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 od 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

2.2 Data explorers, correct data handling and saving

2.2 Data explorers, correct data handling and saving

Total Commander
- shareware
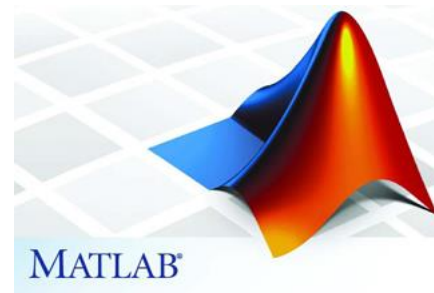- the biggest advantage is the existence of two identical windows
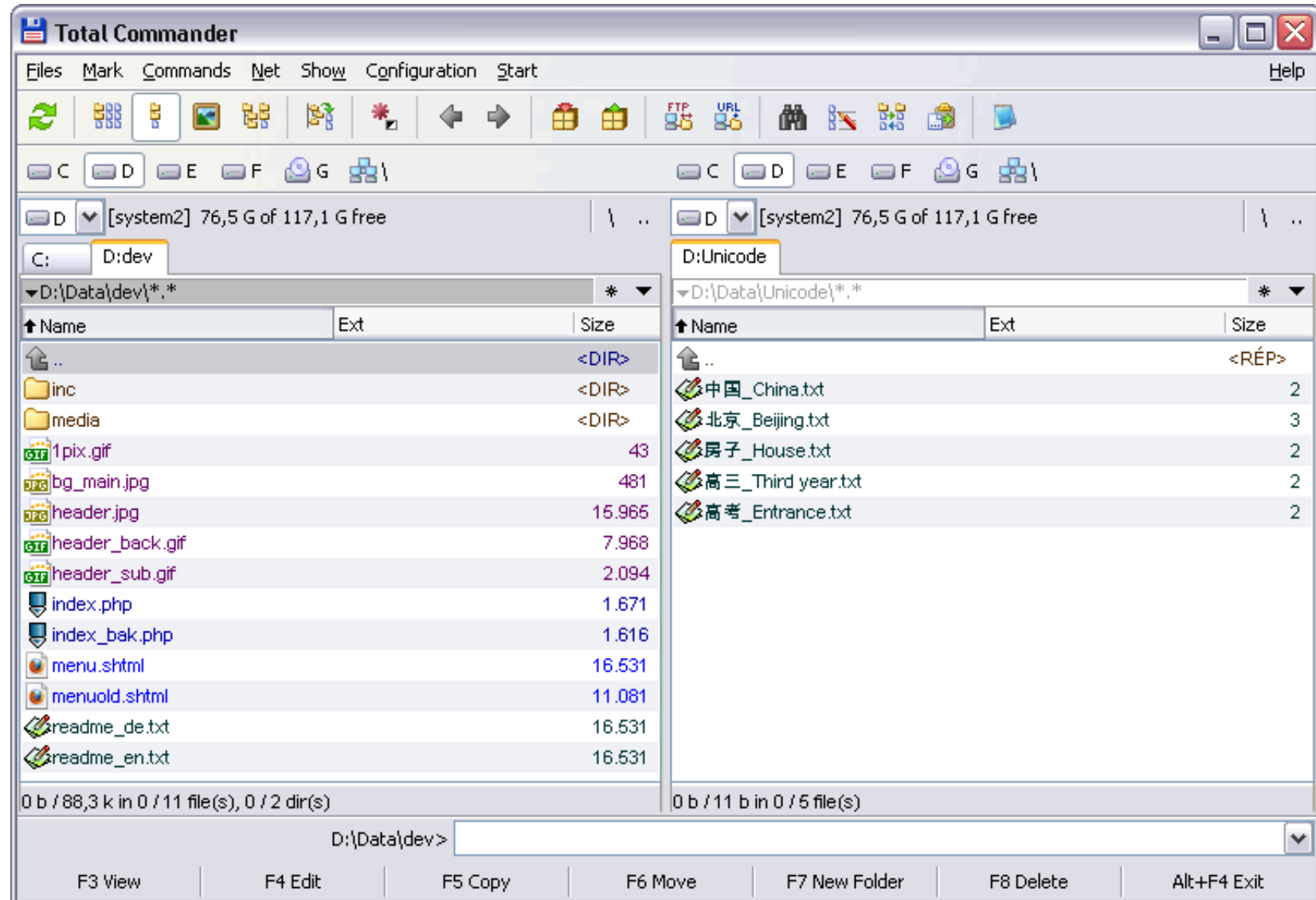
2.3 Spreadsheets, statistical software, graph editors

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
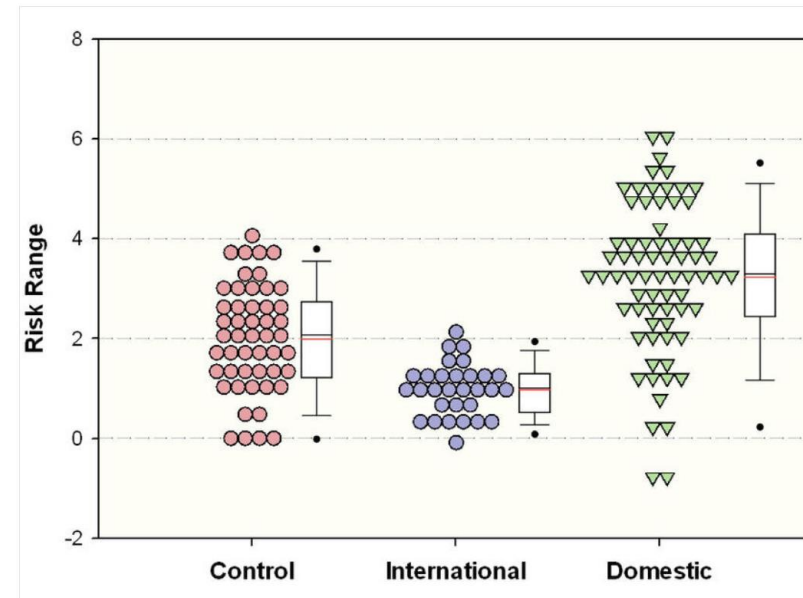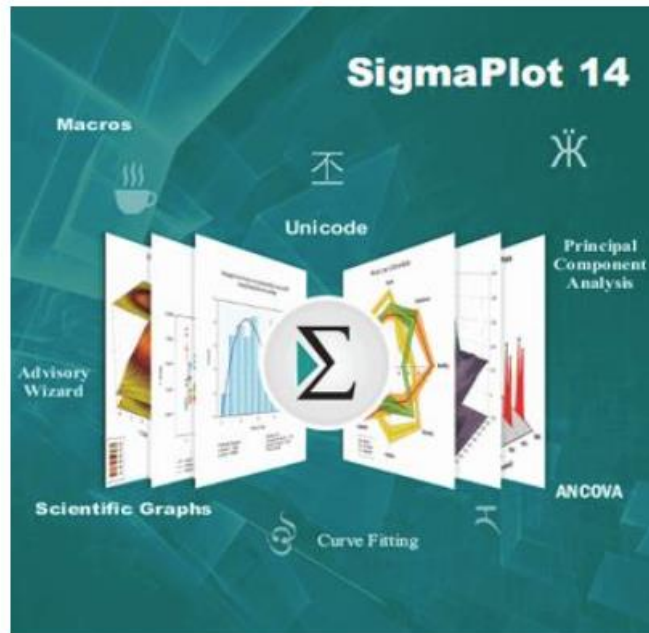
2.3 Spreadsheets, statistical software, graph editors

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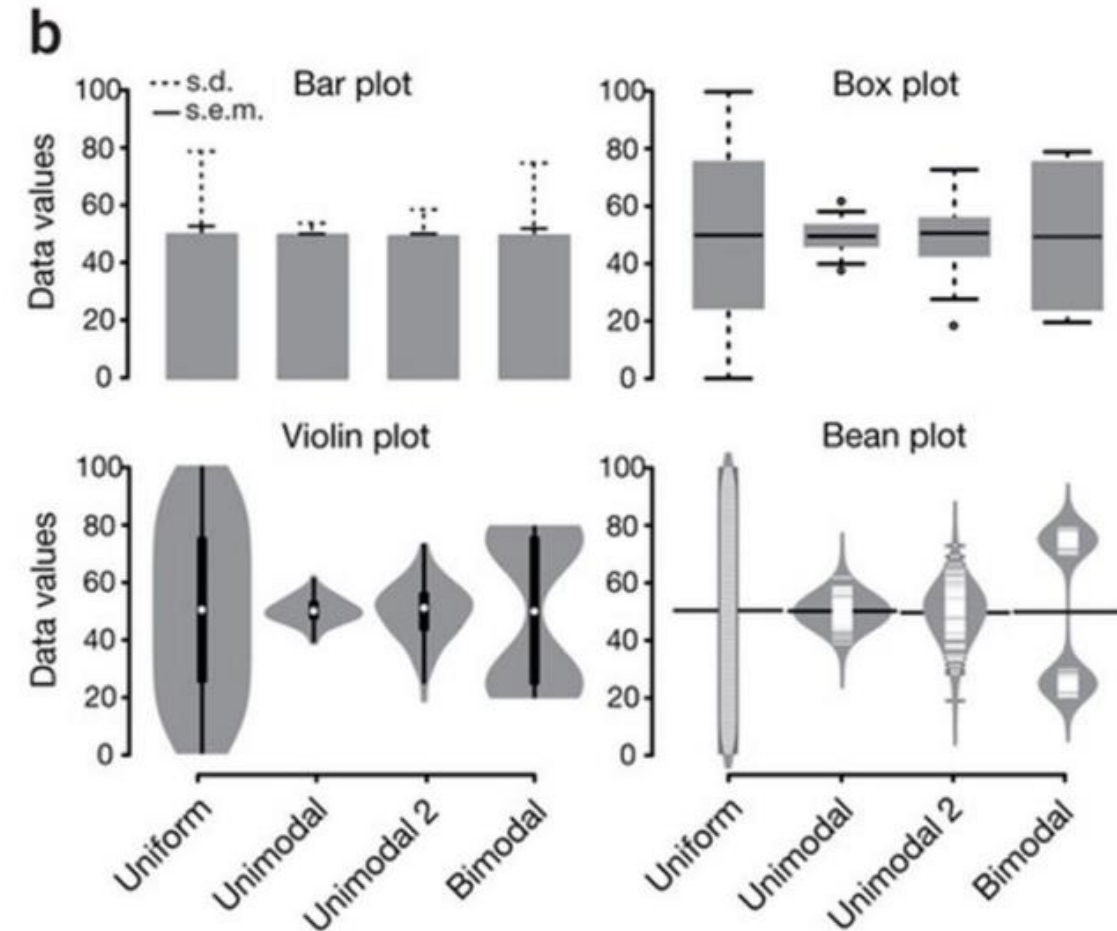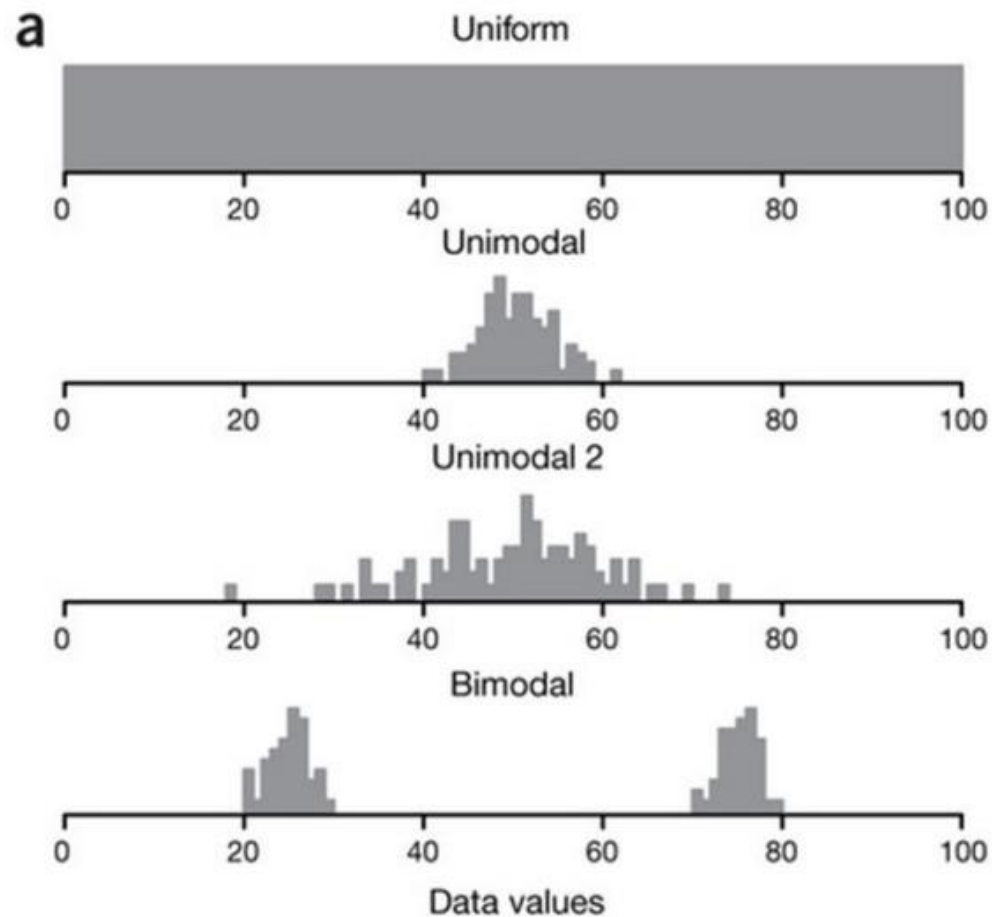
Spitzer M, Wildenhain J, Rappsilber J, Tyers M (2014) BoxPlotR: a web tool for generation of box plots. Nat Methods 11:121–122.

2.3 Spreadsheets, statistical software, graph editors

Creating graphs - line/scatter plots, bar charts, box plots, histograms
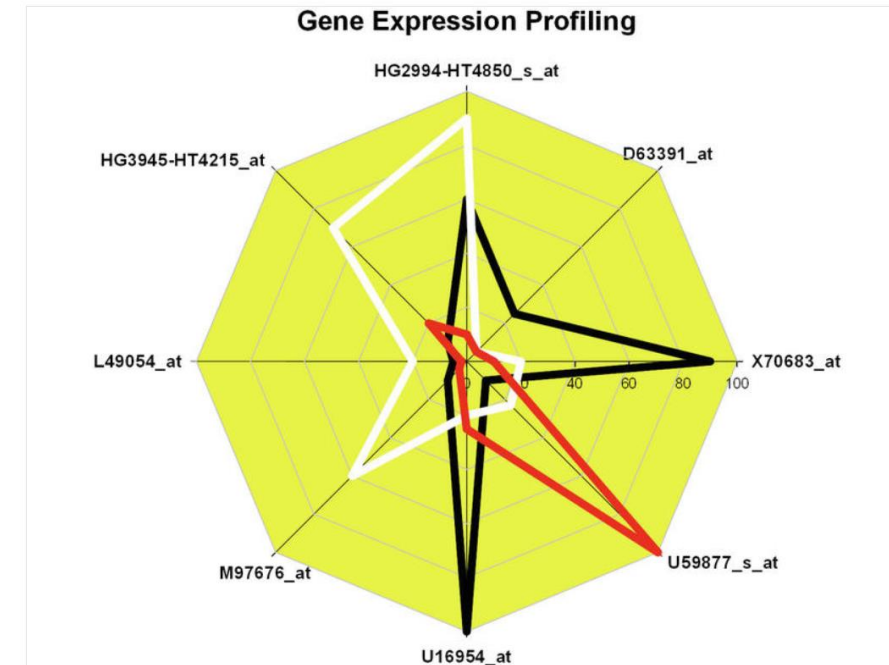
2.3 Spreadsheets, statistical software, graph editors

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
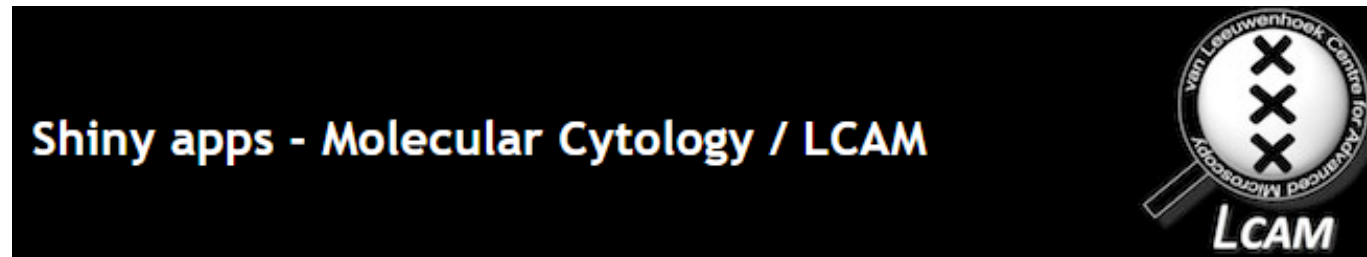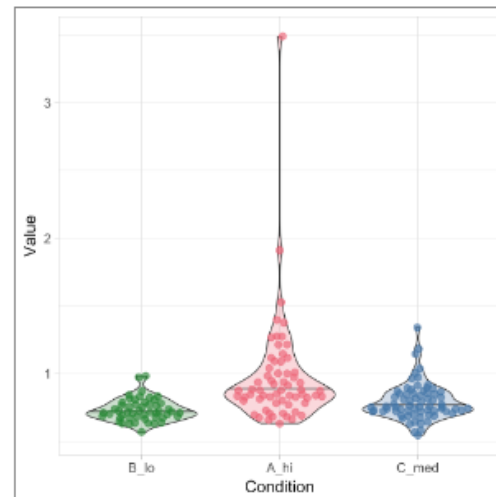
PlotsOfData   SuperPlotsOfData
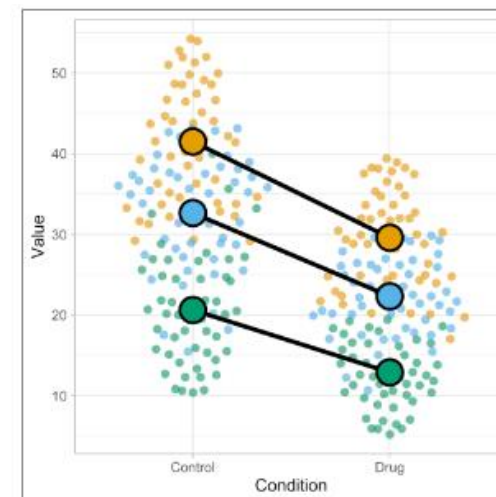
2.3 Spreadsheets, statistical software, graph editors

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

R studio helps new users to start with R
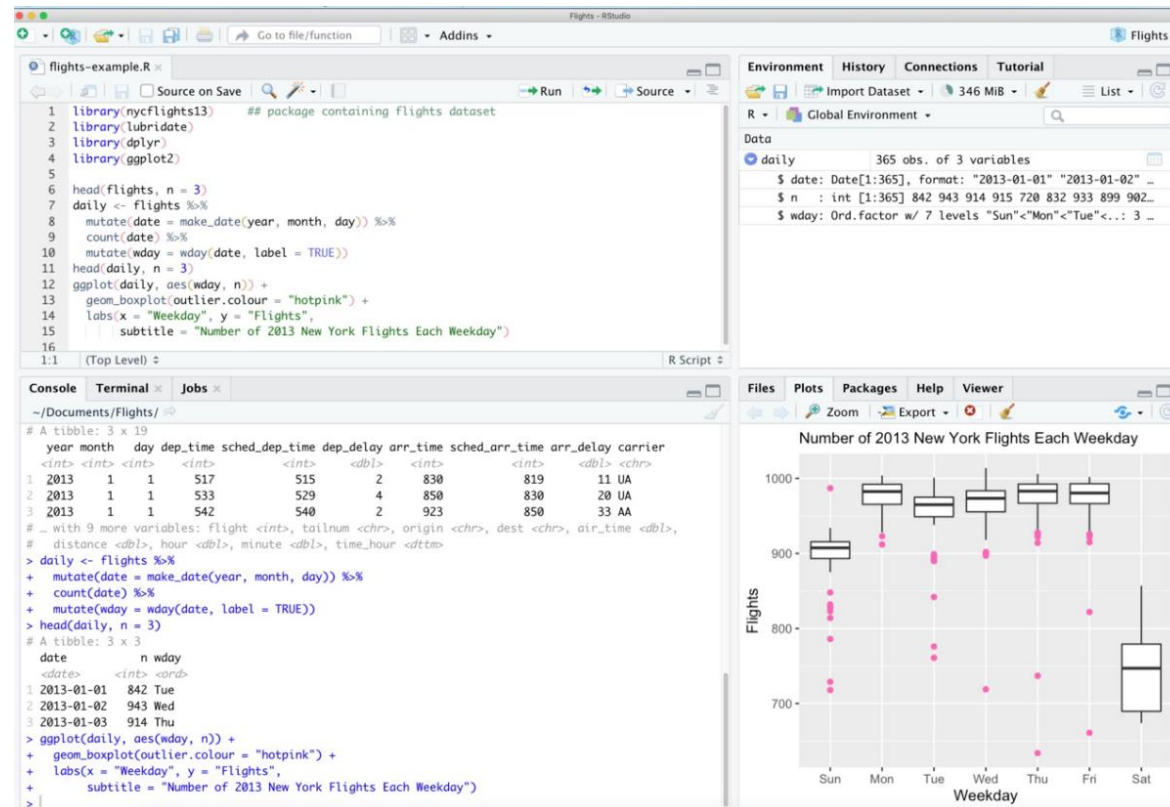
https://en.wikipedia.org/wiki/RStudio#/media/File:RStudio_IDE_screenshot.png

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:
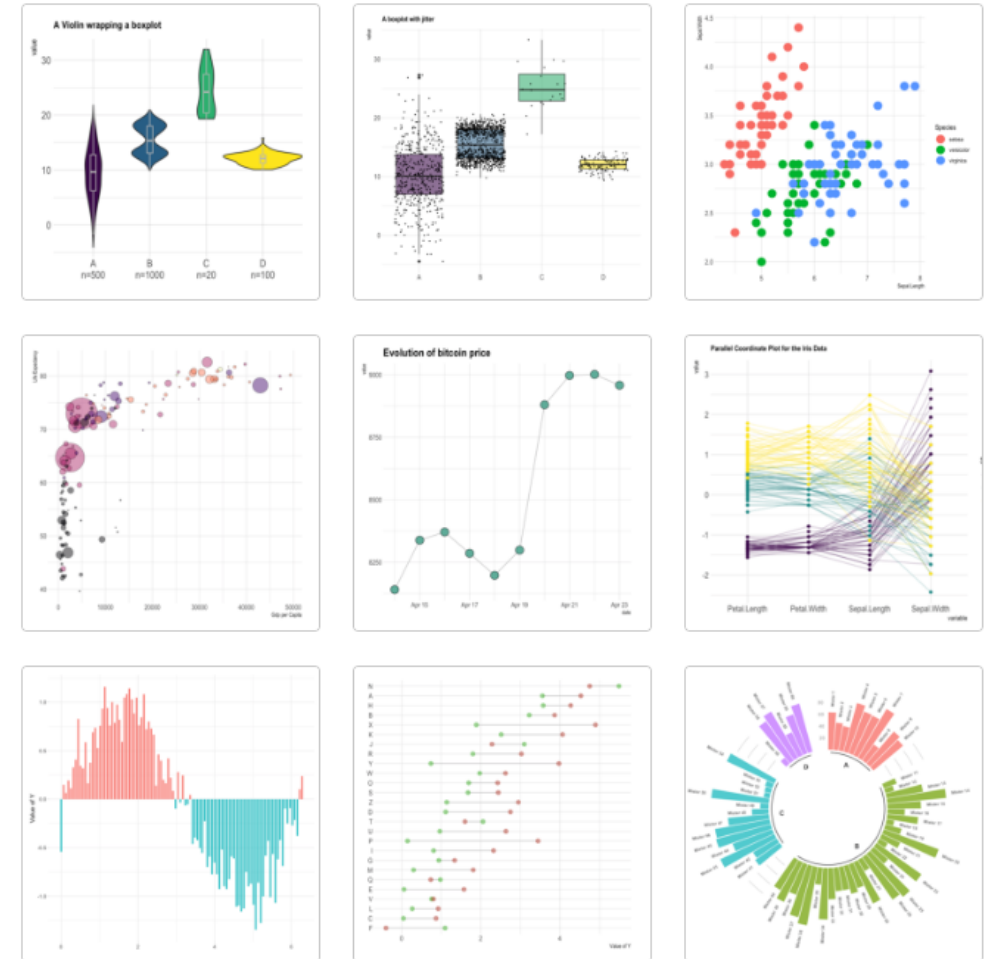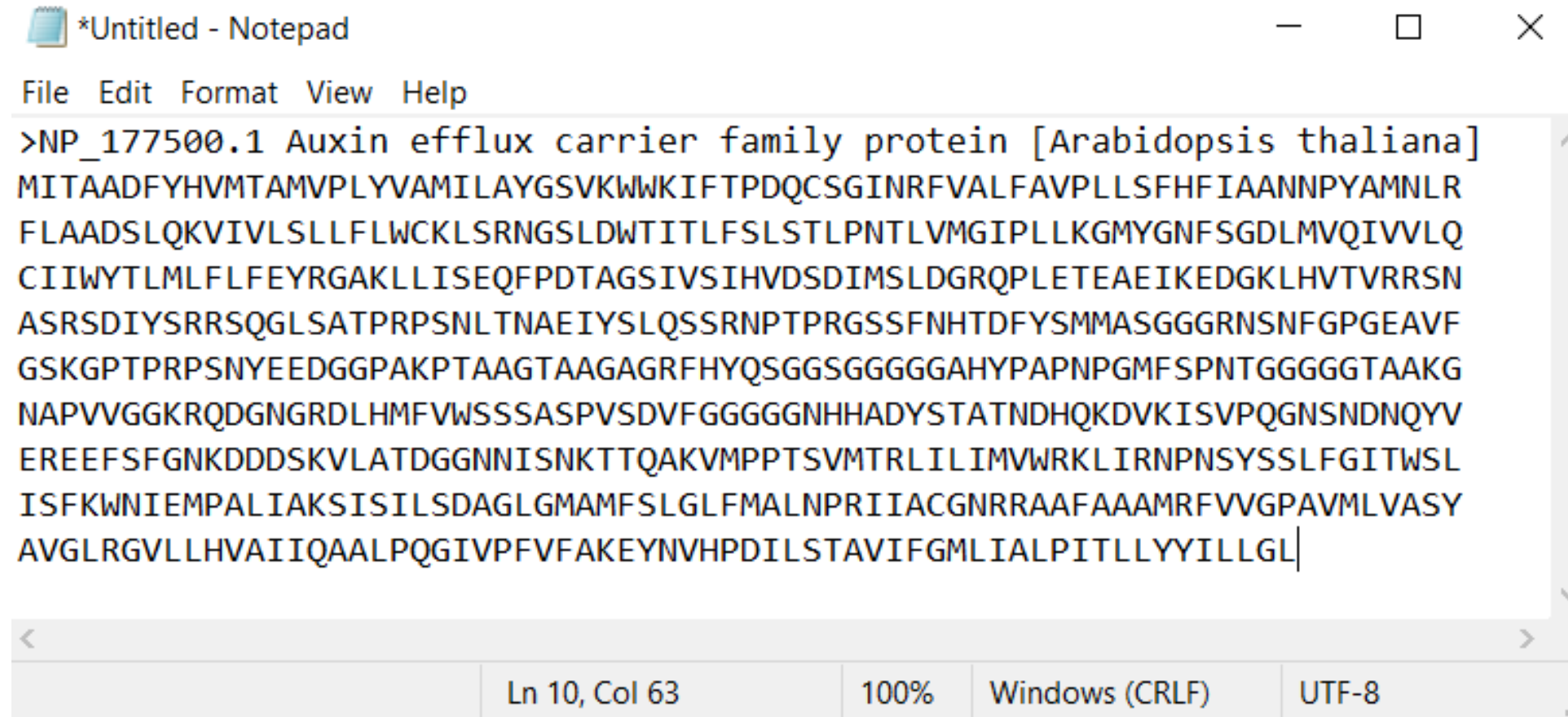
2.4 Processing of structural and sequence data

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

2.4 Processing of structural and sequence data
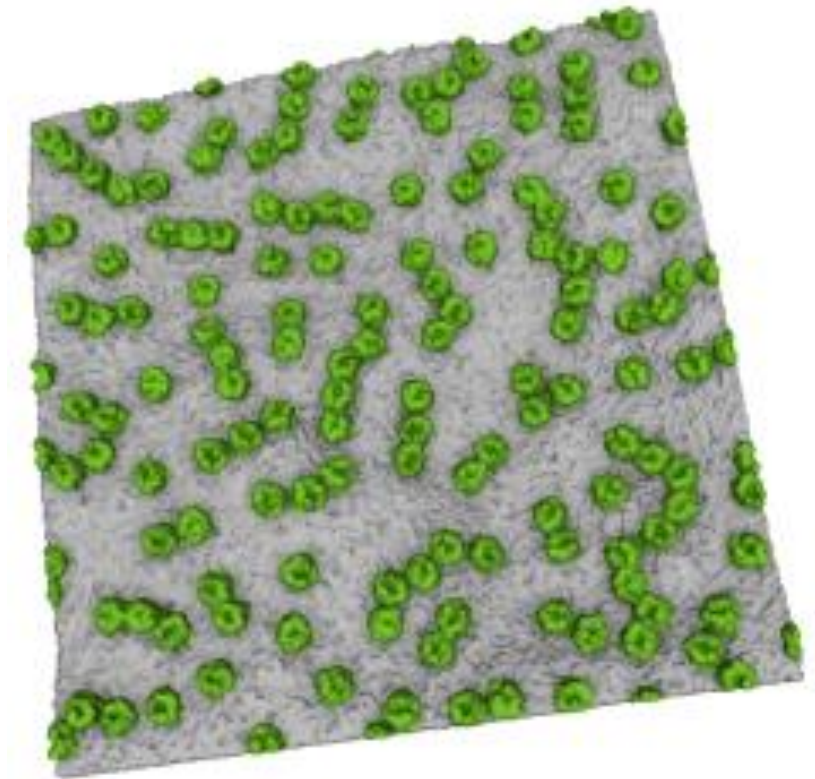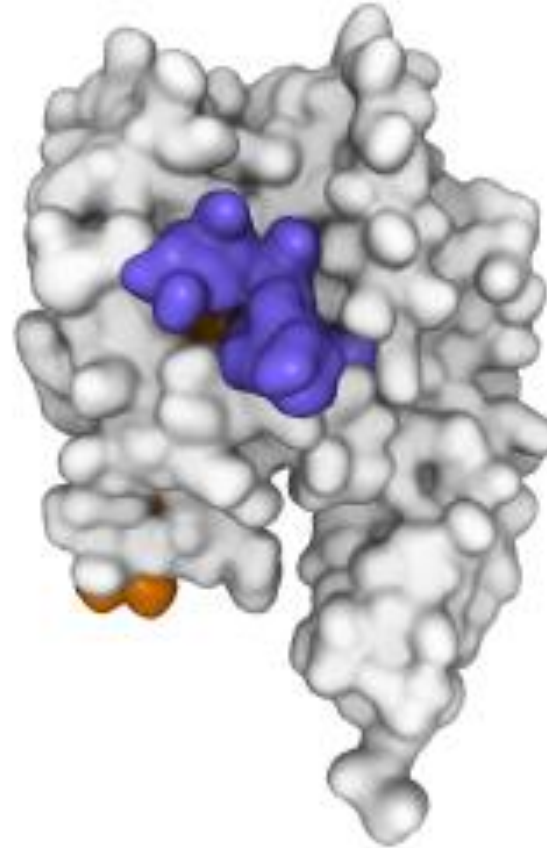
SnapGene

Geneious

2.4 Processing of structural and sequence data

3D structures of macromolecules – Molstar viewer
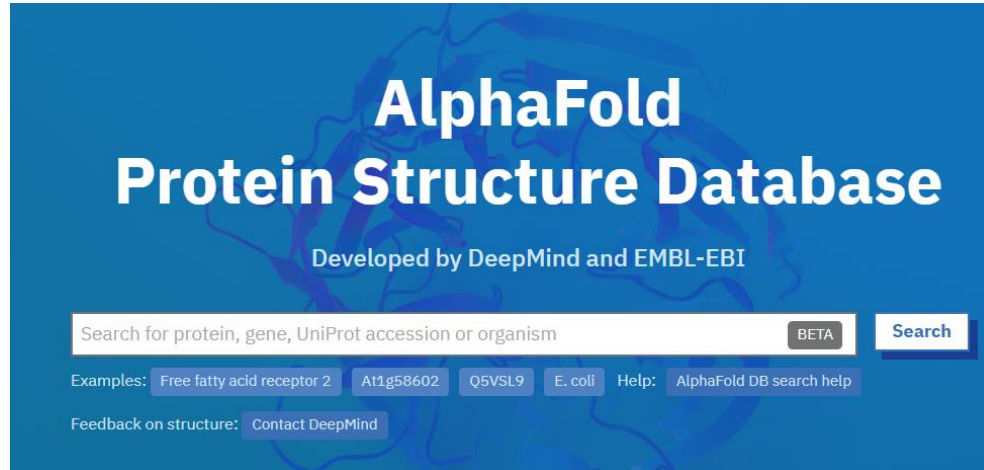

Mol* (/'molstar/) is a modern web-based open-source toolkit for visualisation and analysis of large-scale molecular data

Sehnal et al., Nucleic Acids Res 49:W431–W437, 2021

2.4 Processing of structural and sequence data

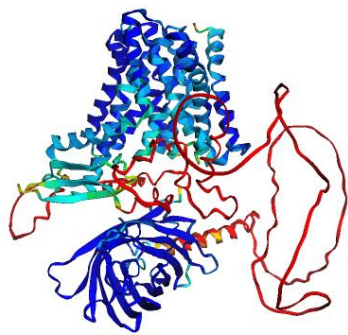3D structures of macromolecules - AI-based solution for „*in silico*" translation
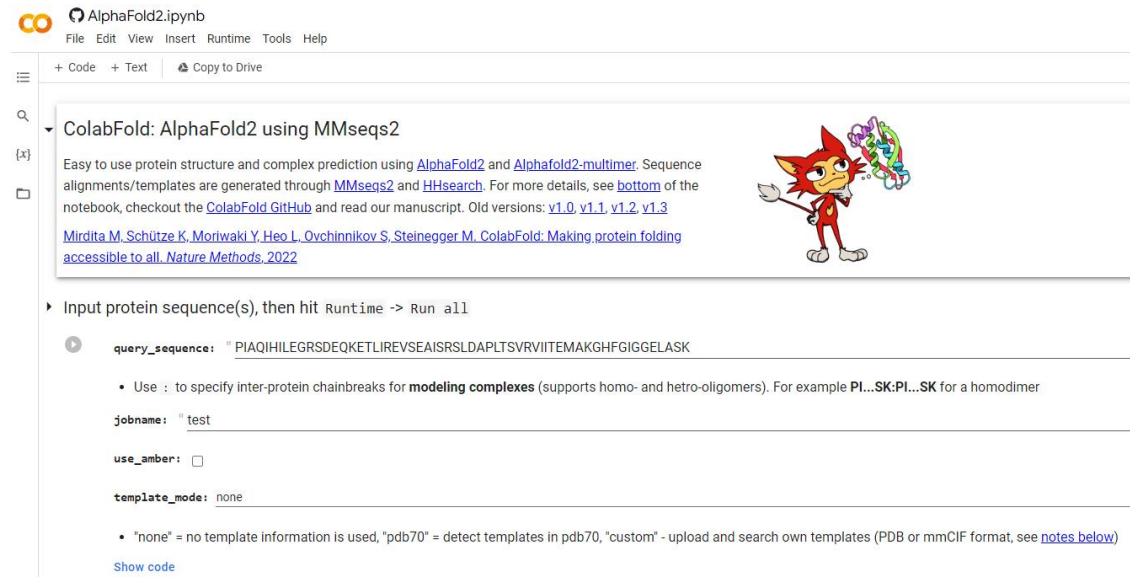
NtPIN2T   NtPIN3bT   NtPIN11T

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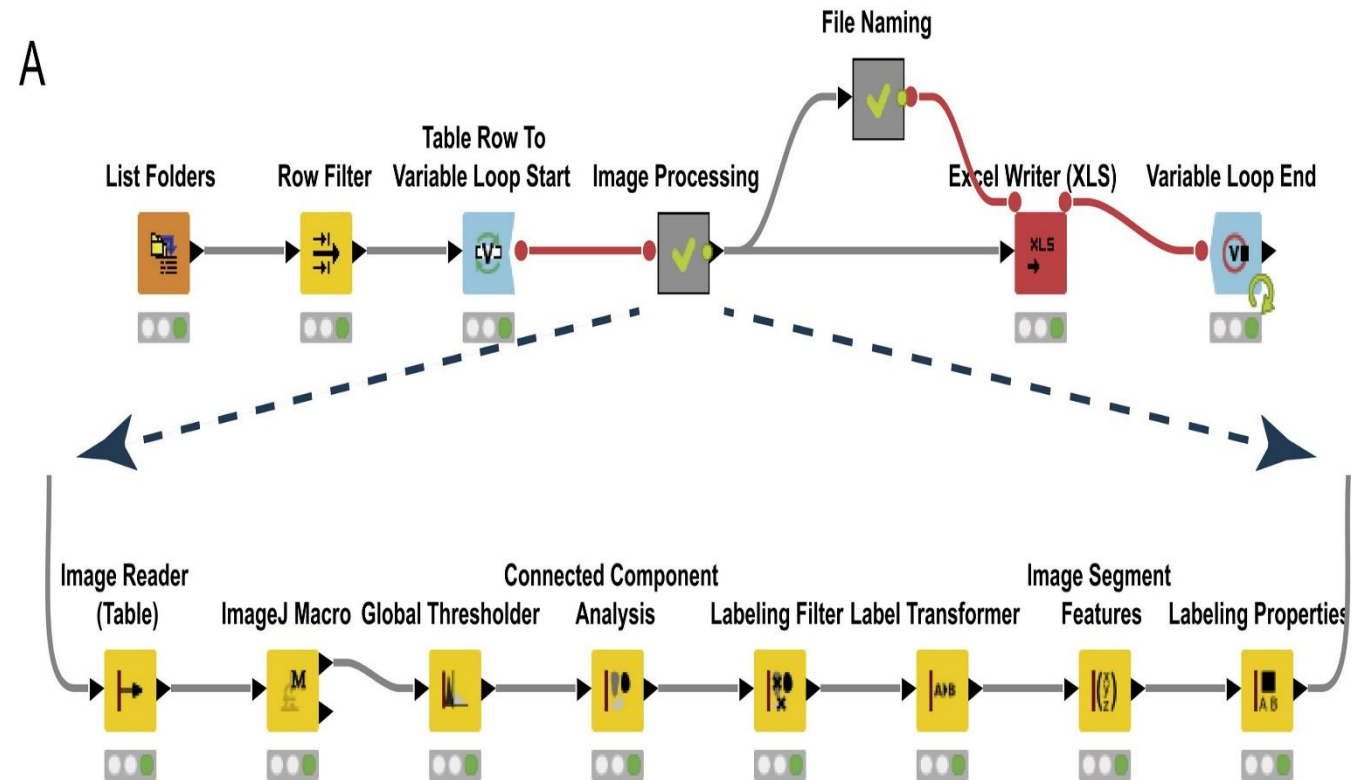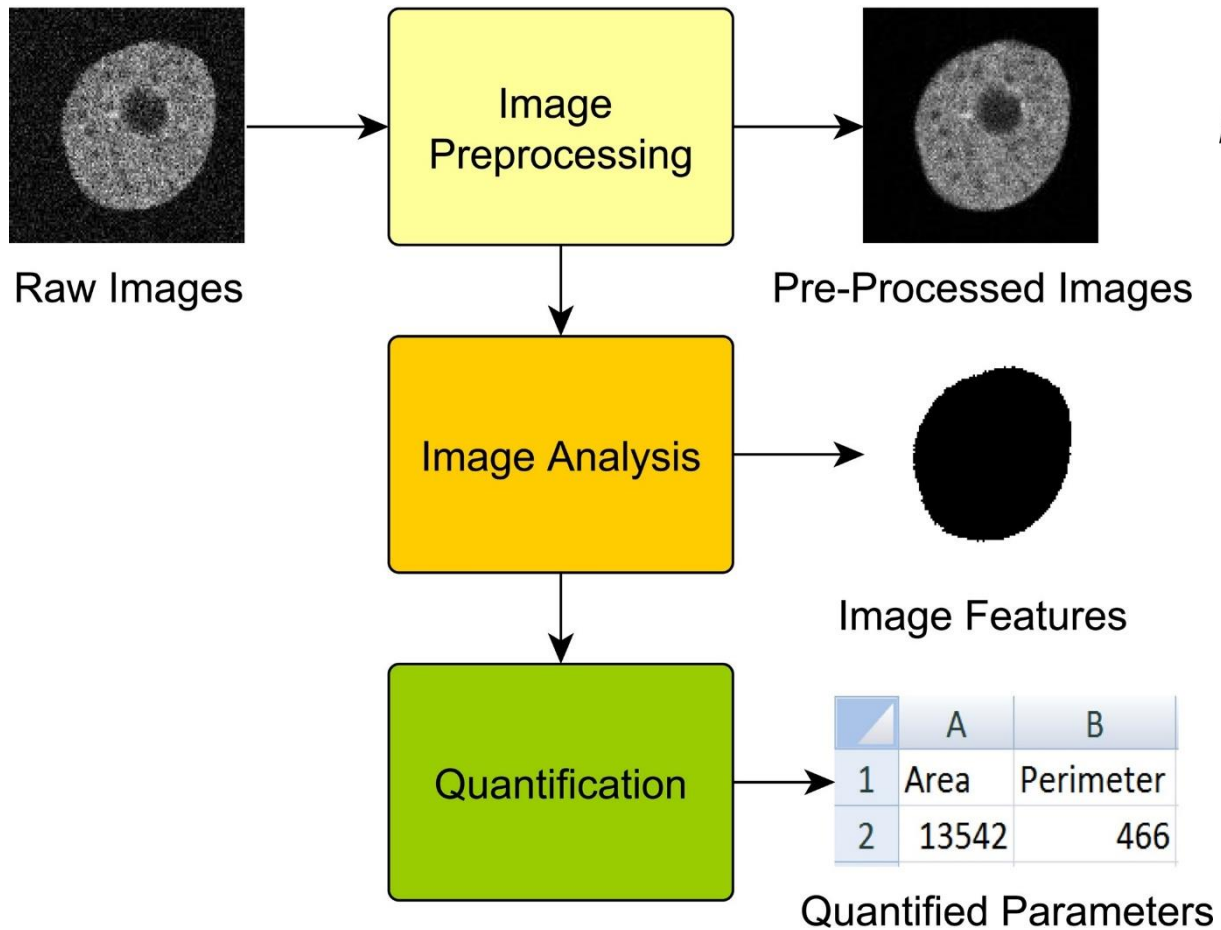
Wollmann T et al., J Biotechnol 261:70-75, 2017

2.5 Image analysis, graphical software, presentation software

Artificial intelligence in image analysis
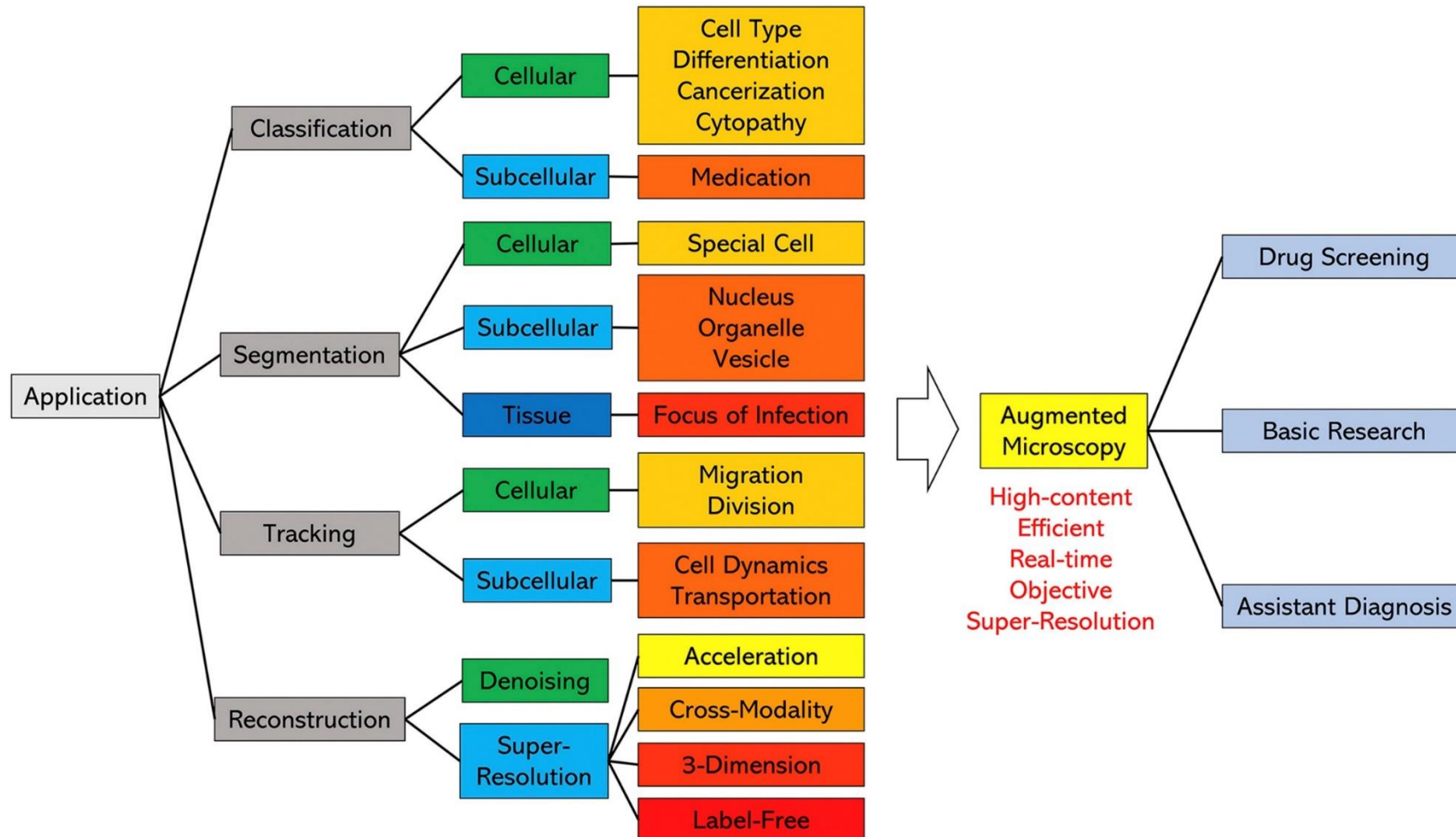
Liu et al., Comput Biol Med 134:104523, 2021

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

2.5 Image analysis, graphical software, presentation software

ImageJ - open source image analysis

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

2.5 Image analysis, graphical software, presentation software

ImageJ (NIH)

Fiji

Imaging tutorials
Scientific
Imaging
Tutorials
(imagej.net)
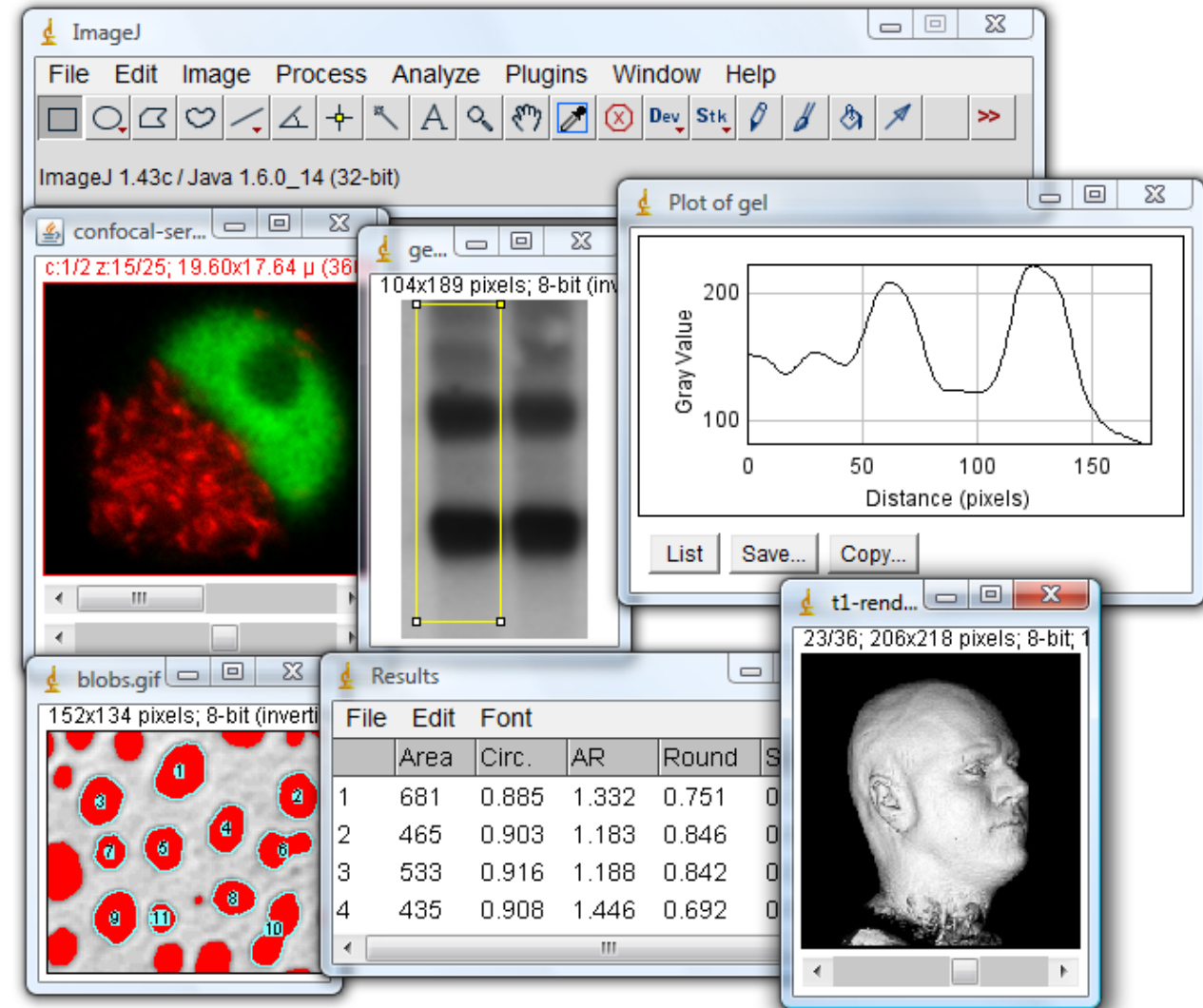
http://rsb.info.nih.gov/ij/docs/concepts.html

2.5 Image analysis, graphical software, presentation software

NIS Elements – good example of commercially available, comprehensive image analysis tool
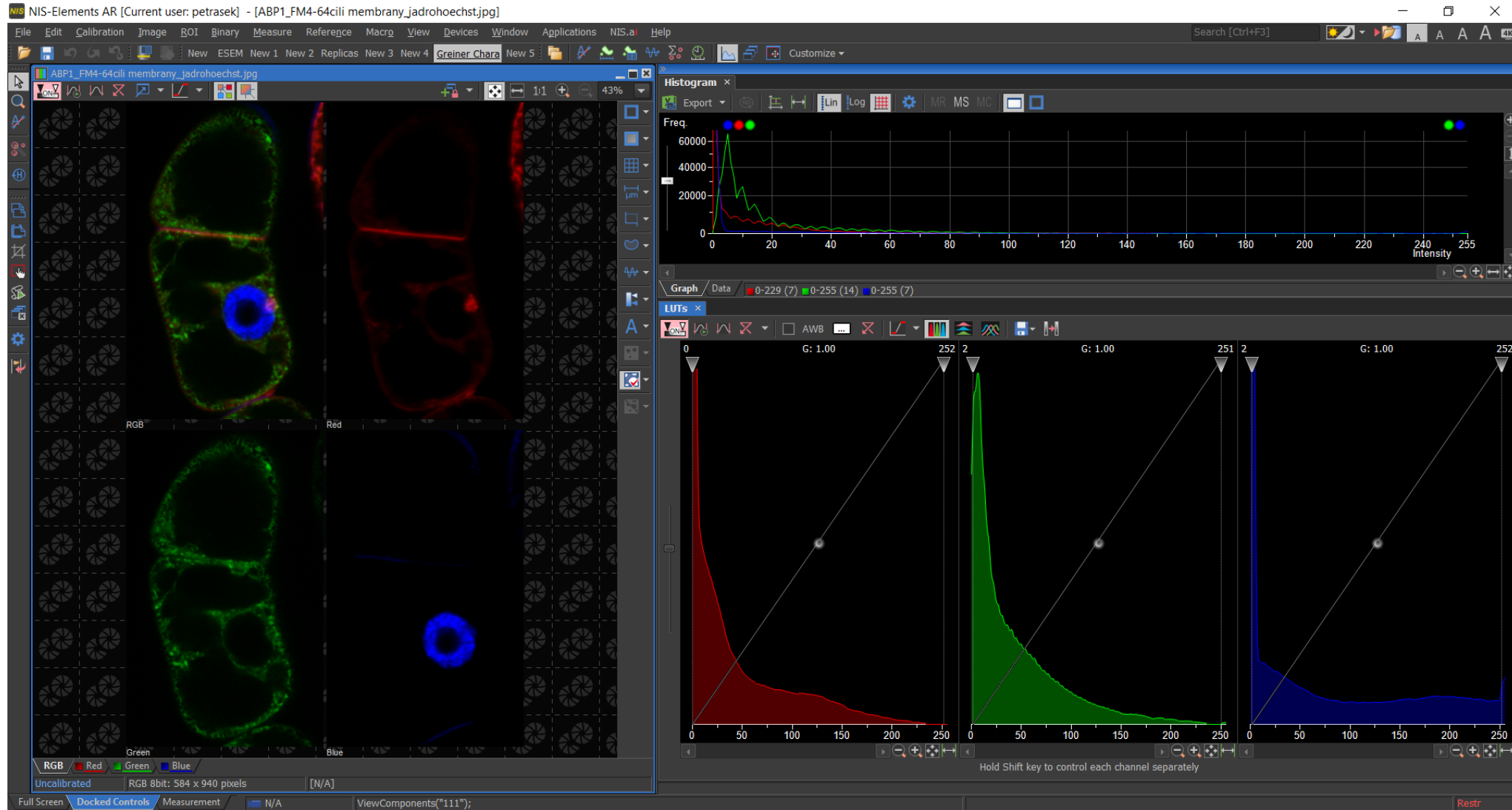
2.5 Image analysis, graphical software, presentation software

NIS Elements – good example of commercially available, comprehensive image analysis tool
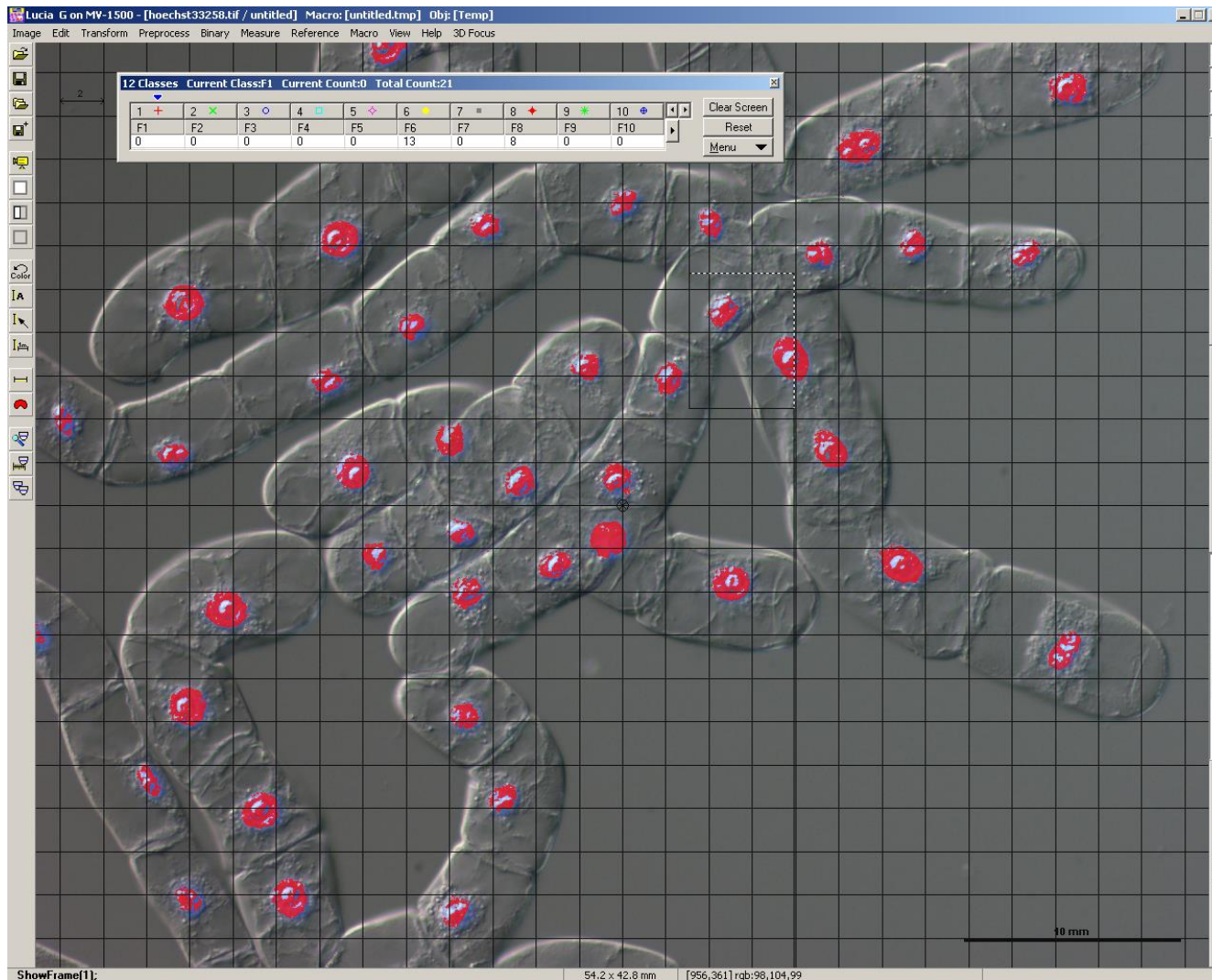
2.5 Image analysis, graphical software, presentation software

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

Leica

Zeiss

Nikon

Olympus
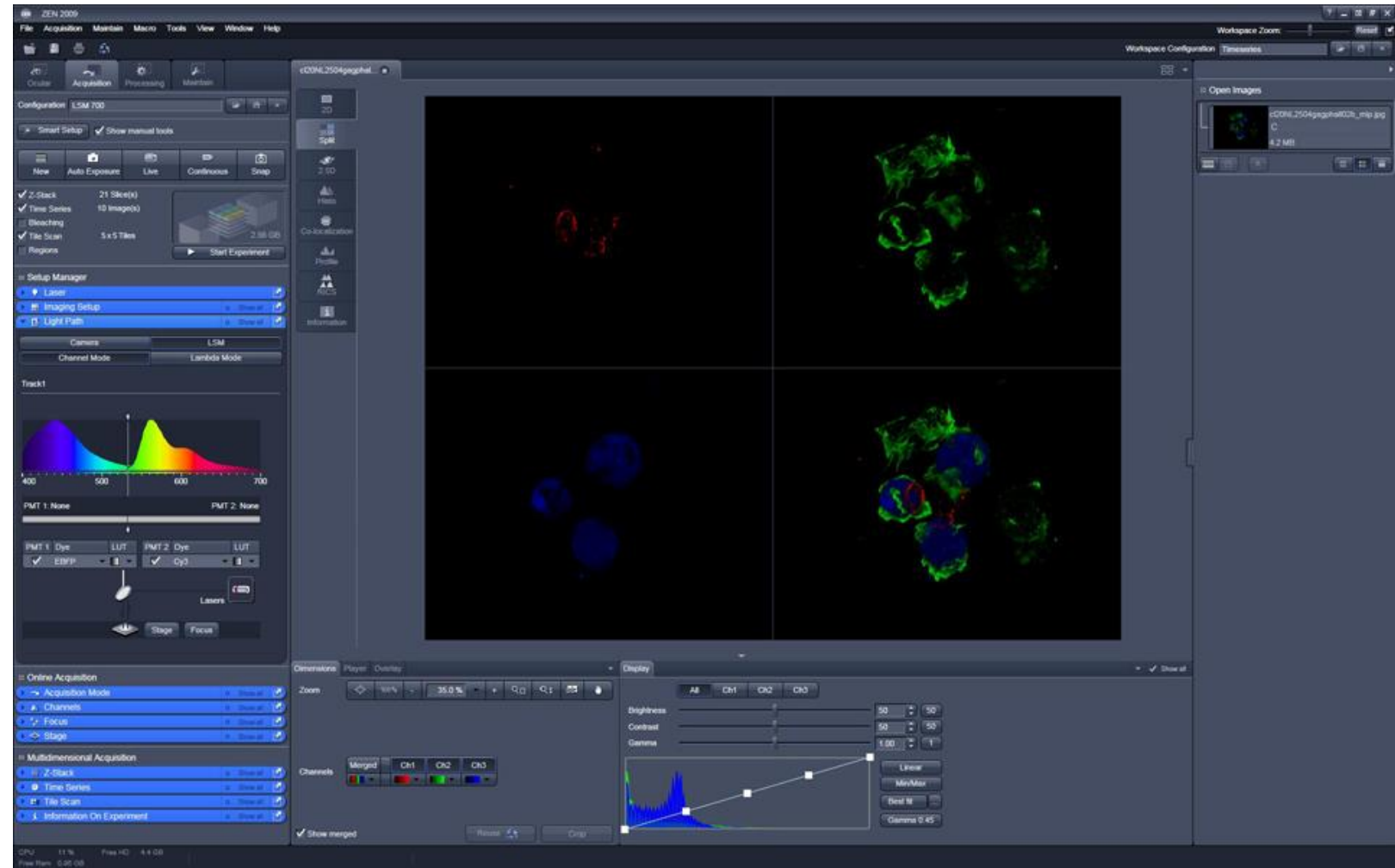
2.5 Image analysis, graphical software, presentation software

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. **L**ook **U**p **T**able,
is crucial for the correct presentation of the image
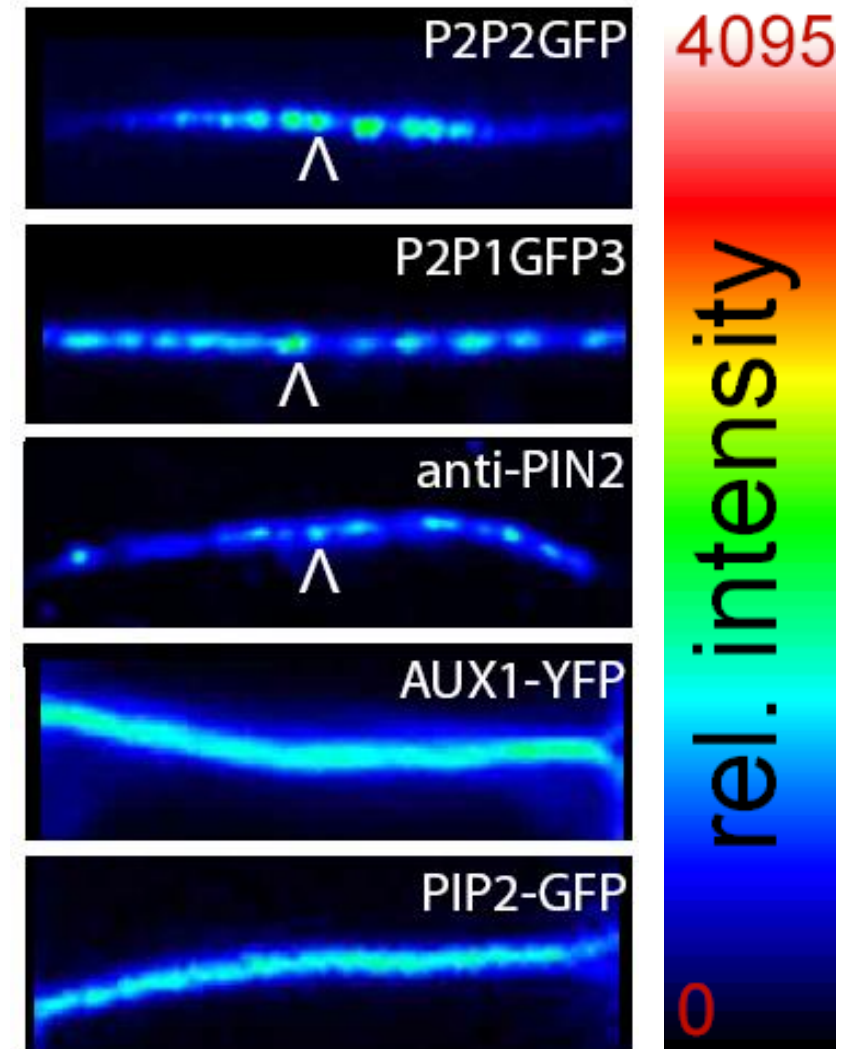
Kleine-Vehn, Mol Syst Biol 7:540, 2011

2.5 Image analysis, graphical software, presentation software
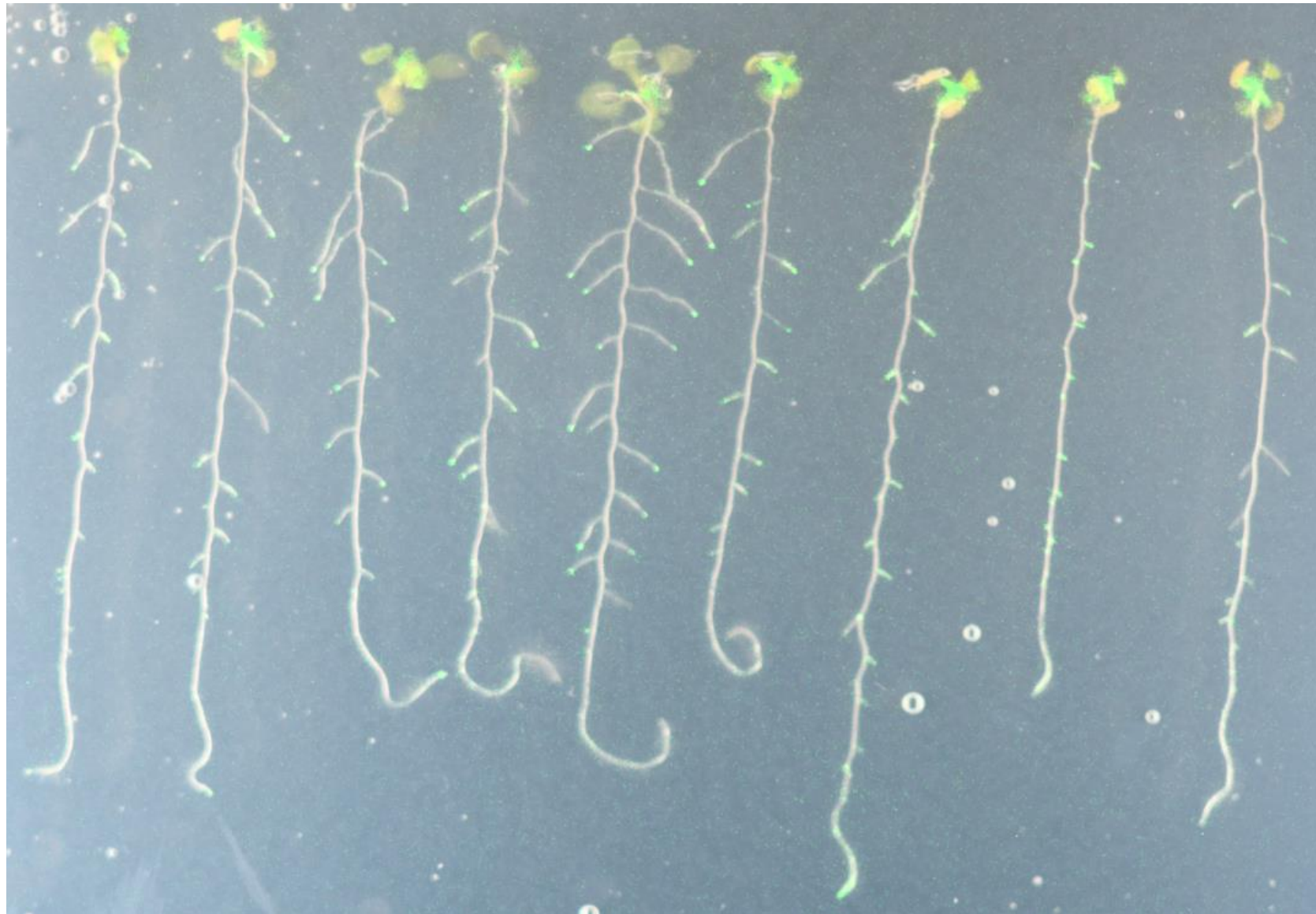
2.5 Image analysis, graphical software, presentation software

2.5 Image analysis, graphical software, presentation software

2.5 Image analysis, graphical software, presentation software
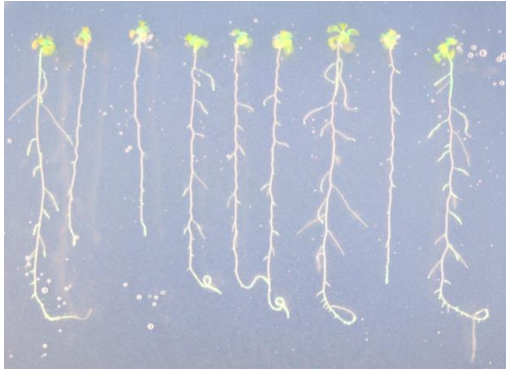
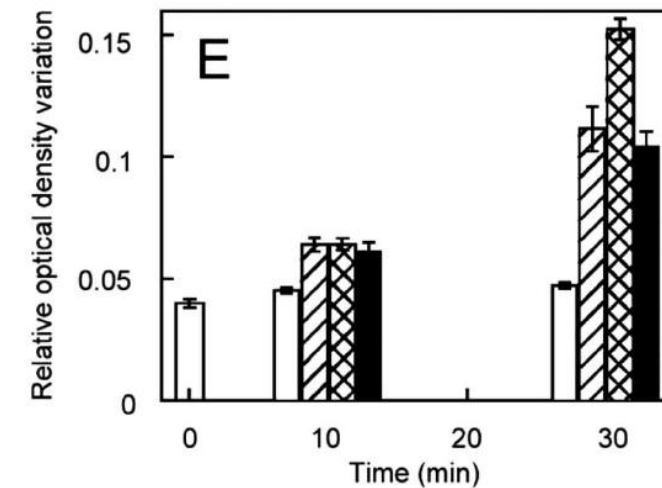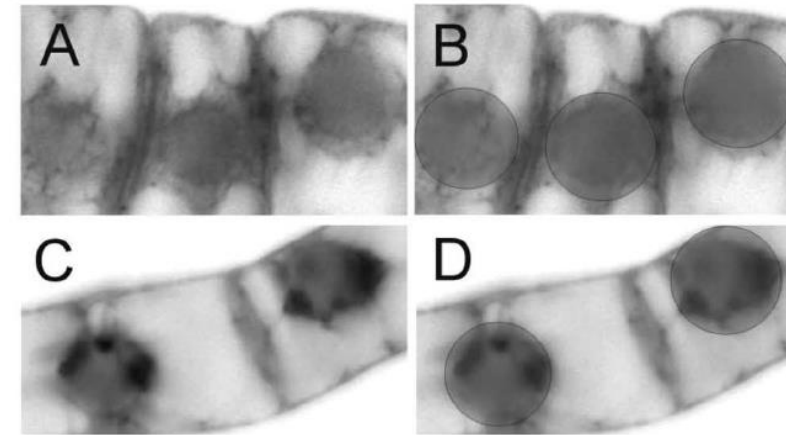# 2.5 Image analysis, graphical software, presentation software
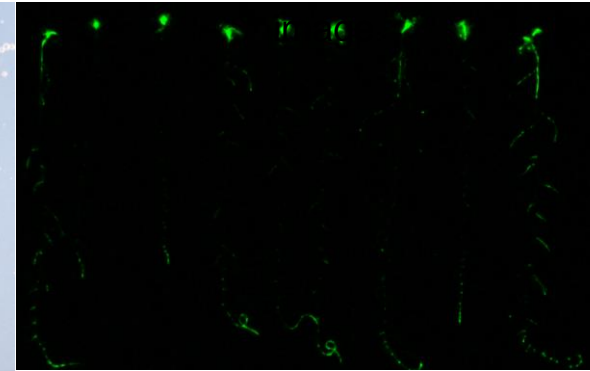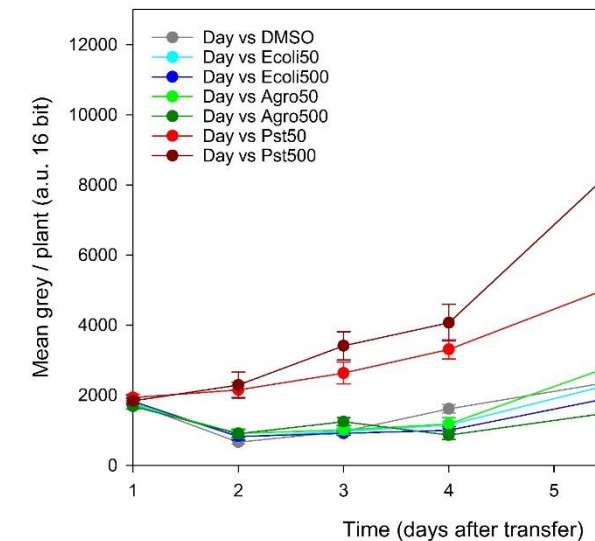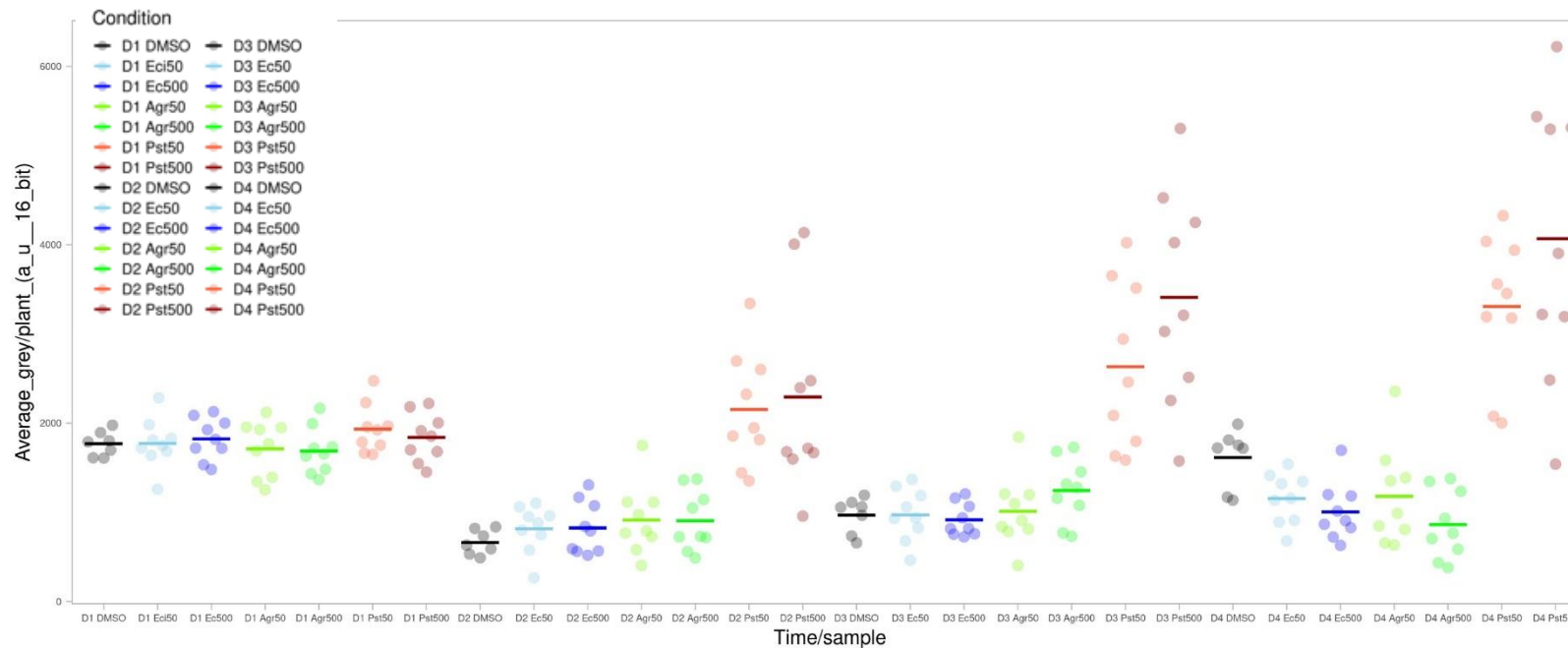
2.5 Image analysis, graphical software, presentation software

Analysis of colocalization

(c) FM 4-64 PIN1-GFP Merged

(e)

Intensity of FM 4-64 (red channel) vs. Intensity of PIN1::PIN1:GFP (green channel)
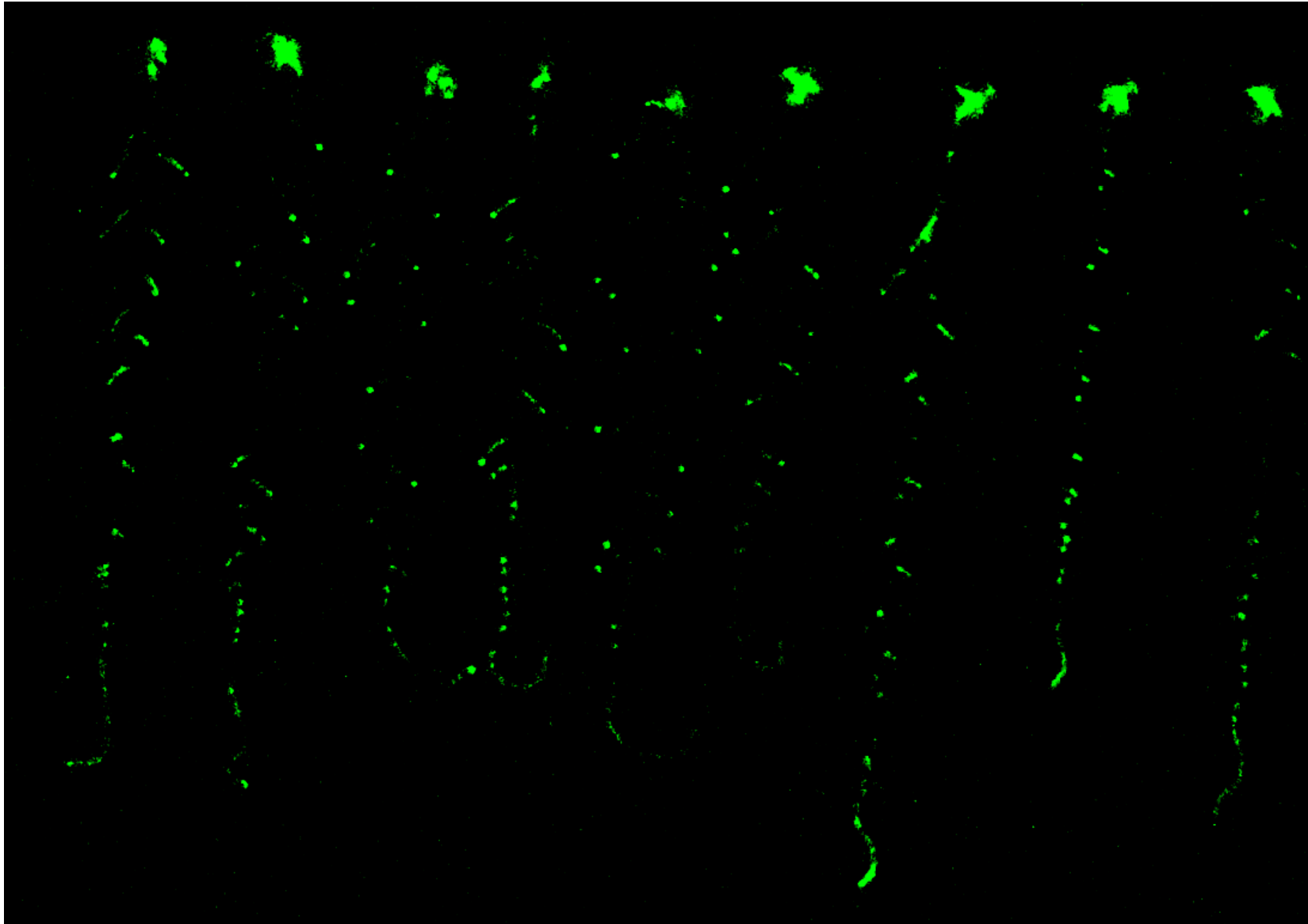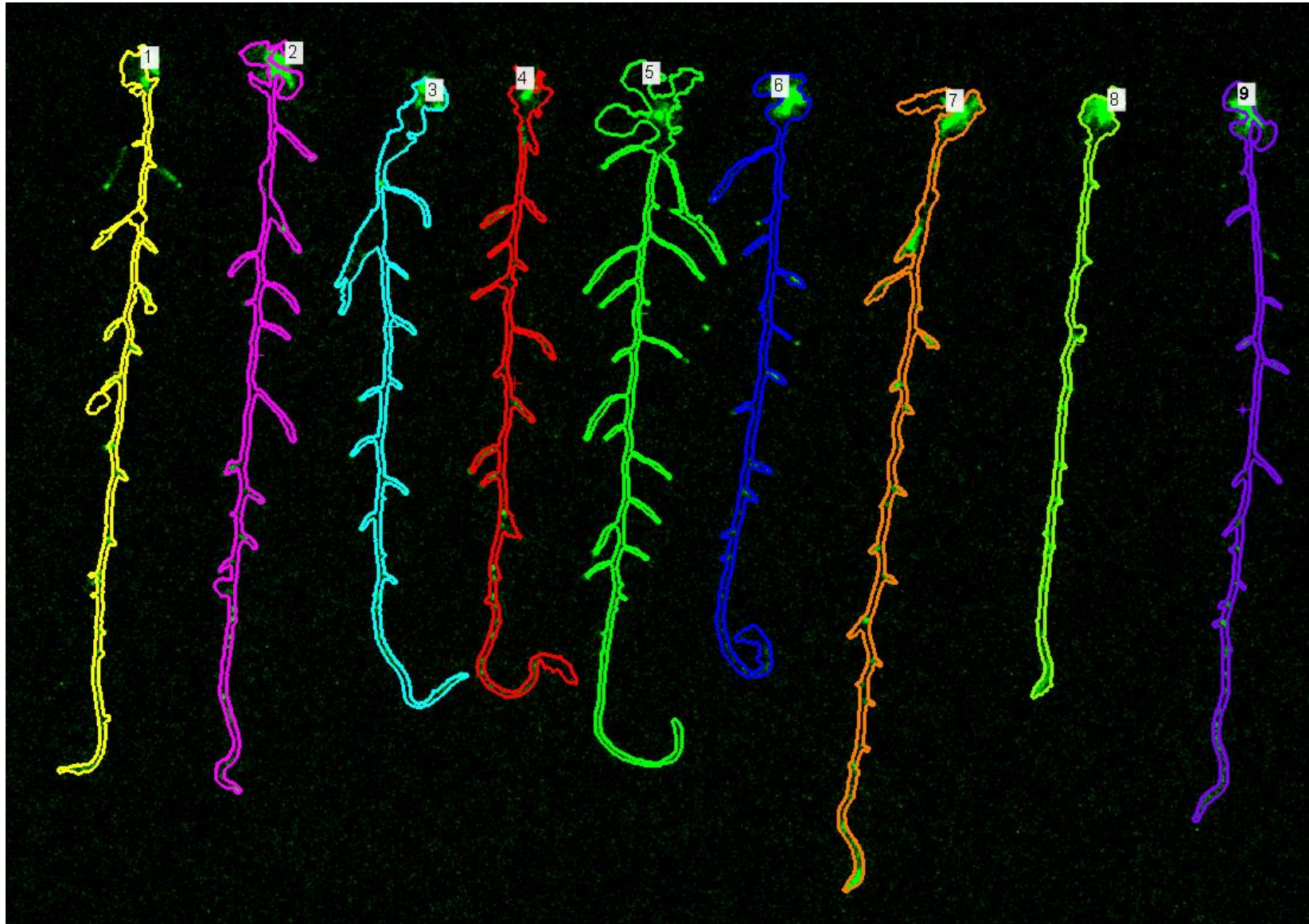
Absolute intensity
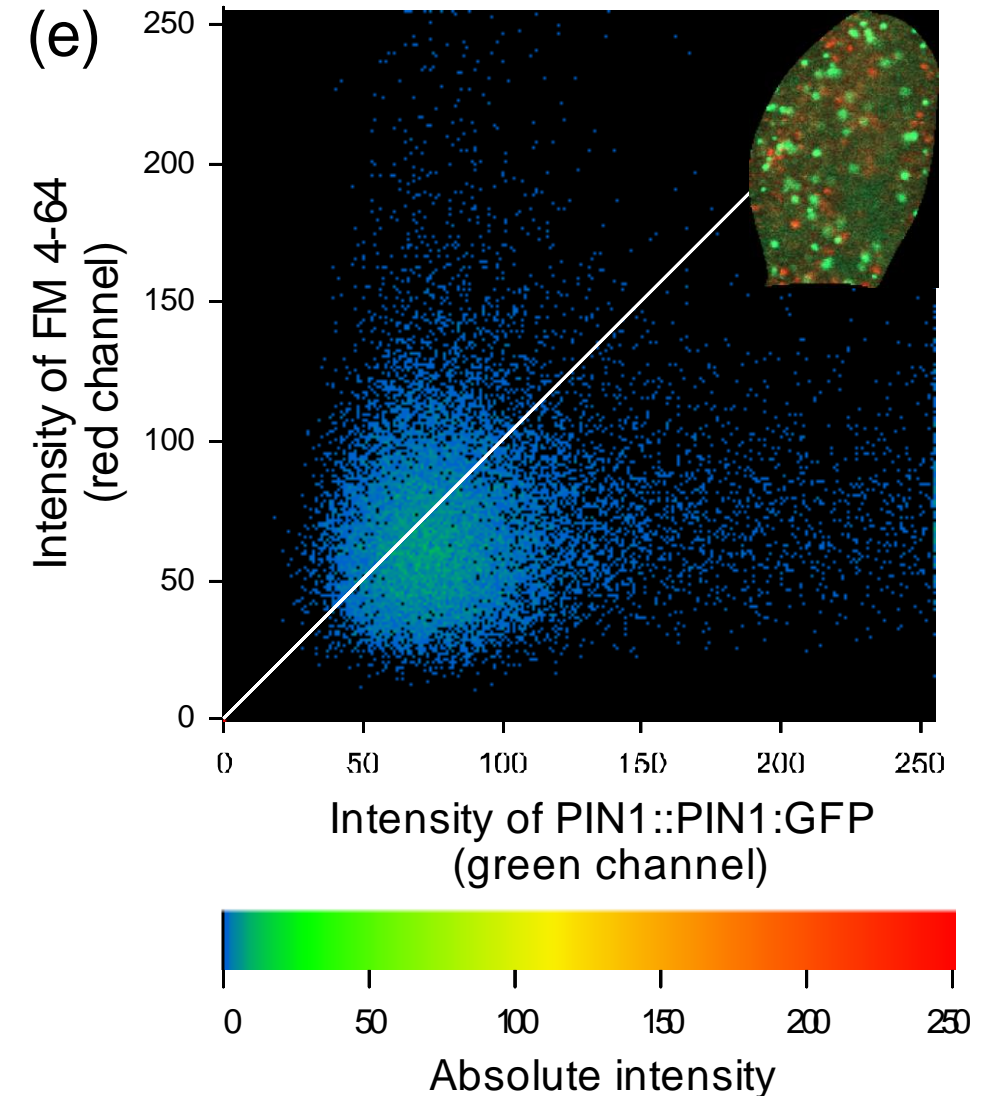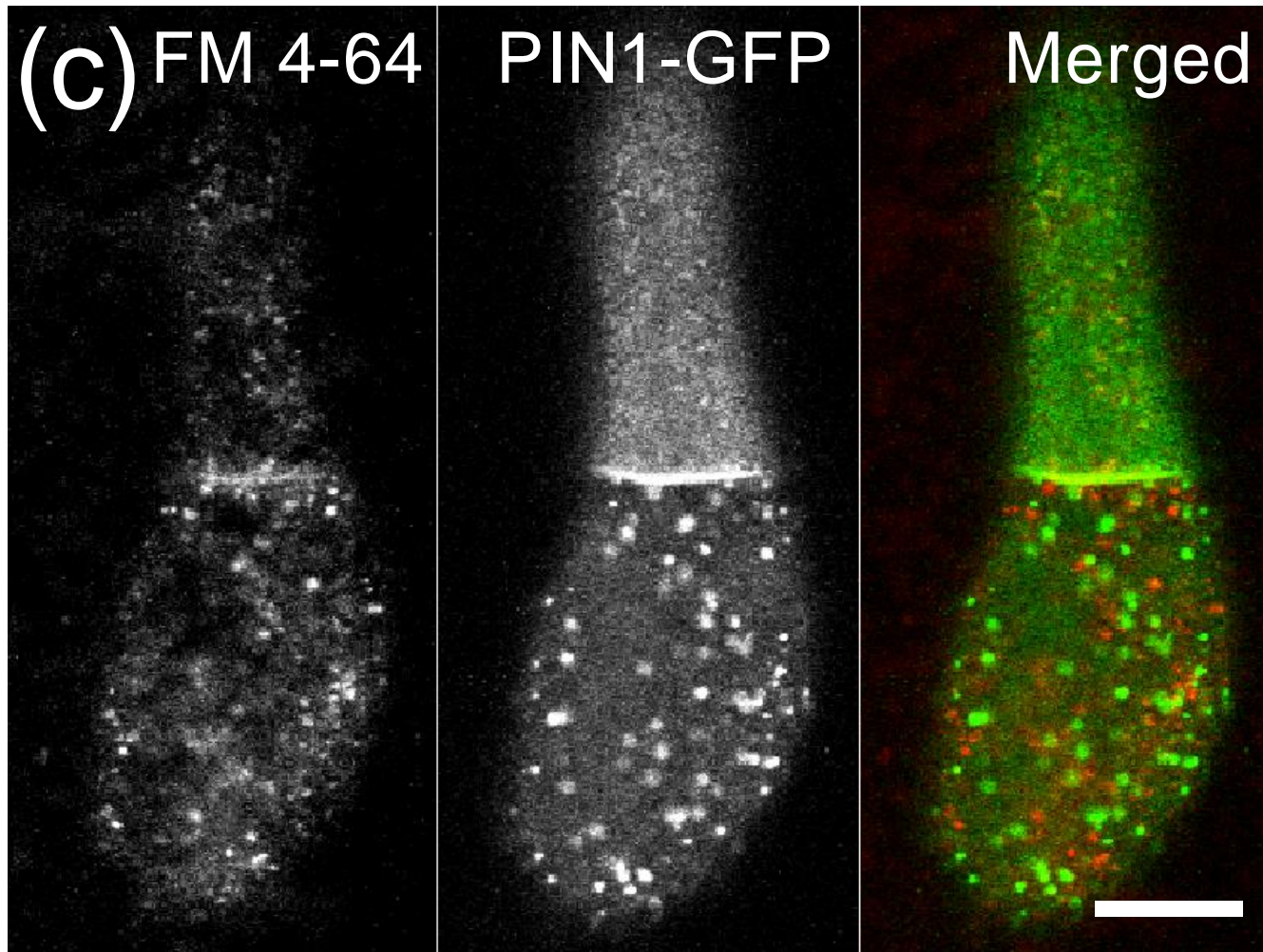
Jelínková et al., Plant Journal 61, 883-891, 2009

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
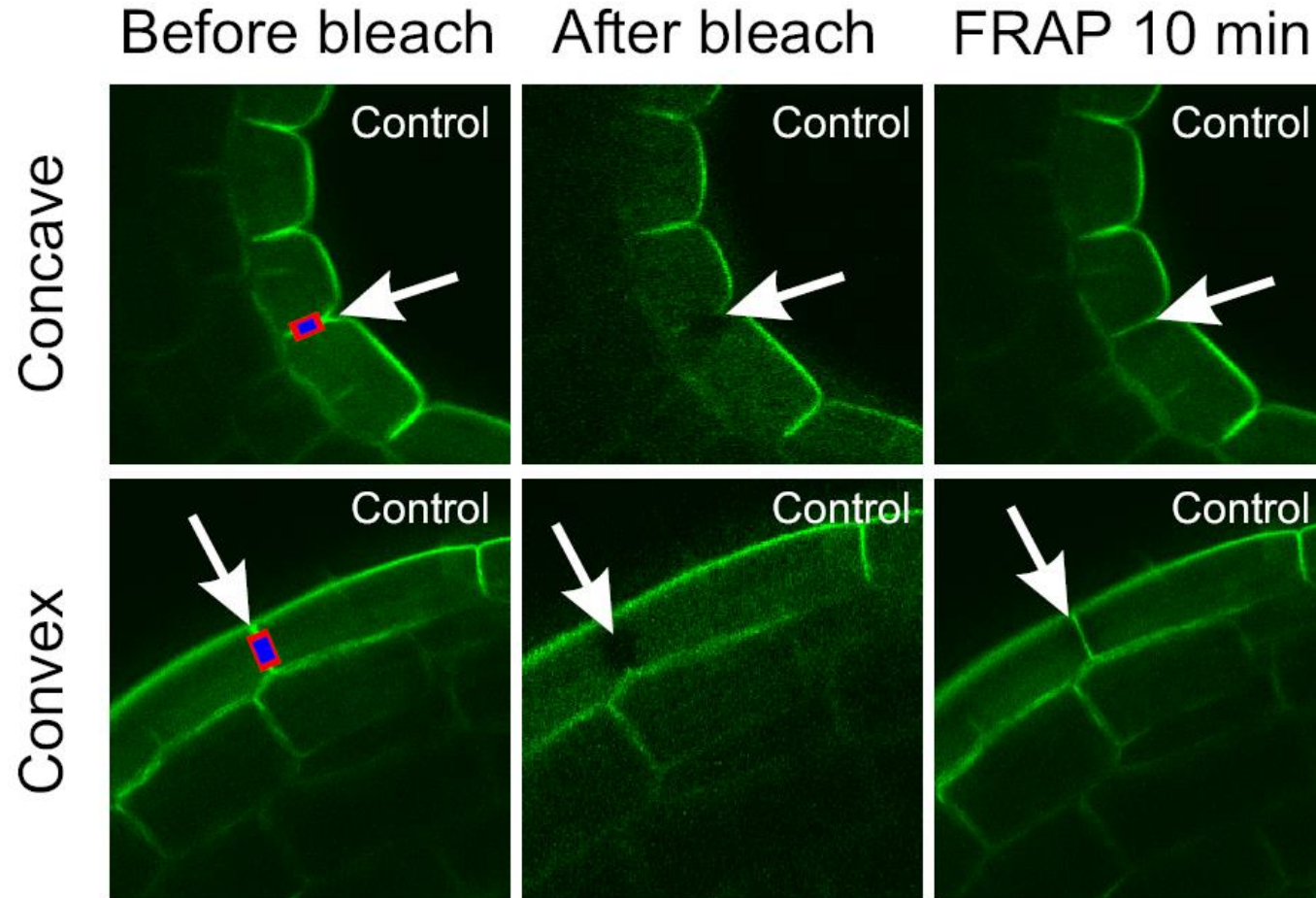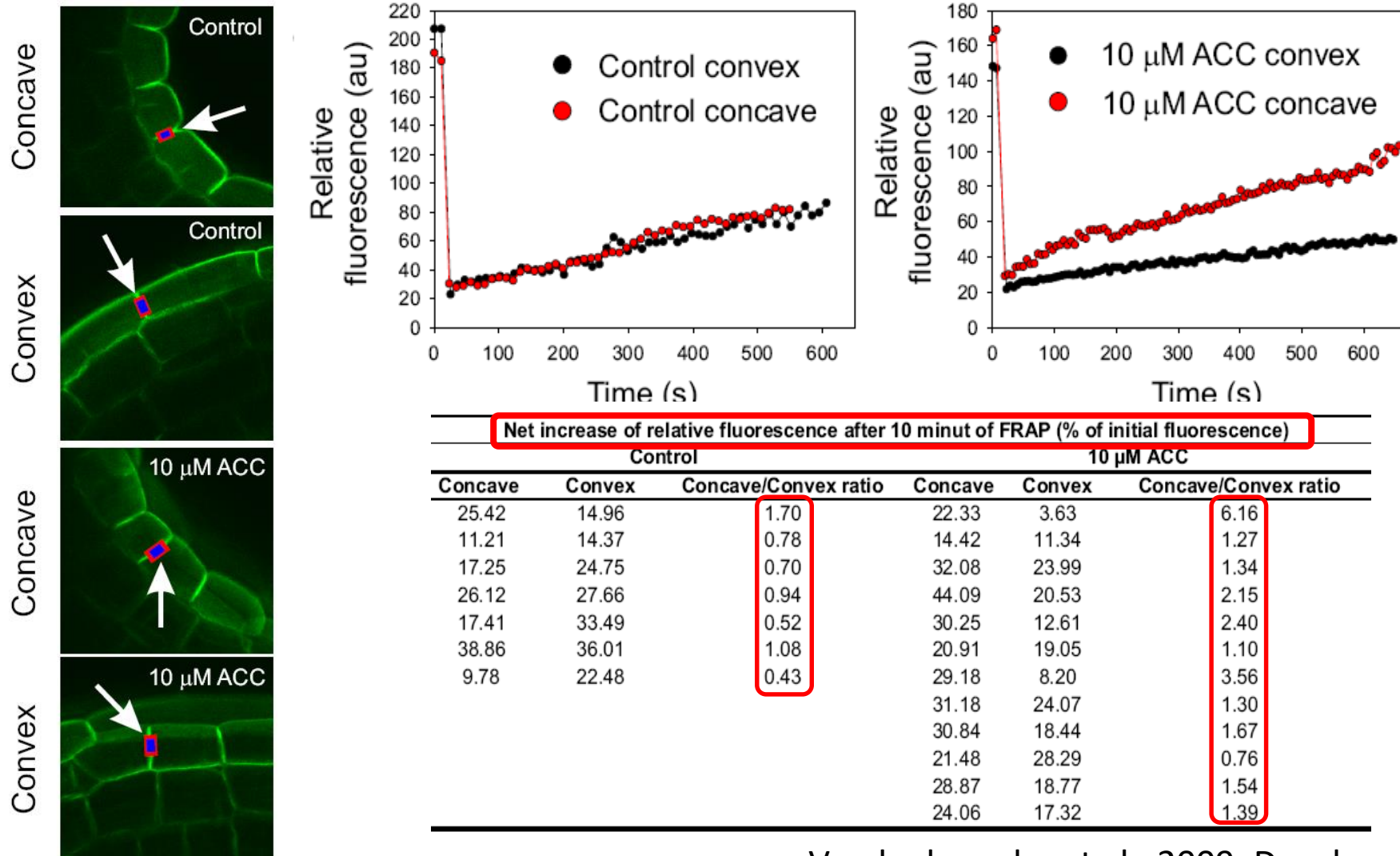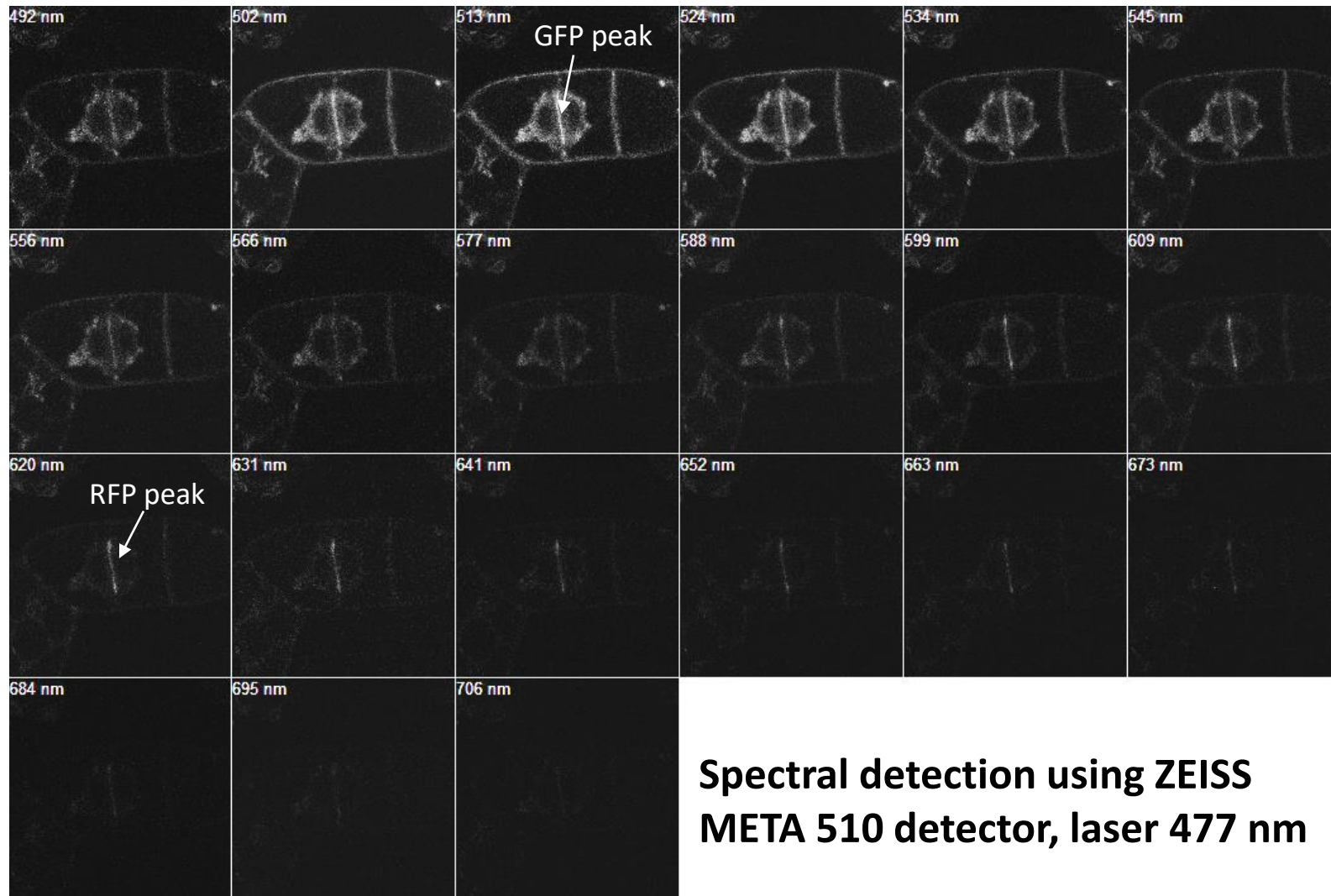
Analysis of intracellular protein dynamics

Net increase of relative fluorescence after 10 minut of FRAP (% of initial fluorescence)					
Control			10 µM ACC		
Concave	Convex	Concave/Convex ratio	Concave	Convex	Concave/Convex ratio
25.42	14.96	1.70	22.33	3.63	6.16
11.21	14.37	0.78	14.42	11.34	1.27
17.25	24.75	0.70	32.08	23.99	1.34
26.12	27.66	0.94	44.09	20.53	2.15
17.41	33.49	0.52	30.25	12.61	2.40
38.86	36.01	1.08	20.91	19.05	1.10
9.78	22.48	0.43	29.18	8.20	3.56
			31.18	24.07	1.30
			30.84	18.44	1.67
			21.48	28.29	0.76
			28.87	18.77	1.54
			24.06	17.32	1.39

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)

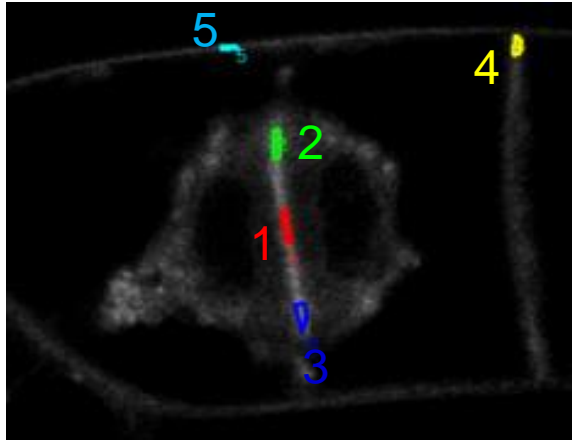**Spectral detection using ZEISS META 510 detector, laser 477 nm**

2.5 Image analysis, graphical software, presentation software

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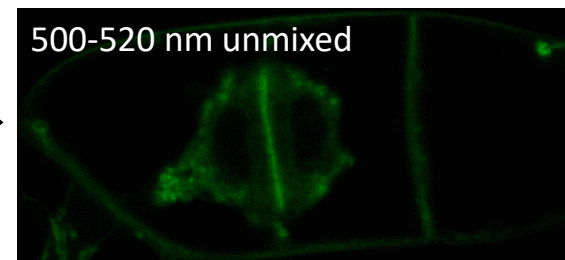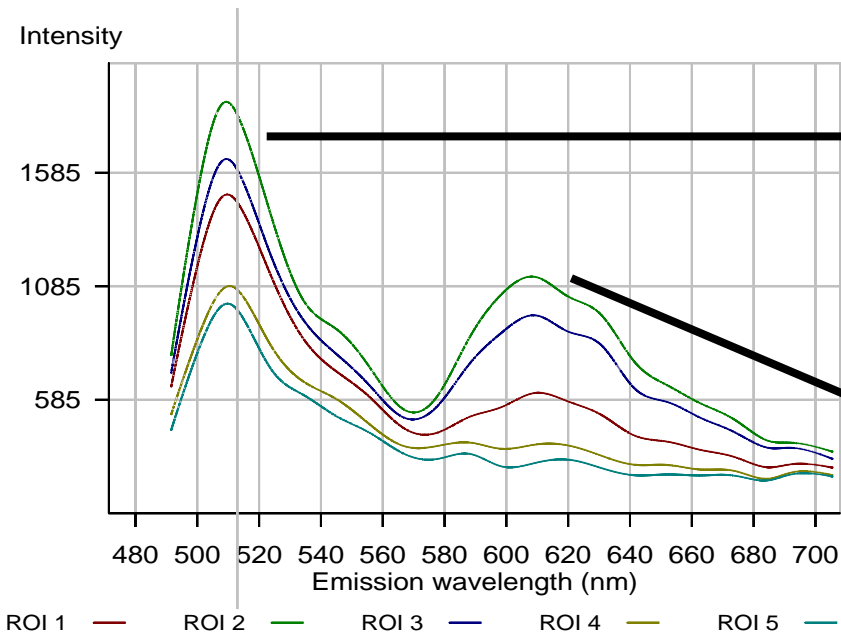

Average fluorescence intensity measured over the emission spectrum in the following region of interests:
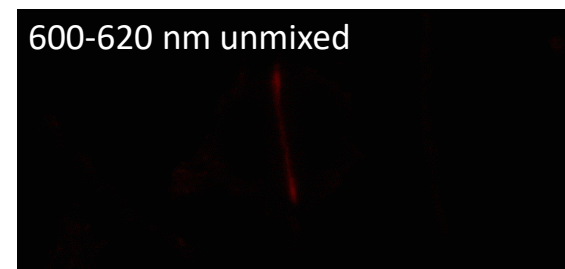
ROI1 - middle cell plate

ROI2, ROI3 - growing ends of the cell plate

ROI4, ROI5 - transversal and longitudinal plasma membranes

500-520 nm unmixed

600-620 nm unmixed

ROI1 - low FRET

ROI2, ROI3 - high FRET

ROI4, ROI5 - no FRET

2.5 Image analysis, graphical software, presentation software

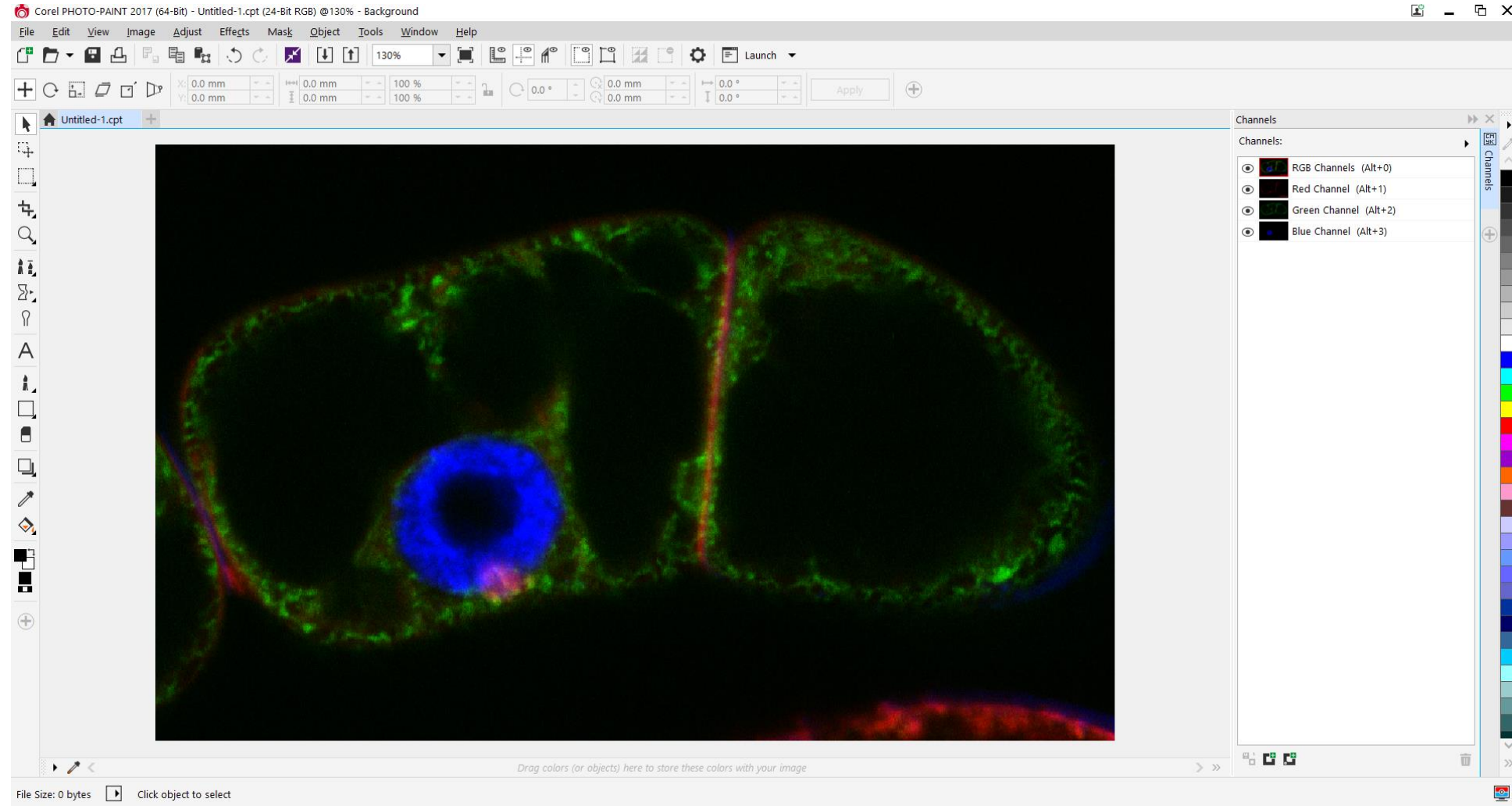Irfan View - the ideal tool for daily viewing of experimental documentation

2.5 Image analysis, graphical software, presentation software

Raster graphics a its editing using software editors
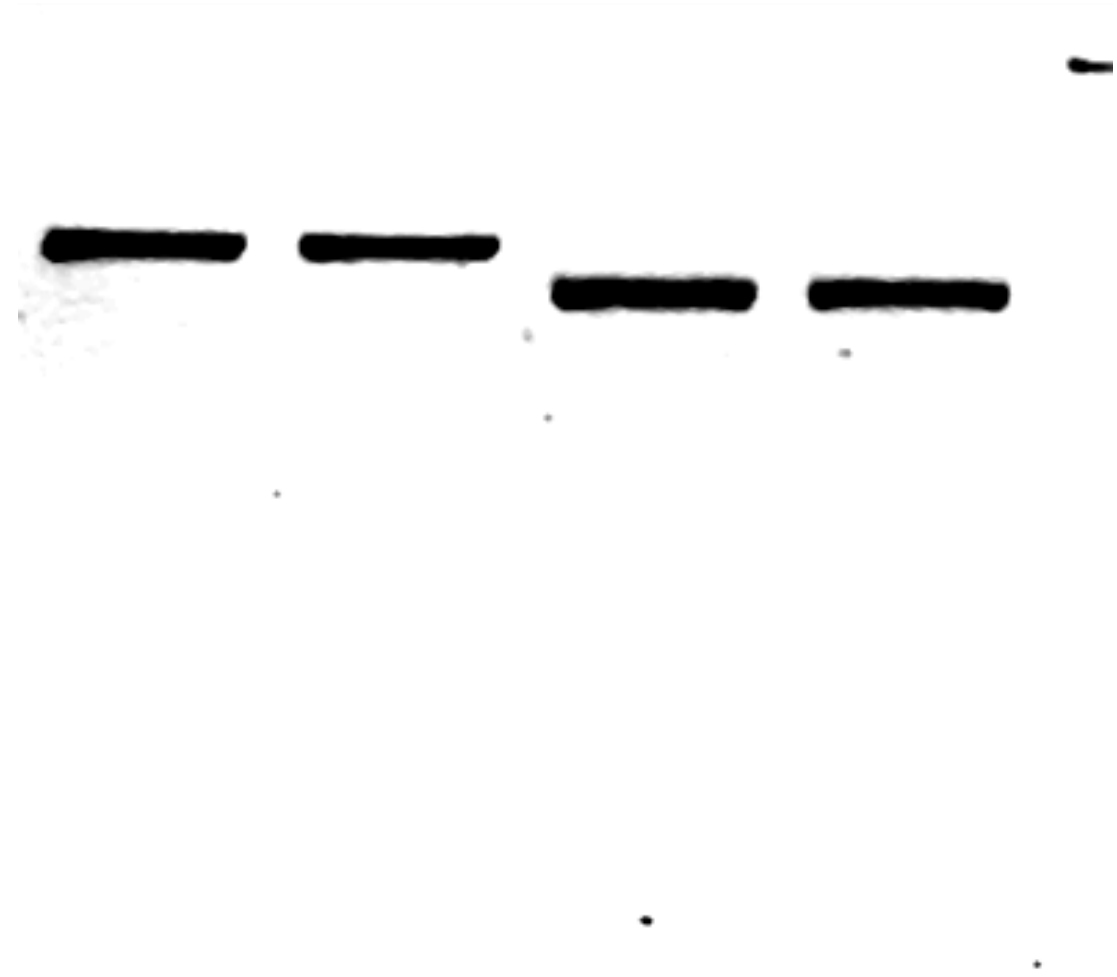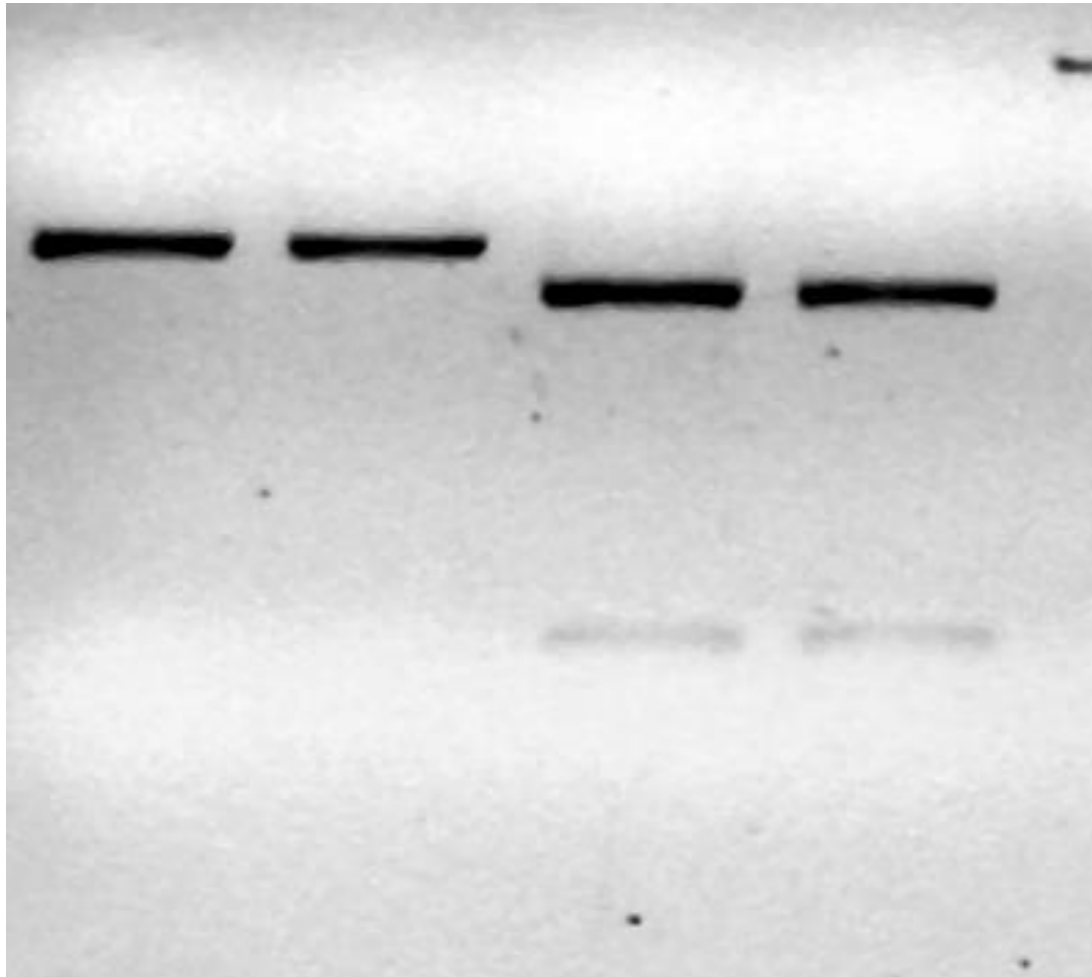
Corel Photo Paint

Adobe Photoshop

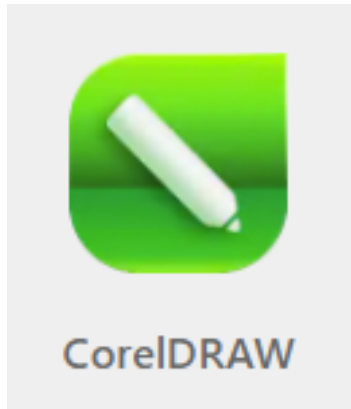# 2.5 Image analysis, graphical software, presentation software

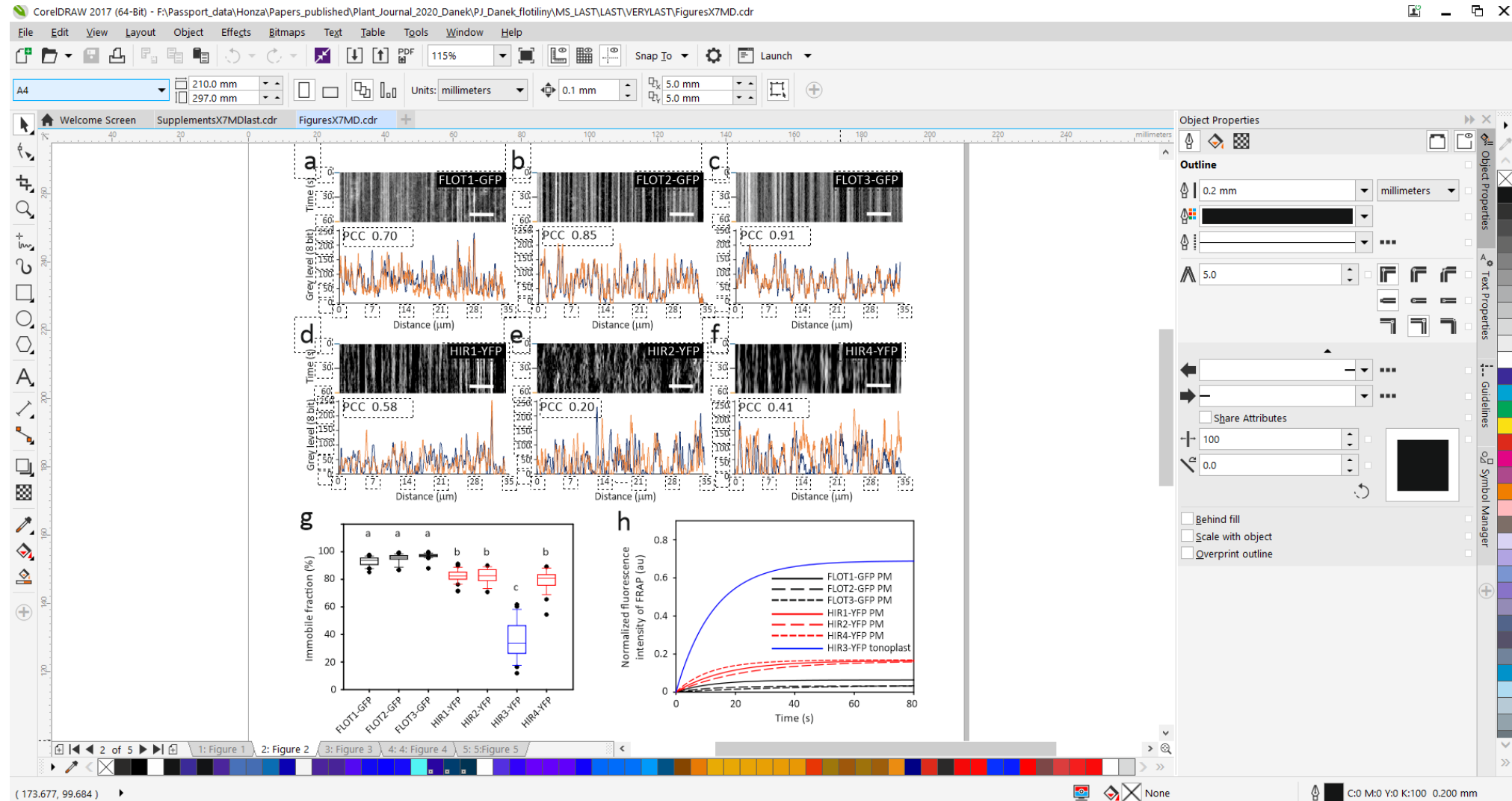## Beware of too extensive modifications!!!!

2.5 Image analysis, graphical software, presentation software

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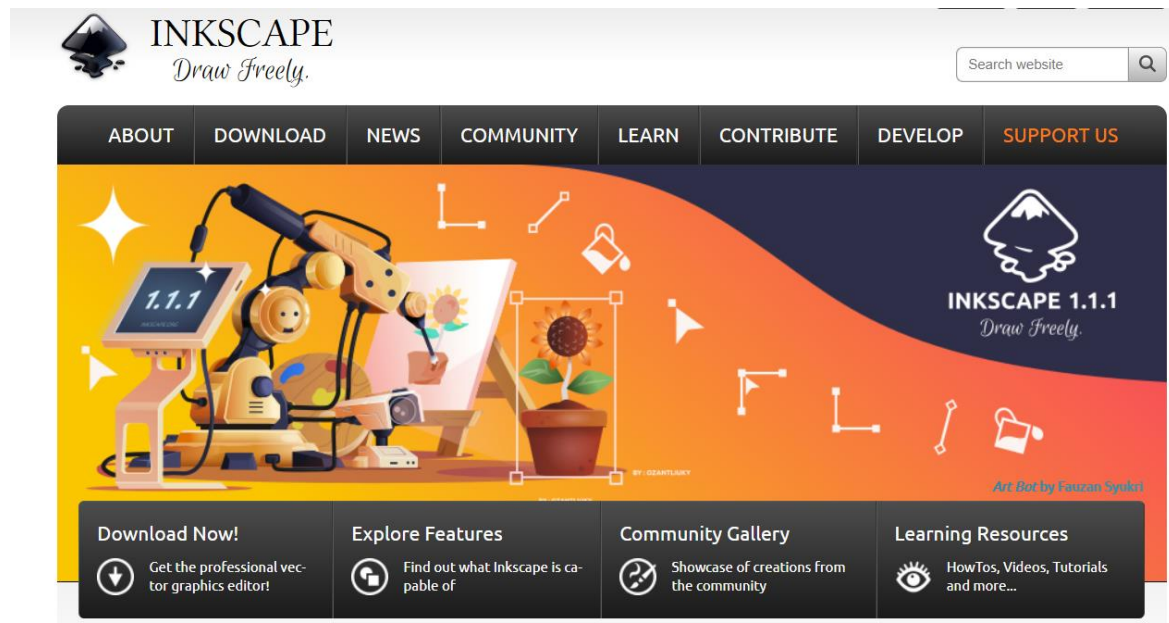
Corel Draw

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
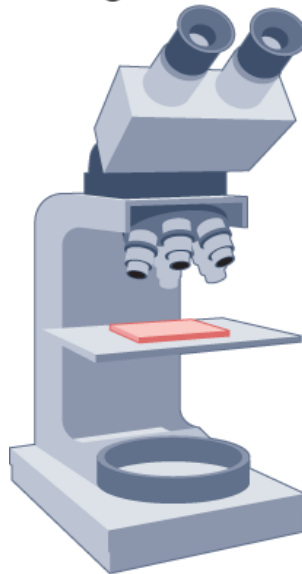Data is well-described and provides clear usage of licences

FAIR Principles - GO FAIR

FAIR principles in bioimaging

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