

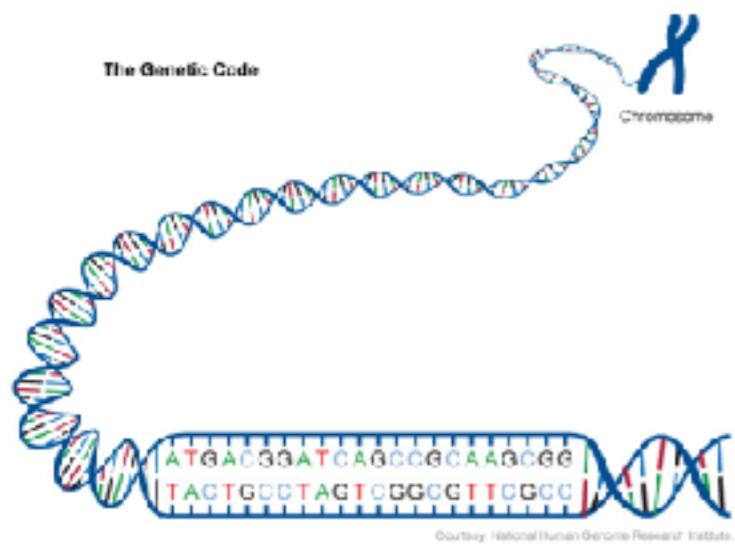
# Introduction: Bioimage Informatics

Thomas Walter

# Overview

- Images in biological research.
- Short review of disciplines related to computational processing/ analysis of images.
- Application examples (A1-A8).

# Images in a Bioinformatics world

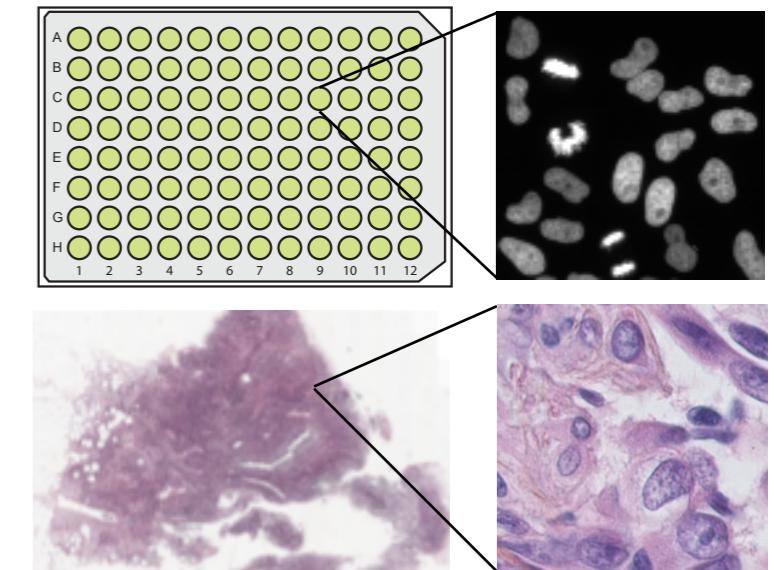
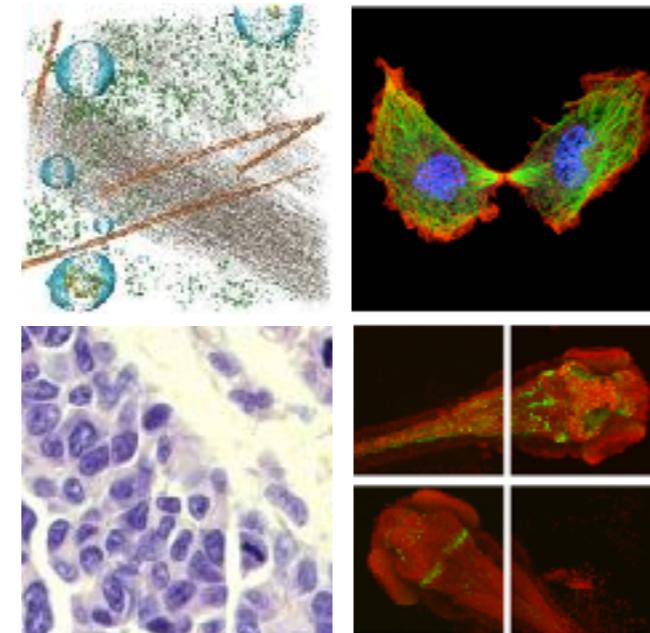


## Big data in bioinformatics

- Genomes
- Transcriptomes
- Proteomes

## Phenotyping by imaging

- Multiple scales
- Morphology
- Spatial organization



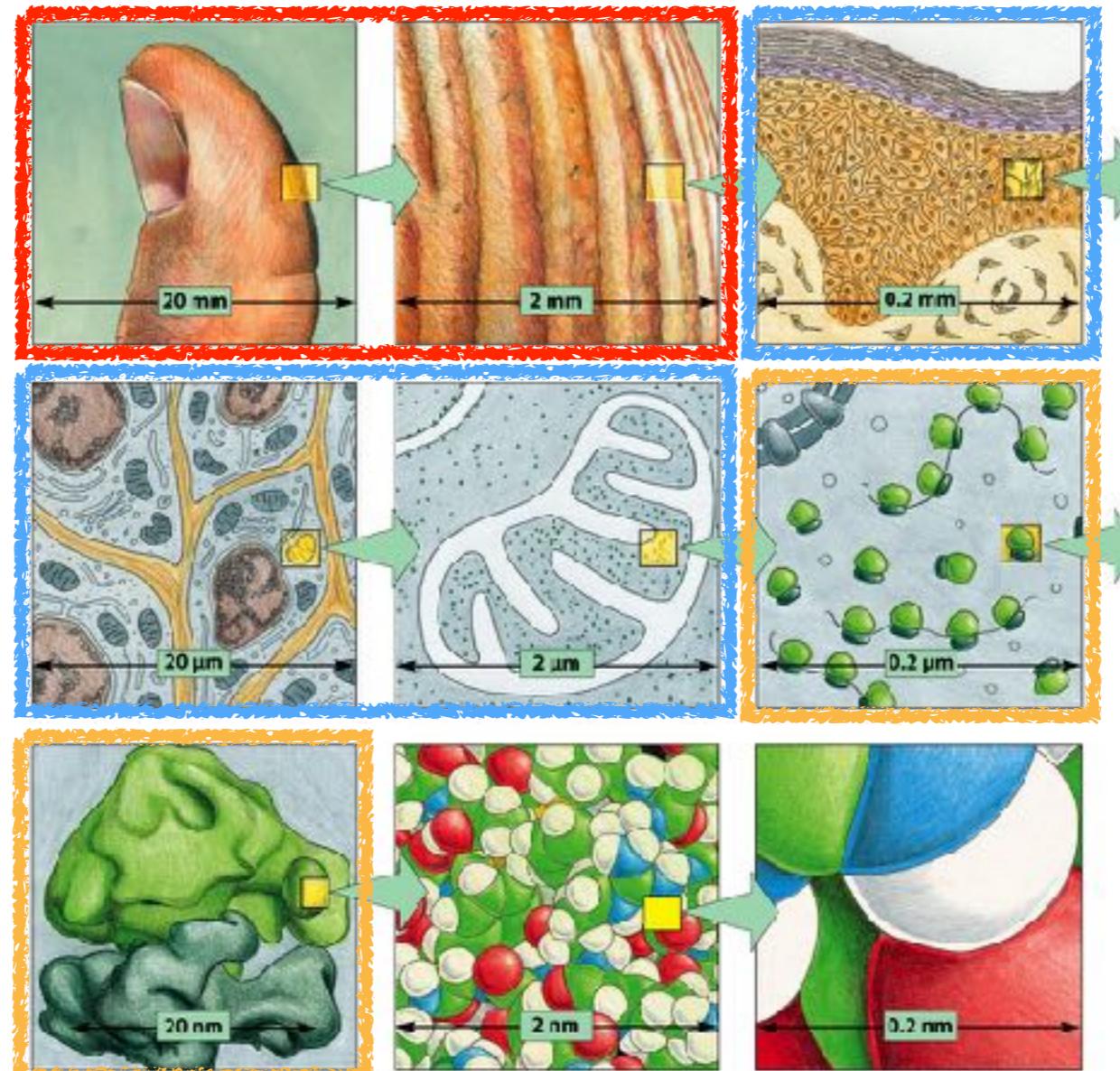
## Big data in bioimaging

- Screening
- Histopathology
- ...

# Images and omics measurements

omics measurements	microscopy
Usually calculated on large cell populations	Usually individual cells measurements
resolution: molecular level (single base pairs)	Subcellular to organism level. Molecular level by specific labelling.
comprehensive measurements (whole genome, whole transcriptome ...)	limited to the visible structures (often highlighted by specific markers)
very sensitive in a sense that small changes (e.g. mutations) can be identified	a change has to have an effect on the cell to be measurable; functional impact
changes in the molecules (structure, abundance, ... )	spatial and temporal organization, morphology

# Seeing is believing



- visible by naked eye
- visible by light microscopy
- visible by electron microscopy

Alberts et al., The molecular biology of the cell

- If we want to study biological systems and their composition, we need to magnify the samples.
- Here, we see a 10x progression between images.
- The birth of cell biology (cell doctrine, 1838) came only with the availability of good light microscopes.
- Today, technical advances in imaging technologies continue having an extremely important impact on scientific discoveries in biology.

# Resolving power as a function of wavelength

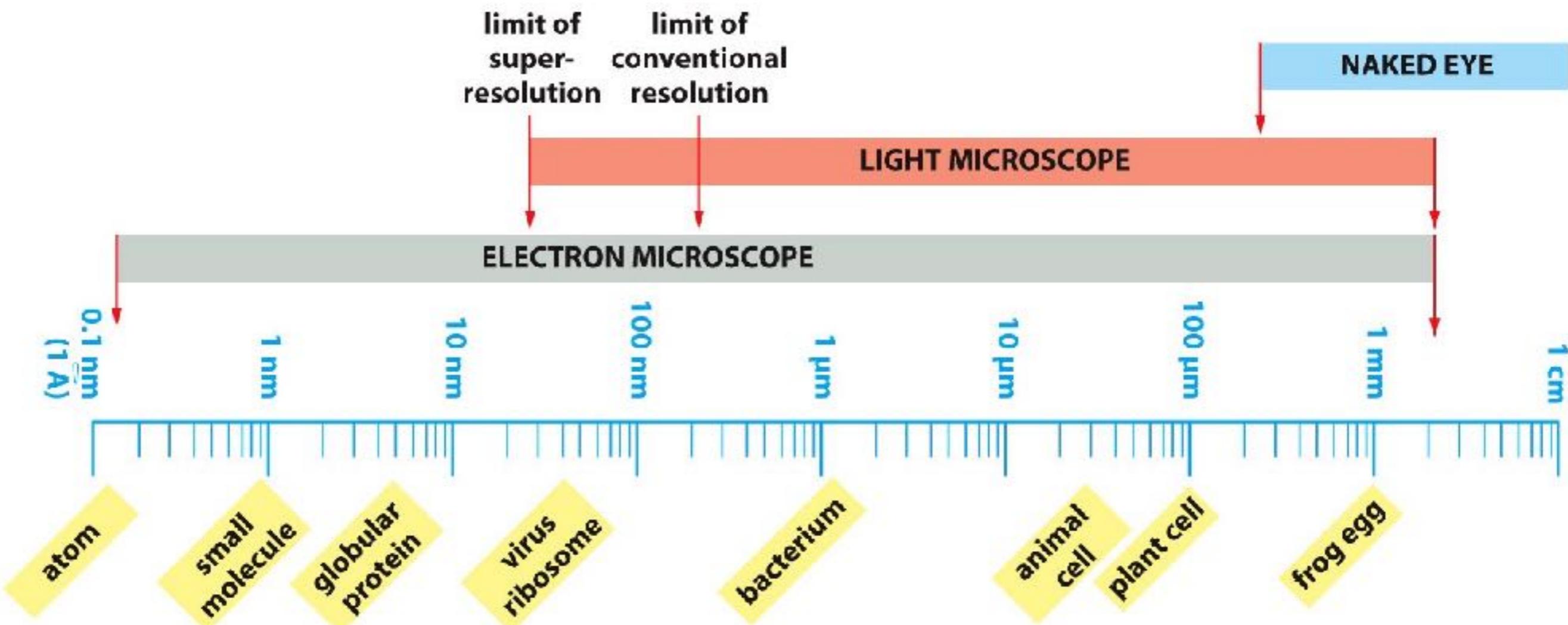
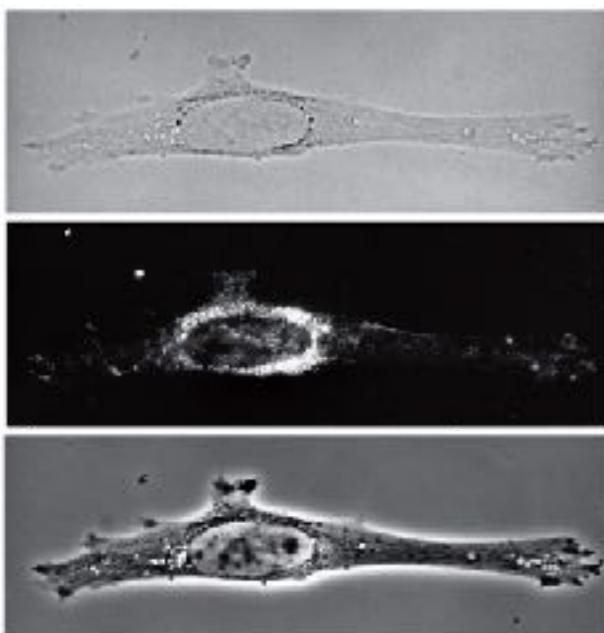
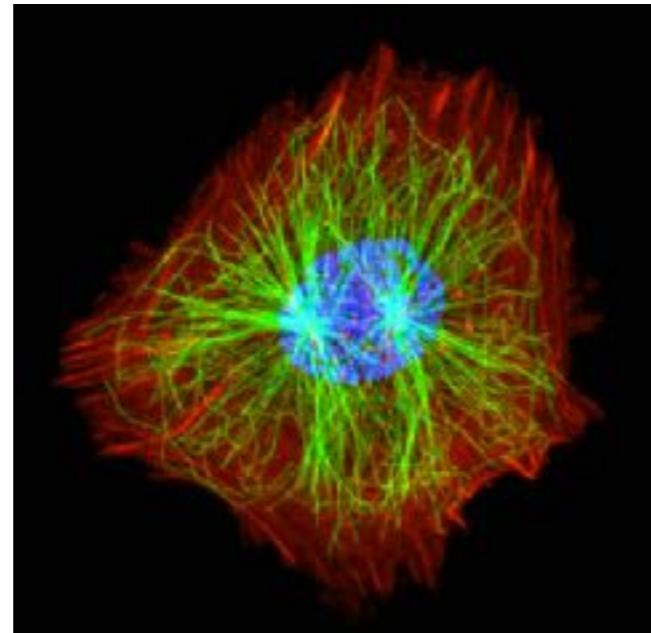


Figure 9-2 Molecular Biology of the Cell 6e (© Garland Science 2015)

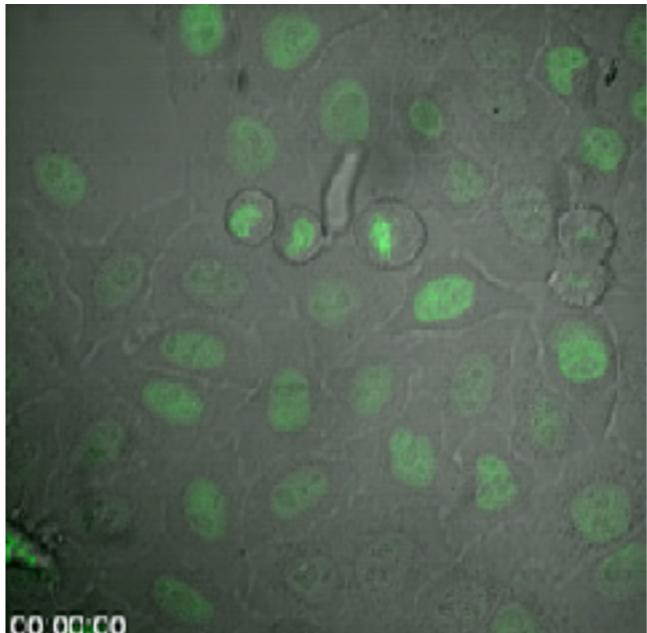
# Diversity of microscopy images



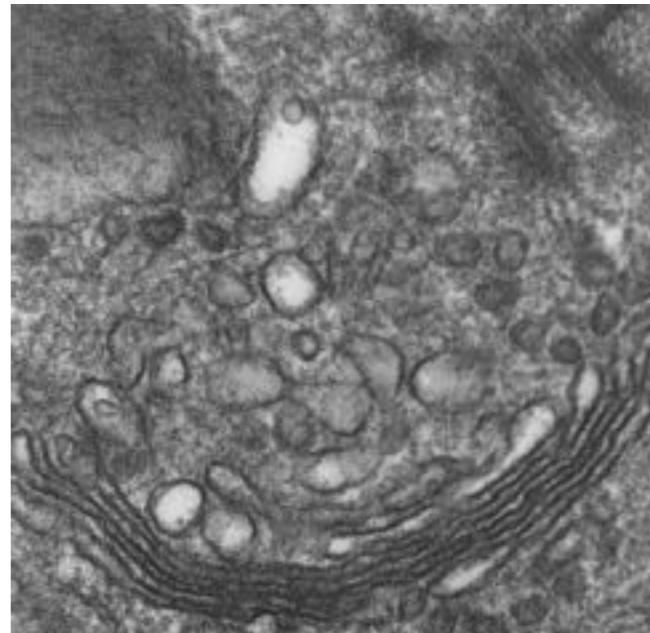
*Label-free*



*Labeling cellular structures*



*Live cell imaging*



*Electron microscopy (TEM)*

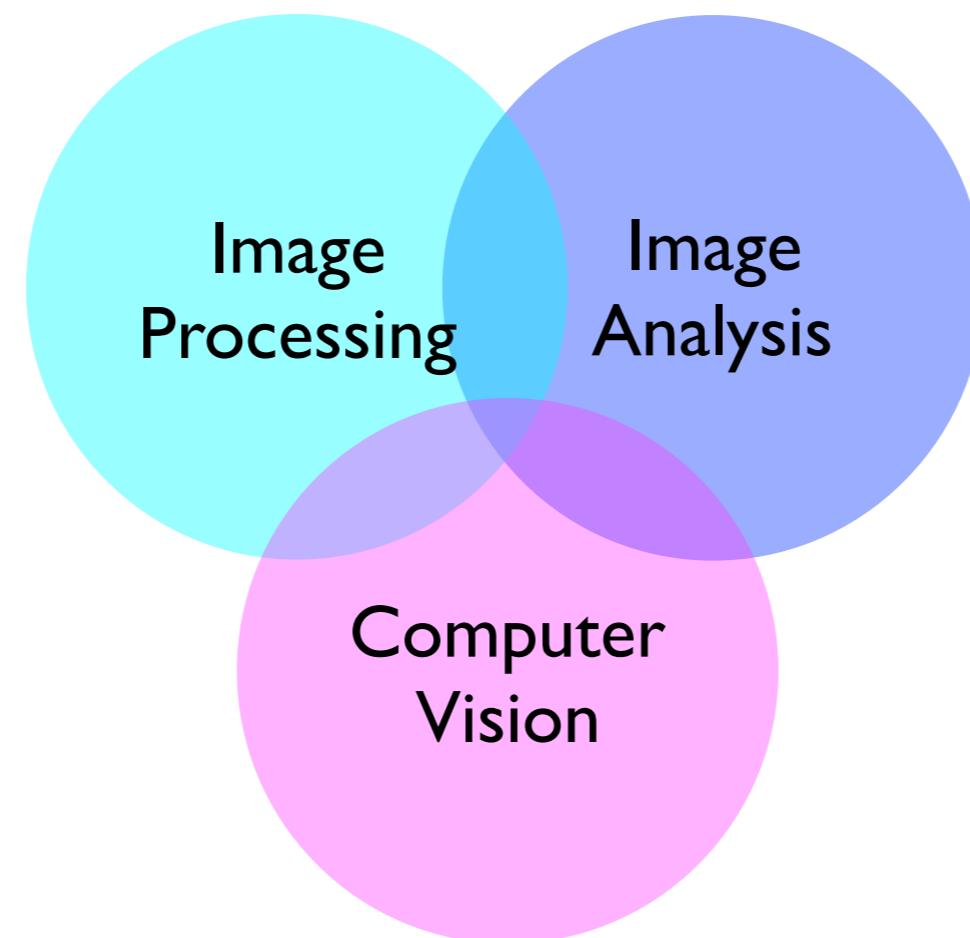
- Light microscopy vs. Electron microscopy
- Fluorescence microscopy vs. label-free approaches
- Live cell imaging vs. fixed
- Cells vs. tissues vs. organisms
- 2D vs. 3D

# In silico disciplines related to images

## **Image Processing**

set of techniques for manipulating and modifying images.

Input and output are images.



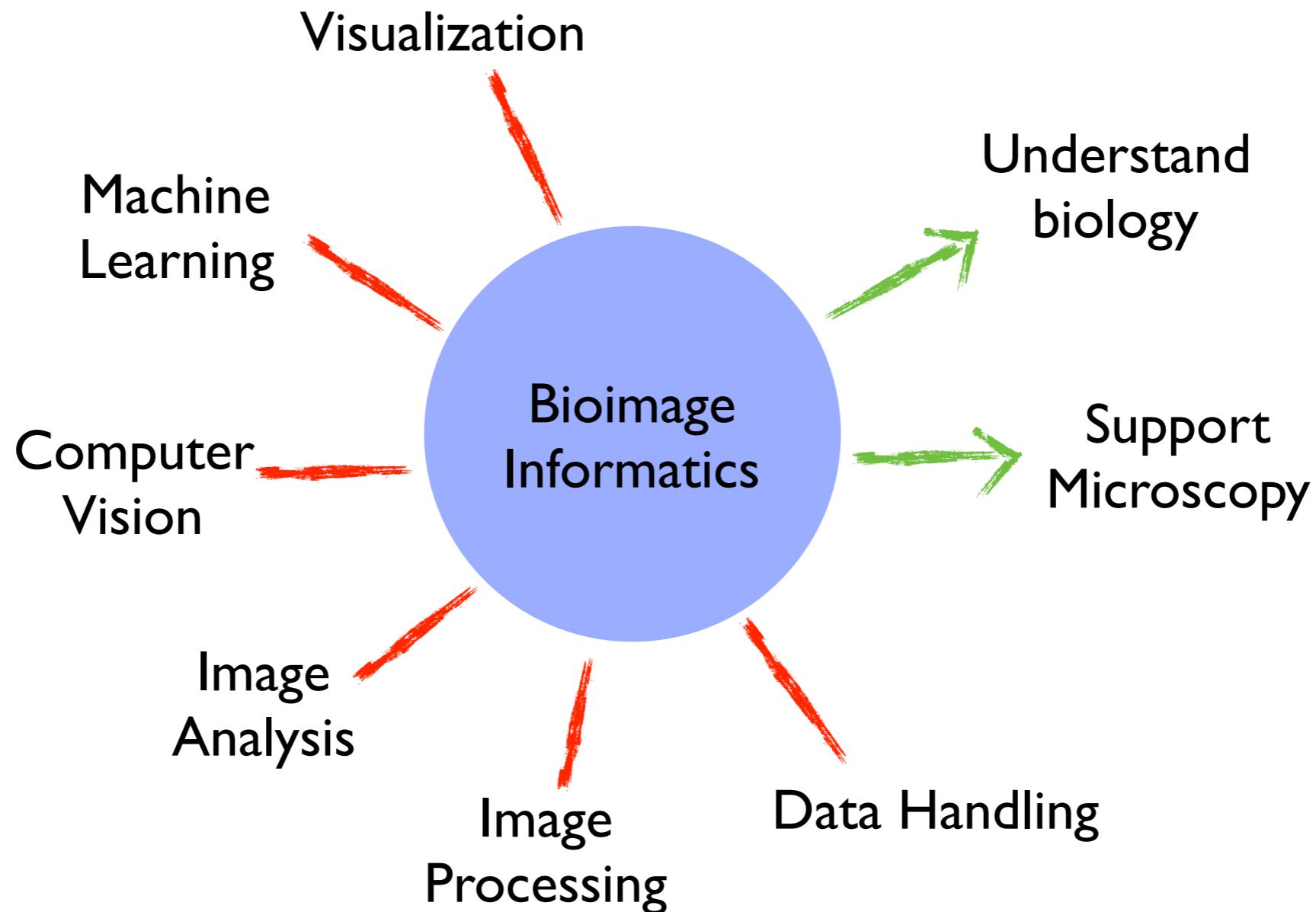
## **Image Analysis**

set of techniques to analyze the content of an image. The input is an image, the output is a set of attributes extracted from the images.

## **Computer Vision**

ultimately aims at emulating human vision, which requires understanding of the content of images. The input is an image and the output an arbitrary data structure that conveys an interpretation of the .

# Bioimage Informatics



## **Bioimage Informatics**

is a discipline at the interface between biology, microscopy, bioinformatics and the computational disciplines of image processing, image analysis, computer vision and machine learning.

# Why do we need Bioimage Informatics?

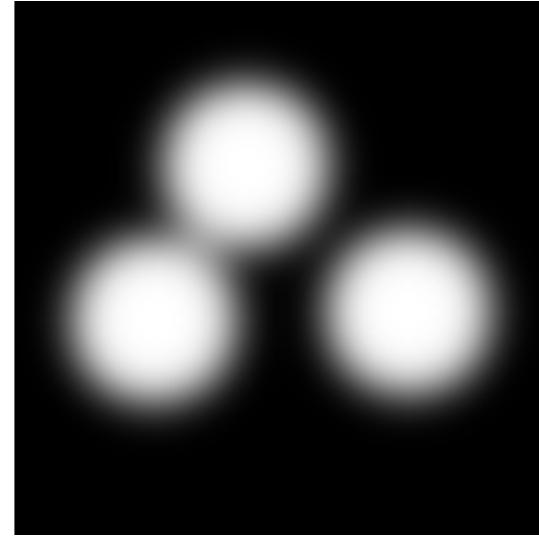
- Image generation (tomographies, super-resolution microscopy)
- Image enhancement (to make hidden structures visible)
- Image quantification:
  - Objectivity
  - Completeness (i.e. quantification of the entire experiment instead of picking single events).
  - Analyzing information that is not appropriate for visual inspection (e.g. particle speed)
- Amount of image data makes it unfeasible to manually analyze all of the images.

# AI: image restoration and inverse problems

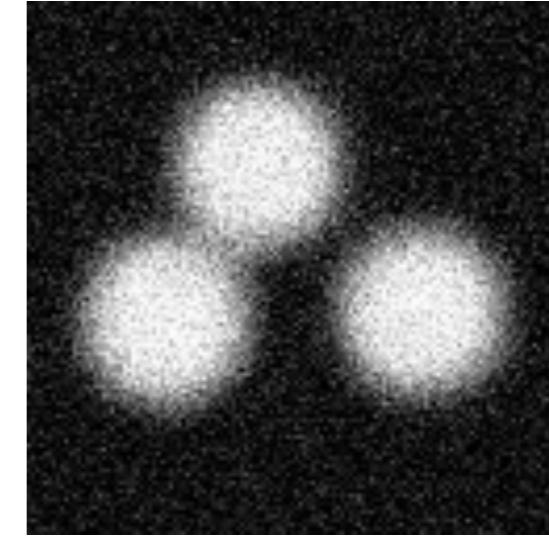
- Microscopy does not produce a perfect reproduction of the reality.



Real objects  $u$



Blurring by the  
optical system



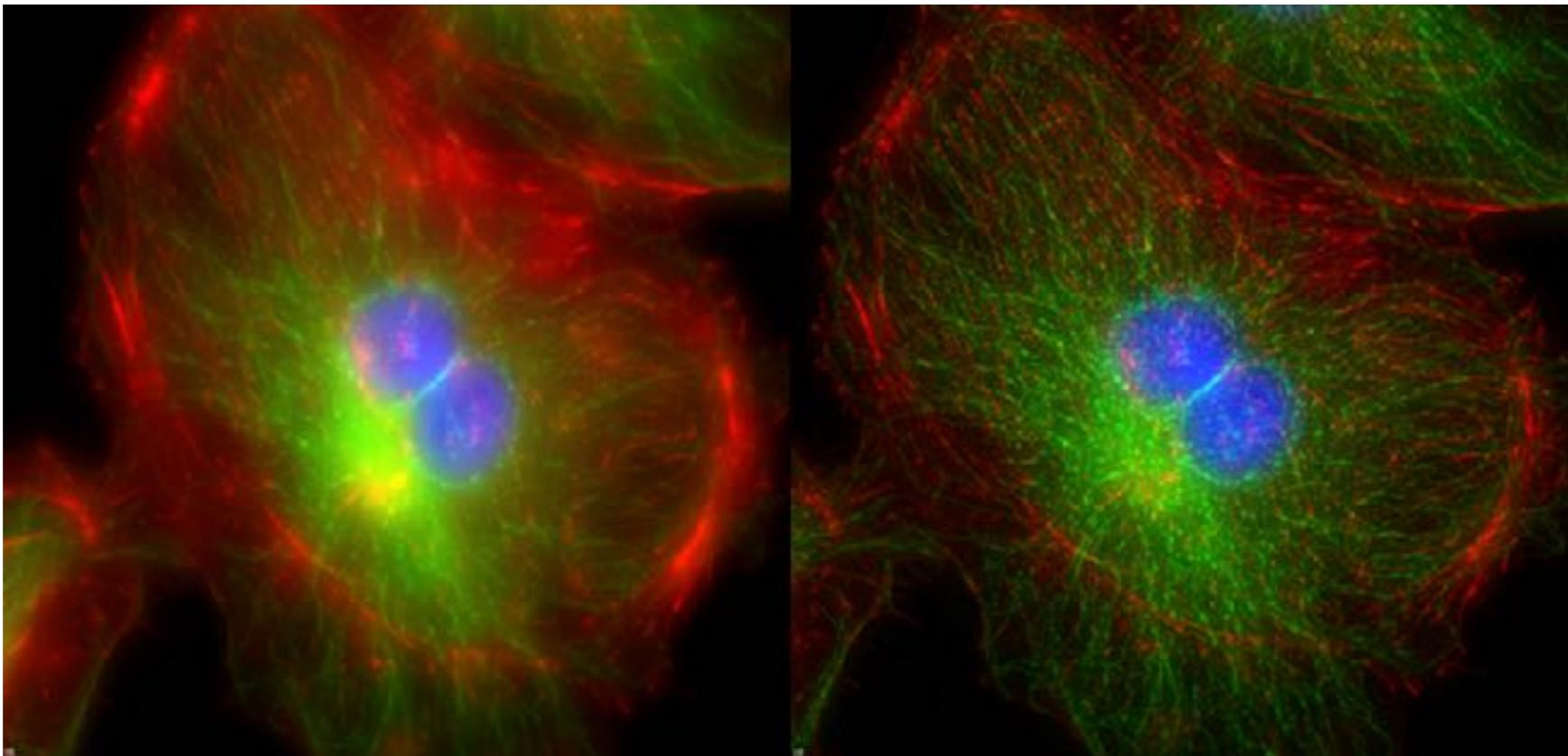
Observation:  $f = Hu + \epsilon$

- Formulation as an inverse problem:

$$f = Hu + \epsilon$$
$$\hat{u} = \arg \min_u D(f, Hu) + \mathcal{R}(u)$$

↑  
Data term      ↑  
Regularization

# AI: image restoration and inverse problems



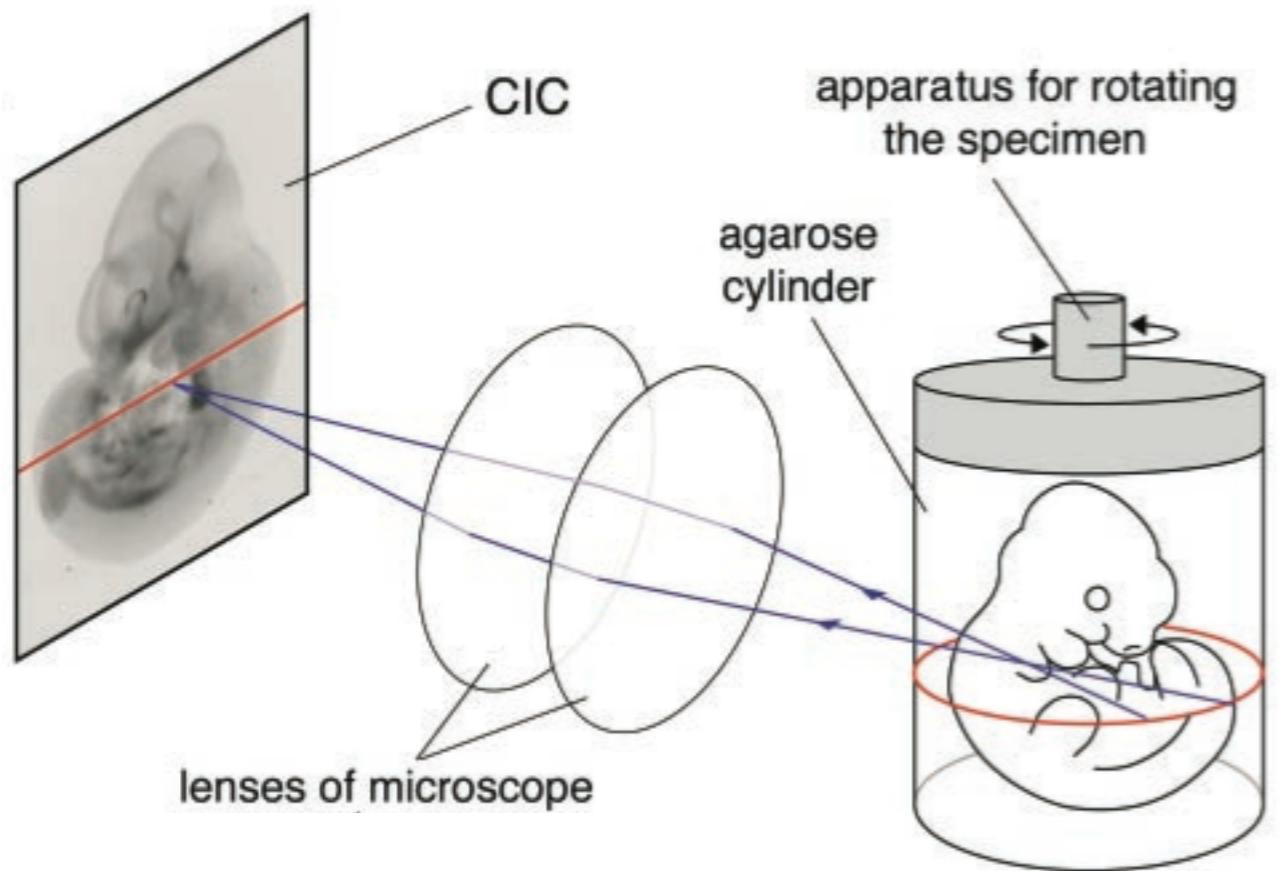
Sibarita et al, 2005

Applications include:

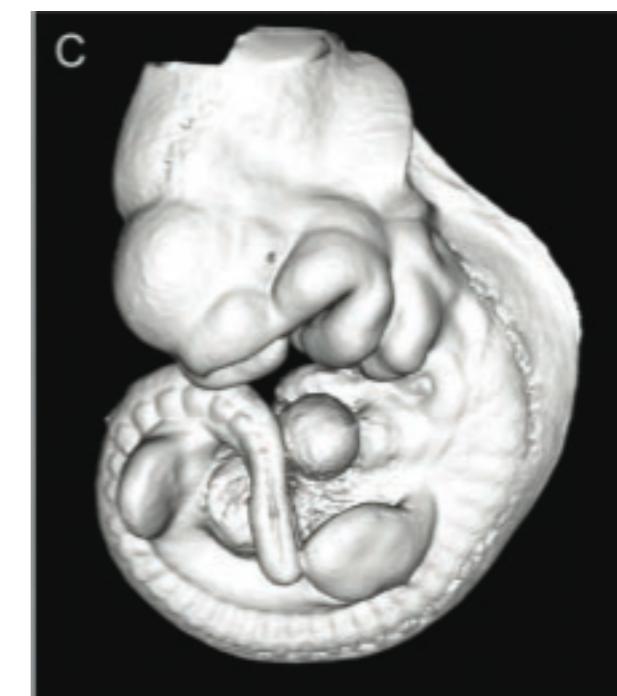
- denoising: finding the most likely noise-free version of the image
- deconvolution: compensation for out-of-focus contributions

## A2: Image Processing for image generation

- Many modern microscopy techniques rely on heavy image processing in order to generate an image.
- Example: optical projection tomography



Principle of optical projection tomography:  
the specimen is turned and optical projections  
from different angles are recorded.

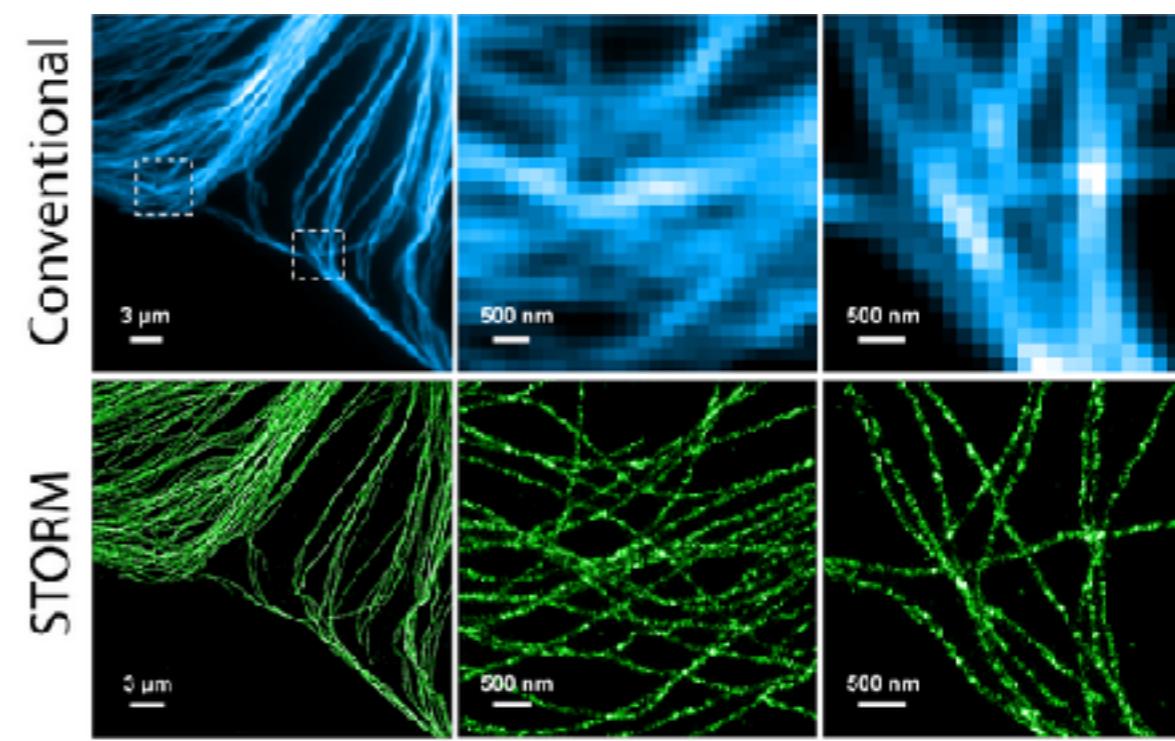
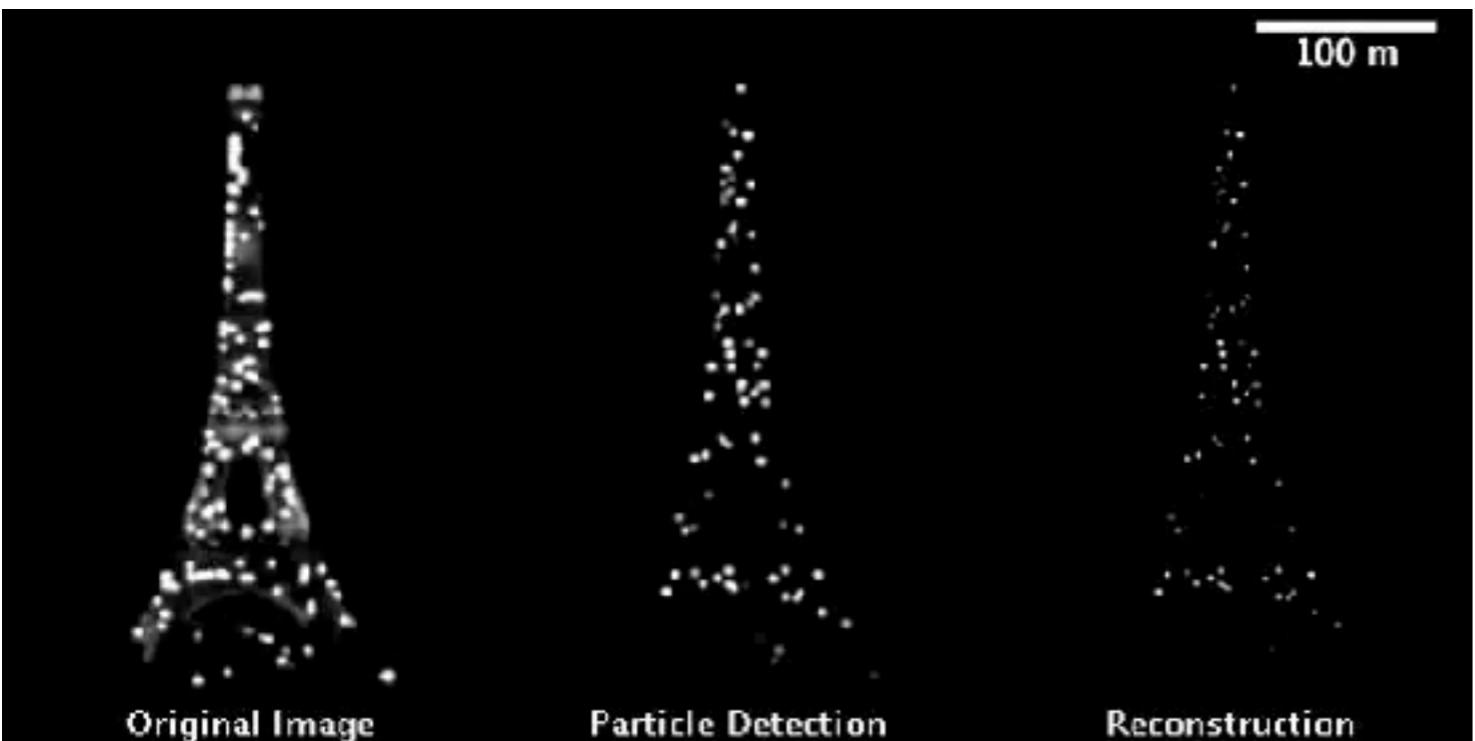


Reconstruction of a 3D volume  
from 400 projections under different  
angles

Sharpe et al, 2002

## A2: Image Processing for image generation

- Many modern microscopy techniques rely on heavy image processing in order to generate the image.
- Example: super-resolution microscopy



Bates et al, 2007

Super-resolution: family of techniques resulting in images with a better resolution to what is physically possible.

Trick: visualize only a small (but changing) set of points, to detect these points with high precision and to reconstruct the entire image from these detections.

Super-resolution microscopy is very popular in biology today and requires heavy processing of images.

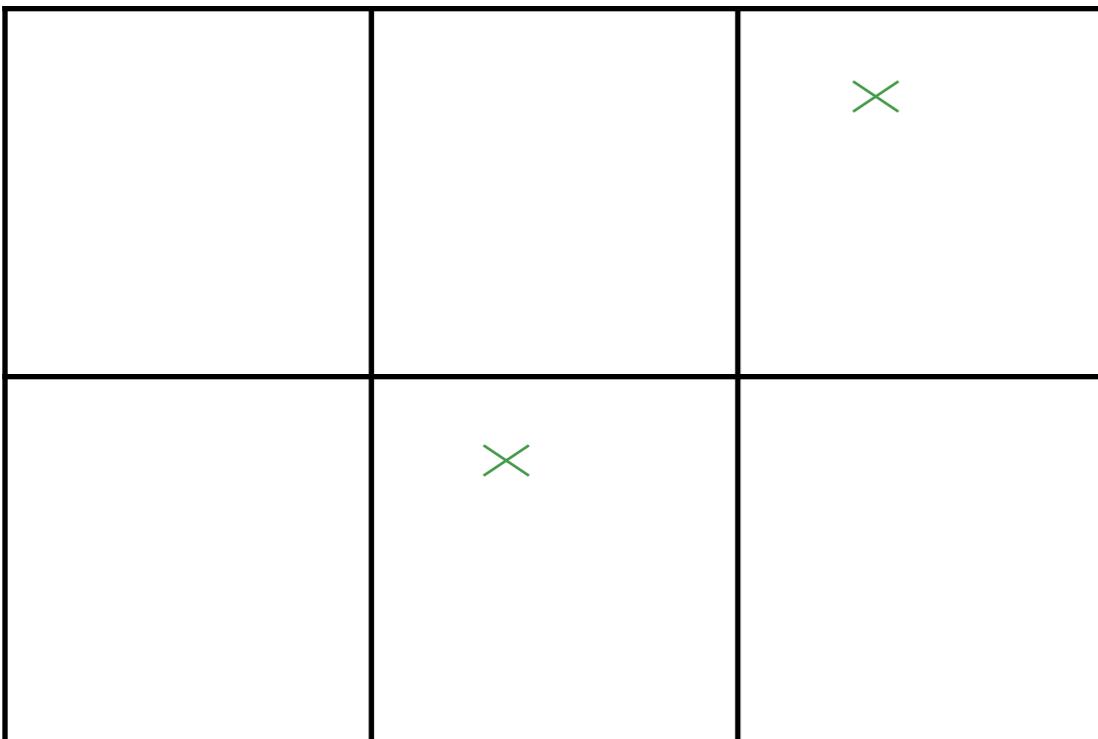
Nobel price in chemistry, 2014.

# A3: the smart microscope

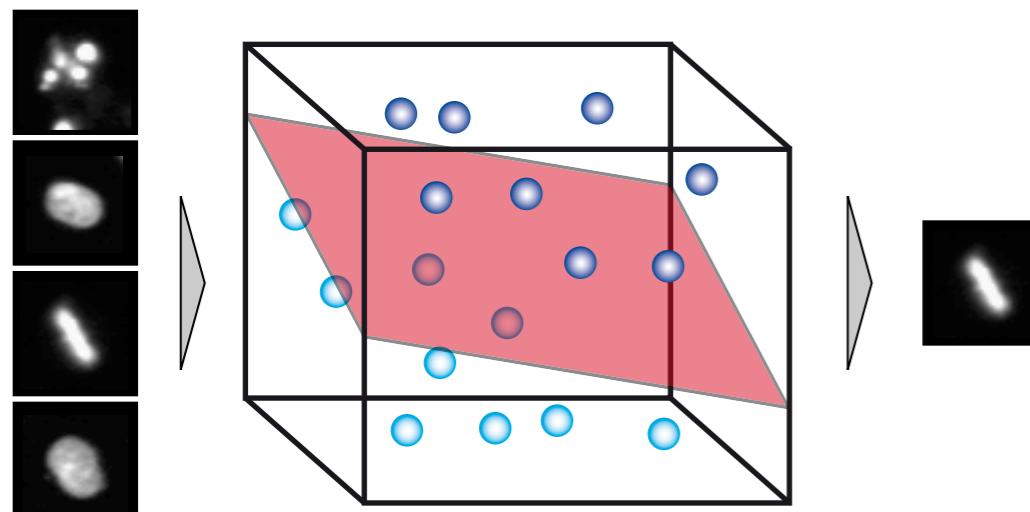
- One important aspect in modern microscopy is automatization.
- When performing imaging experiments on cellular populations, we are often interested only in a small subset of cells (e.g. dividing cells, migrating cells, cells expressing a certain protein, etc.)
- For this, the practitioner needs to manually select the cells of interest and perform the more complex imaging experiment.
- A better solution is the “smart microscope” that can select cells automatically and perform more advanced imaging experiments for these.

# A3: the smart microscope

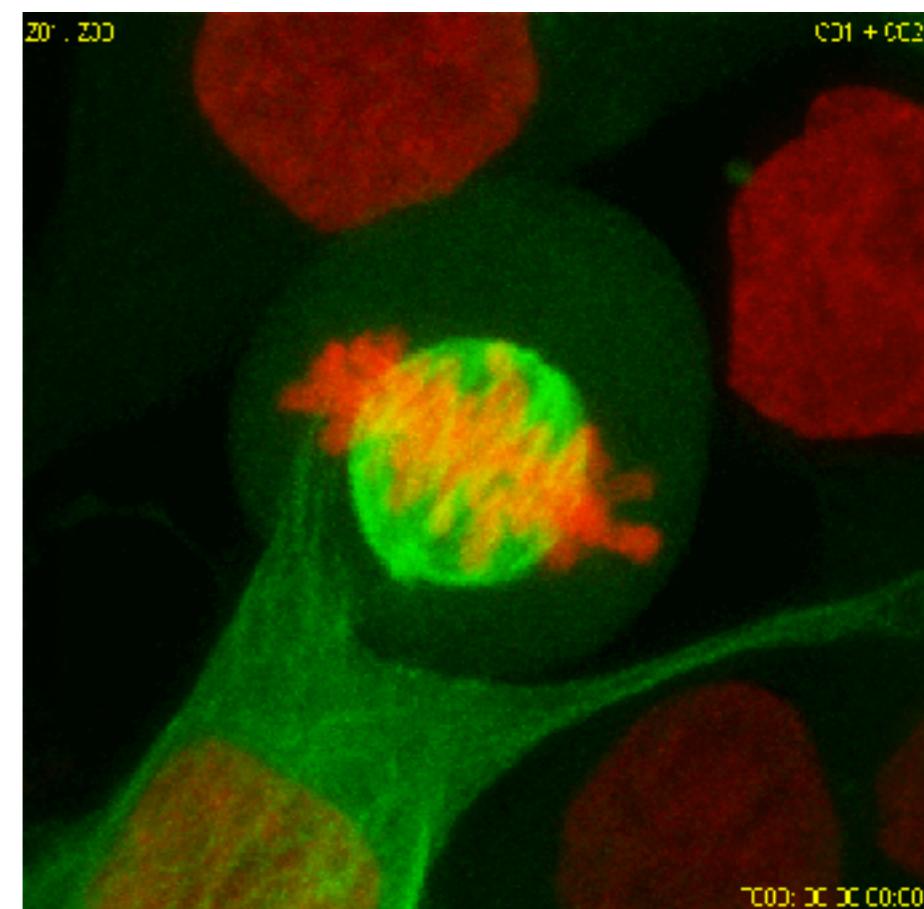
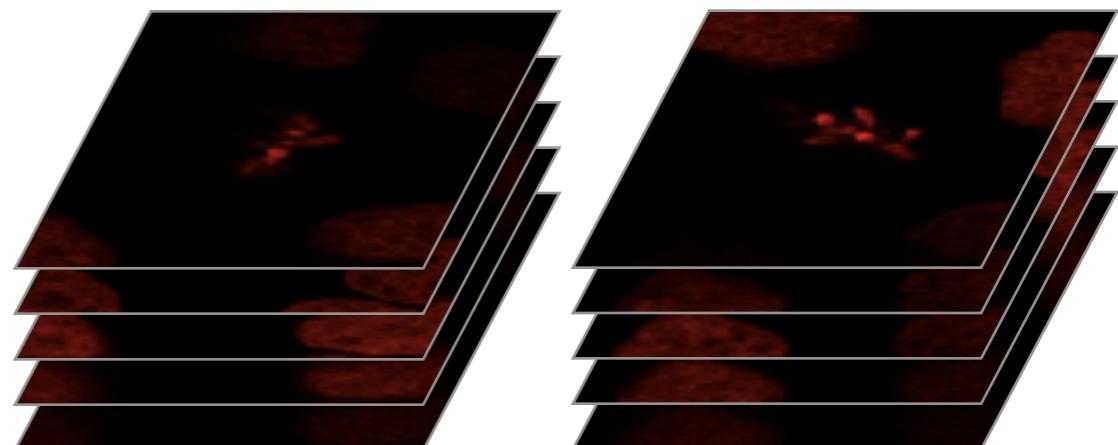
prescan



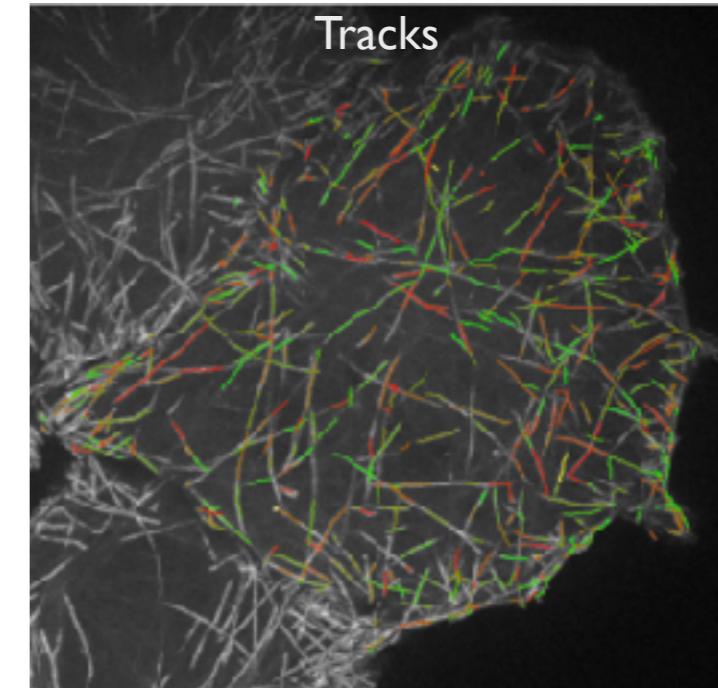
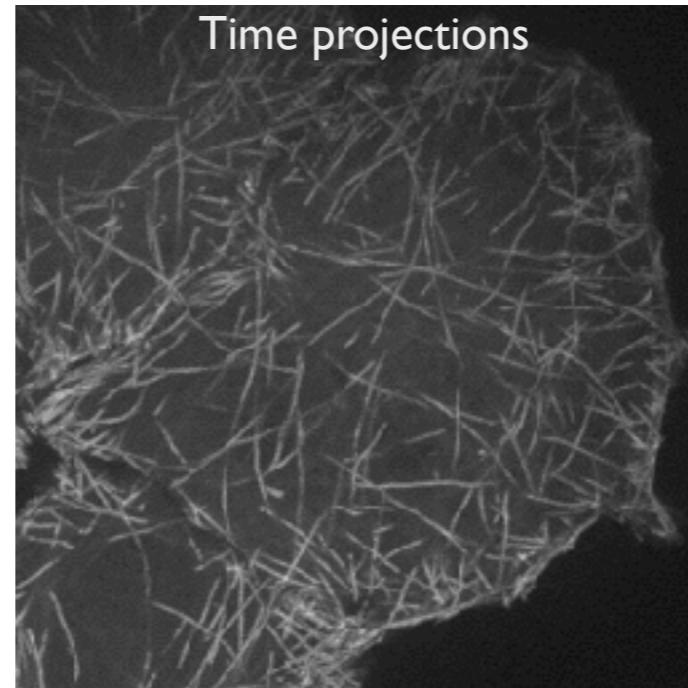
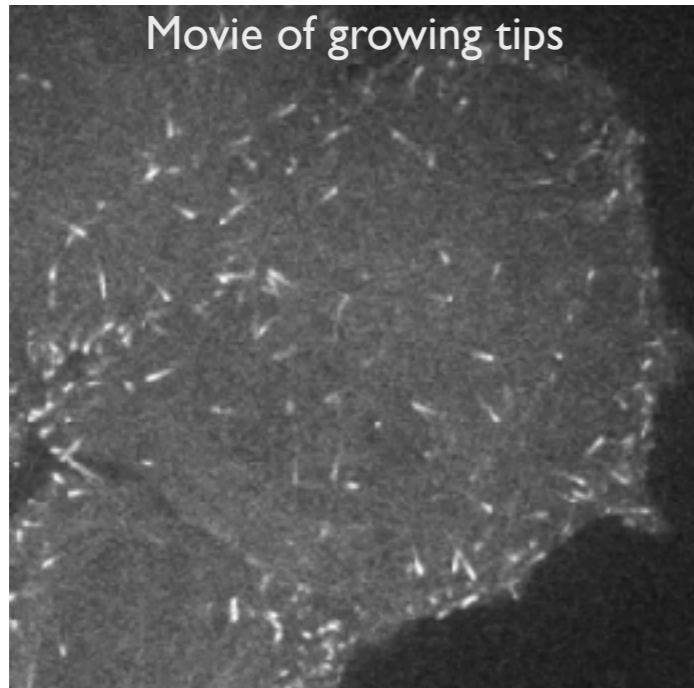
↔  
online classification



high resolution scan for automatically detected cellular events

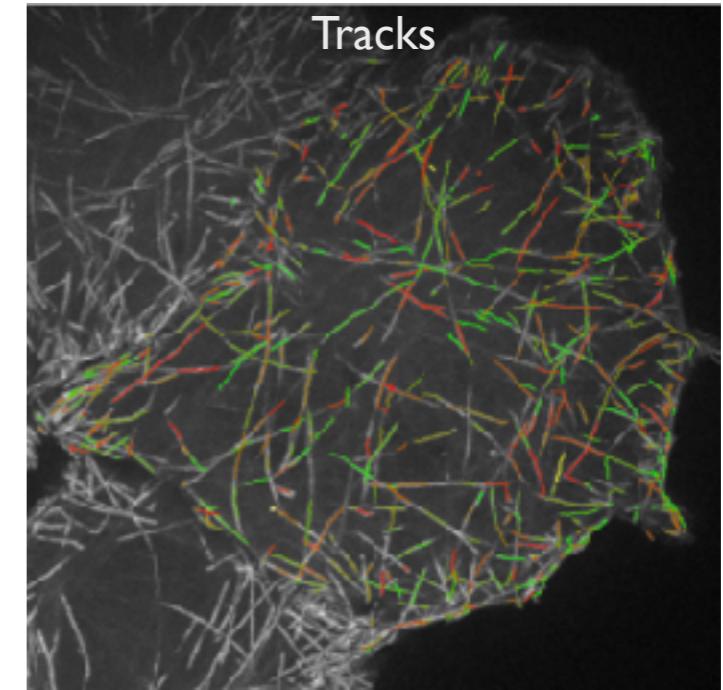
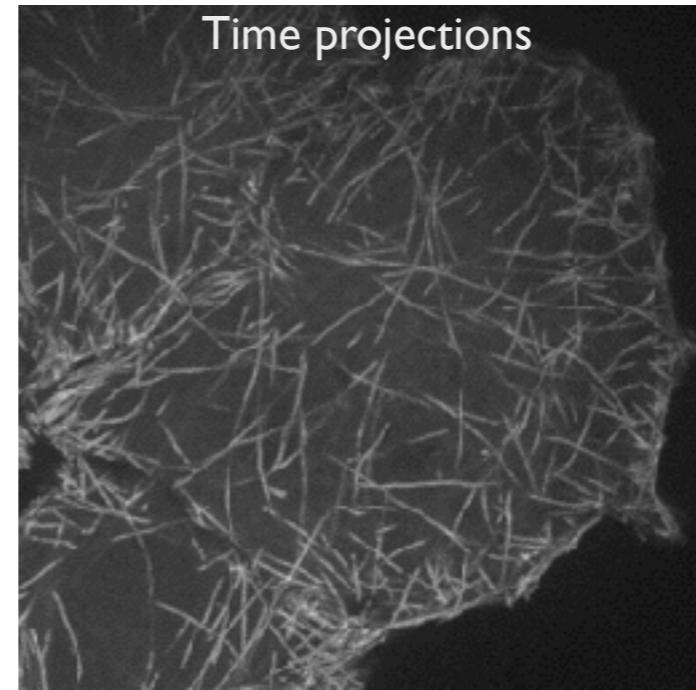
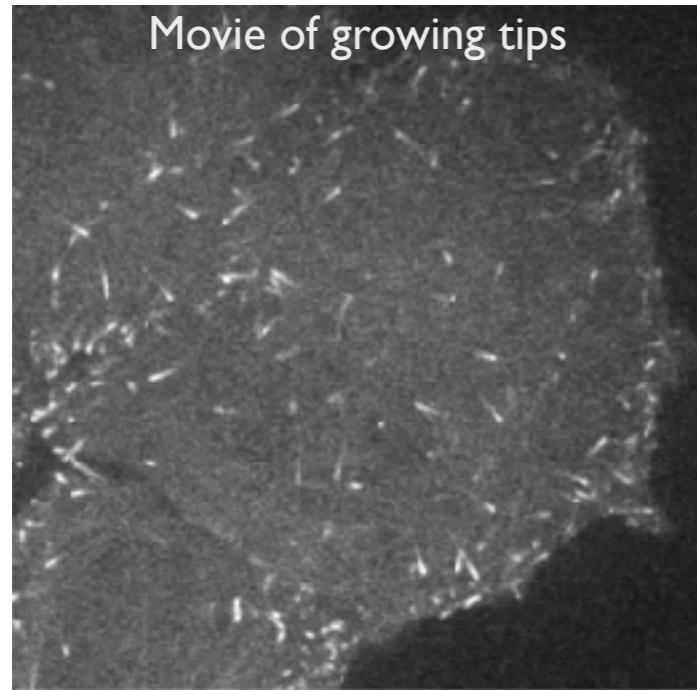


# A4: Measurements of biophysical properties

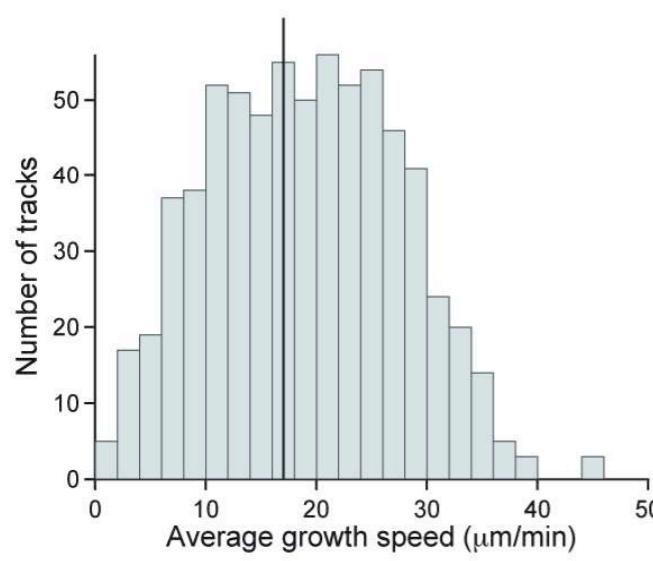


- Most traditional use case of image analysis: to quantify the information in an image.
- This boils down to classical image analysis problems such as segmentation (finding cells, detecting spots, ...), tracking, statistical data analysis, ...

# A4: Measurements of biophysical properties

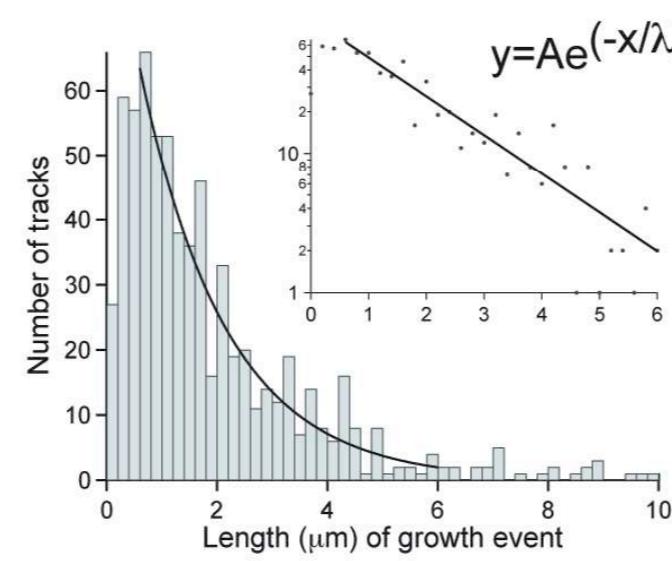


growth speed



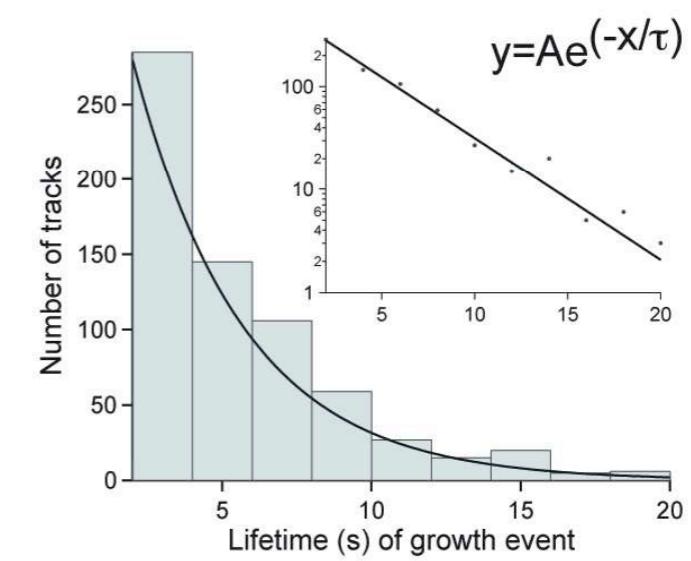
Median (average speed) =  $17.6 \mu\text{m}/\text{min}$

track length



$\lambda$  (characteristic track length) =  $1.6 \mu\text{m}$

lifetime



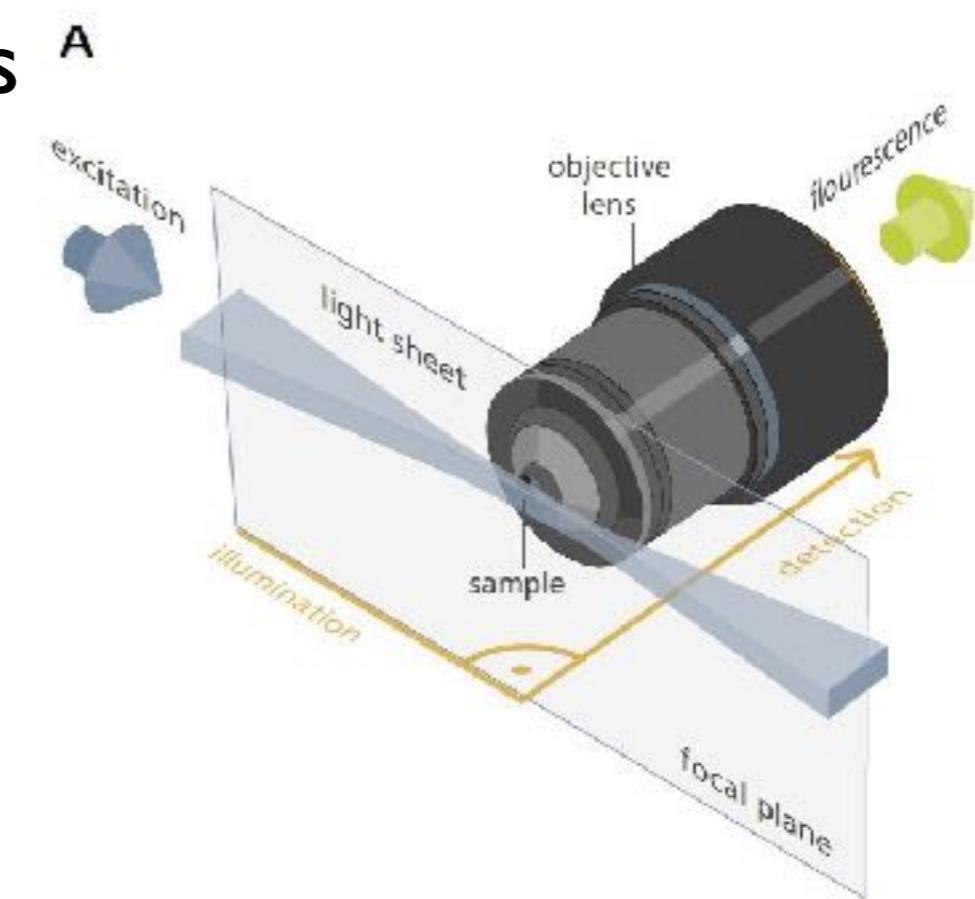
$\tau$  (characteristic track lifetime) =  $3.2 \text{s}$

## A5: handling big data volumes

- There are new microscopy techniques allowing to image large samples in 3D and with high temporal resolution for a long time.

- The trick is to illuminate only the part of the sample that is actually imaged and thereby to avoid phototoxicity.

- This leads to extremely large and complex data sets.
- Challenges include: 3D image reconstruction, handling of TB of image data, analysis and visualization

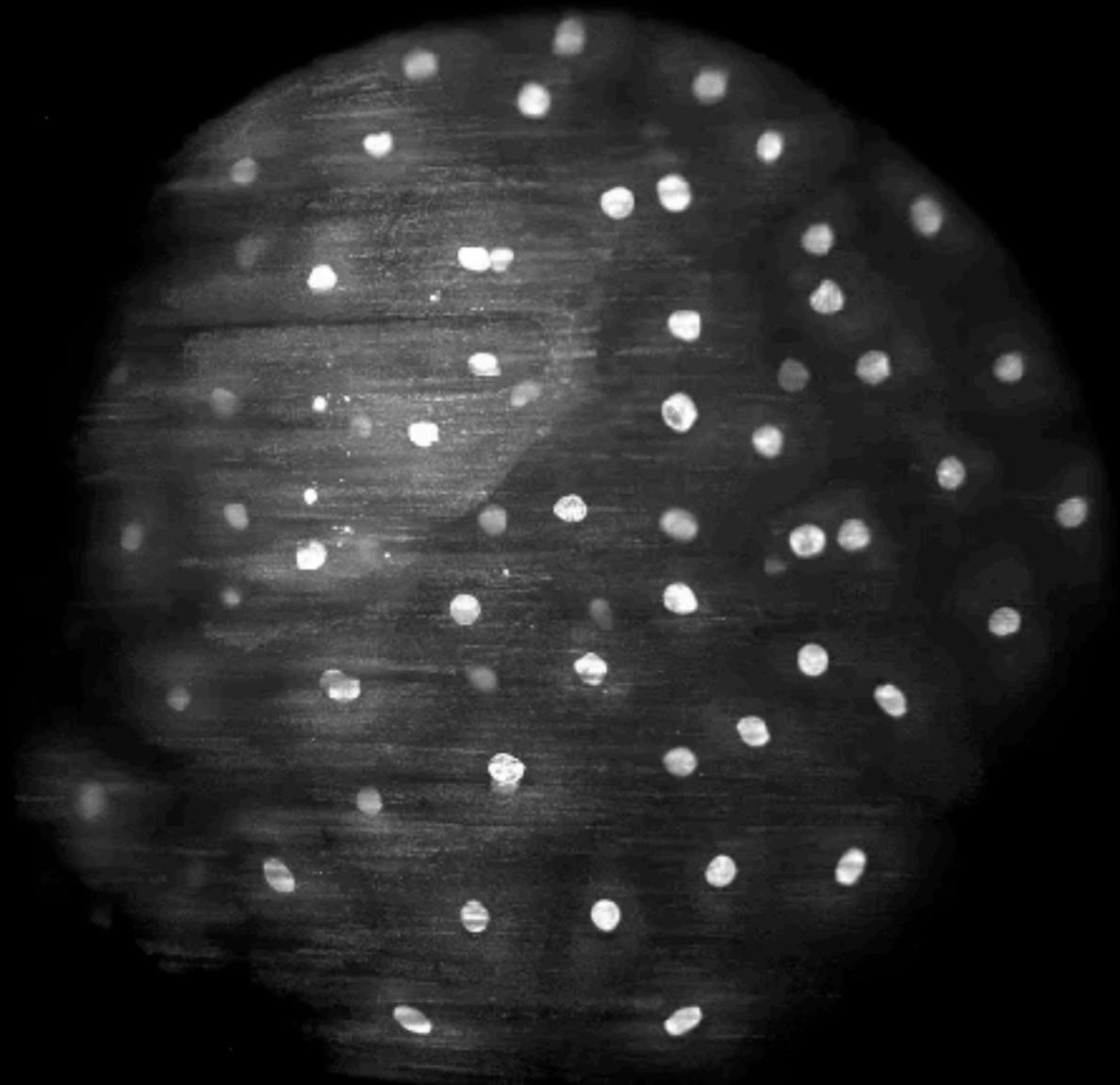


Light Sheet microscopy, Huisken et al., 2004

# A5: handling big data volumes

100 min

future dorsal side



animal view

future dorsal side

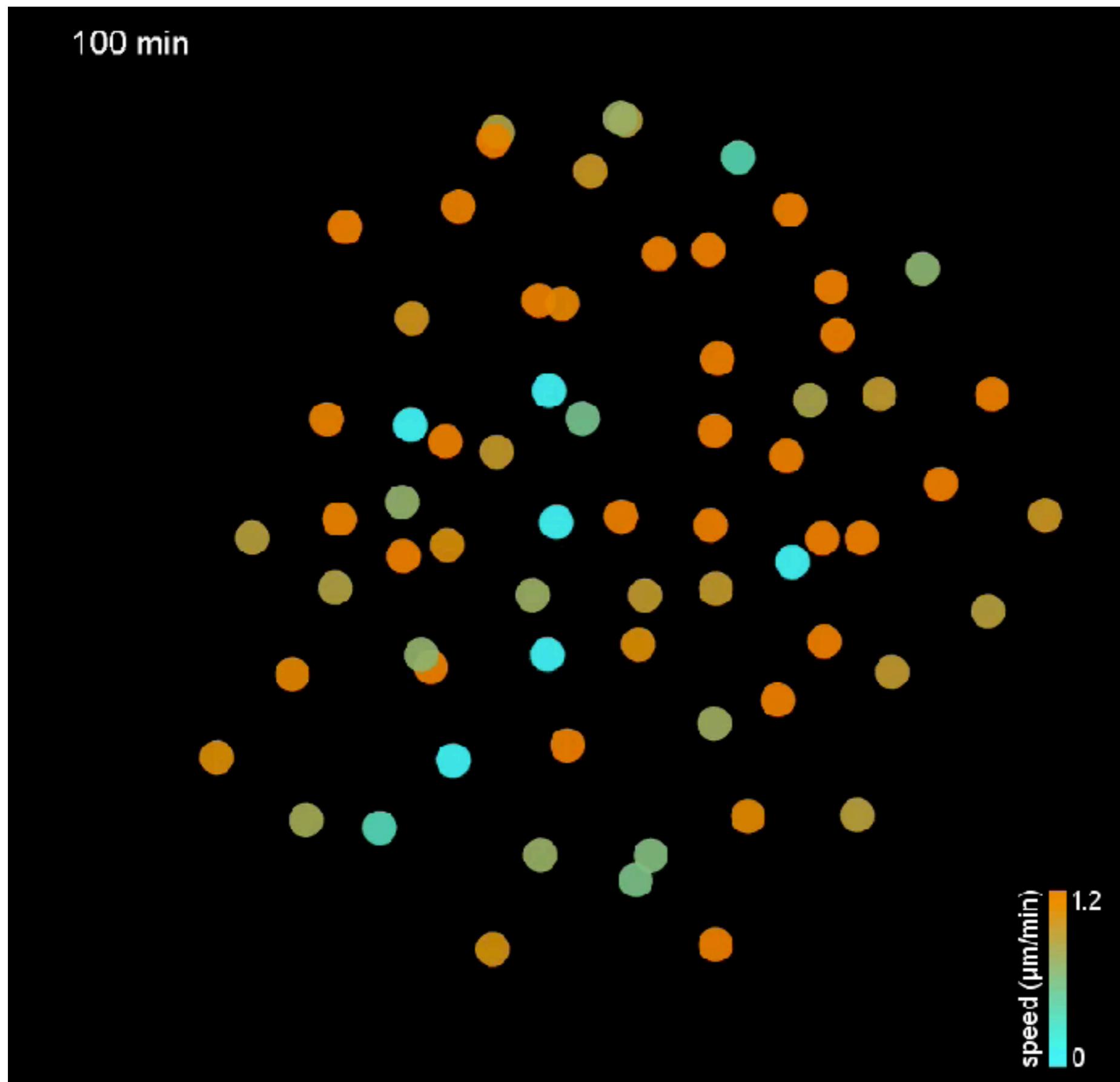


vegetal view

Keller et al., 2008

The movies show the first day in the life of a zebrafish embryo. Cells are stably expressing H2B-GFP. Images are taken with a SPIM microscope. One movie typically contains > 4TB. 19

# A5: handling big data volumes



Keller et al., 2008

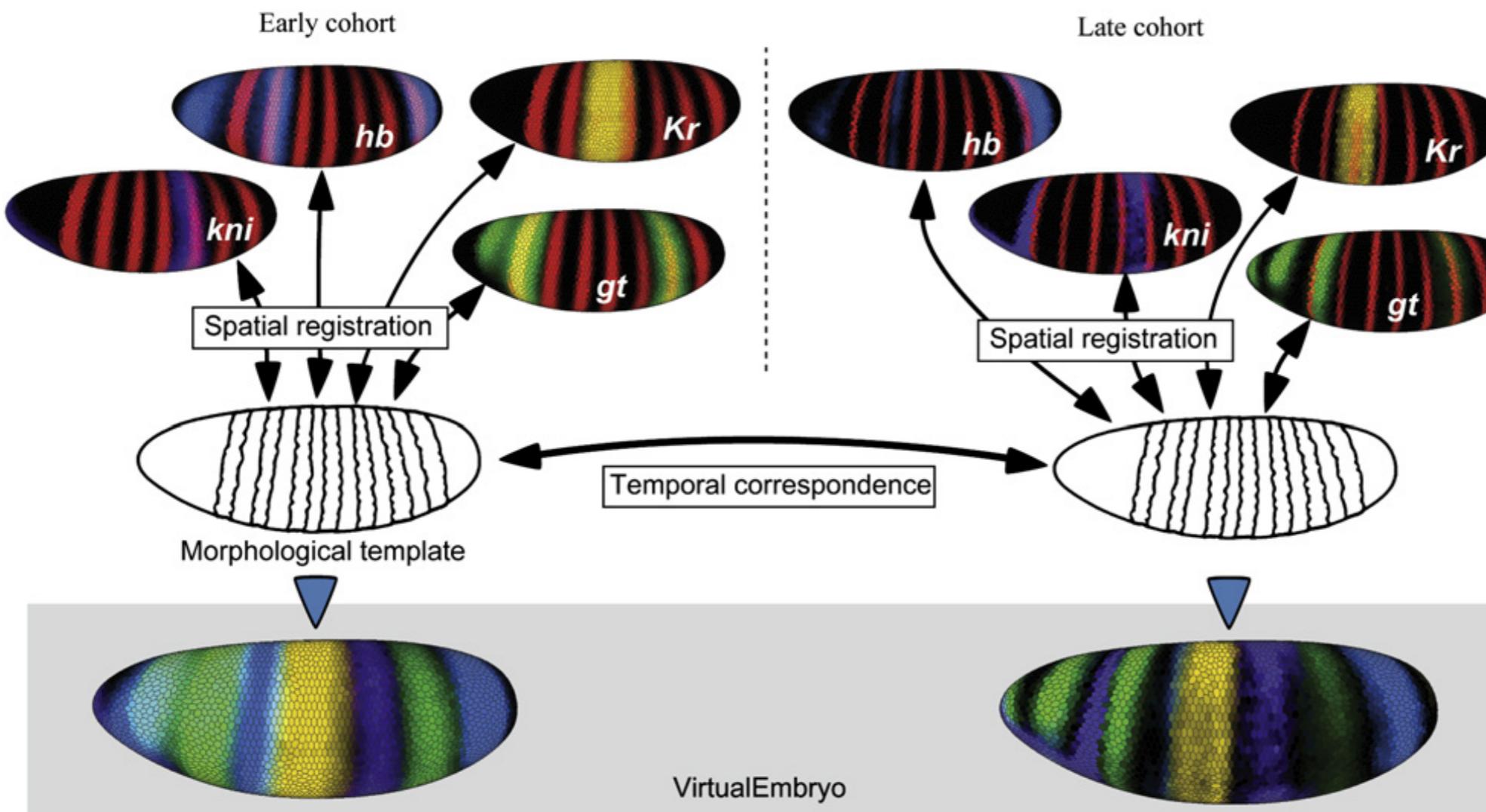
## A6: combining experiments

- Microscopy: only a limited number of markers (fluorescently labeled proteins or mRNA).
- Objective: comprehensive view on biological systems.
- Problem: how can we visualise many markers together?
- Example: transcription patterns in developing embryos
  - We want to understand when which gene is expressed in a developing embryo.
  - Only a small number of markers can be used.
  - Solution: atlas based approach.

# A6: combining experiments

Example: a spatiotemporal gene expression atlas:

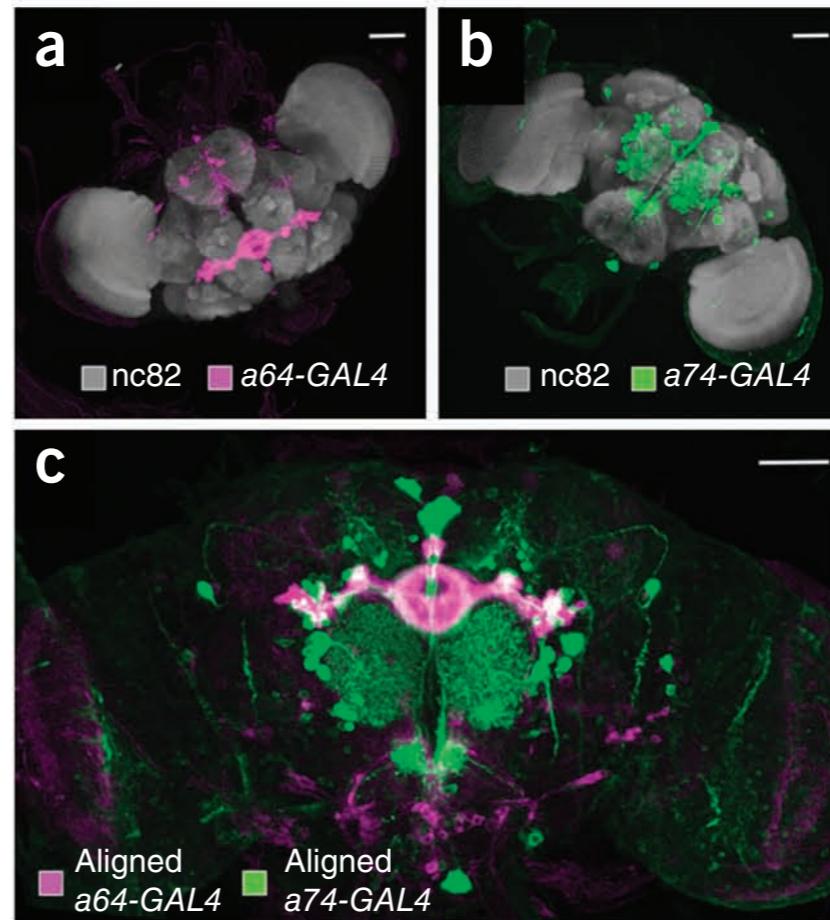
- Each embryo is stained with 3 markers (all nuclei, one reference gene and one gene of interest)
- All embryos are mapped to a virtual embryo (using the reference gene and the nuclei for the spatial mapping).



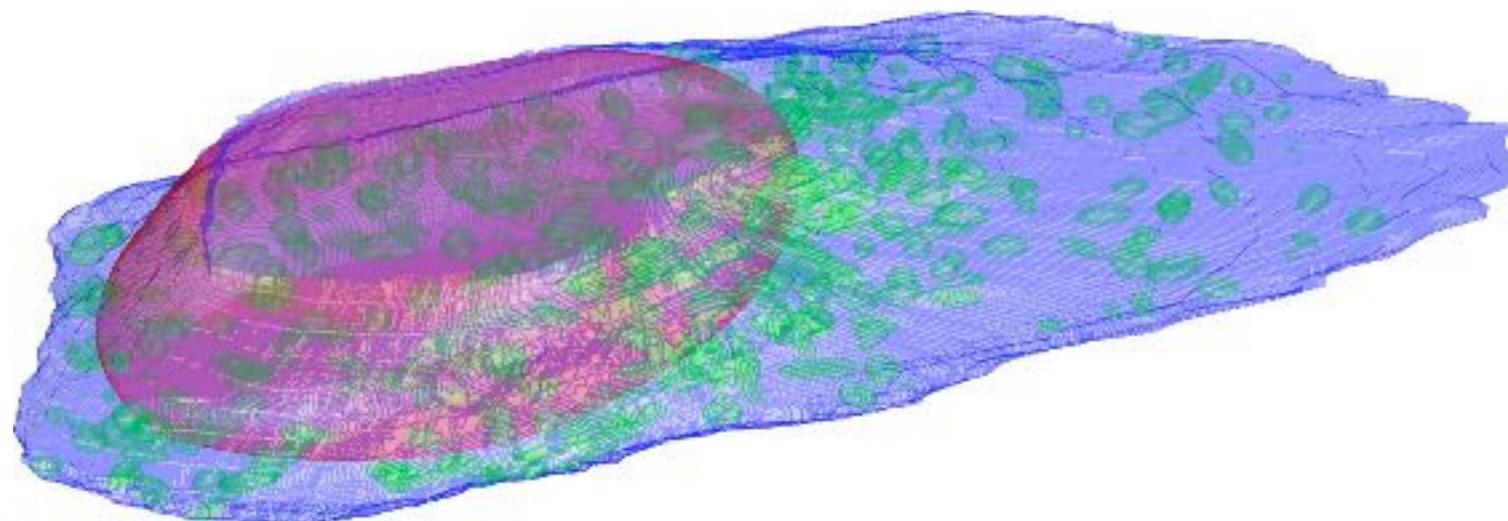
Fowlkes et al., 2008

## A6: combining experiments

- Example: Alignment of drosophila brains to a brain atlas:

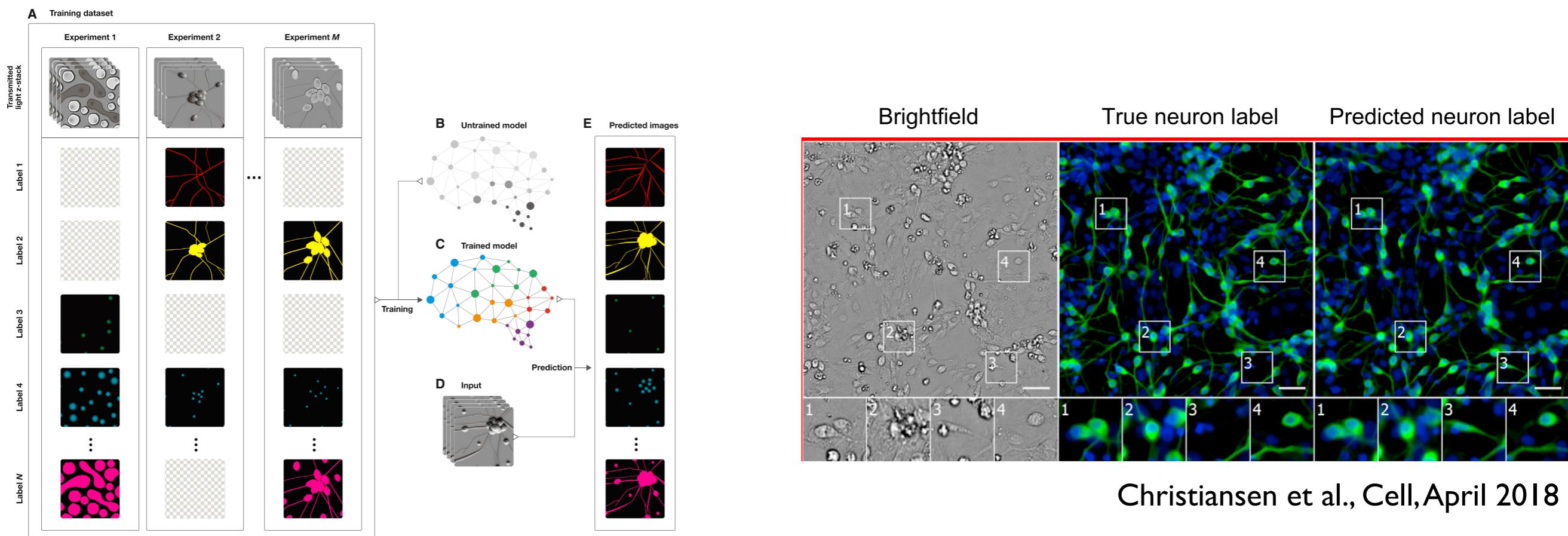


- Example: virtual cell model for subcellular protein localization



# A6: combining experiments

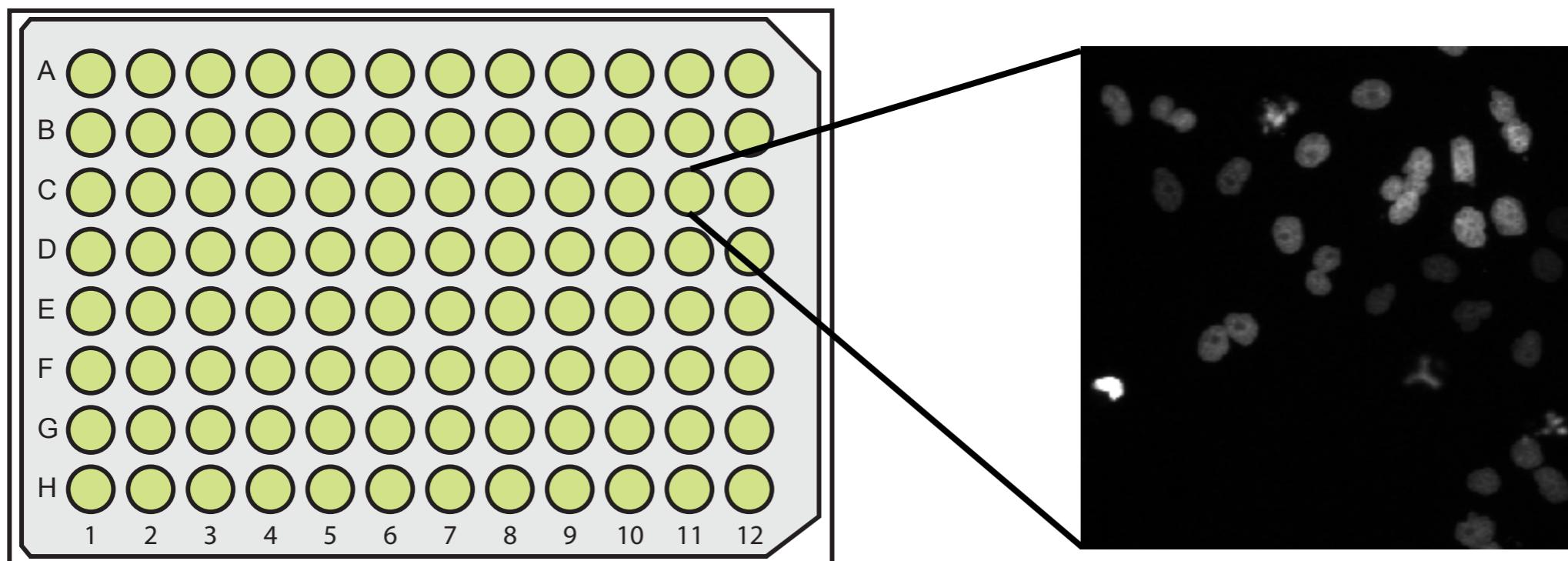
- It is also possible to predict fluorescence labels from label-free images. This also allows for combination of information describing very different aspects of a cellular system.



Christiansen et al., Cell, April 2018

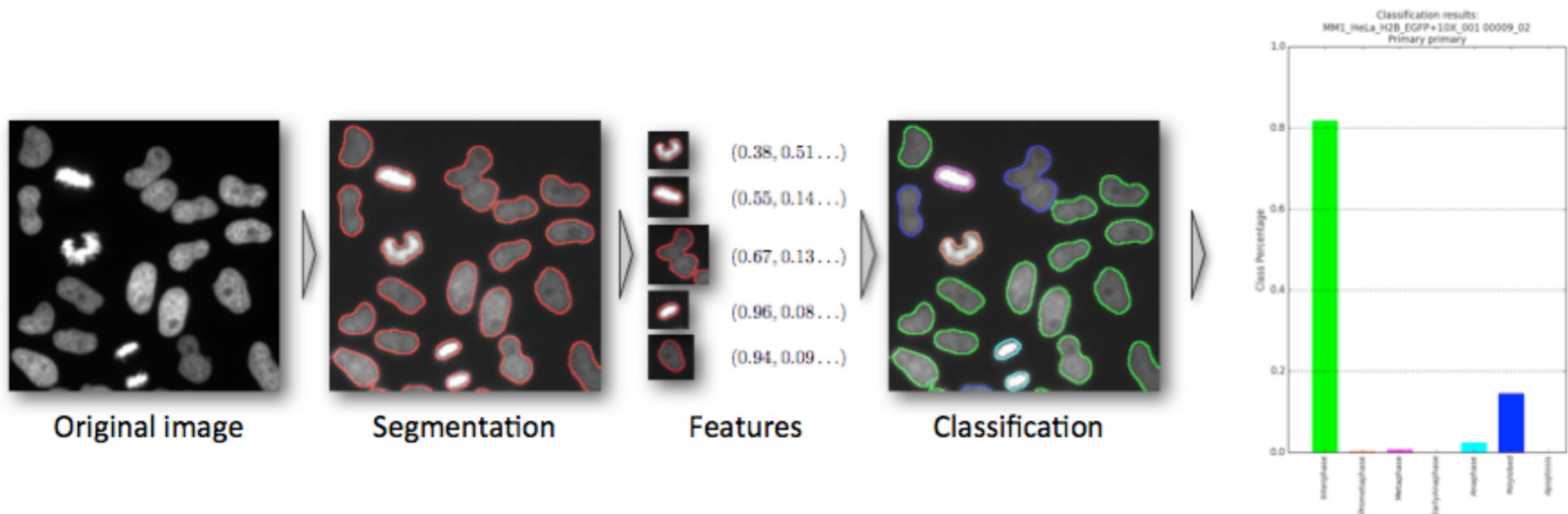
# A7: High-throughput / High Content Screening

- High-throughput:
  - high degree of automation
  - many experiments in parallel under controlled conditions
- High content: high information content of each single experiment.
- Example: drug screen/ RNAi screen

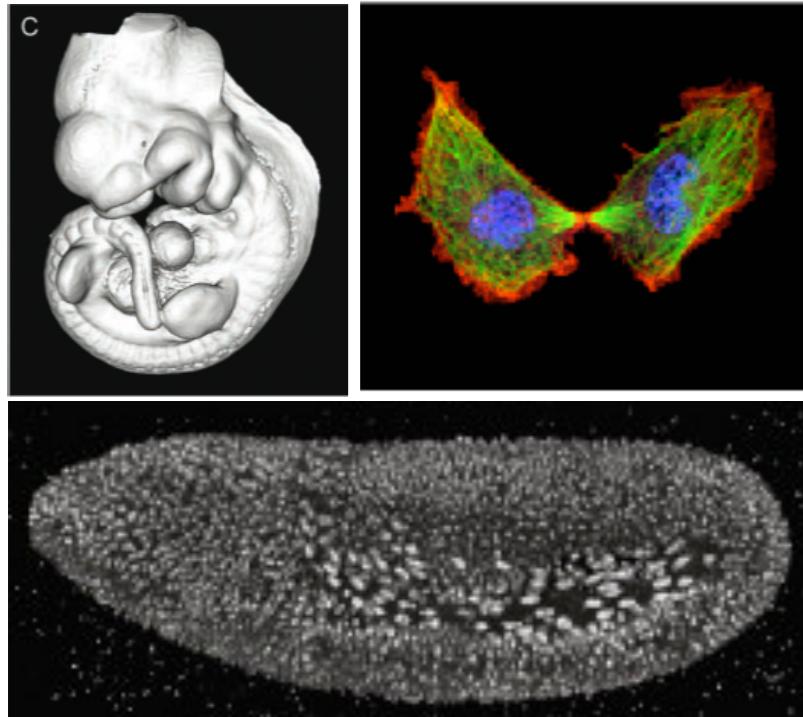


# A7: High-throughput / High Content Screening

- hundreds of thousands of single experiments
- Typical questions:
  - Which experiments deviate significantly from normality?
  - Can we identify conditions with similar effects?
- “Understanding” of cellular phenotypes.



## A8: Image data bases - the problem

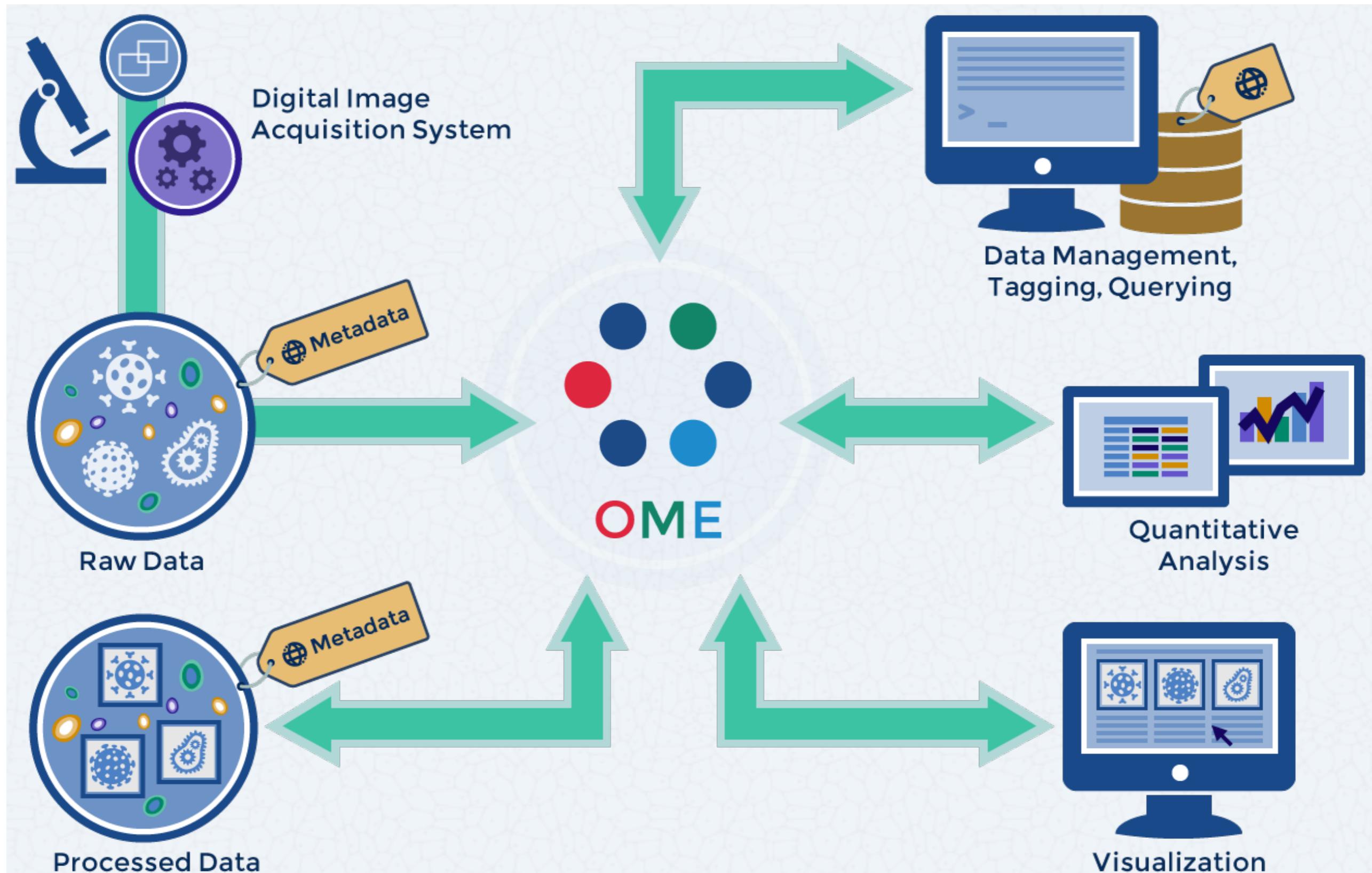


### **Images in biology**

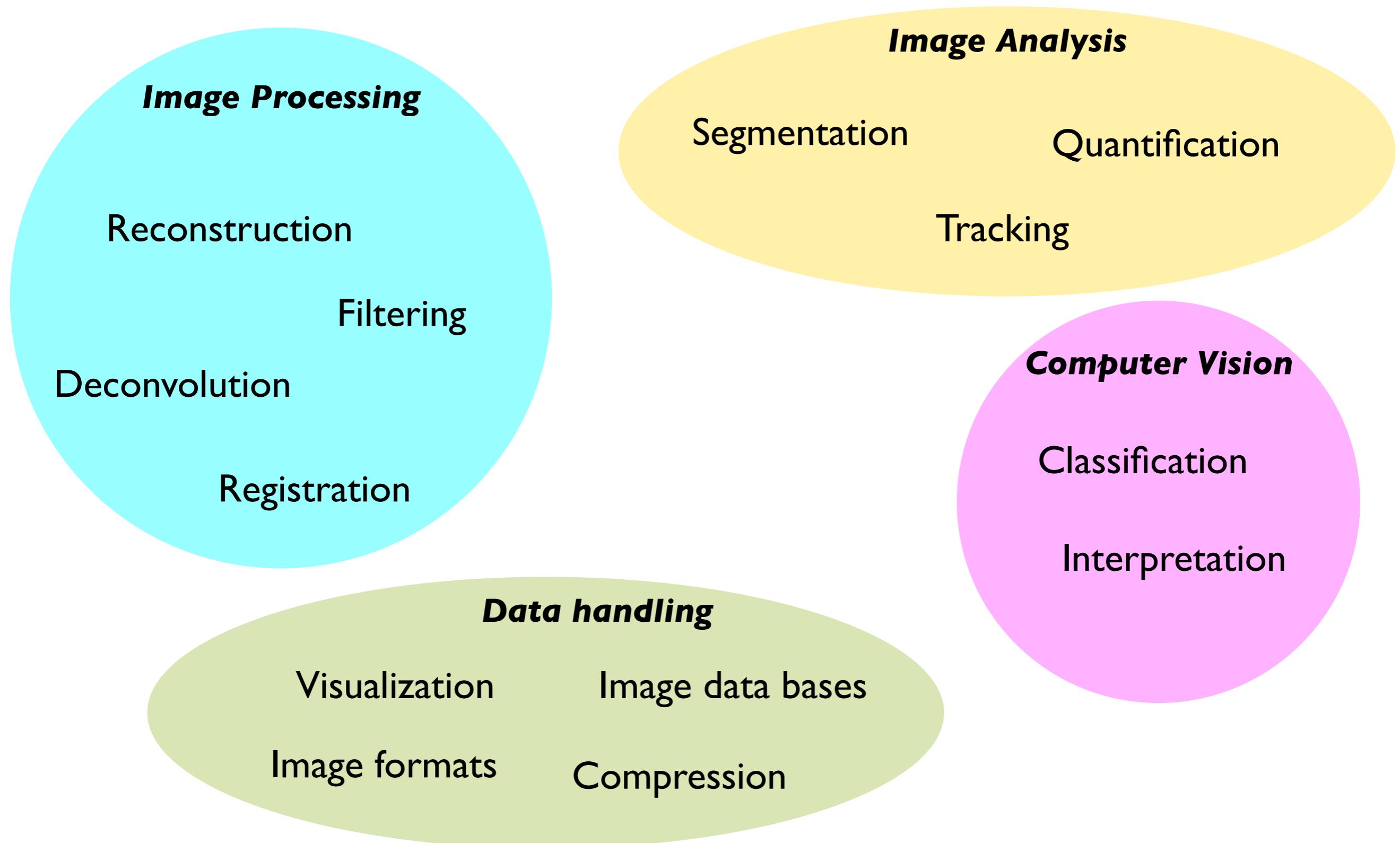
Images are subjectively interpretable measurements.

- Different image formats.
- Link between images and quantitative analysis results and the analysis workflow which produced the results.
- Intelligent browsing of images.
- Images as scientific resource.

# A8: Image data bases - solutions



# Current problems in Bioimage Informatics



# What is Bioimage Informatics?

*From my perspective, it is very reminiscent of the state of bioinformatics in the early 1980s: the exciting, somewhat chaotic free-for-all that is potentially the birth of something new.*

Gene Myers

# References 1/2

- (1) Danuser, G. (2011). Computer Vision in Cell Biology. *Cell*, 147(5), 973–978.
- (2) Myers, G. (2012). Why bioimage informatics matters. *Nature Methods*, 9(7), 659–60.
- (3) Peng, H. (2008). Bioimage informatics: a new area of engineering biology. *Bioinformatics* (Oxford, England), 24(17), 1827–36.
- (4) Peng, H., Bateman, A., Valencia, A., & Wren, J. D. (2012). Bioimage informatics: a new category in Bioinformatics. *Bioinformatics* (Oxford, England), 28(8), 1057.
- (5) Pankajakshan, P., Blanc-féraud, L., Zhang, B., & Kam, Z. (2008). Parametric Blind Deconvolution for Confocal Laser Scanning Microscopy ( CLSM ) -Proof of Concept (pp. 1–45).
- (6) Sibarita, J. (2005). Deconvolution Microscopy. In *Advances in Biochemical Engineering/ Biotechnology* (pp. 201–243). Springer-Verlag Berlin Heidelberg.
- (7) Sharpe, J., Ahlgren, U., Perry, P., Hill, B., Ross, A., Hecksher-Sørensen, J., ... Davidson, D. (2002). Optical Projection Tomography as a Tool for 3D Microscopy and Gene Expression Studies. *Science*, 296, 541–545.
- (8) Conrad, C., Wünsche, A., Tan, T. H., Bulkescher, J., Sieckmann, F., Verissimo, F., ... Ellenberg, J. (2011). Micropilot: automation of fluorescence microscopy-based imaging for systems biology. *Nature Methods*, 8(3), 246–9.
- (9) Keller, P. J., Schmidt, A. D., Wittbrodt, J., & Stelzer, E. H. K. (2008). Reconstruction of zebrafish early embryonic development by scanned light sheet microscopy. *Science* (New York, N.Y.), 322(5904), 1065–9.
- (10) Fowlkes, C. C., Hendriks, C. L. L., Keränen, S. V. E., Weber, G. H., Rübel, O., Huang, M.-Y., ... Malik, J. (2008). A quantitative spatiotemporal atlas of gene expression in the *Drosophila* blastoderm. *Cell*, 133(2), 364–74.
- (11) Peng, H., Chung, P., Long, F., Qu, L., Jenett, A., Seeds, A. M., ... Simpson, J. H. (2011). BrainAligner : 3D registration atlases of *Drosophila* brains. *Nature Methods*, 8(6), 493–500.
- (12) Peng, T., & Murphy, R. F. (2011). Image-derived ,Three-dimensional Generative Models of Cellular Organization. *Cytometry*, (3), 383–391.
- (13) Goldberg, I. G., Allan, C., Burel, J.-M., Creager, D., Falconi, A., Hochheiser, H., ... Swedlow, J. R. (2005). The Open Microscopy Environment (OME) Data Model and XML file: open tools for informatics and quantitative analysis in biological imaging. *Genome Biology*, 6(5), R47.
- (14) Swedlow, J. R., Goldberg, I. G., & Eliceiri, K. W. (2009). Bioimage informatics for experimental biology. *Annual Review of Biophysics*, 38, 327–46.

# References 2/2

- (15) Terjung, S., Walter, T., Seitz, A., Neumann, B., Pepperkok, R., & Ellenberg, J. (2010). High-throughput microscopy using live mammalian cells. In R. D. Goldman, J. R. Swedlow, & D. L. Spector (Eds.), *Live Cell Imaging: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.
- (16) Sbalzarini, I. F., & Koumoutsakos, P. (2005). Feature point tracking and trajectory analysis for video imaging in cell biology. *Journal of Structural Biology*, 151(2), 182–95.
- (17) Sironi, L., Solon, J., Conrad, C., Mayer, T. U., Brunner, D., & Ellenberg, J. (2011). Automatic quantification of microtubule dynamics enables RNAi-screening of new mitotic spindle regulators. *Cytoskeleton (Hoboken, N.J.)*, 68(5), 266–78.
- (18) Christiansen, E. M., Yang, S. J., Ando, D. M., Javaherian, A., Skibinski, G., Lipnick, S., ... Finkbeiner, S. (2018). In Silico Labeling: Predicting Fluorescent Labels in Unlabeled Images. *Cell*, 173(3), 792–803