



Image Analysis of Biological Data using CellProfiler

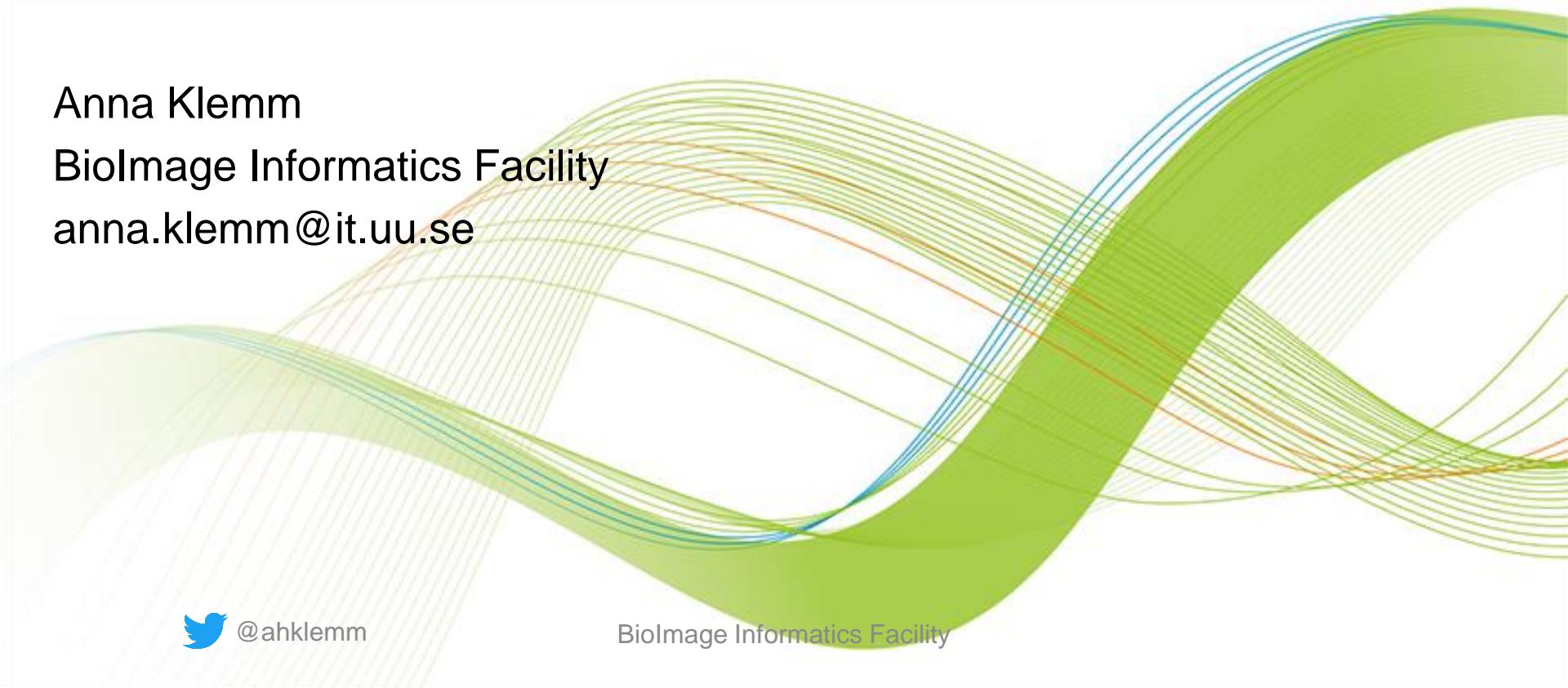
Anna Klemm

BioImage Informatics Facility

anna.klemm@it.uu.se



@ahklemm

A decorative graphic consisting of several overlapping, wavy lines in shades of green, blue, and orange, creating a sense of motion and depth.

BioImage Informatics Facility

Outline of the course

- Introduction to the data set and the image analysis question
- Step-By-Step workflow construction in CellProfiler
- Where to continue: online-resources

Material of this course

- Images and supporting files:
- Youtube-video:

If you get lost open

`catch-up/CellProfiler_Laminar_20200624_stepI.cppproj`

or

`catch-up/CellProfiler_Laminar_20200624_fullPipeline.cppproj`



Cell Atlas Aim:

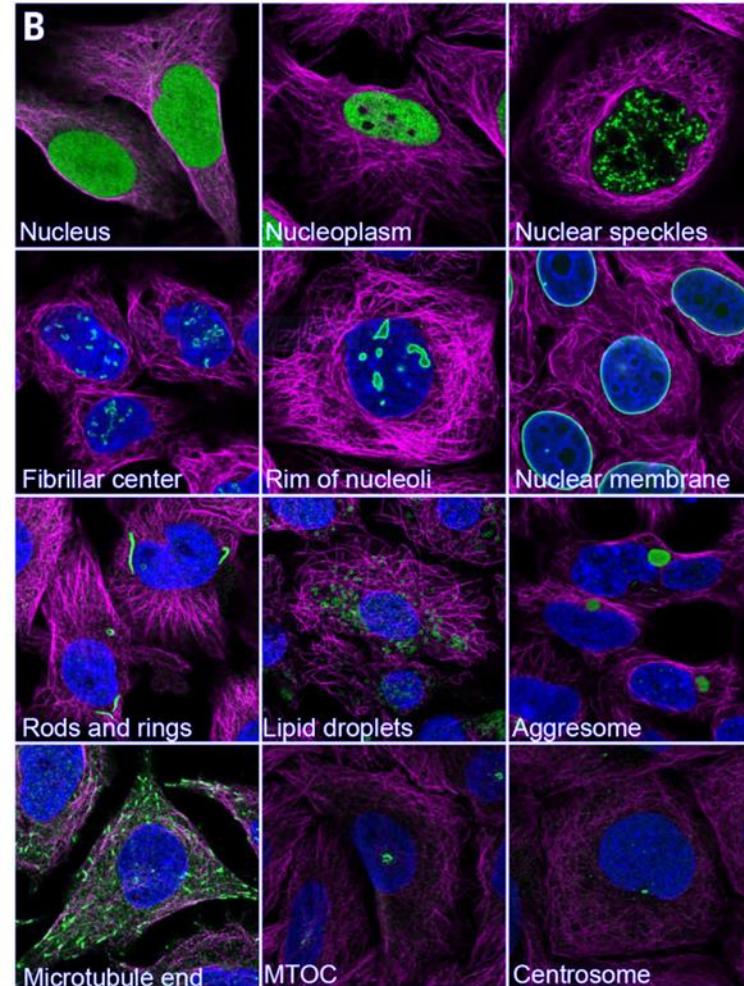
Determine the subcellular location of all cellular proteins.

Experimental Methods:

- Antibody generation against 12.000 human proteins
- Immunostaining, 22 cell lines
- Automated confocal microscopy
- → 82.152 images

Image Analysis Aim:

- Mapping 12.000 human proteins to 30 subcellular structures

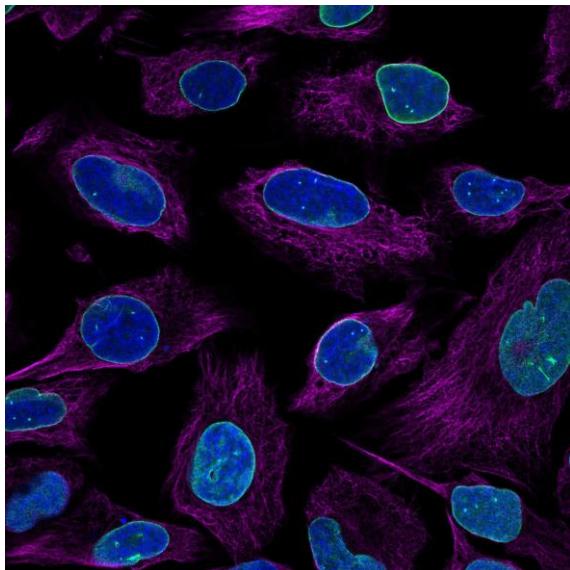


Adapted from Thul, P.J. et al. (2017). A subcellular map of the human proteome. Science 356.

The Data

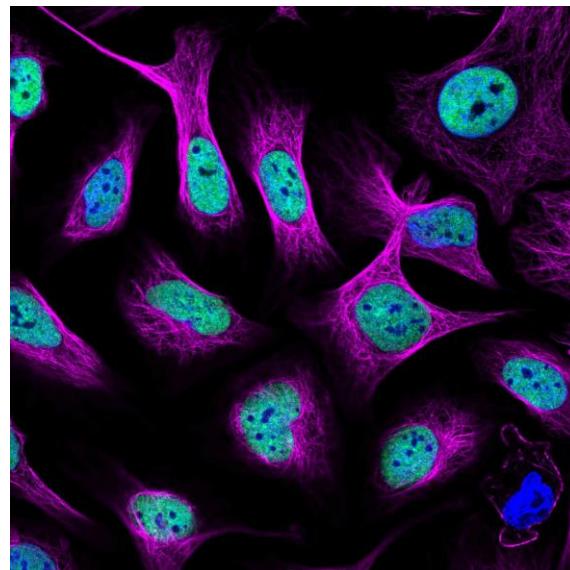
Raw images: Human Protein Atlas

www.proteinatlas.org/ENSG00000113368-LMNB1



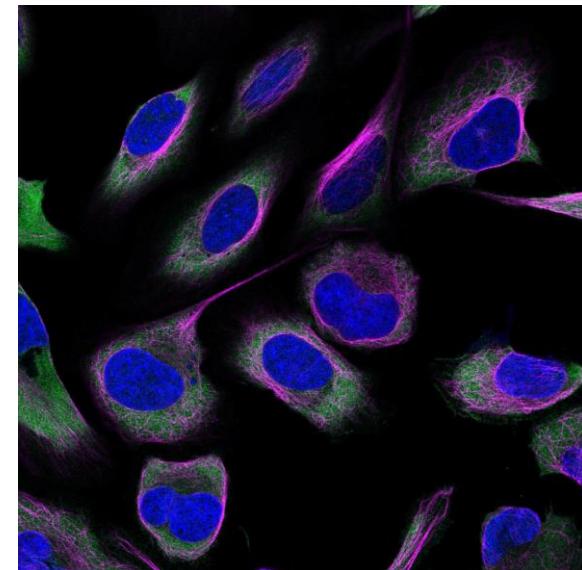
Nuclear Membrane
LMNB1

www.proteinatlas.org/ENSG00000270647-TAF15



Nucleus
TAF15

www.proteinatlas.org/ENSG0000164163-ABCE1



Cytosol
ABCE1

The Task:

Automatically determine the localization of the antibody-detected protein (green).
Microtubule (magenta) and nuclei (blue) should be used as reference channels.

STEP I: ACCUMULATION IN THE NUCLEAR MEMBRANE?

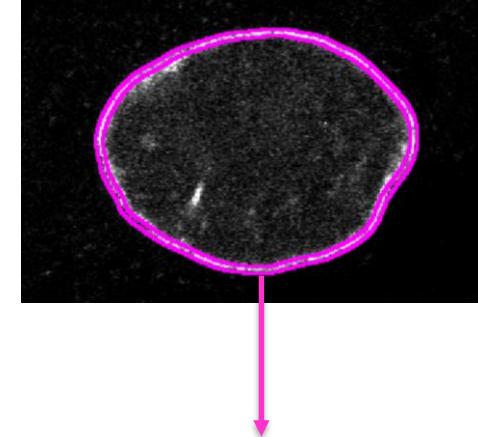
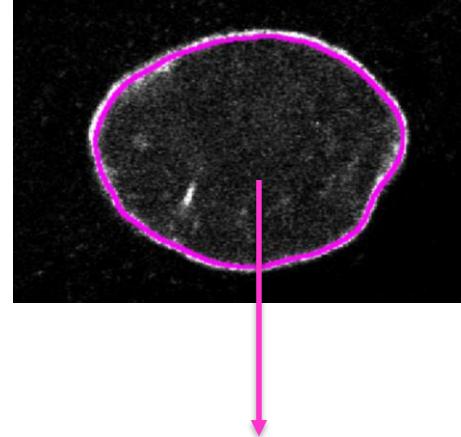
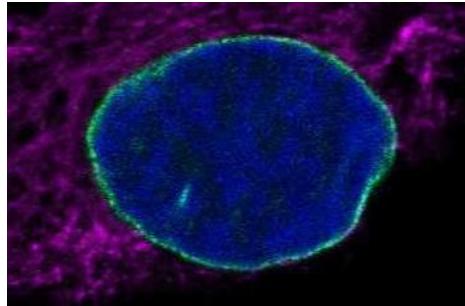


@ahklemm

BioImage Informatics Facility

The Aim: Quantify Signal Accumulation within the Nuclear Membrane

Image source: Human Protein Atlas
v19.proteinatlas.org/ENSG00000113368-LMNB1



Dataset:

- Subset of The Cell Atlas (Human Protein Atlas)
- 3 color stack:
microtubules (magenta),
protein detected by antibody (green), nuclei (blue)

$$\text{Ratio} = \frac{\text{Mean Intensity NucMembrane}}{\text{Mean Intensity Nucleus}}$$

Basic Workflow

- Split Images from a stack into three separate channels.
↓
- Get outlines of nuclei from the DAPI image
↓
- Get outlines of nuclear membrane
↓
- Measure intensity of the signal in both outlines
↓
- Calculate Ratio $I_{\text{NucMembrane}}/I_{\text{Nucleus}}$

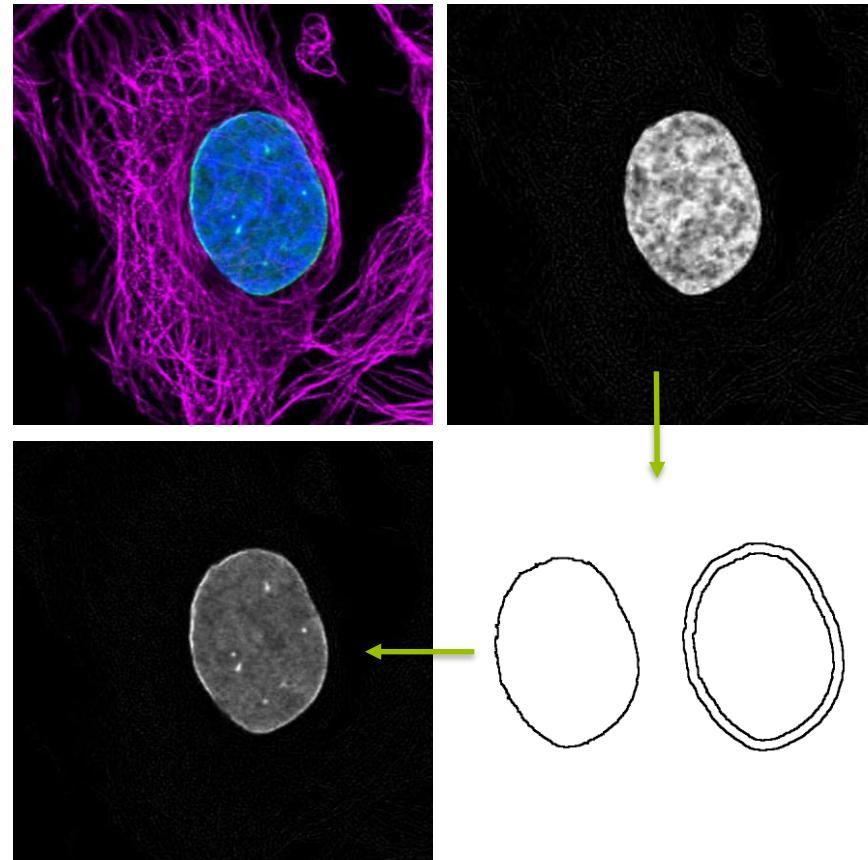


Image source: Human Protein Atlas
v19.proteinatlas.org/ENSG00000113368-LMNB1



- Developed by: Anne Carpenter lab, Broad Institute
- Published in 2006 (Carpenter et al, Genome Biology)
- Cited in more than **9000** scientific papers (June 2020)
- Launched **>100,000** times/year
- Originally based on Matlab, in 2009 completely re-written in Python
- Export pipelines as text files for publication

Cite it when you use it!

Carpenter AE et al. (2006)
CellProfiler: image analysis software for identifying and quantifying cell phenotypes.
Genome Biology 7:R100. PMID: 17076895

Check out details:

<https://cellprofiler.org/citations/>

CP INTERFACE AND FIRST STEPS

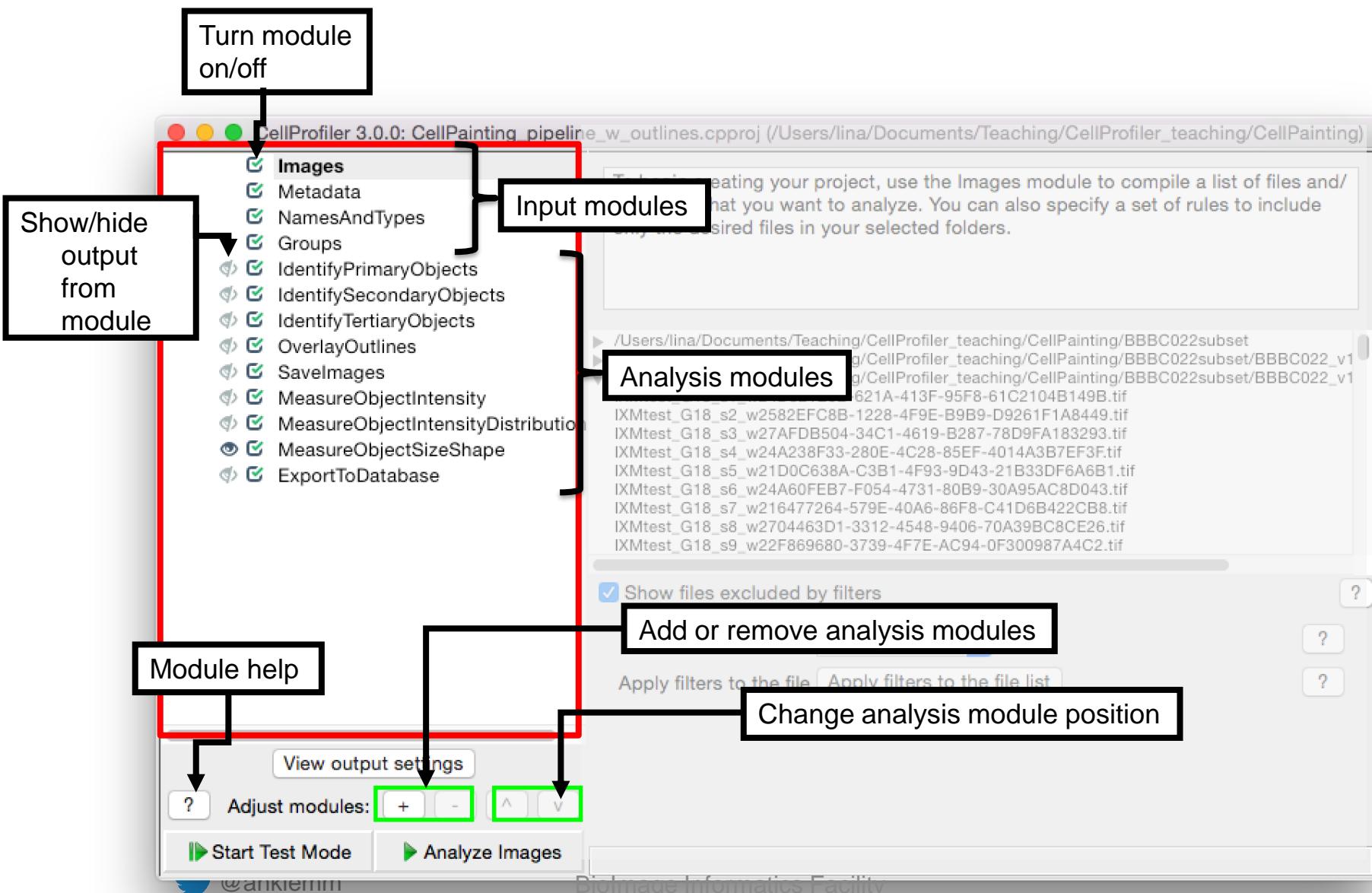


@ahklemm

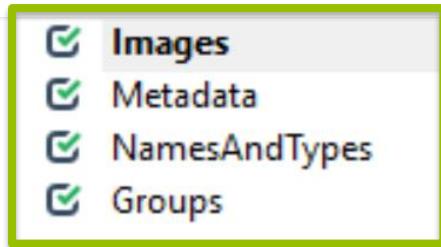
BioImage Informatics Facility

10

CellProfiler Interface



First modules: Import Images



This screenshot shows the 'Images' module settings. It displays a list of files from the path 'C:\Users\Anna\Documents\CellProfiler Teaching\202006'. The files listed are: 1772_B6_4.tif, 1772_B6_7.tif, 1773_H11_1.tif, 1773_H11_2.tif, 1894_G2_3.tif, 1894_G2_4.tif, 1894_G3_4.tif, 1894_G3_5.tif, 235_E7_3.tif, 418_F7_2.tif, and 551_C10_1.tif. Below the file list, there is a checkbox labeled 'Show files excluded by filters' which is checked. Underneath it are two buttons: 'Filter images? Images only' and 'Apply filters to the file list'.

Pipeline

Module: Images

We will use the default settings for the modules *Metadata* and *Groups* in this example.

This screenshot shows the 'NamesAnd Types' module settings. At the top, it says 'Assign a name to All images'. Below that is a section for 'Process as 3D?' with 'No' selected. Then it asks 'Select the image type Color image'. A 'Name to assign these images' field contains 'raw_image'. Under 'Set intensity range from' is a dropdown set to 'Image metadata'. The main area shows a table with columns 'Update' and 'raw_image'. The table lists eight files with their corresponding names assigned:

Update	raw_image
1	1773_H11_1.tif
2	1773_H11_2.tif
3	1894_G3_4.tif
4	1894_G3_5.tif
5	551_C10_1.tif
6	551_C10_2.tif
7	711_D6_1.tif
8	711_D6_2.tif
9	702_AE_1.tif

Module: NamesAnd Types

Basic Workflow

- Split Images from a stack into three separate channels.
↓
- Get outlines of nuclei from the DAPI image
↓
- Get outlines of nuclear membrane
↓
- Measure intensity of the signal in both outlines
↓
- Calculate Ratio $I_{\text{NucMembrane}}/I_{\text{Nucleus}}$

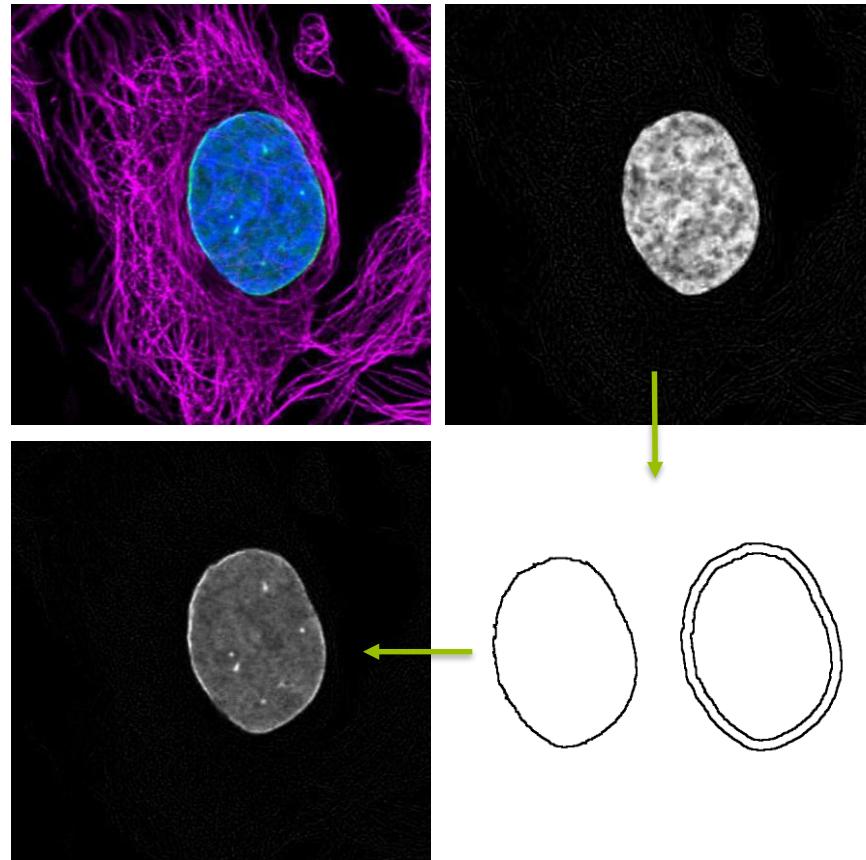


Image source: Human Protein Atlas
v19.proteinatlas.org/ENSG00000113368-LMNB1

ColorToGray - Solution

Select the input image (from NamesAndTypes)

Conversion method

Image type

Channel number

Image name

Channel number

Image name

Channel number

Image name

Module: ColorToGray

- Add and configure the ColorToGray module
- Start the test mode and inspect the outcome of the module

Module

IDENTIFY PRIMARY OBJECTS



@ahklemm

BioImage Informatics Facility

16

Module: IdentifyPrimaryObjects

Use advanced settings? Yes No

Select the input image DAPI (from ColorToGray #05)

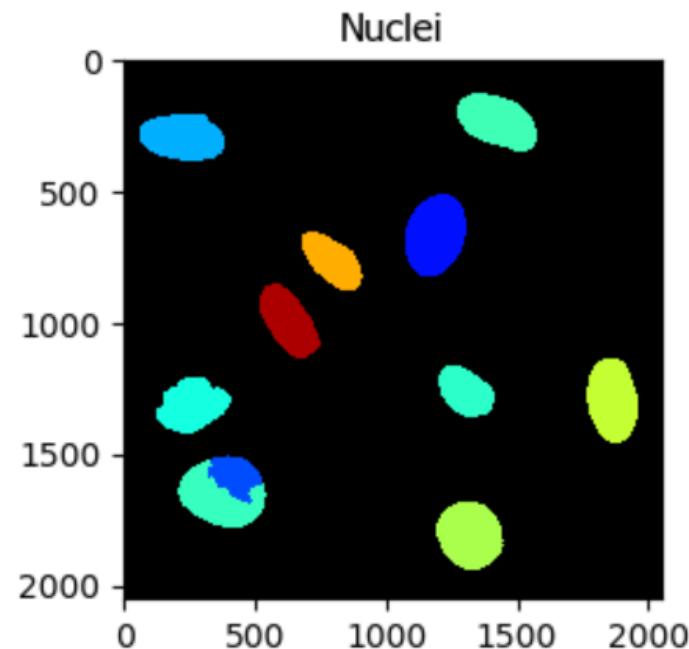
Name the primary objects to be identified Nuclei

Typical diameter of objects, in pixel units (Min,Max) 150 500

Discard objects outside the diameter range? Yes No

Discard objects touching the border of the image? Yes No

Module: ColorToGray



Border of the Image

Use advanced settings? Yes No

Select the input image DAPI (from ColorToGray #05)

Name the primary objects to be identified Nuclei

Typical diameter of objects, in pixel units (Min,Max) 150 500

Discard objects outside the diameter range? Yes No

Discard objects touching the border of the image? Yes No

Discard objects touching the border of the image

Partial objects could result in incorrect measurements

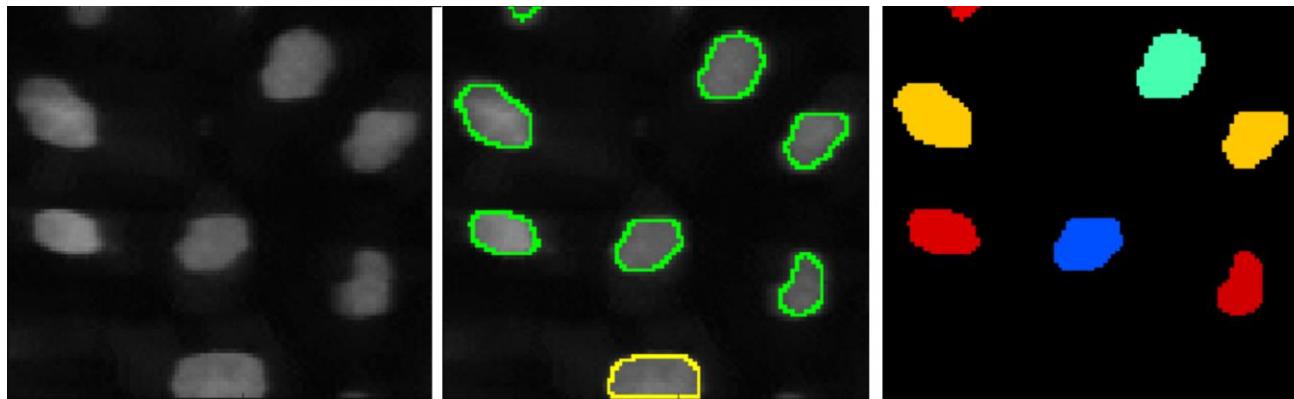


Image Source: C. Wählby, BIIF, SciLifeLab

IdentifyPrimaryObjects- Exercise

- Add the IdentifyPrimaryObjects module and explore the settings
- Which steps would you need to do in Fiji? – Discuss within your group





Which steps would you need to do in Fiji?

- Gaussian Blur
- Threshold
- Fill holes
- "Declumping" of objects by using Watershed
- Analysis of connected components ("Analyze Particles")

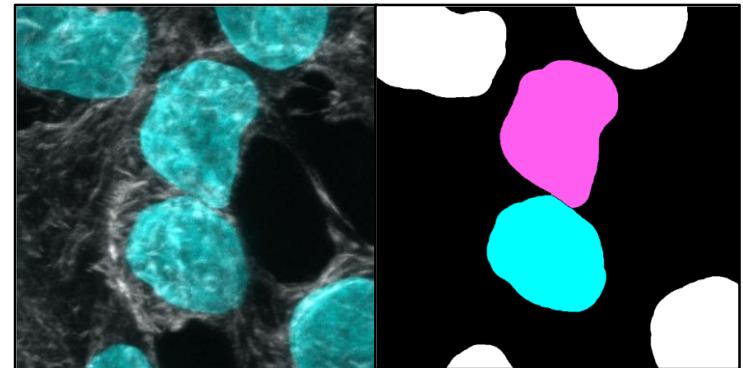


Image source: CFBim, BMC, LMU Munich

Use advanced settings? Yes No

Select the input image DAPI (from ColorToGray #05)

Name the primary objects to be identified Nuclei

Typical diameter of objects, in pixel units (Min,Max) 150 500

Discard objects outside the diameter range? Yes No

Discard objects touching the border of the image? Yes No

IdentifyPrimaryObjects

ADVANCED SETTINGS



@ahklemm

BioImage Informatics Facility

De-Clump Objects

Select the input image DAPI (from ColorToGrayscale)

Name the primary objects to be identified Nuclei

Typical diameter of objects, in pixel units (Min,Max) 150 500

Discard objects outside the diameter range? Yes No

Discard objects touching the border of the image? Yes No

Threshold strategy Global

Thresholding method Minimum cross entropy

Threshold smoothing scale 1.3488

Threshold correction factor 1.0

Lower and upper bounds on threshold 0.0 1.0

Method to distinguish clumped objects Intensity

Method to draw dividing lines between clumped objects Intensity

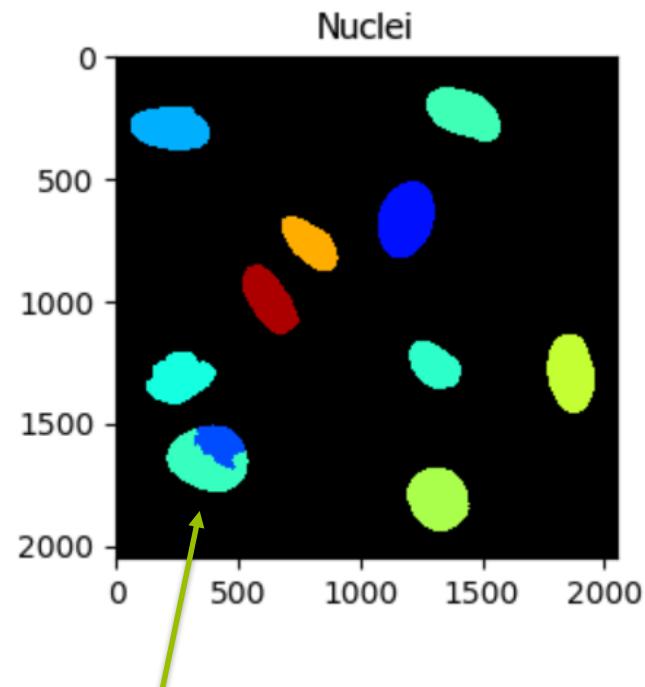
Automatically calculate size of smoothing filter for declumping? Yes No

Automatically calculate minimum allowed distance between local maxima? Yes No

Speed up by using lower-resolution image to find local maxima? Yes No

Fill holes in identified objects? After both thresholding and declumping

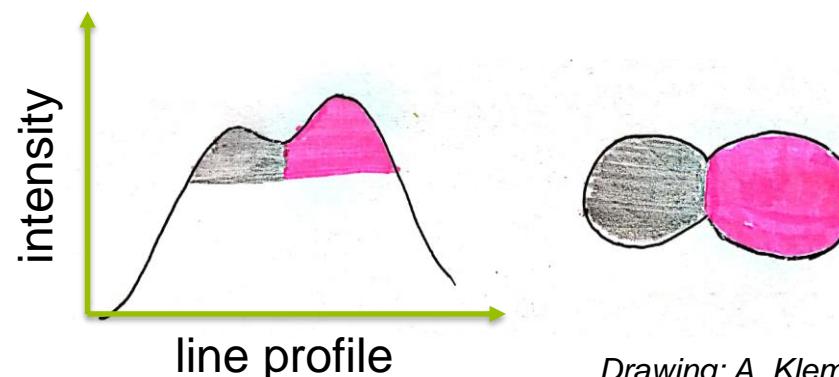
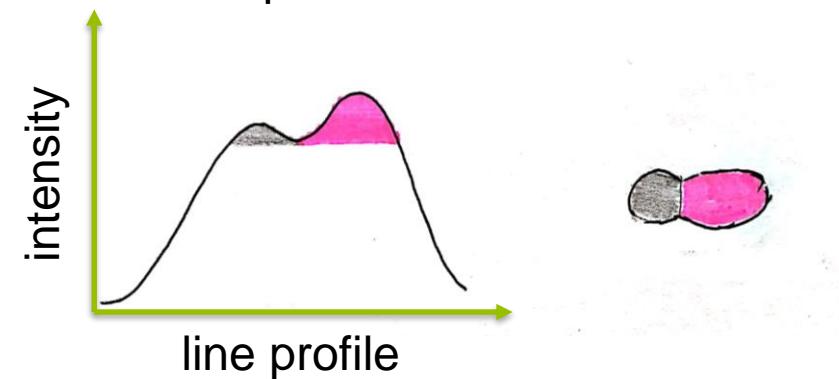
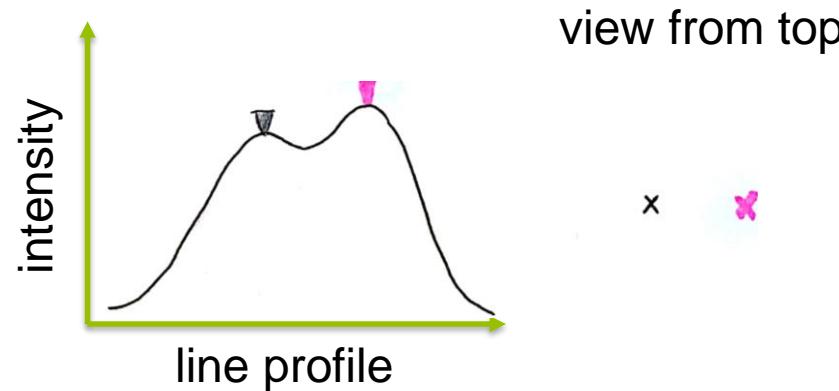
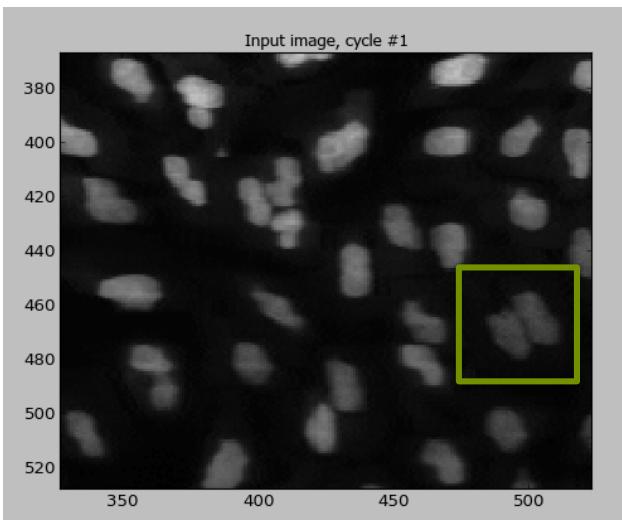
Handling of objects if excessive number of objects identified Continue



IdentifyPrimaryObjects - Watershed

Method to distinguish clumped objects: Intensity

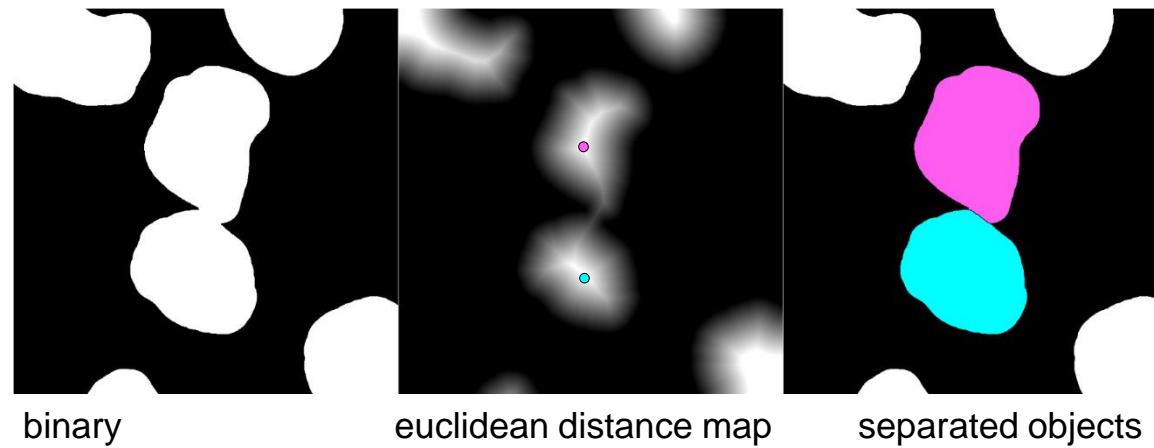
Image Source: C. Wählby, BIIF, SciLifeLab



IdentifyPrimaryObjects - Watershed

Method to distinguish clumped objects: Shape

Image source: CFBim, BMC, LMU Munich



Question 1

What is the CellProfiler Analog to the Fiji-command /Process/Binary/Watershed?

- a. Distinguish clumped objects by **intensity**
- b. Distinguish clumped objects by **shape**

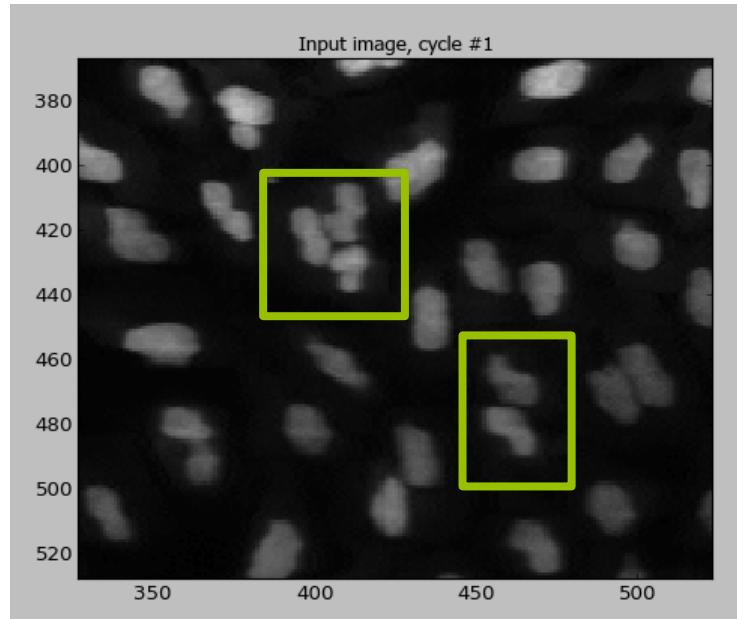
Question 2

When do you expect **intensity-based declumping** to work well, when **shape-based declumping**?

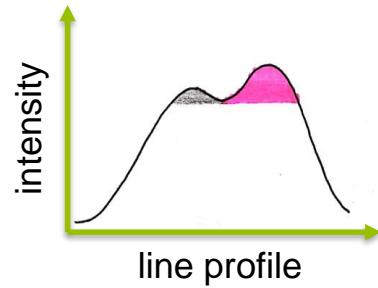


Separating Touching Objects

Image Source: C. Wählby, BIIF, SciLifeLab



- **Intensity:** Works best if objects are brighter at center, dimmer at edges
- **Shape:** Works best if objects have indentations where clumps touch (esp. if objects are round)

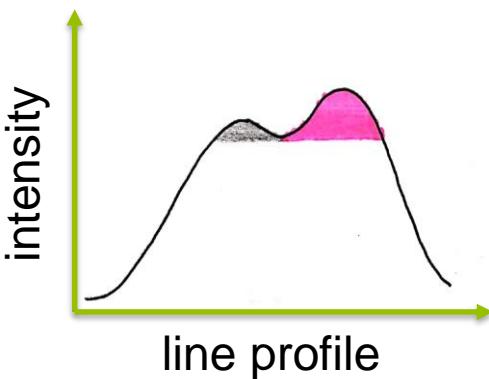


TUNING THE WATERSHED

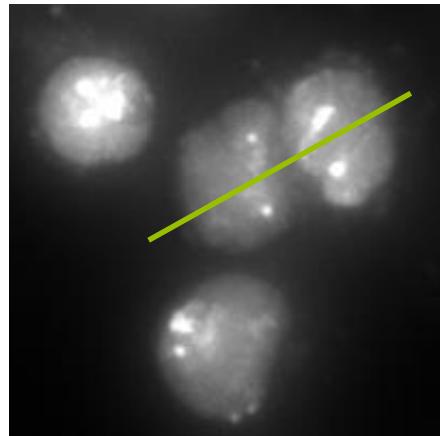
Drawing: Klemm, SciLifeLab

Theory vs. Real Life

Theory



Data

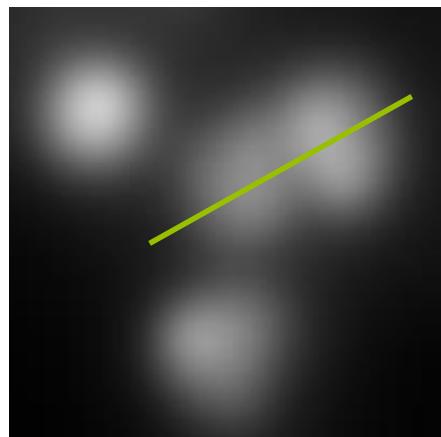


raw data

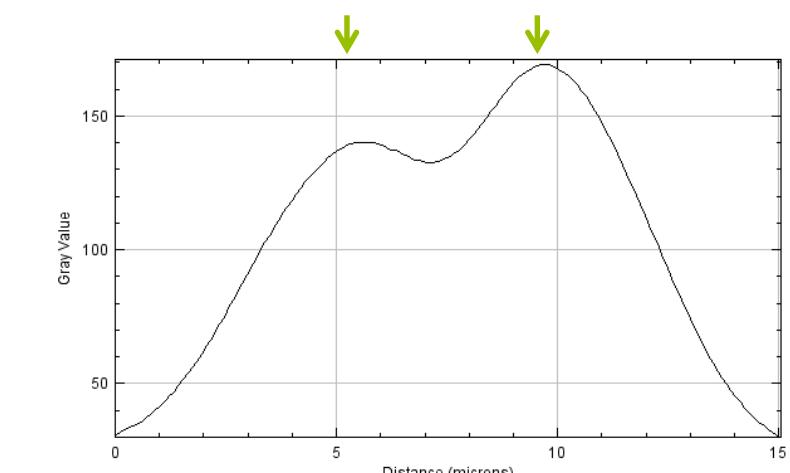
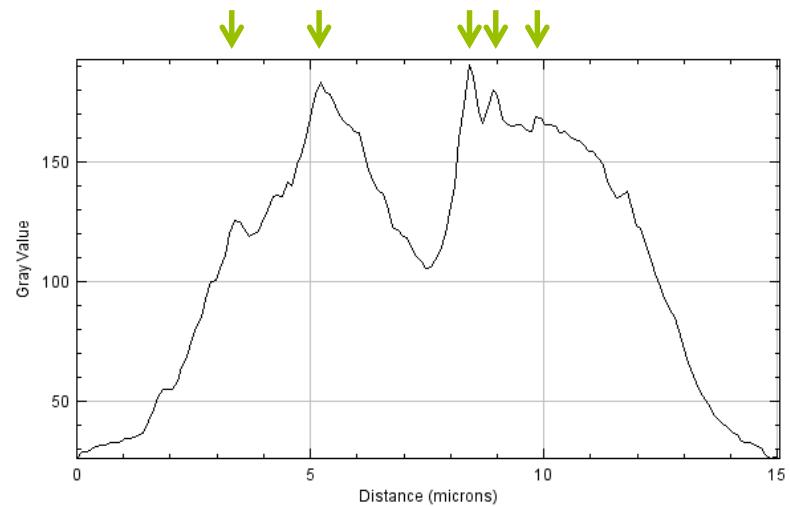
Identify Primary Objects
typical diameter affects size of smoothening filter and minimal allowed distance between maxima.

Drawing: A. Klemm, SciLifeLab

Image Source:
Klemm/Hasse/Sarov, MPI-CBG



Gaussian filtered



Theory vs. Real Life

Select the input image DAPI (from ColorToGrayscale)

Name the primary objects to be identified Nuclei

Typical diameter of objects, in pixel units (Min,Max) 150 500

Discard objects outside the diameter range? Yes No

Discard objects touching the border of the image? Yes No

Threshold strategy Global

Thresholding method Minimum cross entropy

Threshold smoothing scale 1.3488

Threshold correction factor 1.0

Lower and upper bounds on threshold 0.0 1.0

Method to distinguish clumped objects Intensity

Method to draw dividing lines between clumped objects Intensity

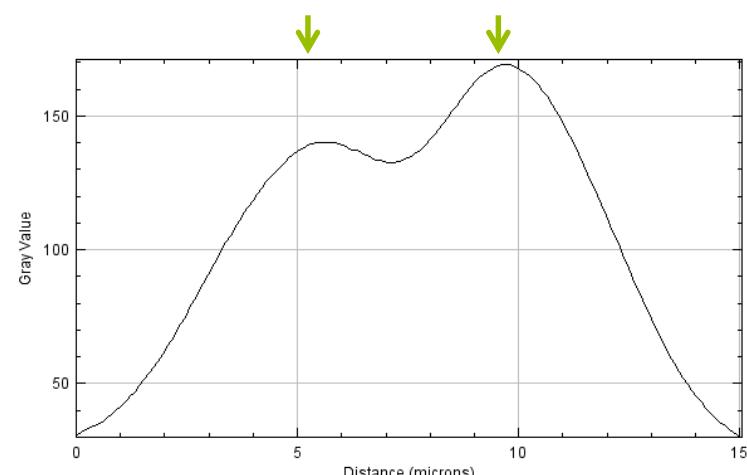
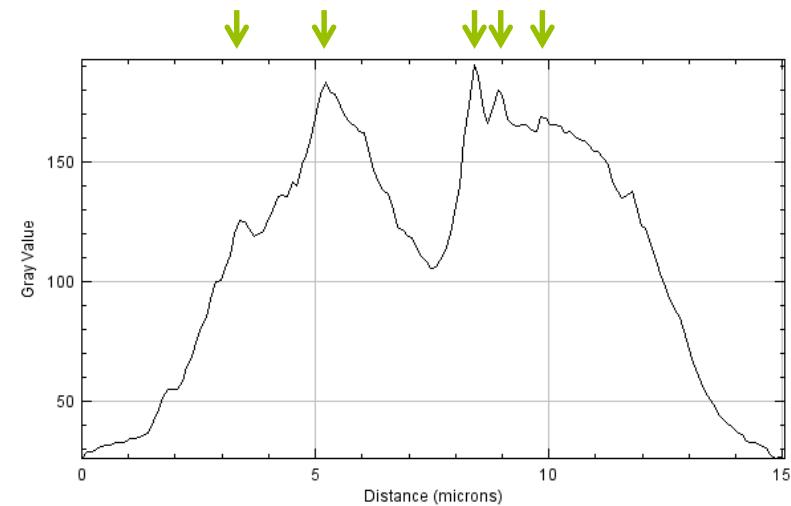
Automatically calculate size of smoothing filter for declumping? Yes No

Automatically calculate minimum allowed distance between local maxima? Yes No

Speed up by using lower-resolution image to find local maxima? Yes No

Fill holes in identified objects? After both thresholding and declumping

Handling of objects if excessive number of objects identified Continue



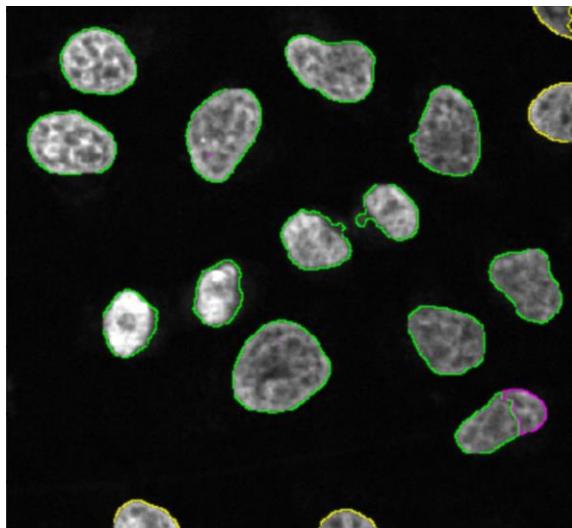
@ahklemm

BioImage Informatics Facility

Tuning IdentifyPrimaryObjects

Problem:

- High intra-nuclear intensity variations
- High inter-nuclear size variations
- Many non-round nuclei

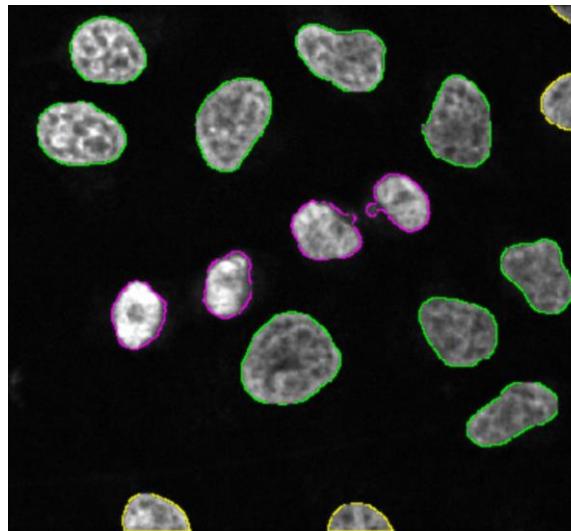


Typical diameter: 40, 200
Smoothing is not sufficient

Image Source:
Klemm/Hasse/Sarov, MPI-CBG

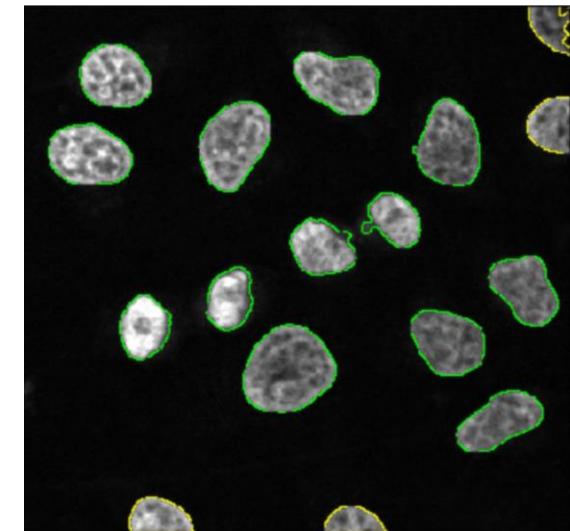
Solution:

Non-automatic adjustment of the size of the smoothing filter.



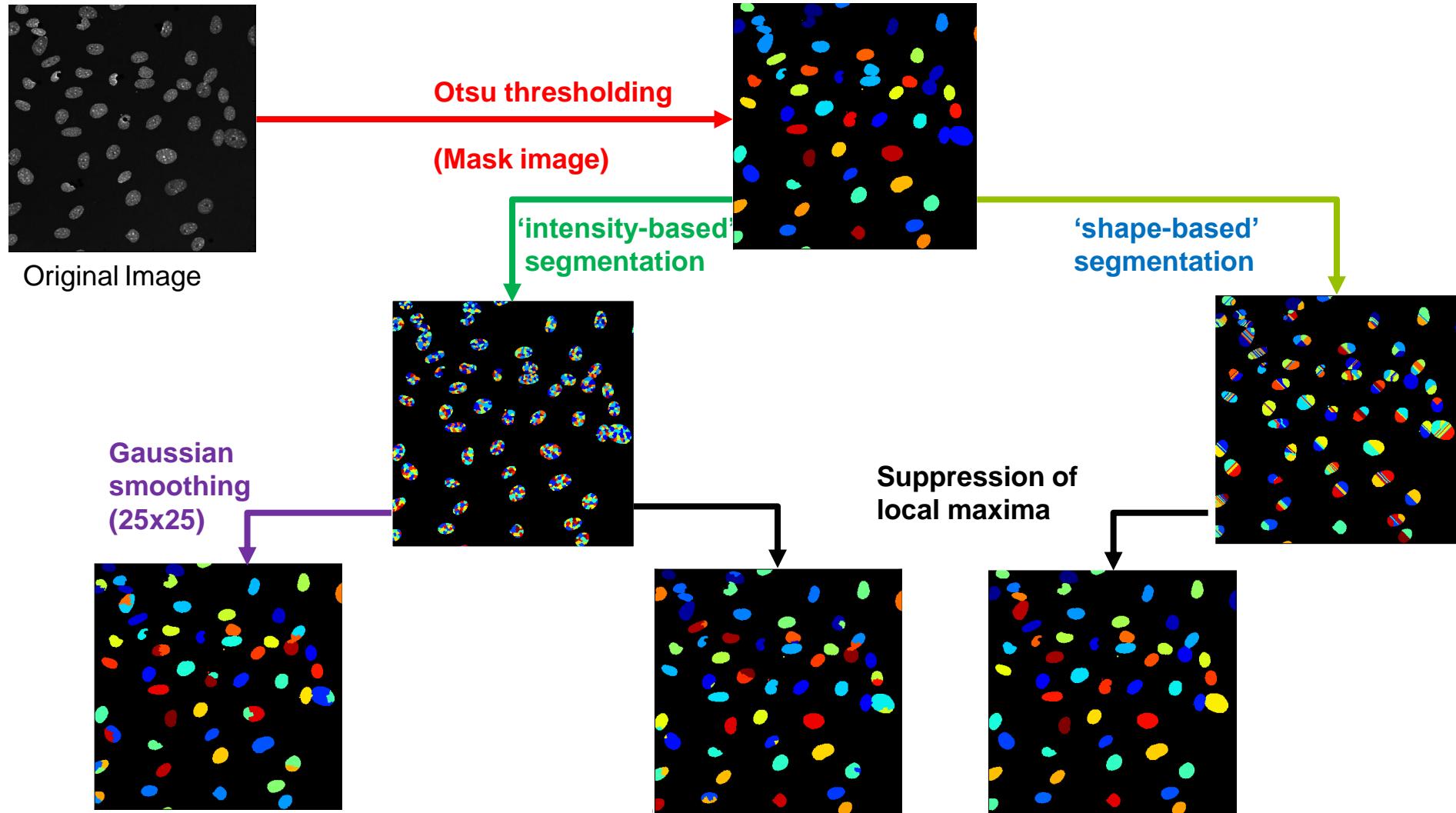
Typical diameter: 50, 200
Smoothing is sufficient, but small objects are discarded

BioImage Informatics Facility



Typical diameter: 40, 200
Manually set smoothing filter size to 50.

Tuning Separation



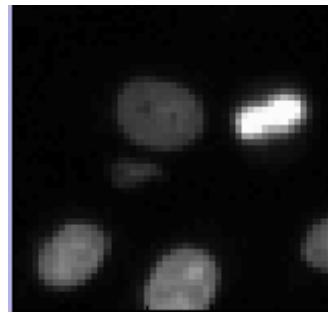
@ahklemm

BioImage Informatics Facility

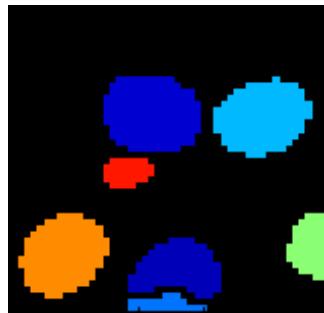
C. Wählby et al Journal of Microscopy, 2004

Additional Separation Settings

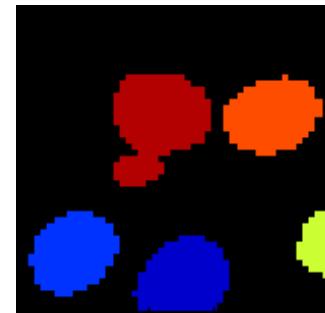
Image Source: C. Wählby, BIIF, SciLifeLab



Original image



Smoothing filter
size = 4



Smoothing filter
size = 8

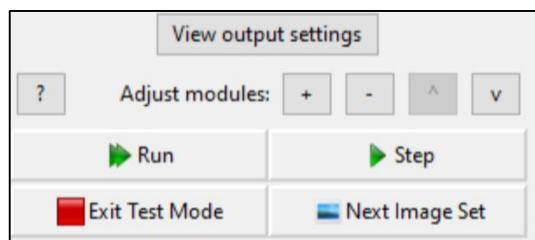
Adjusting parameters may introduce more error than it solves.

The proper settings are usually a matter of trial and error

- The automatic settings are a good starting point, though
- The error often influences the dataset in an unbiased way
- Avoid bias towards one cell-line/condition, e.g. mutants with larger nuclei

IdentifyPrimaryObjects - Exercise

- Open the Advanced Settings of IdentifyPrimaryObjects
- Start the Test Mode
- Try out different parameters (Declumping by Intensity/Shape; Automatic calculate smoothing filter – no, ...)
- See which parameters fit best
- Try out the parameters for the next image by pressing „Next Image Set“



IdentifyPrimaryObjects - Discussion

Module: IdentifyPrimaryObjects

Use advanced settings? Yes No

Select the input image DAPI (from ColorToGray #05)

Name the primary objects to be identified Nuclei

Typical diameter of objects, in pixel units (Min,Max) 150 500

Discard objects outside the diameter range? Yes No

Discard objects touching the border of the image? Yes No

Threshold strategy Global

Thresholding method Minimum cross entropy

Threshold smoothing scale 1.3488

Threshold correction factor 1.0

Lower and upper bounds on threshold 0.0 1.0

Method to distinguish clumped objects Shape

Method to draw dividing lines between clumped objects Intensity

Automatically calculate size of smoothing filter for declumping? Yes No

Automatically calculate minimum allowed distance between objects? Yes No

Speed up by using lower-resolution image to find local maxima? Yes No

Fill holes in identified objects? After both thresholding and declumping

Handling of objects if successive number of objects identified Continue

BACK TO THE WORKFLOW...



@ahklemm

BioImage Informatics Facility

35

Basic Workflow

- Split Images from a stack into three separate channels.
↓
- Get outlines of nuclei from the DAPI image
↓
- Get outlines of nuclear membrane
↓
- Measure intensity of the signal in both outlines
↓
- Calculate Ratio $I_{\text{NucMembrane}}/I_{\text{Nucleus}}$

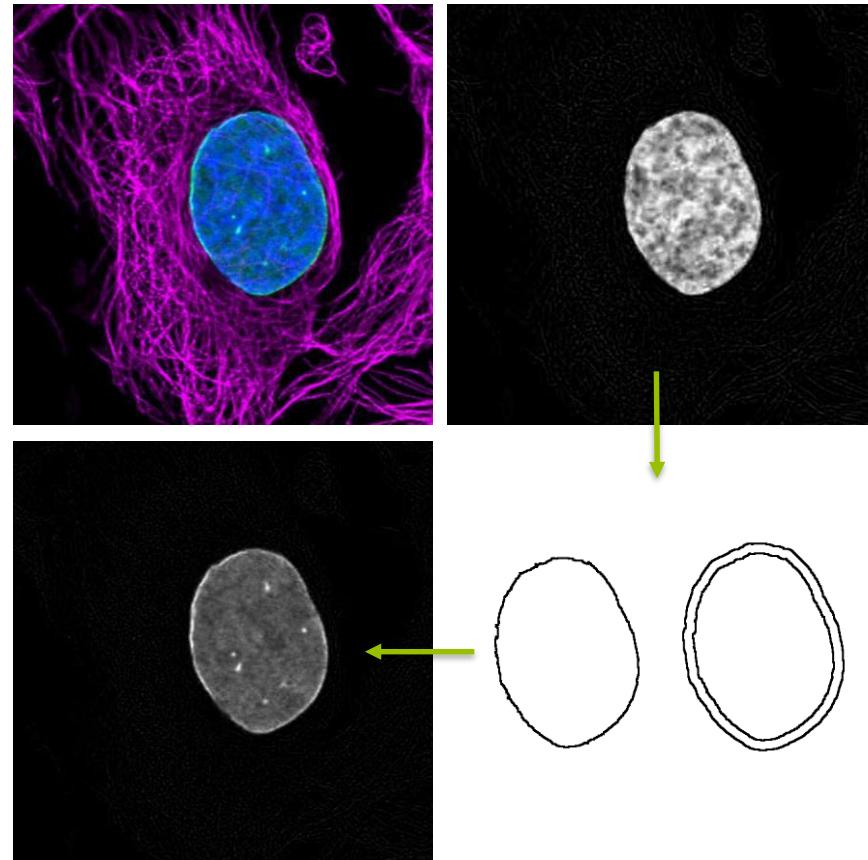


Image source: Human Protein Atlas
v19.proteinatlas.org/ENSG00000113368-LMNB1

Creating a selection around the nuclear membrane

ExpandOrShrinkObjects

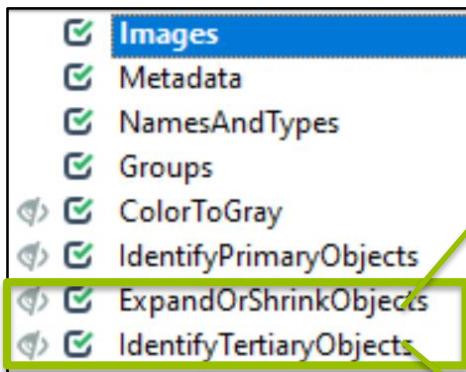
Select the input objects **Nuclei** (from IdentifyPrimaryObjects #06)

Name the output objects **ShrunkenNuclei**

Select the operation **Shrink objects by a specified number of pixels**

Number of pixels by which to expand or shrink **7**

Fill holes in objects so that all objects shrink to a single point? Yes No



IdentifyTertiaryObjects

Select the larger identified objects **Nuclei** (from IdentifyPrimaryObjects #06)

Select the smaller identified objects **ShrunkenNuclei** (from ExpandOrShrinkObjects #07)

Name the tertiary objects to be identified **nuclearMembrane**

Shrink smaller object prior to subtraction? Yes No



Calculating the Ratio $I_{\text{mean NucMembrane}} / I_{\text{mean Nucleus}}$

1)

Select an image to measure **Signal** (from ColorToGray #05)
Add another image

Select objects to measure **Nuclei** (from IdentifyPrimaryObjects #06)

Select objects to measure **nuclearMembrane** (from IdentifyTertiaryObjects #09)
Remove this object
Add another object

2)

Name the output measurement **Ratio_Membrane_Nuclei**
Operation **Divide**

Select the numerator measurement type **Object**
Select the numerator objects **nuclearMembrane** (from IdentifyTertiaryObjects #09)
Category: **Intensity**
Select the numerator measurement **MeanIntensity**
Image: **Signal**

Multiply the above operand by **1.0**
Raise the power of above operand by **1.0**

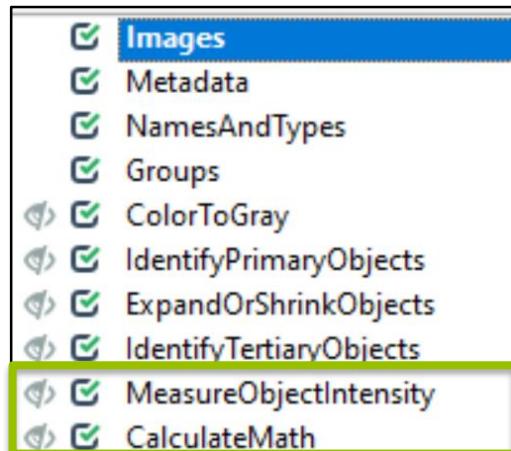
Select the denominator measurement type **Object**
Select the denominator objects **Nuclei** (from IdentifyPrimaryObjects #06)
Category: **Intensity**
Select the denominator measurement **MeanIntensity**
Image: **Signal**

Multiply the above operand by **1.0**
Raise the power of above operand by **1.0**

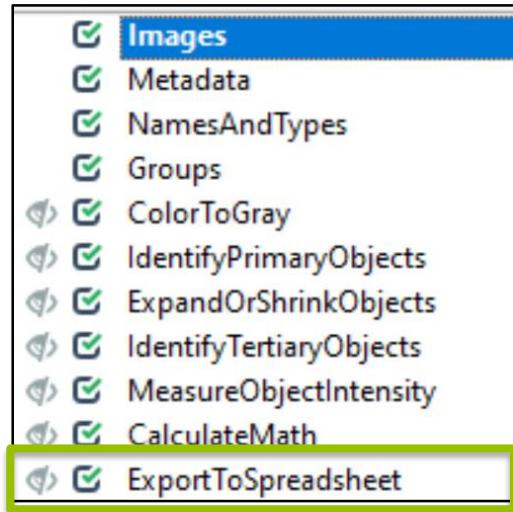
Take log10 of result? Yes No
Multiply the result by **1.0**
Raise the power of result by **1.0**
Add to the result **0.0**

Constrain the result to a lower bound? Yes No
Constrain the result to an upper bound? Yes No

Before we can calculate the ratio, we need to first measure the intensities!



ExportToSpreadsheet



For quick image-based analysis.
This will calculate the mean ratio
(INucMemb/INuc) for all nuclei.

A screenshot of the 'ExportToSpreadsheet' dialog box showing detailed export settings. The right side of the dialog contains various configuration options:

- Select the column delimiter: Comma (",")
- Output file location: Default Output Folder (C:\Users\annkl878\Doc)
- Add a prefix to file names? (radio buttons: Yes, No; Yes is selected)
- Filename prefix: results_
- Overwrite existing files without warning? (radio buttons: Yes, No; No is selected)
- Add image metadata columns to your object data file? (radio buttons: Yes, No; No is selected)
- Representation of Nan/Inf: NaN
- Select the measurements to export (radio buttons: Yes, No; Yes is selected)
- Press button to select measurements (button)
- Calculate the per-image mean values for object measurements? (radio buttons: Yes, No; Yes is selected)
- Calculate the per-image median values for object measurements? (radio buttons: Yes, No; No is selected)
- Calculate the per-image standard deviation values for object measurements? (radio buttons: Yes, No; Yes is selected)
- Create a GenePattern GCT file? (radio buttons: Yes, No; No is selected)
- Export all measurement types? (radio buttons: Yes, No; Yes is selected)

Three green arrows point from the 'ExportToSpreadsheet' option in the first screenshot to the 'Select the measurements to export' radio buttons, the 'Calculate the per-image mean values for object measurements?' radio buttons, and the 'Export all measurement types?' radio buttons in this screenshot.

If you want, select only a subset to
measurements to be exported

Catch-up with the previous steps

- Create the selection covering the nuclear membrane
- Measure the intensities in the selection of nuclear membrane and nuclei
- Calculate the ratio INucMembrane /INuc
- Export to Spreadsheet (for the beginning choose "Select the measurements to export: no")



Workflow

CREATING A CONTROL IMAGE



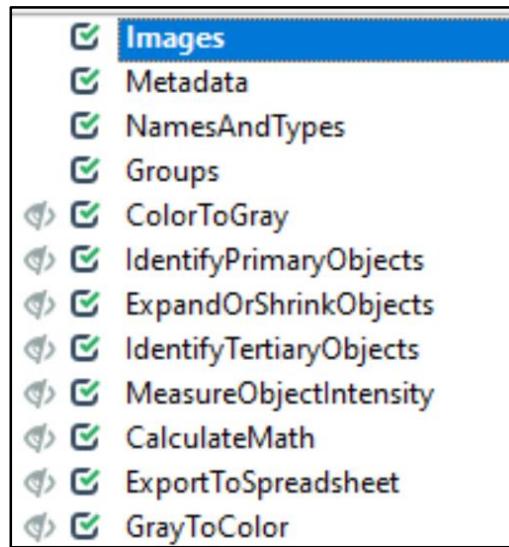
@ahklemm

BioImage Informatics Facility

41

Creating a RGB Image

Input images are 3-channel stack.
We generate a RGB image to define
the color order.



Select a color scheme **RGB** ▾

Select the image to be colored red **MT** ▾ (from ColorToGray #05)

Select the image to be colored green **Signal** ▾ (from ColorToGray #05)

Select the image to be colored blue **DAPI** ▾ (from ColorToGray #05)

Name the output image **raw_RGB**

Relative weight for the red image **1.0**

Relative weight for the green image **1.0**

Relative weight for the blue image **1.0**

Creating and Saving the Control Image

OverlayOutlines

Display outlines on a blank image? Yes No

Select image on which to display outlines raw_RGB (from GrayToColor #13)

Name the output image RawImage_Overlay

Outline display mode Color

How to outline Inner

Select objects to display nuclearMembrane (from IdentifyTertiaryObjects #09)

Select outline color

Add another outline

DisplayDataOnImage

Display object or image measurements? Object

Select the input objects Nuclei (from IdentifyPrimaryObjects)

Measurement to display Category: Number

Measurement: Object_Number

Display background image? Yes No

Select the image on which to display the measurements RawImage_Overlay (from OverlayOutlines)

Display mode Text

Text color

Font size (points) 20

Number of decimals 0

Annotation offset (in pixels) 0

Name the output image that has the measurements displayed RawImage_Overlay_ObjN

Image elements to save Image

SaveImages

Select the type of image to save Image

Select the image to save RawImage_Overlay_ObjN (from OverlayOutlines)

Select method for constructing file names From image filename

Select image name for file prefix raw_image (from NamesAndTypes)

Append a suffix to the image file name? Yes No

Text to append to the image name _Overlay

Saved file format tiff

Image bit depth 8-bit integer

Output file location Default Output Folder

Overwrite existing files without warning? Yes No

When to save Every cycle

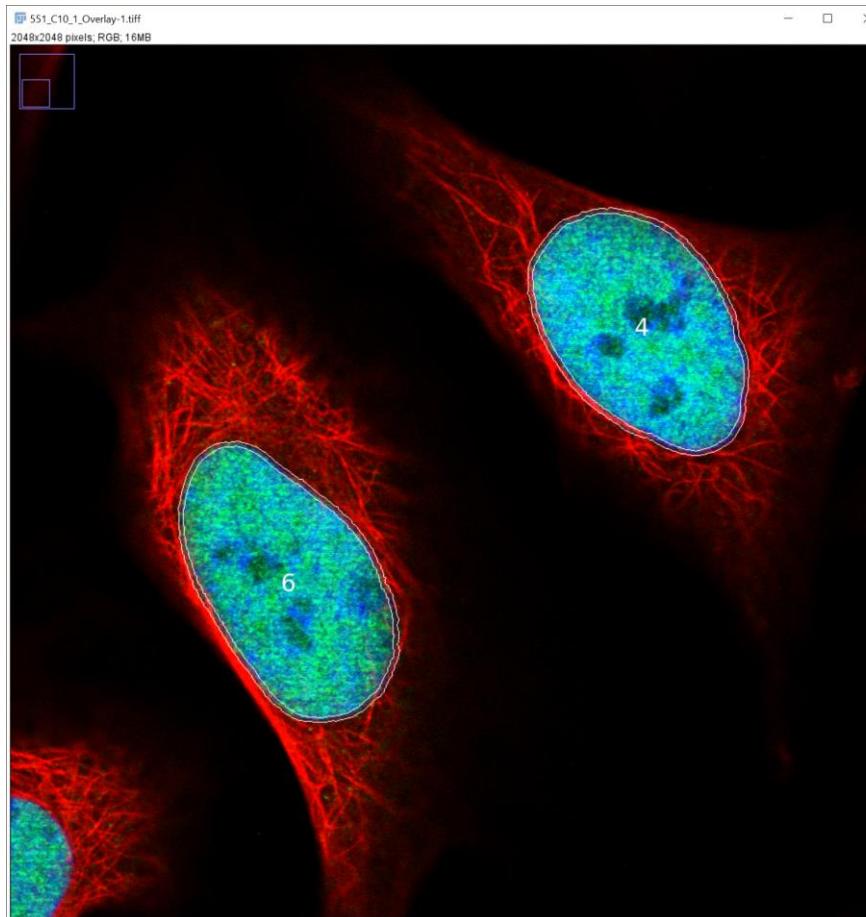
Record the file and path information to the saved image? Yes No

Create subfolders in the output folder? Yes No

- Images
- Metadata
- NamesAndTypes
- Groups
- ColorToGray
- IdentifyPrimaryObjects
- ExpandOrShrinkObjects
- IdentifyTertiaryObjects
- MeasureObjectIntensity
- CalculateMath
- ExportToSpreadsheet
- GrayToColor
- OverlayOutlines
- DisplayDataOnImage
- SaveImages

Control Image

www.proteinatlas.org/ENSG00000121892-PDS5A



- Catch-up: Create and save control images
- Run **Analyze Images**
- Inspect the output

Workflow

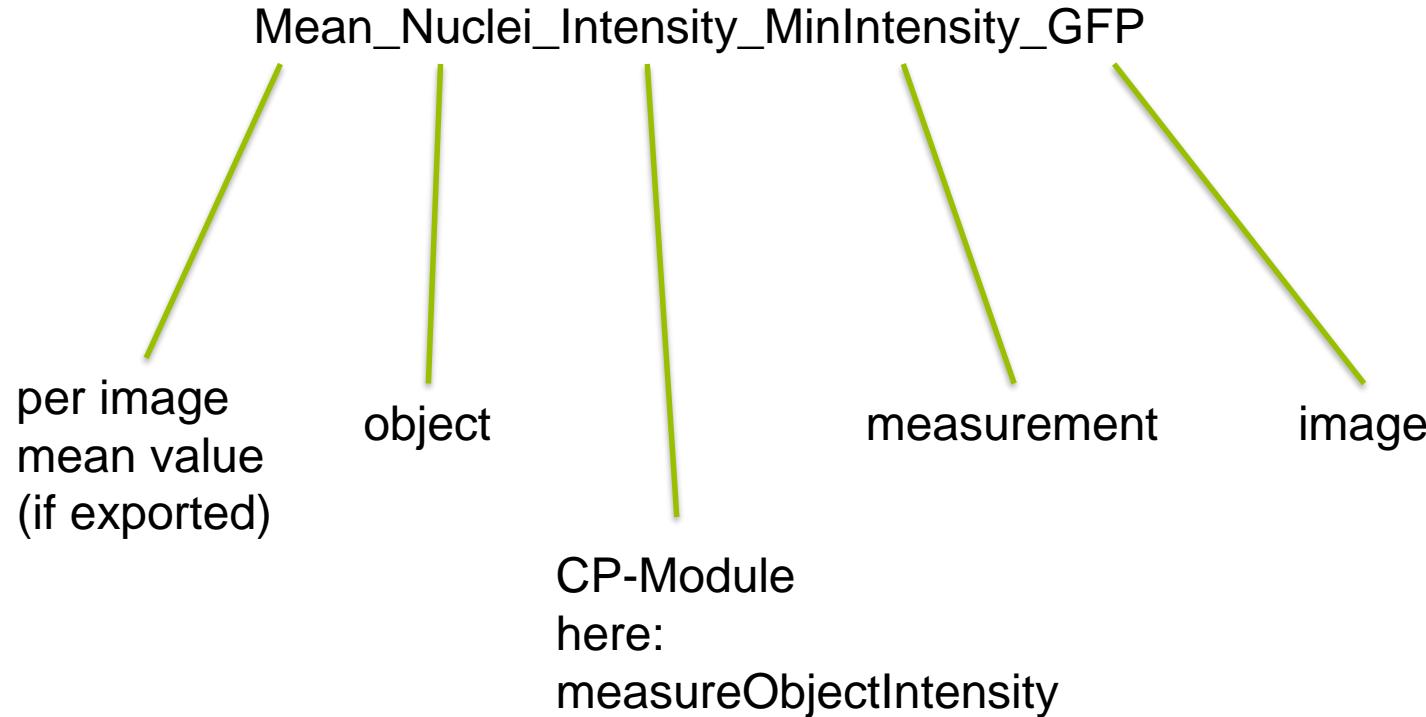
CP OUTPUT



@ahklemm

BioImage Informatics Facility

46



CP output - discussion

Results_Image.csv

Count_Nuclei	Count_ShrunkenNuclei	Count_nuclearMembrane	FileName_raw_image	ImageNumber	Mean_nuclearMembrane_Math_Ratio_Membrane_Nuclei
13	13	13	1772_B6_4.tif	1	1.204
7	7	7	1772_B6_7.tif	2	1.299
5	5	5	1773_H11_1.tif	3	1.169
10	10	10	1773_H11_2.tif	4	1.432
14	14	14	1894_G2_3.tif	5	0.339
13	13	13	1894_G2_4.tif	6	0.319
10	10	10	1894_G3_4.tif	7	1.296
10	10	10	1894_G3_5.tif	8	1.839
13	13	13	235_E7_3.tif	9	2.150
11	11	11	418_F7_2.tif	10	0.930
8	8	8	551_C10_1.tif	11	0.500
8	8	8	566_B12_2.tif	12	3.227
10	10	10	607_D7_2.tif	13	0.174
12	12	12	635_G10_1.tif	14	0.531
6	6	6	669_D7_2.tif	15	2.172
10	10	10	711_D6_1.tif	16	1.201
11	11	11	711_D6_2.tif	17	1.328
10	10	10	736_A11_1.tif	18	0.585
7	7	7	894_G10_3.tif	19	0.290
6	6	6	894_G10_4.tif	20	0.133
11	11	11	931_H7_1.tif	21	1.446
11	11	11	931_H7_2.tif	22	1.000

STEP II: MEASURE MEAN INTENSITY WITHIN CYTOPLASM

Aim 2: measure mean intensity within the cytosol

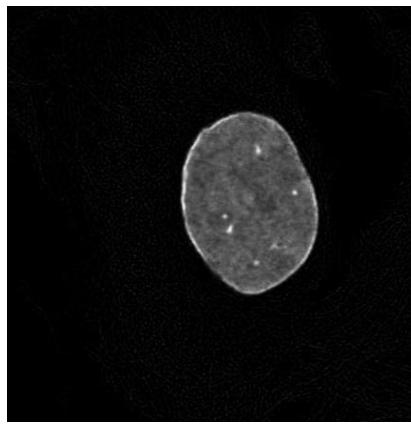
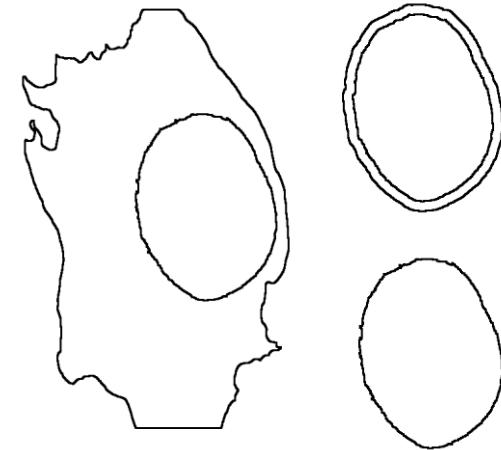
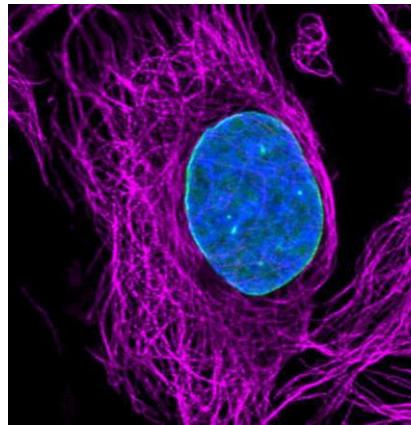
For now we have measured the mean antibody-signal intensity within the **nucleus** and the **nuclear membrane**. Let's also measure the signal intensity within the **cytoplasm**.

Questions:

- How can we segment the cytoplasm?
- Which channel should we use?

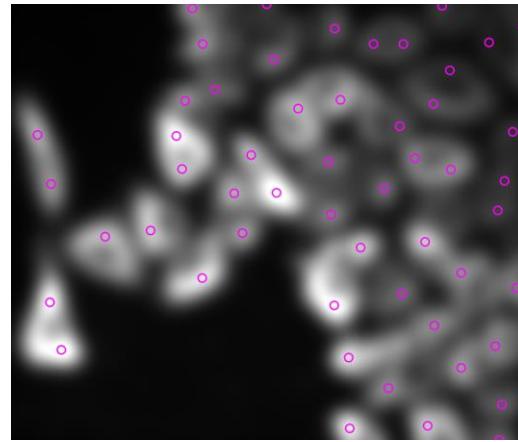
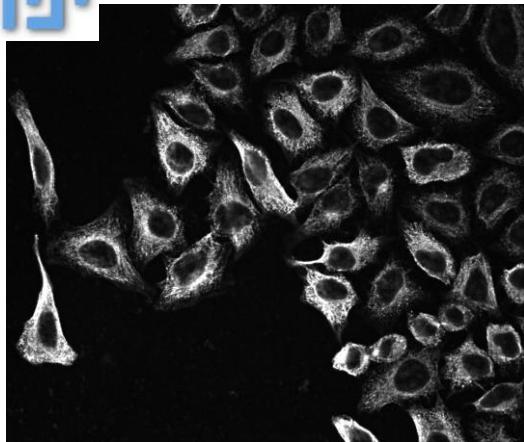
Basic Workflow

- Split Images into three separate channels.
↓
- Get outlines of nuclei from DAPI image
↓
- Get outlines of nuclear membrane
↓
- Get outlines of cytosol from microtubule image
↓
- Measure signal intensity of the antibody (green channel) in all outlines

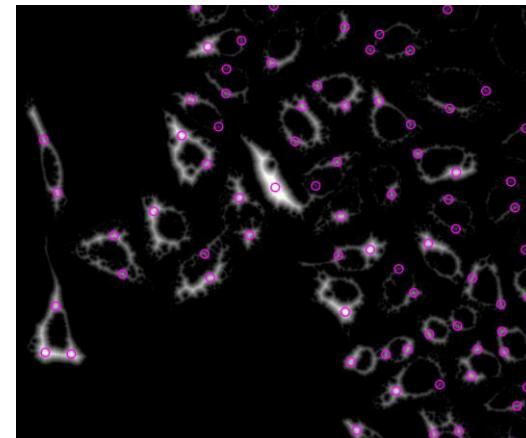
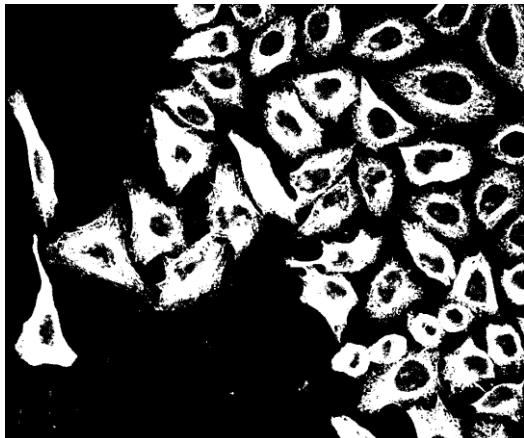


www.proteinatlas.org/ENSG00000113368-LMNB1

Identify Primary Objects on Cells

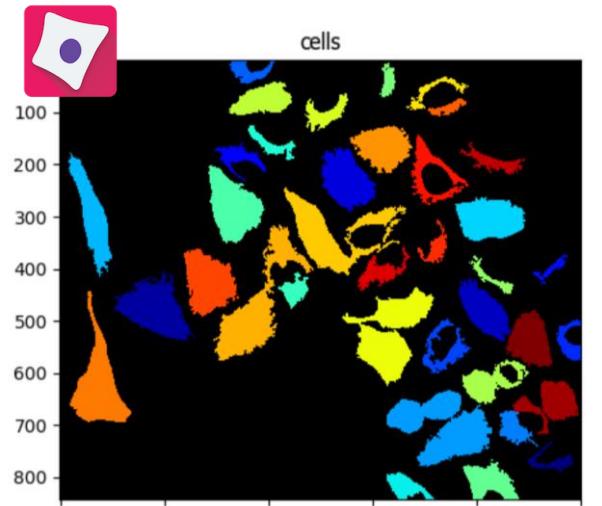


Find Maxima Tolerance 10
after Gaussian Blur r20

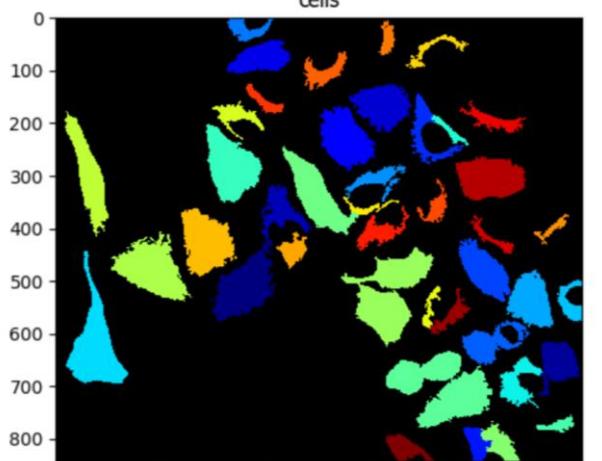


Huang Threshold

Image Source:
Klemm/Hasse/Sarov, MPI-CBG

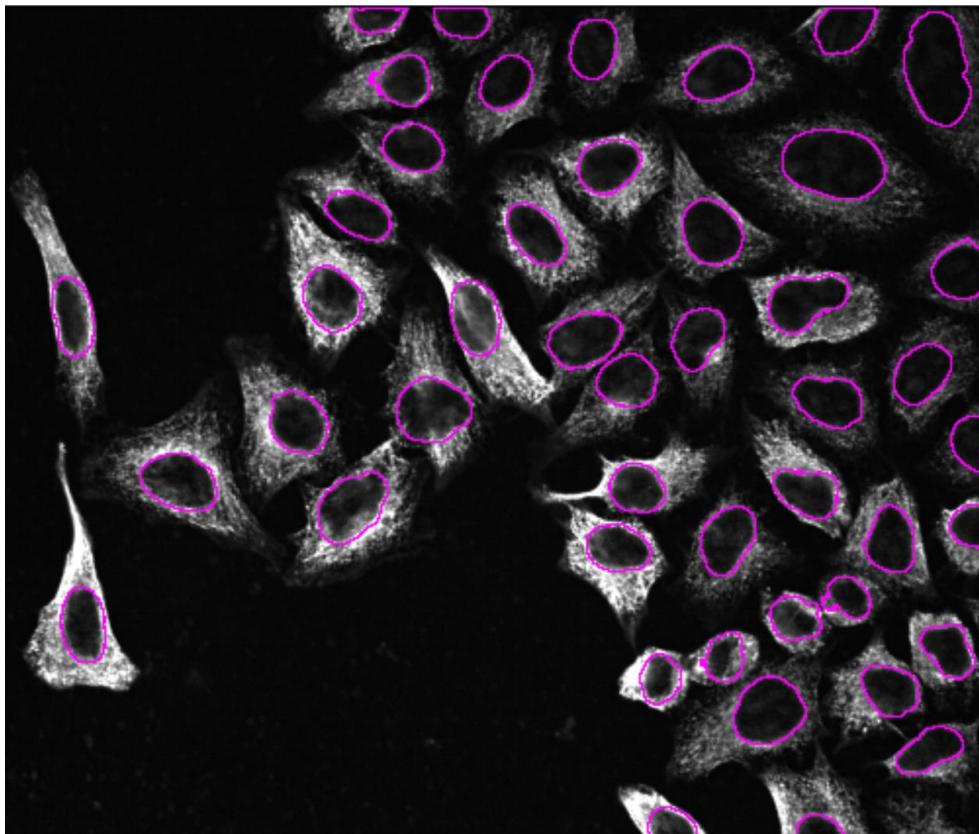


intensity-based declumping



shape-based declumping

Identify Secondary Objects



- **Goal:** Identify cell boundaries
- **Trick:** Nuclei typically more uniform in shape, more easily separated than cells
- **Approach:** Use nuclei as seeds for cell segmentation by using a cell stain channel
- Usually, a whole-cell stain works better for this than a membrane stain

Image Source:
Klemm/Hasse/Sarov, MPI-CBG

IdentifySecondaryObjects– Options I

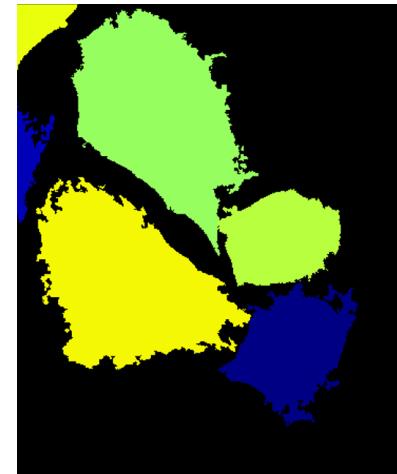
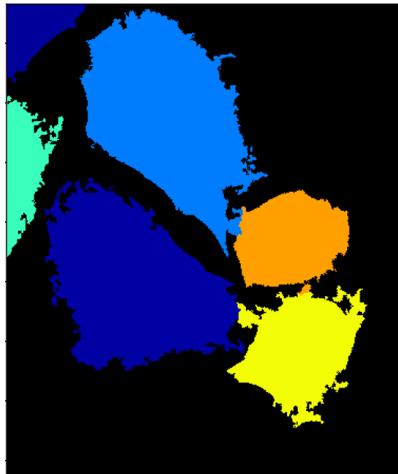
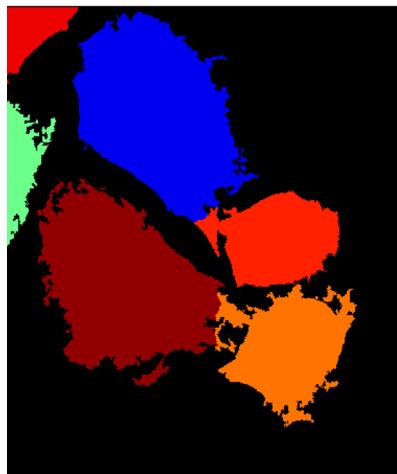
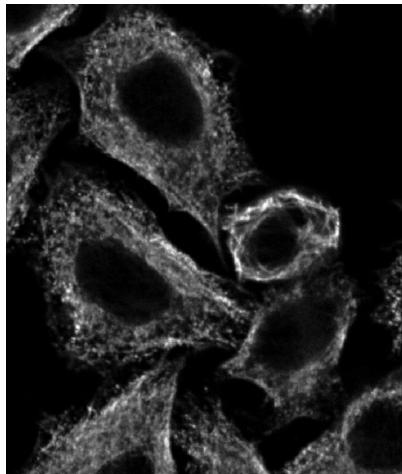


Image Source:
Klemm/Hasse/Sarov, MPI-CBG

Propagation
*“unlike the
Watershed
method, small
gaps are
tolerated”*

Jones et al
ICCV 2005.

**Watershed
Gradient**
*“This method
works best when
the image
intensity drops
off or increases
rapidly near the
boundary
between cells.”*

**Watershed
Image**
*“This method
works best
when there is a
saddle of
relatively low
intensity at the
cell-cell
boundary.”*

IdentifySecondaryObjects – Options - II

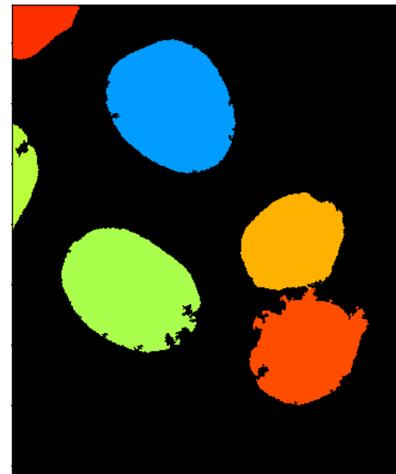
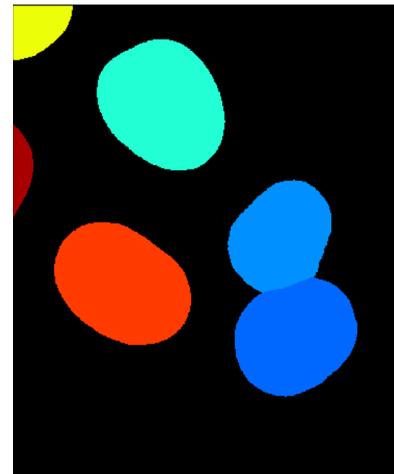
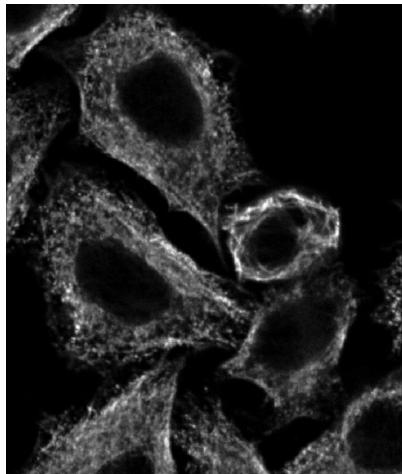


Image Source:
Klemm/Hasse/Sarov, MPI-CBG

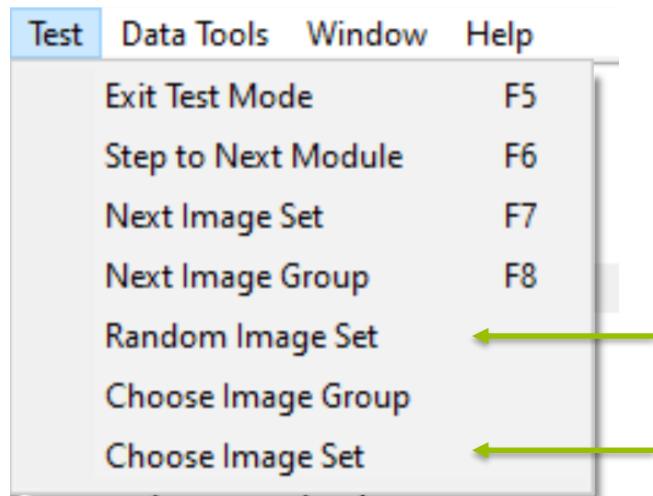
Distance-N

*“In this method,
the image of the
secondary
staining is not
used at all”*

Distance-G

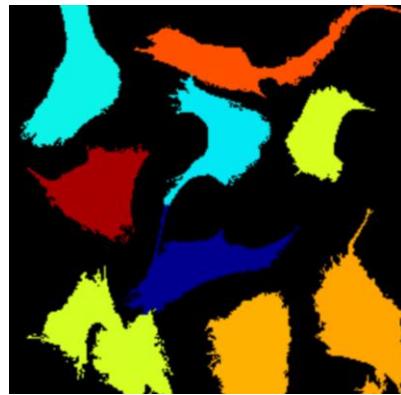
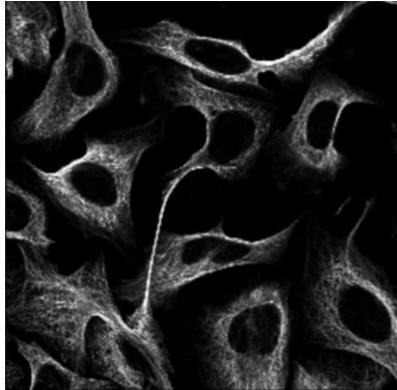
*“Thresholding of
the secondary
staining image is
used to
eliminate
background
regions from the
secondary
objects.”*

- Add the module IdentifySecondaryObjects
- In test mode: check which of the methods for identifying secondary objects works best.
- Try the chosen method on another Random Image Set and a chosen Image Set

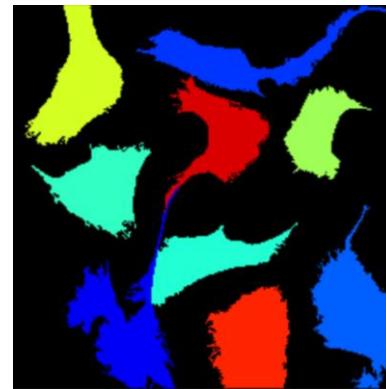


IdentifySecondaryObjects - Discussion

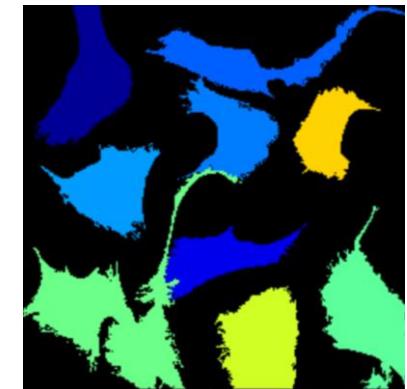
Propagation



Watershed
Gradient



Watershed
Image



www.proteinatlas.org/ENSG00000270647-TAF15

Identify Secondary Object - Solution

Module: IdentifySecondaryObjects

Select the input image MT (from ColorToGray #05)

Select the input objects Nuclei (from IdentifyPrimaryObjects #06)

Name the objects to be identified Cells

Select the method to identify the secondary objects Propagation

Threshold strategy Global

Thresholding method Minimum cross entropy

Threshold smoothing scale 0.0

Threshold correction factor 1.0

Lower and upper bounds on threshold 0.0 1.0

Regularization factor 0.05

Fill holes in identified objects? Yes No

Discard secondary objects touching the border of the image? Yes No

Available modules:

- Images
- Metadata
- NamesAndTypes
- Groups
- ColorToGray
- IdentifyPrimaryObjects
- ExpandOrShrinkObjects
- IdentifyTertiaryObjects
- IdentifySecondaryObjects
- IdentifyTertiaryObjects
- MeasureObjectIntensity
- CalculateMath
- CalculateMath
- ExportToSpreadsheet
- GrayToColor
- OverlayOutlines
- DisplayDataOnImage
- SaveImages

Exercise – Last Steps

- Create a selection of only the cytoplasm (not the entire cell)
- Measure the intensity within the cytoplasm selection in the signal image
- Calculate the ratio $I_{mean\ cytoplasm} / I_{mean\ nucleus}$
- Export to spreadsheet: if you had selected the measurements to export – include the new ratio measurement
- Leave the test mode and analyze all images



Solution - Creating a Selection for the cytoplasm: Tertiary Object

Module: IdentifyTertiaryObjects

Select the larger identified objects (from IdentifySecondaryObjects #07)

Select the smaller identified objects (from IdentifyPrimaryObjects #06)

Name the tertiary objects to be identified

Shrink smaller object prior to subtraction? Yes No

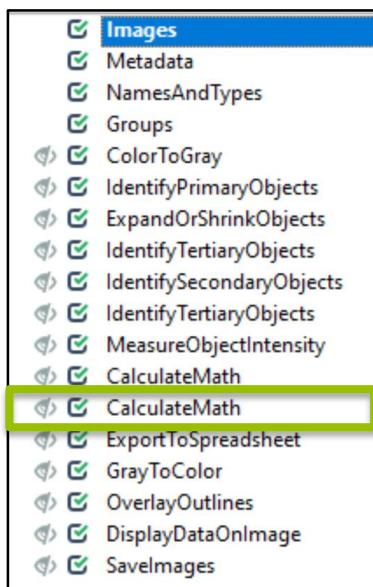
- Images
- Metadata
- NamesAndTypes
- Groups
- ColorToGray
- IdentifyPrimaryObjects
- ExpandOrShrinkObjects
- IdentifyTertiaryObjects
- IdentifySecondaryObjects
- IdentifyTertiaryObjects
- MeasureObjectIntensity
- CalculateMath
- CalculateMath
- ExportToSpreadsheet
- GrayToColor
- OverlayOutlines
- DisplayDataOnImage
- SaveImages

Solution - Modify MeasureObjectIntensity

The screenshot shows the CellProfiler software interface. On the left, a sidebar lists various processing steps: Images, Metadata, NamesAndTypes, Groups, ColorToGray, IdentifyPrimaryObjects, ExpandOrShrinkObjects, IdentifyTertiaryObjects, IdentifySecondaryObjects, IdentifyTertiaryObjects, MeasureObjectIntensity, CalculateMath, CalculateMath, ExportToSpreadsheet, GrayToColor, OverlayOutlines, DisplayDataOnImage, and SaveImages. The 'MeasureObjectIntensity' step is highlighted with a green border and a green arrow points from it to the right panel. The right panel contains three dropdown menus for selecting objects to measure: 'Select an image to measure' set to 'Signal' (from 0), 'Select objects to measure' set to 'Nuclei', and 'Select objects to measure' set to 'nuclearMembrane'. There is also a 'Remove this object' button. Another 'Select objects to measure' dropdown is shown at the bottom set to 'Cytoplasm'.

MeasureObjectIntensity

Solution - Calculate Ratio Imean Cytoplasm / Imean Nuclei



MeasureObjectIntensity

Select the first operand measurement type Object

Select the first operand objects Cytoplasm (from IdentifyTertiaryObjects #11)

Category: Intensity

Select the first operand measurement Measurement: MeanIntensity

Image: Signal

Multiply the above operand by 1.0

Raise the power of above operand by 1.0

Select the second operand measurement type Object

Select the second operand objects Nuclei (from IdentifyPrimaryObjects #06)

Category: Intensity

Select the second operand measurement Measurement: MeanIntensity

Image: Signal

Multiply the above operand by 1.0

Raise the power of above operand by 1.0

Take log10 of result? Yes No

Multiply the result by 1.0

Raise the power of result by 1.0

Add to the result 0.0

Final Workflow

- Split Images from RGB into three separate channels.
 - Get outlines of nuclei from DAPI image
 - Get outlines of nuclear membrane
 - Get outlines of cytoplasm from microtubule image
 - Measure signal intensity of the signal in all outlines
 - Calculate Ratios
 - Export to Spreadsheet
 - Generate and Save control images
-
- 1) Shrink Nucleus Outline
2) Calculate outline of nuclear membrane
= Nuclei– Shrunken Nuclei
- 1) Get outlines of cells
2) Calculate outline of cytoplasm:
= Region cell – Region nucleus

What to do with the output?

- Some calculations can be done in CellProfiler directly, e.g. the calculating the intensity ratio.
- For more complex downstream calculations I prefer to have the CellProfiler output as "raw" as possible. You can then use e.g. Python, R, Knime etc. for more advanced filtering, math, clustering.

When do I prefer using CellProfiler vs other software?

- Batch-processing: large amount of images of same type
- Many cells per image
- “Typical Problem”: CP is extremely good for measuring cells, textures, output, control-images...
- For specialized applications I prefer Fiji
- Existing CP-modules:
 - Illumination Correction
 - Detecting and Counting dots/cell
 - Neurite Detection (Tubeness filter)
 - C.elegans modules
 - ...
 - → Check out Example Pipelines on the CellProfiler website!

Benefits of free and open-source software

- ‘**Reproducible research**’; with an open source solution, you can provide your analysis pipeline as part of the supplementary material of your published paper (along with sample data).
- **Share pipelines** and analysis approaches with collaborators.
- You can bring your methods with you **when moving to a different lab**, or working remotely.
- **Educational value**: anyone can go in and look at the source and learn.
- Commercial companies have trouble keeping up with the most recent developments: the latest things are more likely to be found in the open-source community.
- Very active **user communities**

Please remember to properly cite the software in your publications to facilitate future funding for the developers!!!

<http://forum.image.sc/>

<http://cellprofiler.org/examples/>

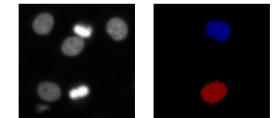
- Many example pipelines!

Specialized Pipelines

In addition to cellular object and feature identification, these pipelines include some of the more specialized modules in CellProfiler for image pre-processing or measurement.

Cell/particle counting and scoring the percentage of stained objects

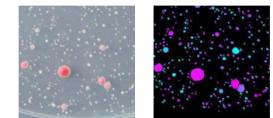
CellProfiler is commonly used to count cells or other objects as well as percent-positives, by measuring the per-cell staining intensity. This pipeline shows how to do both of these tasks, and demonstrates how various modules may be used to accomplish the same result.



[Download](#) (0.2 MB)

Yeast colony classification

This pipeline demonstrates how to classify and count objects on the basis of their measured features. The example identifies uniformly round objects, in this case, yeast colonies growing on a dish. The pipeline also shows how to load a template and align it to a cropped image, as well as how to use illumination correction to subtract for background illumination.



[Download](#) (0.6 MB)

[Tutorial](#)

Yeast patch identification



- Raw images from The Human Protein Atlas / The Cell Atlas,
<https://www.proteinatlas.org/humanproteome/cell>

Thul, P.J., Åkesson, L., Wiking, M., Mahdessian, D., Geladaki, A., Ait Blal, H., Alm, T., Asplund, A., Björk, L., Breckels, L.M., et al. (2017).
A subcellular map of the human proteome. *Science* 356.

THANKS!

Contributors to these slides:

Anna Klemm, Carolina Wählby



@ahklemm

BioImage Informatics Facility