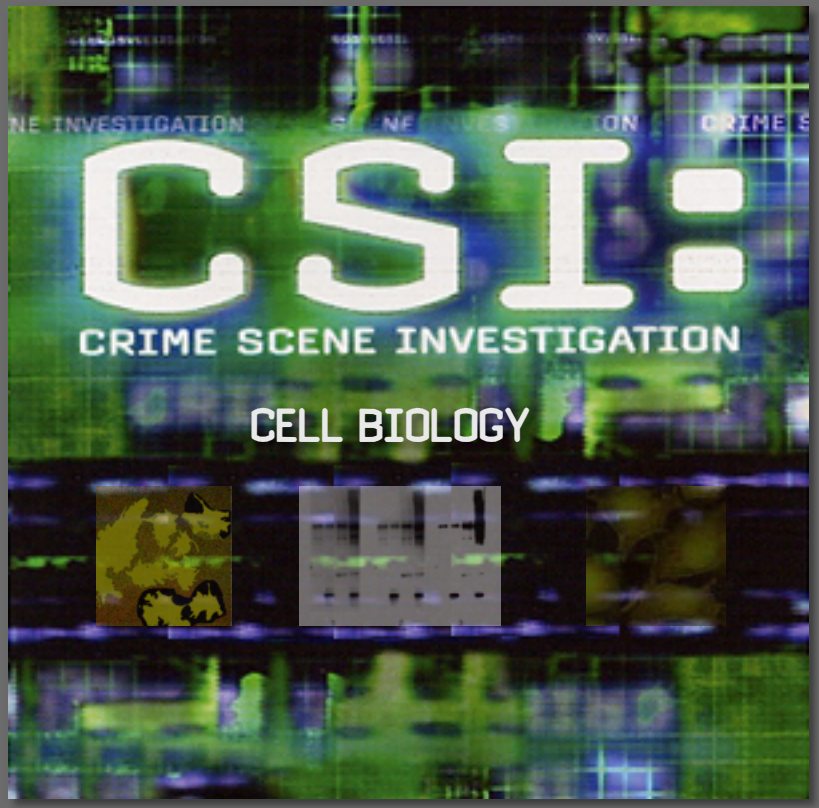
# CSI – Cell Biology:
# The Ethics of Digital Manipulation




**Oliver Biehlmaier**
Imaging Core Facility
Biozentrum
Uni Basel

# Outline

- **Different techniques of image acquisition:**
  - **Imaging systems:** light, confocal, electron microscopes
  - **Detector types:** digital camera, photomultiplier (PMT)
- **Guidelines for imaging ethics**
- **Examples of "bad" imaging ethics**
- **Tools for CSI-cell biology**

# Imaging systems and detector types

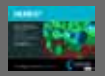
- The most currently used systems:
  - Wide field microscopy
  - Confocal microscopy
  - Electron microscopy
  - PET, MRI, etc.
  - Documentation systems (gels, etc.)
- Detector types:
  - Digital camera
  - Photomultiplier
- Software used for image processing and analysis

# Detector types

- **Digital camera:**
  - widefield microscopes
  - Spinning Disk Confocal microscopes
  - EMs, …
- **Photomultiplier:**
  - confocal & multiphoton microscopes
- **Other** types of **detectors**:
  - MRI-Scanner
  - PET-detectors
  - X-ray detector, …

# Software for image processing and analysis

- **Microscope acquisition software** of all different brands (Zeiss, Leica, Olympus, Nikon,…)
- Software for **3D/4D-rendering**, visualization, measurments:
  - **Imaris**
  - **ImageJ**, etc.



- **PHOTOSHOP** ! ! !

# Imaging Basics – bit depth

| Bit depth | Number of colors/shades | Approximate file size |
|---|---|---|
| $2^1$ | 2 | 0.125 MB |
| $2^4$ | 32 | 0.5 MB |
| $2^8$ | 256 | 1.0 MB |
| $2^{12}$ | 4096 | 1.5 MB |
| $2^{16}$ | 65,536 | 2.0 MB |
| $2^{24}$ | 16,777,216 | 3.0 MB |

*human sensitivity*

Bit-depth relates to the number of colors that can be displayed in the image.

- Images with only two colors are binary, the pixels are either black or white.
- Monitors and imaging hardware are typically limited to displaying grayscale images in 8-bit mode (256 shades of gray).
- Most monitors can only display color images in a maximum of 24-bit mode (true color), due to the limitations of the electronics in the cathode ray tube.
- Even these ranges are greater than the sensitivity of the human eye, which is often stated as only being able to detect 16-32 shades of gray and roughly 2,000,000 colors.
- Larger bit depth only if needed for intensity measurements/analysis

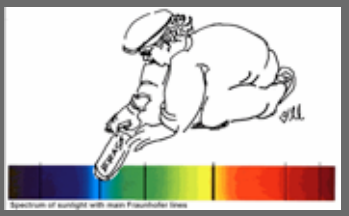
# Basics – common file formats

- **BMP** - windows bitmap
- **EPS** - encapsulated postscript, this format is more useful for vector-based information than pixel-based information
- **GIF** - graphics exchange format, originally copyrighted by CompuServe, used on web pages, has a 256 color palette limitation, not suitable for most scientific images
- **JPEG** - joint photographic experts group, supports 24-bit color, uses a lossy compression technique (discrete cosine function), most often used on web pages, not suitable for most scientific images. The proposed JPEG-2000 format will use the lossless wavelet-based compression technique.
- **PNG** - portable network graphics, supports 48-bit color and 16 grayscale, lossless compression, a relatively new format that is not widely supported yet
- **TIFF** - tagged image file format, originally developed by Aldus Corp. (which was purchased by Adobe Systems) & Microsoft Corp., supports paletted images (up to 8 bit), 8 & (in some programs) 16 bit grayscale as well as 24 bit color, this is probably the most commonly used file format for scientific images, supports lossless LZW compression (although not all programs can open compressed

# Guidelines for image acquisition and processing

1) Treat images as data
2) Save the original
3) Do only simple adjustments
4) Cropping is usually ok
5) Comparing images under identical conditions
6) Only manipulate entire image, NEVER just an area
7) Filters degrade data
8) Cloning degrades data
9) Intensity measurements only on uniformly processed image data
10) Lossy compression degrades data
11) Issues with magnification
12) Issues with pixels

# 1) Treating images as data



Scientific digital images are data that can be compromised by inappropriate manipulations.

- Digital images are composed of a grid of pixels. Each pixel has
  - a specific location relative to its neighbors
  - a scale *(size in microns)* relative to the acquisition system used
  - an intensity value based on the amount of light (energy) that captured.
- Intensity values can be expressed as numerical values (e.g. 8bit numbers). In the case of a three color image: numbers between 0-255 for each color (red, green and blue)
- Images are a grid, each position in the grid has numerical values
  => digital images share many characteristics with a spreadsheet!
  - image manipulations are simply mathematical functions that change the underlying numbers in the image.
- If you change the numbers, you change the image. When you manipulate the image in software, you are changing the underlying numerical values!

# 2) Saving the original

Manipulations of digital images should always be done on a copy of the raw image data. The original *(incl. metadata)* must be retained!

- Processed images can be compared to unprocessed images to
  - ensure that important data was not lost from, or inadvertent artifacts added to, the processed images
  - protect an investigator in the event of an accusation of scientific misconduct
- Keeping a copy of the raw image data file is also a user's greatest protection against an inadvertent mistake in image processing
- Keep raw image data if created in vendor's proprietary file format (incl. metadata)
- When converting image data from a proprietary format, users should save their image data as TIFF (tagged image file format) files.

# 3) Making simple adjustments

Simple adjustments to the entire image are usually acceptable.

- Digital images, like photographic negatives, sometimes need some help so that the viewer can appreciate all the information that is present in the image.
- "Reasonable" adjustments using software tools like brightness and contrast, levels, and gamma are usually appropriate manipulations for digital images.
- Beware, many of the automatic image adjustment tools that are available in commercial software. They are far too aggressive for scientific images and documentation of the changes done is impossible.

# Contrast/brightness adjustments

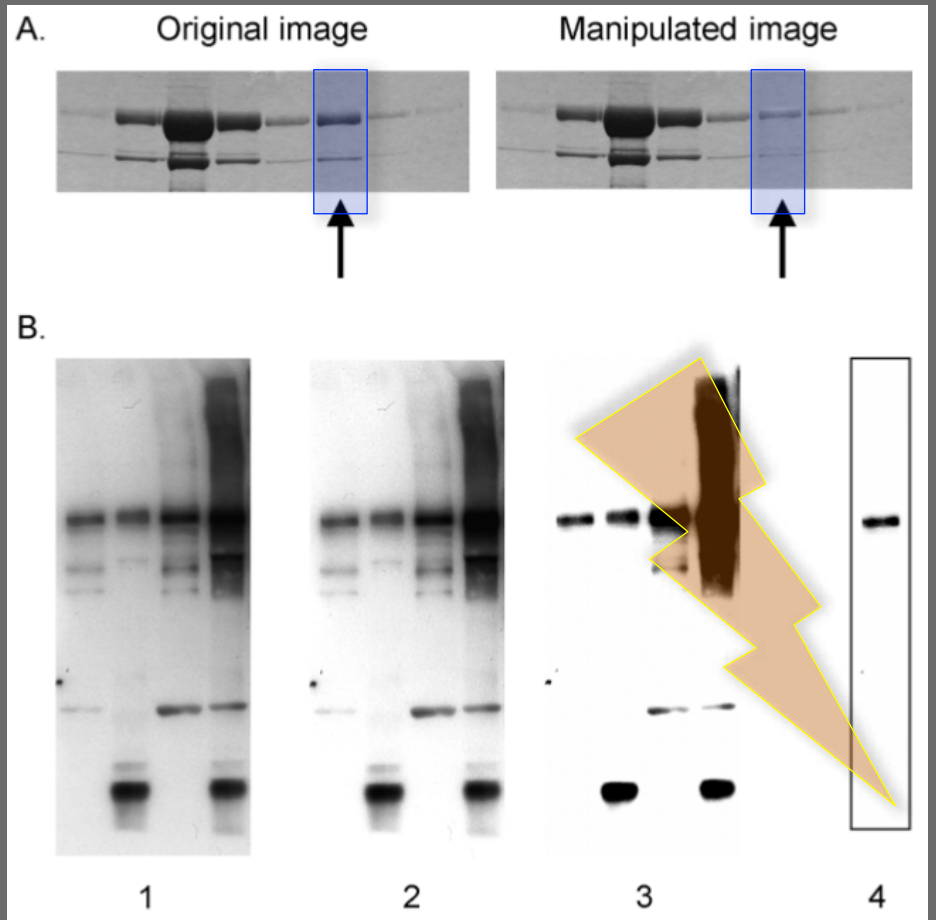
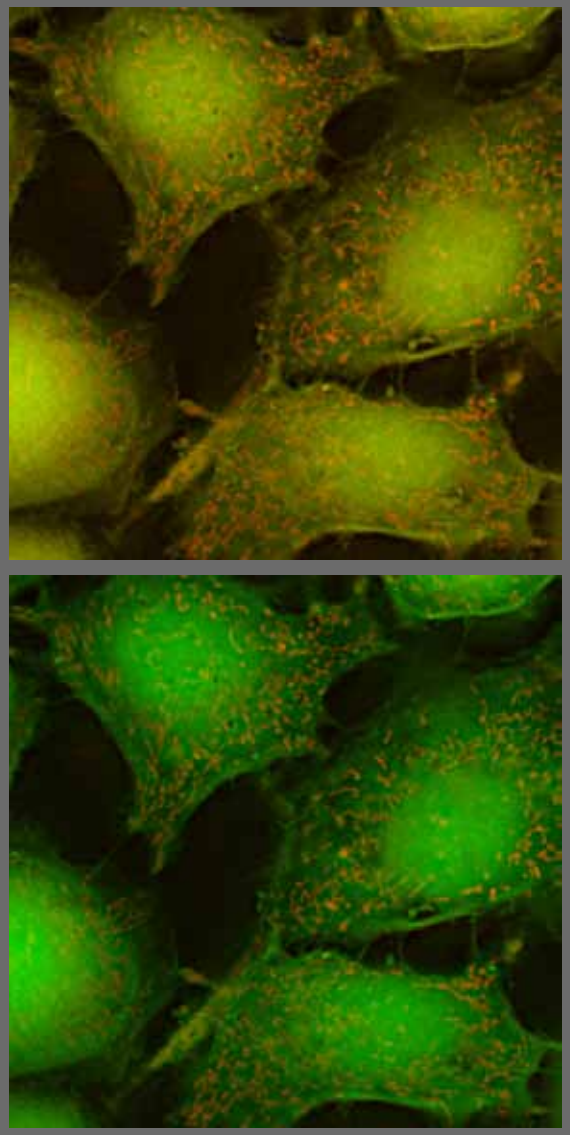



Figure 3. **Manipulation of blots: brightness and contrast adjustments.** (A) Adjusting the intensity of a single band (arrow). B) Adjustments of contrast. Images 1, 2, and 3 show sequentially more severe adjustments of contrast. Although the adjustment from 1 to 2 is acceptable because it does not obscure any of the bands, the adjustment from 2 to 3 is unacceptable because several bands are eliminated. Cutting out a strip of a blot with the contrast adjusted provides the false impression of a very clean result (image 4 was derived from a heavily adjusted version of the left lane of image 1). For a more detailed discussion of "gel slicing and dicing," see *Nature Cell Biology* editorial (2).

***Sometimes images require that the contrast or brightness be adjusted in order to make the background appear less hazy or make features more distinct or noticeable (more aesthetic image).***

- **Contrast and brightness adjustments should not be too drastic**
- **All of the relevant objects in the image should remain true to the observation.**
- **No objects should disappear entirely. Except for very minor ones**
- **Adjustments should be reported in the figure legends.**

# Color adjustments



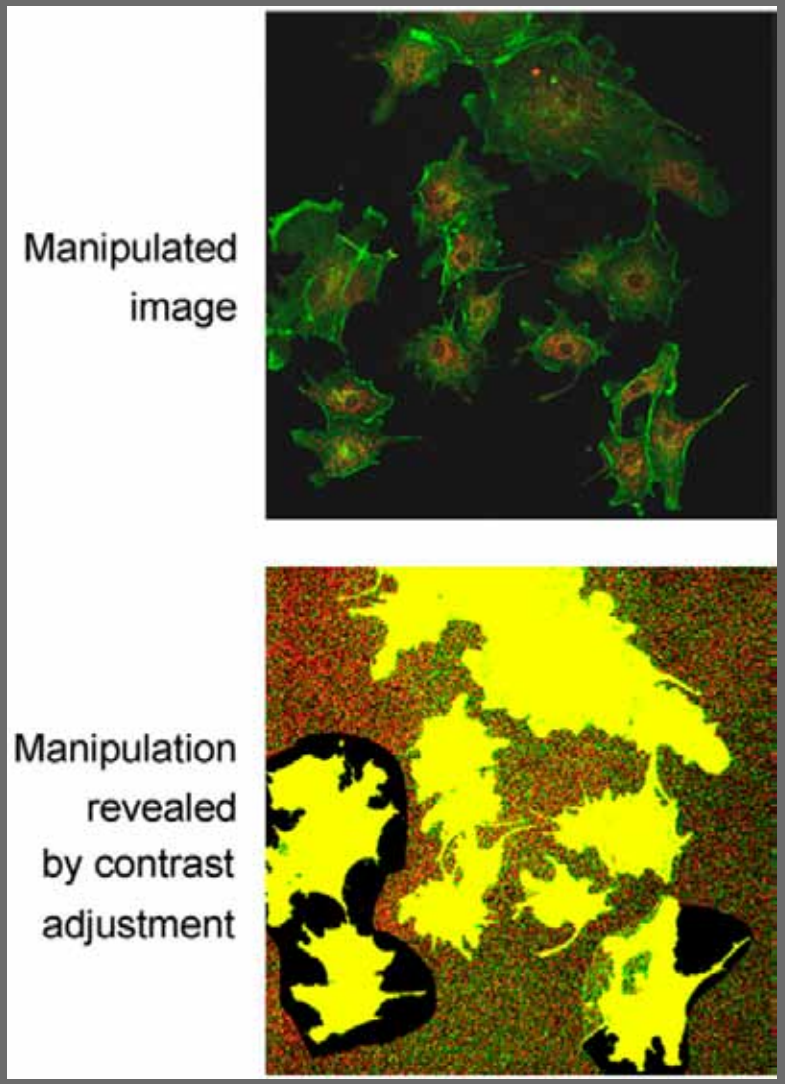
- Adjusting color settings is a not uncommon practice since doing so creates more vibrant and eye catching images.
- Color adjustments should be made only if the color of the image is not a part of the information carried by it and is not relevant to the conclusions drawn from the image

# 4) Cropping is usually ok

Cropping an image is usually acceptable.

- The Crop tool in most image processing programs is used to trim off the outside edges of a digital image. Cropping can be used to
  - make an image smaller (in pixels)
  - change the aspect ration (length to width) of the image
- Photographers have historically cropped images to direct the viewer's eye to a particular subset of a larger image.
  This can be very useful in art photography, but potentially misleading in scientific photography.
- Legitimate reasons for cropping:
  - Centering an area of interest
  - Trimming "empty" space around the edges of an image
  - Removing a piece of debris from the edge of the image
- Questionable forms cropping:
  - Cropping out dead or dying cells, leaving only a healthy looking cell
  - Cropping out gel bands that might disagree with the hypothesis being proposed in the paper

# Cropping




- Cropping can be a useful tool, for focusing attention on a particular object – e.g. one cell - within a larger image by removing irrelevant parts This is usually acceptable
- The line dividing good editing from bad practices is crossed when an image is cropped in a way that changes the context of what remains.
- Cropping is a common when researchers ensure that images remain representative of the observations

# 5) Comparing images under identical conditions

Digital images that will be compared to one another should be acquired under identical conditions, and any post-acquisition image processing should also be identical.

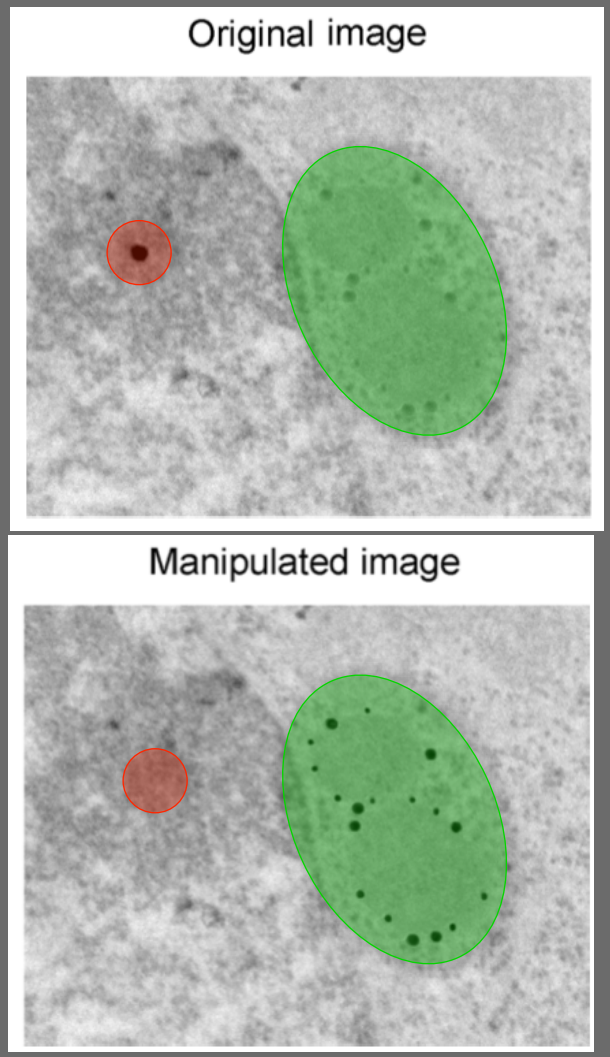
- Accurate comparison of treated and control digital images is difficult or impossible if the images are not acquired under similar conditions.
  - Several types of instruments, like confocal microscopes, allow users to adjust the signal amplification settings on every image.
    => Set the same settings for data that you wish to compare!
- When a group of images is to be compared to one another, the processing of the individual images should be identical
  - Background subtraction
  - White-level balancing
  - Gain of PMTs or EM-CCD-cameras

# 6) Only manipulate the entire image NEVER just an area!

Manipulations that are specific to one area of an image and are not performed on other areas are questionable.

- Most forms of selective enhancement are most likely inappropriate:
  - e.g. selective enhancement of specific bands or lanes within an image of a gel.
- If selective enhancement was performed and properly declared in submitted manuscript, then it would be up to the editors and reviewers to decide if it is appropriate on a case-by-case basis..

# Selective enhancements




*Enhancing certain features of an image and not the whole image, may sometimes seem desirable. While the intentions are sometimes innocuous, the end result is in* general **not good science**.

- Selective enhancements that should be avoided:
  - removing objects present in the original image
  - adding new objects that were not present

  The result of both is a **misstatement** to the reader about what was seen!

# 7) Filters degrade data

Use of software filters to improve image quality is usually not recommended for biological images.

- Commercial digital image manipulation software primarily written for design and print industries, not for scientists.
- Software filters typically change the numerical intensity value of every pixel in the image
  - Convolution kernels perform their mathematical functions using the numerical intensity values in small regions of the image (typically 3x3, 5x5, 7x7 pixel areas).
  - This means that the amount of change to the intensity of an individual pixel that is caused by the convolution kernel can be different in different parts of the image.
- If software filters must be used on scientific image data, the filters should be noted in an article's figure legends or methods section.

# 8) Cloning degrades data

Cloning or copying objects into a digital image,
from other parts of the same image or from a different image, is very questionable.

- Users are often tempted to use tools like Photoshop's clone stamp tool to "clean up" a dirty preparation. Don't do that!!!
- The use of cloning or copying techniques to specifically create objects in an image that did not exist there originally is research misconduct (falsification, fabrication)!!!
- In some instances the combination of two images into a single publication figure is appropriate.
  - The figure needs to make it obvious to the editor, reviewers, and journal readers that the two parts came from separate images.
  - Example: Aligning lanes from two or more gels into one figure.
    (*Most journals require a line or small gap between the combined images to clearly show that they are from separate gels)*

# Cloning objects into an image

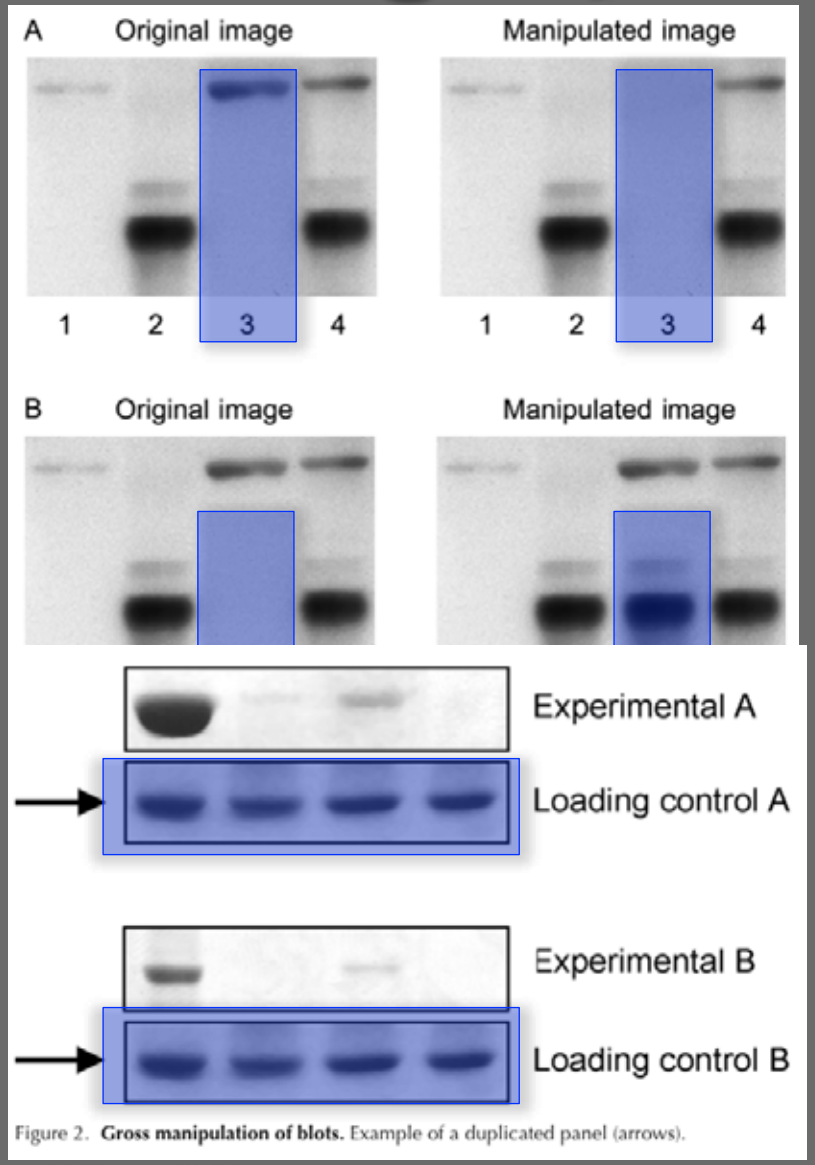
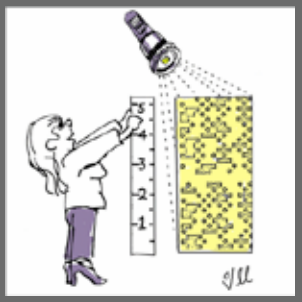



Figure 2. **Gross manipulation of blots.** Example of a duplicated panel (arrows).

- The "clone stamp" (PS) can copy and transfer a sample of pixels from one area of an image to another
- Cloning an object into an image in which it did not previously exist is always highly questionable.

*Whether to add features that were not present in the first place or to clean up the background, is strongly discouraged and highly questionable.*
*If for some reason the tool is used, the manipulation made should be reported in detail in the figure legend.*

# 9) Intensity measurements only on uniformly processed image data



Intensity measurements should be performed on uniformly processed image data, and the data should be calibrated to a known standard.

- Many variables must be considered and controlled for, before the data can be regarded as meaningful
  - uniformity of sample preparation techniques
  - uniformity of optical aberrations and electronic noise, etc.
- Intensity measurements include the analysis of colocalization data (comparison of intensity data from two different wavelengths)
- In general, intensity measurements should be performed on raw data to avoid potential artifacts.
- If normalization, calibration, or any other image processing is performed on the data, it should be performed uniformly across all the data!
  *The procedures should be carefully described in the methods section of the publication.*

# 10) Lossy compression degrades data

Avoid the use of lossy compression.

Image file compression comes in two basic types

1) Loss-less file compression reduces the size of the image file while maintaining the integrity of the image data (e.g. TIF LZW compression)

2) Lossy file compression is provided by the Jpeg file format. Jpeg uses the discrete cosine transform which changes the spatial (XY) resolution of the image and the intensity value of any given pixel.

   - Using any form of lossy compression for images may create problems. The information thrown away during lossy compression is generally that information that is imperceptible to a human eye - not necessarily showing the same characteristics as computer image processing software".

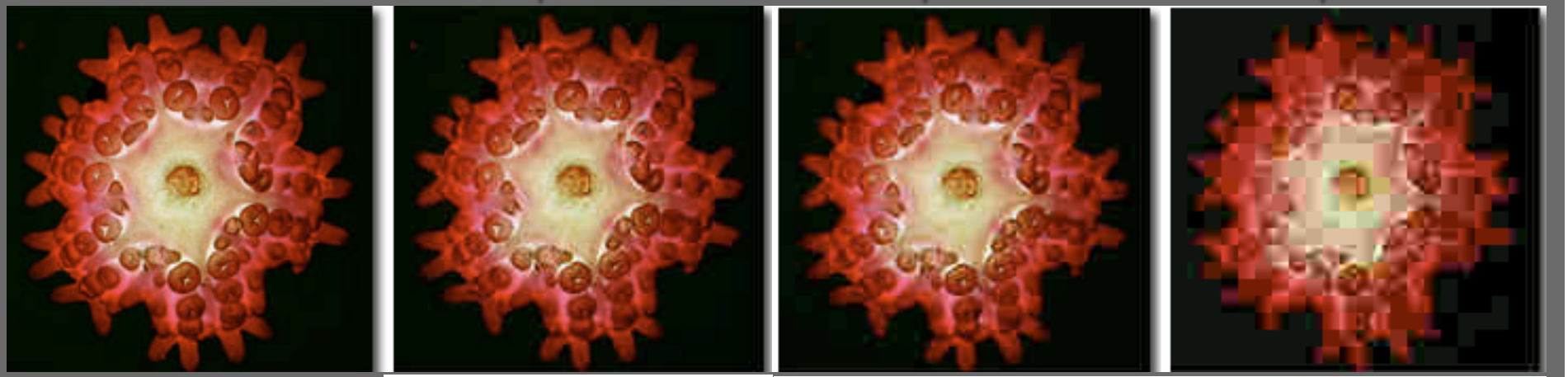
# What jpeg does to your images...

original

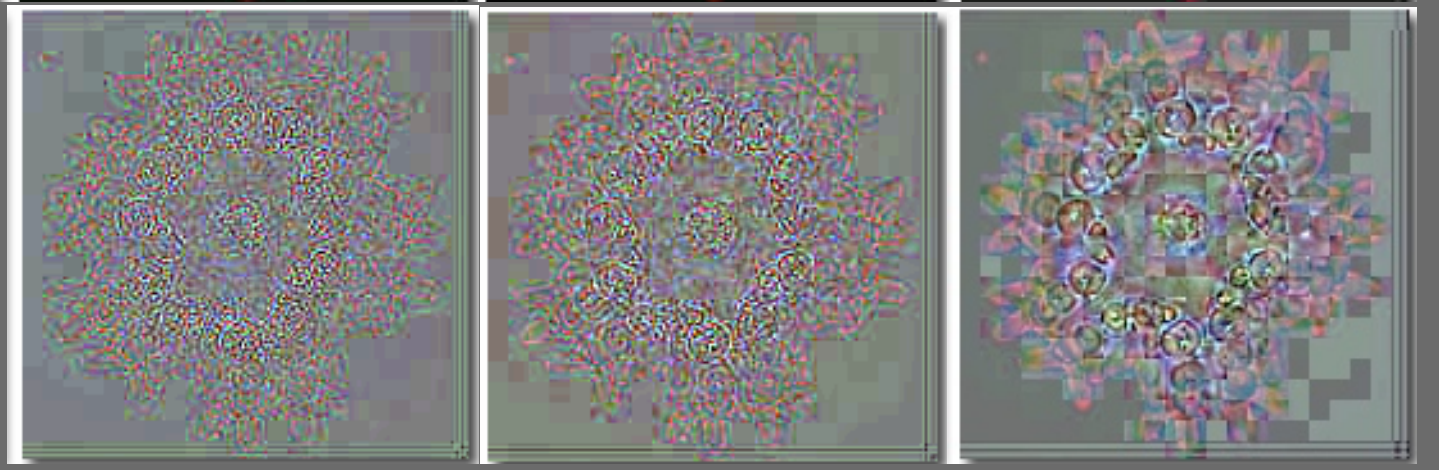
low compression

medium compression

strong compression



difference:
What has been deleted from your image...



generated on: http://micro.magnet.fsu.edu/primer/java/digitalimaging/processing/jpegcompression/

# 11) Issues with magnification

Scientific digital images are data that can be compromised by inappropriate manipulations.

- Each pixel in the image has a scale ( tenths of microns per pixel)
- Ideally, the scale is the same in both the X and Y dimensions; however, this is not always the case.
- In confocal microscopy and other sectioning techniques, the XY pixel also represents a volume, because the image includes a Z dimension
- Important issue with sampling small objects using digital image capture is the need to correctly oversample the object.
  - The Nyquist sampling theorem suggests that a point object should be oversampled at least two times.
    - Undersampling an image can yield artifacts that masquerade as real structures, which can lead to misinterpretation of the image data
    - Oversampling does not lead to artifacts; however, it does not increase the diffraction-limited spatial resolution data to the image.

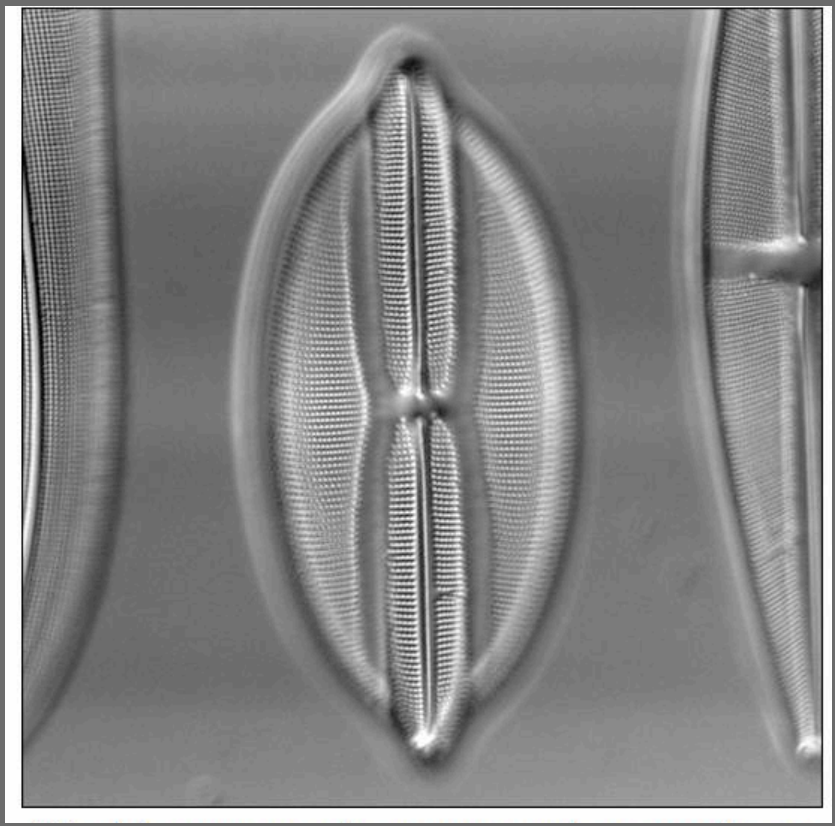
# 12) Issues with pixels

Be careful when changing the size (in pixels) of a digital image...

Changing the size of an image (the number of pixels in X and Y) can introduce aliasing

- Decreasing the image size in pixels reduces the spatial (XY) resolution in an image.
  - If the size reduction is not by a power of two, the software program performs an interpolation to create both a new XY resolution as well as the intensity values of each pixel.
  - If the image has regular, repeating structures, size reductions can create moiré artifacts.
- Increasing the image size causes the software program to interpolate the new XY resolution and pixel-intensity values
  - No increase in spatial resolution, but: specific features become more difficult to resolve because aliasing artifacts tend to make the edges of features less distinct.
- If you change the size: Insert a magnification scale bar prior to the changes
- Only resize once

# Resizing



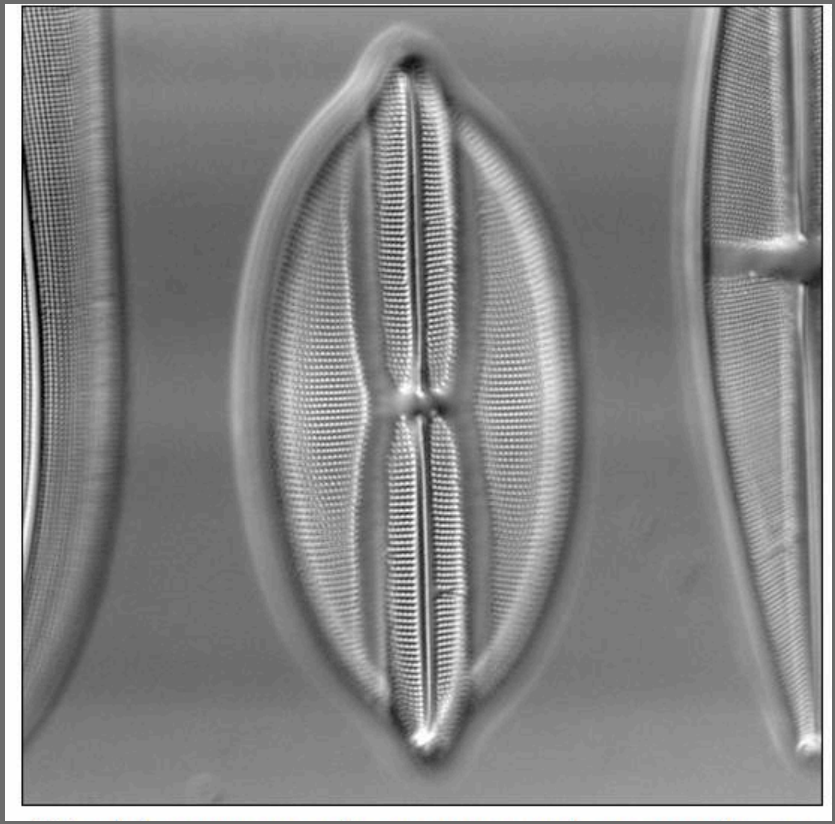
original picture

Two different factors can cause image data to degrade during resizing:

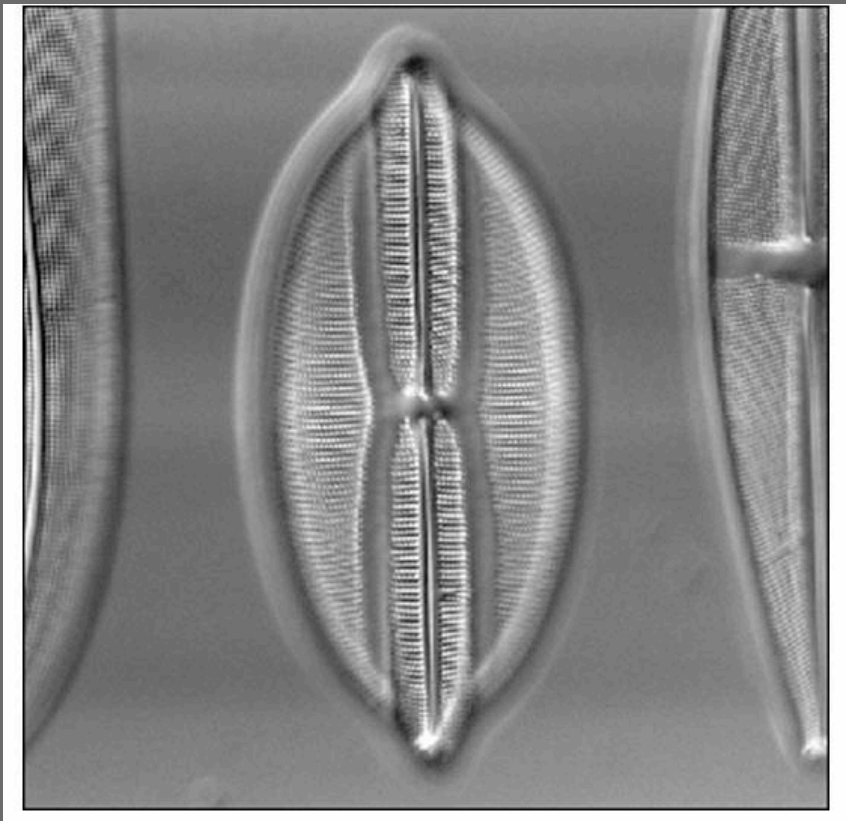
1. The aspect ratio, which if inadvertently changed, changes the shapes of the objects in the image.
2. Under-sampling, which degrades an image and can result in pixilation, called "anti-aliasing," which distorts the details and borders of objects.

*Care should be taken not to change the aspect ratio or to under-sample. As with nearly every other alteration, it is important that whatever is done, it is explained in the figure legend of the image.*

# Resizing

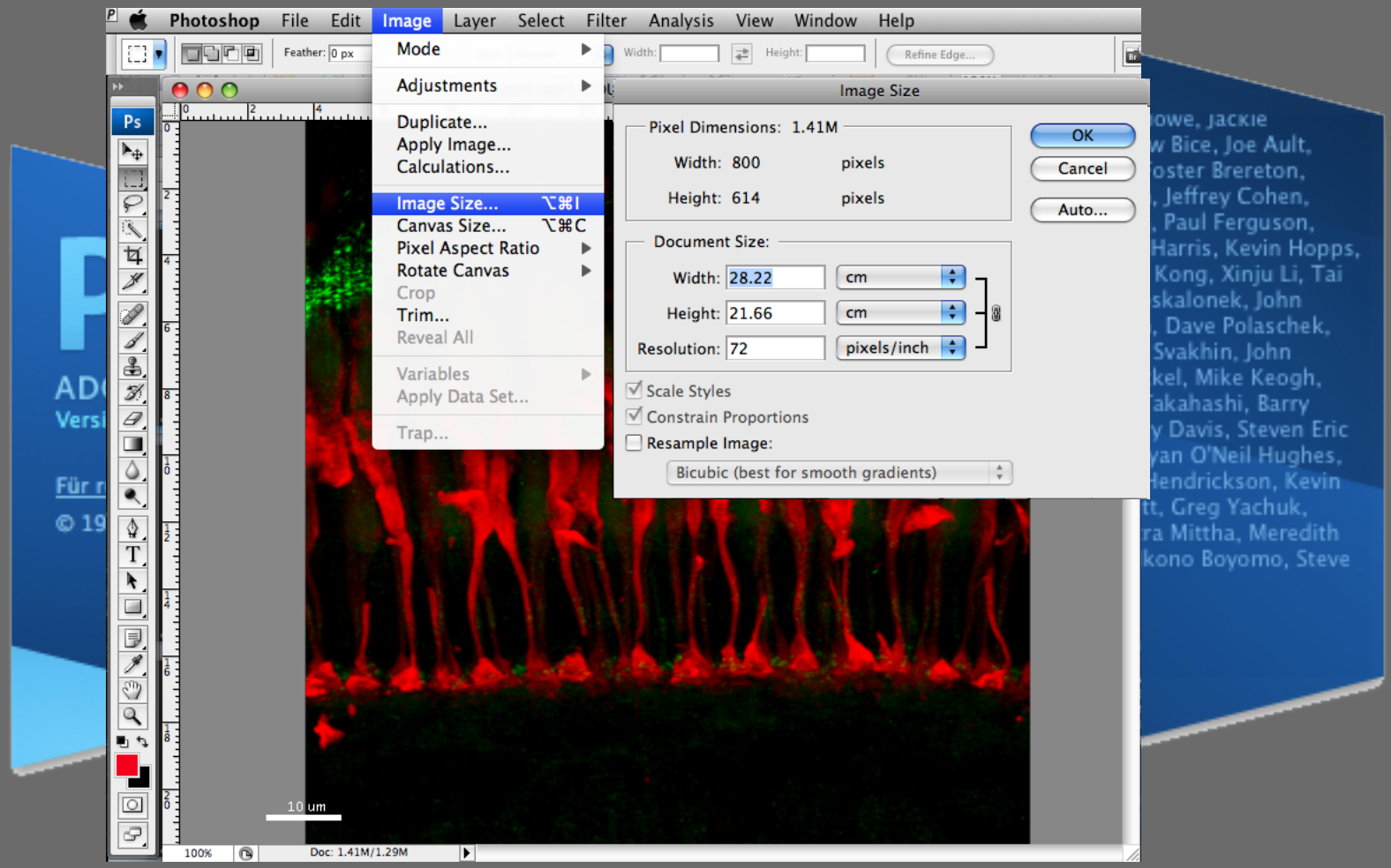


original picture



resized picture

# Caution with ADOBE Photoshop

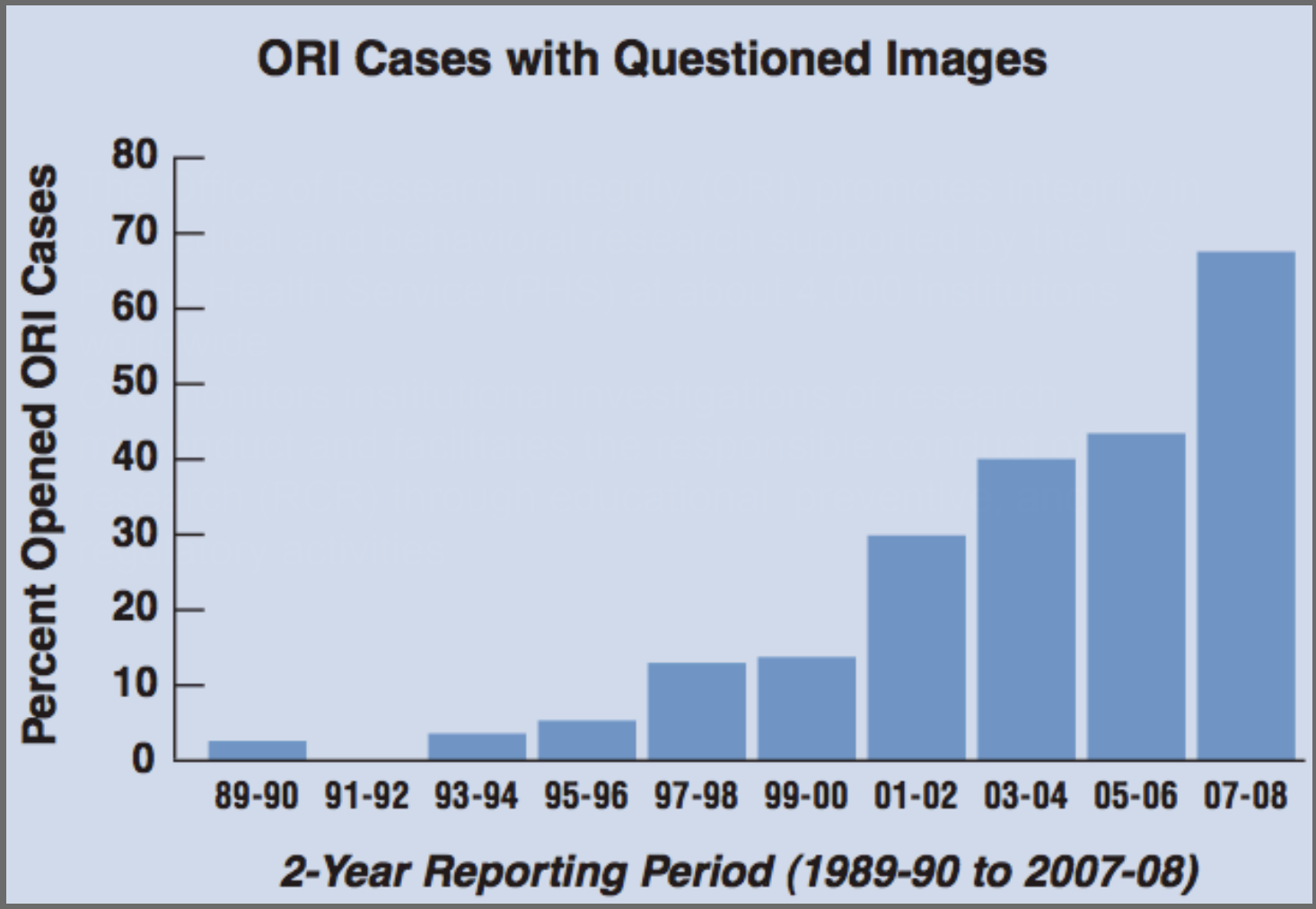


# Office of Research Integrity (ORI)

**Table 1. ORI findings in cases, by misconduct groups (June 1992 through November 2005)**

| ORI Findings | Total number of cases |
|---|---|
| Plagiarism alone | 8 |
| Plagiarism and Falsification | 6 |
| Plagiarism, Falsification, and Fabrication | 5 |
| Plagiarism and Fabrication | 0 |
| Falsification alone | 59 |
| Fabrication alone | 30 |
| Falsification and Fabrication | 53 |
| Other | 1 |
| **Total number of ORI cases with findings** | 162 |

# Office of Research Integrity (ORI)



# Case 1

## Researchers reprimanded for cleaning up figures

- A prominent German neuroscientist and a Swiss colleague were reprimanded by the German funding agency DFG for altering two figures in a paper published in 1998 in the Journal of Neurochemistry, according to Science (302:763).
  - Heinz Breer, University of Hohenheim, and Johannes Noe, University of Zurich, found guilty of research misconduct by a DFG committee that investigated an allegation made by a former postdoc in Breer's lab.
  - The committee concluded that the researchers committed misconduct because they failed to disclose that they **had cleaned up the primer bands** in the Southern Blots to make them look more dramatic. The alterations did not affect the paper's conclusions and the researchers remain eligible for DFG funding.

# Case 2

## German University Investigating Allegation against Cancer Researchers

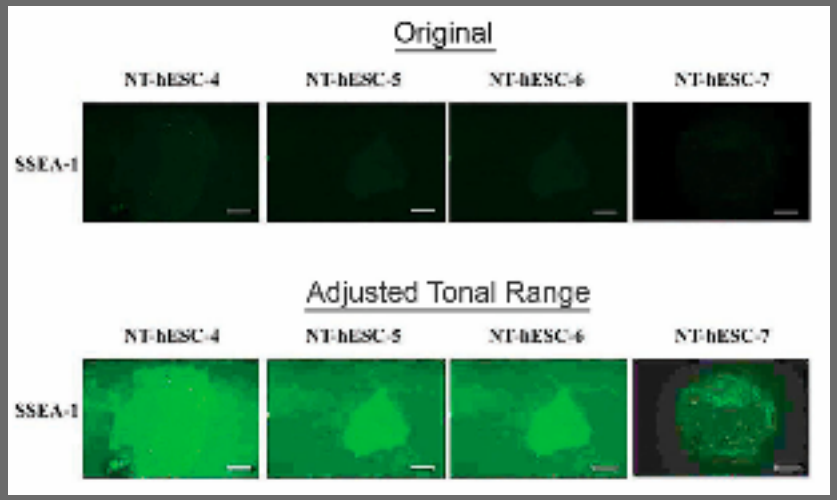
- The University of Gottingen in Germany is investigating an allegation of research misconduct against a researcher at that institution and another at the University of Tubingen who, according to press reports, are attempting to produce a cancer "vaccine."
- The clinical research involves treating kidney cancer patients with a vaccine composed of tumor cells fused with immune cells. Regression of secondary tumors was reported in an article last year.
- The allegations assert that a photo in another paper submitted for publication purportedly showing the fused tumor cells had been **downloaded from the Internet**. In addition, the investigation is looking into charges of patient endangerment. Recruitment into the study has been suspended.

# Case 3

Hwang Woo-Suk - was a professor of theriogenology and biotechnology at Seoul National University (dismissed on March 20, 2006)

- Rose to fame after claiming a series of remarkable breakthroughs in the field of stem cell research. Until November 2005, he was considered one of the pioneering experts in the field of stem cell research.
- Best known for two articles published in the journal Science in 2004 and 2005 where he fraudulently reported to have succeeded in creating human embryonic stem cells by cloning.
- Both papers were later editorially retracted after they were found to contain a large amount of fabricated data.
- Admitted to various lies and frauds.

# Hwang case example




- The image in the top row is from the third row of Supplemental Figure S1B in that paper.
- It purports to show negative staining for a particular cell surface marker in four different cell lines.
- A simple adjustment of tonal range in Photoshop clearly shows that **the two middle images are identical**.
- Minor differences in pixel structure are due to image compression.
- Detecting this duplication should have led to request the original data from the authors.

# Tools for CSI-cell biology:

## Plagiarism Detection Resources

- The Plagiarism Resource Center at the University of Virginia - Site provides free software to detect plagiarism.
- eTBLAST - a text similarity engine, which accepts a query and then compares it to a collection of other text, especially Medline. A variety of post-processors analyze the "hits" to provide added value to the user. Applications include reference searches, novelty assessment and publication ethics.
- Deja vu - A database of highly similar citations identified by eTBLAST. The database includes over 70,000 pairs of citations and notes from manual inspection of some full text article pairs.

# Tools for CSI-cell biology:

## Forensic "actions" for Photoshop

1. **Dark or Light Areas Discontinuities:** Histogram equalization, which may reveal areas of erasure in dark areas, or areas of whitening in bright areas. The process is most effective if any surrounding contrasting border is first cropped from the region of interest.

2. **Visualize Background:** This Action set enhances background by embossing (shadowing) details and permits false coloring of the results.

3. **Colorize Shapes & Details:** Details in the unique fine structure surrounding features in question may serve to test whether two features are the same. False coloring and sliding the Hue setting facilitates detection of small differences in gray shades that would be otherwise imperceptible.

4. **Solarize Margins & Edges:** This Action set uses a non-linear contrast adjust to amplify differences in the midrange and the dark end of the histogram. The option exists to also false colorize the results.

5. **Color Overlay Two Images:** This Action overlays two images, and color codes each so that individual features in the overlap can be linked back to their respective image.

# Acknowledgements & Links


- Cellular Imaging Core @ The University of Arizona
  http://swehsc.pharmacy.arizona.edu/exppath/index.php
- Mike Rossner,
  JCB editor, Executive Director, The Rockefeller University Press, NYC, USA
- George McNamara,
  Analytical Imaging Core, University of Miami, USA

# Links @ http://www.biozentrum.unibas.ch/imcf

Imaging Core Facility (IMCF)

## Imaging Ethics - CSI Cell Biology

**Guidelines & Information**

- Digital imaging ethics guidelines
  by Lantz CR, Douglas W, Cromey MS - Cellular Imaging Facility Core, University of Arizona
- Online learning tool for research integrity & image processing
  This site explains what is appropriate in image processing in science and what is not. It also shows how best practices in handling images intersects with other best practices.
- Image manipulation: CSI - cell biology
  Pearson H
  One of the initial articles bringing up the issue of image manipulation and scientific miscunduct.
- Imaging Forensics Interesting links and tutorial concerning imaging forensics on the homepage of the computer science group of Hany Farid at Dartmouth College
- Singapore statement on research integrity

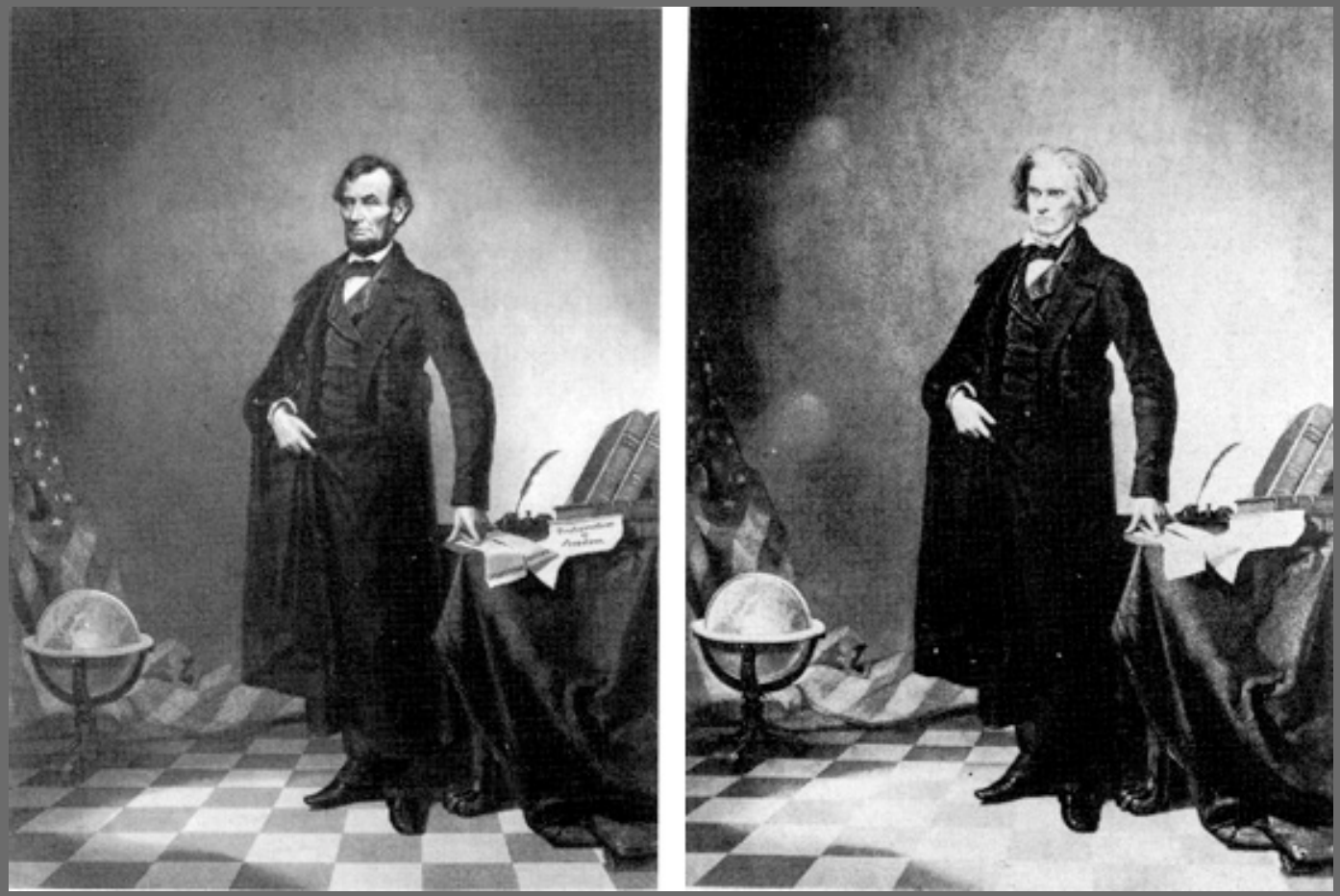
**Research Integrity institutions**

- Office of Research Integrity (ORI)
- European Forum For Good Clinical Practice

**Interactive Movie on Research Misconduct**

- The LAB - Avoiding Research Misconduct
  In "The Lab: Avoiding Research Misconduct", you become the lead characters in an interactive movie and make decisions about integrity in research that can have long-term consequences. The simulation addresses Responsible Conduct of Research topics such as avoiding research misconduct, mentorship responsibilities, handling of data, responsible authorship, and questionable research practices.

# manipulated photography

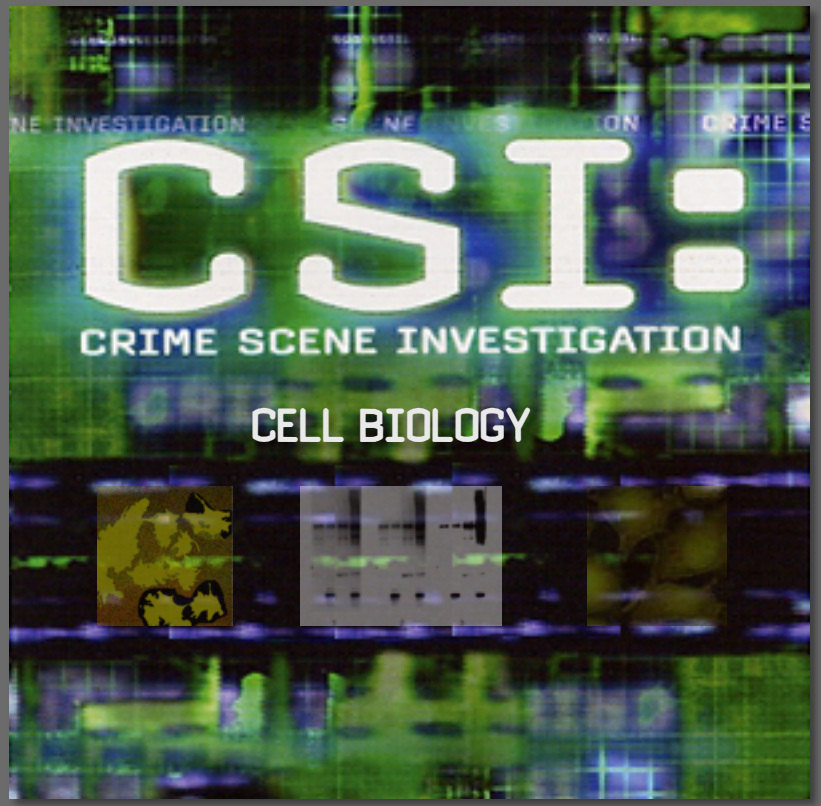


# manipulated photography

Беседа Л. И. Брежнева с В. Брандтом

ЯЛТА, 17. (ТАСС). Сегодня в районе Ореанды состоялась продолжительная беседа между Генеральным секретарем ЦК КПСС Л. И. Брежневым и Федеральным канцлером ФРГ В. Брандтом. В ходе беседы были затронуты актуальные вопросы развития отношений между Советским Союзом и Федеративной Республикой Германии и международные проблемы, представляющие взаимный интерес, в особенности вопросы укрепления европейской безопасности. Беседа будет продолжена во второй половине дня.

# Thank you!




**Oliver Biehlmaier**
Imaging Core Facility
Biozentrum
Uni Basel
oliver.biehlmaier@unibas.ch
www.biozentrum.unibas.ch/imcf