

## 2.1 Organization of work, type of data

---

Why is it important to process experimental data continuously?

Avoid this approach:

"I'm going to experiment this year. Next year I will evaluate what I have got".

### Types of data:

- **Numerical data** - raw data or data converted to the numerical form, may be obtained with simple tools (ruler etc.) or more complex instrumentation (spectrophotometer, etc.)
- **Graphical data** - very frequent type of data, their quantification is very important
- **Structural data** - analytical data, outputs from sequencing machines, mass spectrometers, etc.

## 2.1 Organization of work, type of data

---

### Software used in modern laboratory:

- **Software controlling machines** - very often quite complex. For the user, the most important is the form of the output (table, text)
- **General, „office“-type software** - text and spreadsheet editors, graphical software, presentation software, typically products of Microsoft at the Charles University
- **Dedicated software** - very important in the field of image analysis and processing of sequence data
- **Open source software** - very good tools, there are often problems with the compatibility
- **Online software** - mainly in bioinformatics and structure analysis, but there are also a plenty of software available for various routine office applications

## 2.1 Organization of work, type of data

---

**Windows Explorer** or similar software (Windows Commander)

Individual folders should be named according to their date, ideally as **YYMMDD (211104)**. The date **must correspond well** to the one in the laboratory book.

**Never modify folder with raw data** - copying to the folder with other files accompanying the experiment.

**Back up data** - external HDD, shared laboratory HDD drives, or online **cloud repositories**.

**Do not postpone backuping!!!**

Experiment must be always finished with **complete documentation**, do not start the new one **without finishing the previous**.

# 2.2 Data explorers, correct data handling and saving

## Windows Explorer

The screenshot shows a Windows Explorer window titled '050420'. The left pane displays a folder tree with the following structure:

- Protokoly
  - 2002
  - 2003
  - 2004
  - 2005
    - kvasinky
    - MDRs
    - PBA
    - PINs**
      - 050302
      - 050308
      - 050316
      - 050317
      - 050323
      - 050406
      - 050420
      - 050614
      - 050624
      - AdrianakvanifikaceGE
      - materiály pro alignme

The right pane shows a table of files in the '050420' folder:

Název	Velikost	Typ	Změněno
2005 04 20 PIN7 _ 3H IAA vs NPA a BFA .JNB	312 kB	SigmaPlot 9.0 Note...	22.4.2005 13:54
2005 04 20 PIN7 _ 3H IAA vs NPA a BFA .xls	135 kB	List aplikace Microso...	21.4.2005 18:14
050420.doc	30 kB	Dokument aplikace ...	20.4.2005 16:02

## 2.2 Data explorers, correct data handling and saving

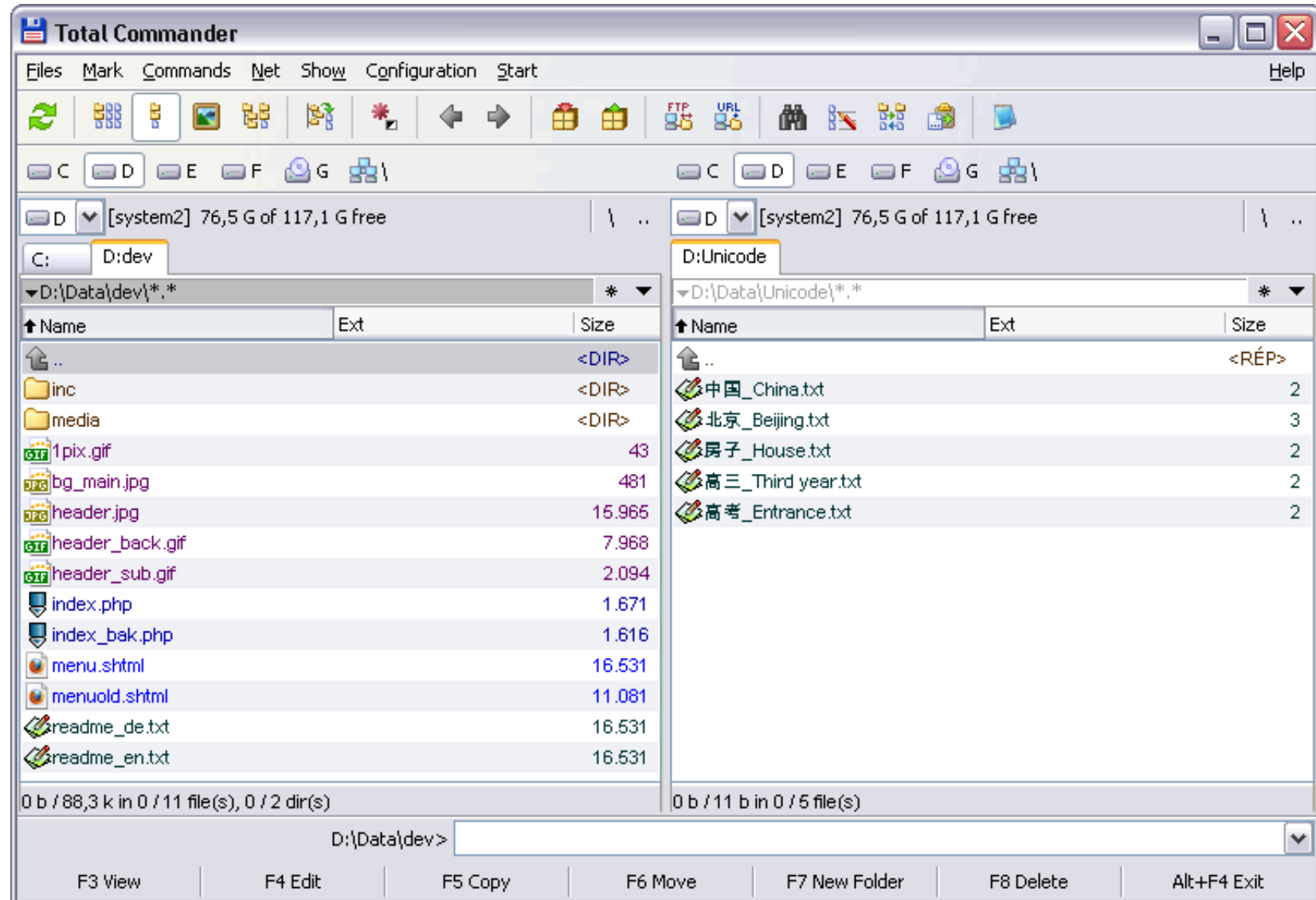
070420_PIN1GFP_ACC_AVG	070420-0h	
070423_PIN1GFP_pINTAMPINGFP_	070422-48h	
070424_PIN1GFP_ACCxAVG	070423-72h	
070427_PIN1-GFP_DR5GFP_AVG_2	070420.doc	26 kB
070503_PIN1-GFP_AVG_ACC_STS	report070420.cdr	17 158 kB
070504_PIN1-GFP_SN_ACC_AVG_s	report070420.pdf	1 066 kB
070509_PIN1-GFP_ACC_AVG_exp	Záloha_report070420.cdr	17 158 kB
070510_DR5-GFP_ACC_AVG_stac		
070510_PIN1-GFP_DR5GFP_MCP		
070514_PIN1-GFP_Et_gasing_1ppn		
070605_PIN1-GFP_ACCxAVG		
070607_BY-2wt_BRs1_2_3_Kohout		
070608_DR5rev-GFP_Et_gasing_1p		
070614_Et-gas_AVG_PIN1-GFP		
070614-PIN1GFP-BRs1_2_3_Khou		
070618_DR5_BRs1_2_3_Kohout		
070619_Et-gas-AVG_PIN1-GFP		
070621_Et-gas-AVG_PIN1-GFP		
070626_Et-gas-AVG_PIN1-GFP_int		
070627_PIN3-GFP_test		
070702_PIN1_trafficking		
071105_BRs_Kohout_GVGPIN7		
071110_BRs_Kohout_GVGPIN7_d2		

## 2.2 Data explorers, correct data handling and saving



### Total Commander

- shareware
- the biggest advantage is the existence of two identical windows



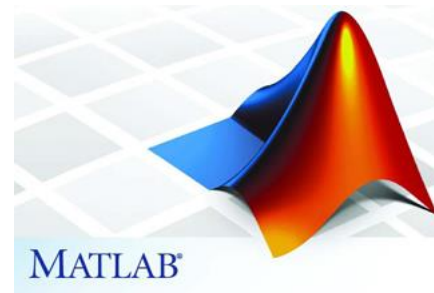
## 2.3 Spreadsheets, statistical software, graph editors

Bulk data processing, their editing and graphical output



**Microsoft Excel** - suitable for all basic operations including statistics and various graphical outputs

**Dedicated mathematical or statistical software** - [Statgraphics](#), [NCSS](#), [Matlab](#), or the open source alternative [R](#)



# 2.3 Spreadsheets, statistical software, graph editors

Keeping a complete log of the experiment, automation of numerical operations

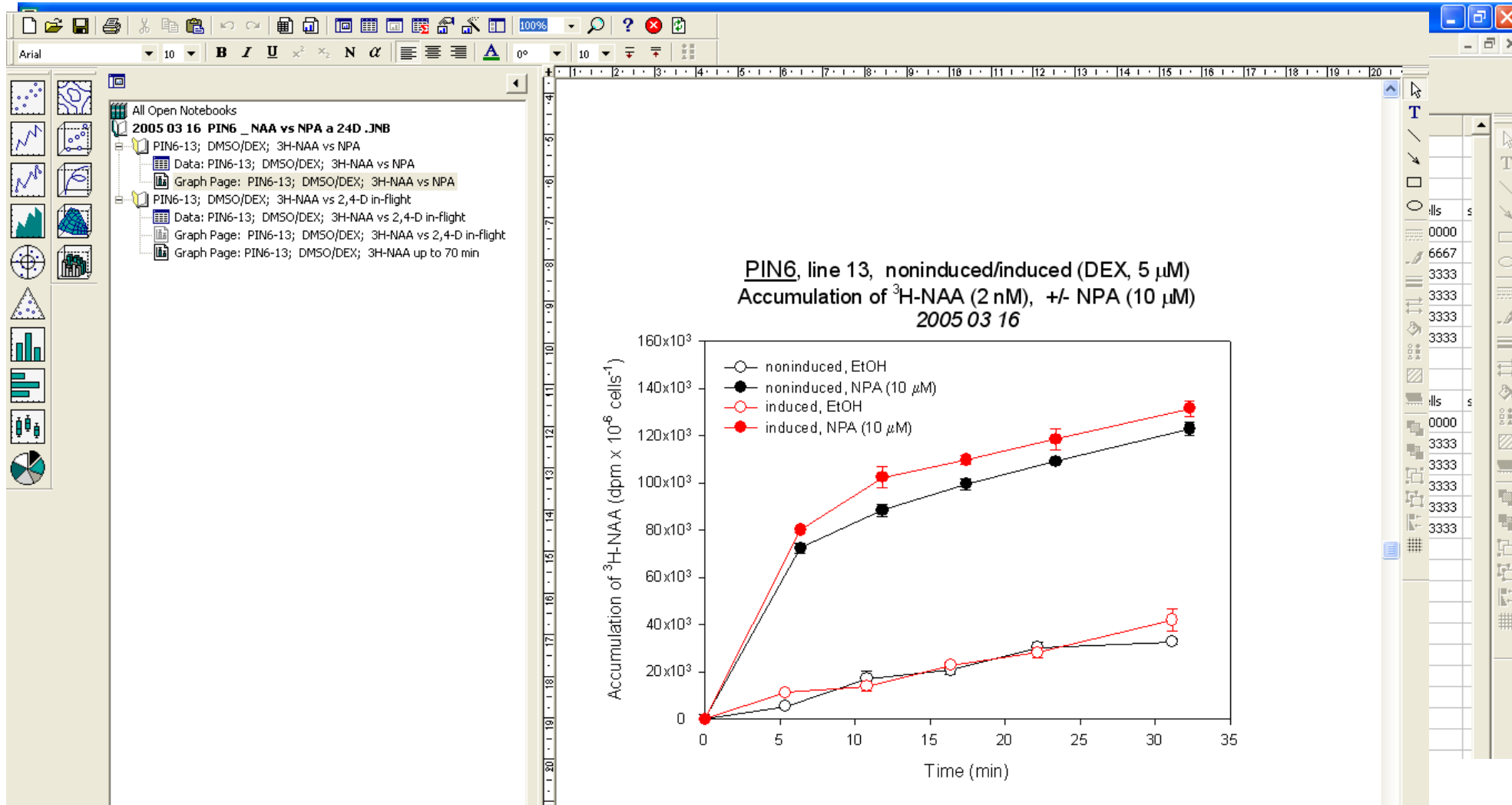


	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1														
2														
3														
4														
5														
6														
7														
8														
9														
10														
11														
12														
13														
14														
15														
16														
17														
18														
19														
20														
21														
22														
23														
24														
25														
26														
27														
28														
29														
30														
31														
32														
33														
34														



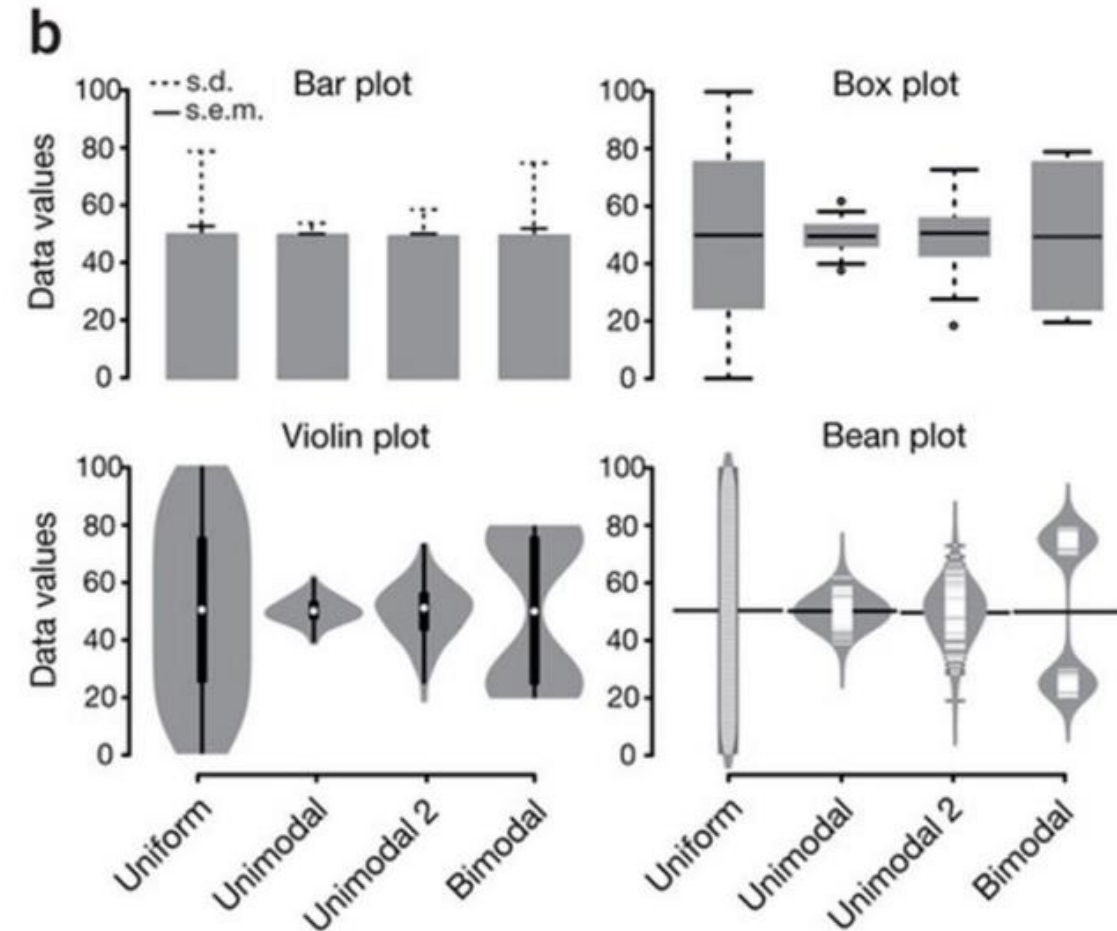
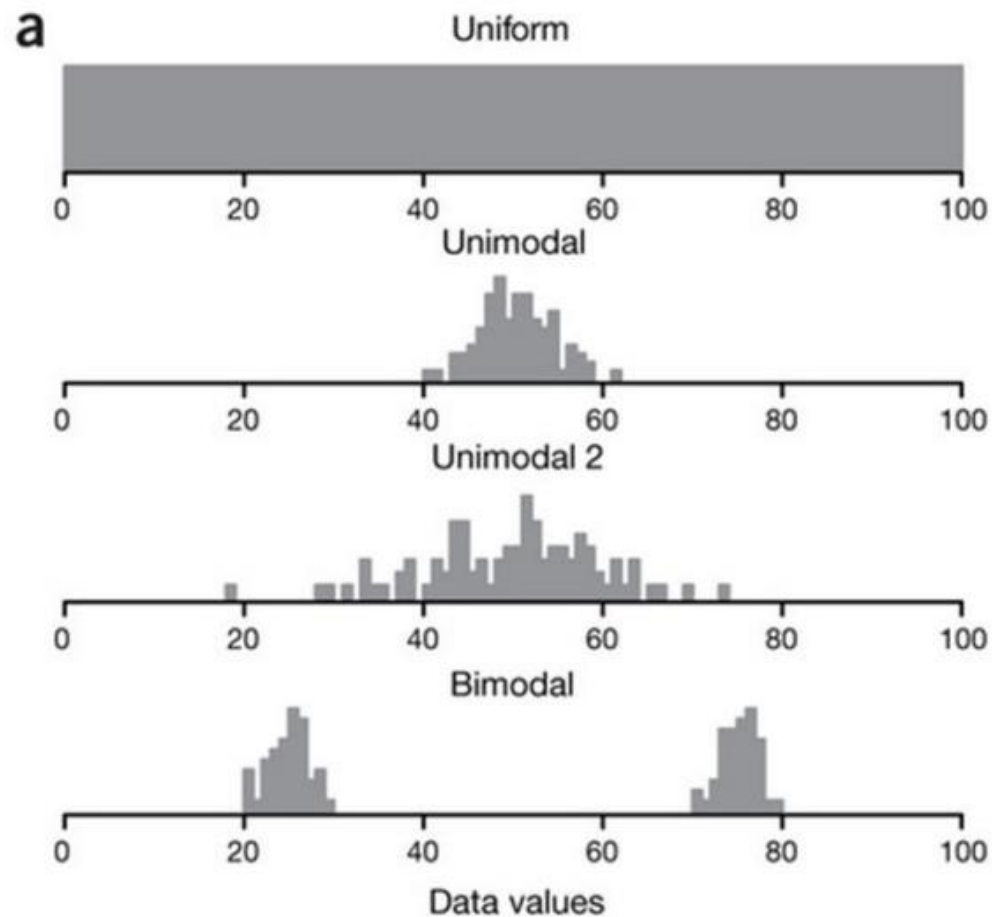
## 2.3 Spreadsheets, statistical software, graph editors

Keeping a complete log of the experiment, automation of numerical operations



## 2.3 Spreadsheets, statistical software, graph editors

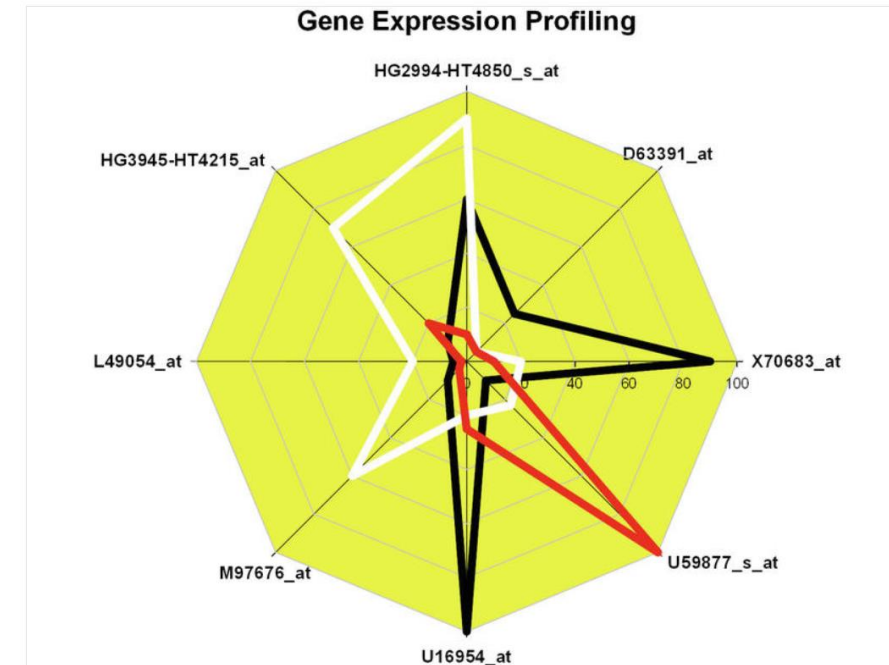
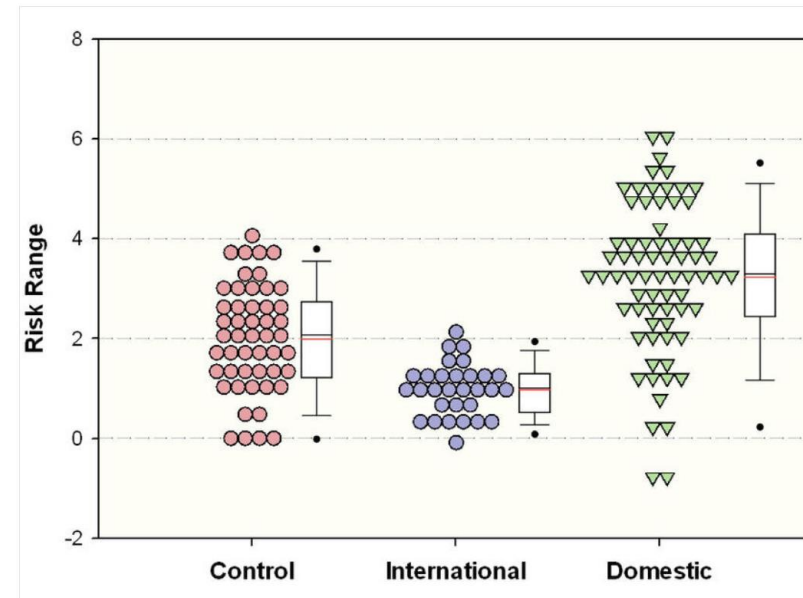
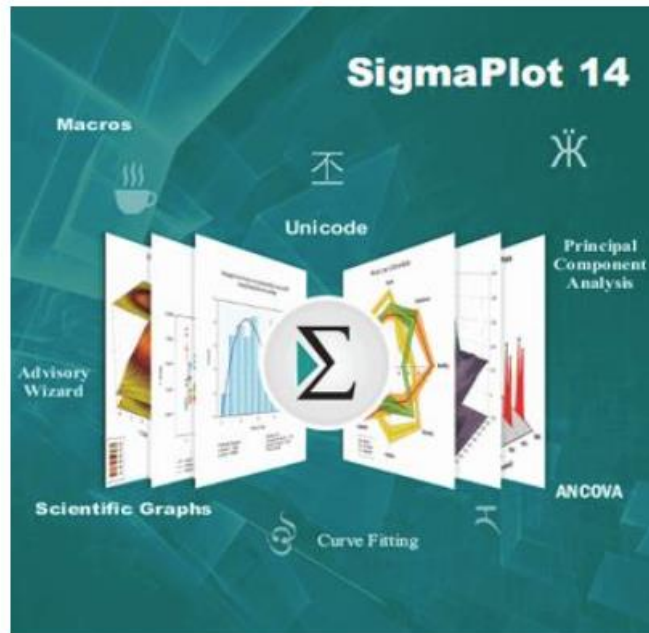
Bar plot versus box plot - what is better?



Spitzer M, Wildenhain J, Rappsilber J, Tyers M (2014) BoxPlotR: a web tool for generation of box plots. Nat Methods 11:121–122.

## 2.3 Spreadsheets, statistical software, graph editors

Creating graphs - line/scatter plots, bar charts, box plots, histograms

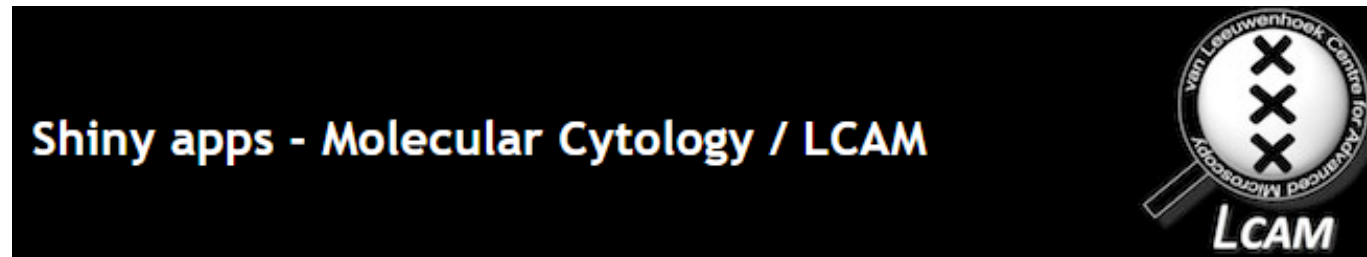


## 2.3 Spreadsheets, statistical software, graph editors

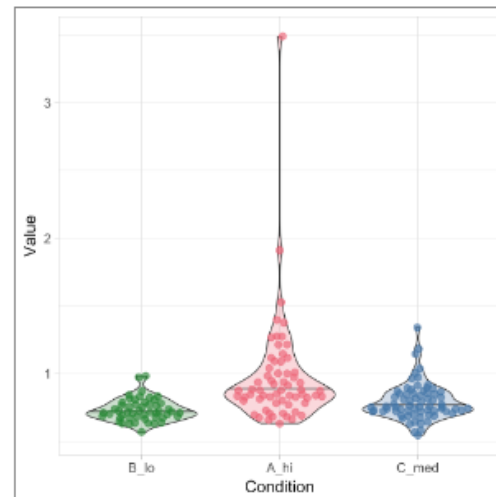
Creating graphs - line/scatter plots, bar charts, box plots, histograms

In biology, it is good to plot **all data** together with **their summaries and statistics**

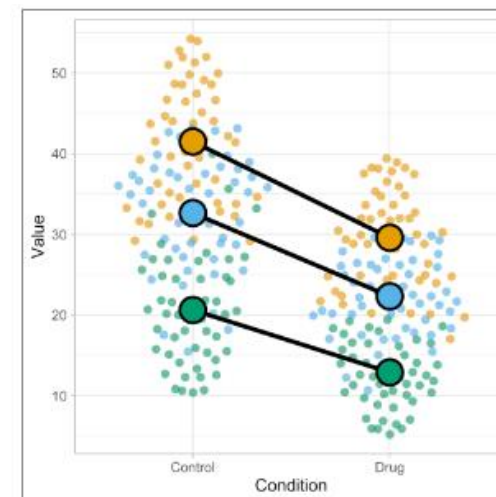
Shiny apps - R-based web tools for creating plots



PlotsOfData



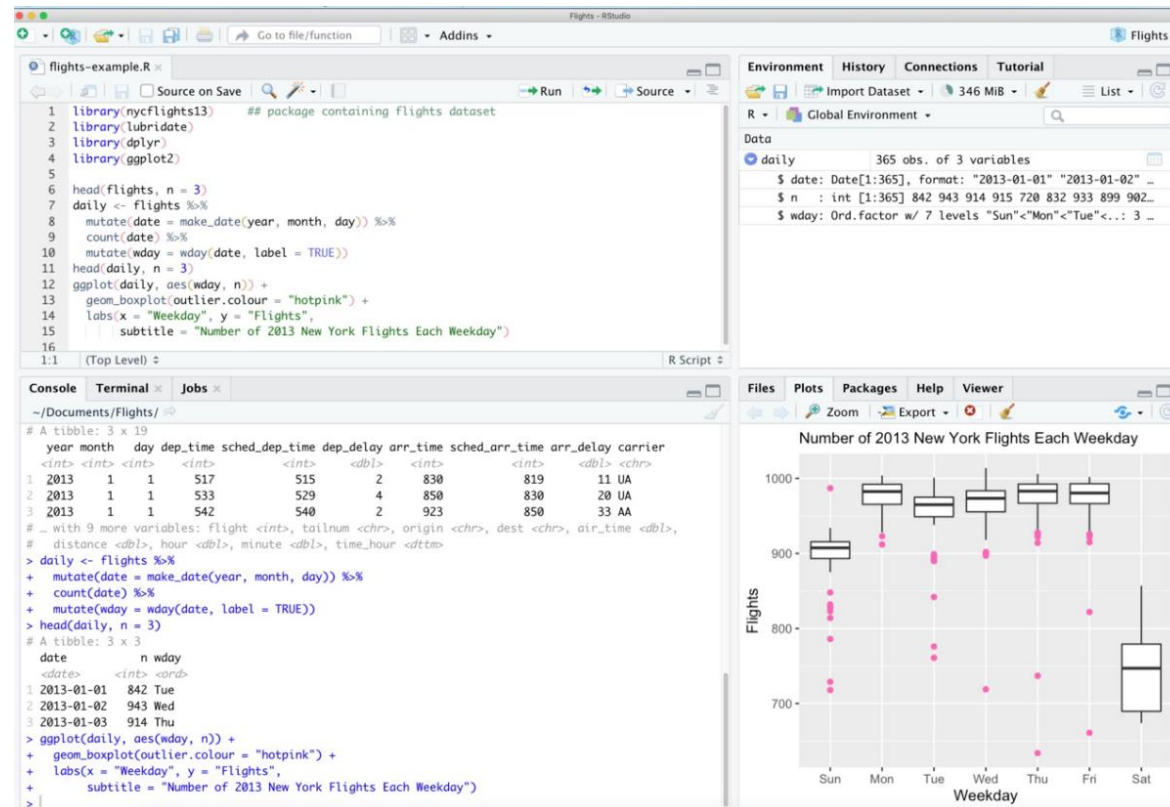
SuperPlotsOfData



## 2.3 Spreadsheets, statistical software, graph editors

Creating graphs - line/scatter plots, bar charts, box plots, histograms

R studio helps new users to start with R



[https://en.wikipedia.org/wiki/RStudio#/media/File:RStudio\\_IDE\\_screenshot.png](https://en.wikipedia.org/wiki/RStudio#/media/File:RStudio_IDE_screenshot.png)

# 2.3 Spreadsheets, statistical software, graph editors



## R packages for data visualization

# ggplot2



`ggplot2` is a R package dedicated to data visualization. It can greatly improve the quality and aesthetics of your graphics, and will make you much more efficient in creating them.

`ggplot2` allows to build almost any type of chart. The R graph

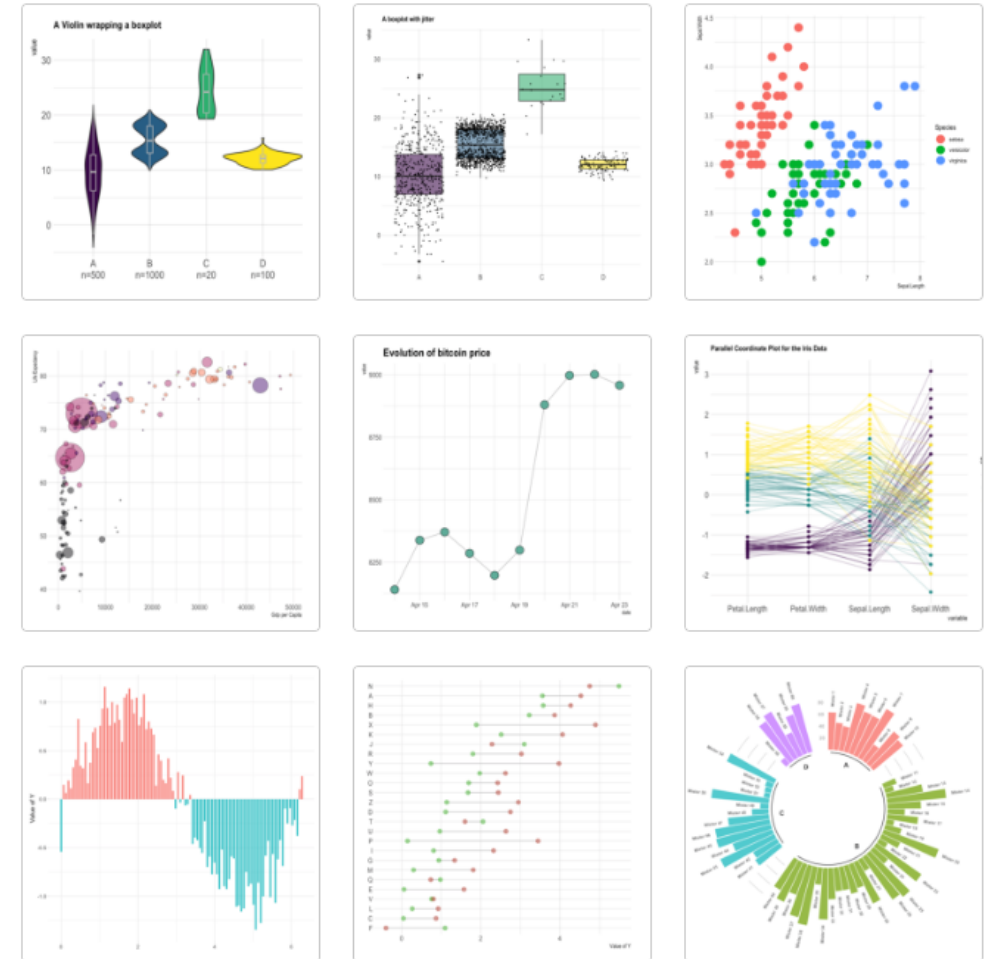
gallery focuses on it so almost every section there starts with `ggplot2` examples.

This page is dedicated to general `ggplot2` tips that you can apply to any chart, like customizing a title, adding annotation, or using faceting.

If you love `ggplot2`, you will love my [productive r workflow](#) project where I show how it interacts with Quarto, Git and Github! ❤️

### AN OVERVIEW OF `GGPLOT2` POSSIBILITIES

Each section of the gallery provides several examples implemented with `ggplot2`. Here is an overview of my favorite examples:

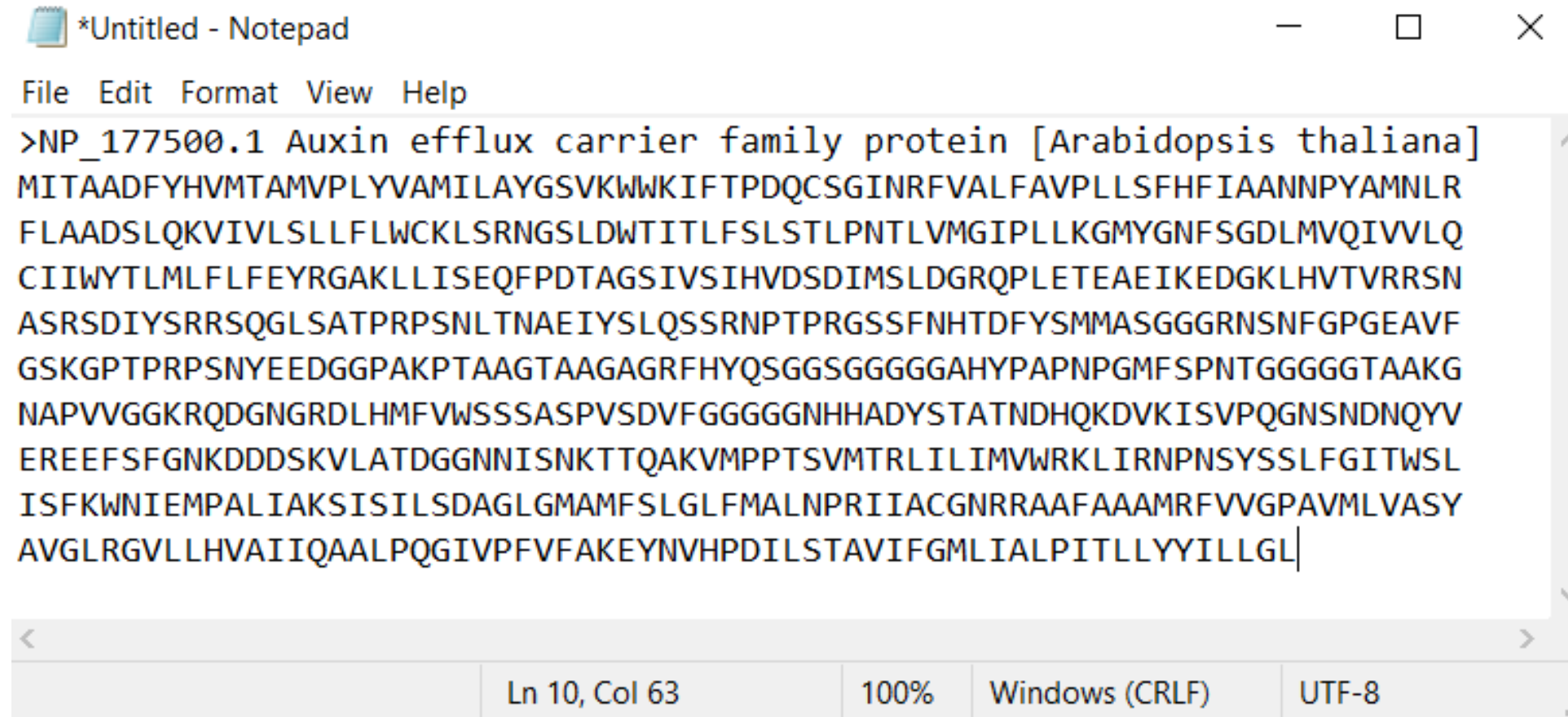


## 2.4 Processing of structural and sequence data

**Notepad** - a digital notebook, which is still an invaluable helper

It is important not to use formatting to write **nucleotide or amino acid sequences**

The most commonly used is the so-called **FASTA** format



```
*Untitled - Notepad
File Edit Format View Help
>NP_177500.1 Auxin efflux carrier family protein [Arabidopsis thaliana]
MITAADFYHVMTAMVPLYVAMILAYGSVKWVKIFTPQCSCGINRFVALFAVPLLSFHFIAANNPYAMNLR
FLAADSLQKVIVLSLLFLWCKLSRNGSLDWTITLFSLSLTPNTLVMGIPLLKGMYGNFSGDLMVQIVVLQ
CIIWYTLMLFLFEYRGAKLLISEQFPDTAGSIVSIHVSDIMSLDGRQPLETEAEIKEDGKLHVTVRRSN
ASRSDIYSRRSQGLSATPRPSNLTNAEIYSLQSSRNPTPRGSSFNHTDFYSMMASGGGRNSNFGPGEAVF
GSKGPTPRPSNYEEDGGPAKPTAAGTAAGAGRFHYQSGGSGGGGGGAHYPAPNPGMFSPNTGGGGGTAAKG
NAPVVGGKRQDGNRDLHMFVWSSSASPVSDFVGGGGGNHHADYSTATNDHQKDKVKSVPQGNNDNQYV
EREEFSFGNKDDDSKVLATDGGNNISNKTTQAKVMPPTSMTRLILIMVWRKLIRNPNSYSSLFGITWSL
ISFKWNIEMPALIAKSISILSDAGLGMAMFSLGLFMALNPRIIACGNRRAAFAAAMRFVVGPAVMLVASV
AVGLRGVLLHVAIIQAALPQGIVPFVFAKEYNVHPDILSTAVIFGMLIALPITLLYYILLGL|
Ln 10, Col 63 100% Windows (CRLF) UTF-8
```

# 2.4 Processing of structural and sequence data

## SnapGene

## Geneious



### Restriction Analysis

Get a restriction profile of your DNA based on cutting sites, frequency. Vector NTI Advance offers the most comprehensive REN analysis scheme than any other softwares.



### Digestion/Ligation Cloning

Create recombinants from individual fragments that have compatible ends, and modify the ends if they are not compatible. NEW in version 11, this is all done in a graphical manner.



### Gateway® Cloning

Using the intuitive Gateway Cloning Wizard, you can rapidly clone DNA fragments into multiple expression vectors simultaneously.

[Read more](#)



### TOPO® Cloning

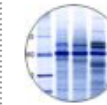
The integrated TOPO® Cloning Wizard greatly facilitates vector construction using the most comprehensive collection of vectors made by Invitrogen.

[Read more](#)



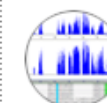
### Design Primers for:

- Amplification
- Sequencing
- Multiplex PC
- Hybridization



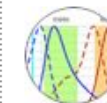
### Gel Simulation

- Run digested DNA on Agarose or Acrylamide gels as if you would in the lab
- Get prepared before loading the real stuff



### Multi-Sequence Alignment

- Fast, accurate alignment by ClustalW algorithm
- Alignment statistics, phylogenetic guide tree, repeat identification



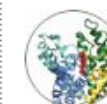
### Contig Assembly

- Fast, accurate assembly by the CAP3 algorithm
- Flexible ends trimming and vector contamination trimming
- Chromatogram data analysis and editing; consensus creation using Quality Values



### Use BioAnnotator to analyze proteins for:

- PFAM, ProSite, BLOCKS, Proteolytic Cleavage, Spidey, Sim4



### View Protein Structure in 3-D

- View protein structures in stereo mode
- Freedom to configure how you

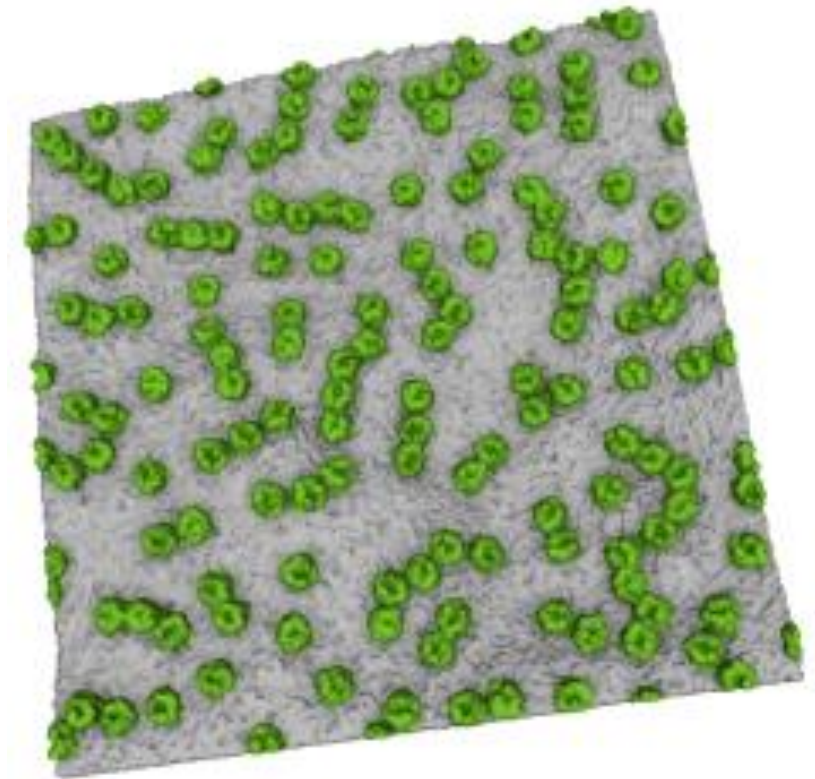
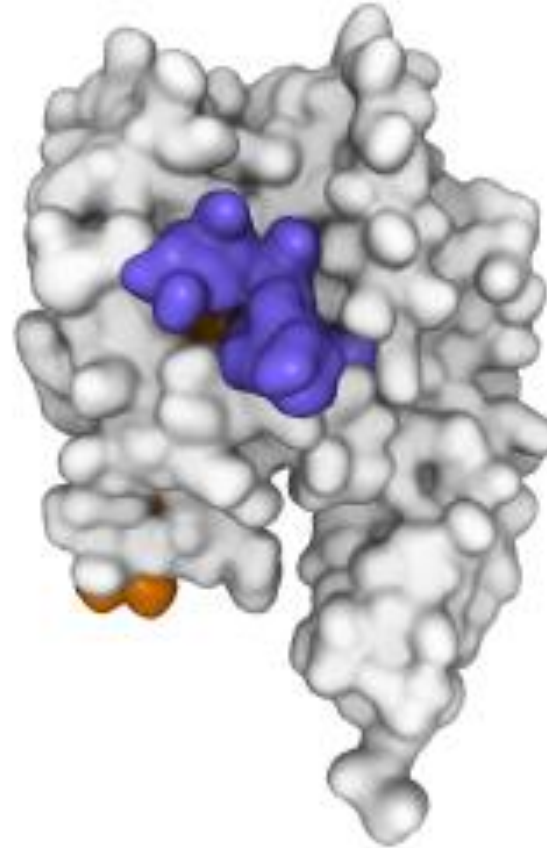


## 2.4 Processing of structural and sequence data

3D structures of macromolecules – Molstar viewer

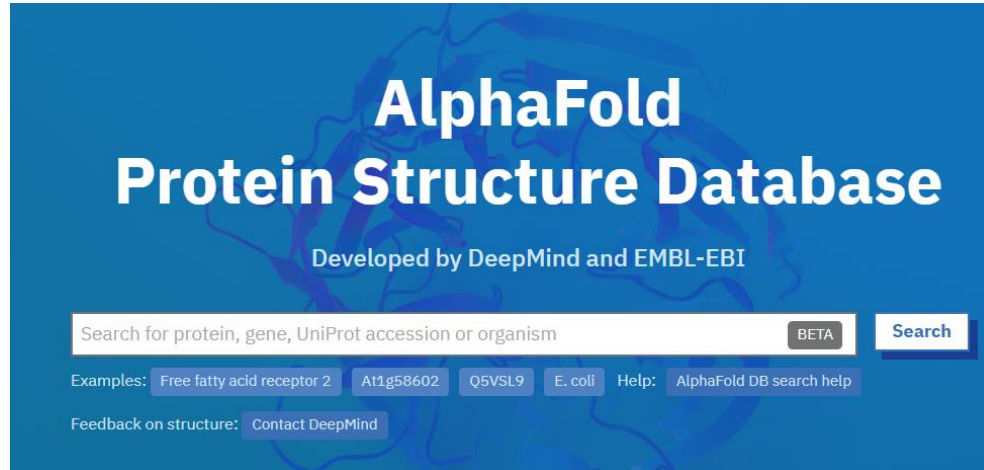


Mol\* (*'molstar'*) is a modern web-based open-source toolkit for visualisation and analysis of large-scale molecular data



# 2.4 Processing of structural and sequence data

## 3D structures of macromolecules - AI-based solution for „*in silico*“ translation



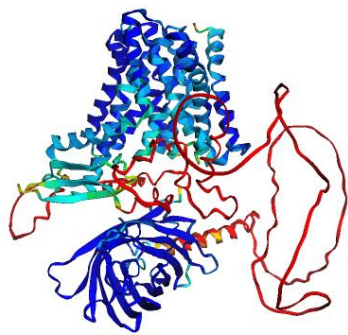
**AlphaFold**  
Protein Structure Database

Developed by DeepMind and EMBL-EBI

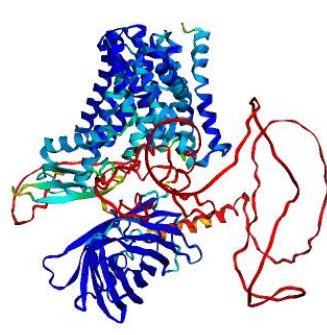
Search for protein, gene, UniProt accession or organism  BETA Search

Examples: [Free fatty acid receptor 2](#) [At1g58602](#) [Q5VSL9](#) [E. coli](#) [Help: AlphaFold DB search help](#)

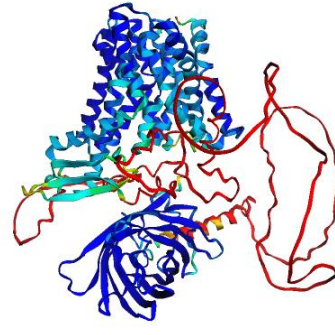
Feedback on structure: [Contact DeepMind](#)



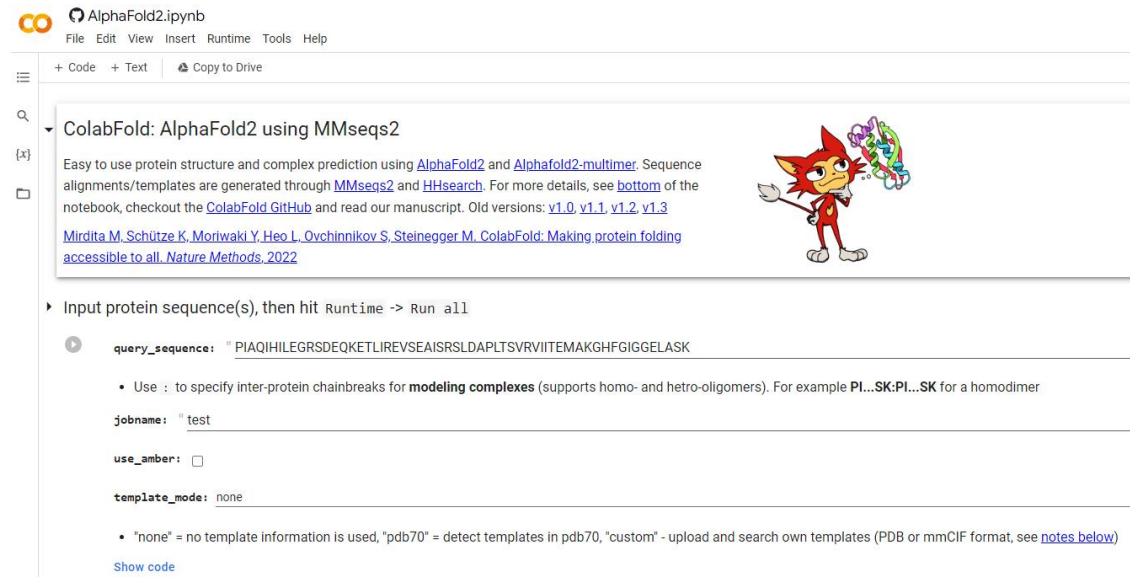
NtPIN2T



NtPIN3bT



NtPIN11T



AlphaFold2.ipynb  
File Edit View Insert Runtime Tools Help

+ Code + Text Copy to Drive

ColabFold: AlphaFold2 using MMseqs2

Easy to use protein structure and complex prediction using [AlphaFold2](#) and [AlphaFold2-multimer](#). Sequence alignments/templates are generated through [MMseqs2](#) and [HHsearch](#). For more details, see [bottom](#) of the notebook, checkout the [ColabFold GitHub](#) and read our manuscript. Old versions: [v1.0](#), [v1.1](#), [v1.2](#), [v1.3](#)  
[Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. ColabFold: Making protein folding accessible to all. Nature Methods. 2022](#)

Input protein sequence(s), then hit Runtime -> Run all

query\_sequence: "PIAQIHILEGRSDEQKETLIREVSEAIRSLDAPLTSVRVIITEMAKGHFGIGGELASK"

- Use : to specify inter-protein chainbreaks for **modeling complexes** (supports homo- and hetero-oligomers). For example **PI...SK:PI...SK** for a homodimer

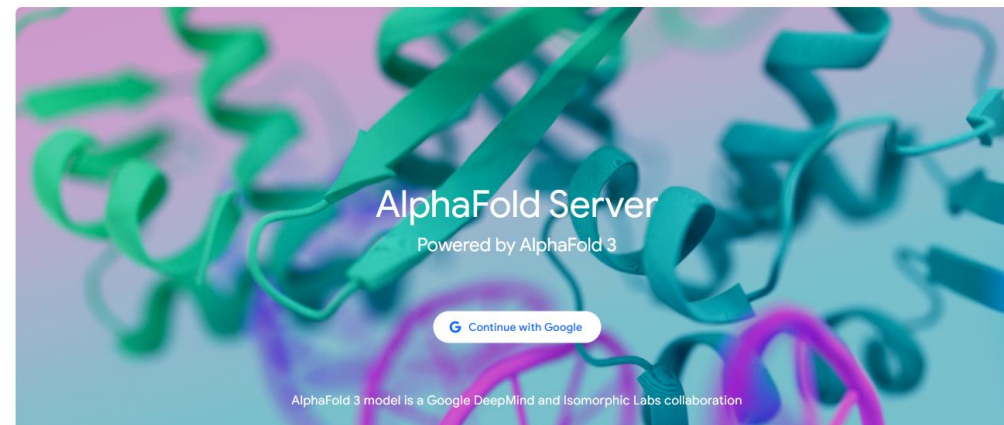
jobname: "test"

use\_amber:

template\_mode: none

- "none" = no template information is used, "pdb70" = detect templates in pdb70, "custom" - upload and search own templates (PDB or mmCIF format, see [notes below](#))

Show code



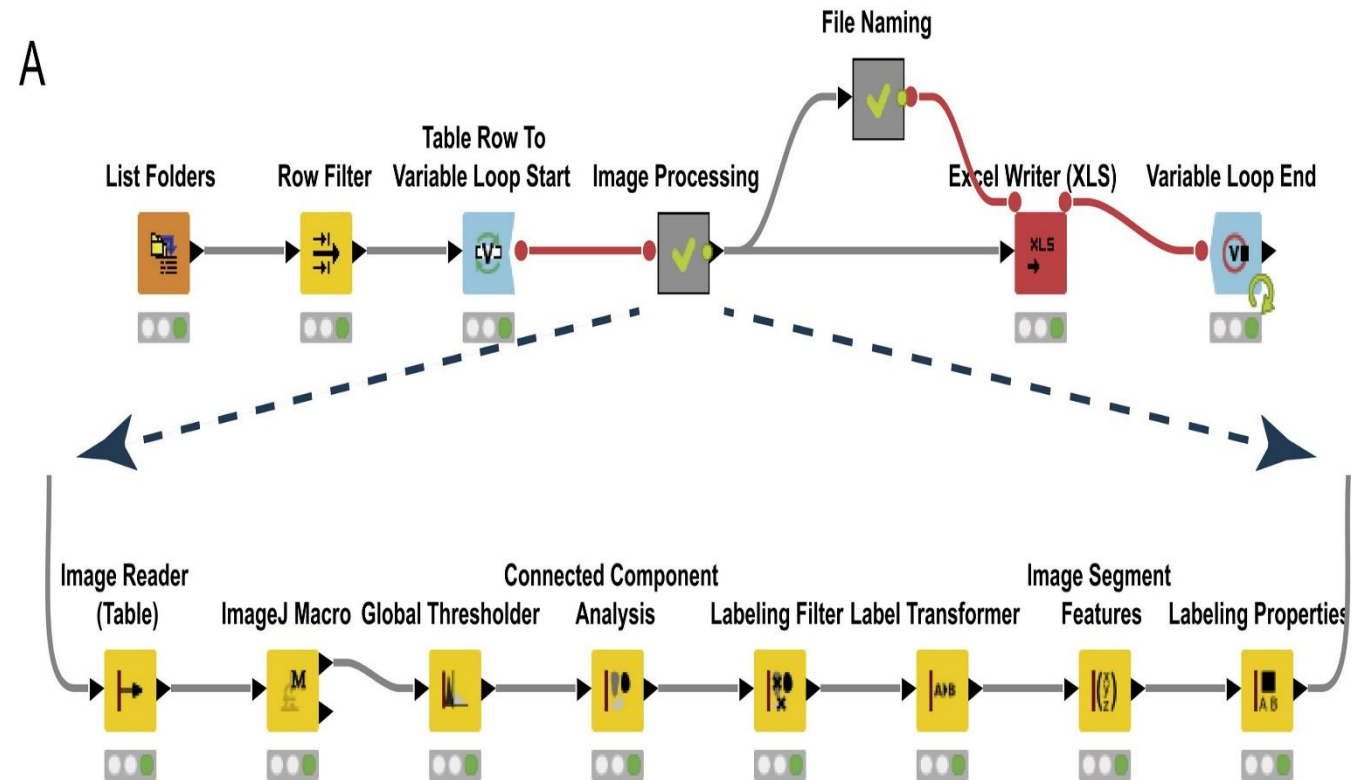
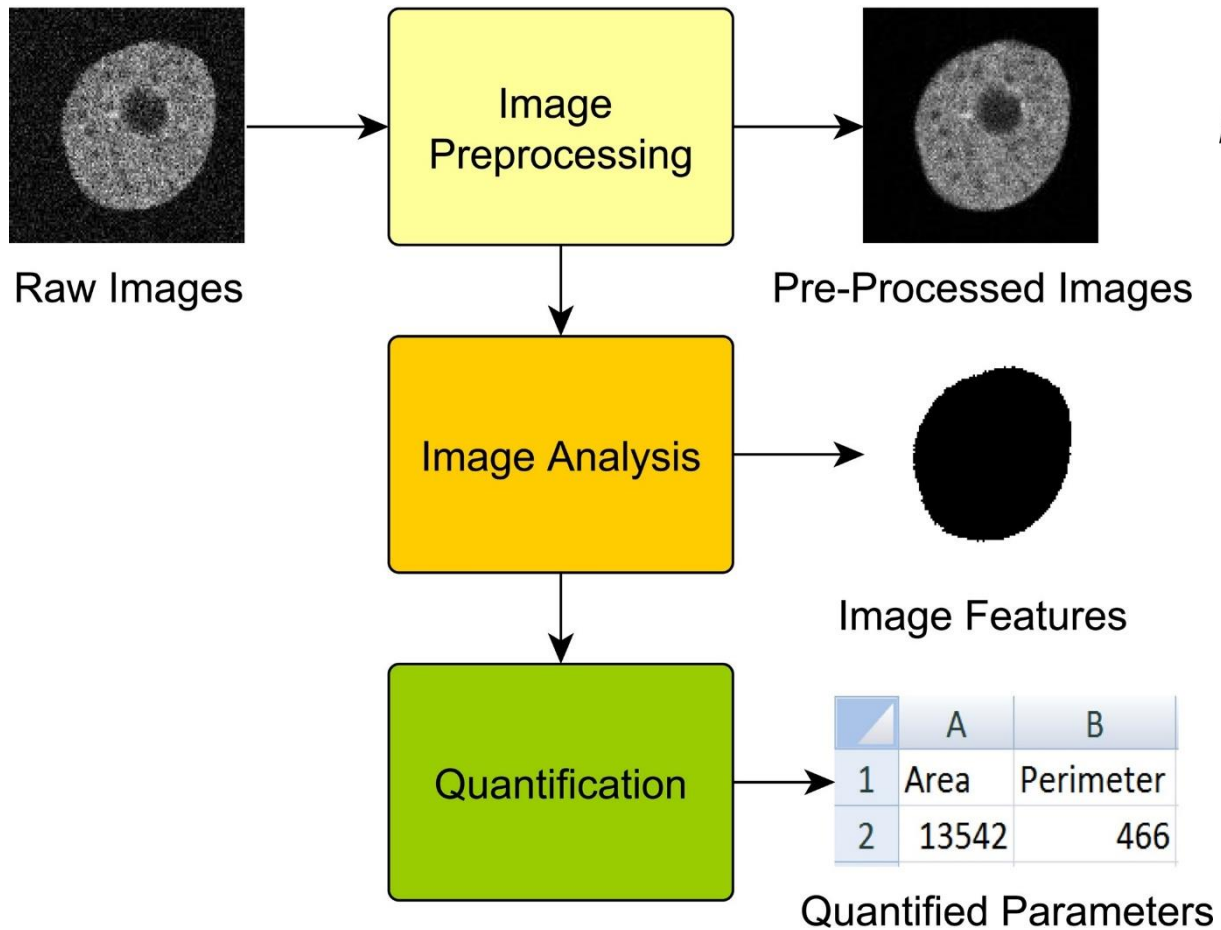
AlphaFold Server  
Powered by AlphaFold 3

[Continue with Google](#)

AlphaFold 3 model is a Google DeepMind and Isomorphic Labs collaboration

## 2.5 Image analysis, graphical software, presentation software

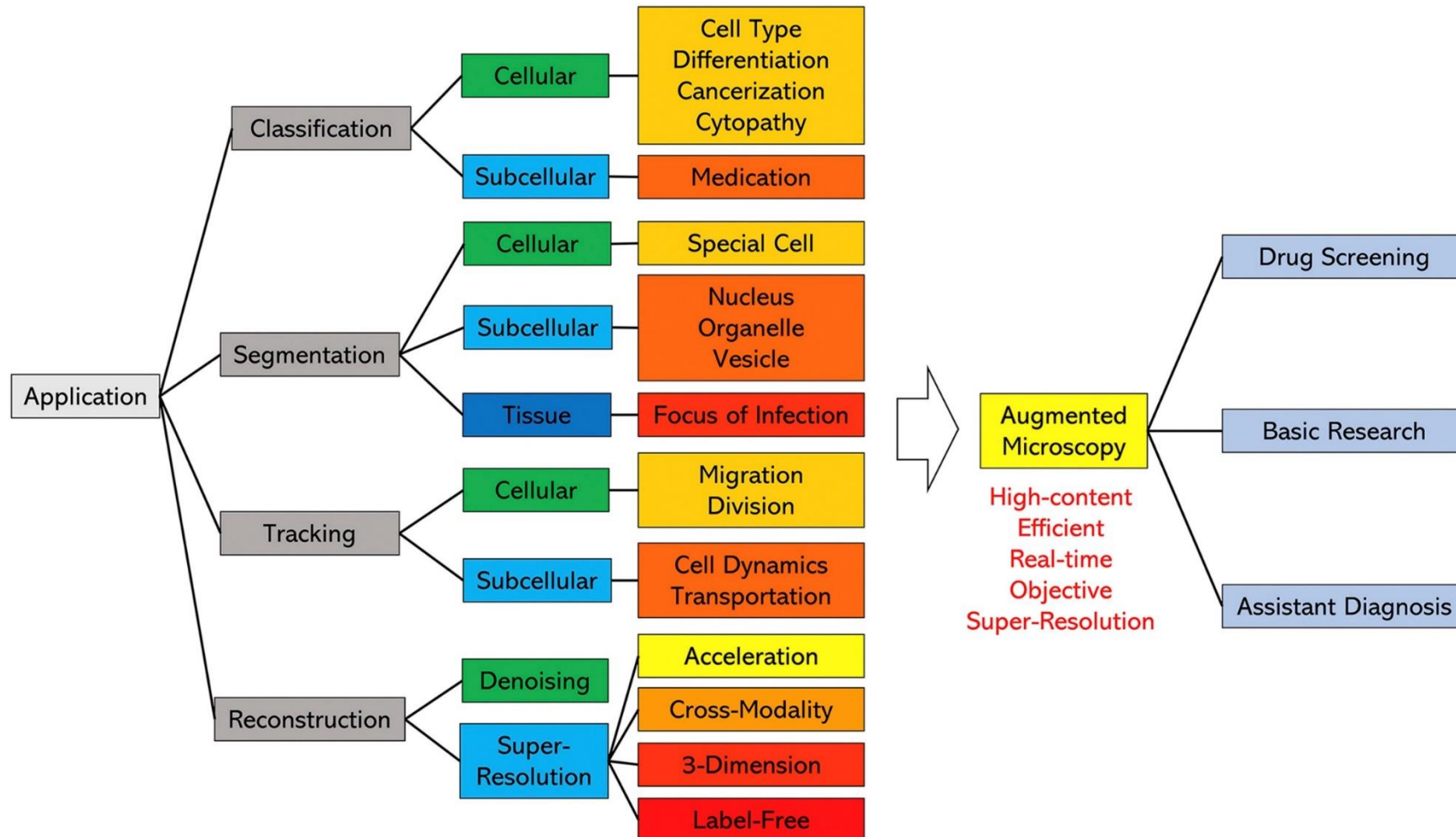
The task of **image analysis** is to **evaluate features that are often not apparent at first glance**



Wollmann T et al., J Biotechnol 261:70-75, 2017

# 2.5 Image analysis, graphical software, presentation software

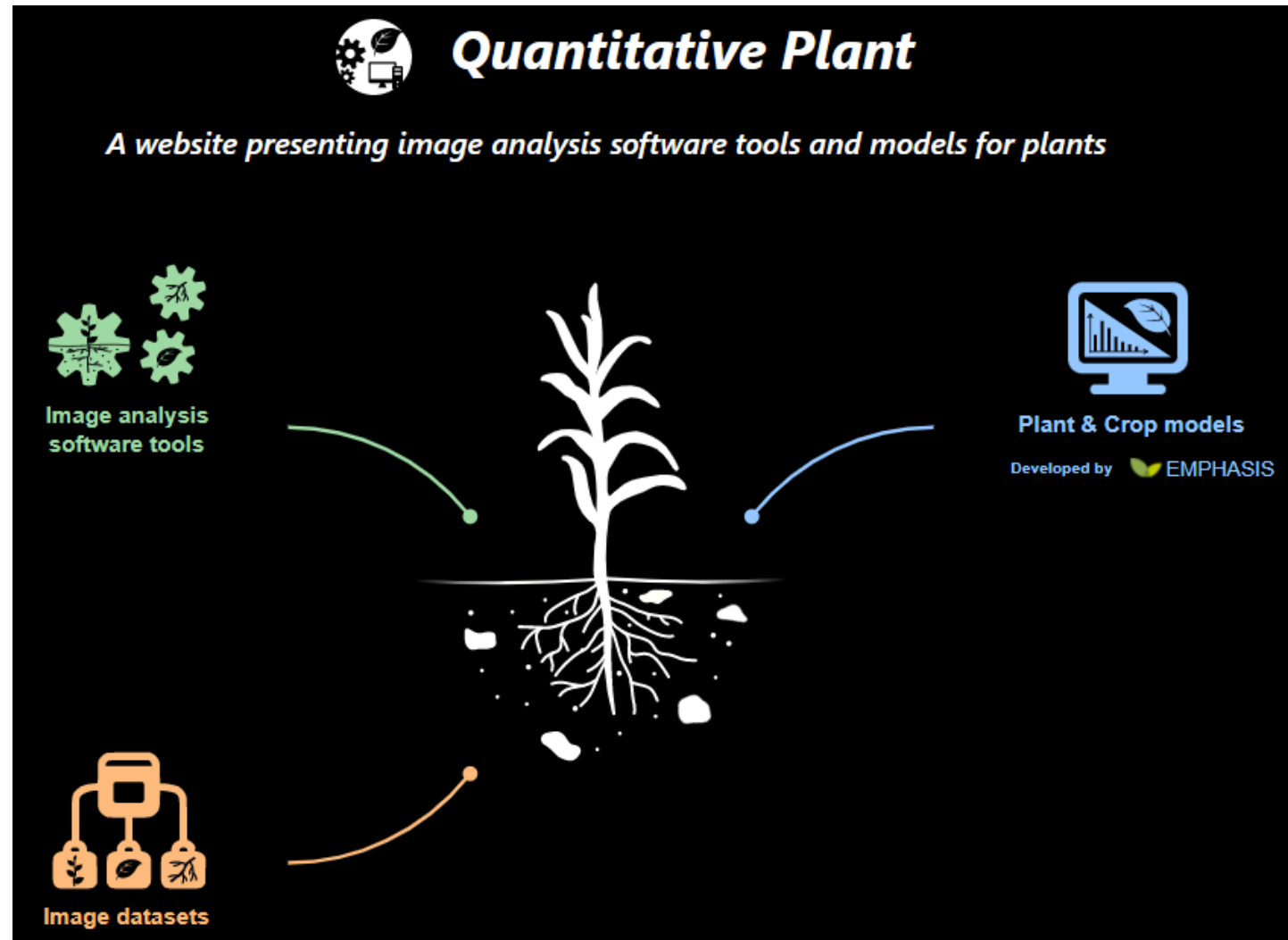
## Artificial intelligence in image analysis



Liu et al., Comput Biol Med 134:104523, 2021

## 2.5 Image analysis, graphical software, presentation software

Summary of image analysis tools for plant research



# 2.5 Image analysis, graphical software, presentation software

## Summary of image analysis tools for plant research

The screenshot displays the Quantitative Plant website interface. At the top, there is a navigation bar with links for 'Image software tools', 'Image datasets', 'Plant and Crop models', 'API', 'Submit', 'About', 'Cite us', and a search bar. Below the navigation bar, there are several filter dropdown menus: 'root-system', 'Measurements', 'Operating system', 'License', and 'Automation level'. The main content area is a grid of tool cards. Each card features a representative image of the tool's output, the tool's name, its category, supported operating systems (Windows, macOS, Linux), a release date, and a list of measurable parameters.

Tool Name	Category	Release Date	Measurable Parameters
ARIA	root-system	2014-10	length, convex-hull, shape, count
DART	root-system	2011-06	length, topology, insertion
DIRT	root-system	2014-09	length, shape, diameter
DynamicRoots	root-system	2015-06	length, topology, growth, count, gravitropism
ER			
GLO-Roots			

# 2.5 Image analysis, graphical software, presentation software

## ImageJ - open source image analysis



Acquisition

Algorithm, Filters, ...

Processing

Automation

Analysis

Image J

Image J ecosystem



<https://cif.unil.ch/cif-wiki/several-image-processing-workflow-examples/>



# 2.5 Image analysis, graphical software, presentation software

[ImageJ \(NIH\)](#)

**ImageJ**  
Image Processing and Analysis in Java

[Fiji](#)

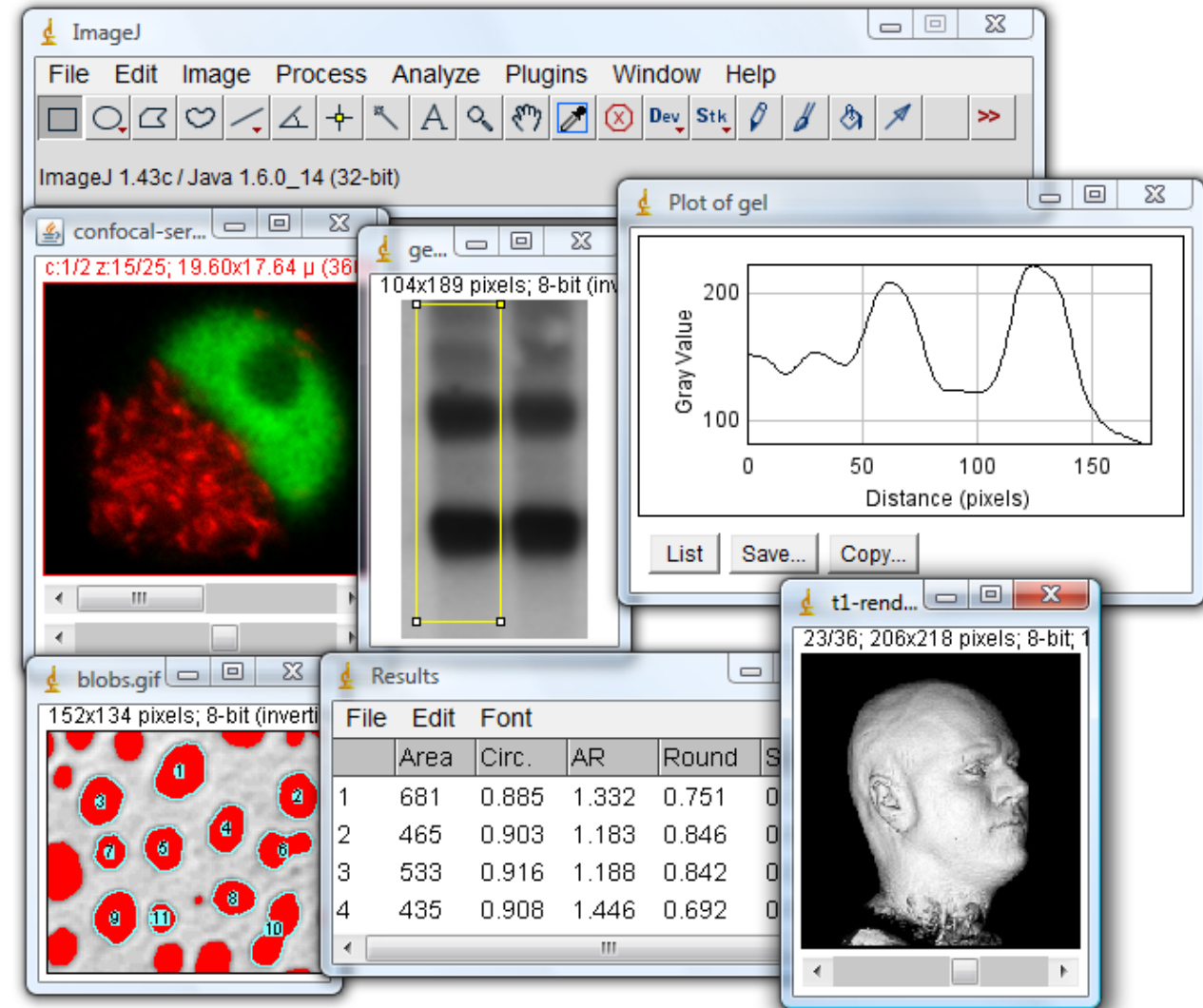
Imaging tutorials

[Scientific](#)

[Imaging](#)

[Tutorials](#)

[\(imagej.net\)](http://imagej.net)

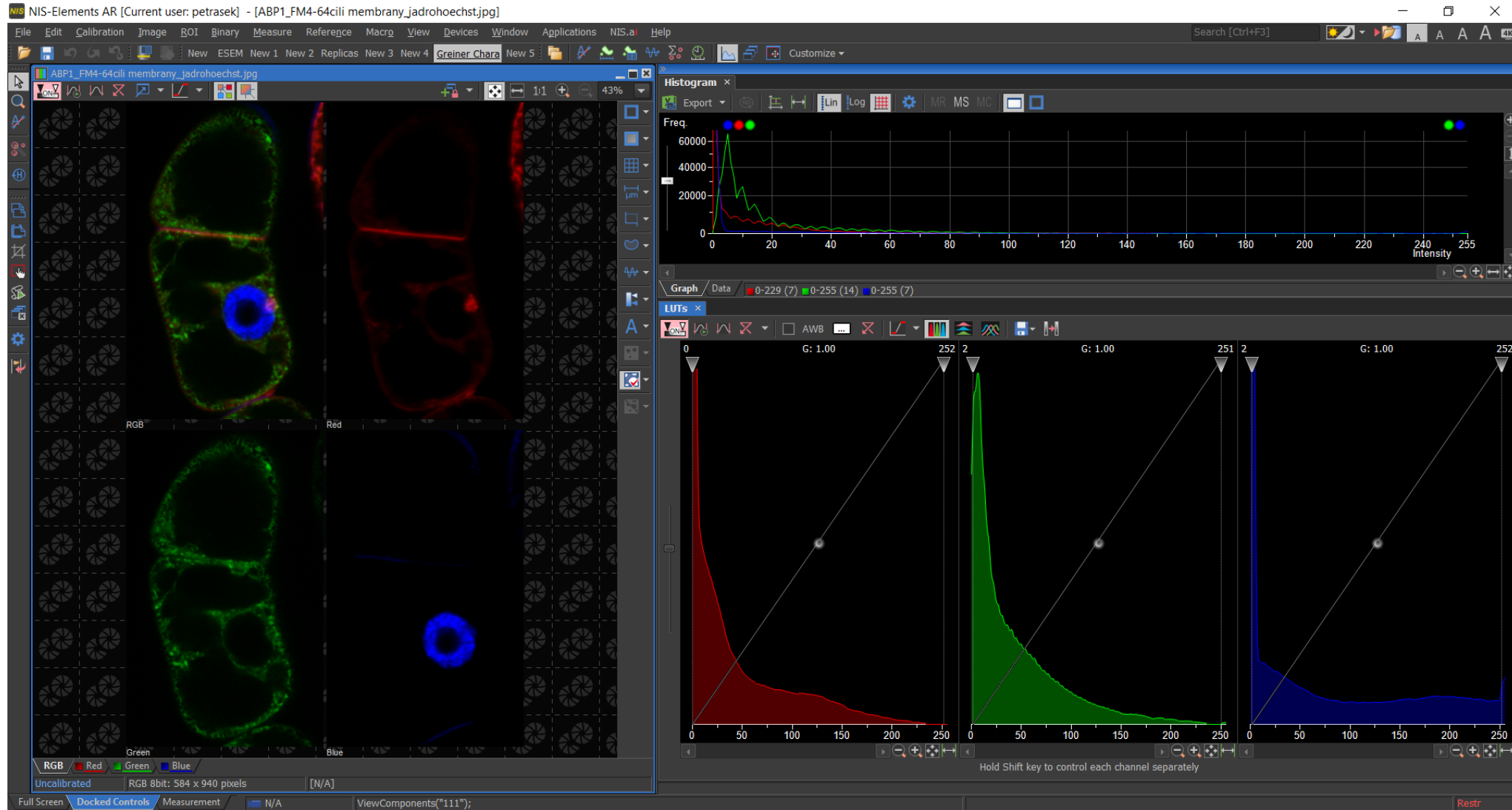


<http://rsb.info.nih.gov/ij/docs/concepts.html>



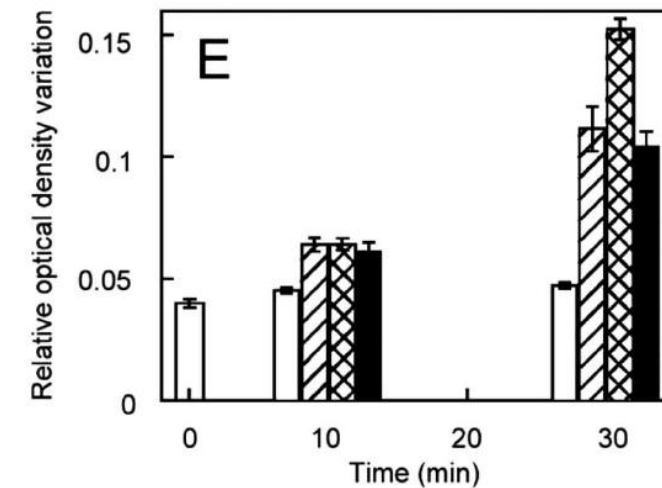
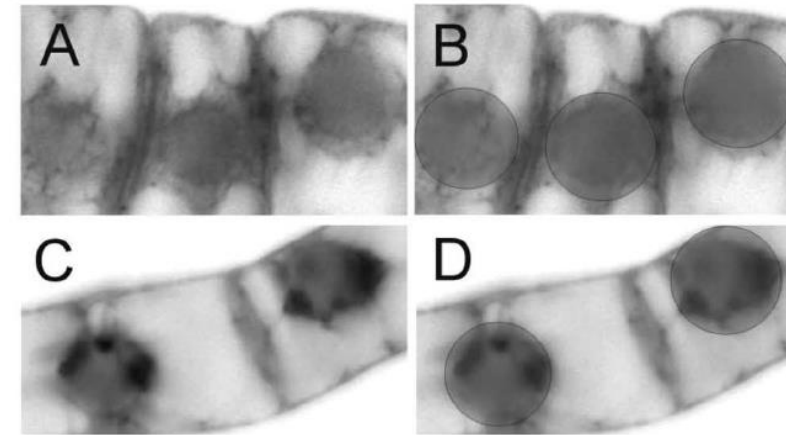
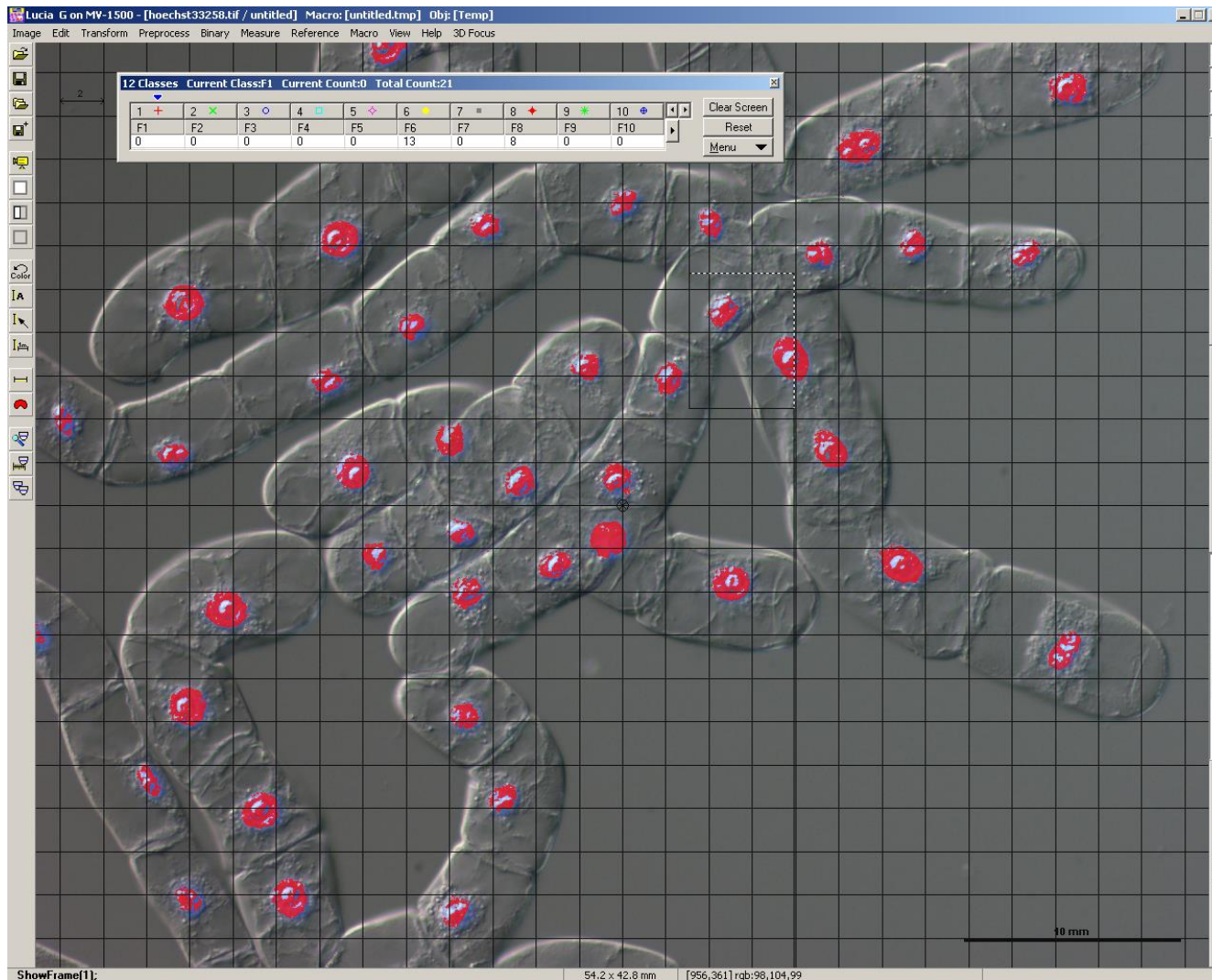
# 2.5 Image analysis, graphical software, presentation software

NIS Elements – good example of commercially available, comprehensive image analysis tool



# 2.5 Image analysis, graphical software, presentation software

NIS Elements – good example of commercially available, comprehensive image analysis tool



# 2.5 Image analysis, graphical software, presentation software

Graphical software for grabbing and processing of micro- and macroscopical images

Leica



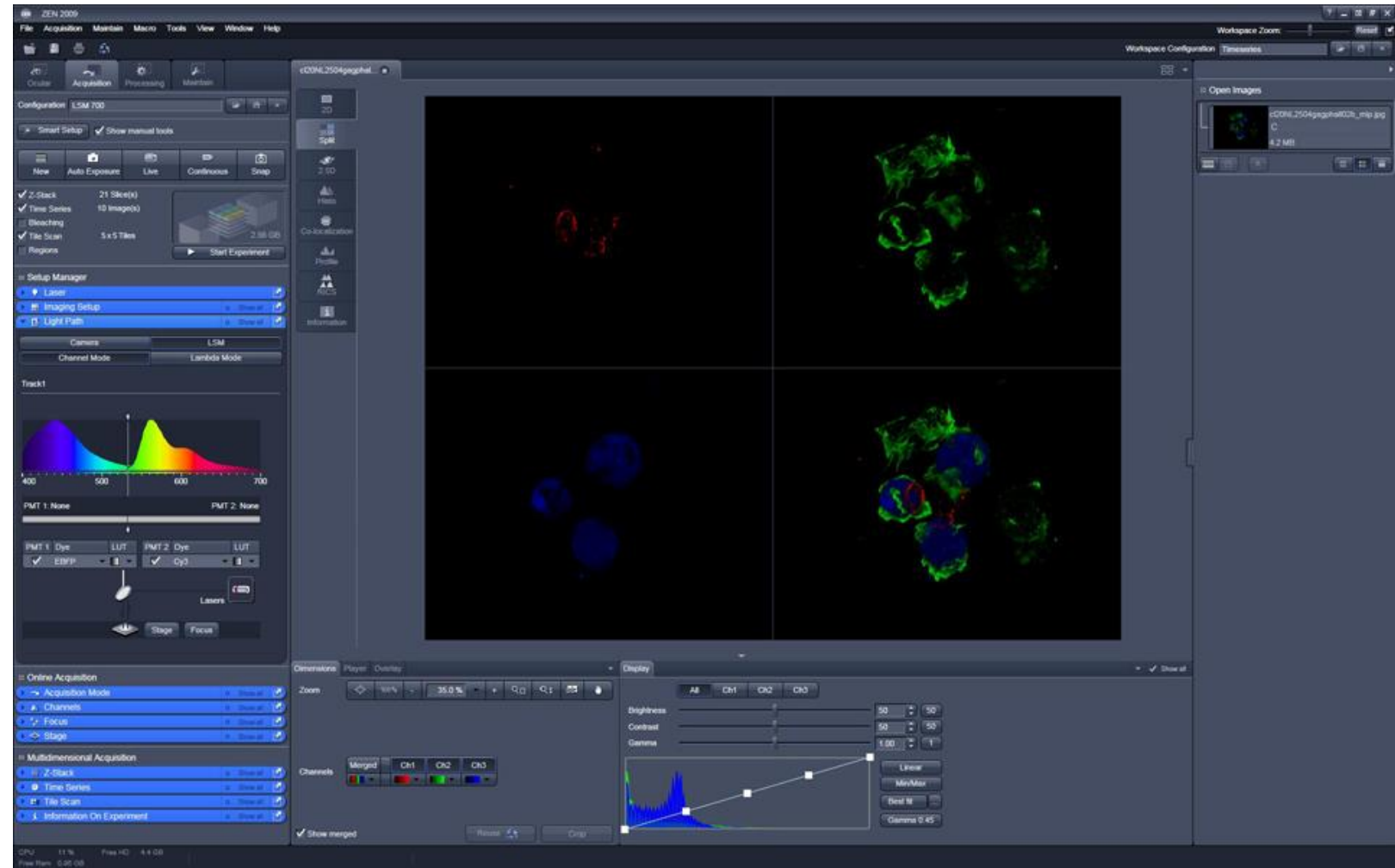
Zeiss



Nikon



Olympus



## 2.5 Image analysis, graphical software, presentation software

### Colour depth in biological imaging

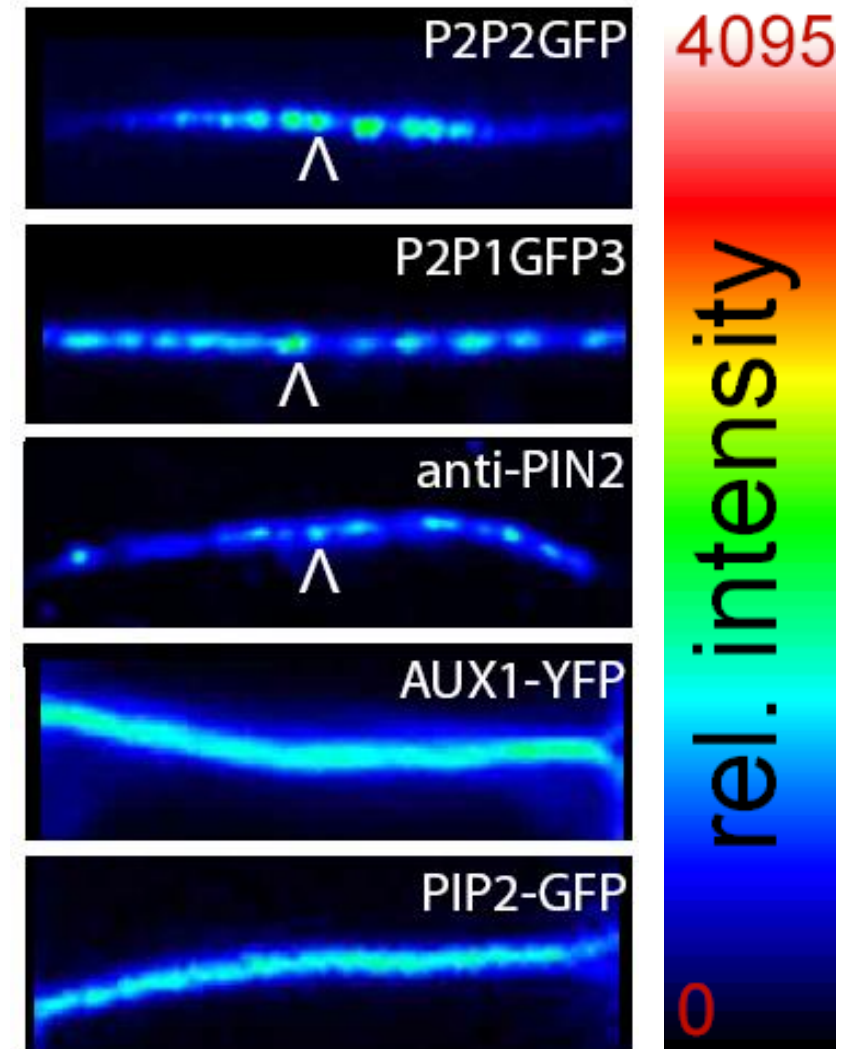
8 bit - 256 levels of grey ( $2^8$ )

12 bit - 4096 levels of grey ( $2^{12}$ )

16 bit - 65536 levels of grey ( $2^{16}$ )

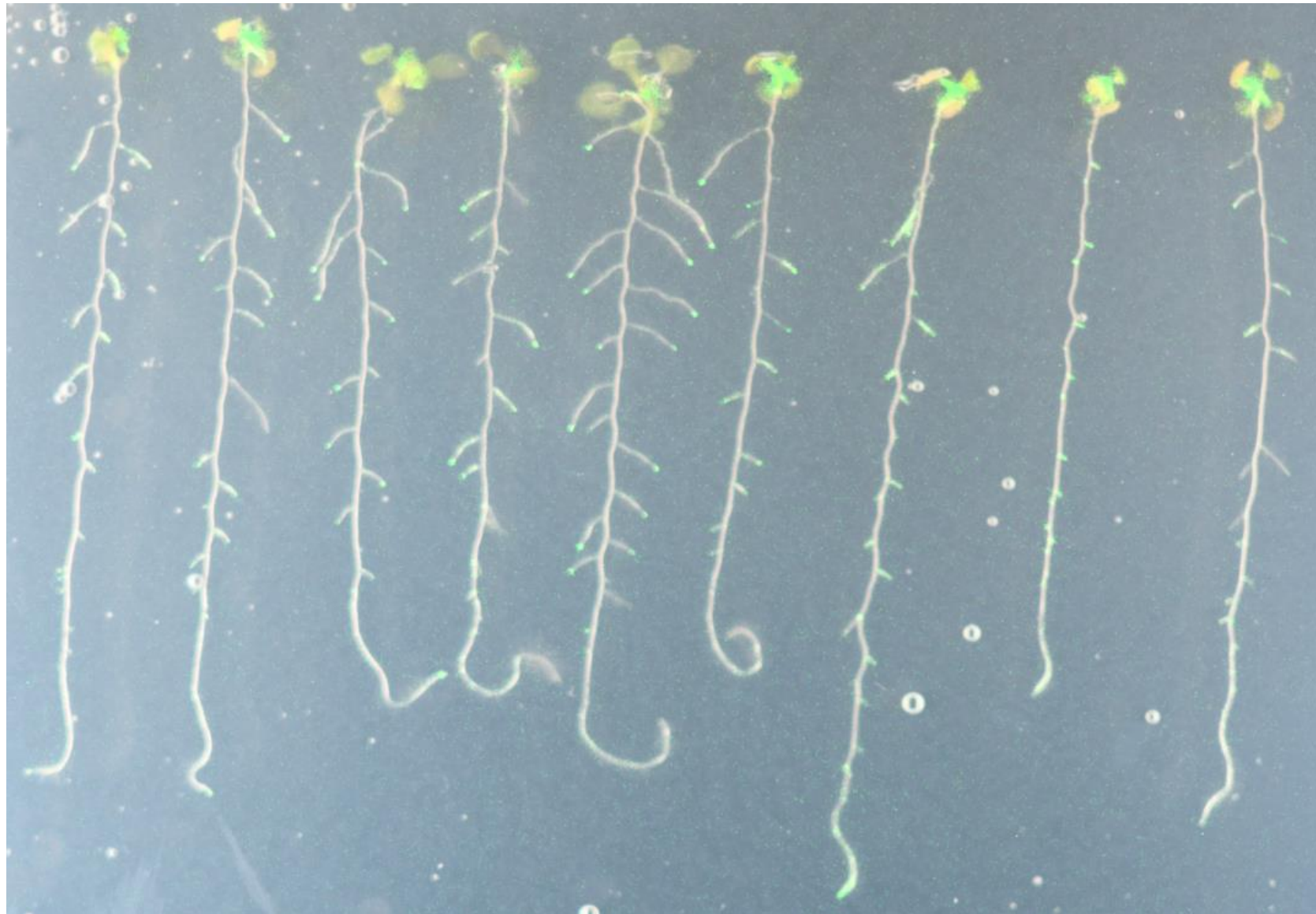
24 bit - 16 777 216 levels of grey ( $2^{24}$ )

A suitable **LUT**, i.e. **Look Up Table**,  
is crucial for the correct presentation of the image



Kleine-Vehn, Mol Syst Biol 7:540, 2011

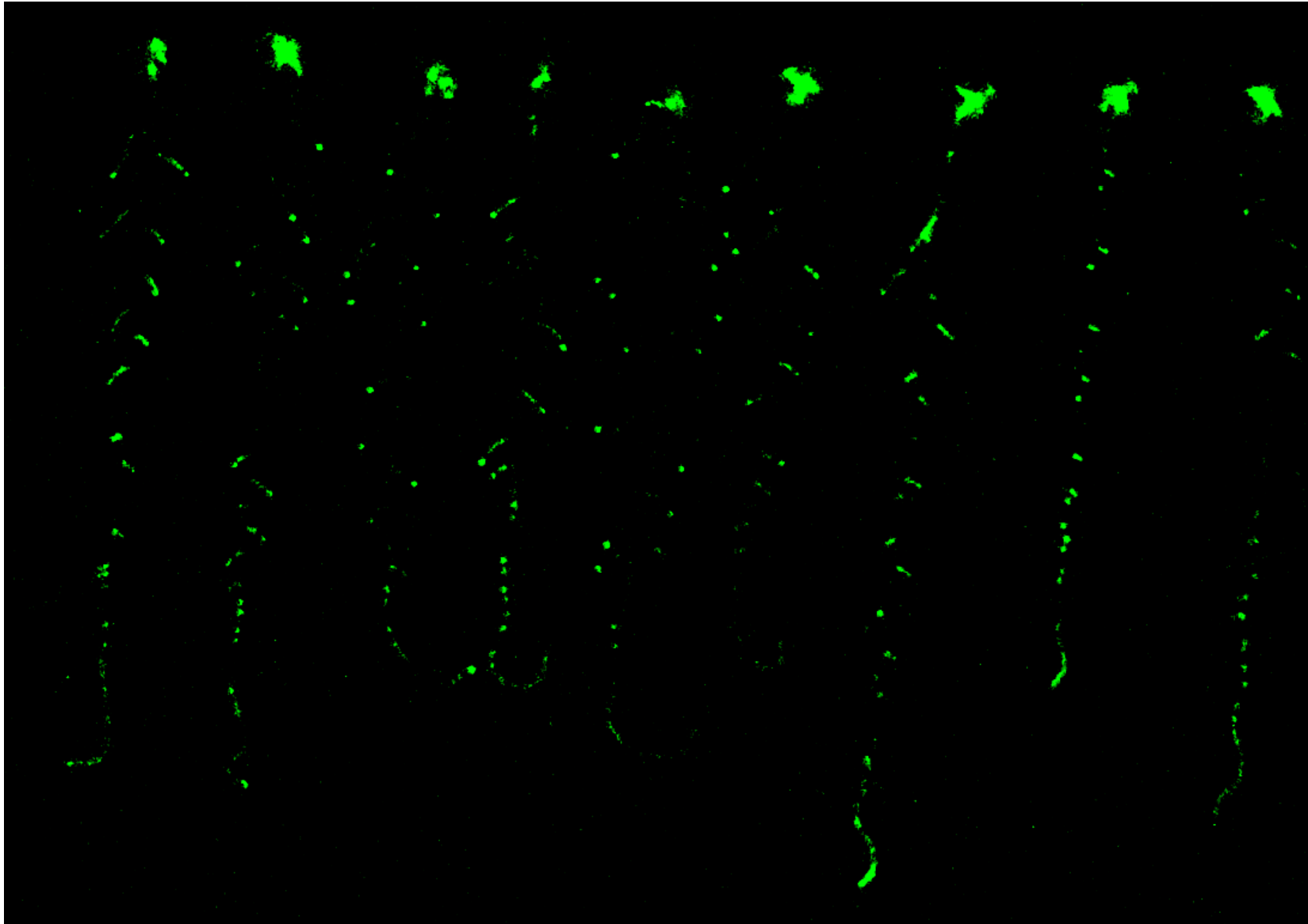
## 2.5 Image analysis, graphical software, presentation software



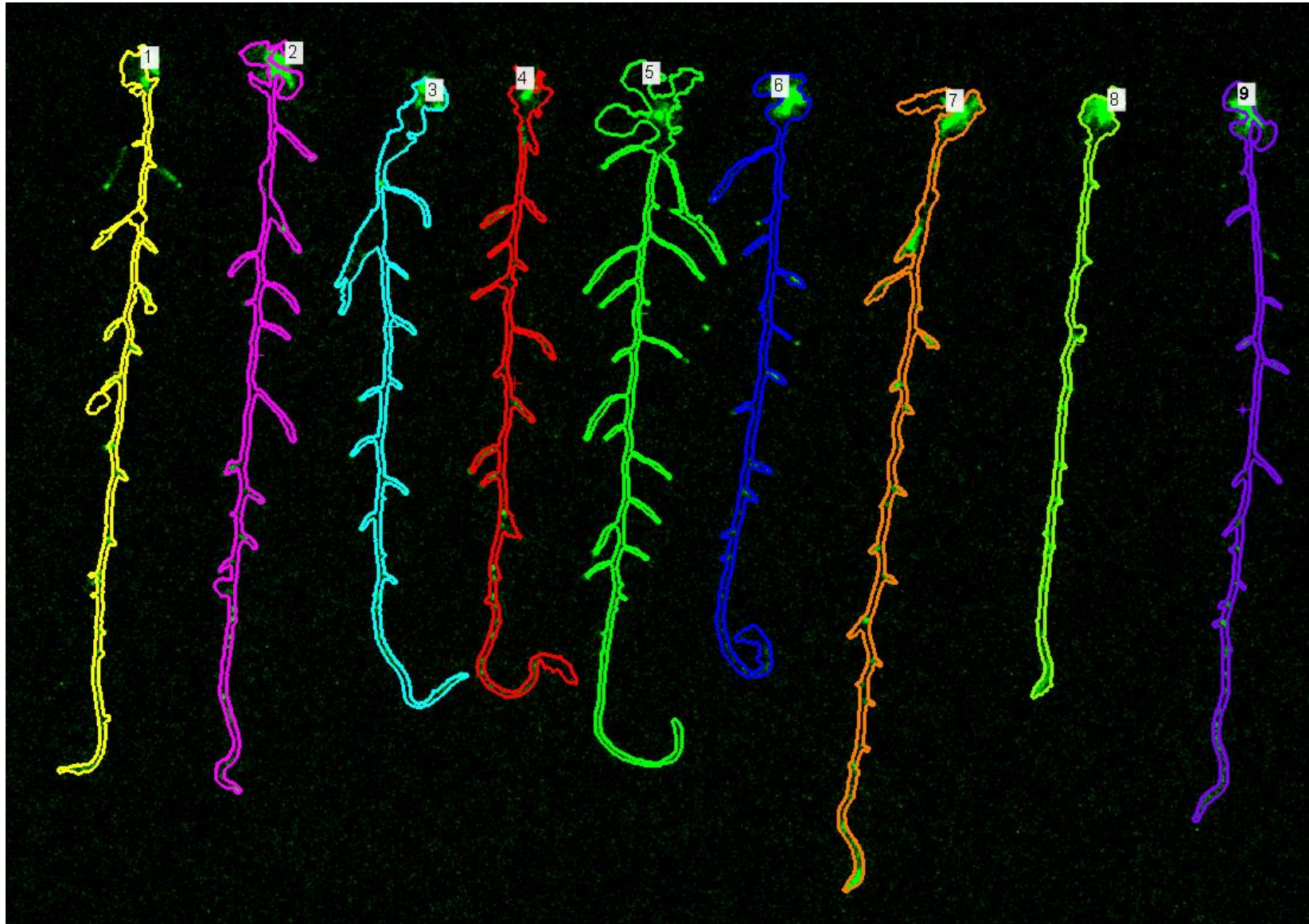
## 2.5 Image analysis, graphical software, presentation software



## 2.5 Image analysis, graphical software, presentation software



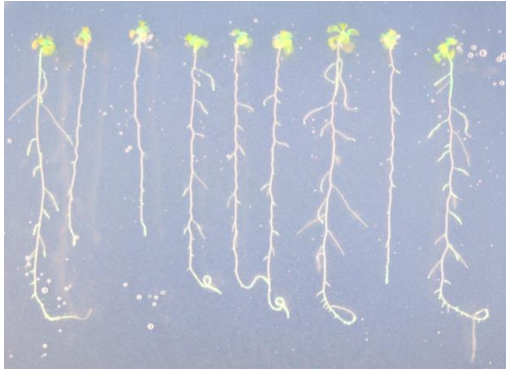
## 2.5 Image analysis, graphical software, presentation software





# 2.5 Image analysis, graphical software, presentation software

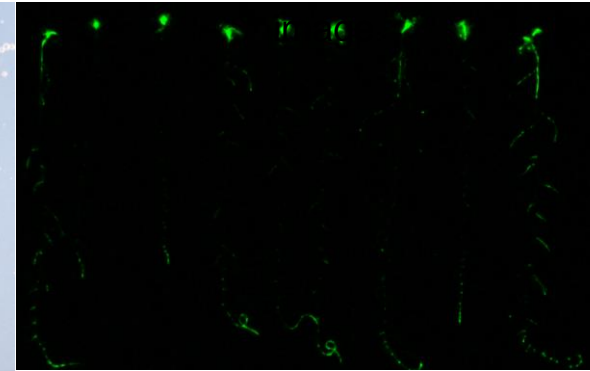
Bright field image



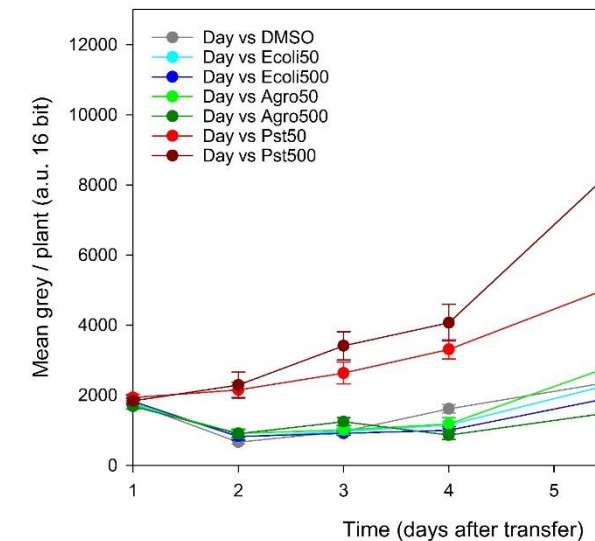
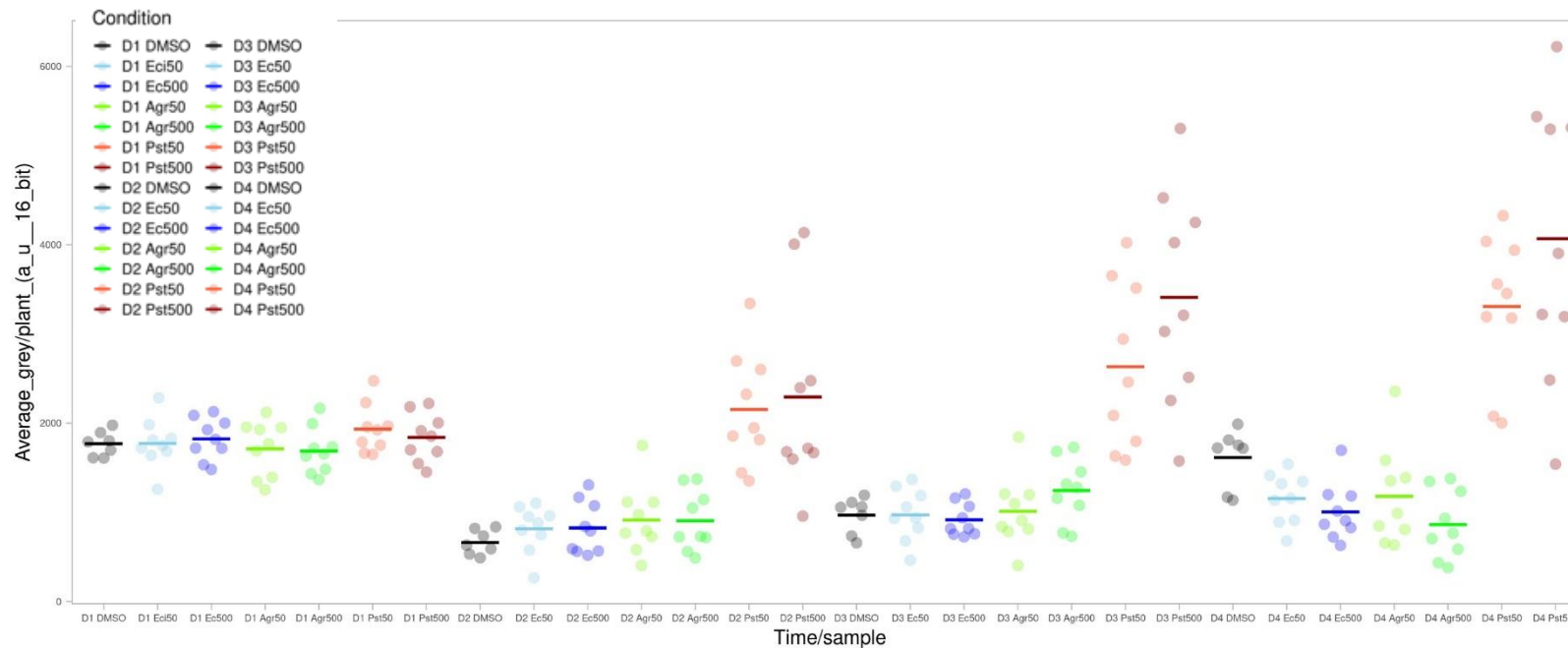
ROI threshold image



Luminescence

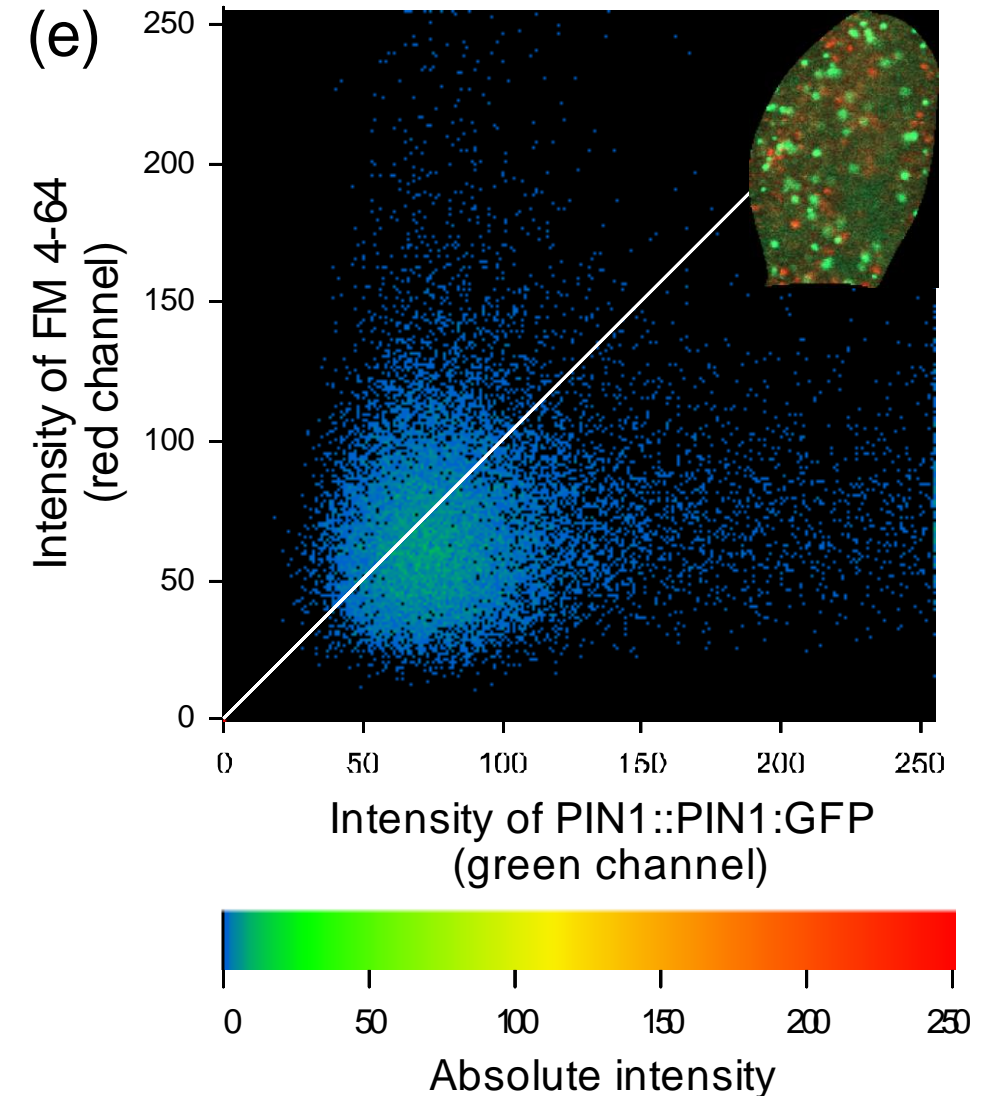
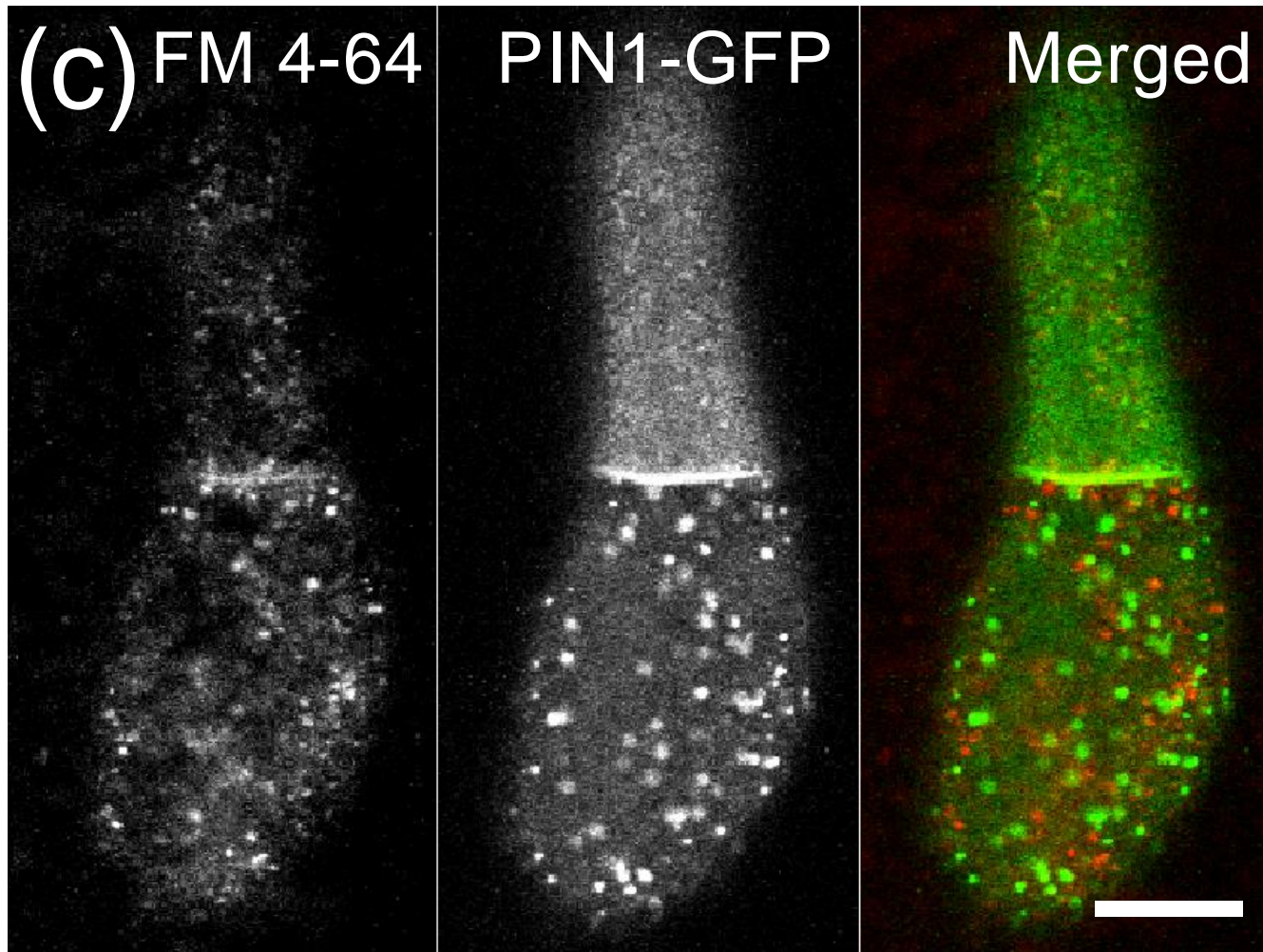


ROIs at luminescence



# 2.5 Image analysis, graphical software, presentation software

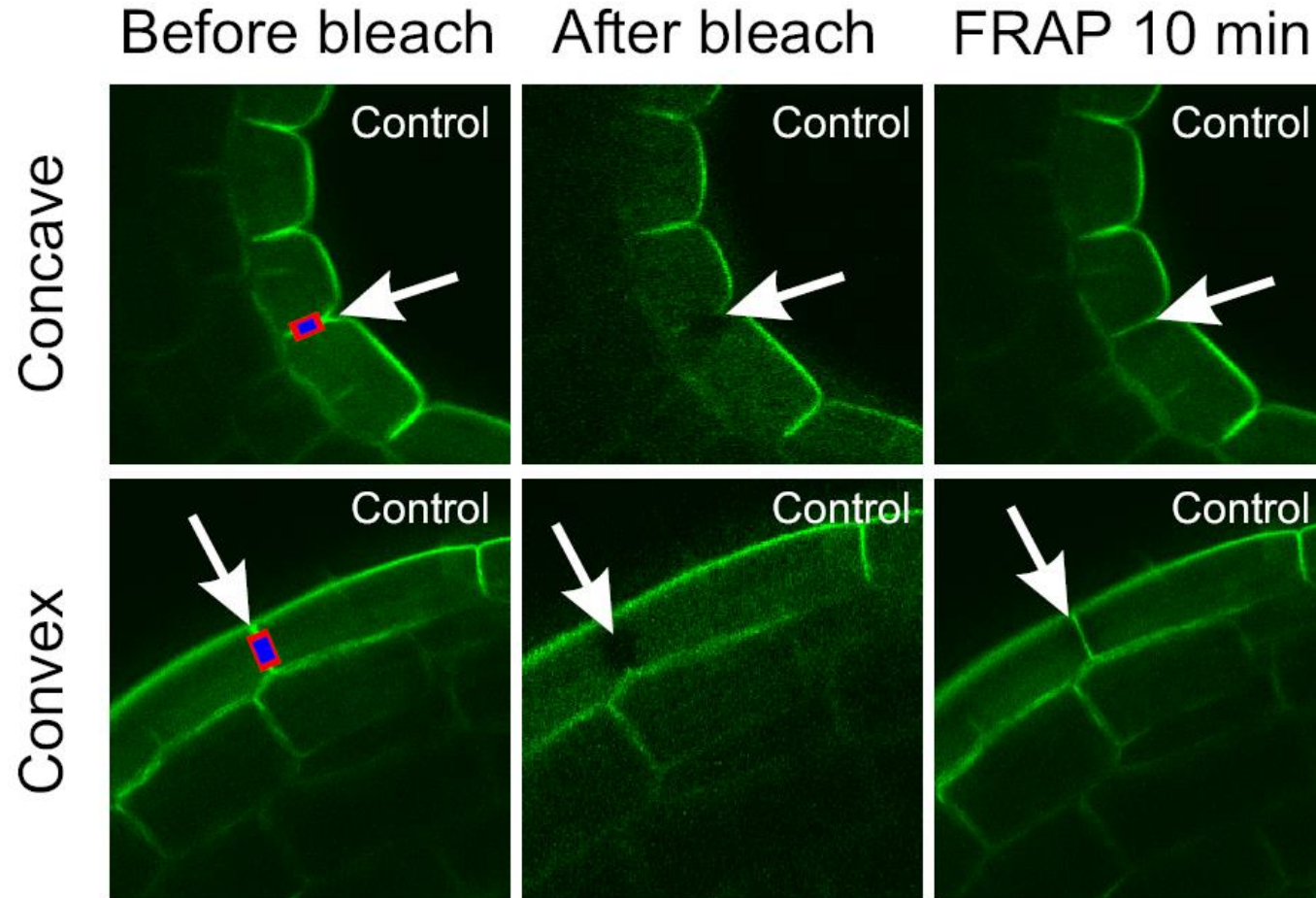
## Analysis of colocalization



Jelínková et al., *Plant Journal* 61, 883-891, 2009

## 2.5 Image analysis, graphical software, presentation software

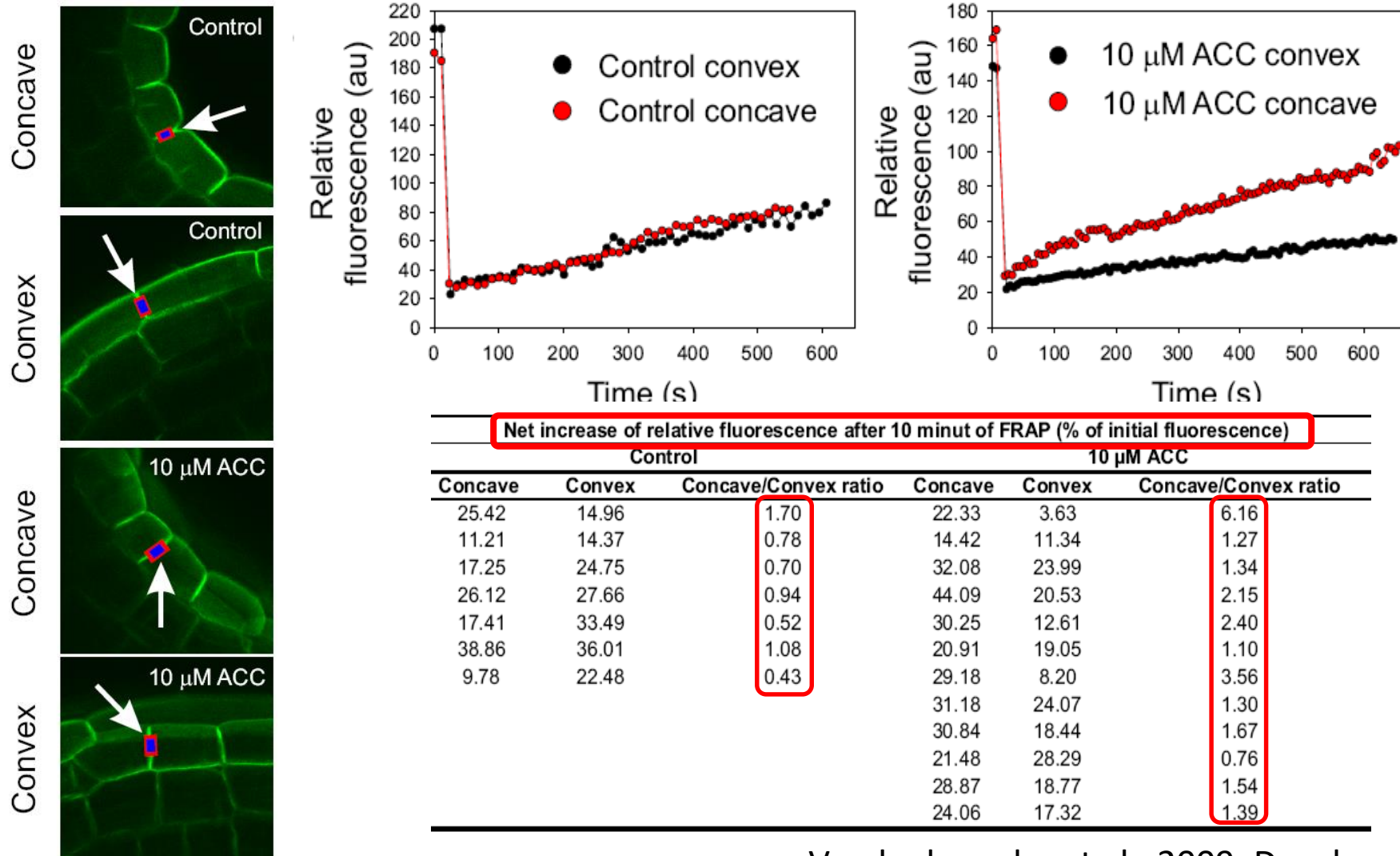
### Analysis of intracellular protein dynamics



Vandenbussche et al., 2009, Development 137, 597-606

# 2.5 Image analysis, graphical software, presentation software

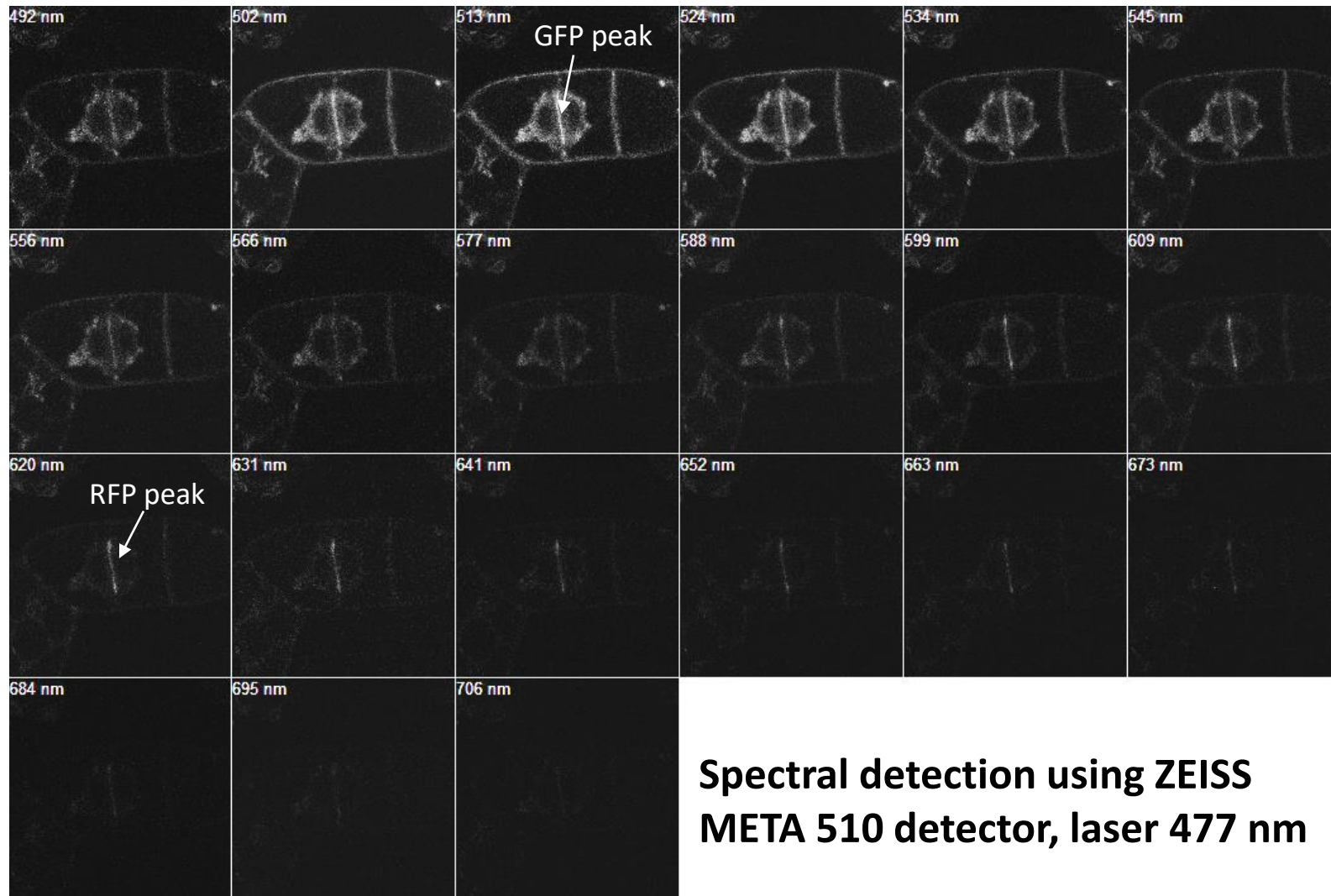
## Analysis of intracellular protein dynamics



Vandenbussche et al., 2009, Development 137, 597-606

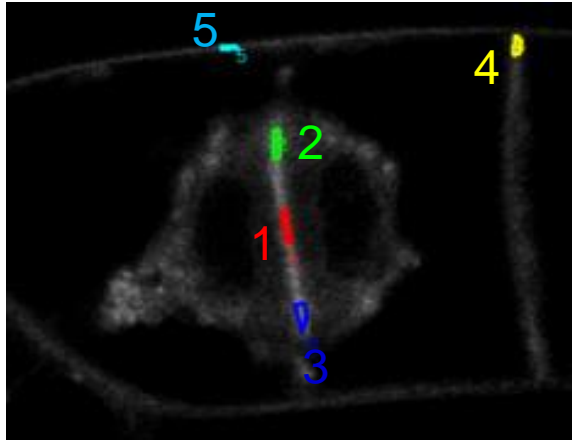
## 2.5 Image analysis, graphical software, presentation software

### Interaction between two proteins - FRET analysis *in vivo* (PIN1-ADL1)



# 2.5 Image analysis, graphical software, presentation software

## Interaction between two proteins - FRET analysis *in vivo* (PIN1-ADL1)

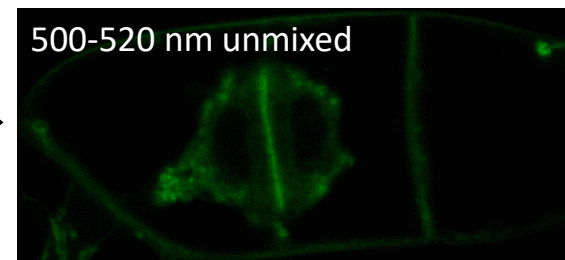
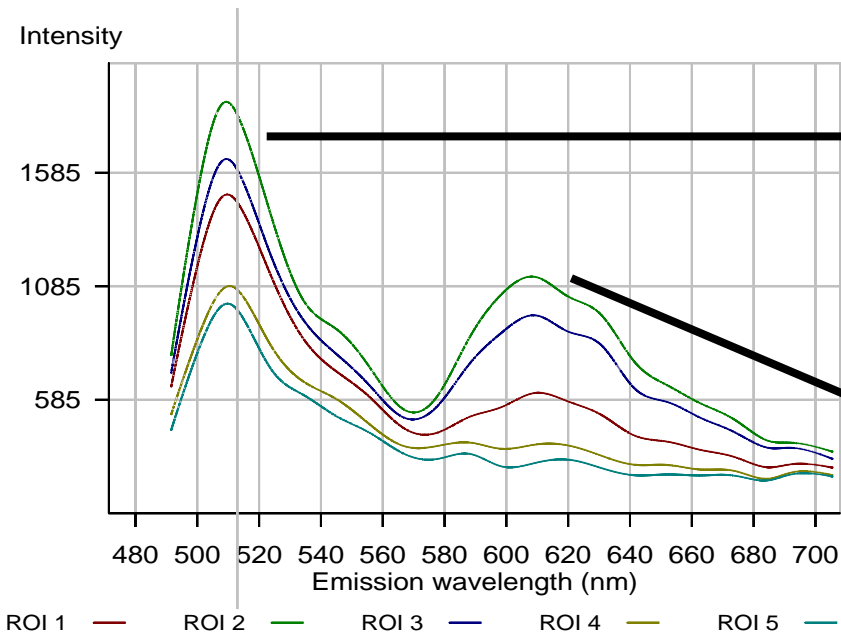


Average fluorescence intensity measured over the emission spectrum in the following region of interests:

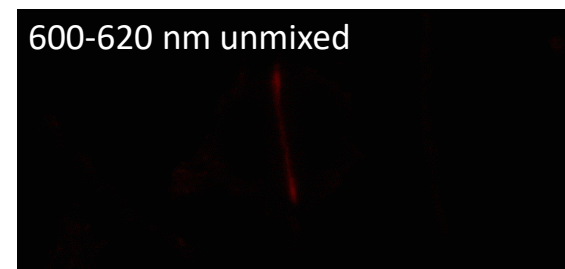
**ROI1** - middle cell plate

**ROI2, ROI3** - growing ends of the cell plate

**ROI4, ROI5** - transversal and longitudinal plasma membranes



**ROI1** - low FRET



**ROI2, ROI3** - high FRET

**ROI4, ROI5** - no FRET

# 2.5 Image analysis, graphical software, presentation software

[Irfan View](#) - the ideal tool for daily viewing of experimental documentation

## IRFANVIEW GRAPHIC VIEWER

- ▶ Fast and compact ( just 5 MB )
- ▶ Freeware for non-commercial use
- ▶ Supports Windows XP, Vista, 7, 8, and 10
- ▶ 32 and 64 bit version
- ▶ Multi language support
- ▶ Unicode support
- ▶ Designed to be simple but powerful

[More information about IrfanView](#)

GET IRFANVIEW (VERSION 4.58)

The program is available in 32 and 64 bit.  
Which version should I download?  
[See 64-bit info.](#)

32-BIT	64-BIT
<b>DOWNLOAD</b> Current version 4.58	<b>DOWNLOAD</b> Current version 4.58

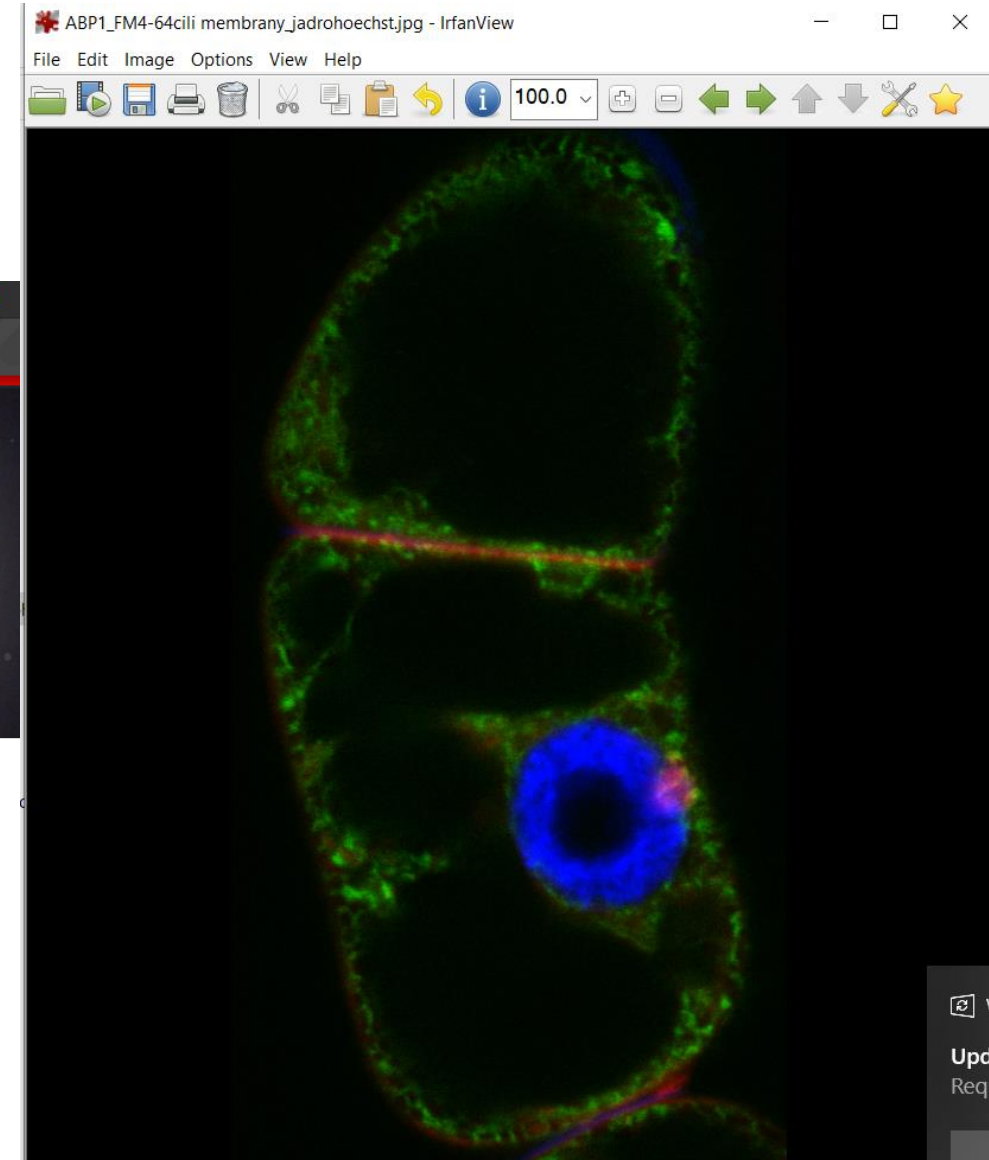
USEFUL PARTNER SITES

Stempel bestellen:  
[www.stempelservice.de](http://www.stempelservice.de)

scams.info:  
[www.scams.info](http://www.scams.info)

Lunacy, Sketch for Windows:  
[icons8.com](http://icons8.com)

Top 5 UI/UX design agencies:  
[unlabeled.com](#)



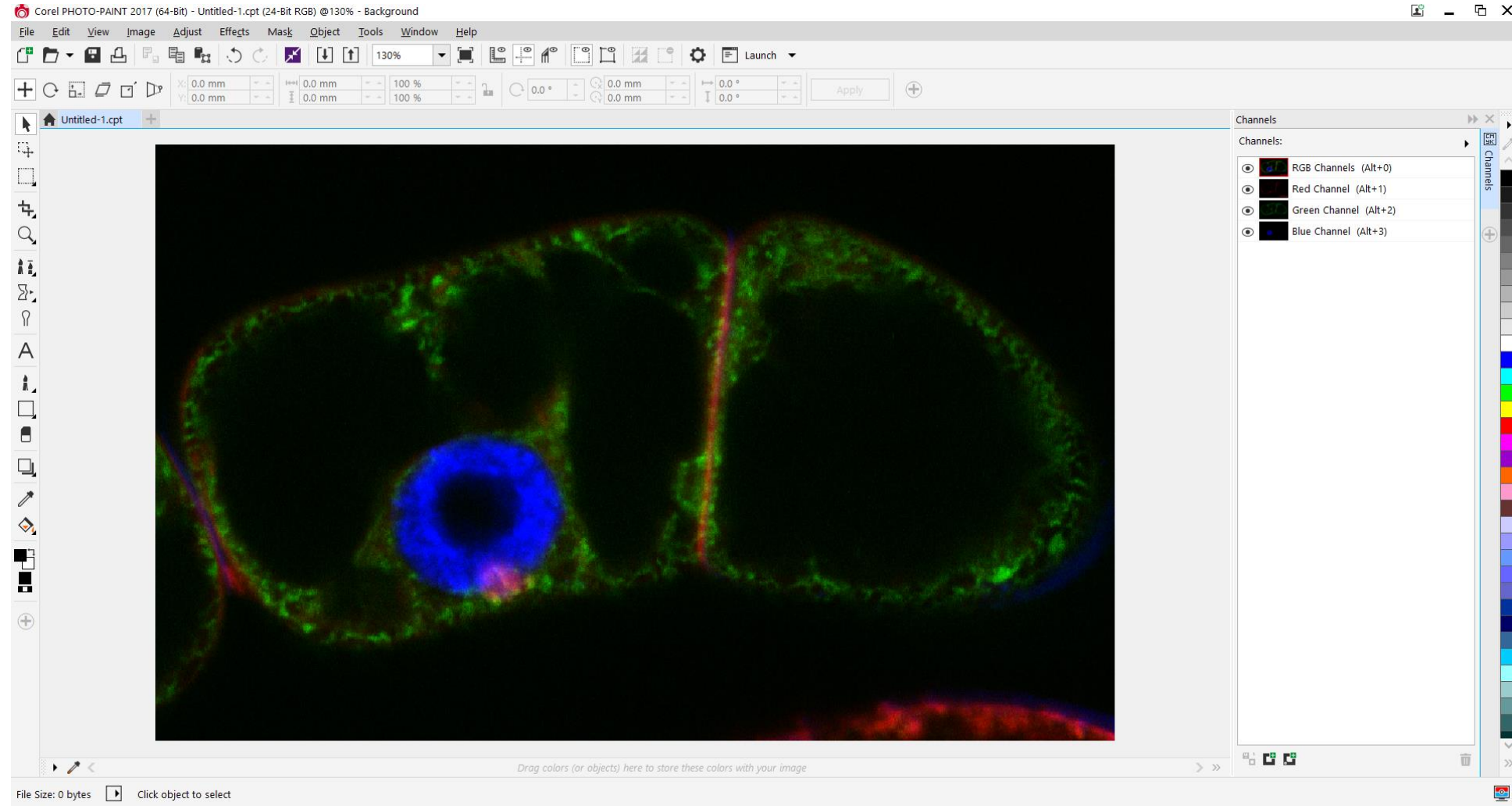
# 2.5 Image analysis, graphical software, presentation software

## Raster graphics and its editing using software editors

### Corel Photo Paint



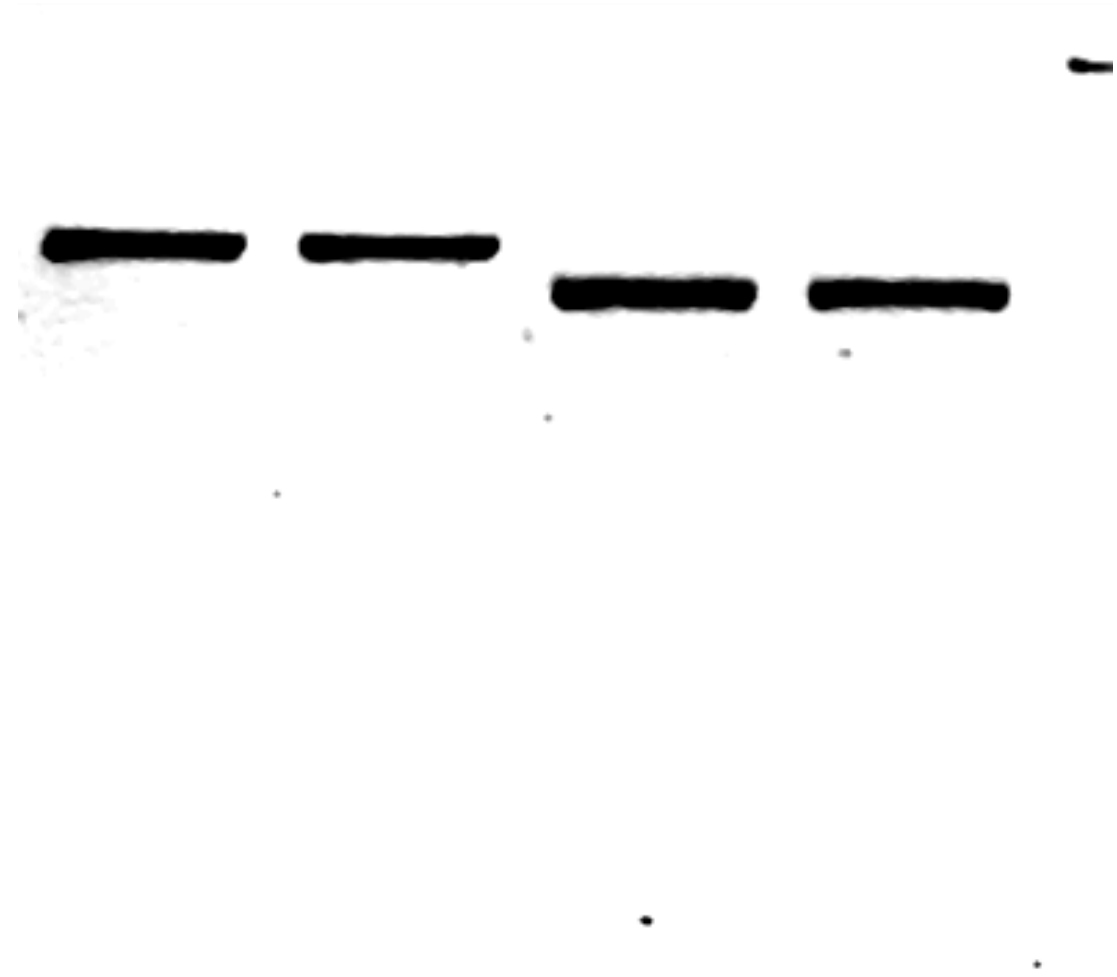
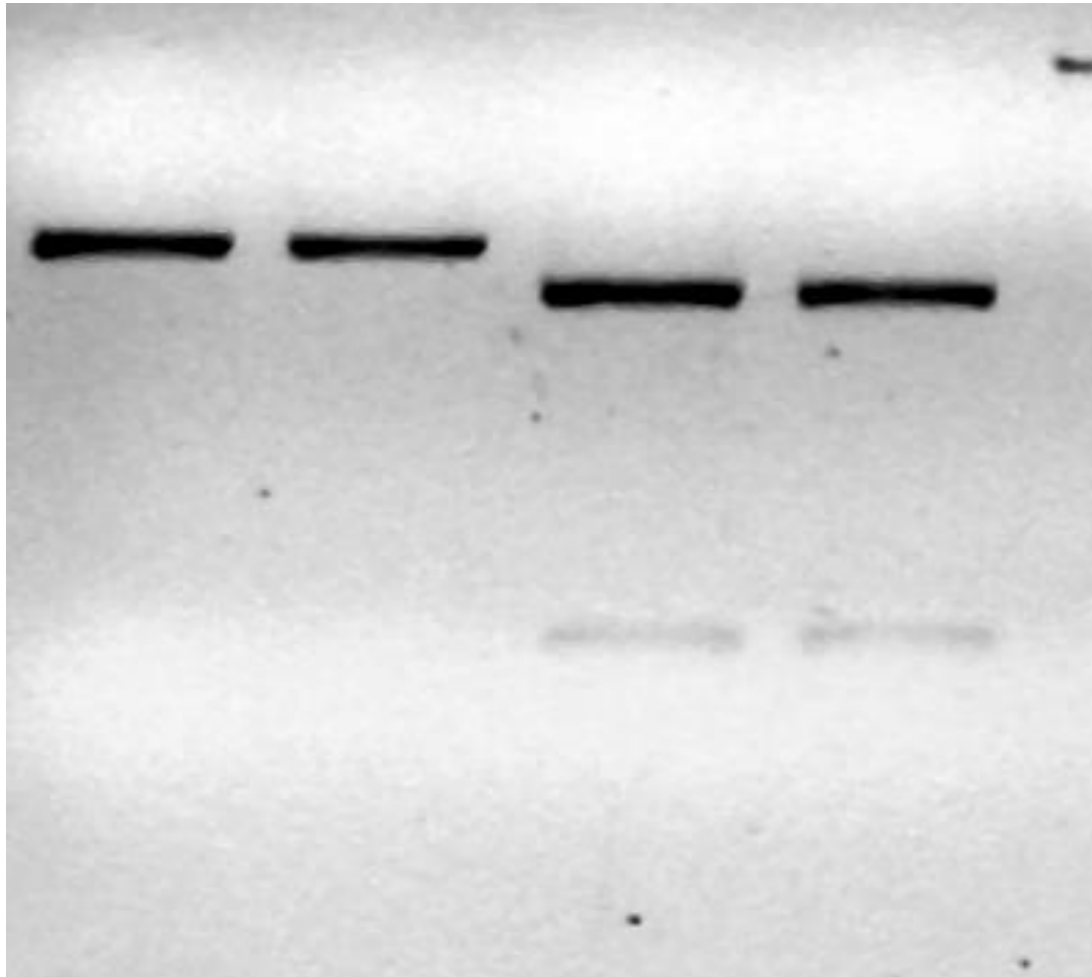
### Adobe Photoshop





# 2.5 Image analysis, graphical software, presentation software

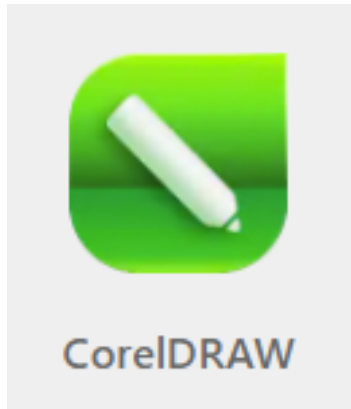
Beware of too extensive modifications!!!!



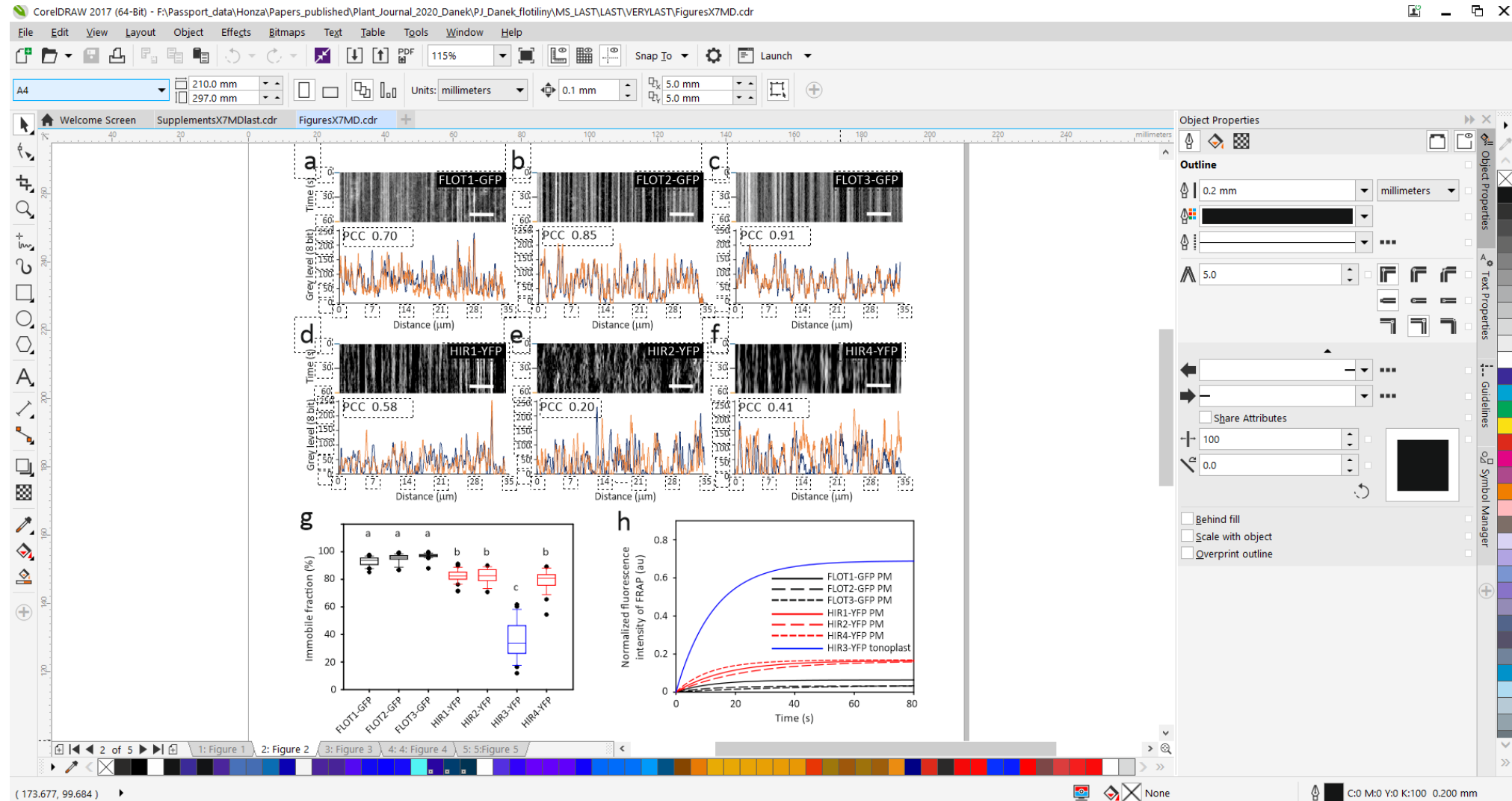
# 2.5 Image analysis, graphical software, presentation software

Vector graphics combined with raster graphics - ideal for creating images for scientific publications

Corel Draw



Adobe Illustrator



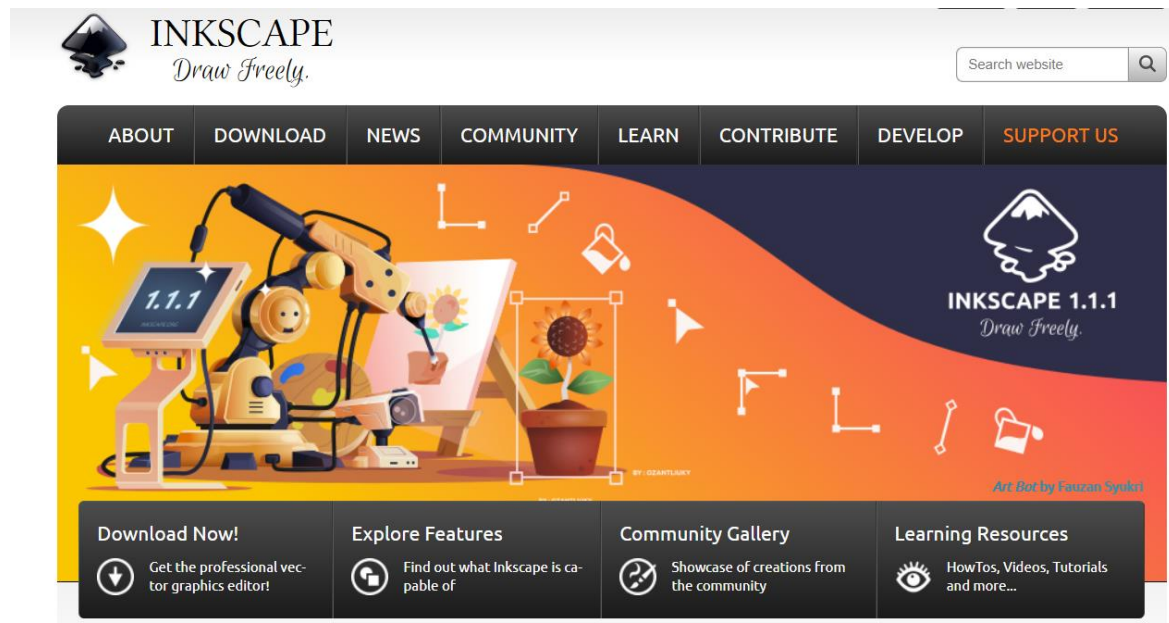
## 2.6 Open source graphical software

Both vector and raster graphics can be edited using freely available programs, but for the purpose of presenting scientific outputs these tools are rather not suitable

GIMP



Inkscape



# Data management

FAIR principles - Findable, Accesible, Iteroperable and Reusable



## FINDABLE

Data has rich metadata and unique identifier



## ACCESSIBLE

Data can be easily downloaded or used by using standard protocols



## INTEROPERABLE

Metadata use an accessible and standard language



## REUSABLE

Data is well-described and provides clear usage of licences

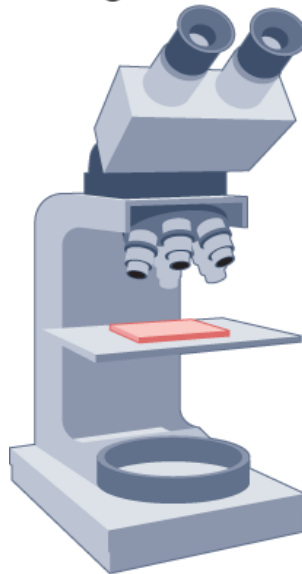
[FAIR Principles - GO FAIR](#)

# FAIR principles in bioimaging

## Community-developed checklists for publishing images and image analyses

### Acquisition of microscopy images

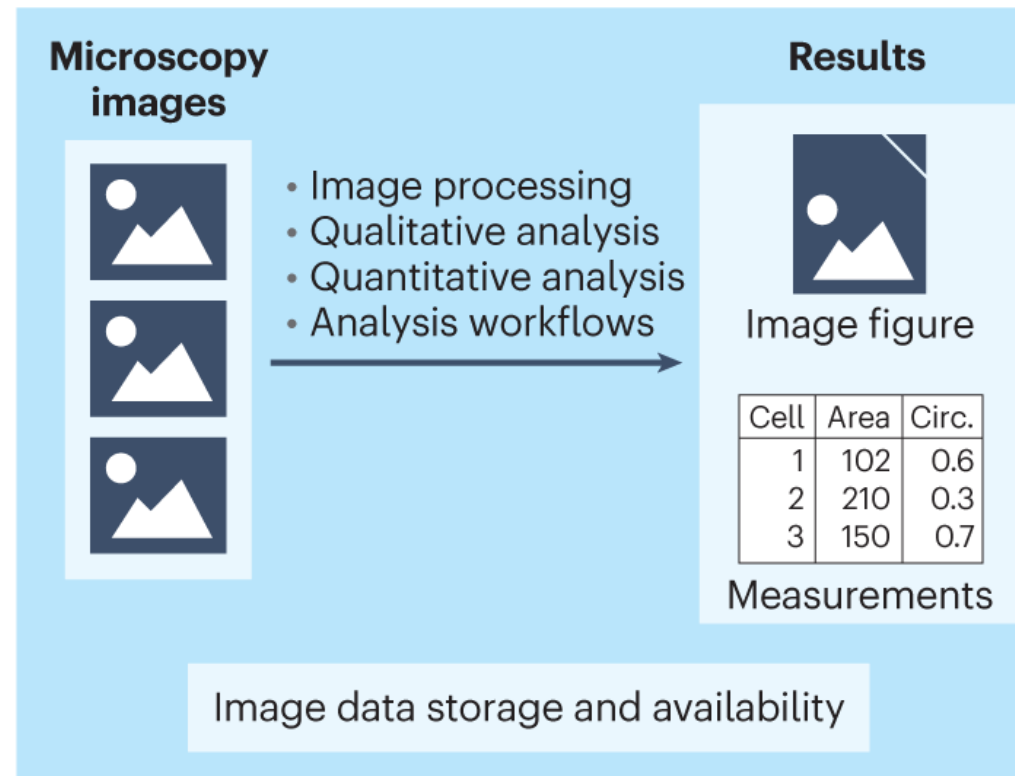
Image, microscope setup, experimental design



### Foundation

Experimental design and good scientific practice

### Publishing of microscopy image figures and image analyses



<https://doi.org/10.1038/s41592-023-01987-9>