(Quantitative Imaging for) Colocalisation Analysis

… or

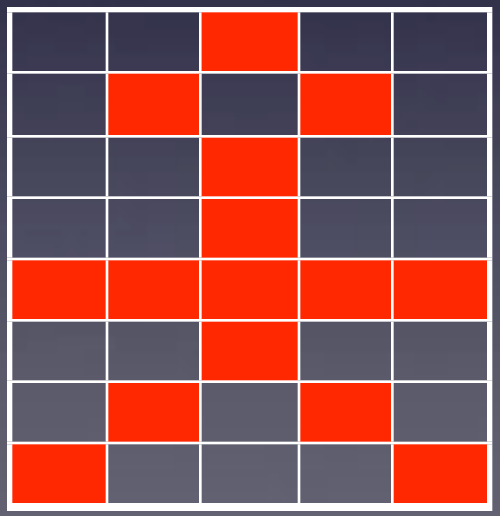
Why Colour Merge / Overlay Images are EVIL!

Special course for
DIGS-BB PhD program

CBG
Max Planck Institute
of Molecular Cell Biology
and Genetics

What is an Image anyway..?

- An image is a representation of reality (not real)
  - Image of a point is not a point (Point Spread Function)
  - Pixelated by detector (CCD or point scanner)



A digital image of ???

Image Analysis
(Brain or Computer)

A stick man?
How do I know?
How can computer know - algorithm?

What is an Image anyway..?

- Images contain information (not just pretty pictures)
  - Manipulate Image = Changed Info
    (Brightness / Contrast - Extreme Caution!!!)
  - Image data can be quantified / measured / analysed
  - You cant add lost info back.
  - Meta data (What, Where, When, How)

A digital image
How many objects?
How “bright” is it?
How big is it?
What is it?
etc.

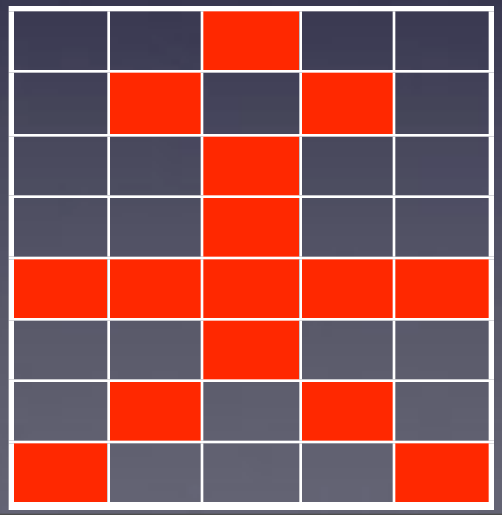
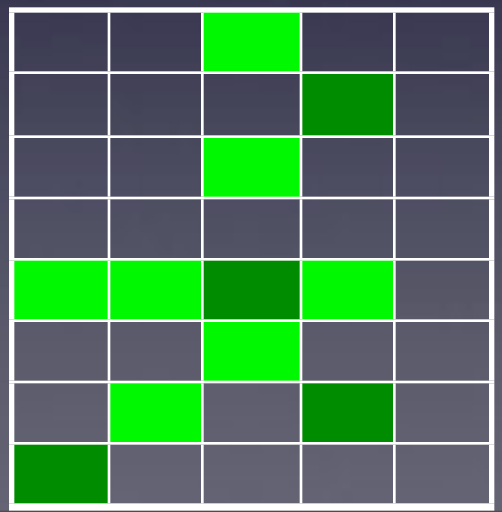


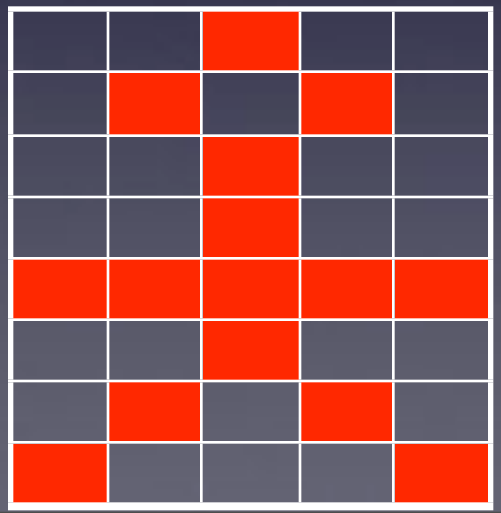
Image Data? What is it?

- Intensity is related to what? Something physical?
  - Dye concentration Or is it? Why not?
  - Noisy Images? Averaging? Pixel Time?
- Comparison of 2 colours/dyes -
  Biology / BioChemistry / Interaction ?
- Shapes, Movement, Structure?

A digital image
With 2 channels / colours

What can you say here?

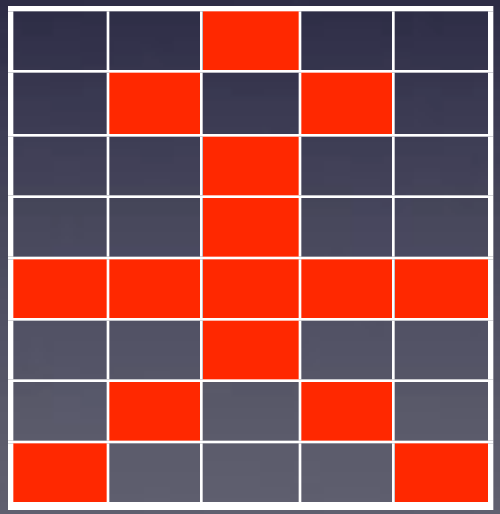




Photographer or Spectroscopist?

- We can show you how to take pretty pictures (Art)
- We can teach you to get useful information (Science)

- You have to choose which you want to be!



This

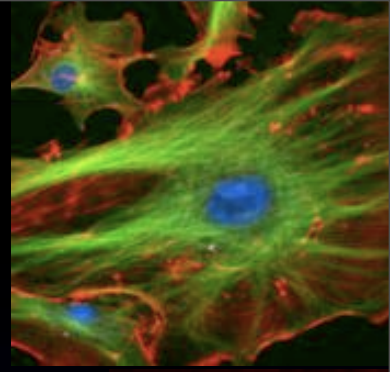
Is simply a way to "Visualise"

This

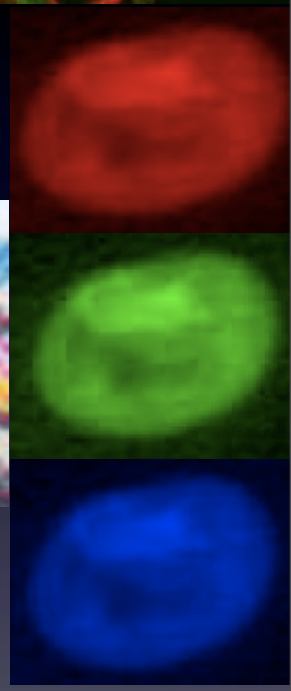
0	0	1	0	0
0	1	0	1	0
0	0	1	0	0
0	0	1	0	0
1	1	1	1	1
0	0	1	0	0
0	1	0	1	0
1	0	0	0	1

Publishing Images

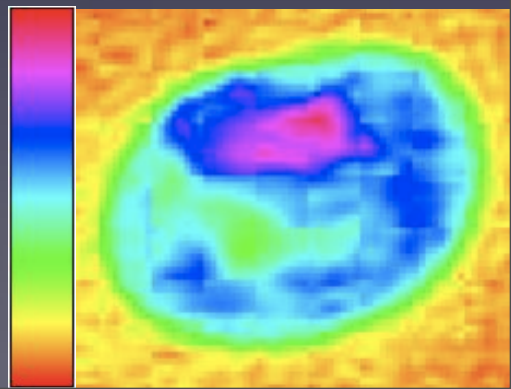
or “how Photoshop can ruin your career”



- Which image? Prettiest? Representative?
- CCD/PMT sees intensities differently than your eye/brain
  - LUT? Gamma correction? Calibrate Monitor - we have the tools!
- RBG colour space is not what we print!
  - RGB - Visualise (LCD, CRT)
  - CYMK - Print
  - Journal Image ≠ Screen Image

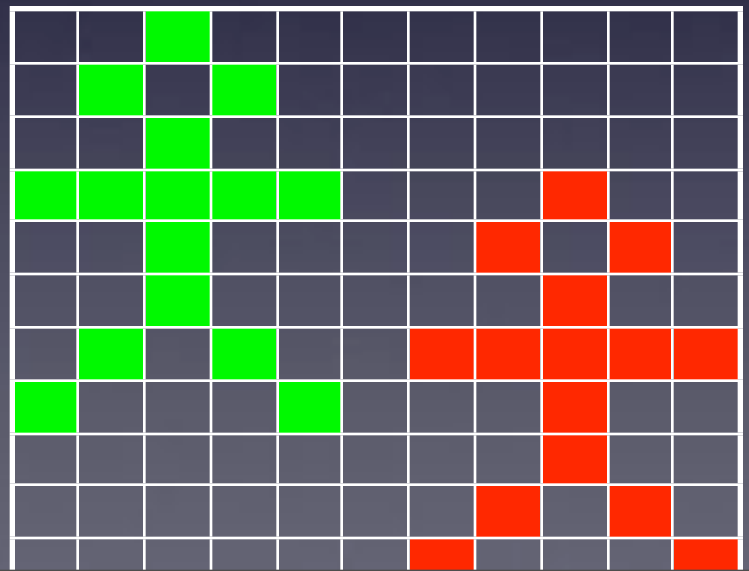


- Author instructions - image format? TIFF CYMK
- Materials and Methods - exact image processing done
- Image = data Don’t corrupt information!
  - PDF - reviewer can check image processing results!
  - Compression - Lossless ok - Lossy (JPEG) very bad
  - You wouldn’t do it to any other kind of data



Quantitative Image Analysis?
…what does that mean?

- Pretty pictures are great for journal covers...
- Movies are great for visual presentation of images...
- Interactive 3D visualisation, data exploration...
- But for meaningful biological conclusions...
  - Scientists need numerical results from image data
  - Need to measure many objects
  - Need statistics from many images
  - Computers become useful!

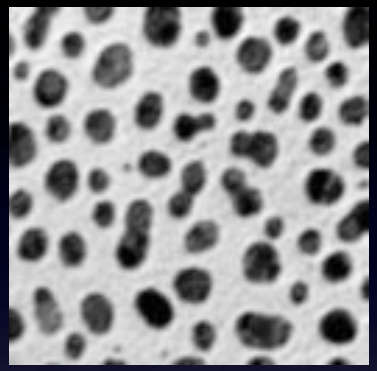


Quantitative Microscopy - First Think...

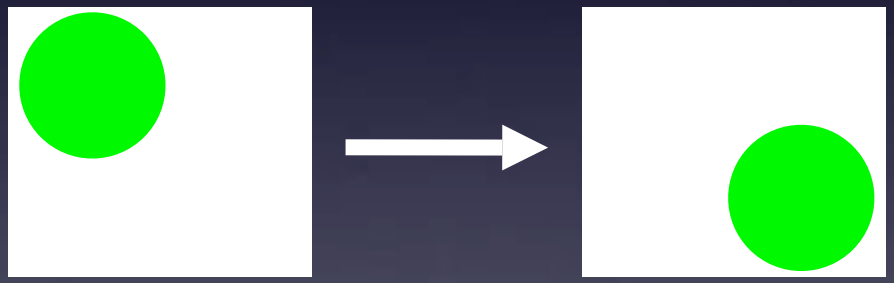
- Choosing experimental and image processing methods:
  - What BIOLOGY am I trying to see or measure?
  - Do I need 3D information? Resolution? Object size?
  - Choose / Optimise microscope system to use!
  - Statistics? How many images / data points / experiments?
  - **Controls!!!**

Experimental Design - First Think...

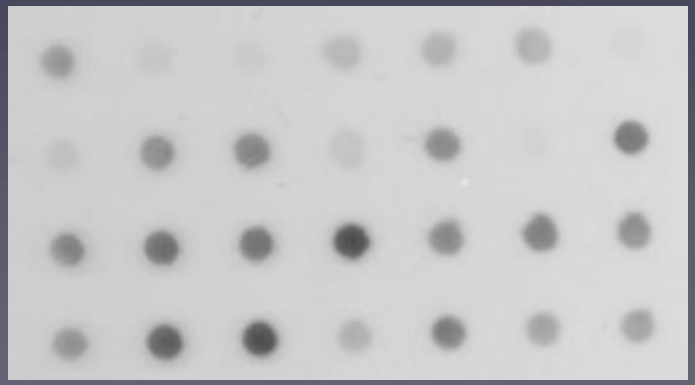
- Quantitative Experiments?
  - Am I trying to measure the size/shape of some type of object(s)



- Am I trying to see movement over time?

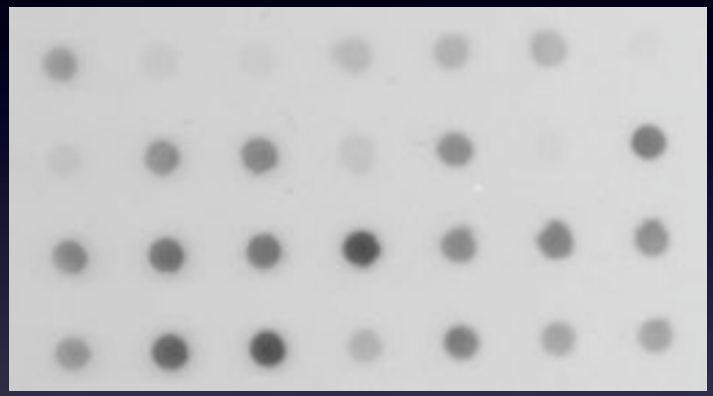


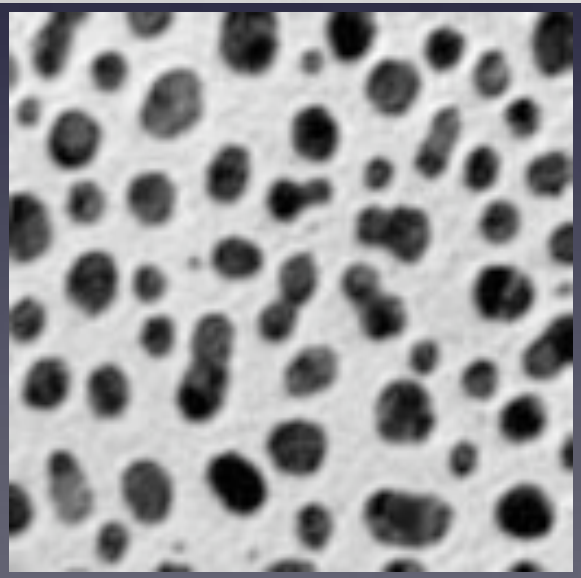
- Am I trying to measure a number, amount or concentration?



Am I trying to measure the number of some type of object?

- Can I define how my objects appear in images?
- Segmentation

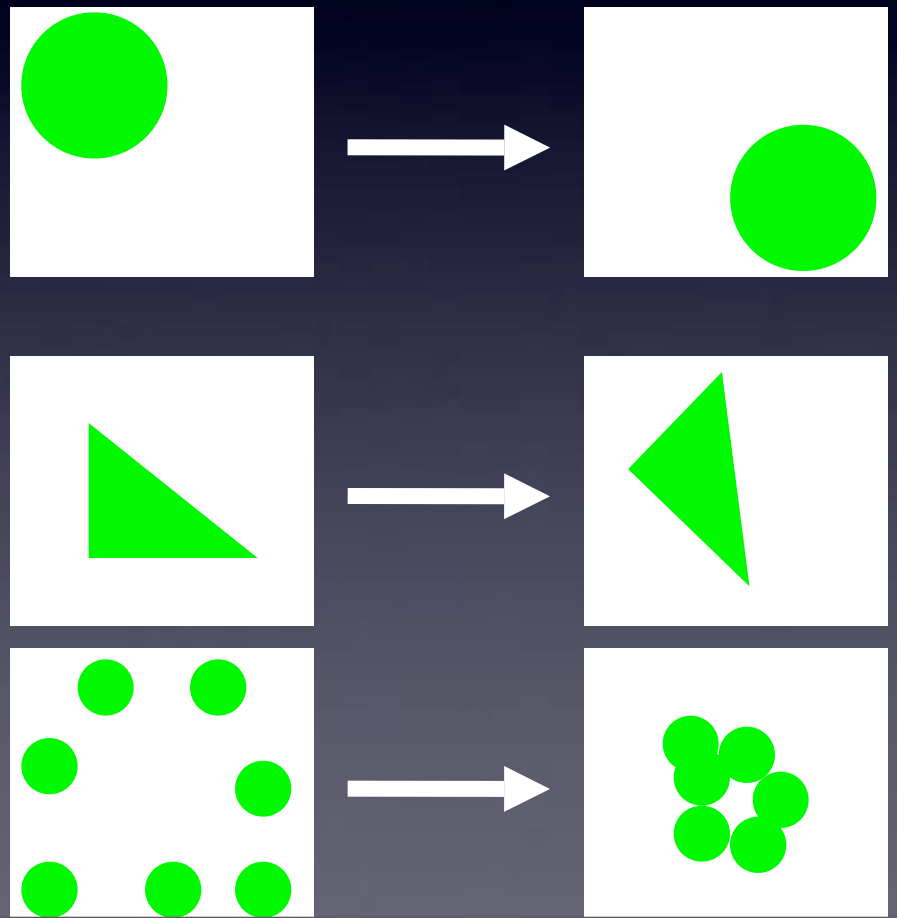




- Image intensity - threshold
- Size - threshold
- Shape - circularity etc.

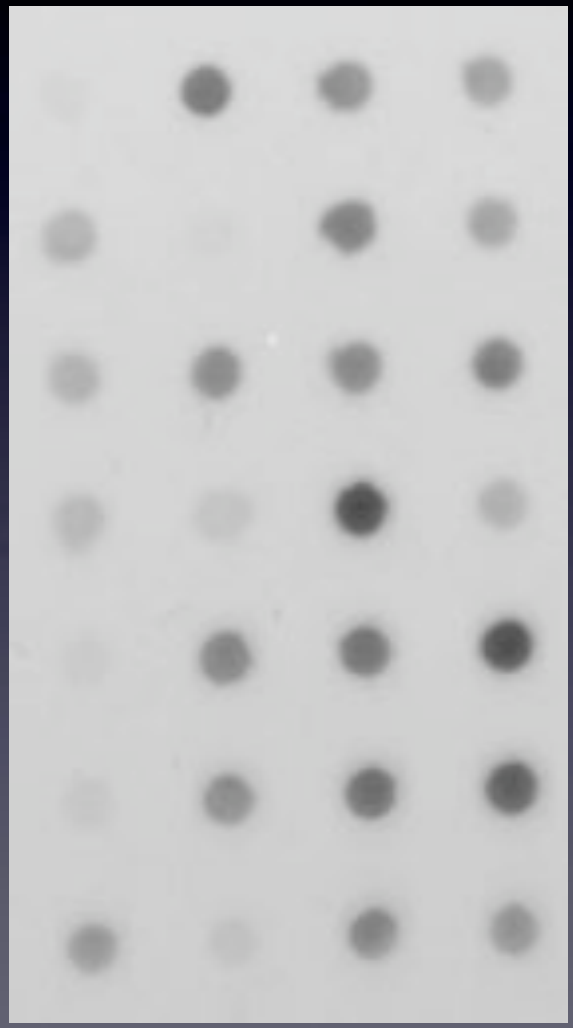
Am I trying to see something move over time?

- Can I define what movement is?
  - Linear - A to B?
    - Direction
    - Speed
    - Velocity
  - Rotation
  - Clustering



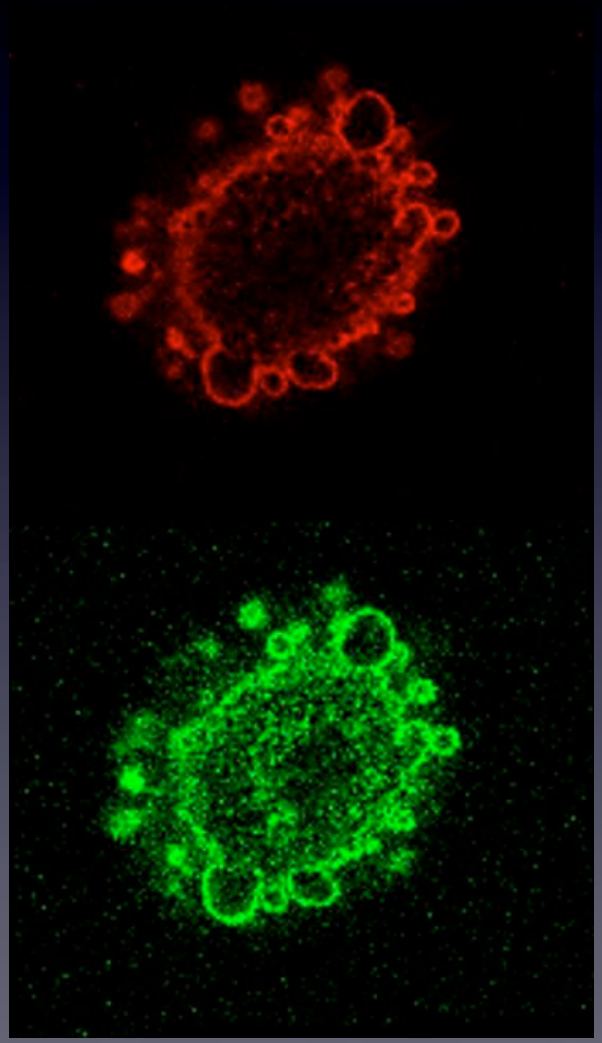
Am I trying to measure an amount or concentration?

- Does that have a Biological meaning?
- Absolute or Relative?
- Can I calibrate my image intensity vs. something else / itself?
  - eg. Fluorescence signal vs. Quantitative Assay or Baseline / Control
  - Fluorescence response might not be linear!

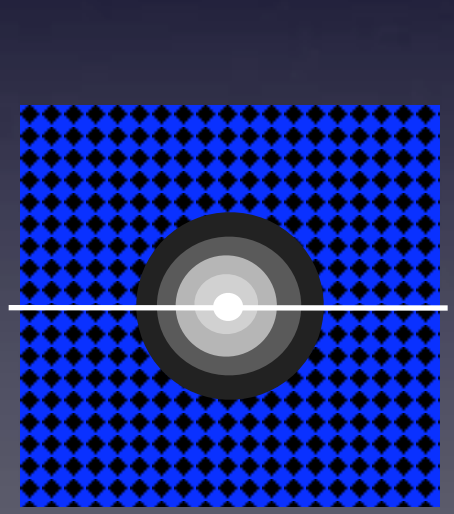


Am I trying to measure an "image parameter"?

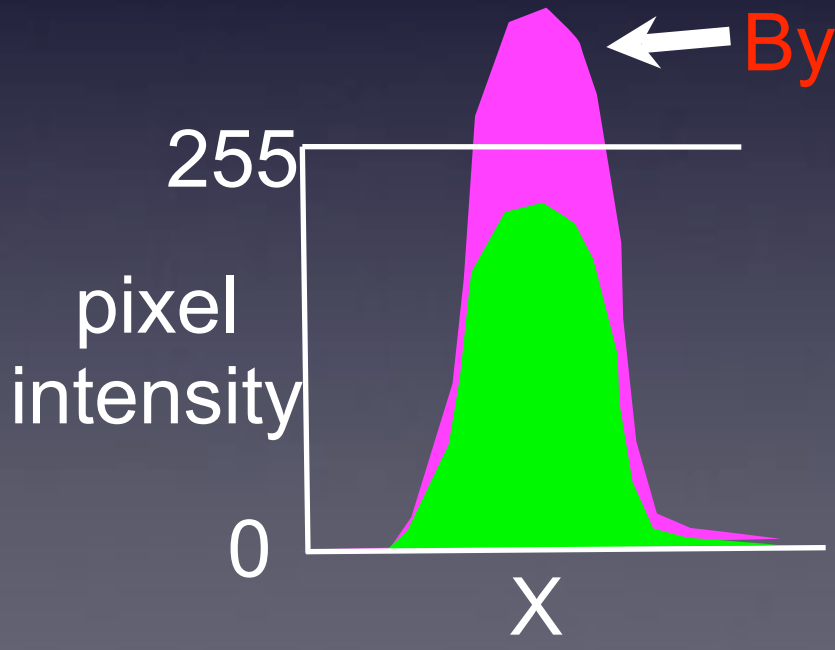
- Does that have a Biological meaning?
- Absolute or Relative?
  - Total / Mean / SD of signal
  - Background
  - Signal : Noise
  - Texture (smooth/spotty)
  - Colocalisation between "colours" / channels

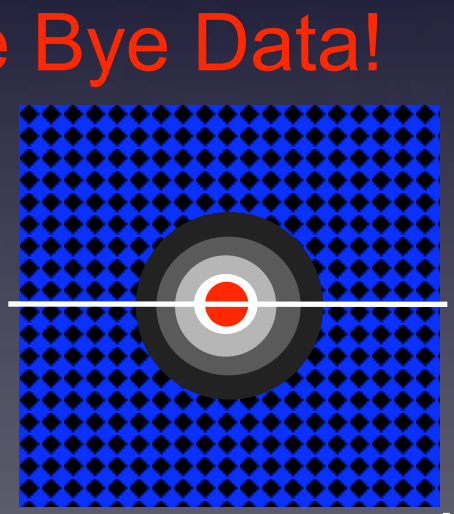


- Signals within the range of the detector?
  - Your eyes lie! You can't see low intensities close to black! Use Range Indicator / HiLo / OU and spectrum CLUTs
  - Adjust so brightest part is within detector range.
  - Remember to check z dir. also.
  - Don't over expose the image! Why not? Lost Info!



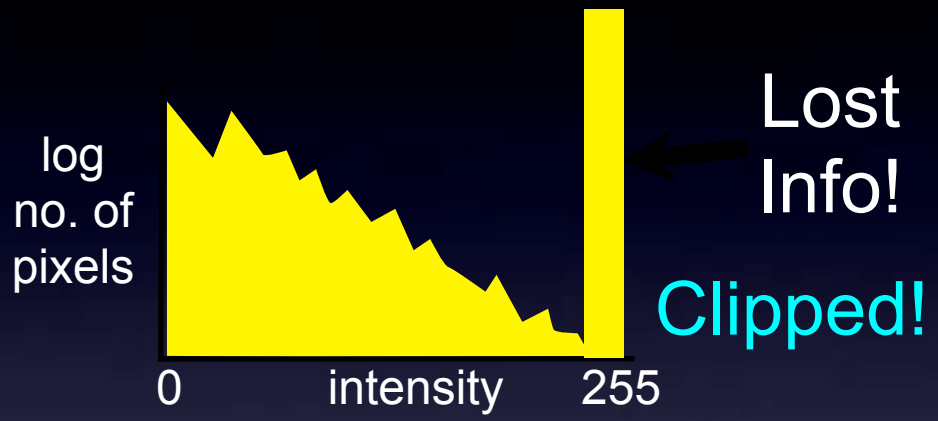
in range

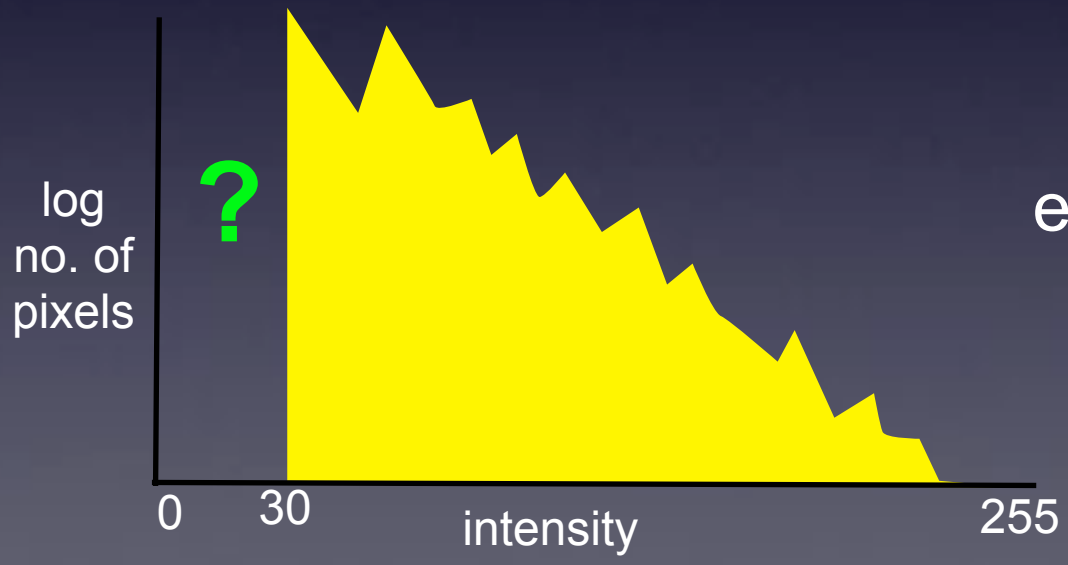




overexposed saturated

Image Intensity Histograms - Use them!

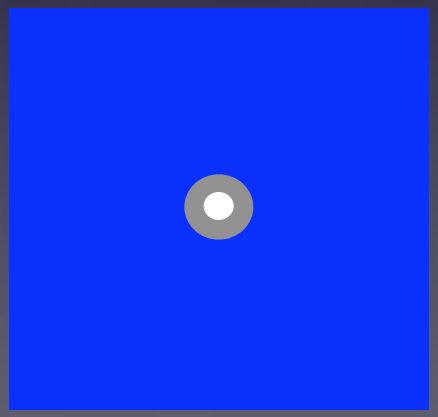




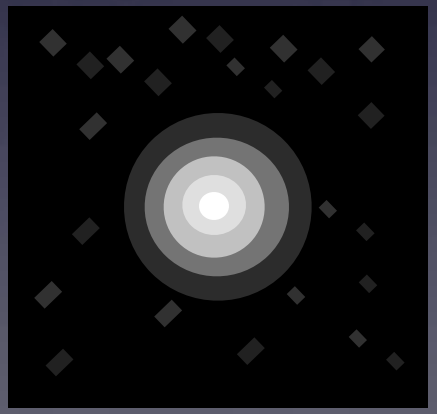
bad detector offset

In Histograms:
easily see problems
for image
quantification!

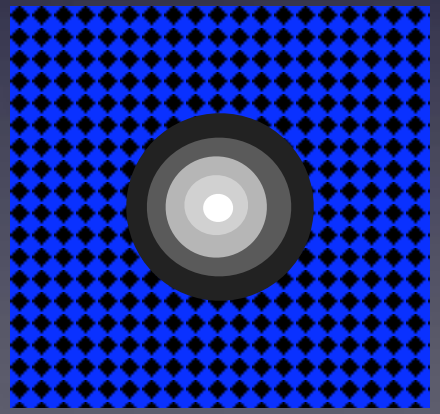
- Signal within the range of detector?
  - Offset / Zero Background - Set properly.
  - Why? “background” ≈ zero, but keep low intensity info
    - What is “Background”? You decide!
  - Range indicator / HiLo CLUT - background black and blue ~50:50
  - (0 = Blue, 1 = Black, 254 = White, 255 = Red)



too high



too low

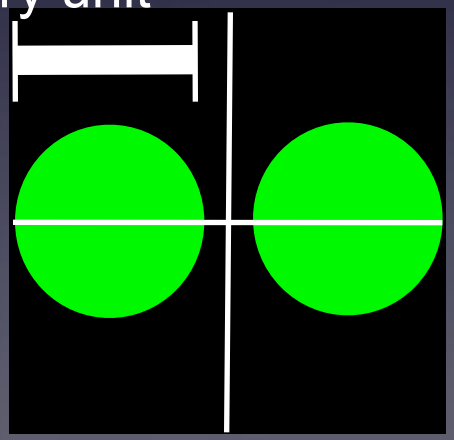


correct

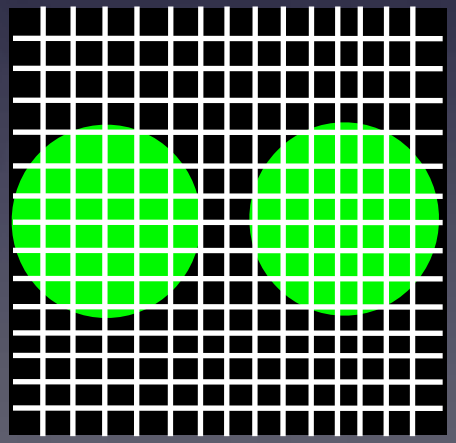
Pixel Size / Resolution

- "Correct" image size (64x64, 512x512, 2048x2048)?
  - Get all information microscope can resolve, but files not too big
  - Proper spatial sampling (Nyquist sampling theory)
  - 2.3-3 pixels over optical resolution distance. (x, y and z)
  - Adjust zoom and image size.
  - Auto Pinhole or 1 Airy unit

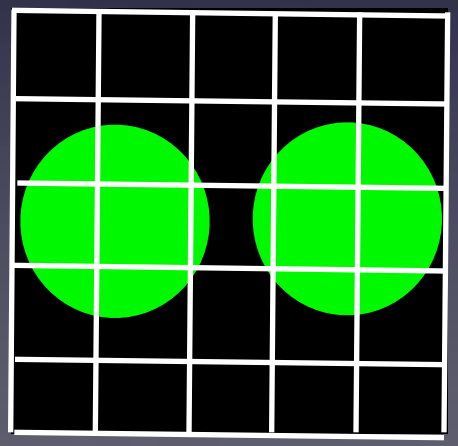
1 Airy unit



under sampled



over sampled

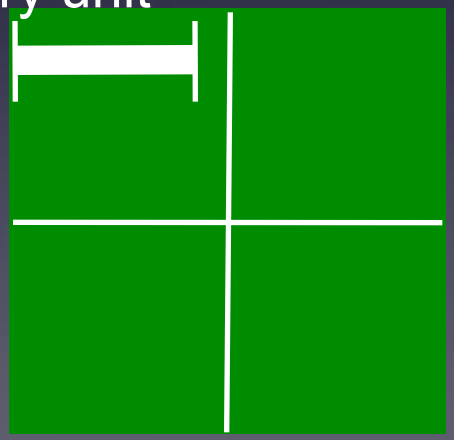


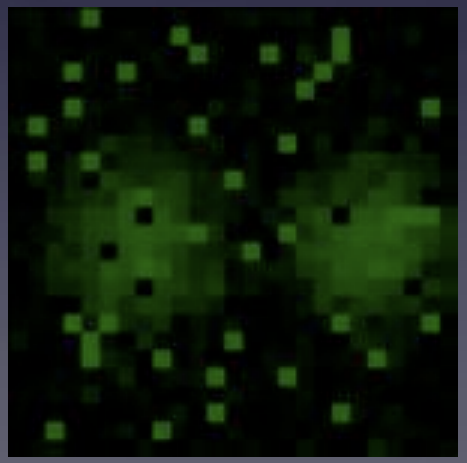
correct sampling

Pixel Size / Resolution

- “Correct” image size (64x64 or 2048x2048 - or something else)?
  - Get all information microscope can resolve, but files not too big
  - Proper spatial sampling (Nyquist sampling theory)
  - 2.3-3 pixels over optical resolution distance. (x, y and z)
  - Adjust zoom and image size.
  - Auto Pinhole or 1 Airy unit

1 Airy unit

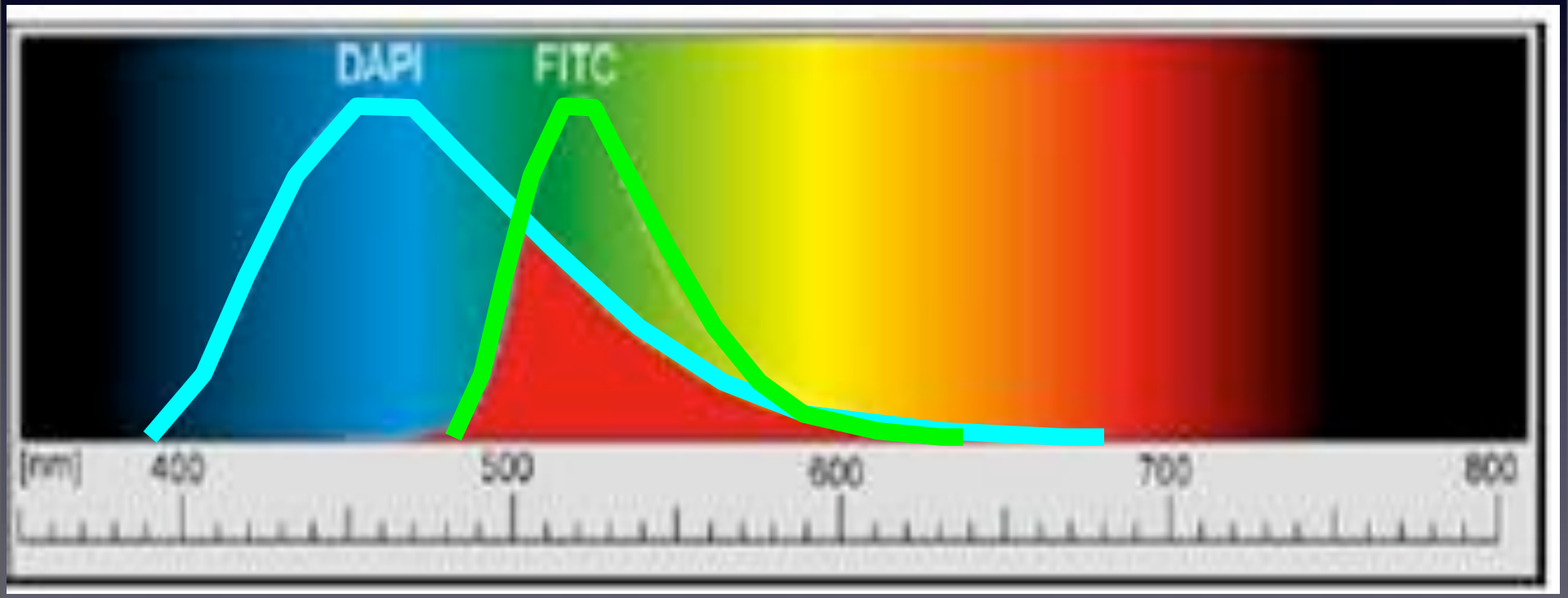




under sampled over sampled correct sampling

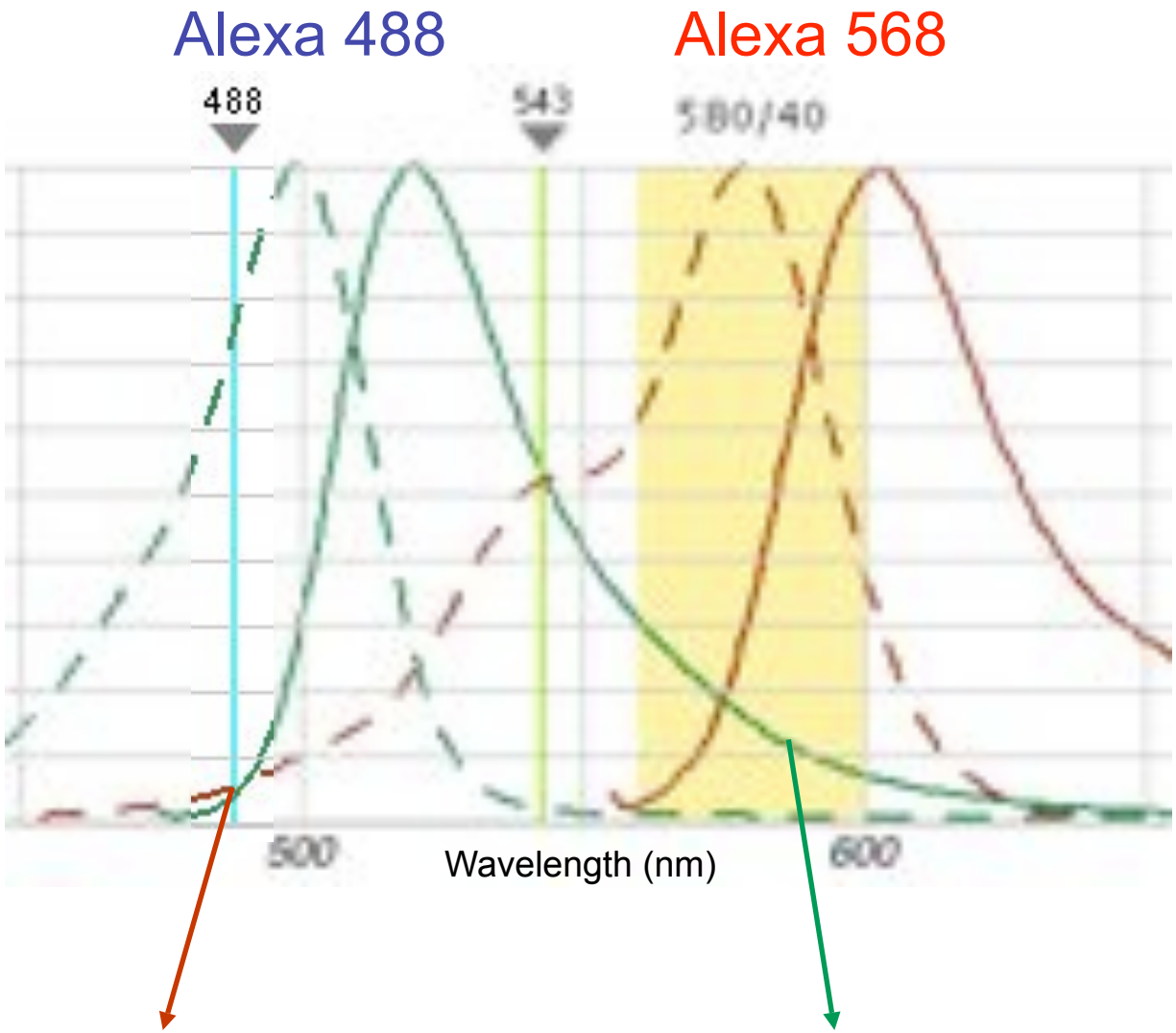
Avoid Emission Bleed Through and

- Dye selection / Filter selection
  - Emission bleed through and/or excitation crosstalk...
  - Means you get: Overlapping emission - Quantitative? No!



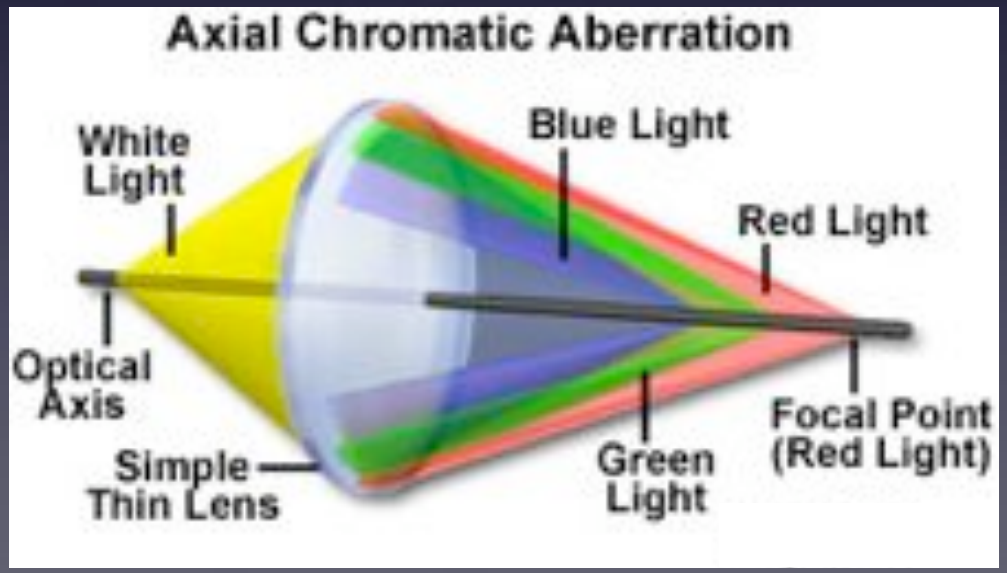
- Use multi tracking (Zeiss) / sequential (Olympus)

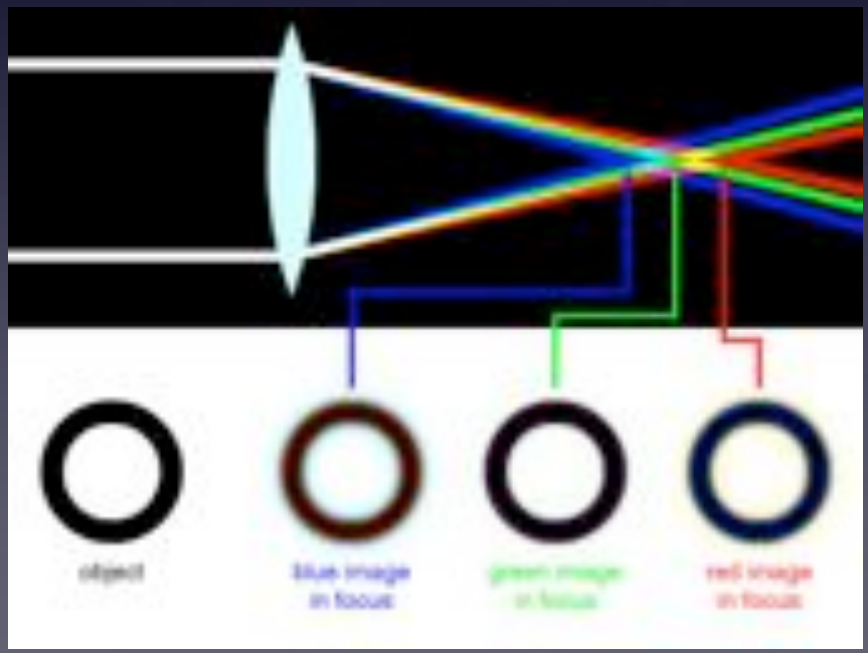
Beware ! Crosstalk and Bleed Through



Watch Out - More Holes To Fall Into:

- Correct objective lens / microscope setup for task
  - N.A / Resolution.
  - Apochromat for different colours (UV)
  - Calibrate Scanner / Check with multi-colour beads



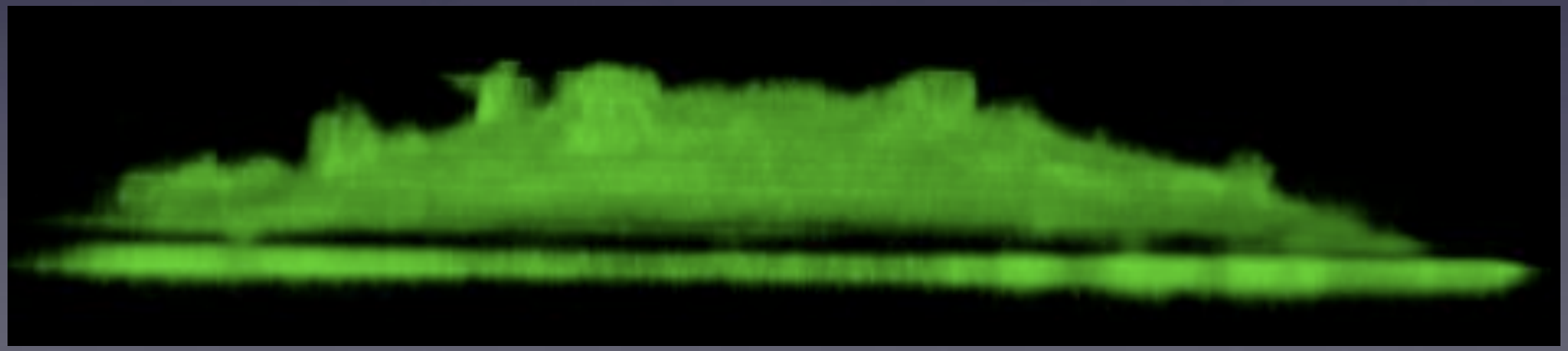


Watch Out - More Holes To Fall Into:

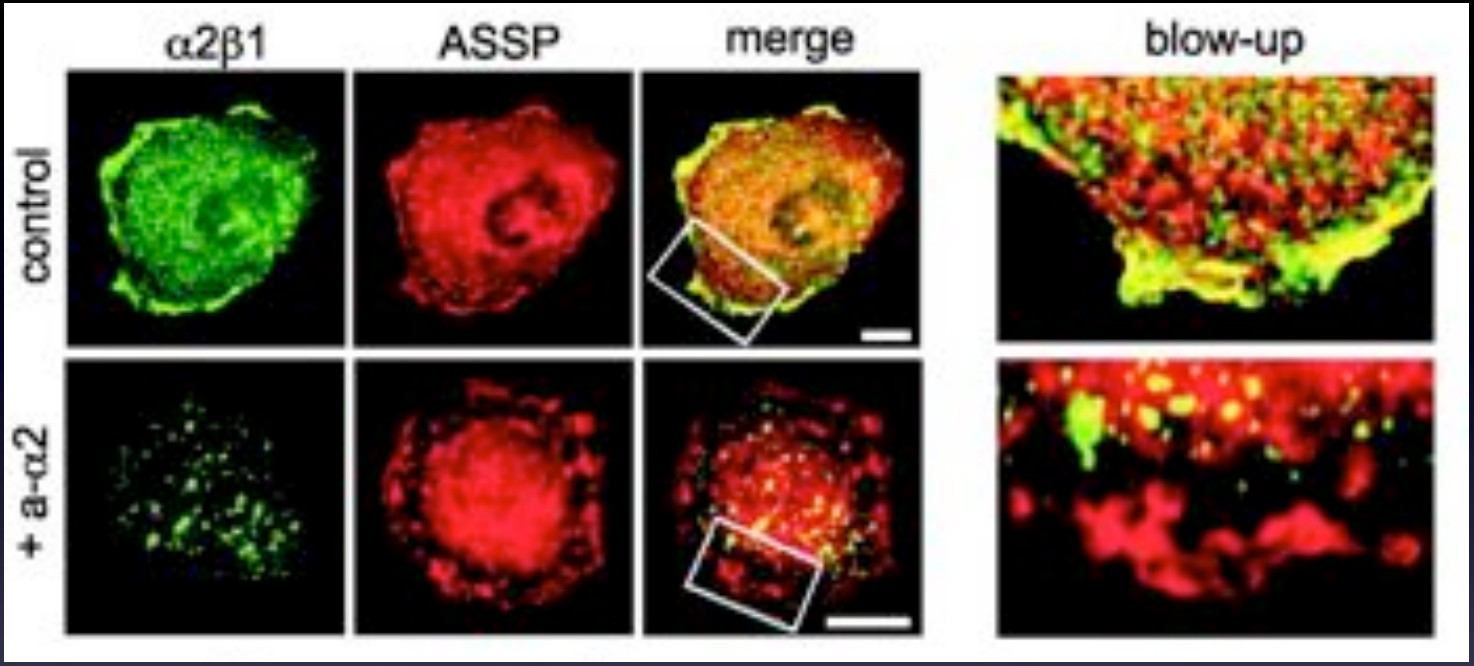
- Required bit depth - 8 bit often enough for LSCM imaging… and colocalisation analysis.
  - More bits only for quantitative experiments where small intensity differences are measured.
  - 12 bit - bigger files than 8 bit.
    (OlympusFV1000, 12 bit only. Zeiss 8,12. Leica 8,12,16.)
  - 16 bit file is 2x bigger in RAM / on disk, than 8 bit !
  - CCD - some cases 12 bit might give better coloc info.

Watch Out - More Holes To Fall Into:

- Laser power - don't bleach area before imaging it.
  - Bleached sample
    - Lower signal : noise
    - Lost information
  - Set the HV and Offset quickly (Auto HV)
  - Live imaging, bleaching - big problem
    Use low laser power (but more noise)



Colocalisation/Correlation

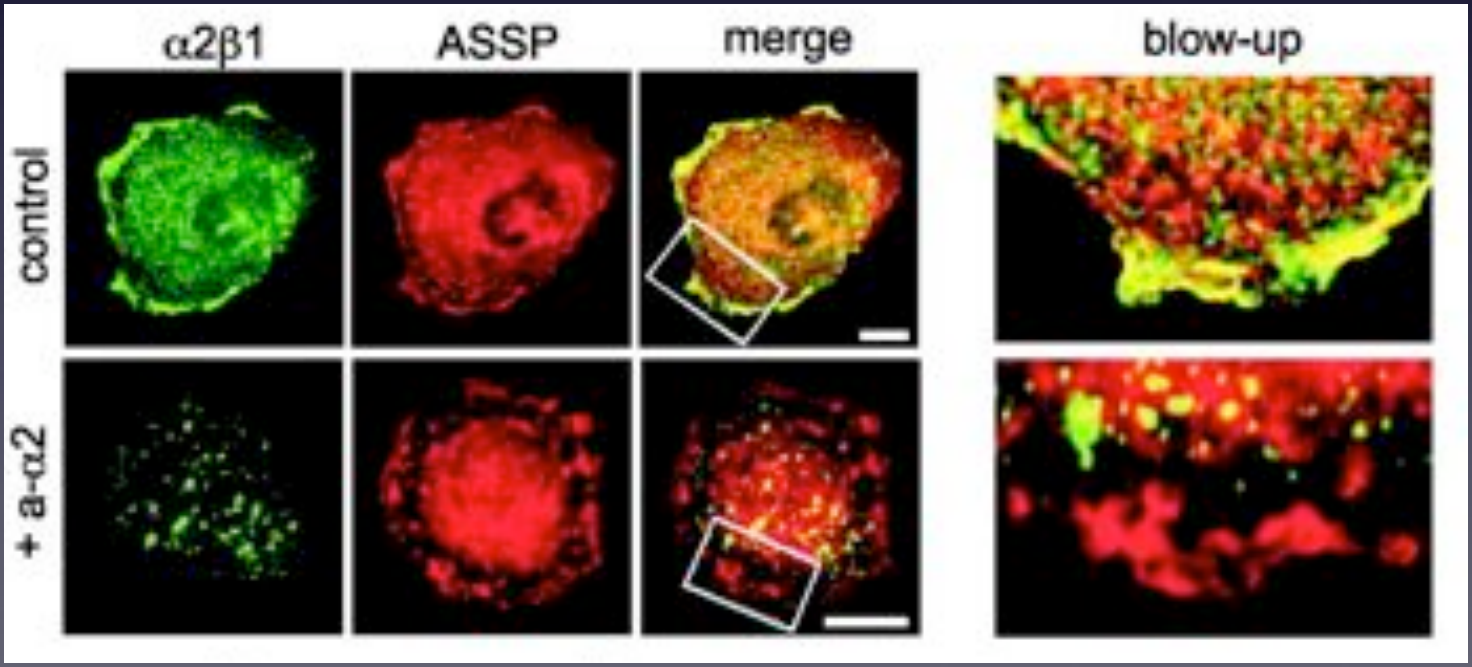


The past:
"I see yellow - therefore there is colocalisation"
but published images "look" over exposed.
No colocalisation definition + No stats = No Science.

From Now On: 3D. Quantification. Correlation. Statistics.
Complementary methods: BioChemical, Optical (FRET, FLIM)

Colour Merge Images? Only for Art!

- Channel Merge Images? What are they good for?
  - Apart from looking pretty... not much.
  - Scientific conclusions from the image below?
  - Colour blind people - see green and red the same!

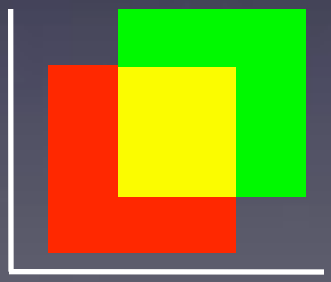


Colour Merge + Projection = Danger!

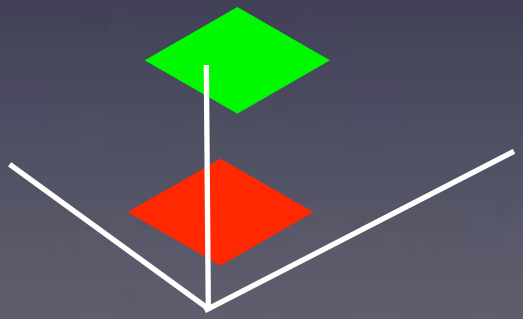
Never make colour merge / overlay images from projections of 3D / z stacks... why not?

Lose 3D info - are the objects overlapping in 3D, or is one in front of the other one, in the z-stack.

False overlaps!!! Easy to make false interpretation



colour merged projection



3 D

What does “Colocalisation” mean anyway…?

- That depends who you ask…
- … and what **BIOLOGY** you are thinking about

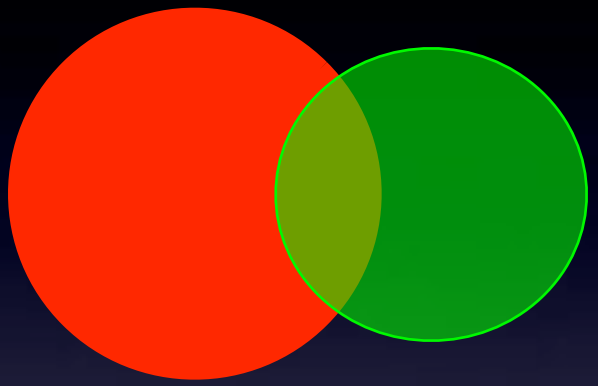
Colocalisation/Correlation? Think about the biology!

- What is the biological/biochemical question?
  - Are you looking for Co-Compartmentalisation?
  - Are you looking for exclusion / anti correlation?
  - Are you looking for interacting molecules?
    - Then you must also do biochemsitry (Immuno Co-precip, Fluo Correlation Spectroscopy)
    - FRET / FLIM might be very informative

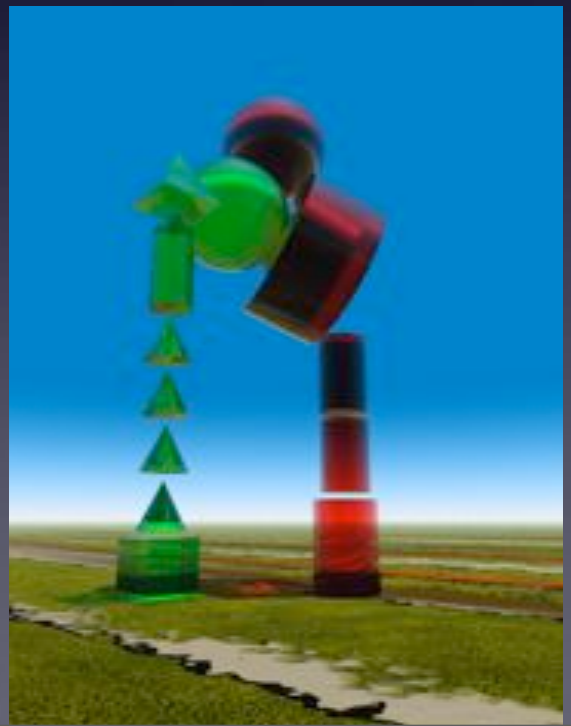
Colocalisation / Correlation / Concurrence?

"Colocalisation" covers two qualitatively different conditions:

1) that objects have both fluorophores present
(Object Based Coloc)
Segmentation needed.
Biology?



2) there is some relationship between the **intensities** of the fluorophores in a pixel.
(Pixel Intensity Based Coloc)
Interaction - BioChemistry?



Colocalisation / Correlation / Concurrence?

2 fluorophores are there in a pixel

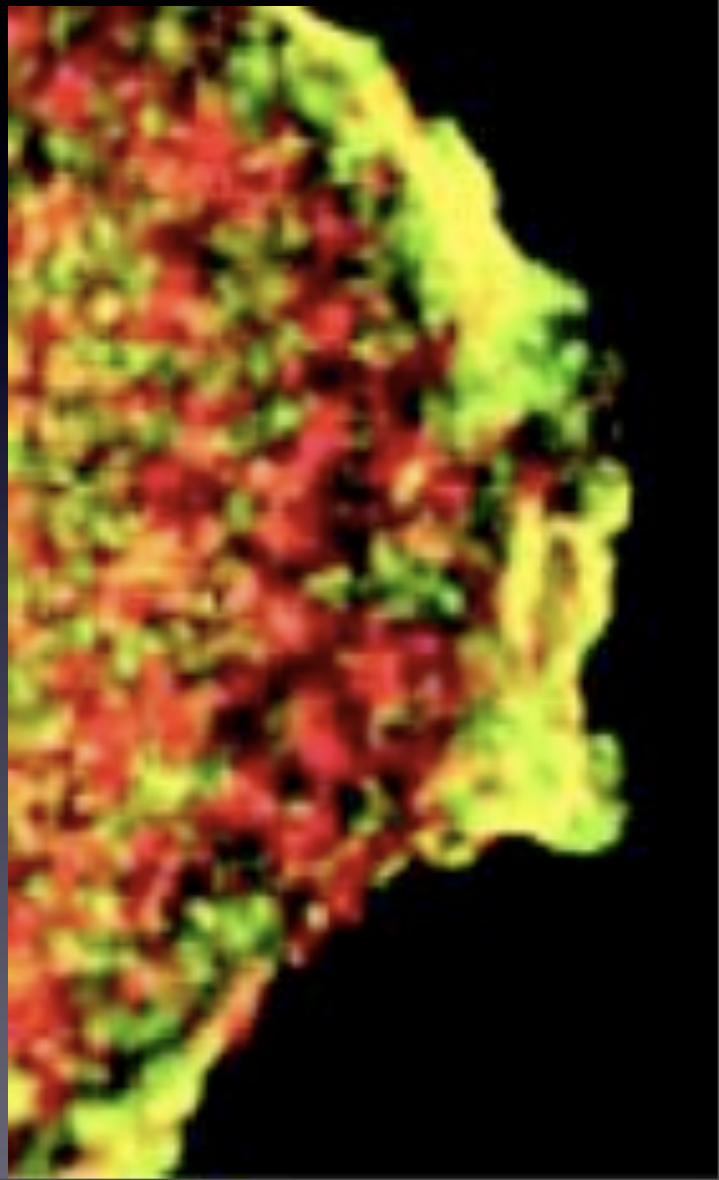
Binary information

Is it Random?

Is it Real?

Little or no biological meaning?

…unless you are confident about how to segment objects out from the background.

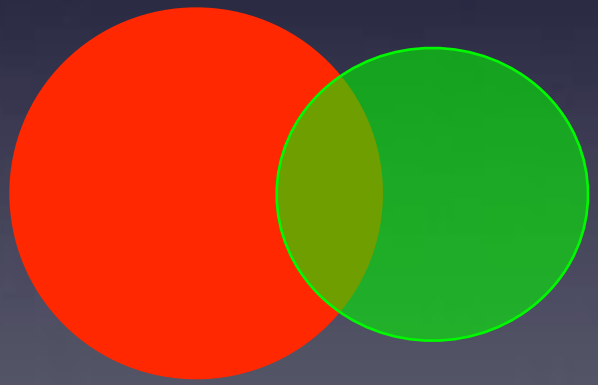


Definition of Terms

- “Concurrence” = “co-presence” “there is red and green”
- “Colocalisation” = Relationship between channel intensities
  - Eg. “Red is only found with Green”
  - Special case - “Correlation”
    - **Intensity** Correlation over Space

Define what is Colocalisation/Correlation?

Colocalisation is #1

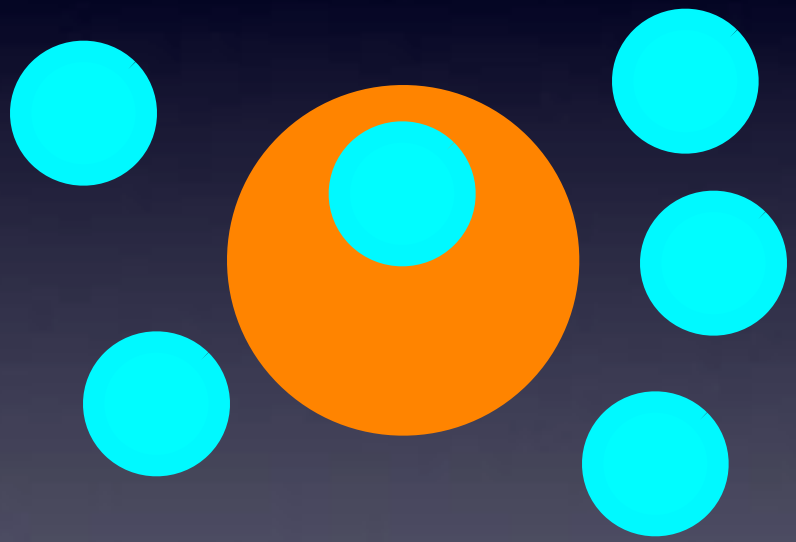


2 objects overlap
Binary information
No intensity information

Concurrence?
Image Segmentation!

Biological Meaning?

Colocalisation is #2

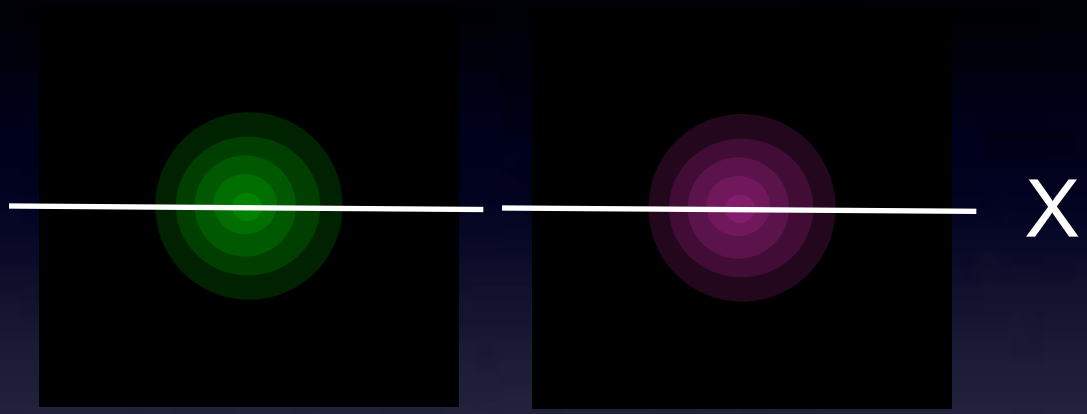


Some objects appear to overlap
with another object
Binary information
No intensity information

Colocalisation?

Biological Meaning?

Colocalisation is: #3



Intensity profiles overlap

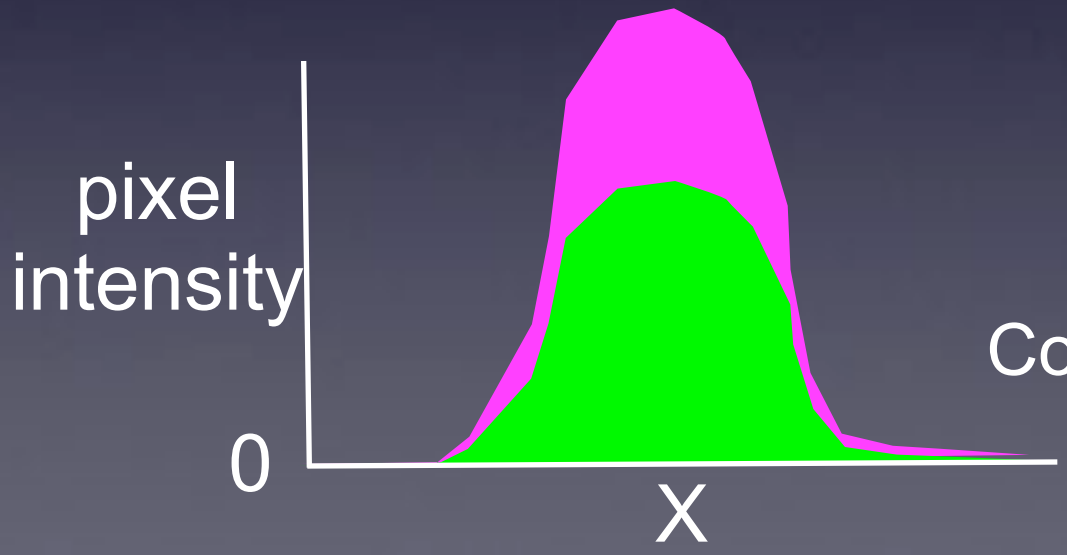


Image “Correlation”

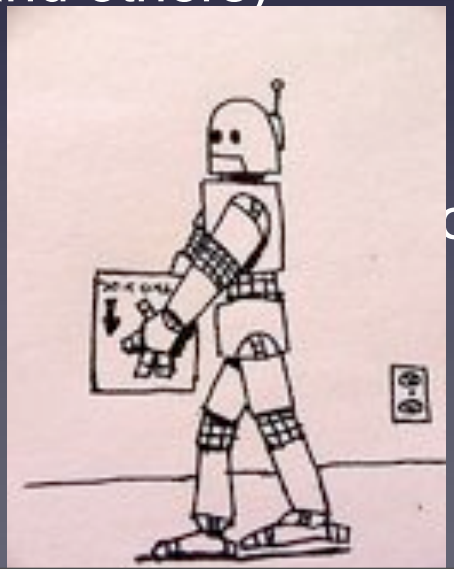
Biological Meaning?
Co-compartmentalisation?
Physical interaction?

Colocalisation/Correlation -Think about:

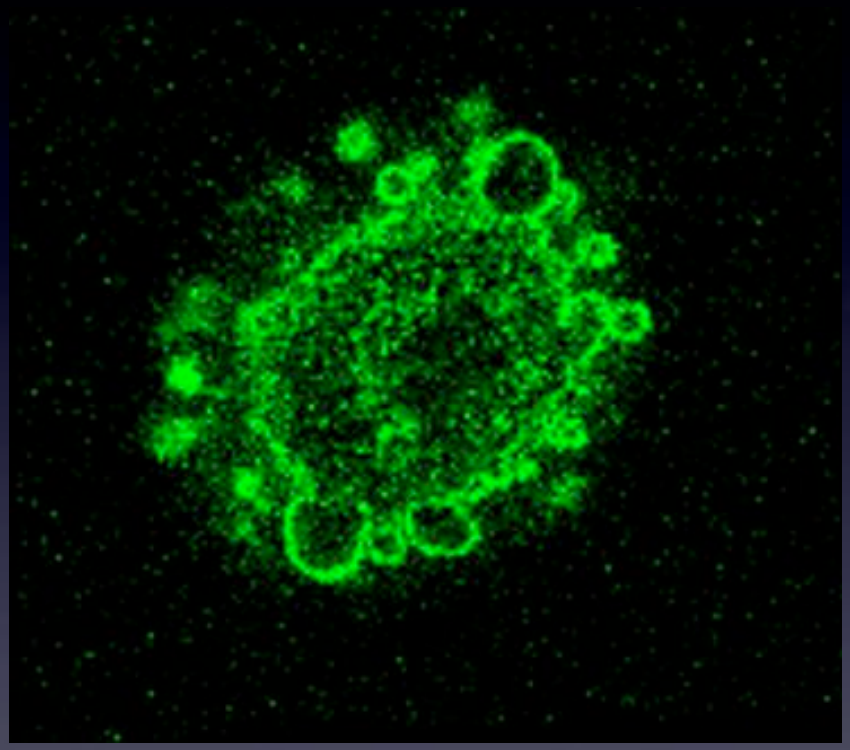
- Are your “objects” smaller than optical resolution?
  - Vesicles? Small Organelles?
    - Check channel overlap with sub resolution beads!
- Are your objects large?
  - Large single homogenous blobs?
  - Large reticular networks / membranes
  - Resolution required?
- Complementary “correlation” methods
  - Fluorescence correlation spectroscopy (FCS in live cells)
  - Flow Cytometry? Multiple markers in a cell. Good stats.

Colour Merge Images = Bad
… so what should I do instead?

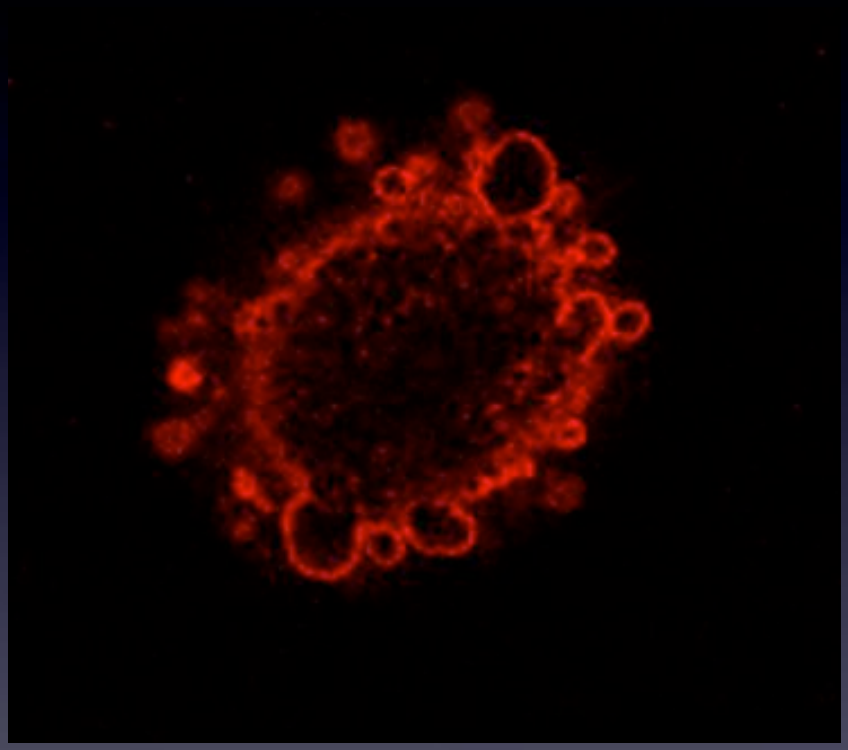
- "Colocalisation Analysis"
- Statistical Significance of Colocalisation
  - Single image - random / insignificant.
  - Statistical P value (significance), Manders coefficients, and Scatter Plot. (ImageJ, BioImageXD, Imaris and others)
- But remember…
  - Don't merge projections of stacks ou lose 3D info, false coloc)
  - Don't believe your eyes, they lie. Machines don't make mistakes…



Colocalisation Analysis



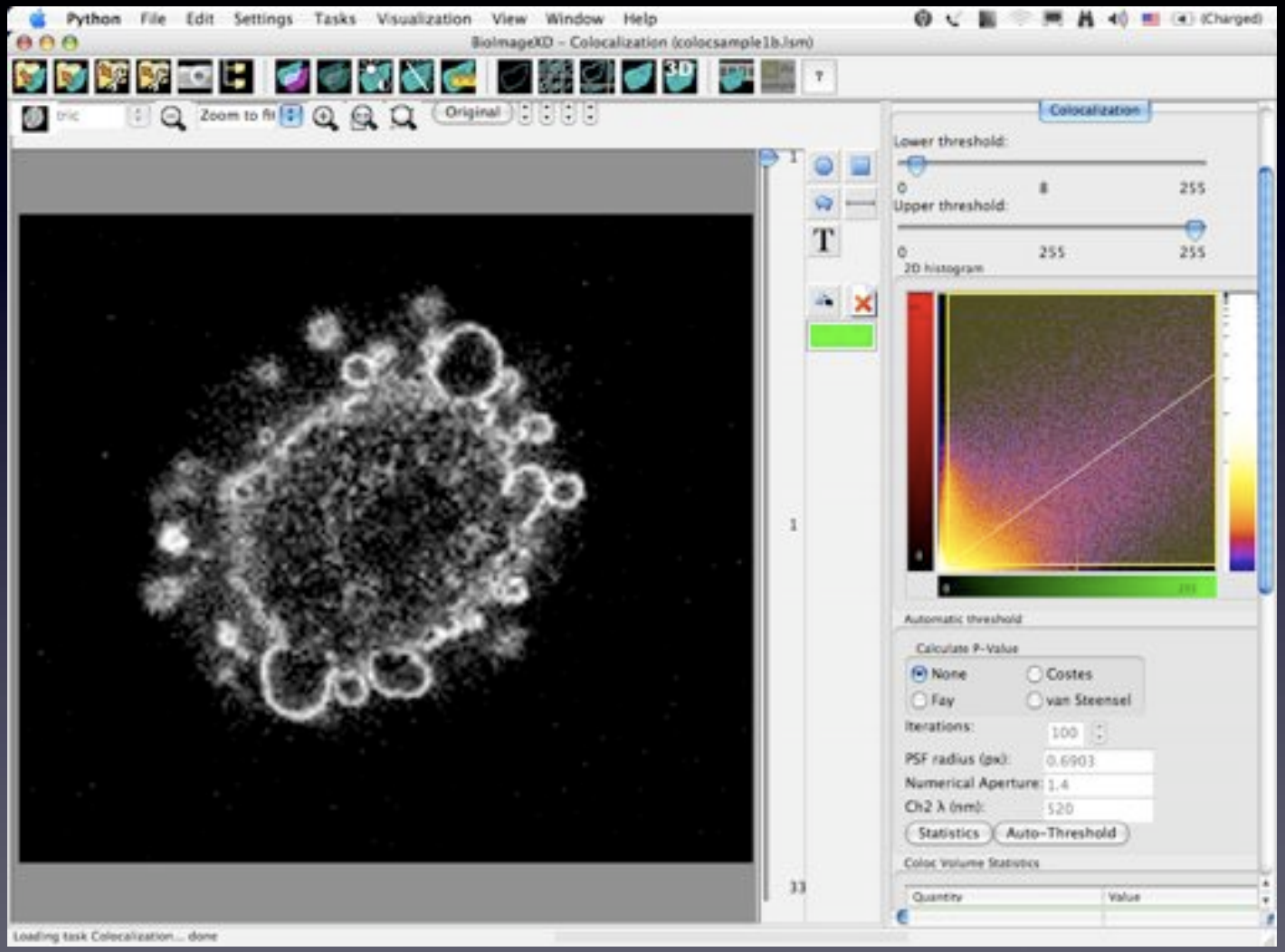
vs.



How can I measure the amount of colocalisation or rather “correlation” between these two images?

BioImageXD, ImageJ and others have methods to do that!

Colocalisation Analysis



Scatter plot
2D histogram
Publish it?

Coloc stats:
Pearsons *r*
M_1, M_2,
Costes P-val,

Automatic thresholding

Coloc Stats - Costes et al. 2004 Biophysical J. vol 86 p3993

Pearson's Image Correlation Coefficient (Manders et al., 1993)

$$r = \frac{\sum_i (R_i - R_{av}) \cdot (G_i - G_{av})}{\sqrt{\sum_i (R_i - R_{av})^2 \cdot (G_i - G_{av})^2}}$$

Don't panic - it's not that complicated!

Correlation between images, *r* ranges from -1 to +1

+1 means full correlation (images are the same)

0 means no correlation (random)

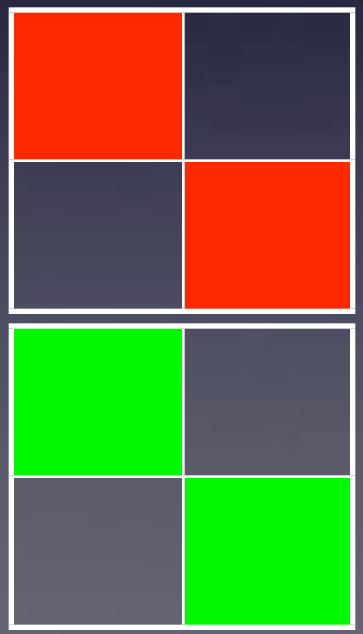
-1 means full anti correlation (no red where there is green)

Pearson's Image Correlation Coefficient

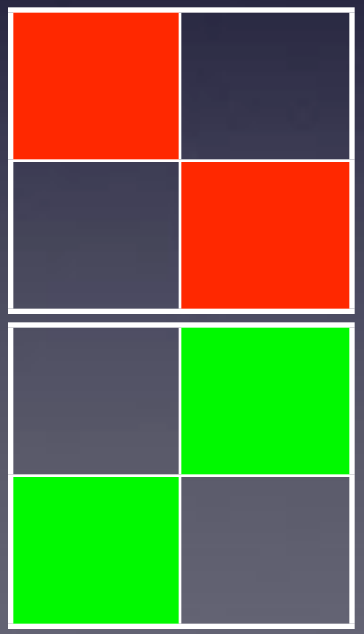
In English…per pixel and summed for the whole image:

$$r = \frac{\text{sum of (red intensity - average red) x (green intensity - average green)}}{\text{sqrt of squares of above}}$$

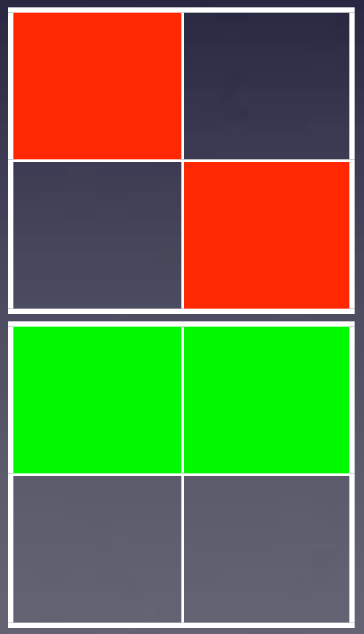
r = +1



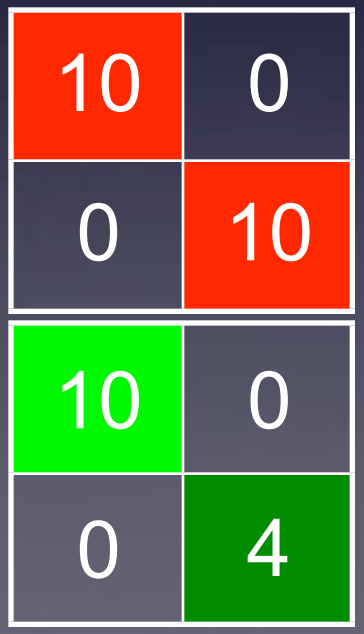
r = -1



r = 0



r > 0



Pearson's Correlation Coefficient is good because…

- Not sensitive to diff intensity of the 2 images. Why?
- If red is 1/2 as bright as green…
  - Still get Pearson's *r* of 1, as the correlation is the same,
    - only realtive intensity is different.

r = +1

10	0
0	10

10	0
0	10

r = +1

5	0
0	5

10	0
0	10

Pearson's Correlation Coefficient is BAD because…

- A single coeffiicent *r* describes the whole sitiation
  - but it's different from the perspective of either colour
- Pearsons *r* is ambiguous because of the strong influence of the ratio of the number of objects in both channels.
  - There is different "stuff" in the 2 colour channels.
- We need a way to decribe colocalisation from the perspective of BOTH colour channels:
  - **Manders' Coefficients M1 and M2**

Manders Coefficients

$$M_R = \frac{\sum_i R_{i,coloc}}{\sum_i R_{i,total}}$$

$$M_G = \frac{\sum_i G_{i,coloc}}{\sum_i G_{i,total}}$$

Biologically meaningful coloc coefficients:

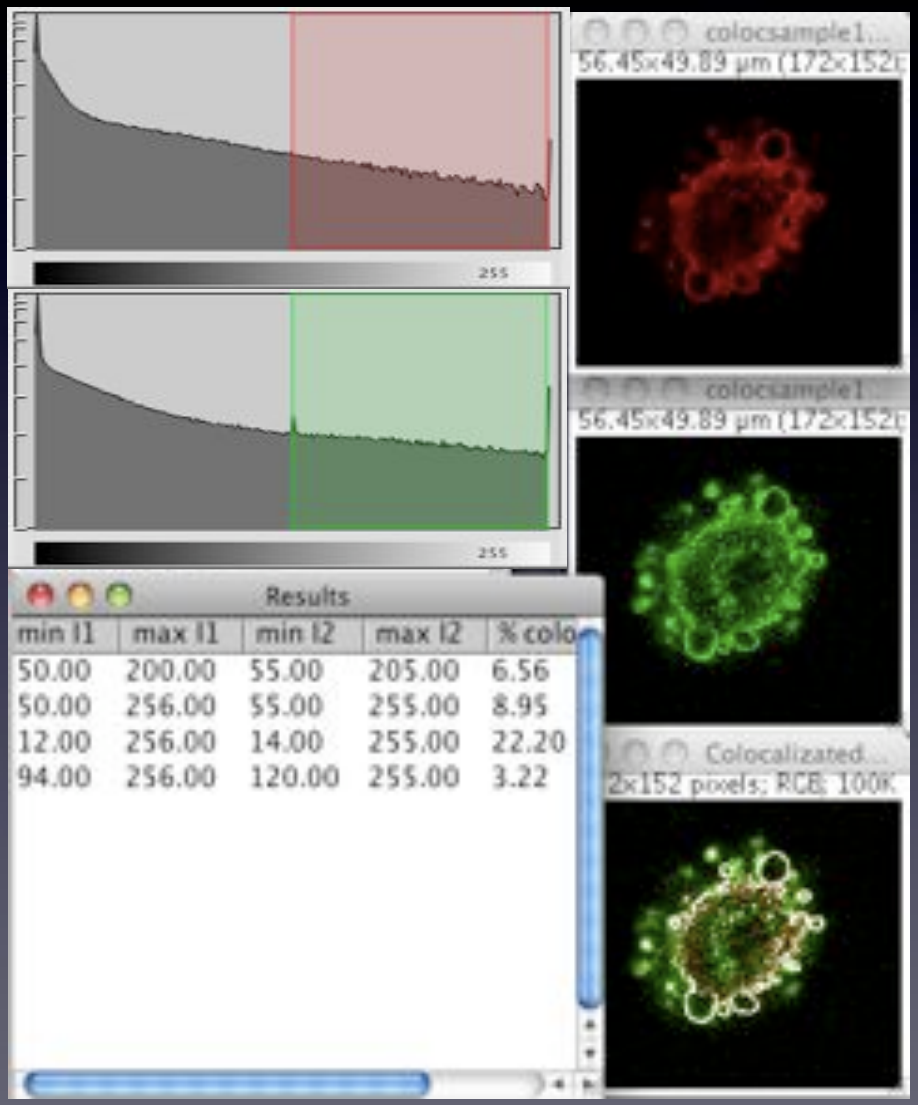
Proportion of each dye colocalised with the other (Manders et al., 1993)

$R_{i,coloc}$ = colocalised red signal

$R_{i,total}$ = total red signal

Great! … but how do I know which pixels are colocalised and which are not…?

“Thresholding” and “% colocalisation”



The calculated “% colocalisation” depends on what thresholds you set.

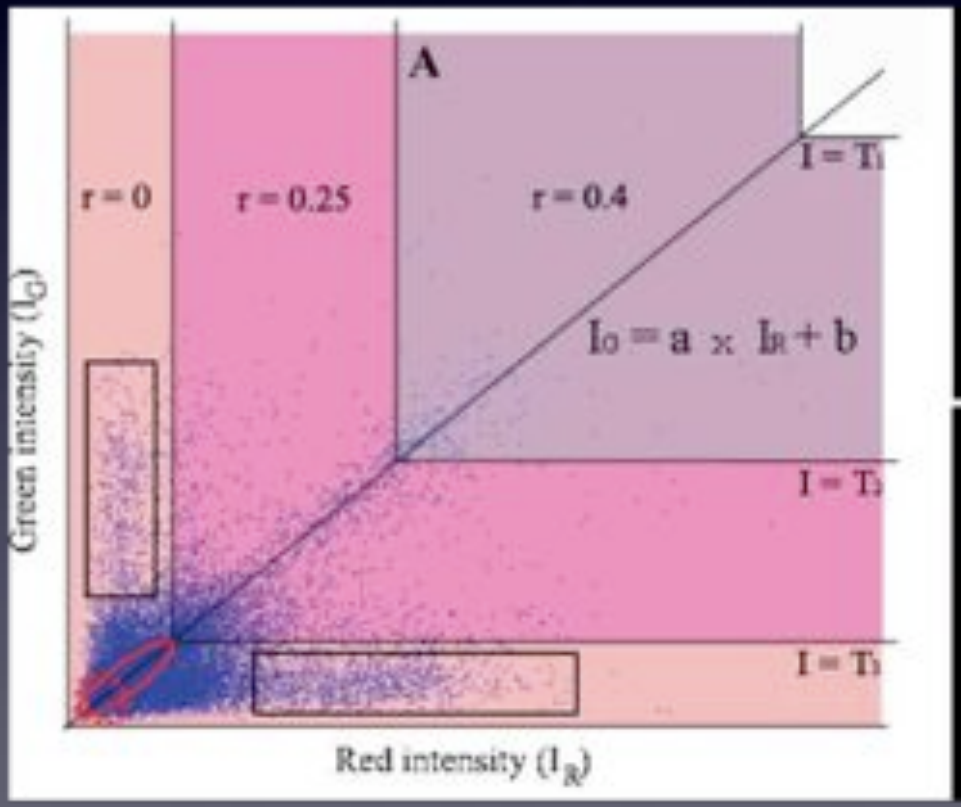
… so how should one set them?

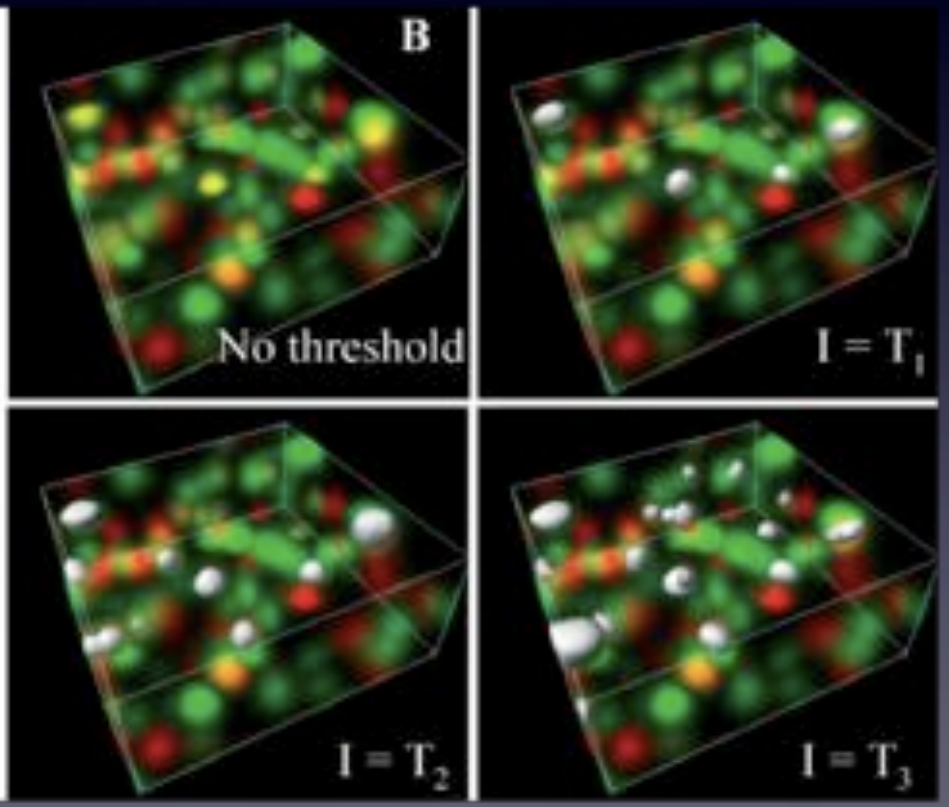
..until you get the result you want?

No science here!

Automatic Thresholding?

- How should I set the thresholds of the 2 channels?
  - Manually? No! Subjective user bias, not reproducible...
  - Need a robust reproducible method!
  - Find thresholds where Pearson correlation below thresholds <= 0

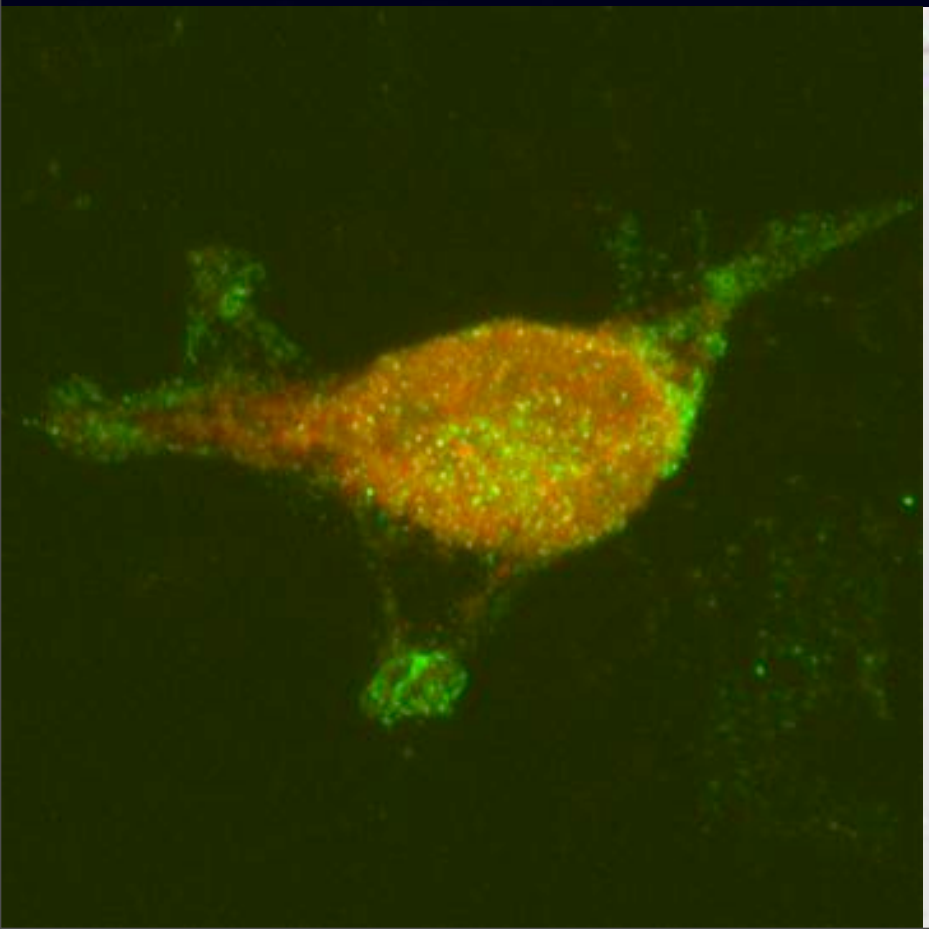


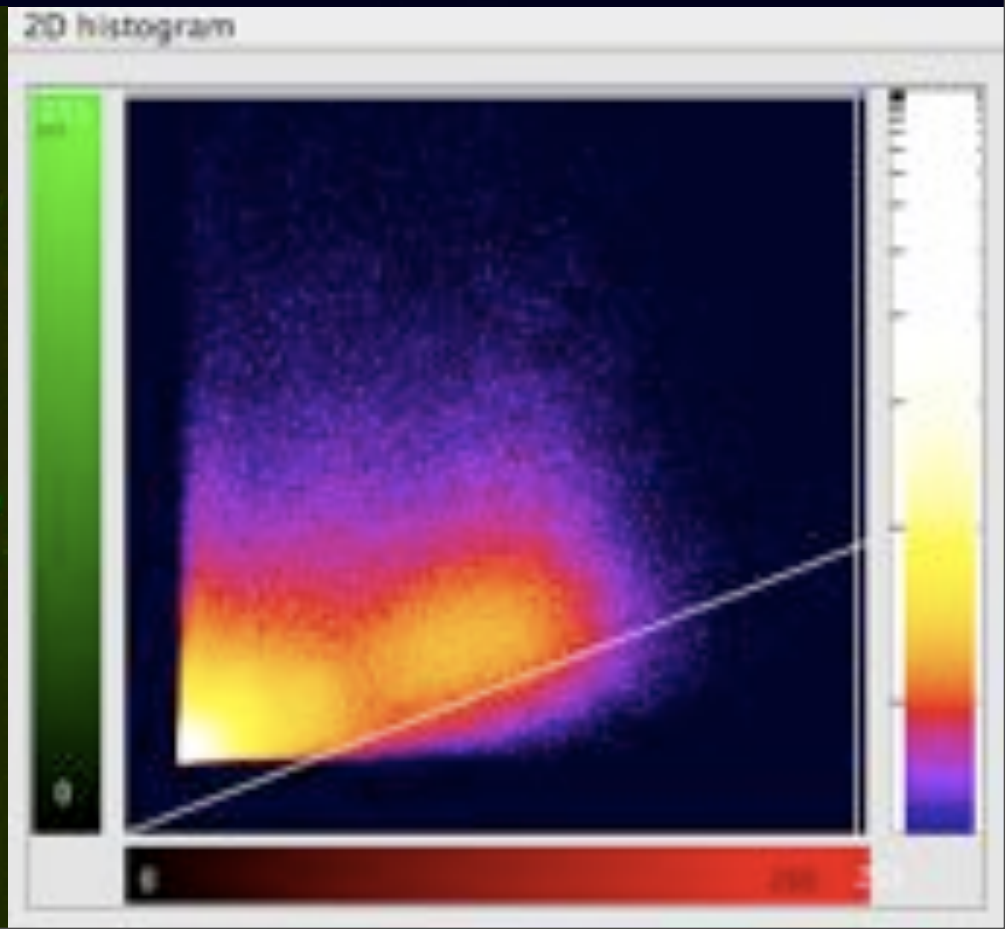


Auto Threshold - Costes et al. 2004 Biophysical J. vol 86 p3993

2D Histograms / Scatterplots

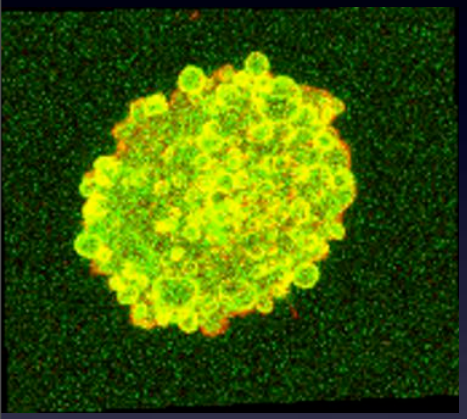
- Display 2 colour channel image data in 2D:
  - colour merge / overlay or 2D histogram?
    - 2D histogram: Ch1 - y axis (left), Ch2 - x axis (bottom)



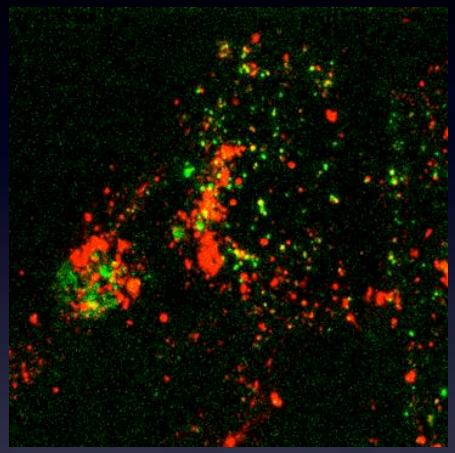


2D Histograms / Scatterplots

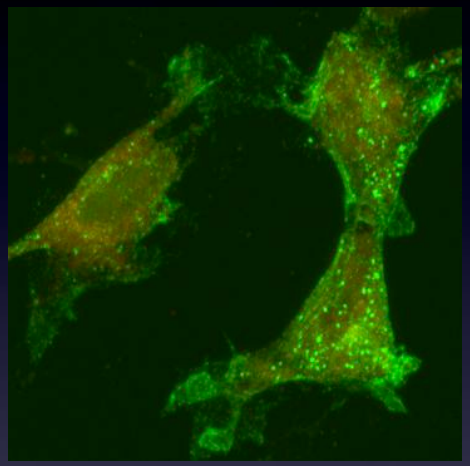
- See correlation qualitatively - better than colour merge
- See problems from imaging:



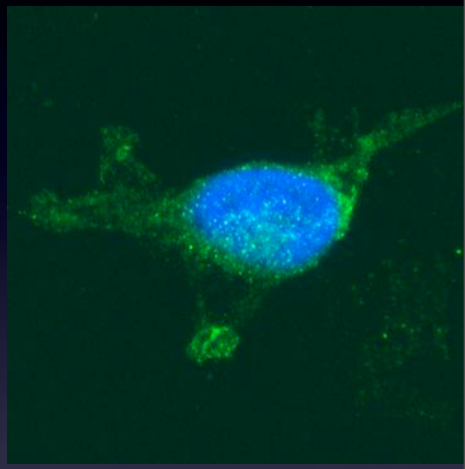
Saturated
Noisy



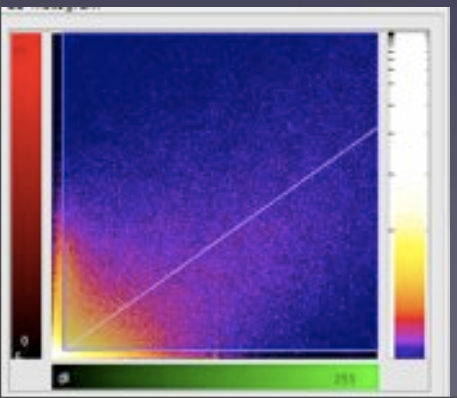
Saturated
No correlation?

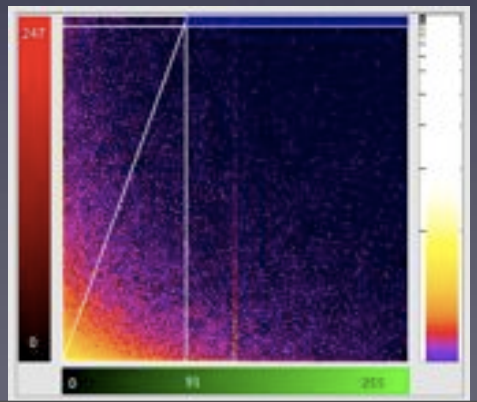


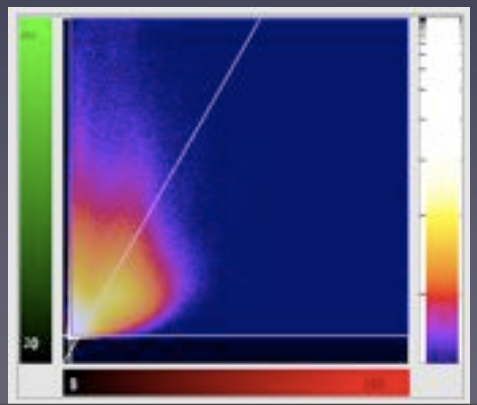
Wrong offset

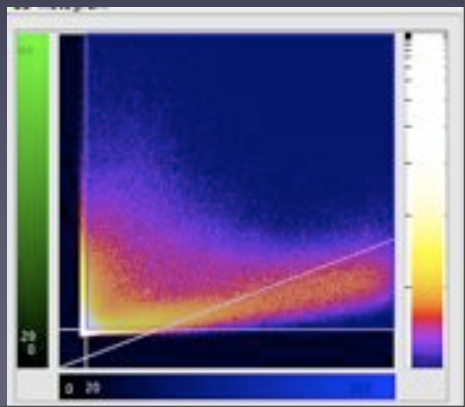


Wrong offset
Bleed through



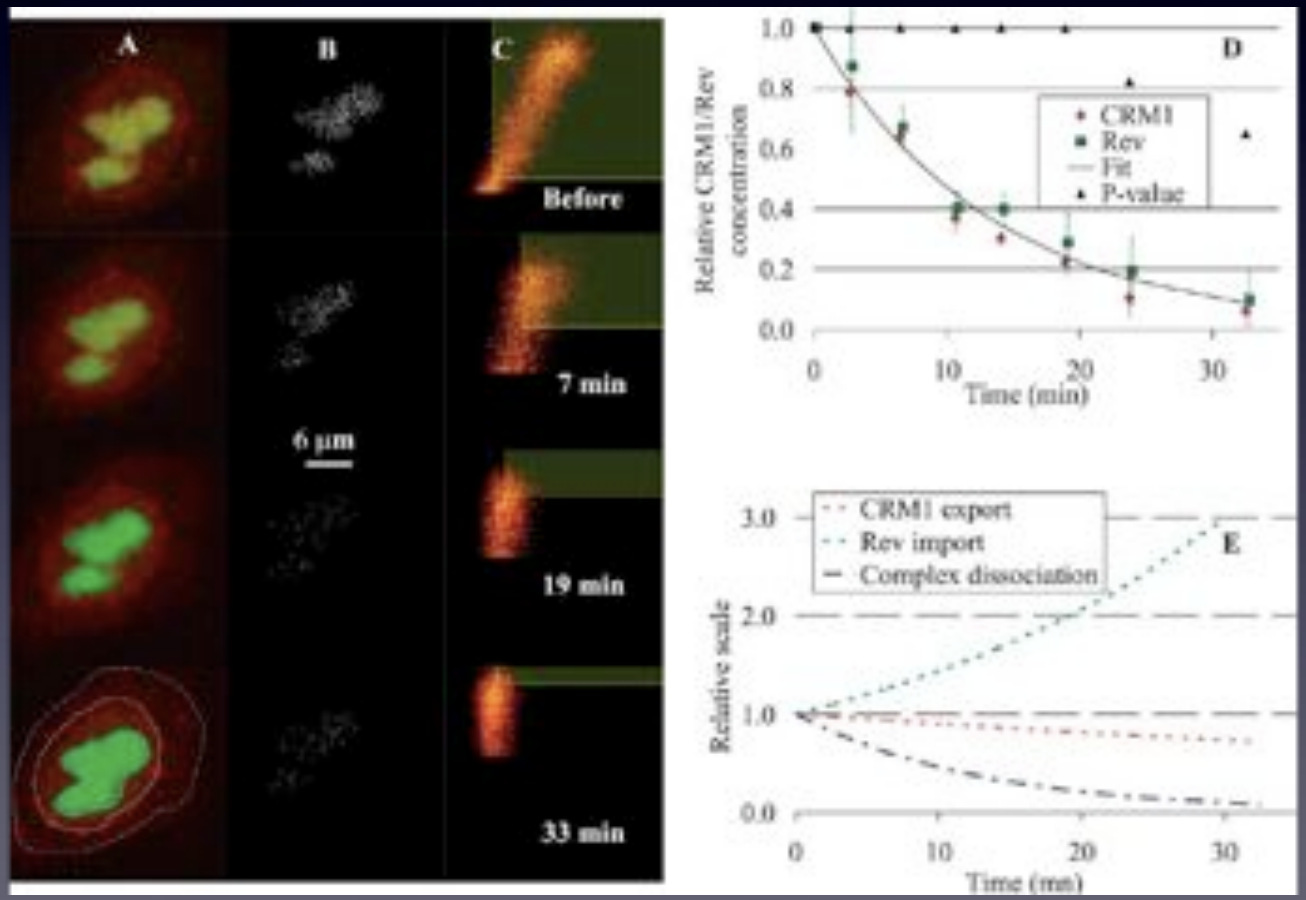






Automatic Thresholding?

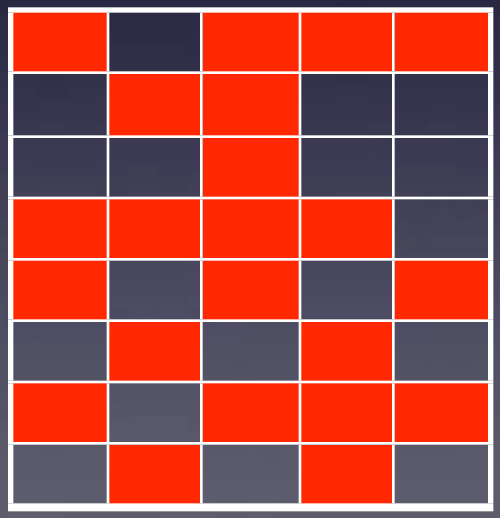
- Does it work in a biological experiment? Yes!
  - Time course of Rev-CRM1 dissociation, nucleolus to nucleus
  - The dissociation rate constant kd =1.25 ± 0.31 x 10^{-3} s^{-1}



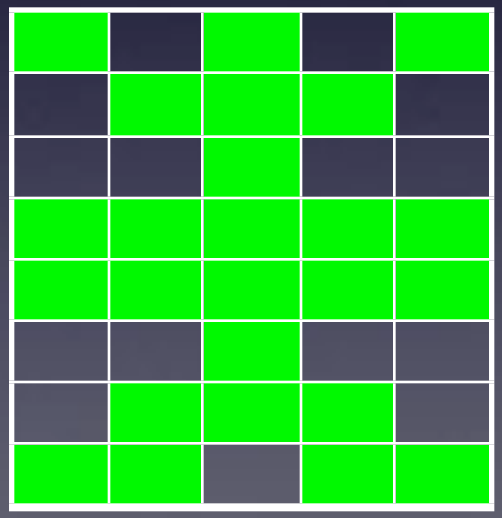
auto threshold - Costes et al. 2004 Biophysical J. vol 86 p3993

One more thing…

- Statistical significance!
  - Are coloc results better than random chance?
  - A busy image might give high correlation and Manders
    - Lots of signal = larger chance of random signal overlap.



vs.



17 / 40 pixels overlap !!!

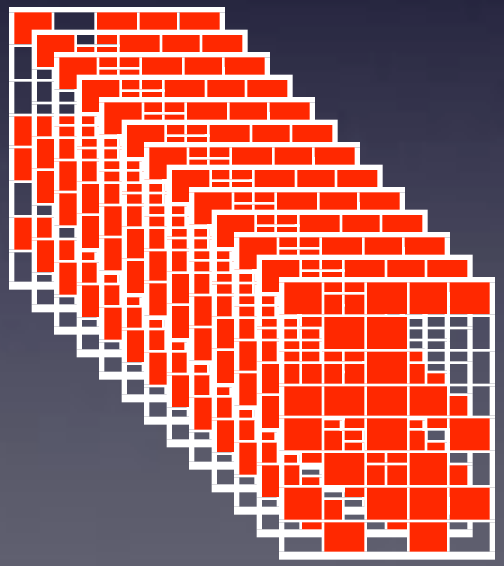
Is that significant or just random?

Statistical confidence P - Costes et al. 2004 Biophysical J. vol 86 p3993

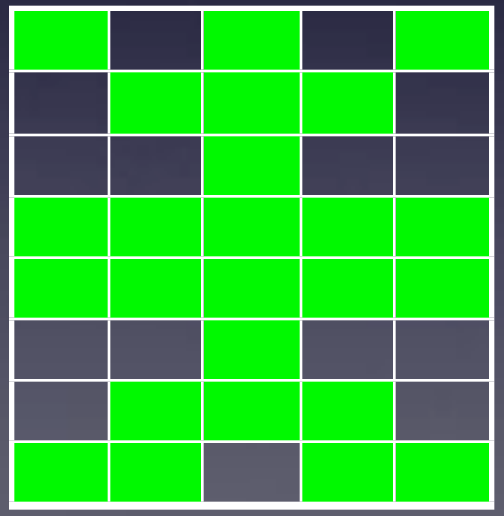
Costes Method - Randomisation…

- Measure Pearson's correlation for:
  - Randomised 1st channel image data (PSF sized chunks)
  - Repeat 100 times
  - How many randomised have <= correlation than real image.
    - If > 95% of randomised are worse, then we believe Manders.

P = 0.5 = 50% (no)
P = 0.95 = 95% (yes)
P = 1 = 100% (YES!)
confidence



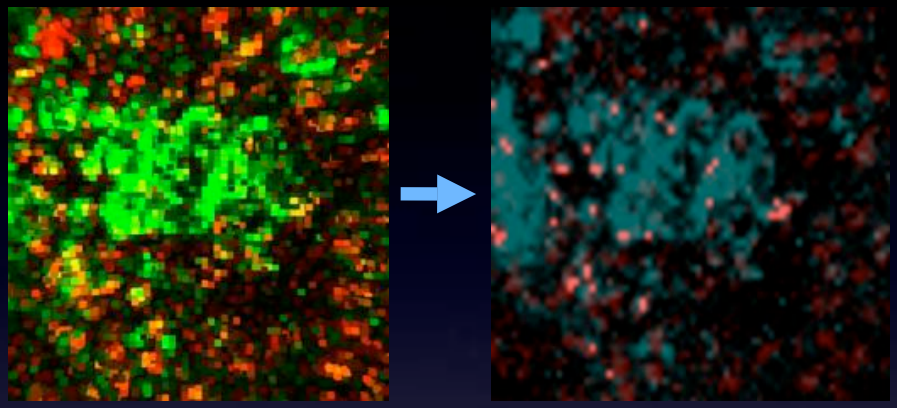
vs.



Statistical confidence P - Costes et al. 2004 Biophysical J. vol 86 p3993

Colocalisation example: virus entry to caveolae

10 min P.I.

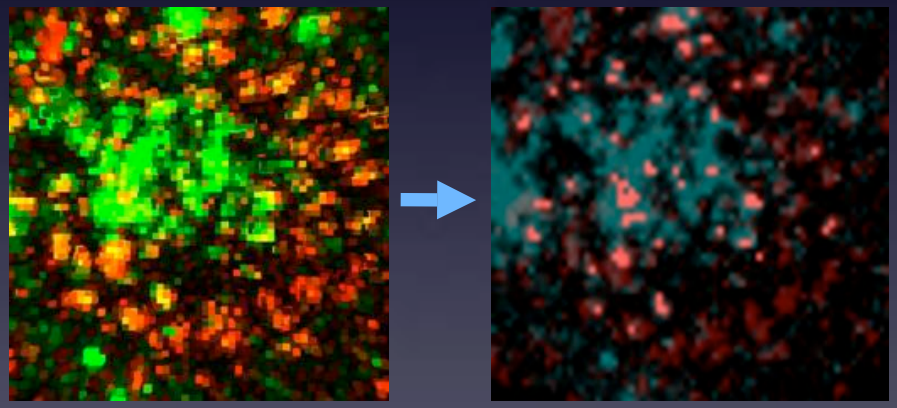


32% of virus colocalized

Costes P-value 0.00

0% chance it's real

20 min P.I.



39% of virus colocalized

Costes P-value 1.00

100% chance it's real

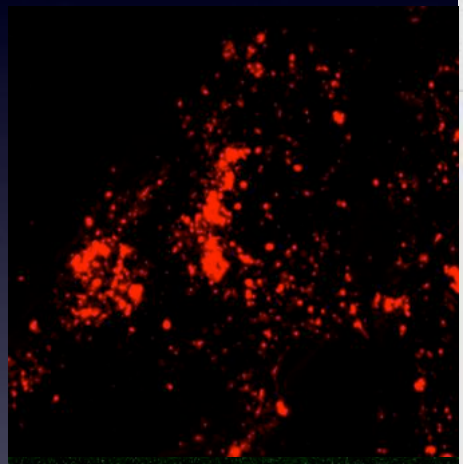
Without significance test, we wrongly assume virus is colocalised with caveolae at 10 min P.I.

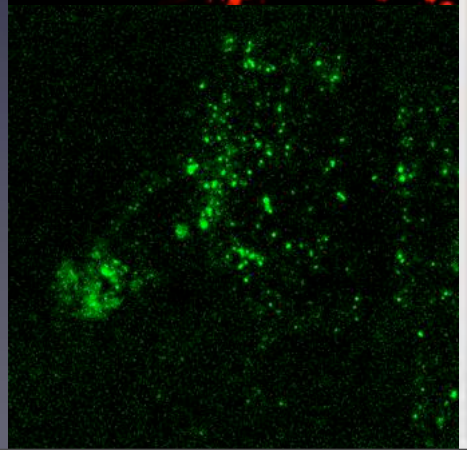
It is not! Only at 20 min is there signficant correlation.

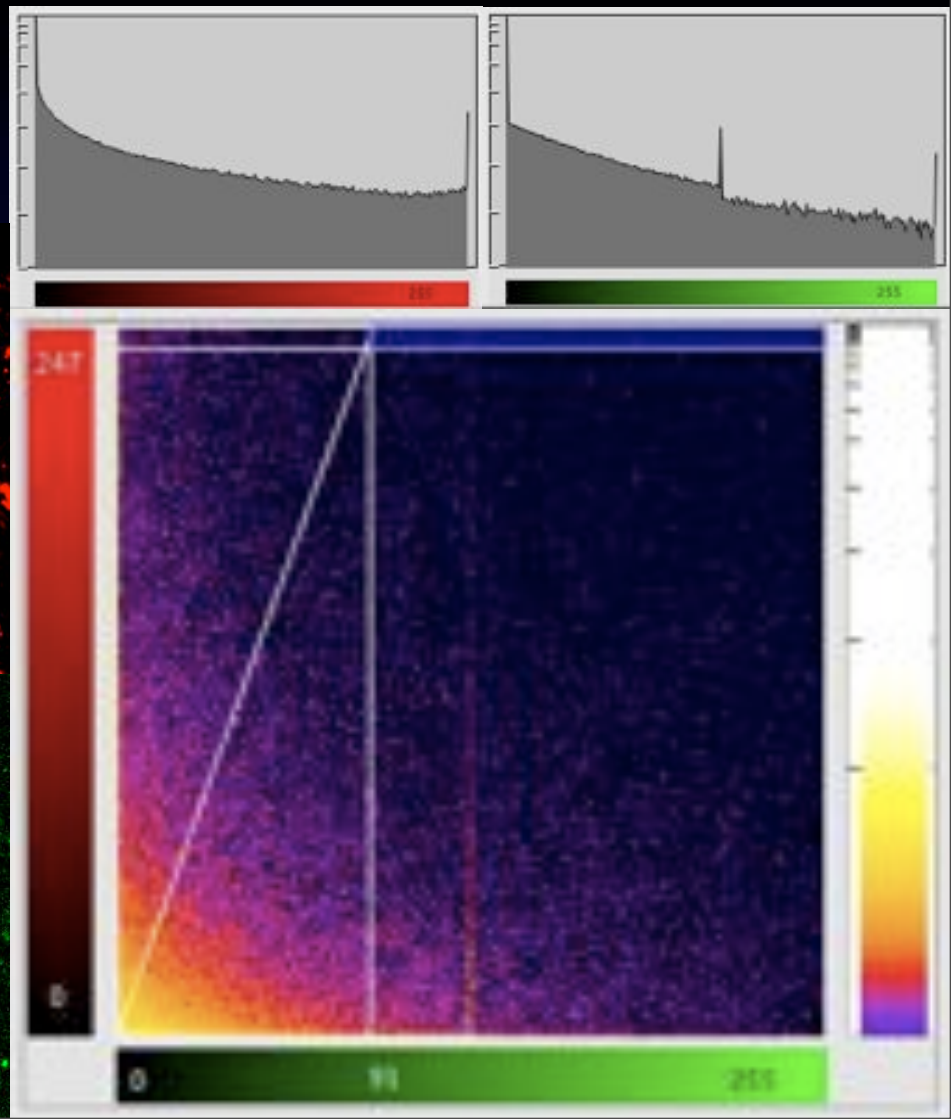
Examples:
No Correlation?

Pearson r 0.024
M1 0.0354
M2 0.0471

Why high Thresholds?





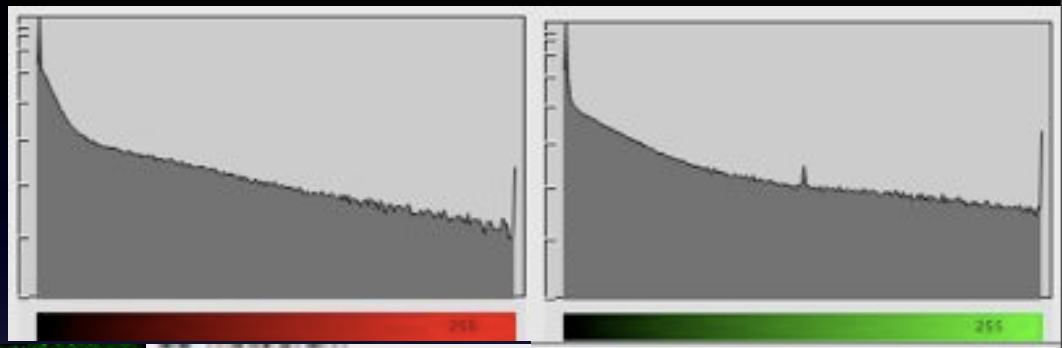


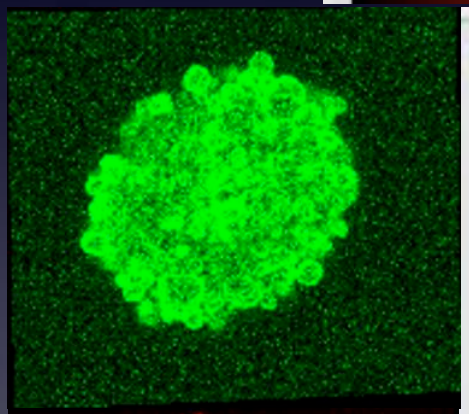
Noisy Saturated Images
Good Correlation?

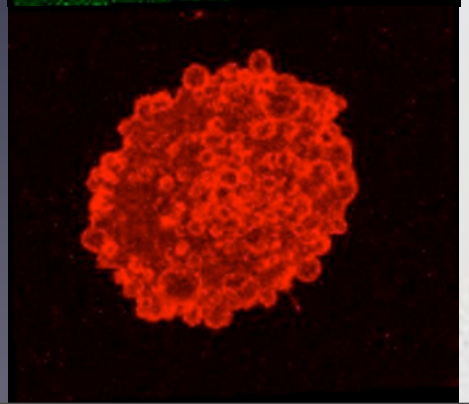
Pearson r 0.747
M1 0.7291
M2 0.7420

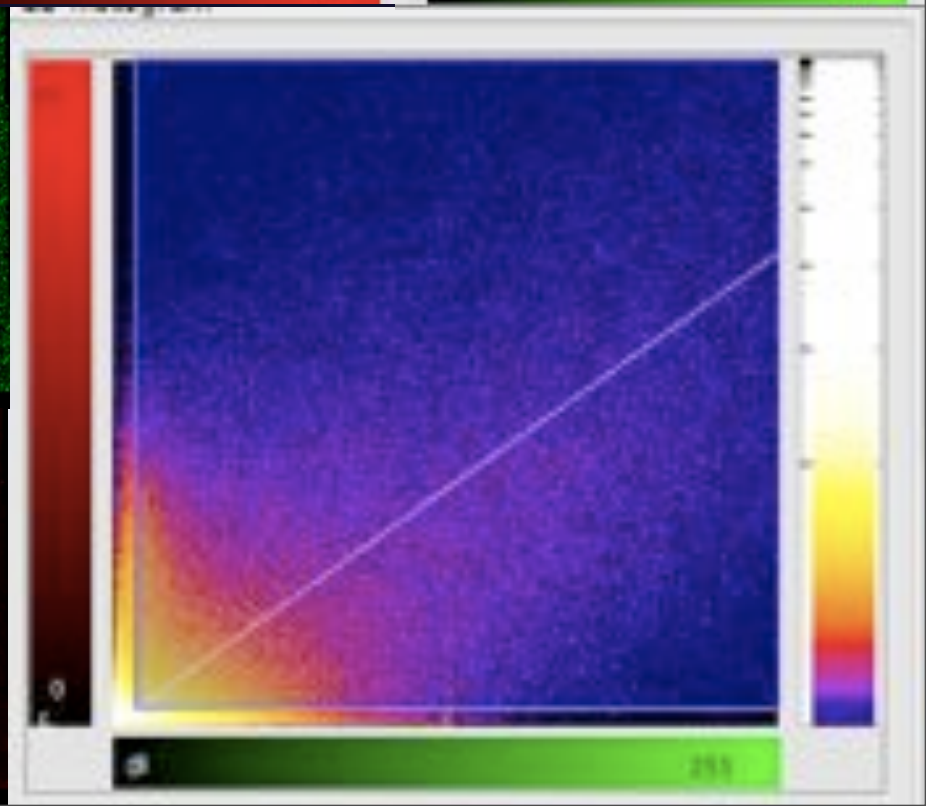
Thresholds
Include
noise?

Badly
Saturated!







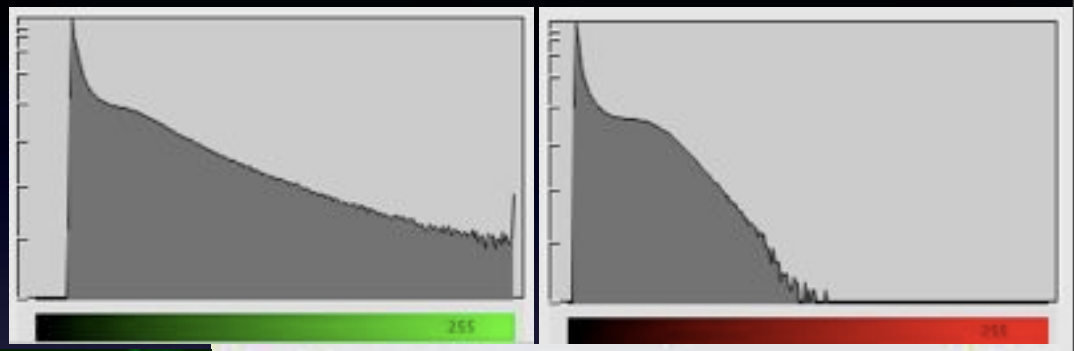


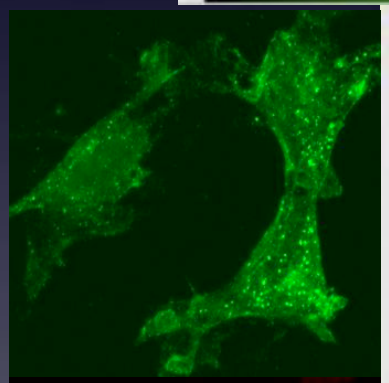
Bad detector settings
Good Correlation?

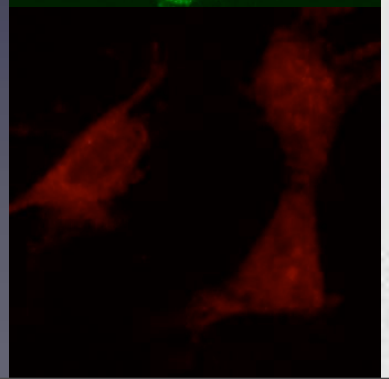
Pearson r 0.68
M1 0.77
M2 0.63

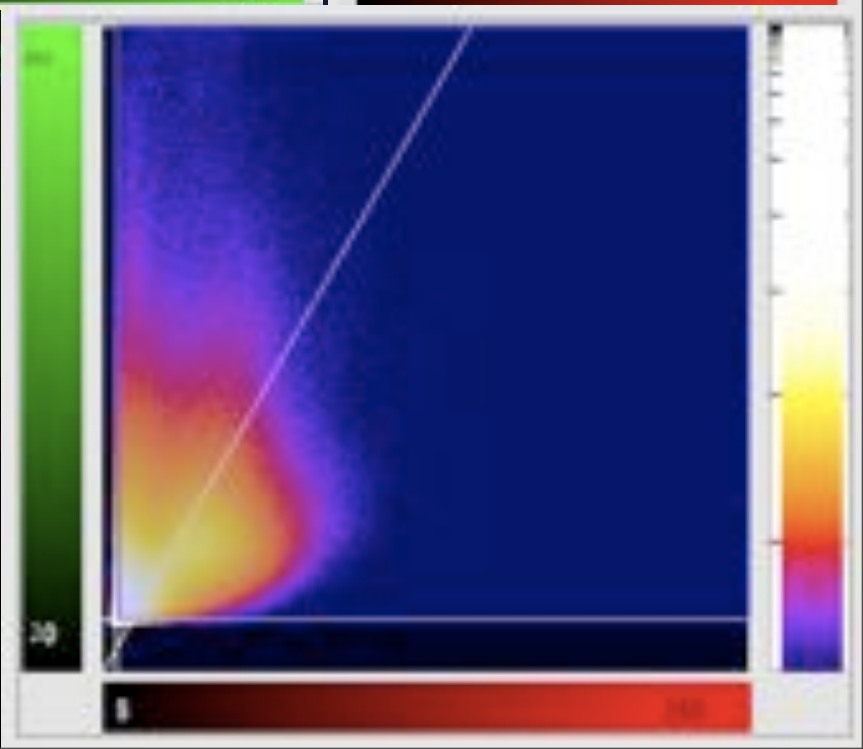
Offset wrong
+ Saturated

Thresholds
Handle it?
No?

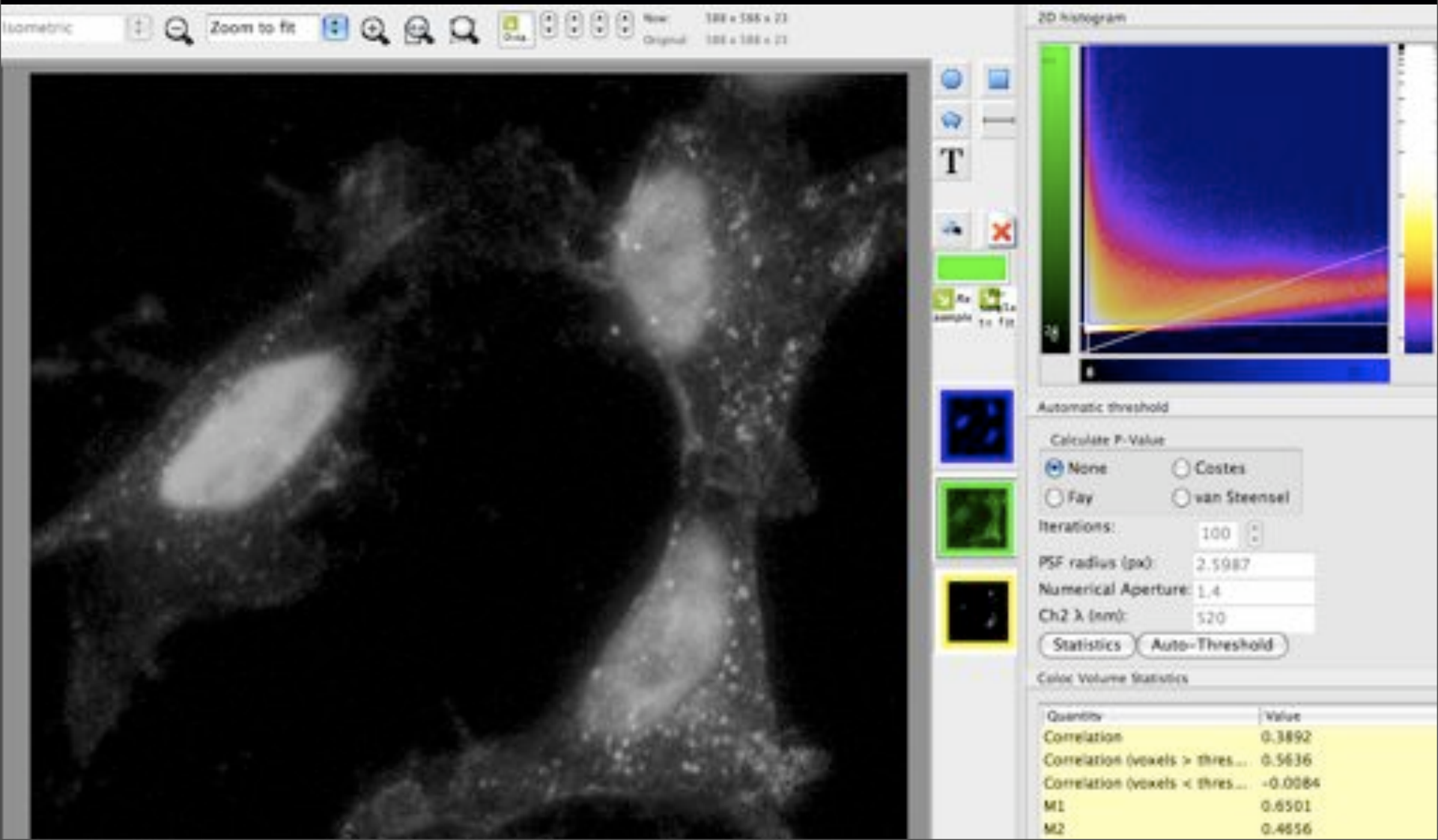




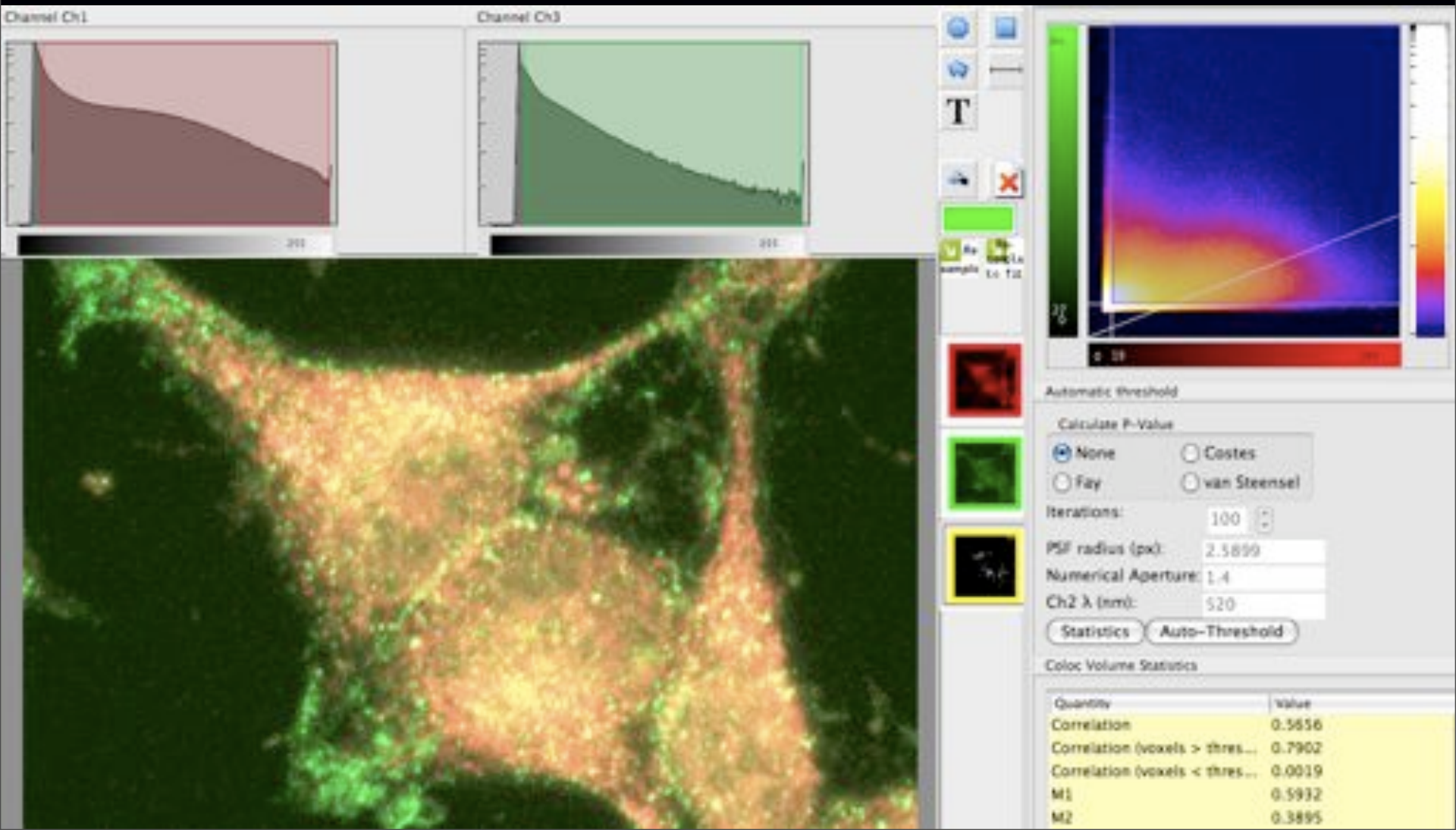




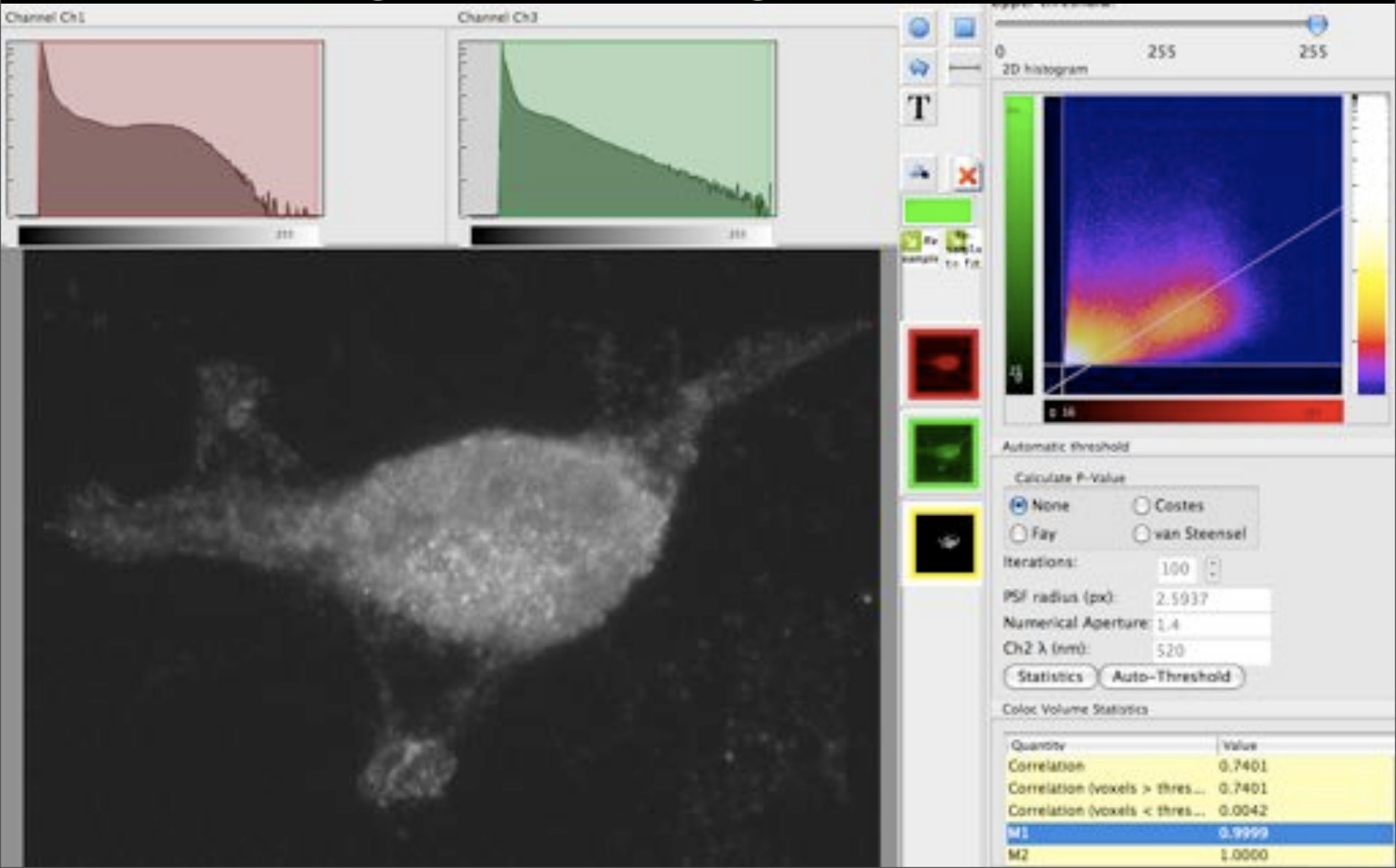
Bleed Through!
DAPI into GFP



Bad detector settings
Good Correlation? Bleed through?



Bad detector settings…
…gives wrong results!!!



Software for Colocalisation

ImageJ - Fiji : Colocalisation plugins
http://pacific.mpi-cbg.de/wiki/index.php/ColocalizationAnalysis

BioImageXD (Coloc Task - Pixel Intensity and Object based methods)

Imaris (Coloc module)

... and others.....


Thanks to: MPI-CBG LMF and IPF
Heino, Pahajoki,
Kankaanpää, Marjomäki
Uuksalainen, Paavolainen,
TEKES,
Thanks for listening