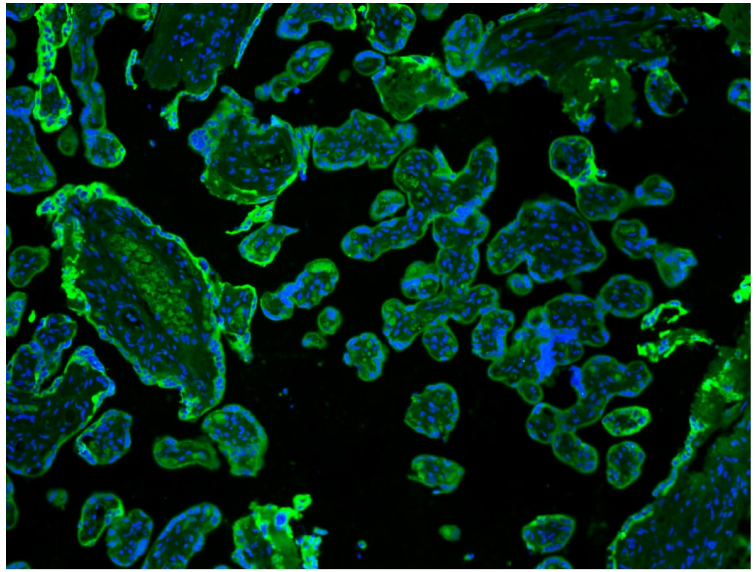
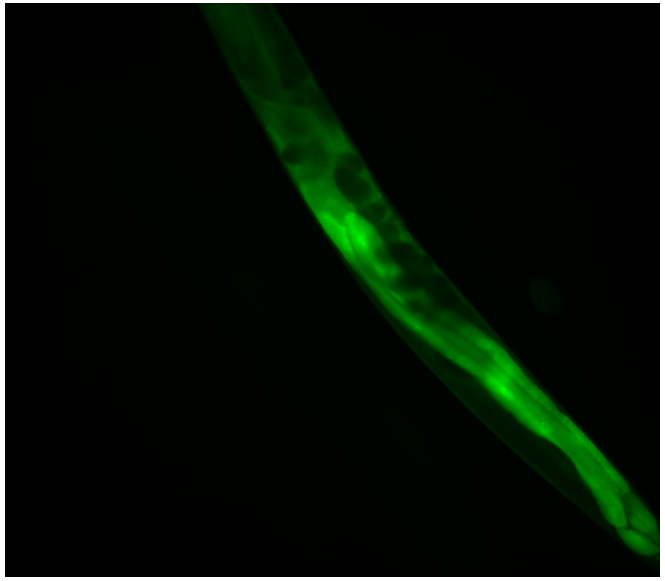
Image Data Analysis in H.sapiens and C.elegans





Dr. Alexander K. Seewald

SEEWALD SOLUTIONS

(Our) Relevant Collaborations for Image Data Analysis

2009-2012
Med.Univ.Vienna, Austria
Tissue Gnostics GmbH
Funded by FFG Bridge










2007-2010
Univ. Colorado@Boulder, USA
Self-funded





2011-
IMP Vienna

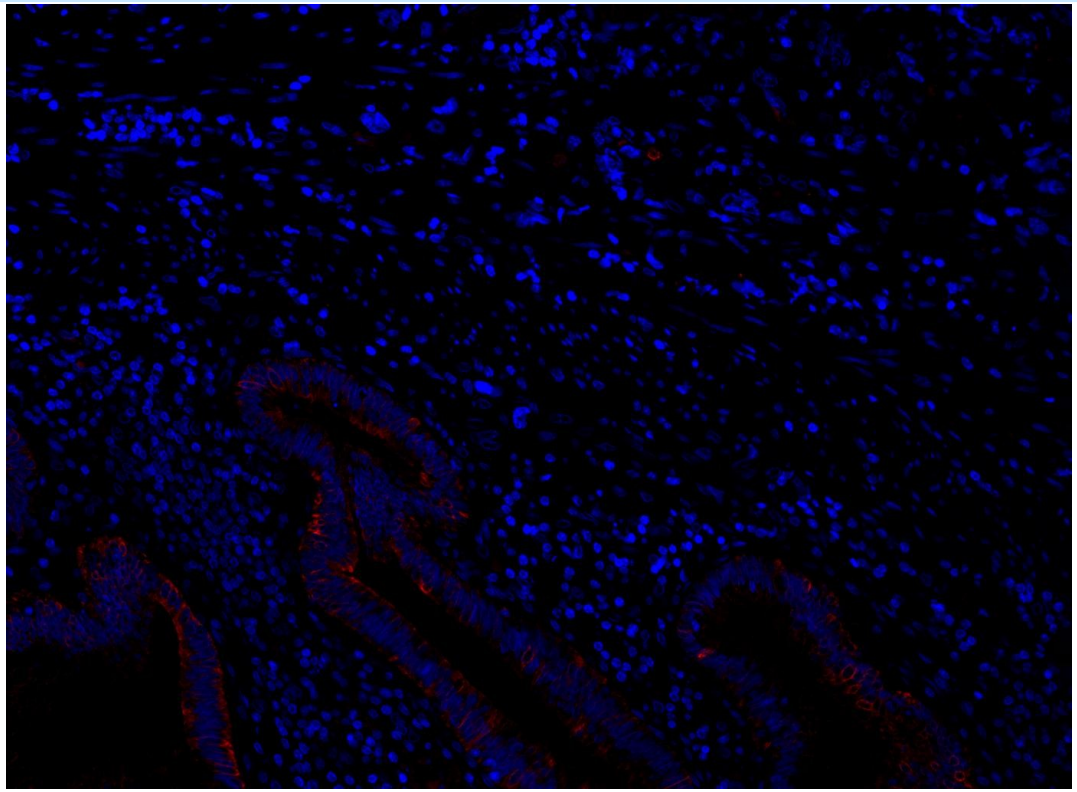

alex@seewald.at
www.seewald.at

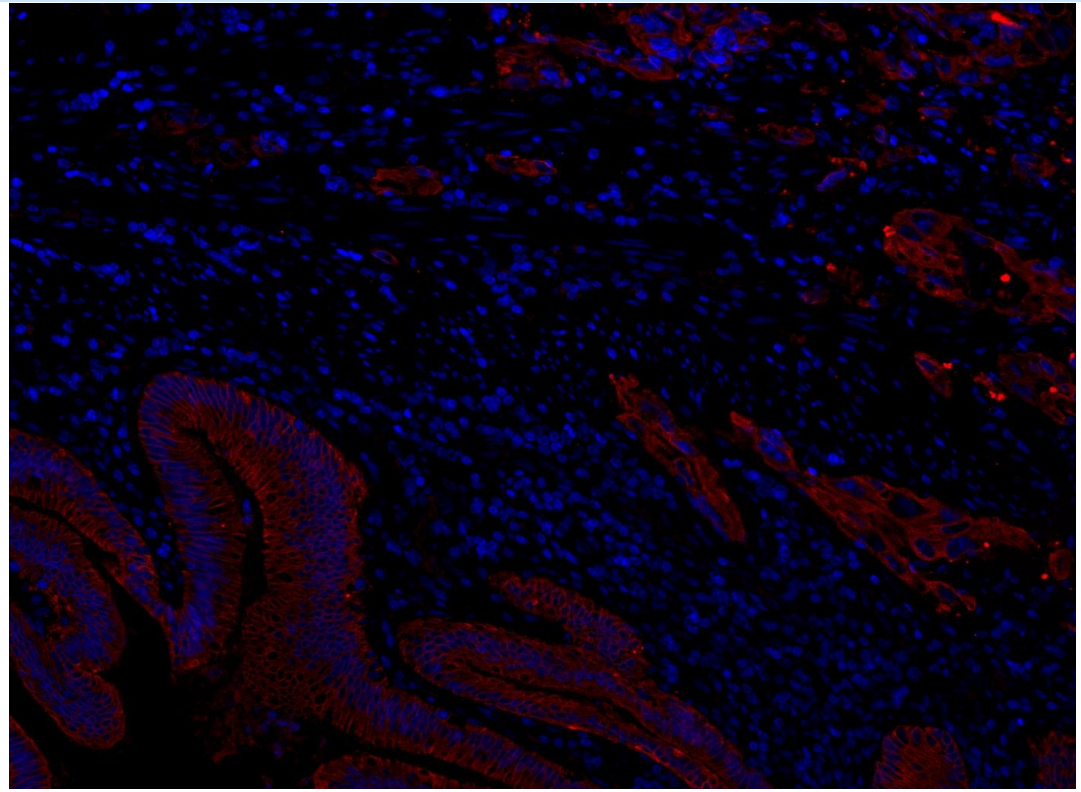
Video 1 - How Tissue Samples Are Created

SEEWALD SOLUTIONS

alex@seewald.at
www.seewald.at

Staining tissue is a complex process!





- Tissue characteristics (thickness, density, age, photobleaching, ...)
- Quality of antibodies (too old, too new, wrong kind, cross-reactions, ..)
- Crosstalk between multiple stainings, Protocols, MTA characteristics...

SEEWALD SOLUTIONS

alex@seewald.at
www.seewald.at

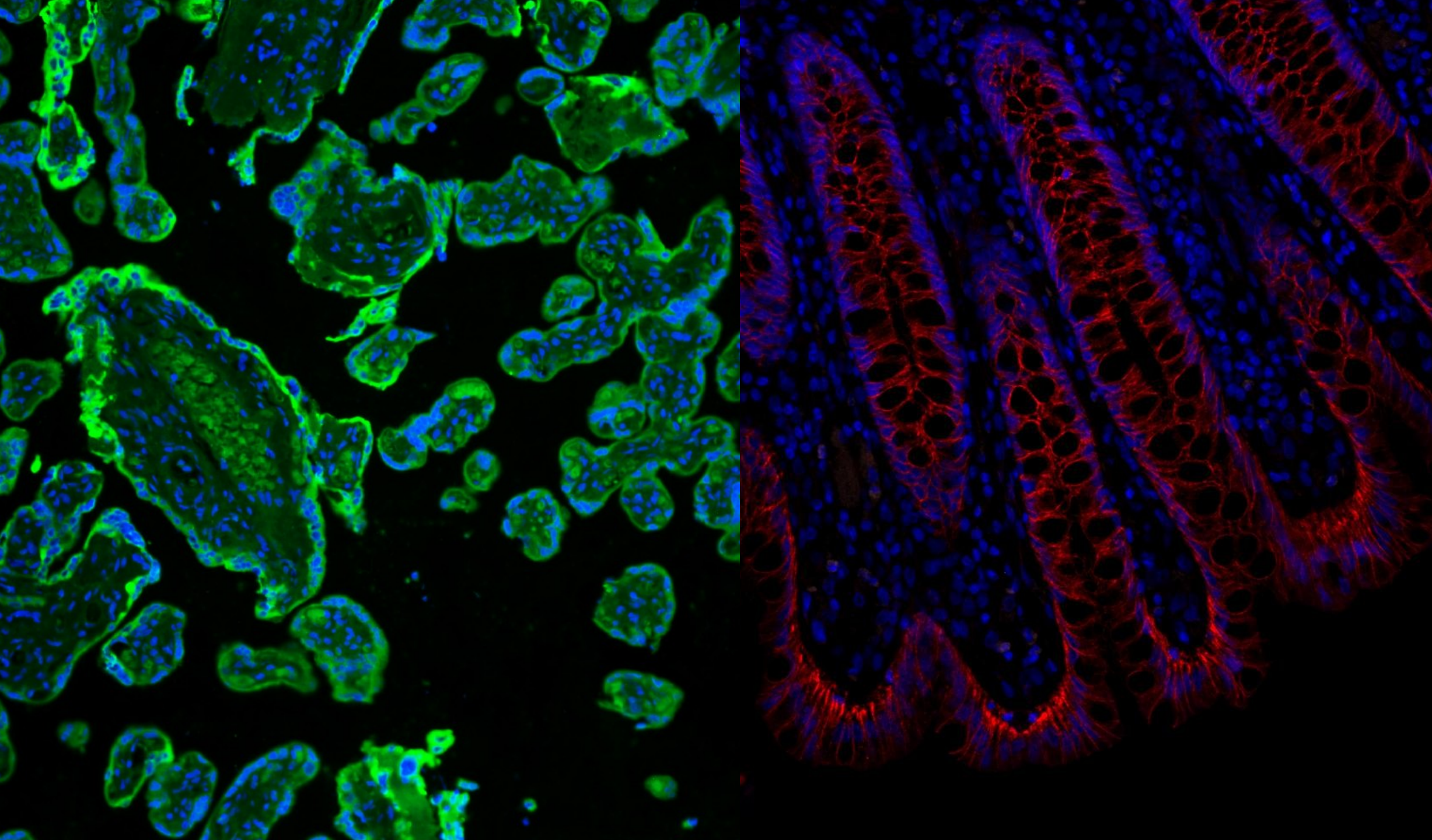
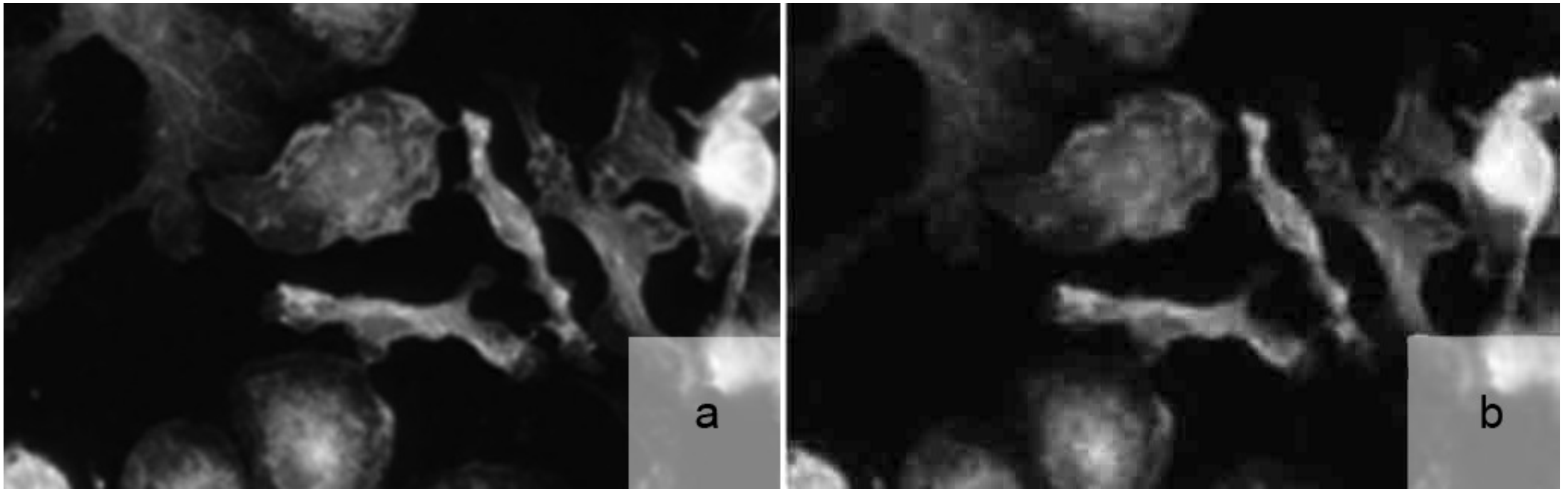


Image Formats

Use only lossless formats (TIFF, PNG)!

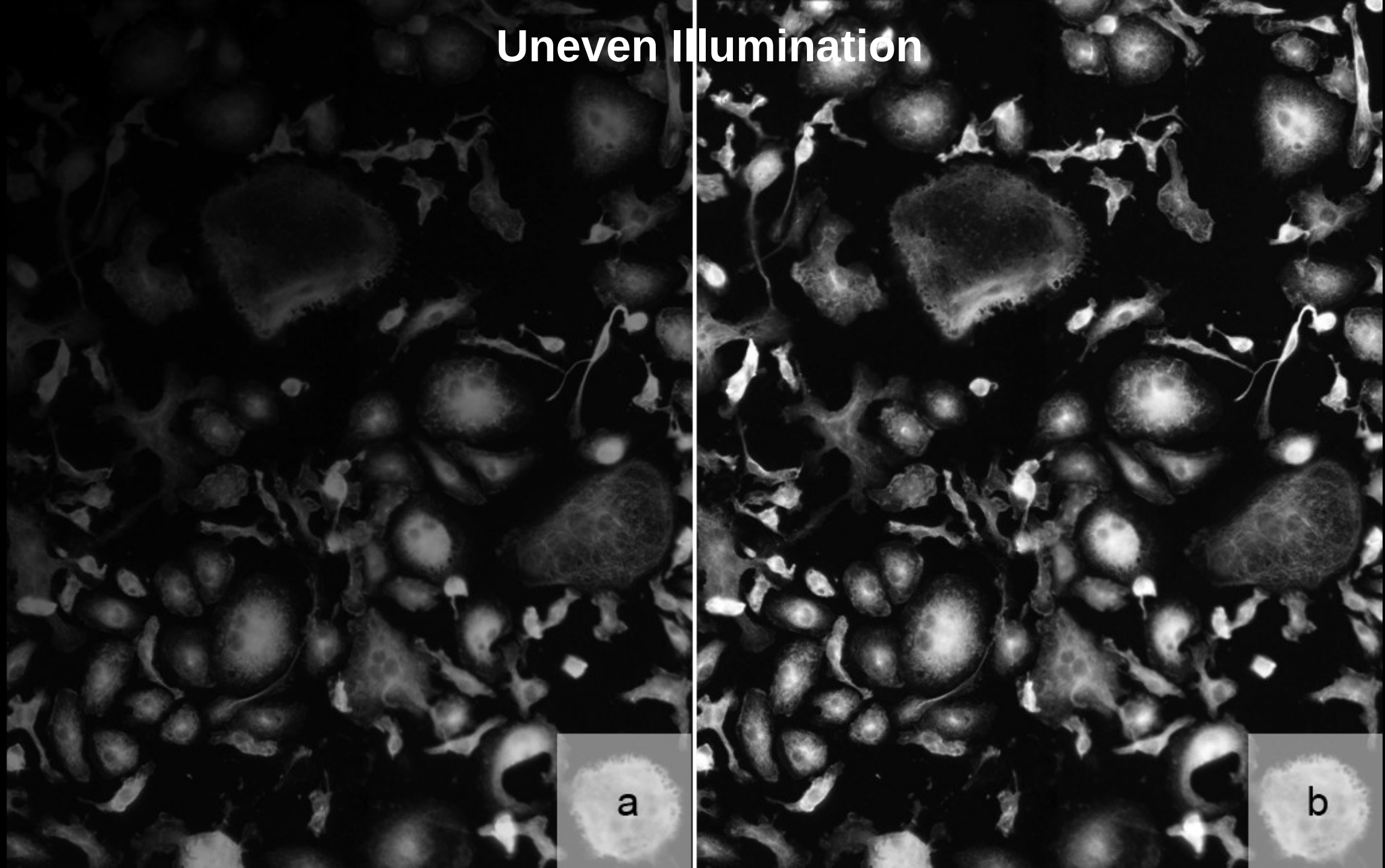


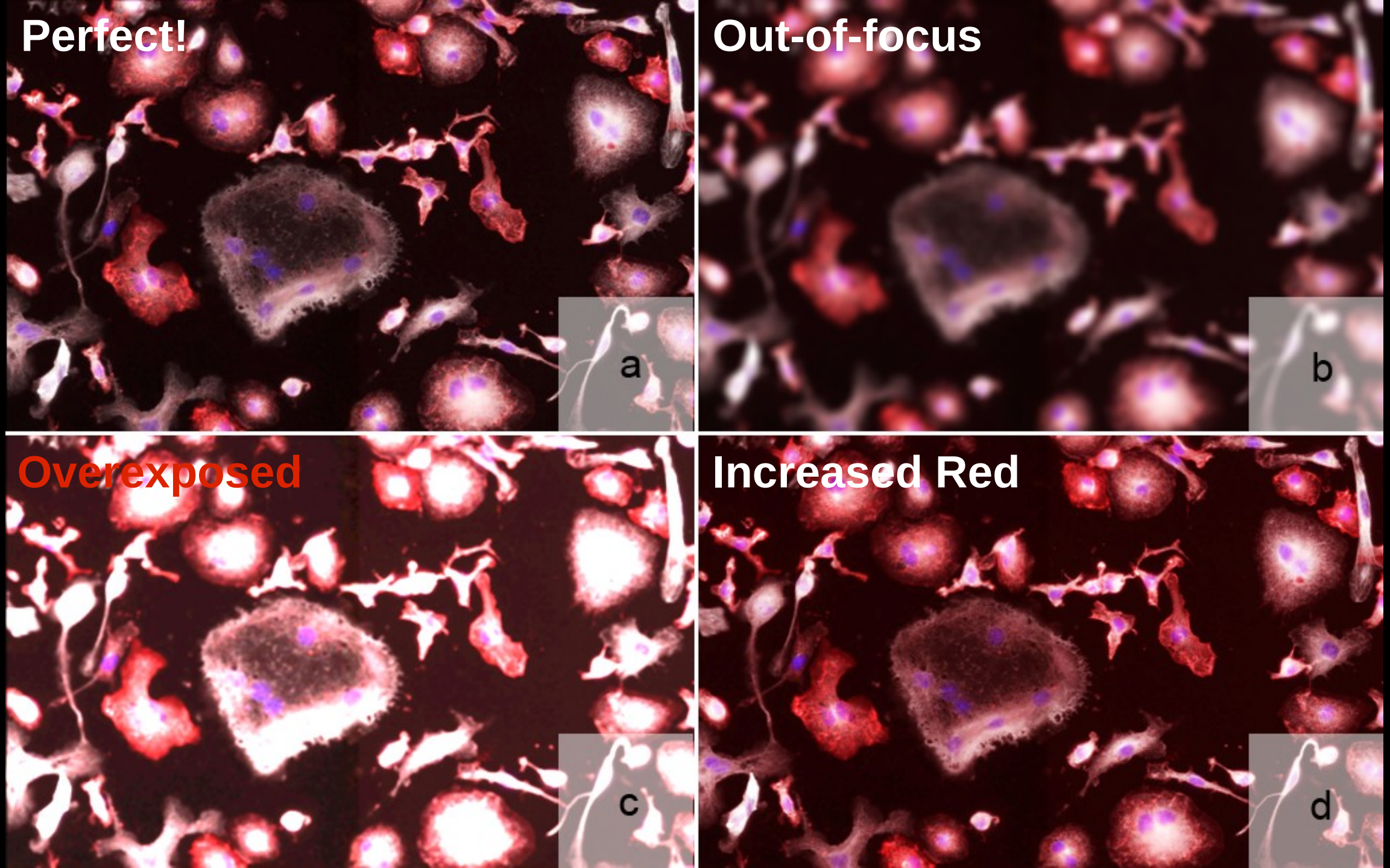
Especially JPEG is to be avoided! Removes information not visible to human eye which is still useful to image processing, introduces artefacts... **Also: No normalization of input images!**

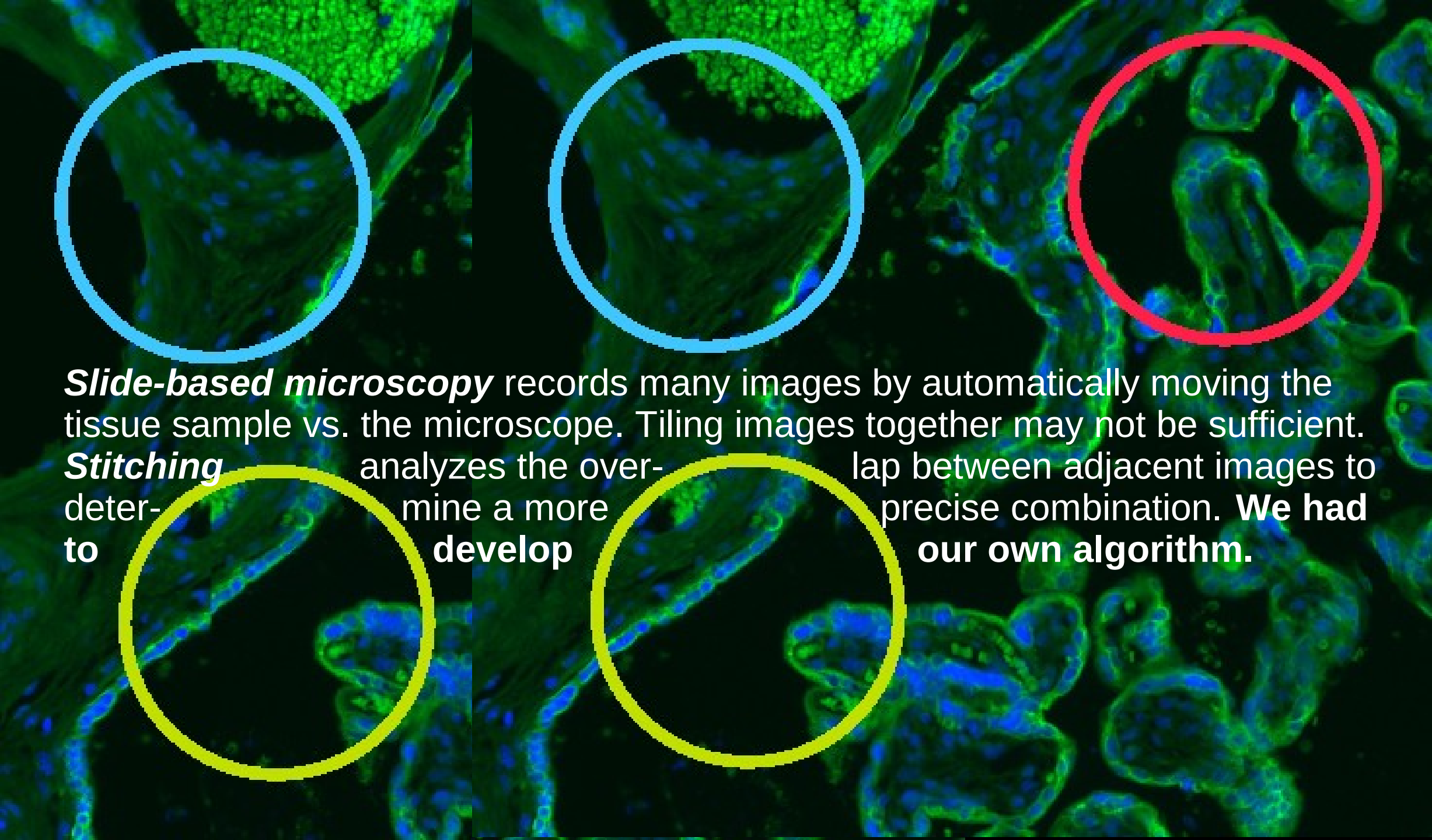
SEEWALD SOLUTIONS

alex@seewald.at
www.seewald.at

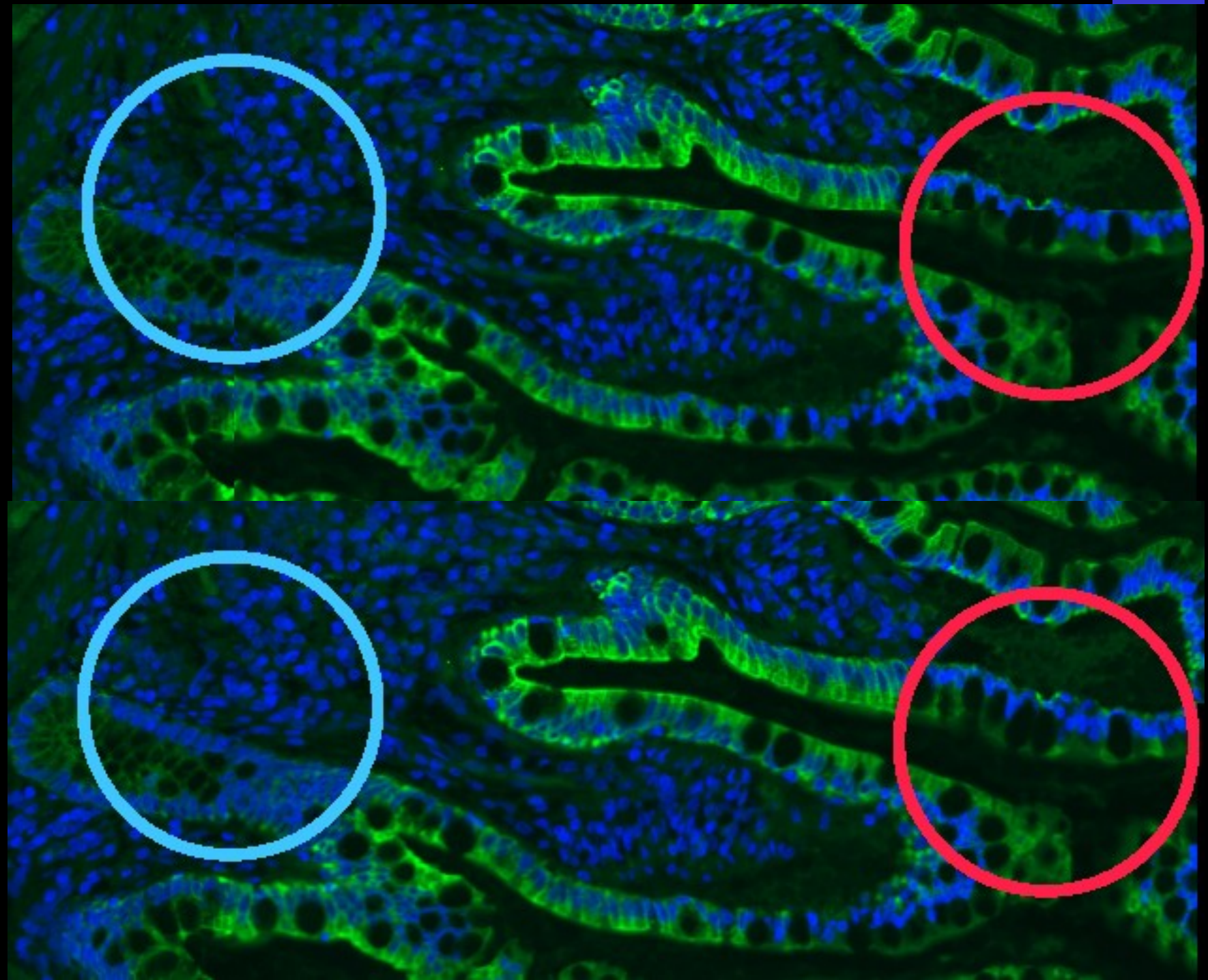
Uneven Illumination



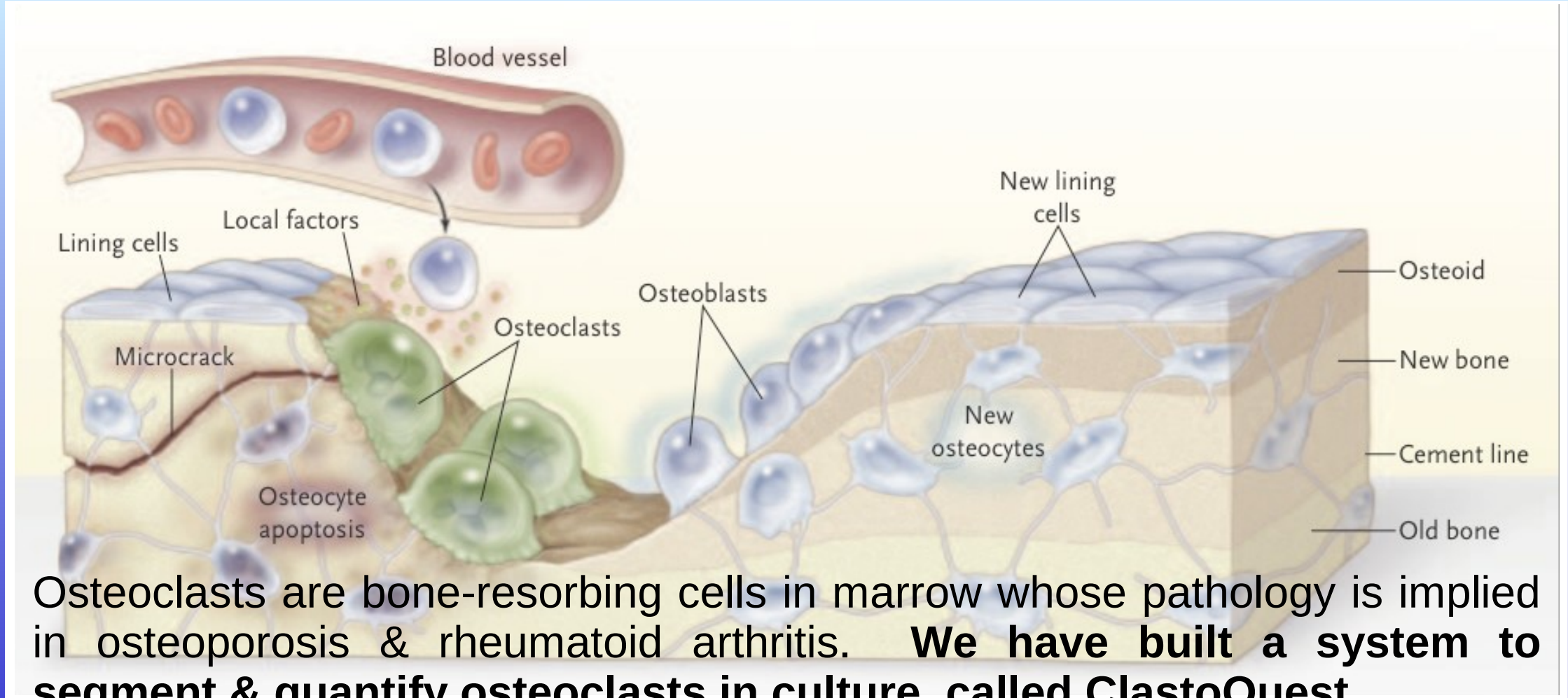

Perfect!
a
Out-of-focus
b
Overexposed
c
Increased Red
d



Slide-based microscopy records many images by automatically moving the tissue sample vs. the microscope. Tiling images together may not be sufficient. ***Stitching*** analyzes the overlap between adjacent images to determine a more precise combination. **We had to develop our own algorithm.**



ClastoQuest (1)

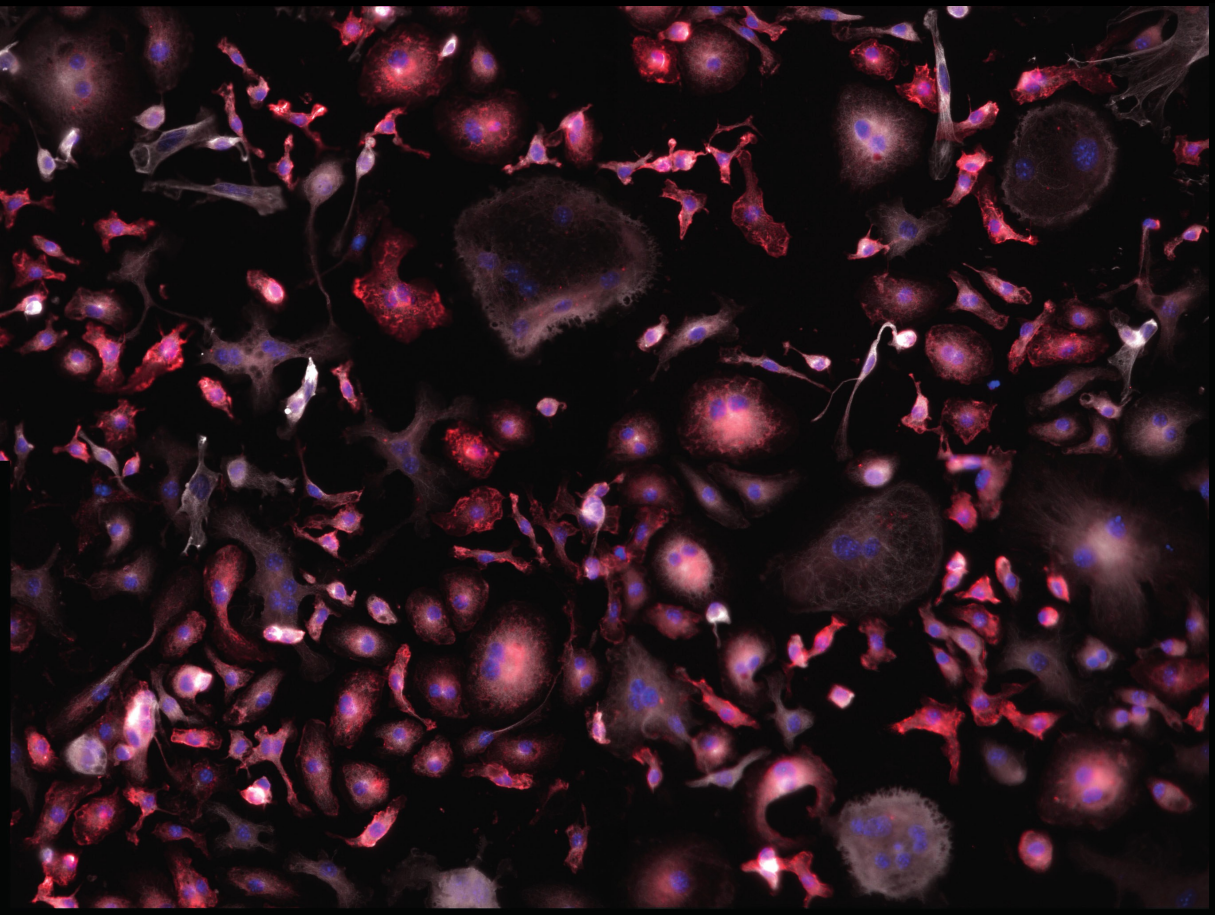


Osteoclasts are bone-resorbing cells in marrow whose pathology is implied in osteoporosis & rheumatoid arthritis. **We have built a system to segment & quantify osteoclasts in culture, called ClastoQuest.**

SEEWALD SOLUTIONS

alex@seewald.at
www.seewald.at

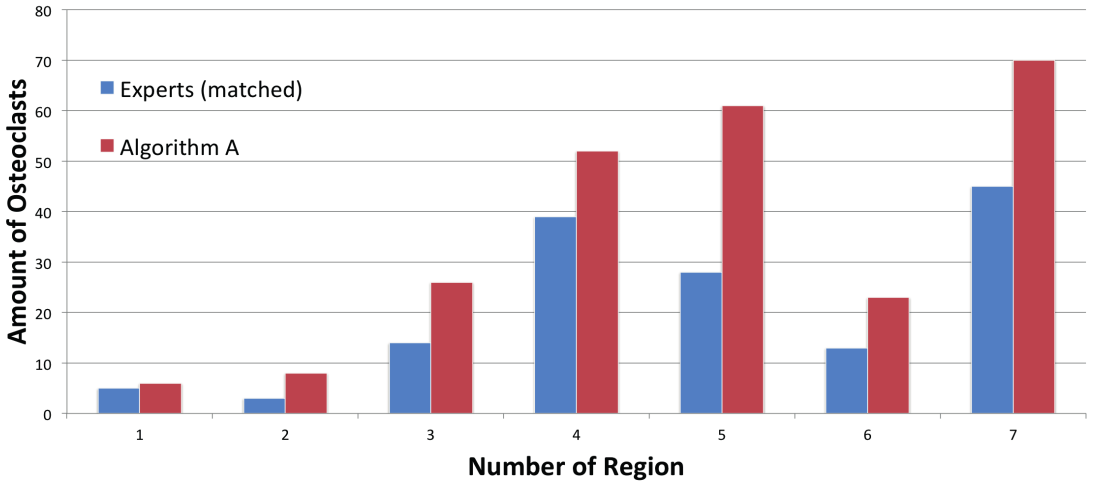
ClastoQuest (2) - Sample



Triple Staining: white = cells, blue = nuclei, red = precursor, not osteoclast

Detection works by counting nuclei (>=3) and computing red average area

ClastoQuest (3) - Algorithm

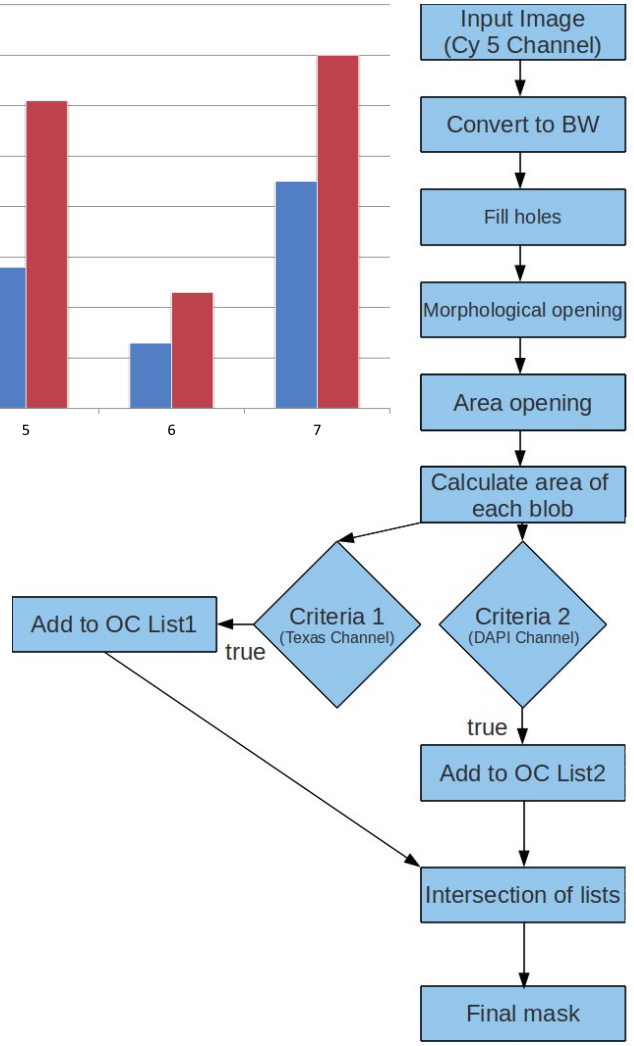


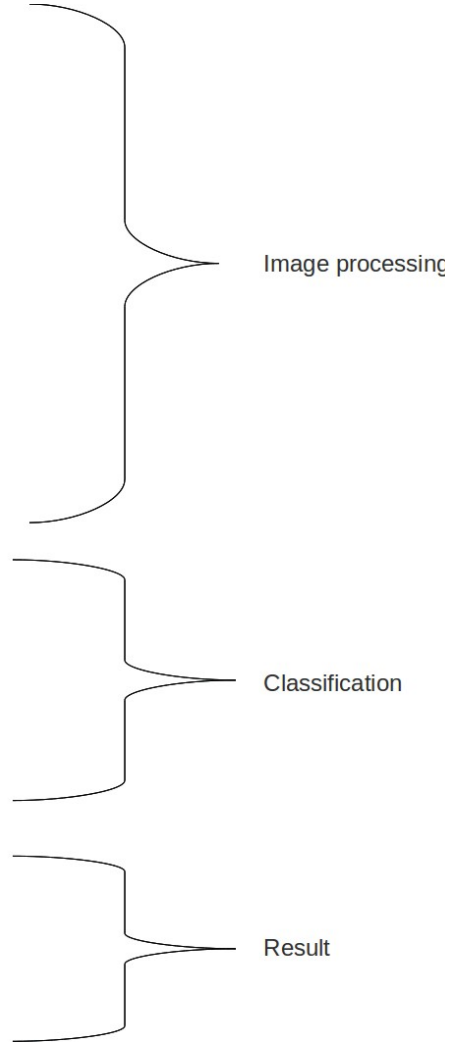
- Criterion 1 computes the ratio of the macrophage marker inside the cell versus the total area of the cell

$$\frac{amount(macrophage\ marker)}{area\ of\ cell} < T_1$$

- Criterion 2 counts the amount of DAPI staining inside the cell

$$amount(DAPI) > T_2$$





SEEWALD SOLUTIONS

alex@seewald.at
www.seewald.at

ClastoQuest (4) - Current Work

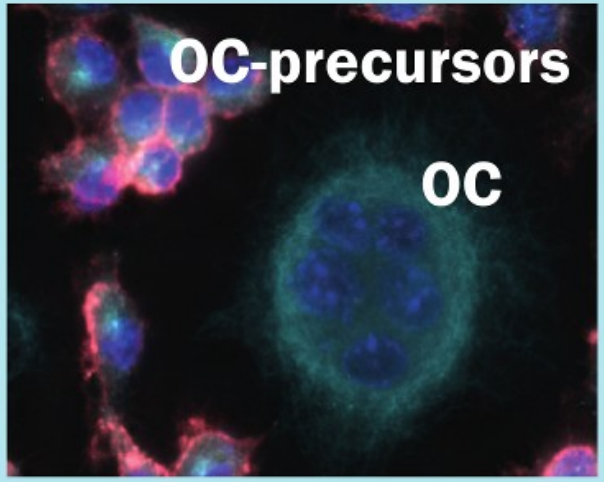


Figure 5

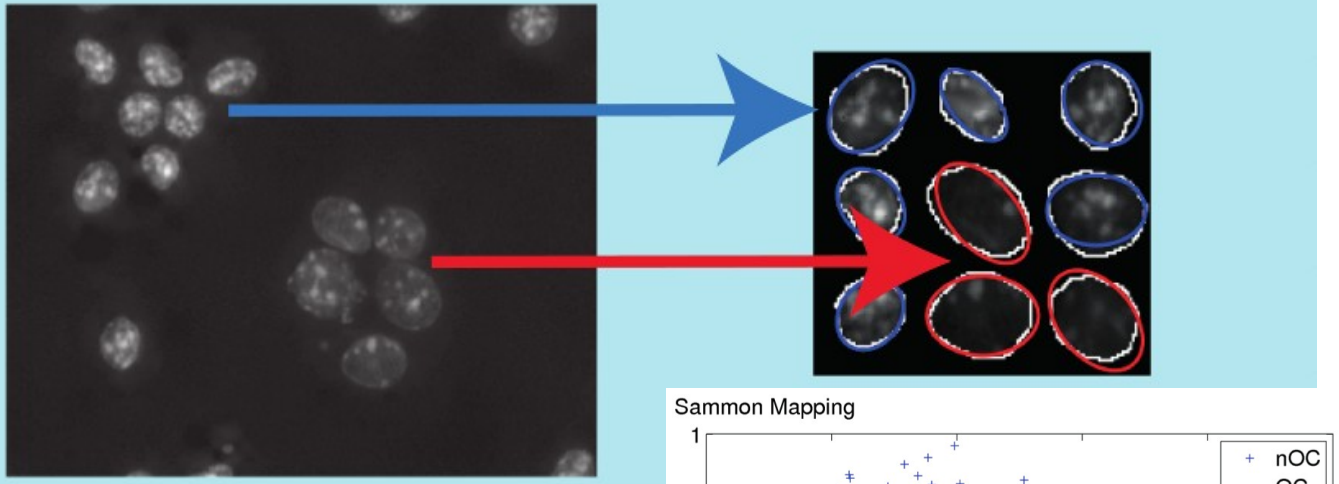
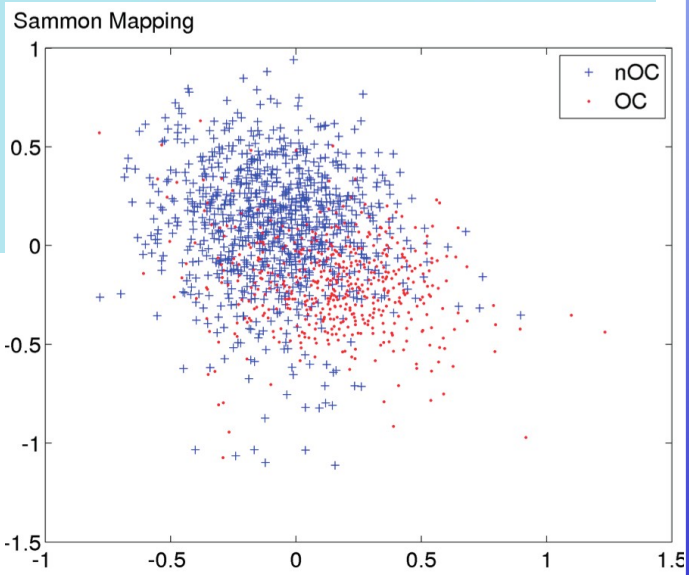


Figure 6



Shape and texture features computed from nuclei of osteoclasts (OC) and non-osteoclasts (nOC) allow to determine class with some confidence (Logistic regression 10CV: 83% accuracy, 75% on holdout test sample) May be combined with ClastoQuest for additional precision. **Dataset**

SEEWALD SOLUTIONS

alex@seewald.at
www.seewald.at

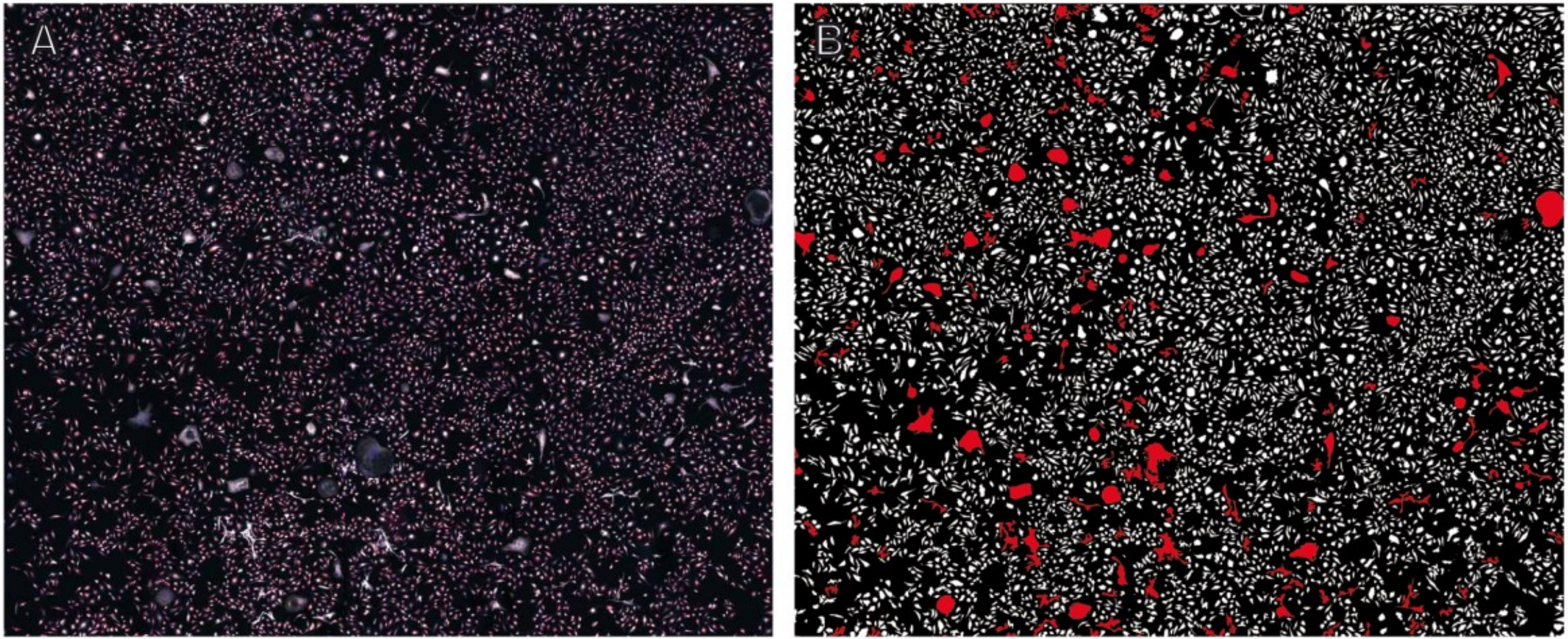
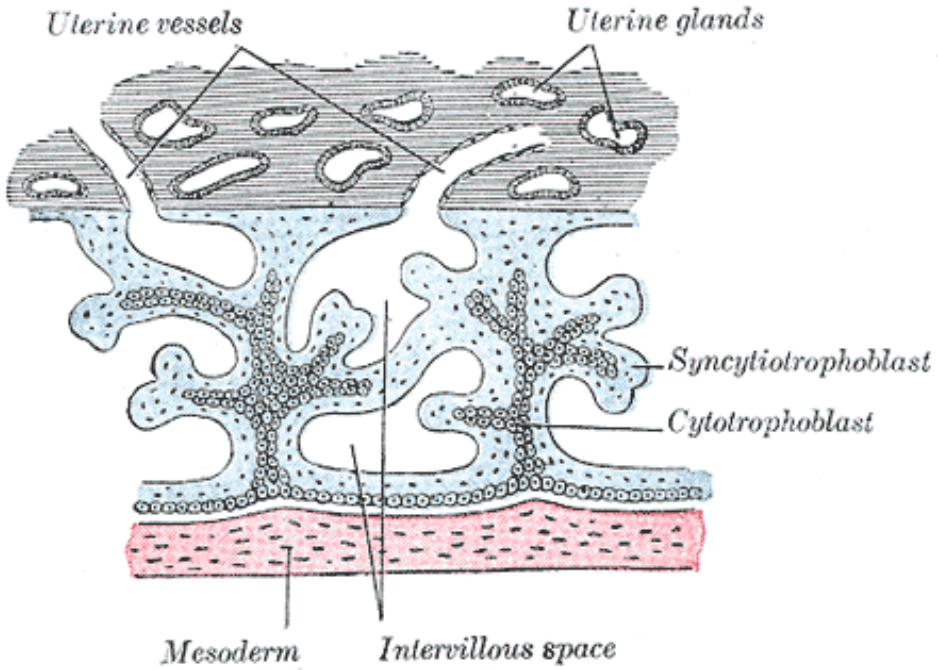


Figure 4: Image analysis of a whole sample-region consisting of 100 (10 x 10) FOVs. A: Immunofluorescence image. B: Result of cell segemntation and analysis: non-OC are marked in white, OC are marked in red.

Syncytiotrophoblast (1)

Primary chorionic villi

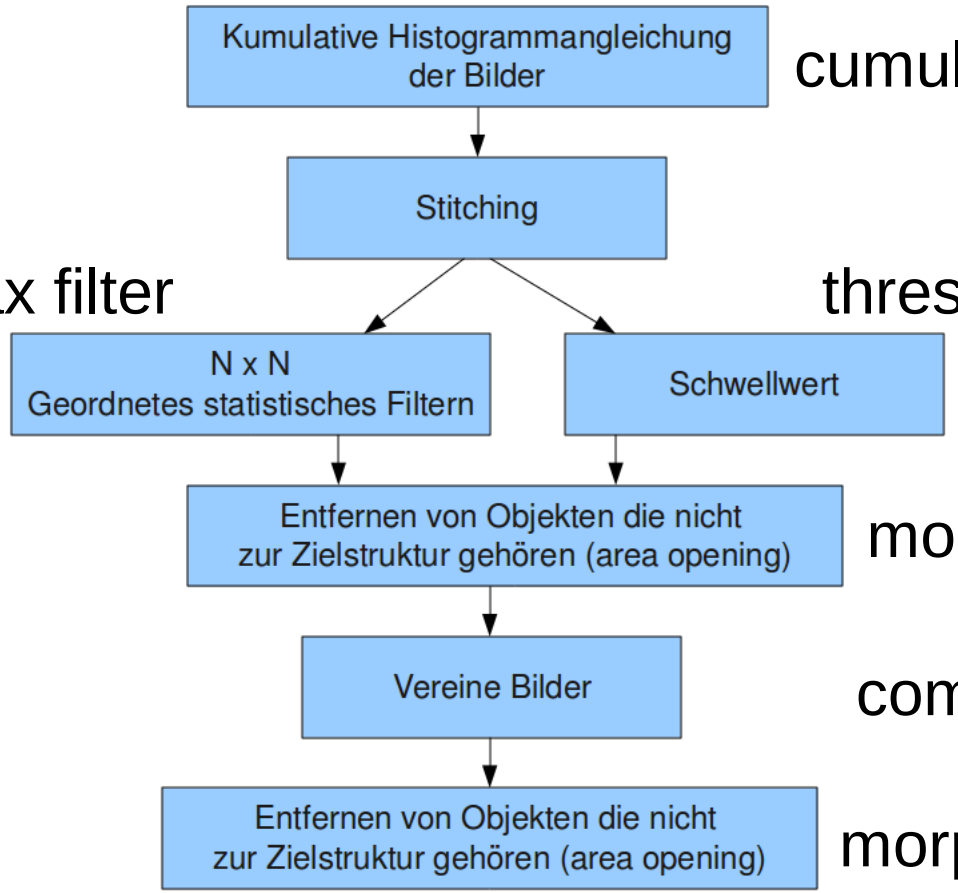


Syncytiotrophoblasts are multinucleated cells within the placenta of embryos at the surface of chorionic villi. Chorionic villi are part of the border between maternal and fetal blood during pregnancy. **We have built a system to segment villi & syncytiotrophoblast and applied it to protein quantification of Receptor for Advanced Glycated End products(RAGE)**

SEEWALD SOLUTIONS

alex@seewald.at
www.seewald.at

Syncytiotrophoblast (2) - Algorithm



cumulative histogram

max filter

thresholding

morphologic opening

combine via OR

morphologic opening

&

Exhaustive Parameter-Optimization

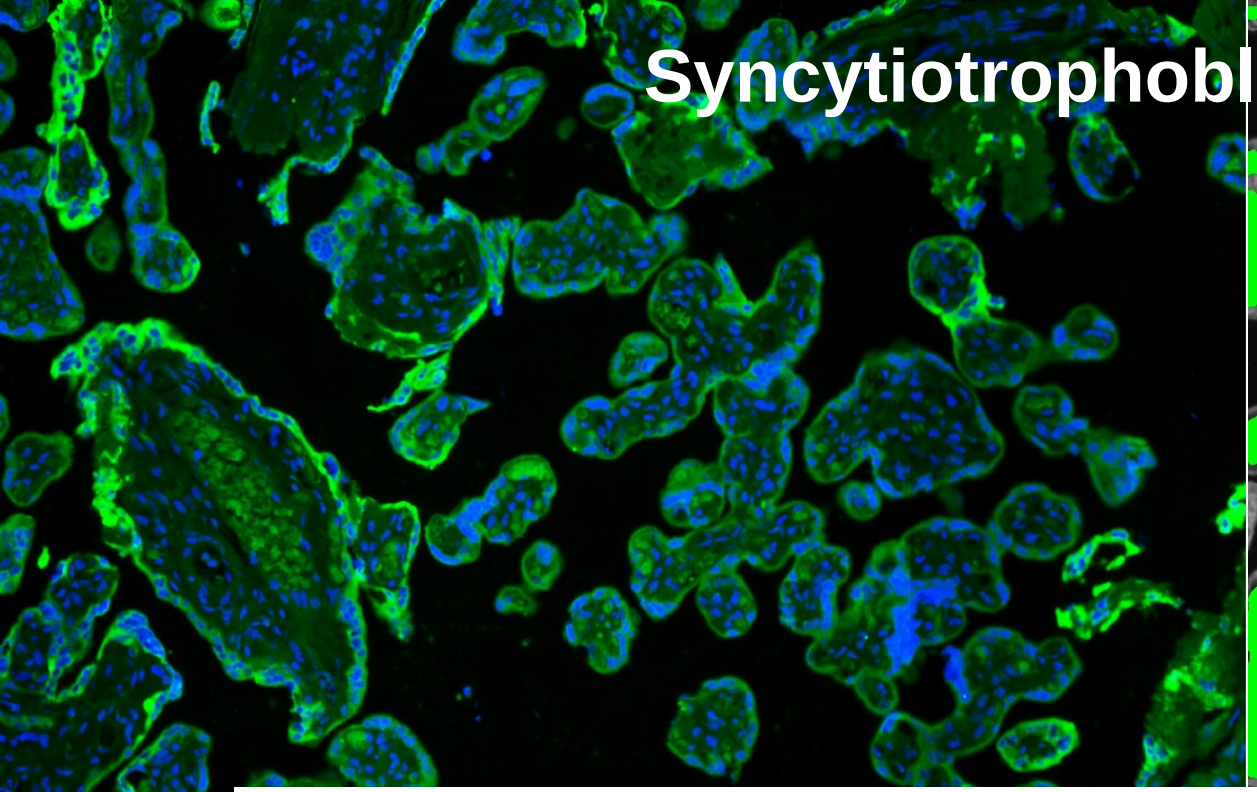
Around 8,100 sets of 7 parameters were tested.

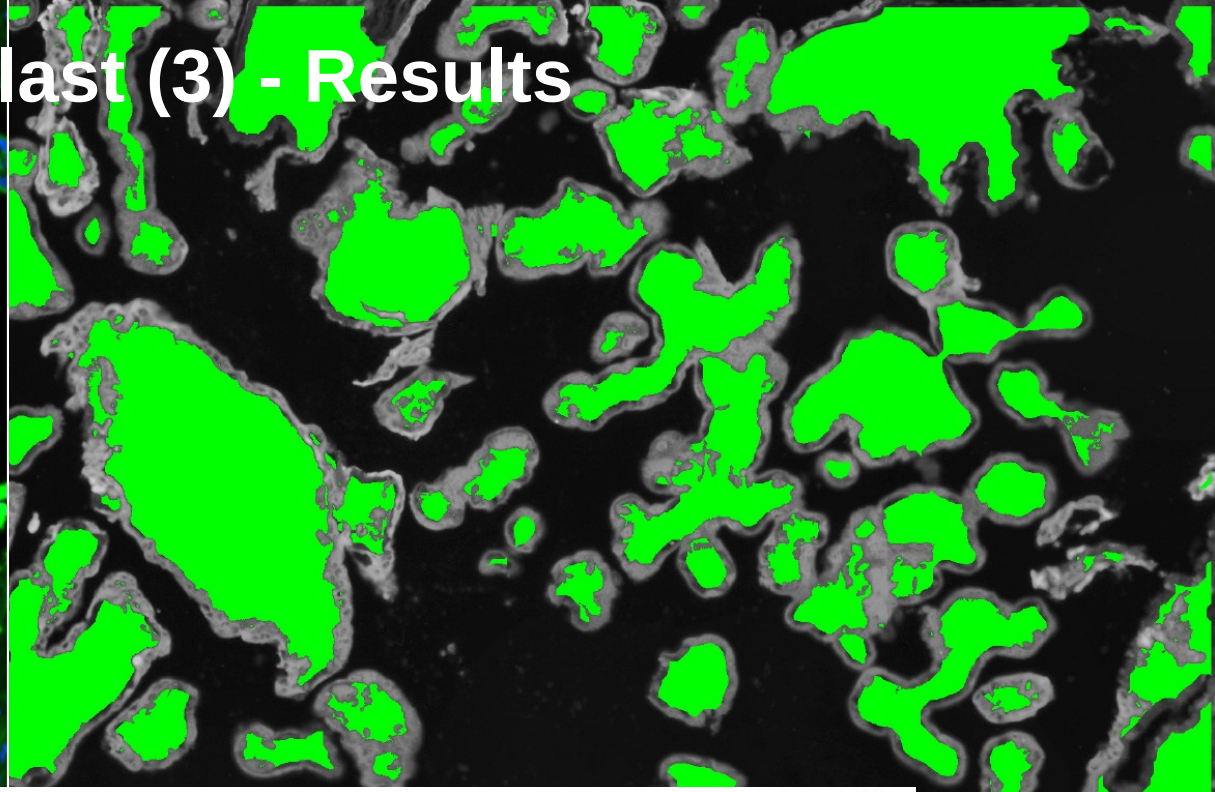
Improved F1 from 0.78 to 0.93

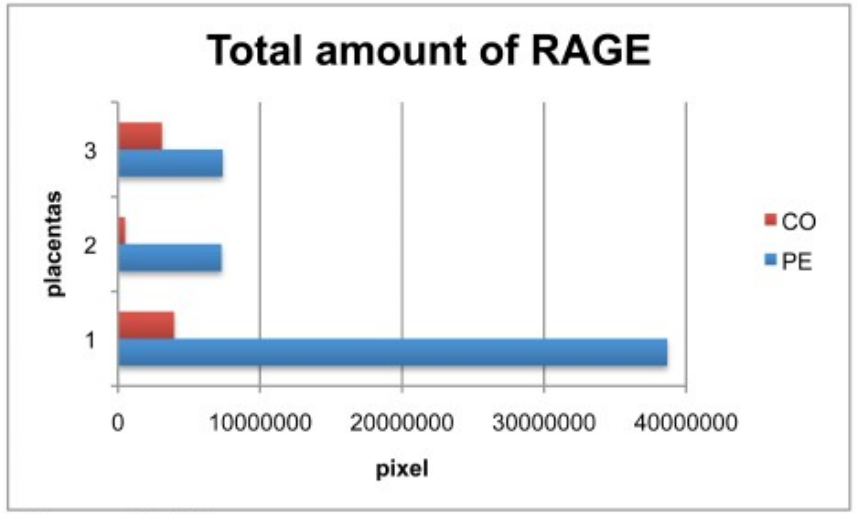
SEEWALD SOLUTIONS

alex@seewald.at
www.seewald.at

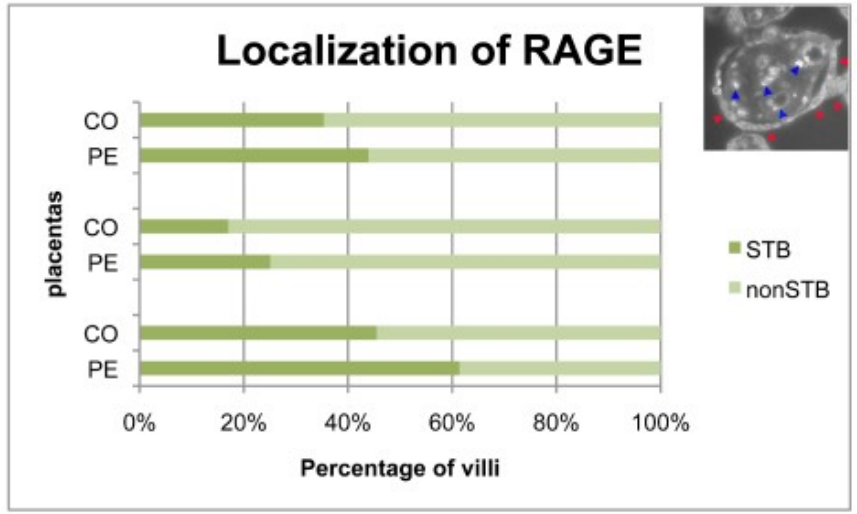
Syncytiotrophoblast (3) - Results







8



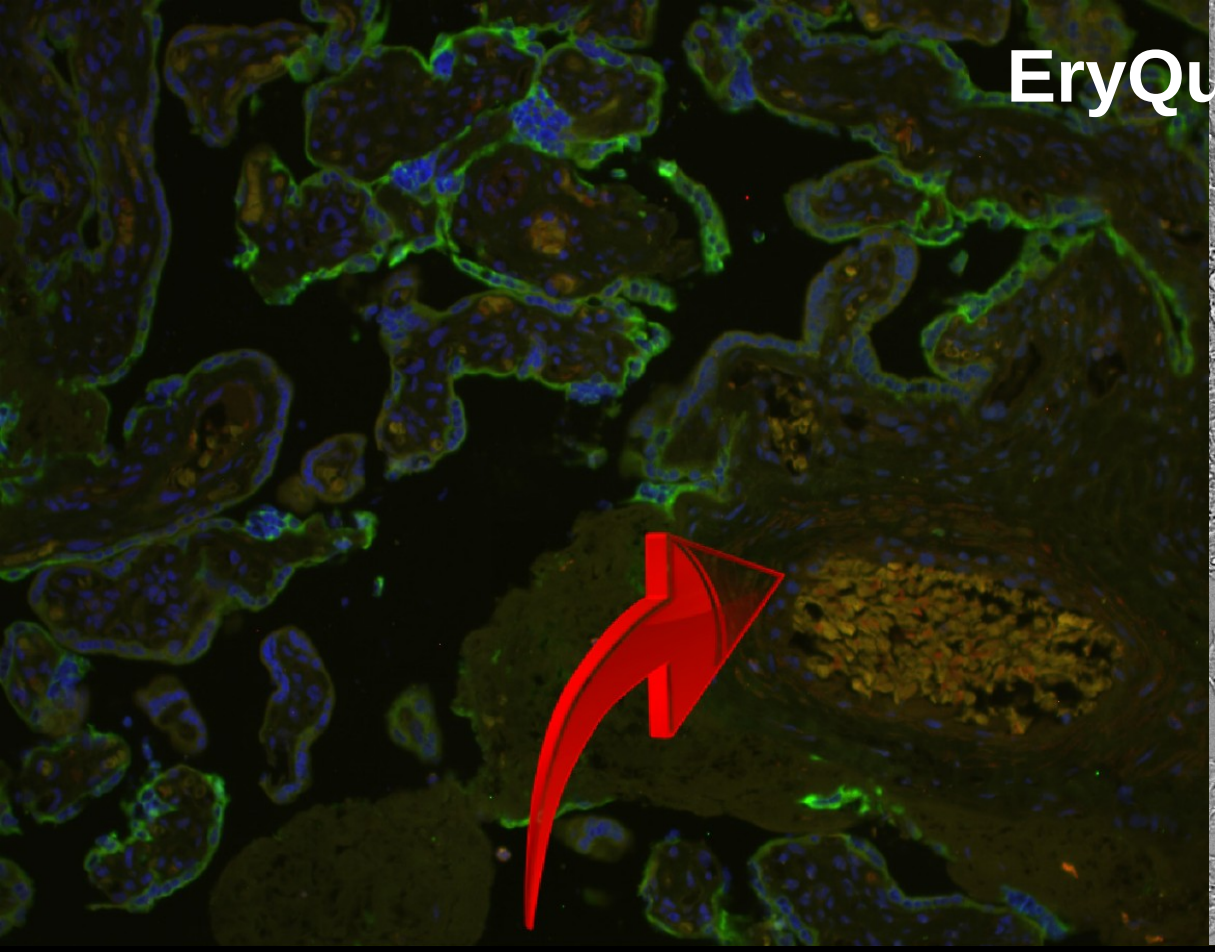
9

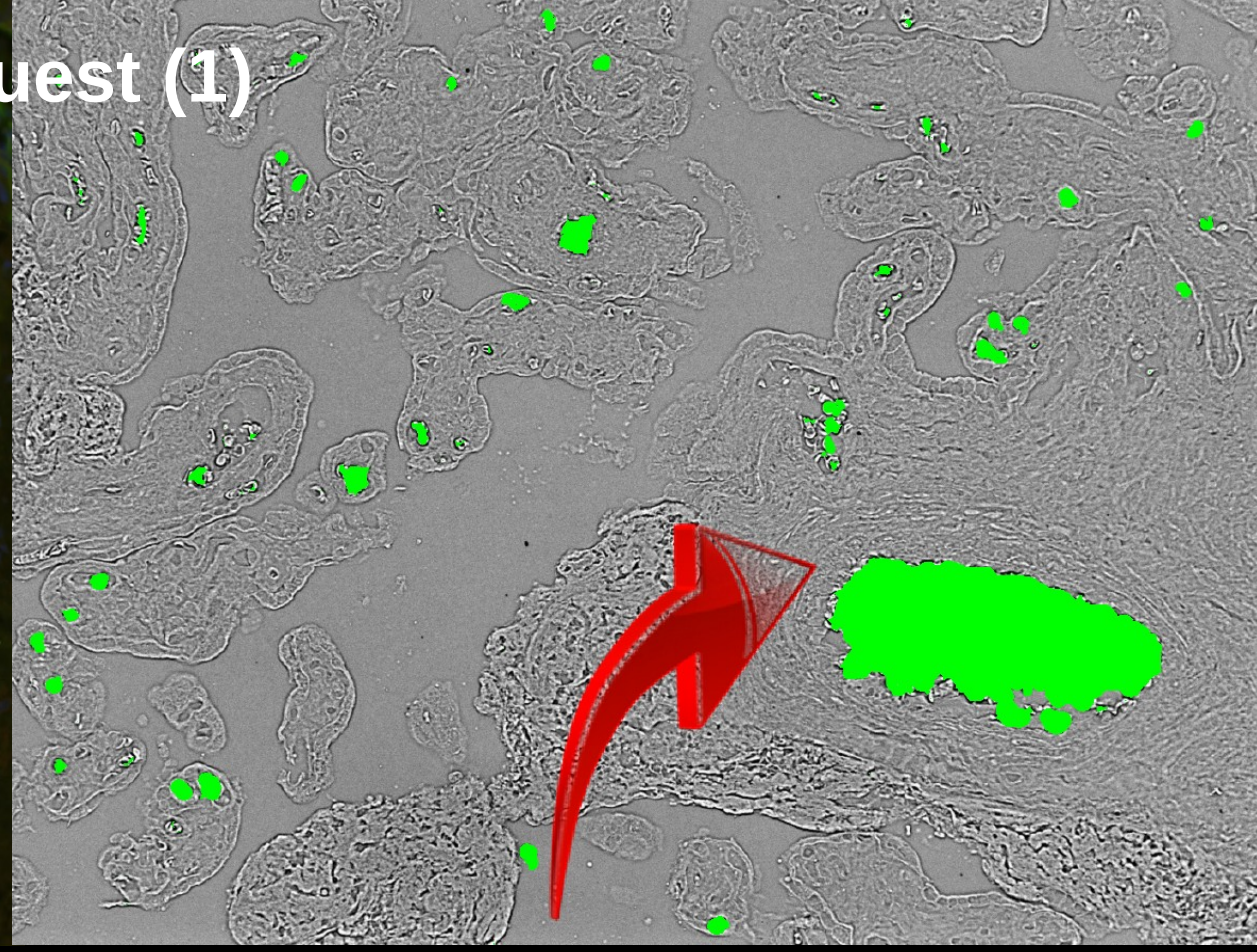
10

Figure 8-10

The total amount of RAGE proteins quantified in PE and CO placentas is shown in Figure 8, respectively localization within the villi in Figure 10. RAGE can be found in STB (red arrowheads) or other cell types (blue arrowheads) visualized in Figure 9.

EryQuest (1)





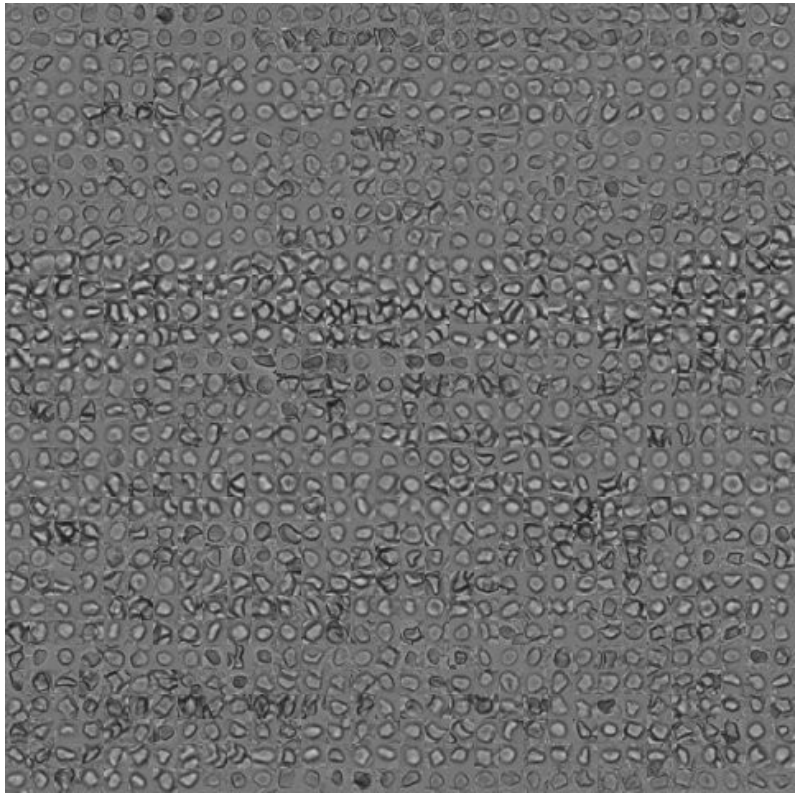
Erythrocytes have high autofluorescence. May lead to noisy immunofluorescent measurements. **Using only ground-truth data, we „taught“ the computer to remove erythrocytes from images ([Viola & Jones, 2001] = OpenCV Haartraining)**

Essential for RAGE analysis from previous slide!

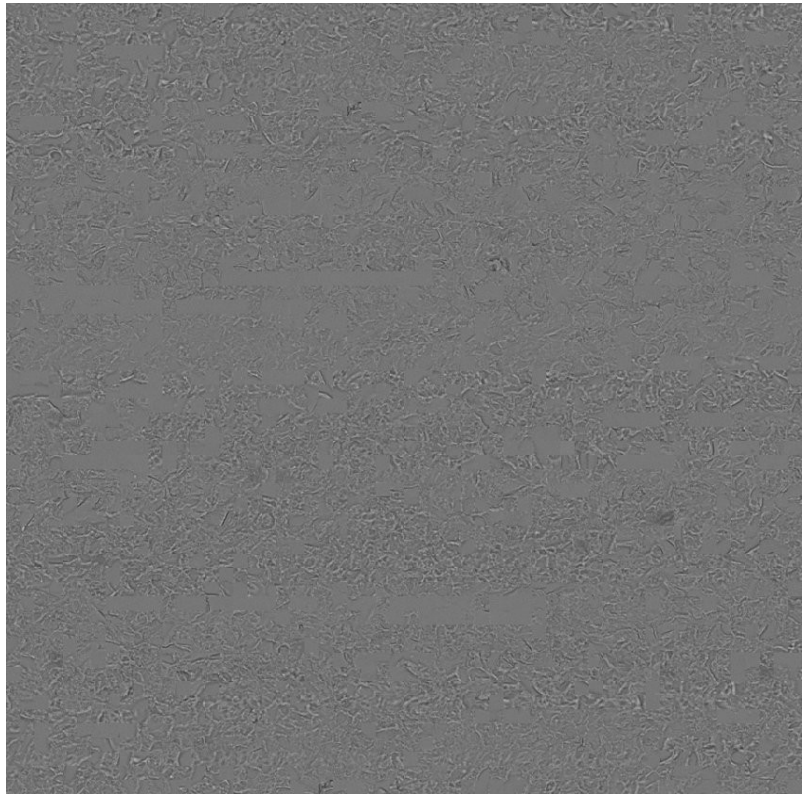
EryQuest (2) - Training

Training Data

>4000 erythrocytes have been marked up manually as well as...



... >2000 regions containing no erythrocytes (tissue, background)

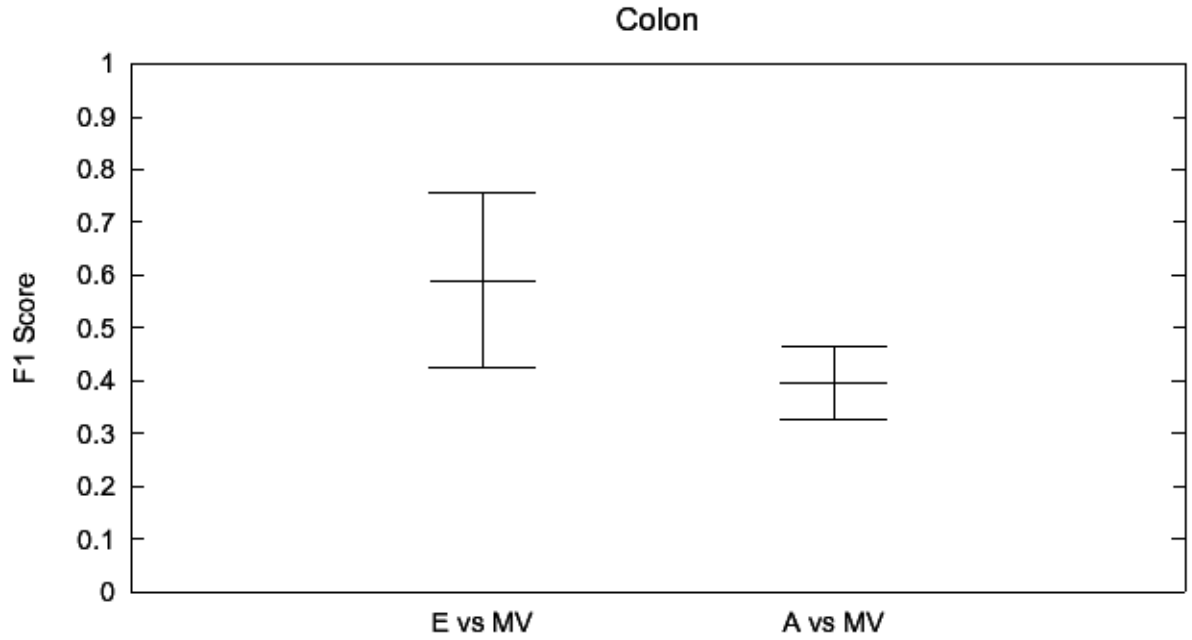


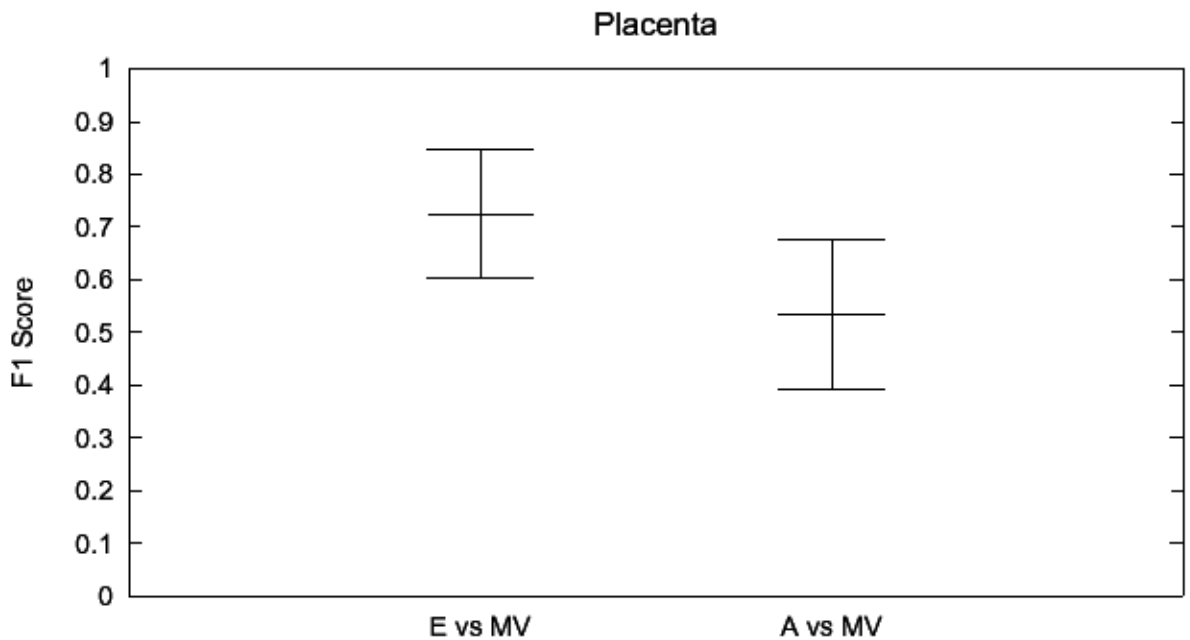
&

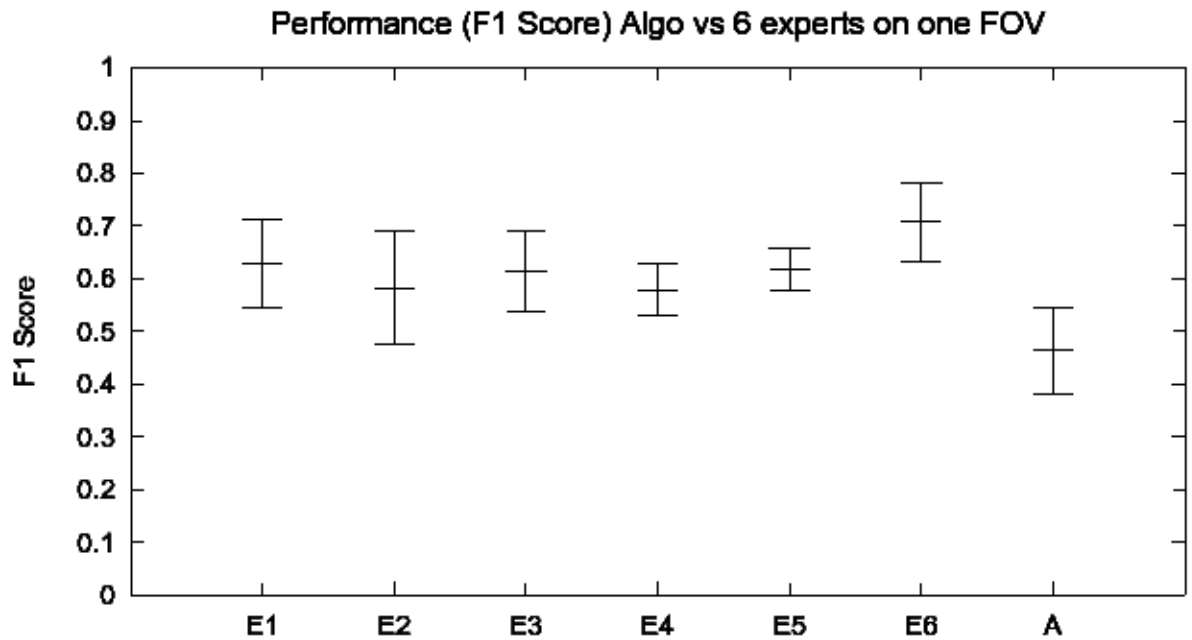
- Exhaustive Haar-Train-Param.Opt (1.8 million masks)
- >=2 Experts for each FOV, many train data variants
- Extend Haar-Test with conf. values
- Run four times with lossless *k**90° rotations...

SEEWALD SOLUTIONS

alex@seewald.at
www.seewald.at


Colon
F1 Score
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
E vs MV
A vs MV


Placenta
F1 Score
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
E vs MV
A vs MV


Performance (F1 Score) Algo vs 6 experts on one FOV
F1 Score
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
E1
E2
E3
E4
E5
E6
A

C. Elegans Protein Localization (1)

Seewald AK, Cypser J, Mendenhall A, Johnson T (2010) Quantifying Phenotypic Variation in Isogenic Caenorhabditis elegans Expressing Phsp-16.2::gfp by Clustering 2D Expression Patterns, PLoS ONE 5(7): e11426. doi:10.1371/journal.pone.0011426.

Quantifying Phenotypic Variation...

Analyzing changes in appearance / phenotype...

in Isogenic Caenorhabditis elegans...

in small nematodes (worms) which all have the same genetic code (clones)

Expressing Phsp-16.2::gfp...

which express a GFP reporter that binds to heatshock protein 16 (transgenic)

by Clustering 2D Expression Patterns

by extracting 2D expression patterns that are independent of worm pose AND clustering these patterns using hierarchical clustering methods.

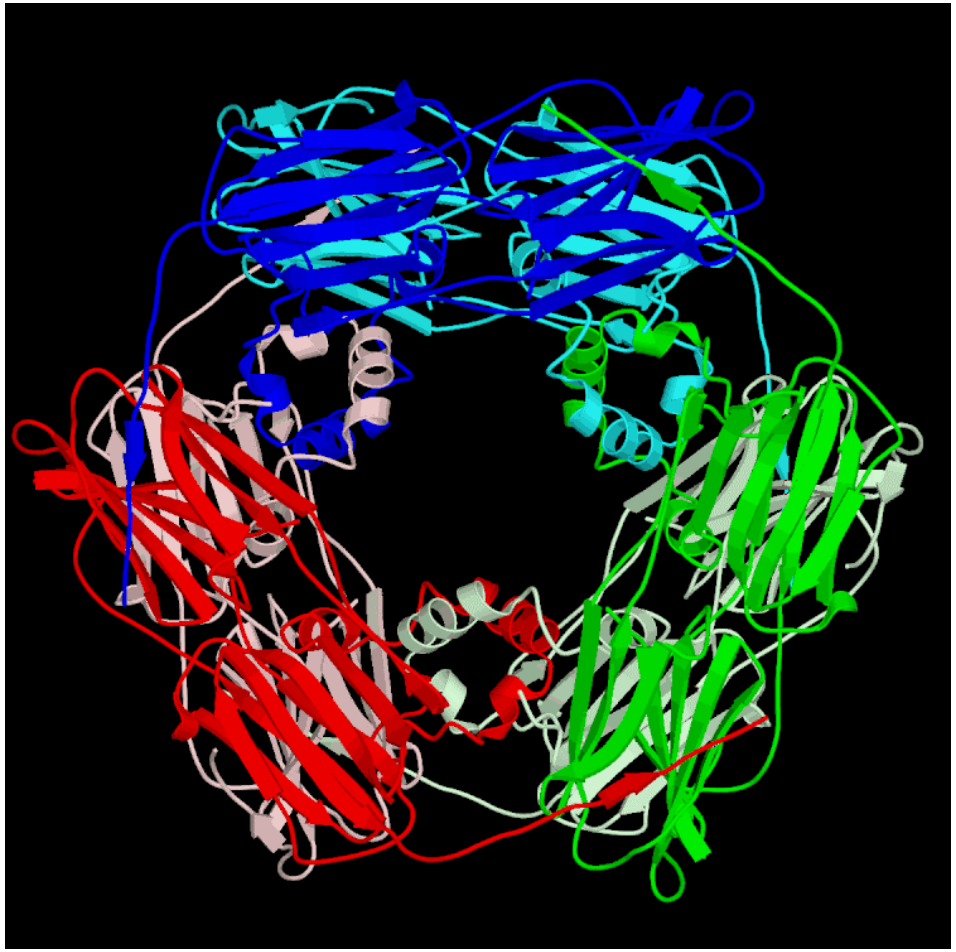
SEEWALD SOLUTIONS

alex@seewald.at
www.seewald.at

C. Elegans Protein Localization (2)

Heat Shock Protein 16 – Increases expression (also) when organism is exposed to high temperatures

- HSP are named by molecular weight (=16kD). Expressed in intestine and pharynx. Induced in response to heat shock or other environmental stresses.
- Interacts with intra-cellular human beta amyloid peptide (Alzheimer plagues)
- High expression correlates with worm longevity acc. to earlier studies.



SEEWALD SOLUTIONS

alex@seewald.at
www.seewald.at

C. Elegans Protein Localization (3)

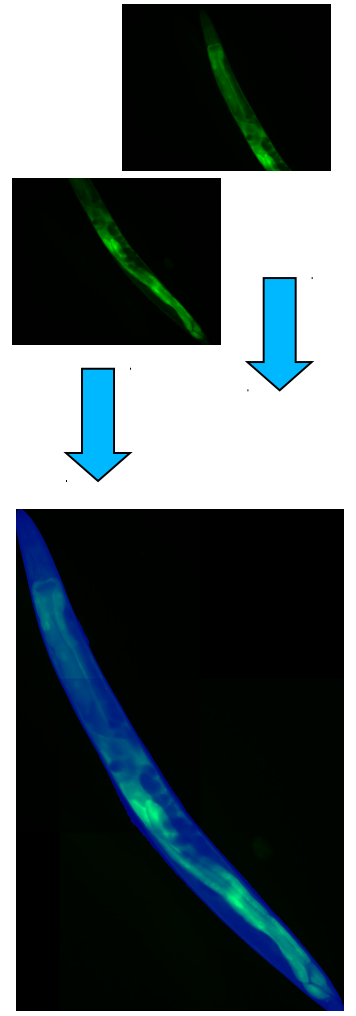
Merging anterior/posterior image & extracting worm

(pixelClassification)

- Machine learning from manually tagged sample images (i.e. „Ground Truth“)
- Threshold optimization by testing minimum circularity and area of largest blob
- Closure (erode, dilate), fill internal holes with circularity below threshold
- Heuristic search for breaks in contour, which are repaired with straight lines and filled on the inside

(meshAB)

- Image correlation coefficient for combining head and tail images (~ simplified stitching)



SEEWALD SOLUTIONS

alex@seewald.at
www.seewald.at

C. Elegans Protein Localization (4)

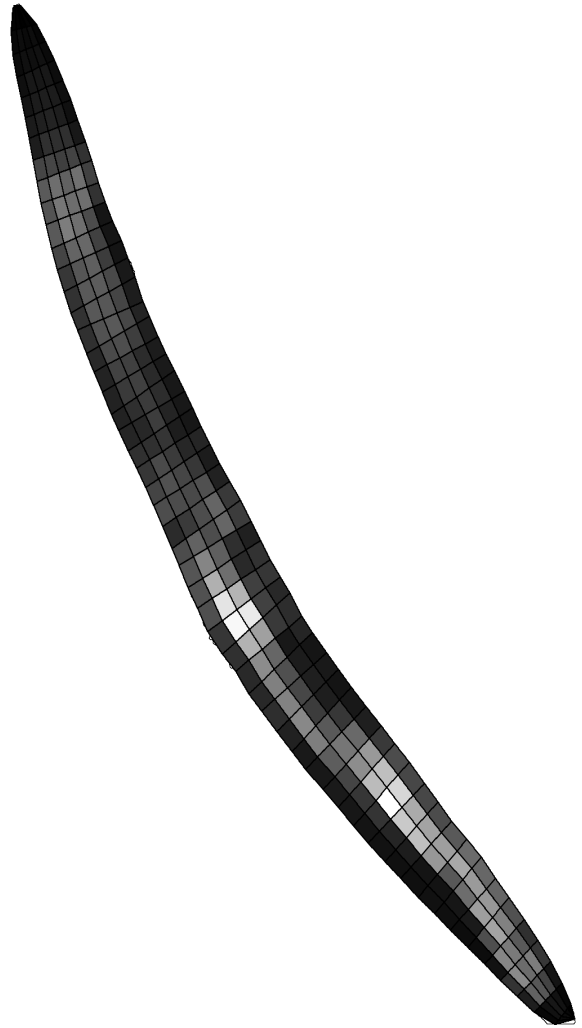
Computing 2D expression patterns (1)

(meshAB)

- Computing head/tail/vulva position from worm pose via heuristics (~ 50%)
- Manual markup in remaining cases

(sampleCE)

- Worm backbone via „thinning“
- Search left/right from backbone in perpendicular direction to local curvature for worm border
- Split left/right and top/bottom into the desired number of tiles
- Compute average GFP intensity per tile

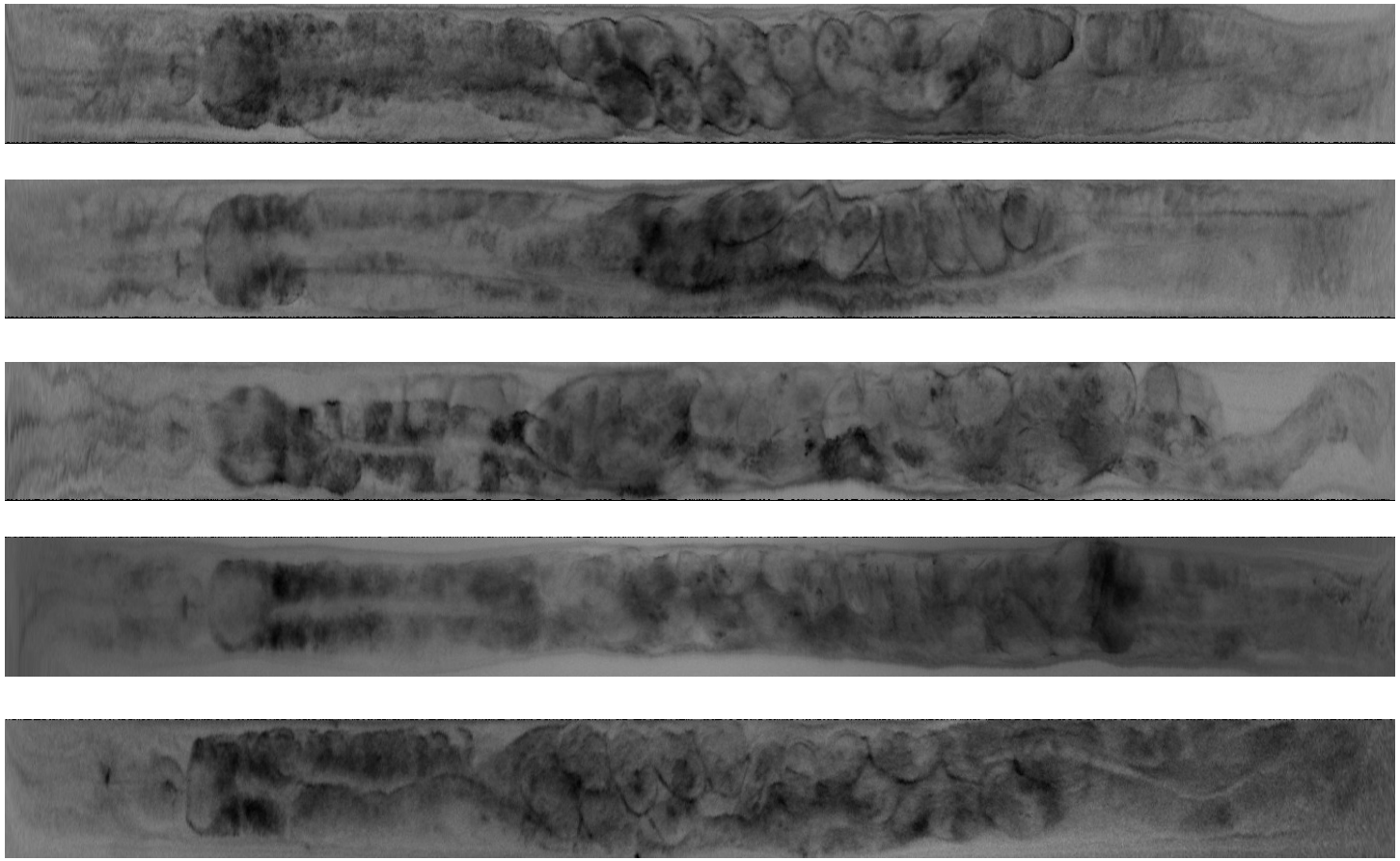


SEEWALD SOLUTIONS

alex@seewald.at
www.seewald.at

C. Elegans Protein Localization (5)

Computing 2D expression patterns (2)



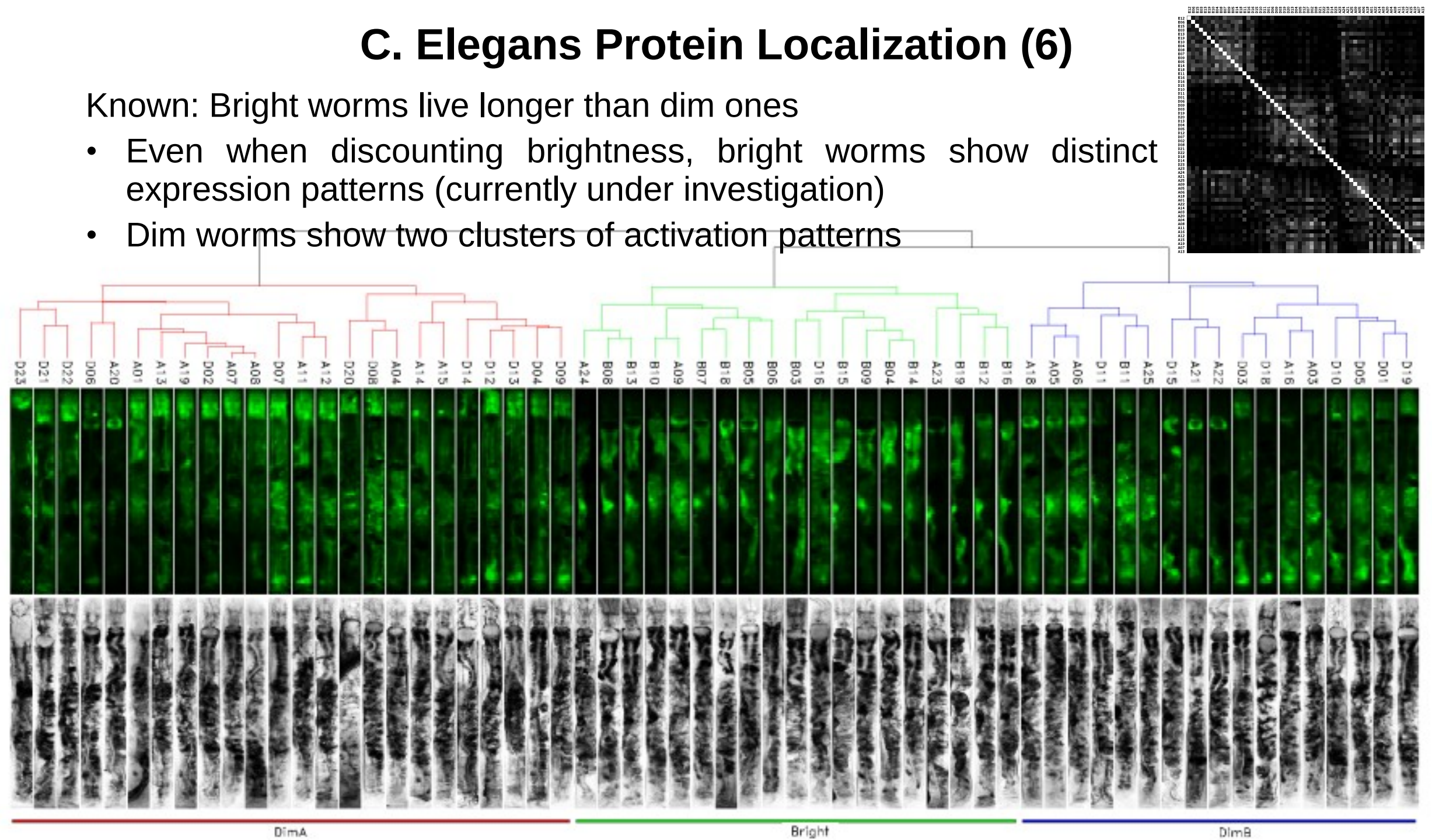
SEEWALD SOLUTIONS

alex@seewald.at
www.seewald.at

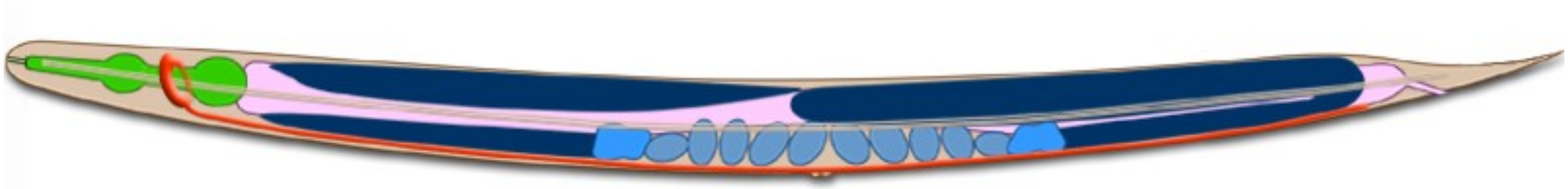
C. Elegans Protein Localization (6)

Known: Bright worms live longer than dim ones

- Even when discounting brightness, bright worms show distinct expression patterns (currently under investigation)
- Dim worms show two clusters of activation patterns



Future Work (1)



- Bottleneck is image acquisition – each worm has to be taken from culture medium, anesthesized, cleaned and imaged (ca. 30min per worm)
- Resolution is too coarse for observing single cells

→ *Culturing worms on chamber-slides, using slide-based microscopy & automated imaging*

- Lots of problems with different microscope settings, air bubbles, finetuning,...

→ *„Closed-loop“ system (microscope, moveable slide and image analysis - coupled system)*

SEEWALD SOLUTIONS

alex@seewald.at
www.seewald.at

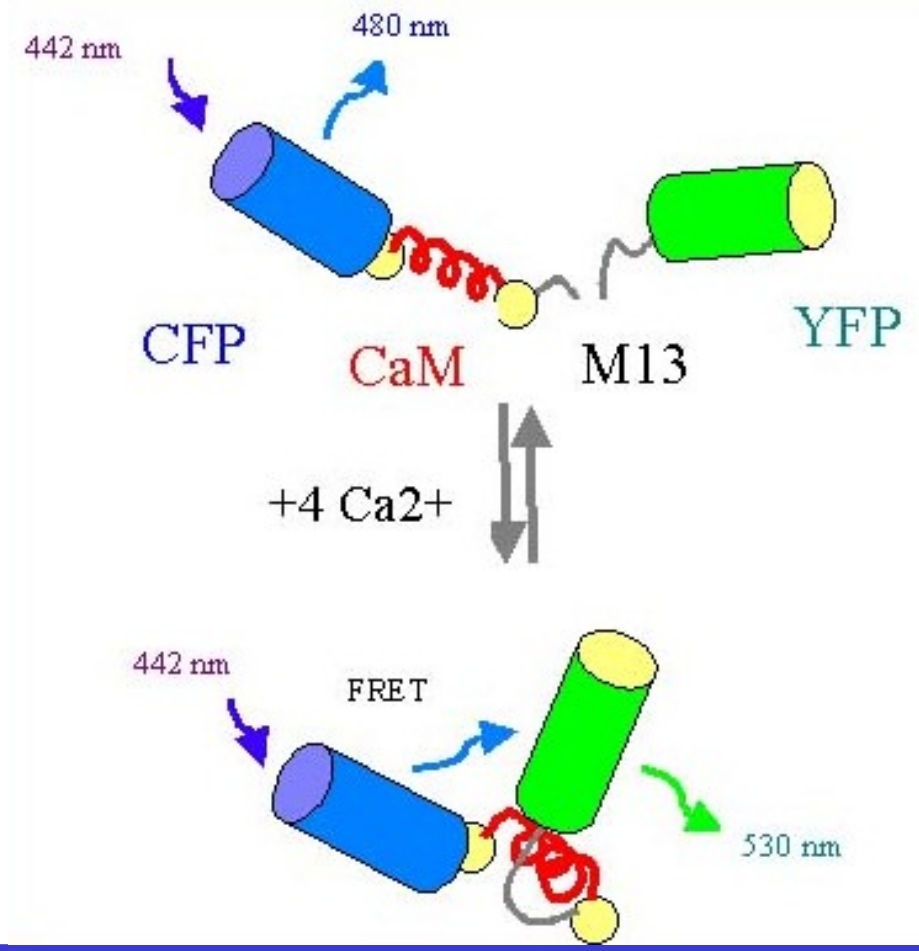
Future Work (2)

Cameleon: measures Ca2+ level = nerve cell activity

CFP emits a 480nm photon on excitation with 442nm.

High Ca2+ concen-trations lead to conformation changes and the photon is absorbed by YFP and re-emitted as 530nm.

Proportion between 480nm and 530nm response used as signal for Ca2+ level.



SEEWALD SOLUTIONS

alex@seewald.at
www.seewald.at

Video 2 - Worm Dedistorting in Real Time

SEEWALD SOLUTIONS

alex@seewald.at
www.seewald.at

Thanks for listening!

Science2Business 2011 Award (3rd Place)

SEEWALD SOLUTIONS