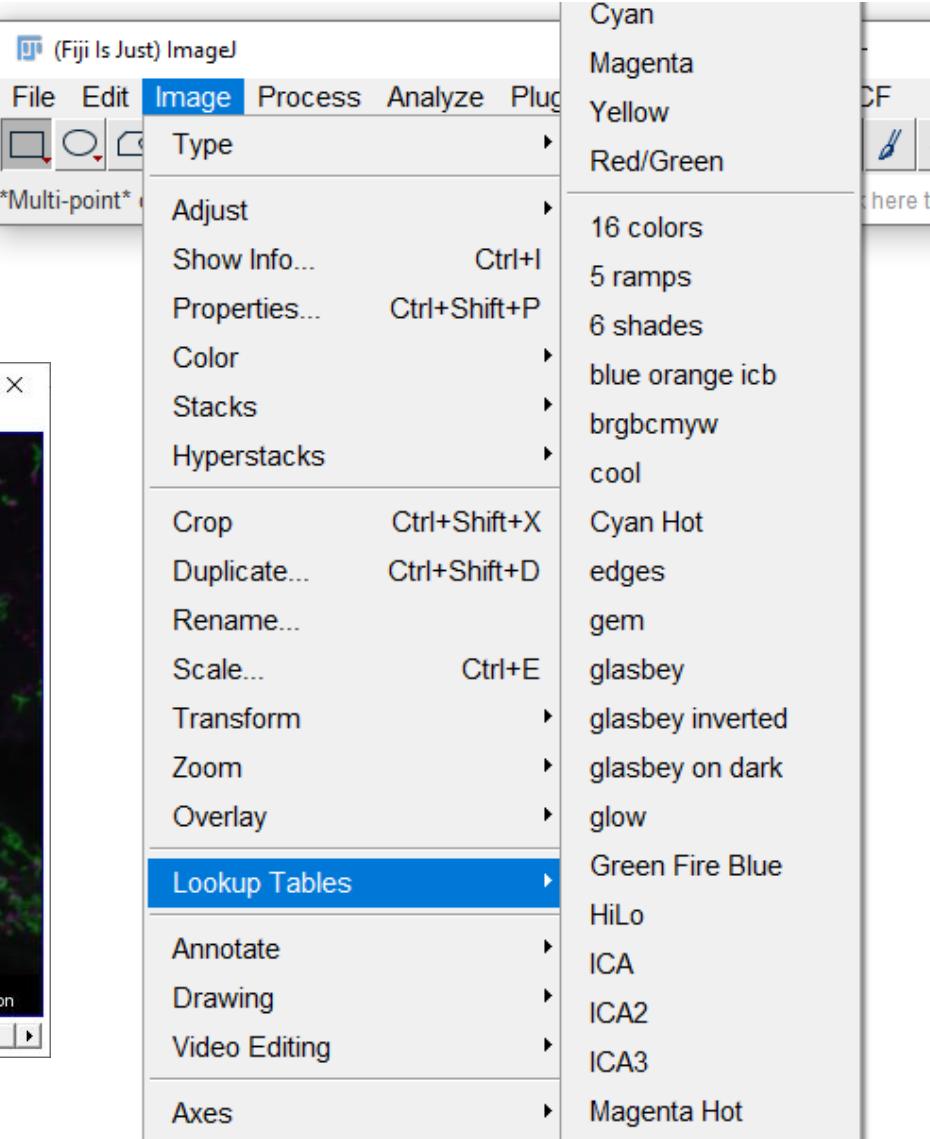
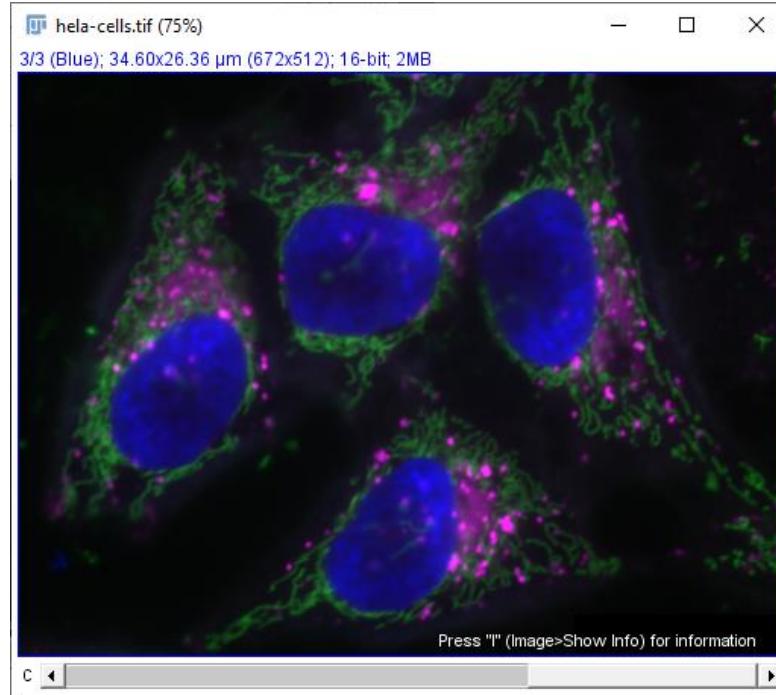
# Channel selection

- Intensity measurements depend on the selected channel!



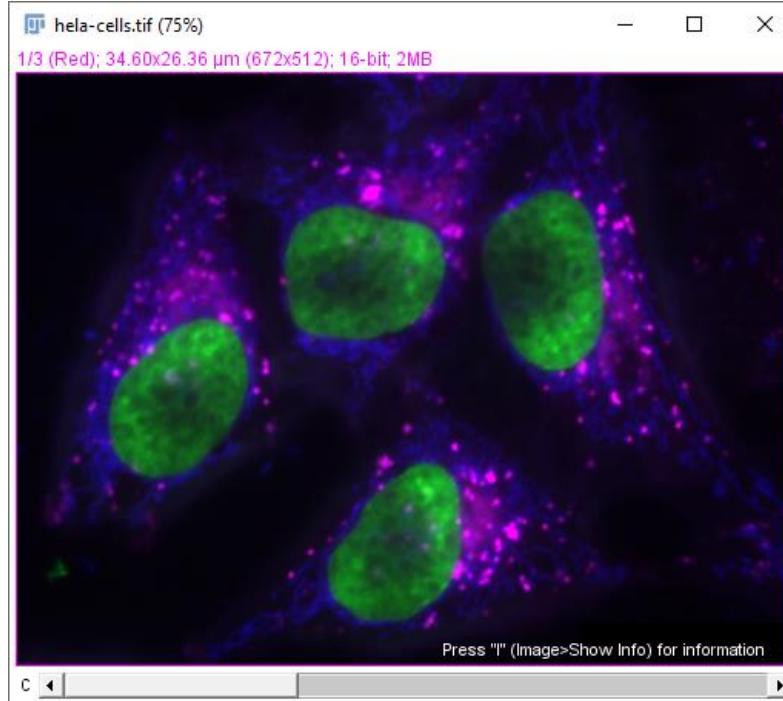
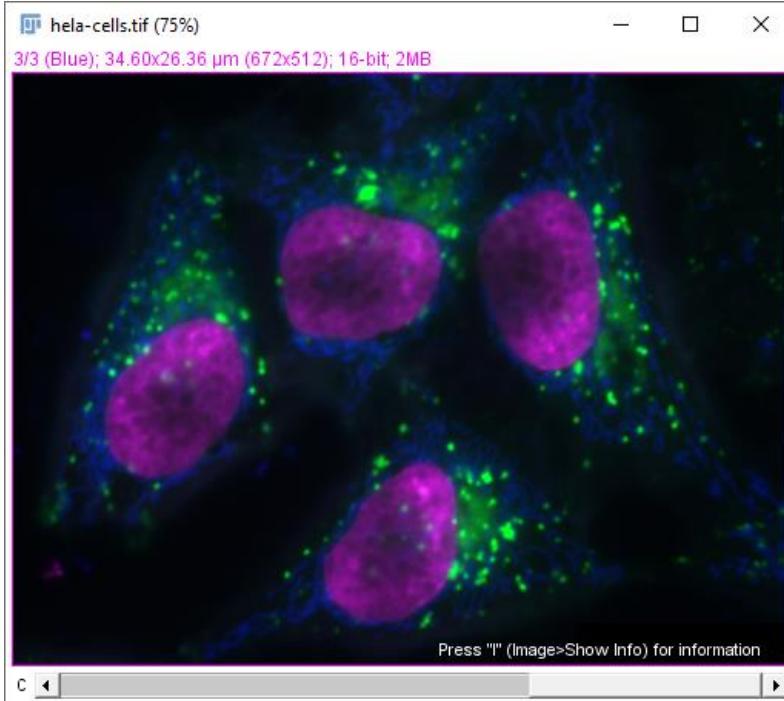
# Lookup tables

- Choose visualization of your color tables wisely!
- Think of people with red/green blindness!

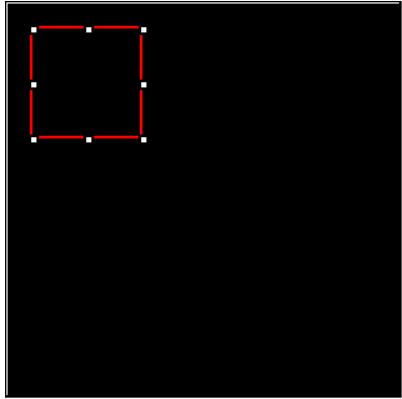


# Lookup tables

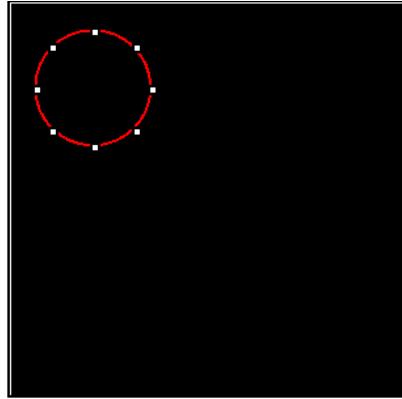
- Lookup tables are a free choice.
- They may (or may not) express which wavelengths has been imaged before.
- Hint: Try not to do arts.



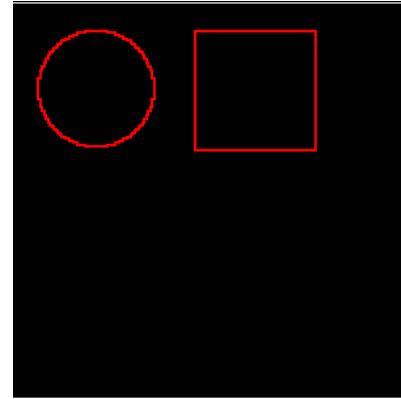
# Regions of interest (ROI)



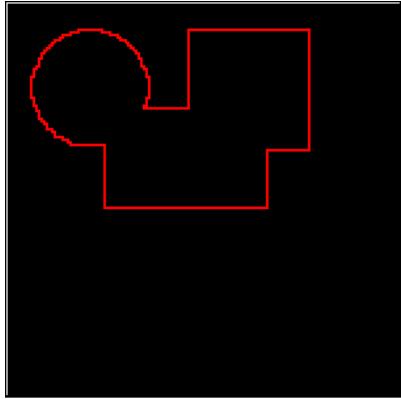
draw a rectangle  
using the rectangle tool



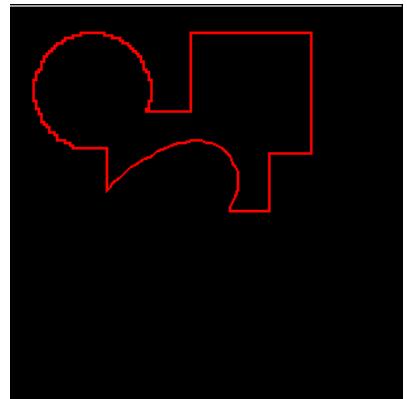
draw a circle  
using the circle tool



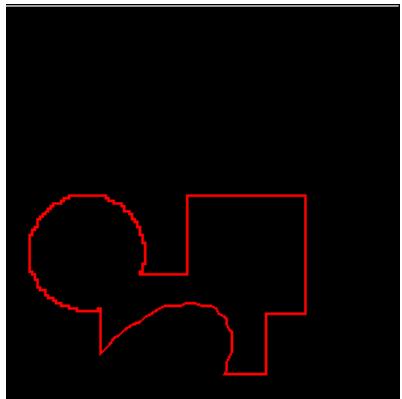
add objects to the ROI  
by holding SHIFT



unite objects to the ROI  
by holding SHIFT



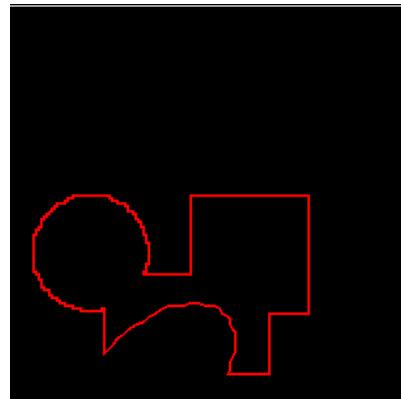
subtract objects to the  
ROI by holding ALT



move the ROI  
by clicking inside and  
move the mouse



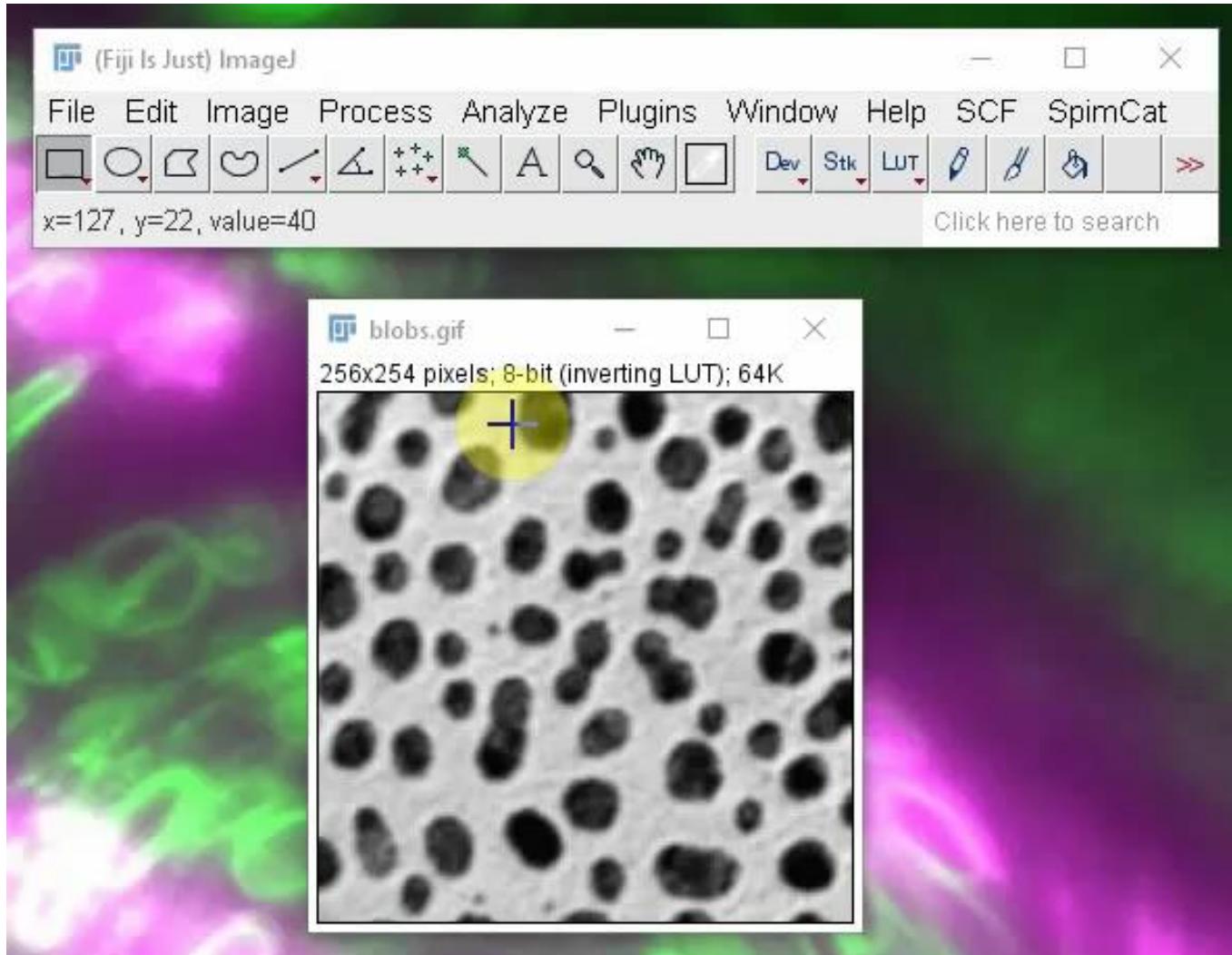
delete the ROI  
by clicking outside



Recover the last ROI by using menu  
Edit > Selection > Restore Selection  
(SHIFT+CMD+E)

# Fijis user interface

- The brush tools is very popular and very hidden

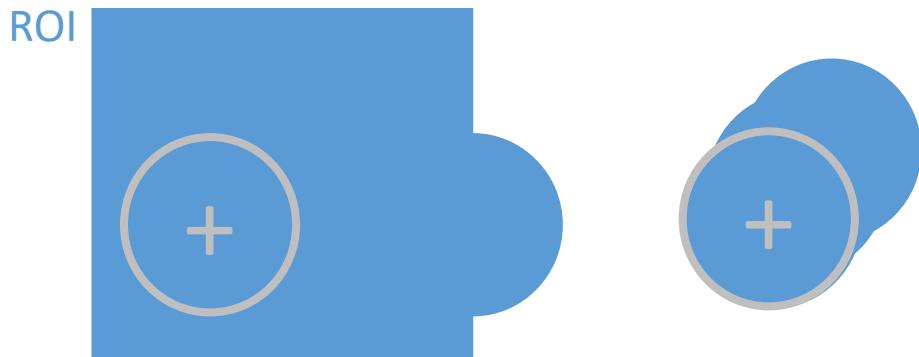


# Regions of interest (ROI)

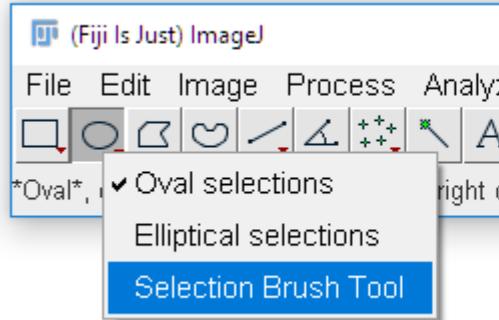
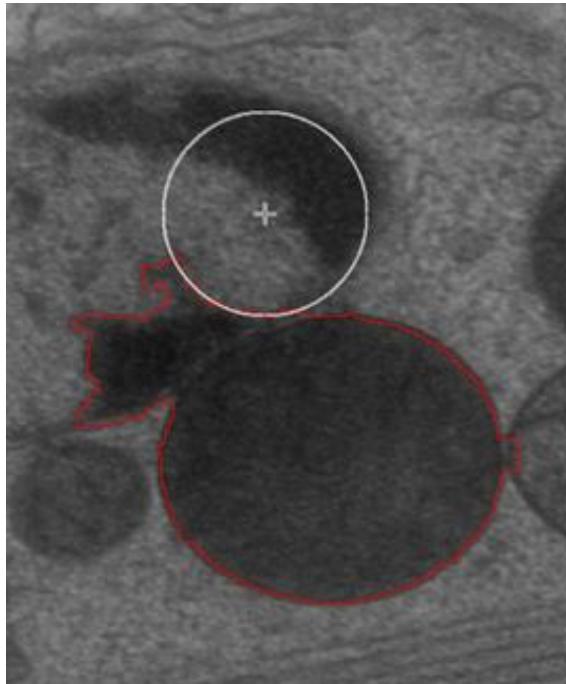
- The brush tools is very popular and very hidden

ADD pixels to an ROI:

- Approach it from inside



- Alternatively: hold SHIFT

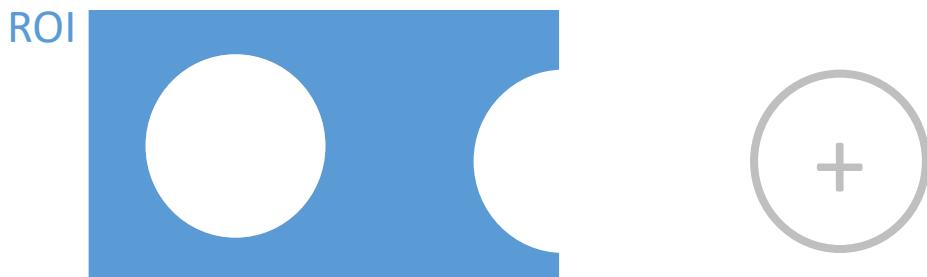


# Regions of interest (ROI)

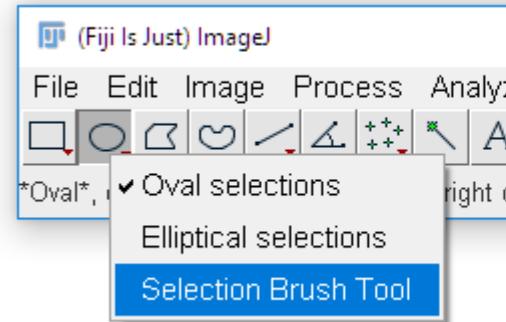
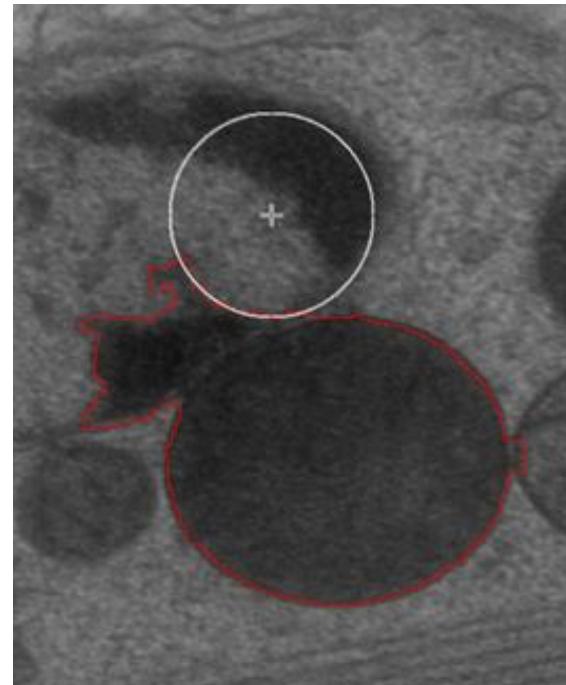
- The brush tools is very popular and very hidden

REMOVE pixels from an ROI:

- Approach it from outside

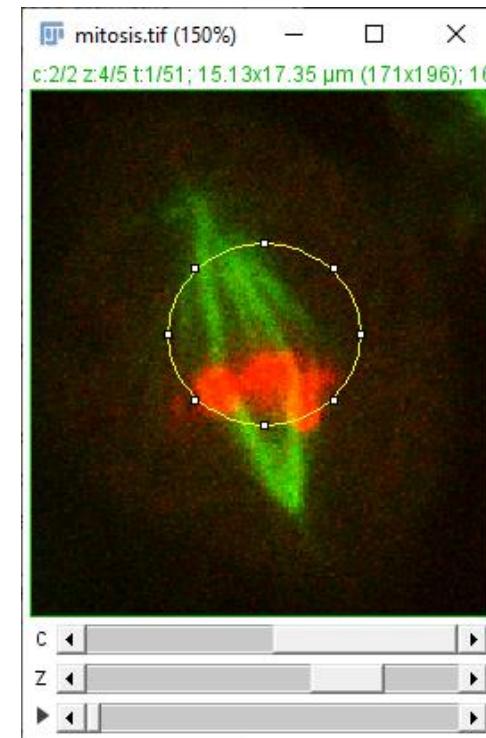
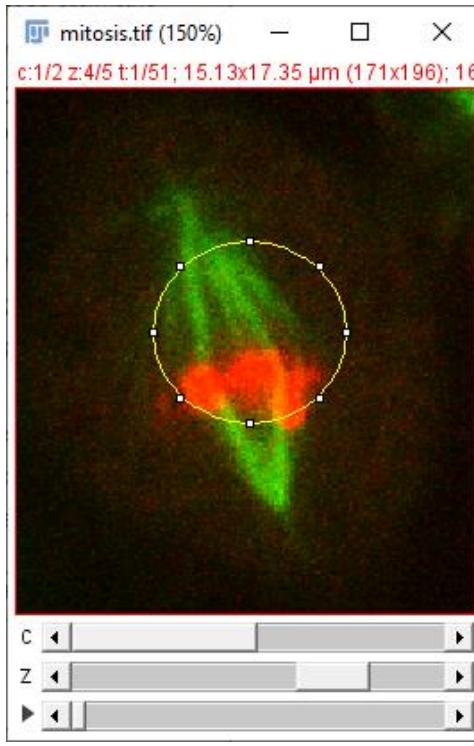
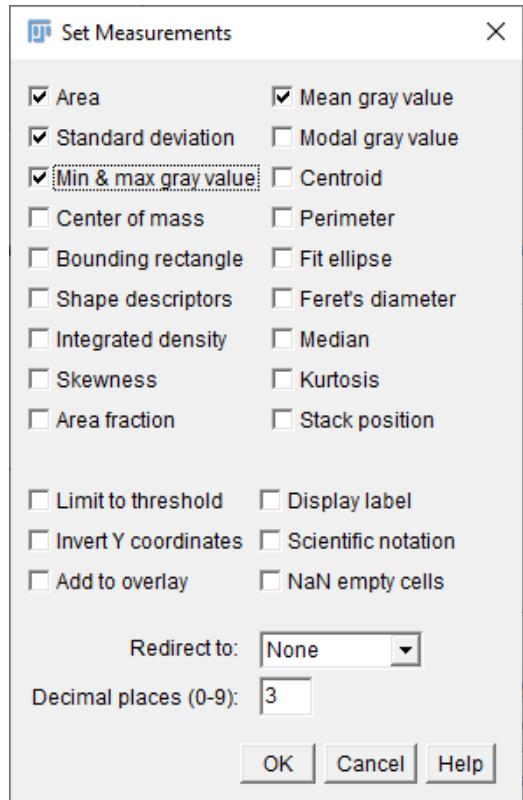


- Alternatively: hold ALT



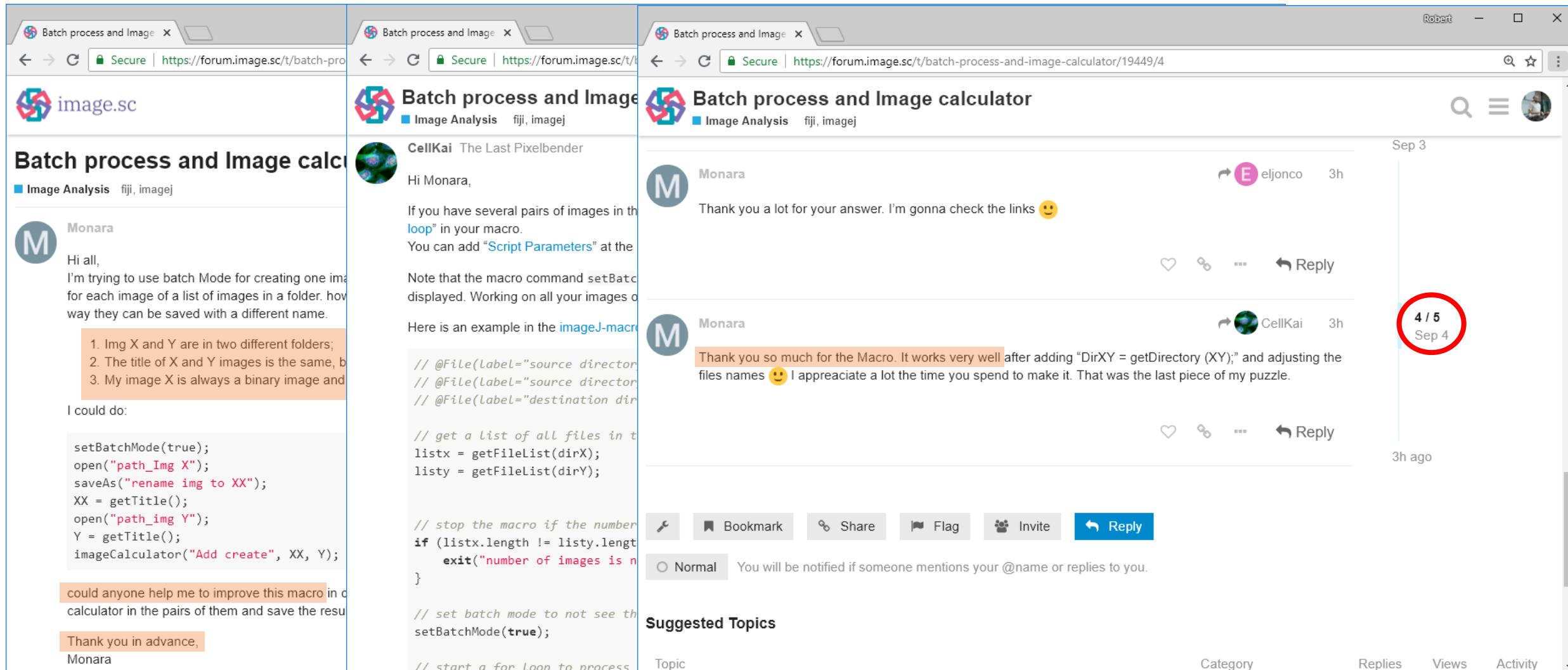
# Measurements

- ImageJ supports a range of measurements of regions of interest
- Configure them using the menu Analyse > Set measurements
- Measure them using the M key.



	Area	Mean	StdDev	Min	Max
1	30.201	4544.906	4844.726	1783	47693
2	30.201	5800.947	2331.946	2150	13738

- Visit <http://image.sc>, an online Q&A platform for the image sciences hosted by University Wisconsin and the Broad institute of Harvard and MIT.



The screenshot shows three browser tabs from the image.sc forum:

- Batch process and Image calculator**: Monara asks how to use batch mode to save images with different names. CellKai responds with a macro example.
- Batch process and Image calculator**: Monara thanks CellKai for the macro.
- Batch process and Image calculator**: CellKai replies, noting the macro works well after some adjustments.

A red circle highlights the "4 / 5" rating and the date "Sep 4" on the third post.

**Monara's message:**

Hi all,  
I'm trying to use batch Mode for creating one image for each image of a list of images in a folder. how way they can be saved with a different name.

1. Img X and Y are in two different folders;  
2. The title of X and Y images is the same, b  
3. My image X is always a binary image and

I could do:

```
setBatchMode(true);
open("path_Img X");
saveAs("rename img to XX");
XX = getTitle();
open("path_img Y");
Y = getTitle();
imageCalculator("Add create", XX, Y);
```

could anyone help me to improve this macro in calculator in the pairs of them and save the resu

Thank you in advance,  
Monara

**CellKai's response:**

CellKai The Last Pixelbender

Hi Monara,

If you have several pairs of images in the "loop" in your macro.  
You can add "Script Parameters" at the top of the macro.

Note that the macro command `setBatchMode(true)` is displayed. Working on all your images one by one.

Here is an example in the [imageJ-macros](#) repository:

```
// @File(Label="source directory")
// @File(Label="source directory")
// @File(Label="destination directory")

// get a List of all files in the source directory
listx = getFileList(dirX);
listy = getFileList(dirY);

// stop the macro if the number of files is not equal
if (listx.length != listy.length)
    exit("number of images is not equal");

// set batch mode to not see the progress bar
setBatchMode(true);

// start a for loop to process
```

**Monara's reply:**

Monara

Thank you a lot for your answer. I'm gonna check the links 😊

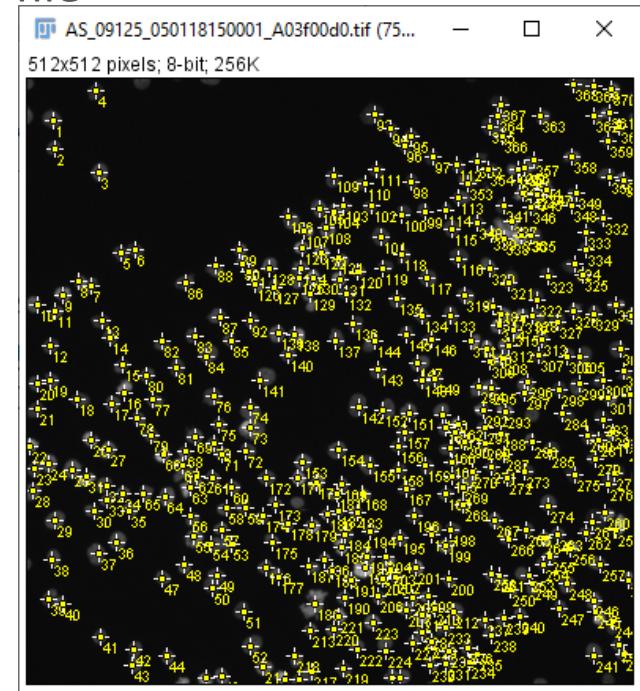
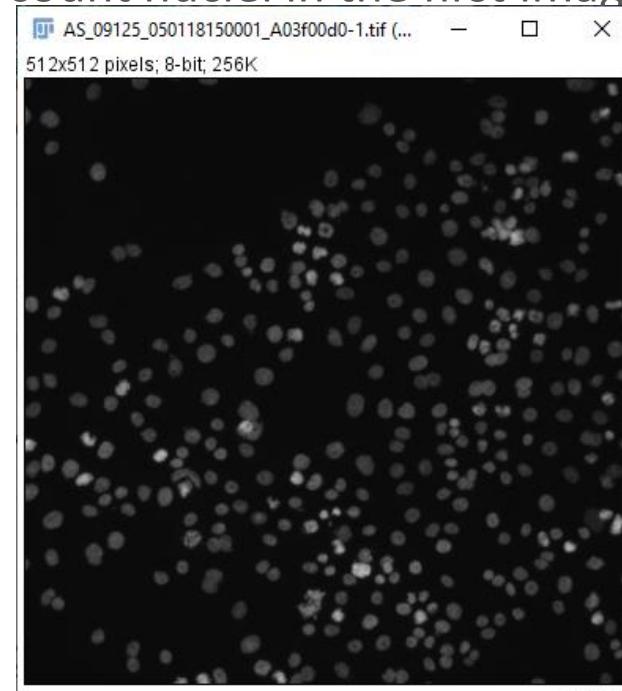
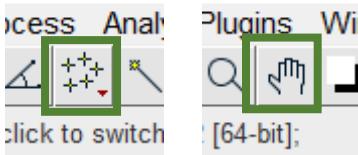
**CellKai's reply:**

CellKai

Thank you so much for the Macro. It works very well after adding "DirXY = getDirectory (XY);" and adjusting the file names 😊 I appreciate a lot the time you spend to make it. That was the last piece of my puzzle.

# Homework

- Download and install Fiji
  - <http://fiji.sc/Downloads>
- Download the BBBC001 dataset (TIF), unzip it and count nuclei in the first image file
  - <https://bbbc.broadinstitute.org/BBBC001>
  - AS\_09125\_050118150001\_A03f00d0.tif
- Hints:
  - Use zoom (+/- key)
  - Use the multipoint-tool and the pan tool

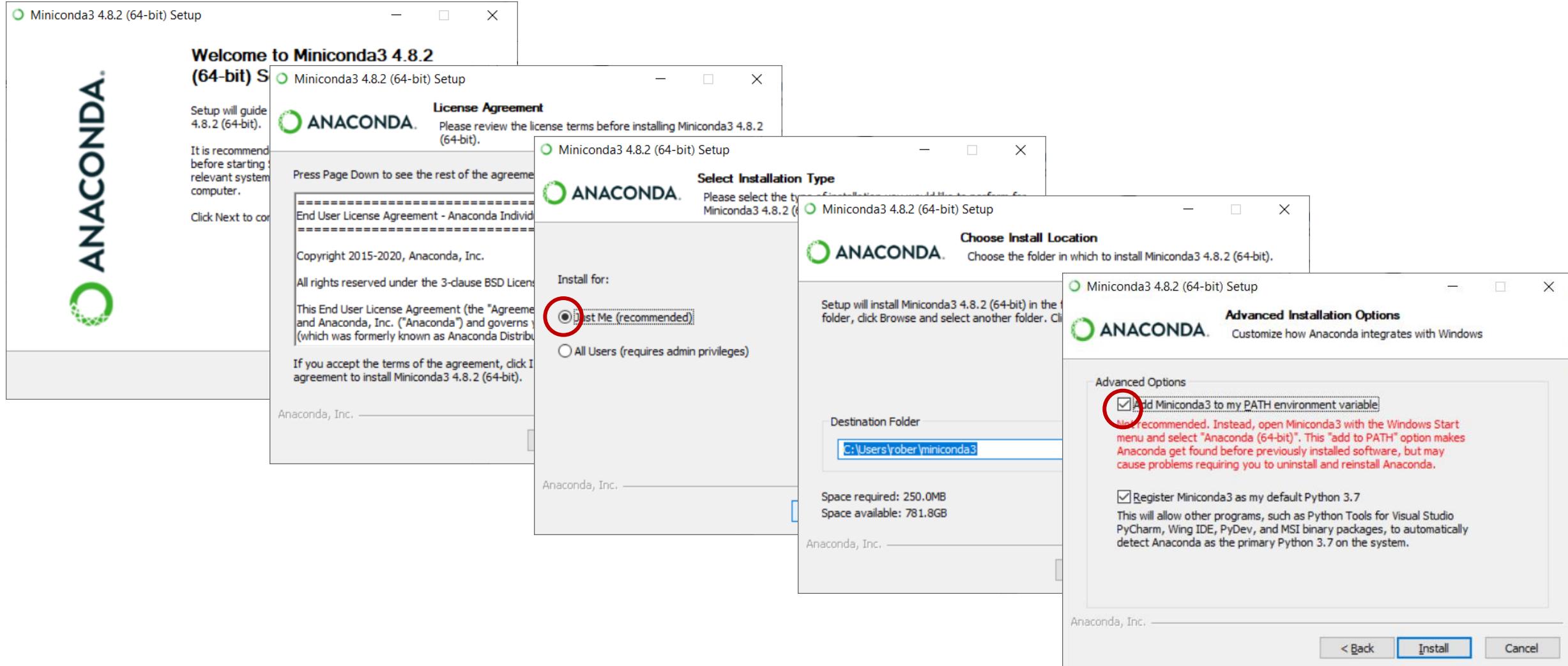


- Enter your count here: <https://docs.google.com/spreadsheets/d/1Ek-23Soro5XZ3y1kJHpvaTaa1f4n2C7G3WX0qddD-78/edit?usp=sharing>
- Be a good scientist: Don't cheat! ;-)
- We will repeat the cell count in the coming weeks with automated methods and compare!

# Homework: Install miniconda and test python

Detailed instructions:

- <https://github.com/BiAPoL/Bio-image Analysis with Python/blob/main/conda basics/01 conda environments.md>



Today, you learned

- *Bio-image analysis*
  - Quantitative
  - Objective
  - Reproducible
  - Repeatable
  - Reliable
- What are images, pixels, voxels, regions of interest
- Visualisation: Lookup tables
- Explore image data
- Do basic measurements
- Reminder: Use the etherpad to ask questions anonymously!

Coming up next

- Python programming I
  - Variables and operators
  - Lists, tuples and dictionaries
  - Tables

```
# going through arrays pair-wise
measurement_1 = [1, 9, 7, 1, 2, 8, 9, 2, 1, 7, 8]
measurement_2 = [4, 5, 5, 7, 4, 5, 4, 6, 6, 5, 4]

for m_1, m_2 in zip(measurement_1, measurement_2):
    print("Paired measurements: " + str(m_1) + " and " + str(m_2))
```

```
Paired measurements: 1 and 4
Paired measurements: 9 and 5
Paired measurements: 7 and 5
Paired measurements: 1 and 7
Paired measurements: 2 and 4
Paired measurements: 8 and 5
Paired measurements: 9 and 4
Paired measurements: 2 and 6
Paired measurements: 1 and 6
Paired measurements: 7 and 5
Paired measurements: 8 and 4
```