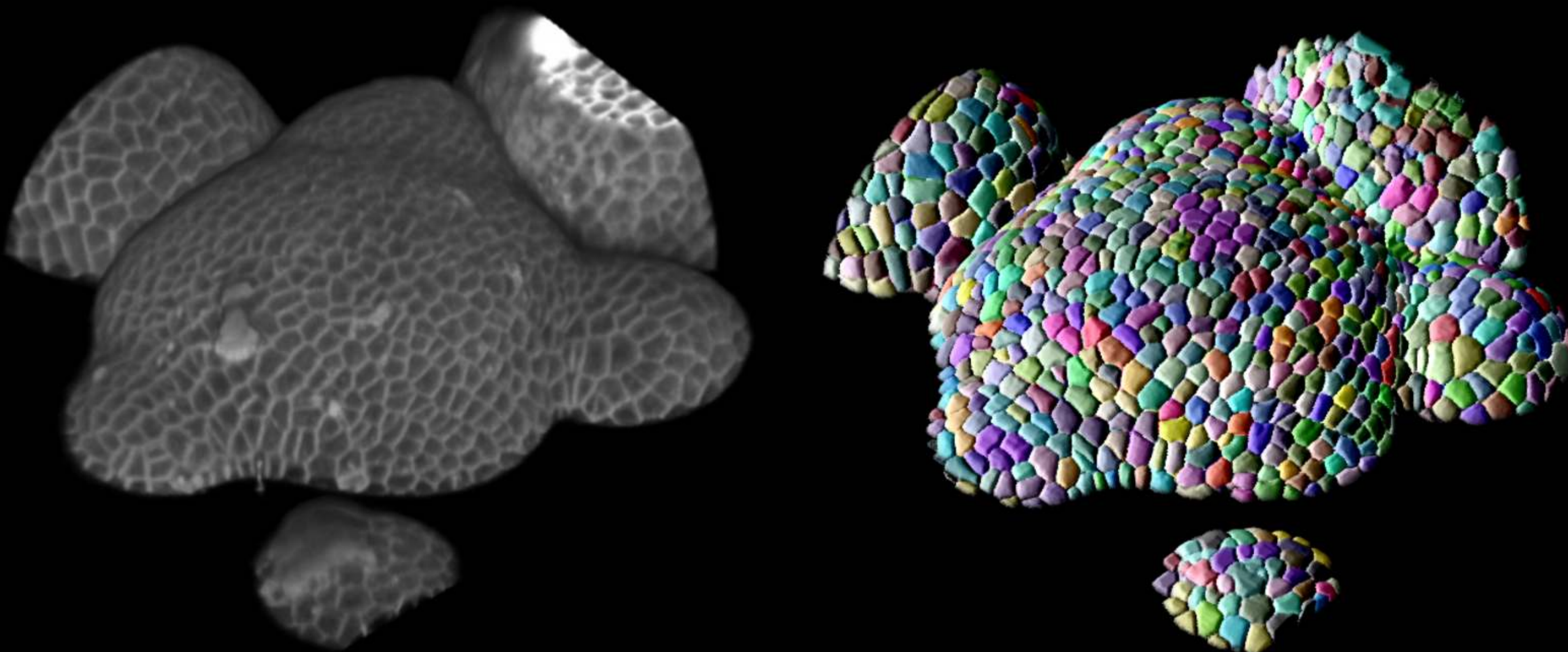
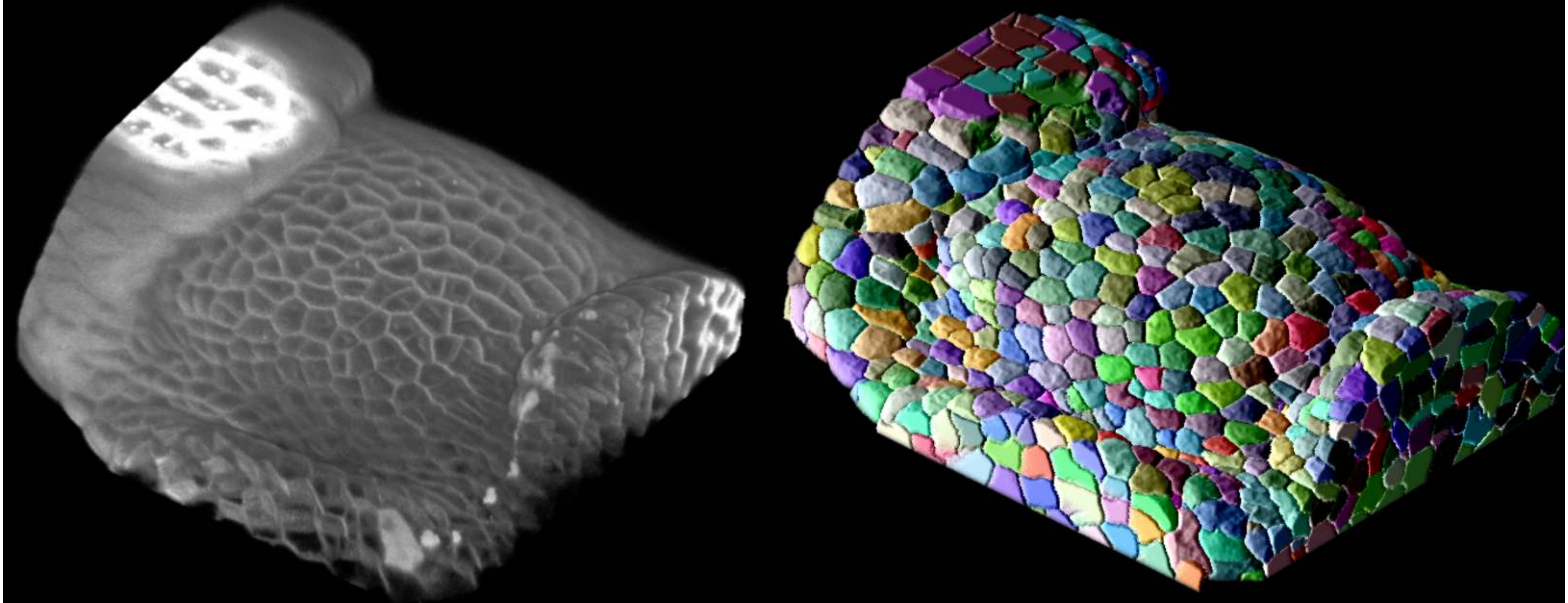
Biological Image Processing

Alexandre Cunha
CAMBIA – Beckman Institute

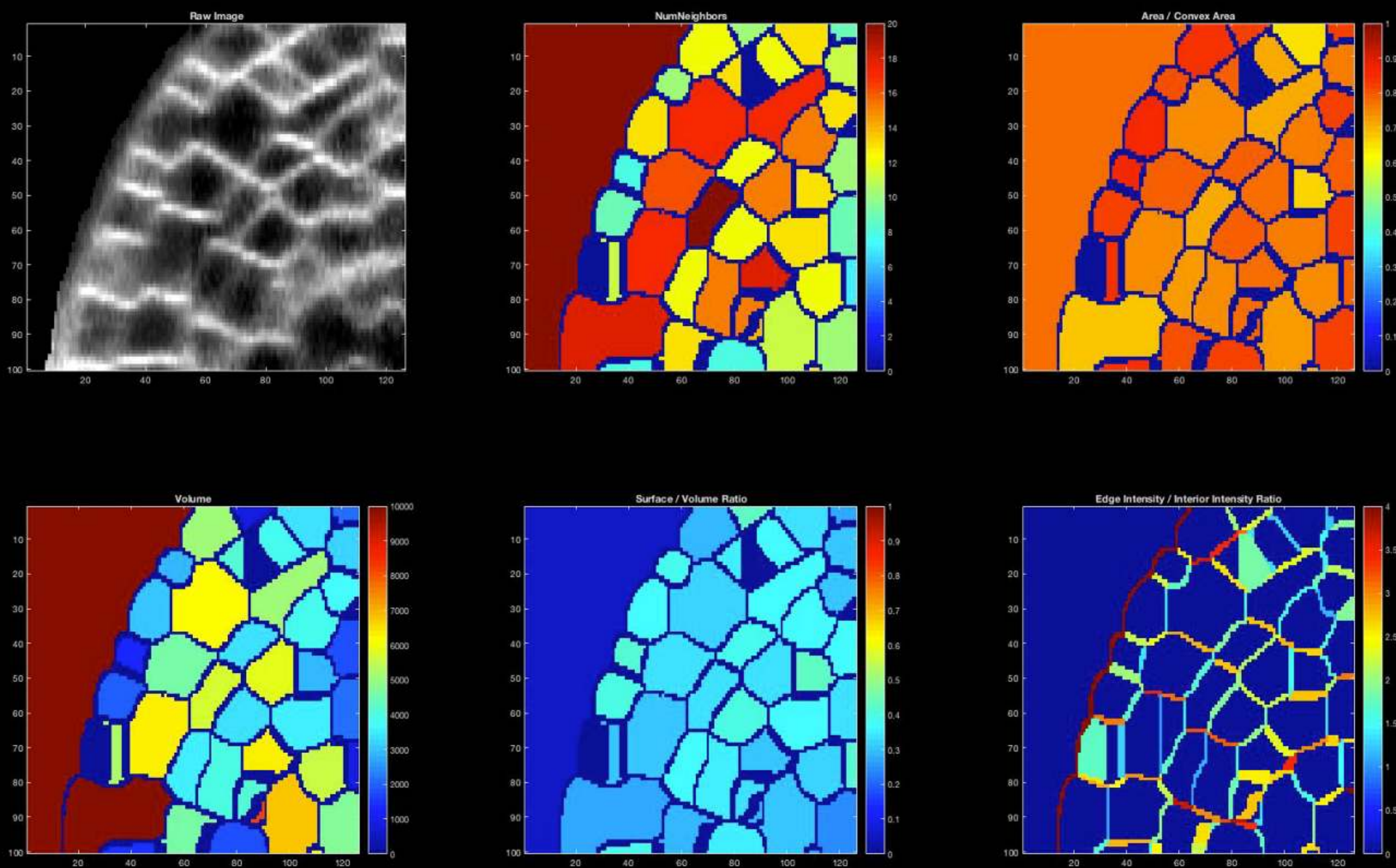
Meristem segmentation (Meyerowitz lab) – J. Stegmaier



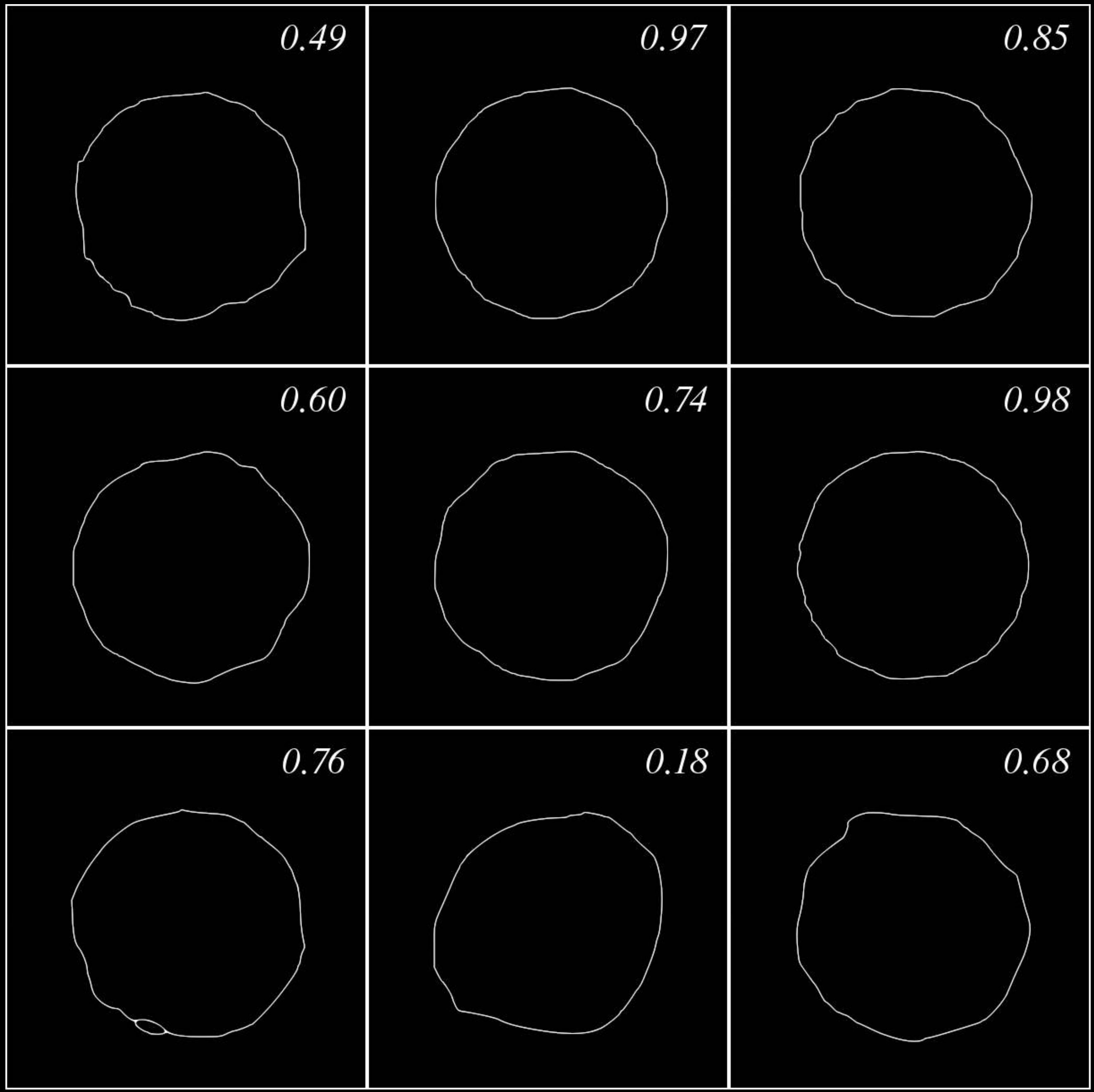
Floral meristem segmentation (Arabidopsis)



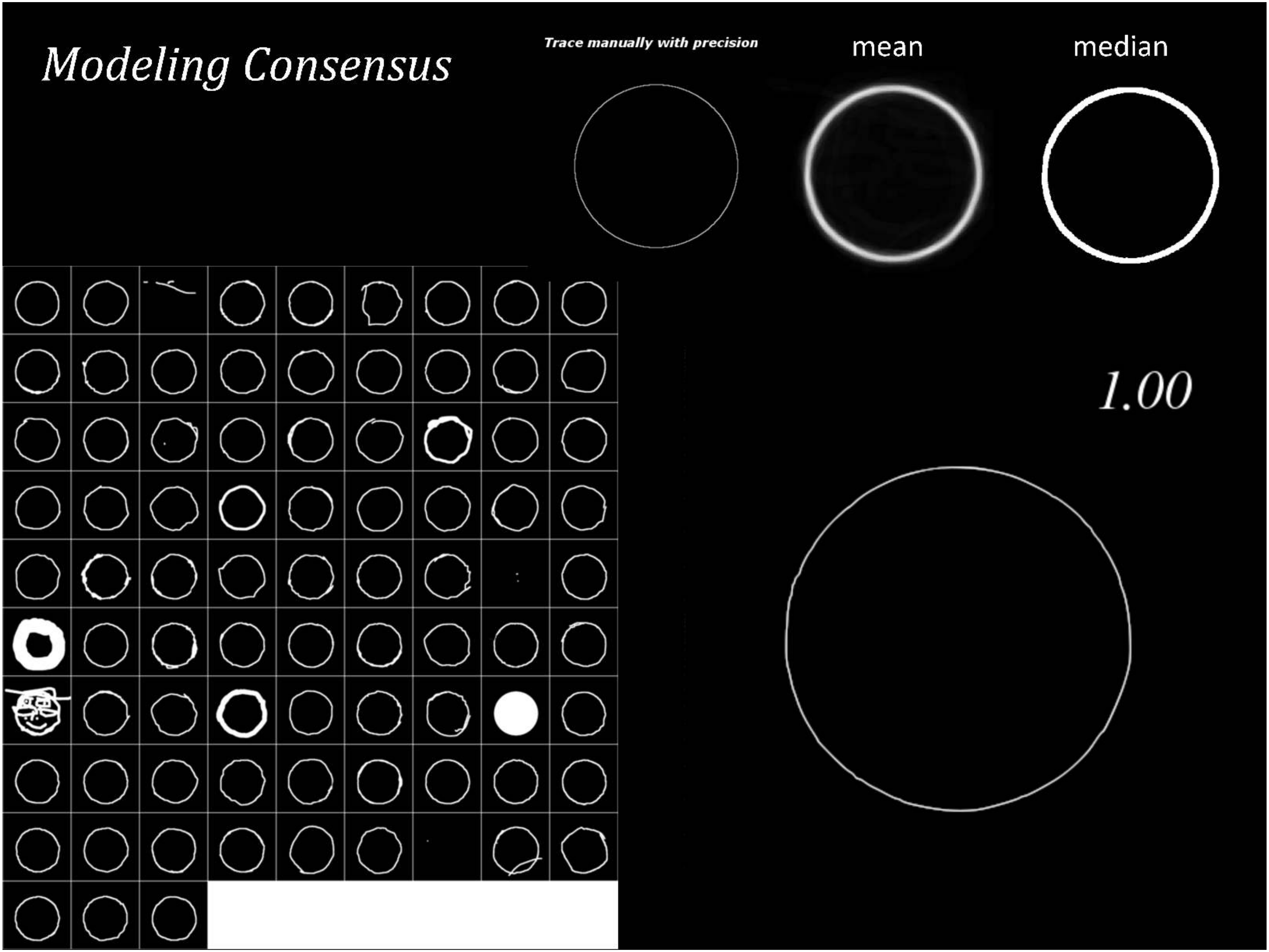
Using cell properties for error elimination/reduction



Modeling Consensus with good and bad apples



Modeling Consensus

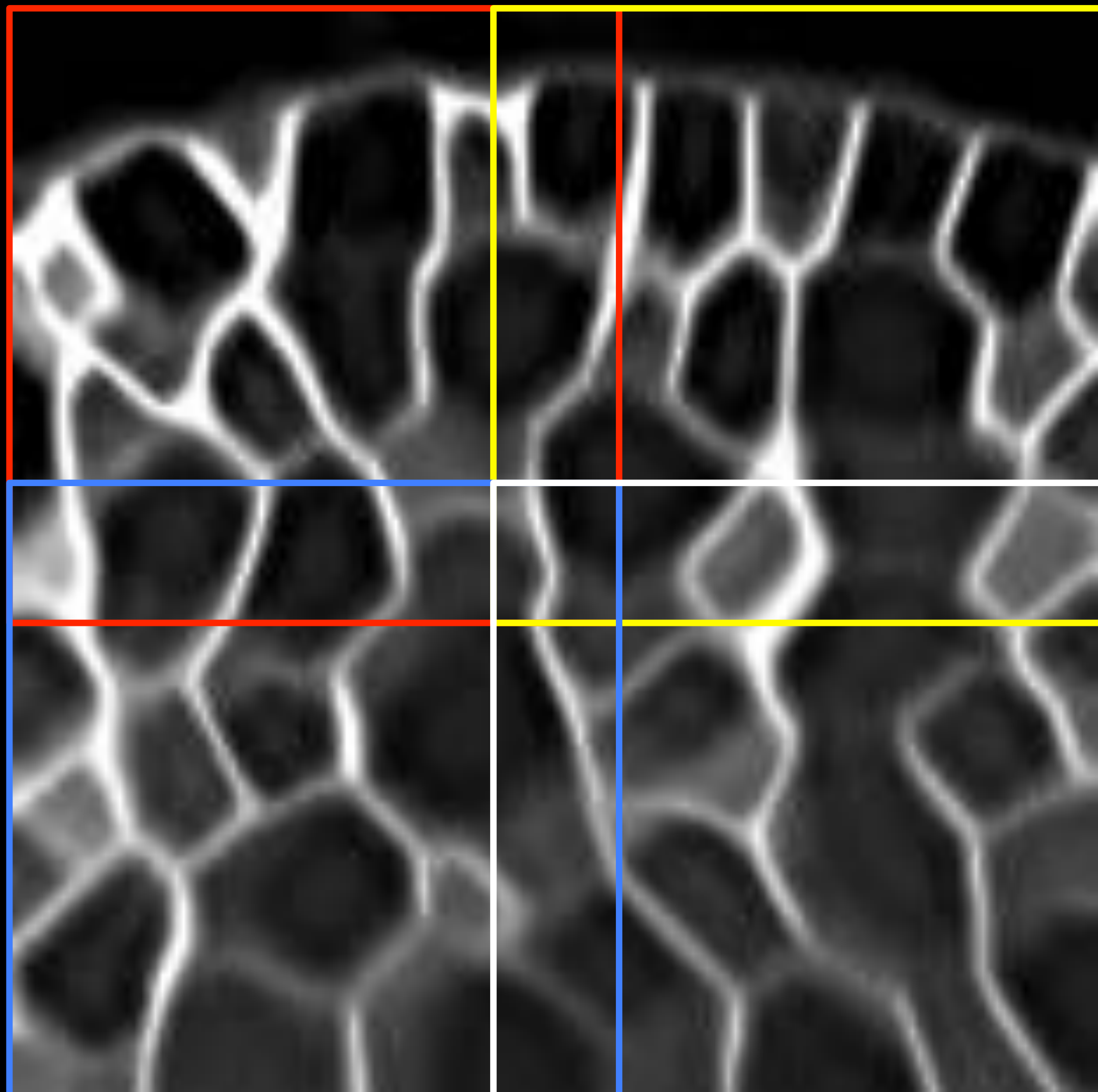


Modeling Consensus

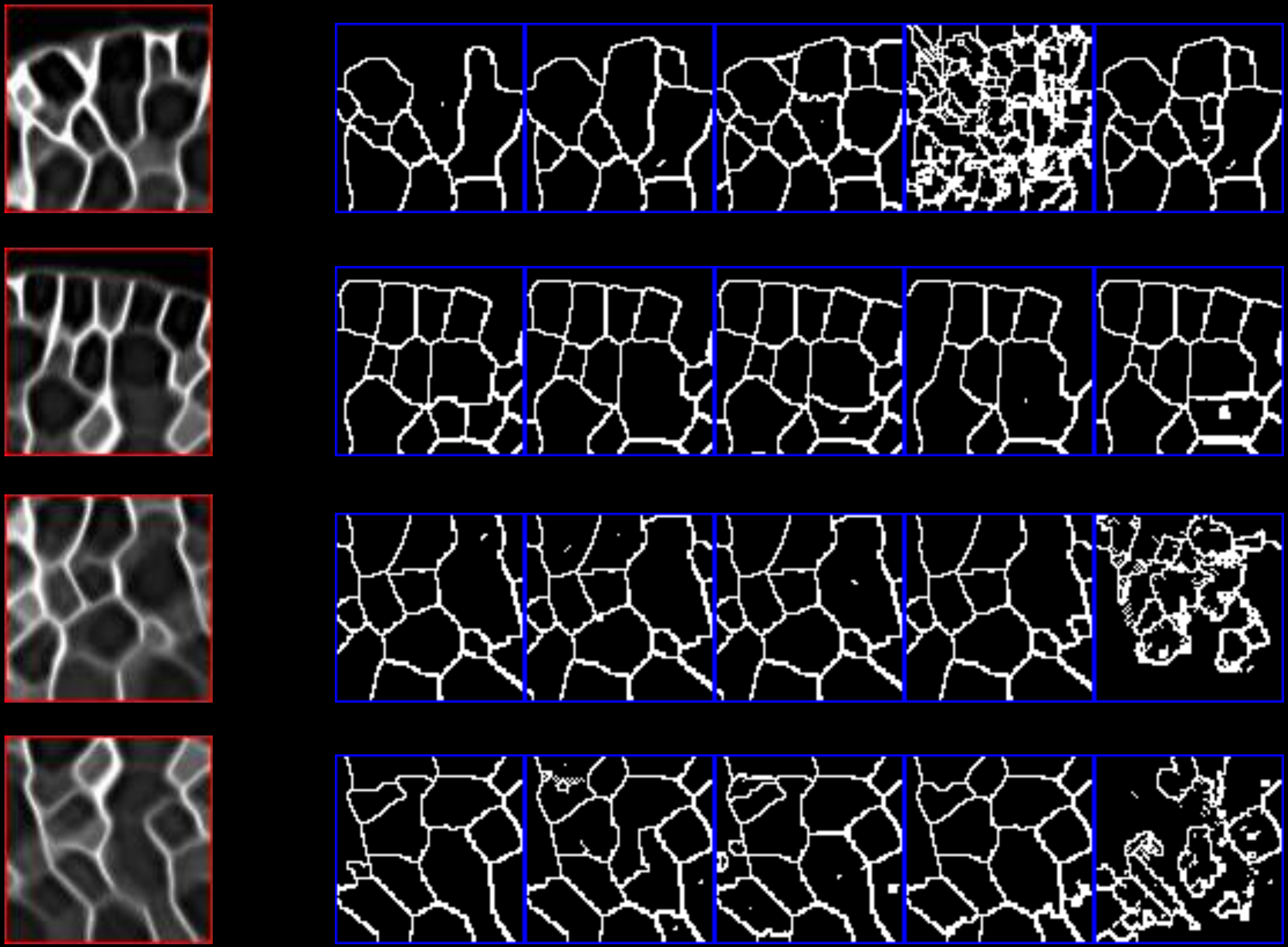
CoSe – Collaborative Segmentation – generates multiple segmentations for the same image.
A consensus segmentation judiciously combine these into one or a few solutions with ***superior quality***.

A high school in Ithaca working in the segmentation of images of meristem, sepal, and pavement cells, ran by Adrienne Roeder.

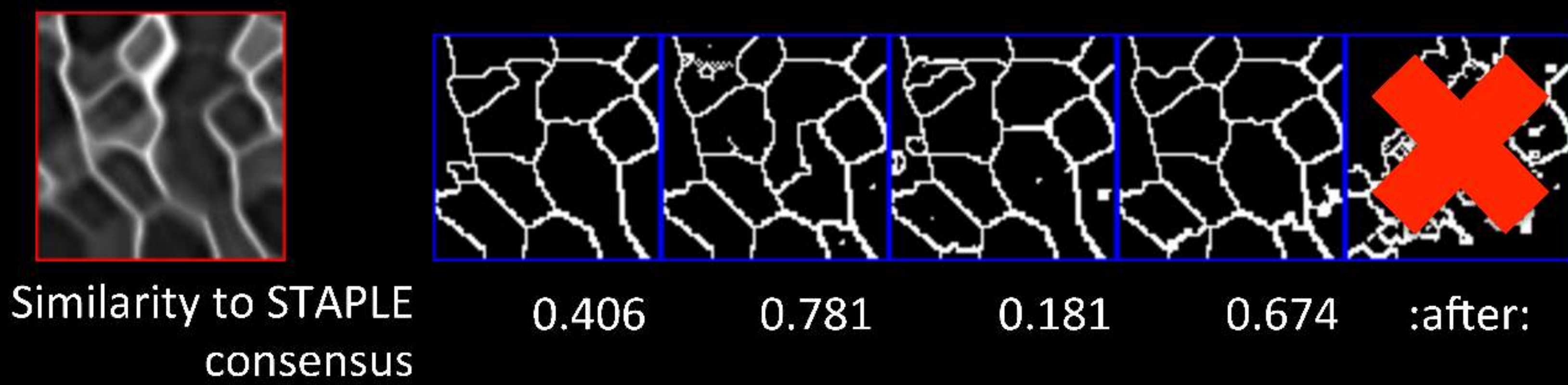
Divide & Conquer

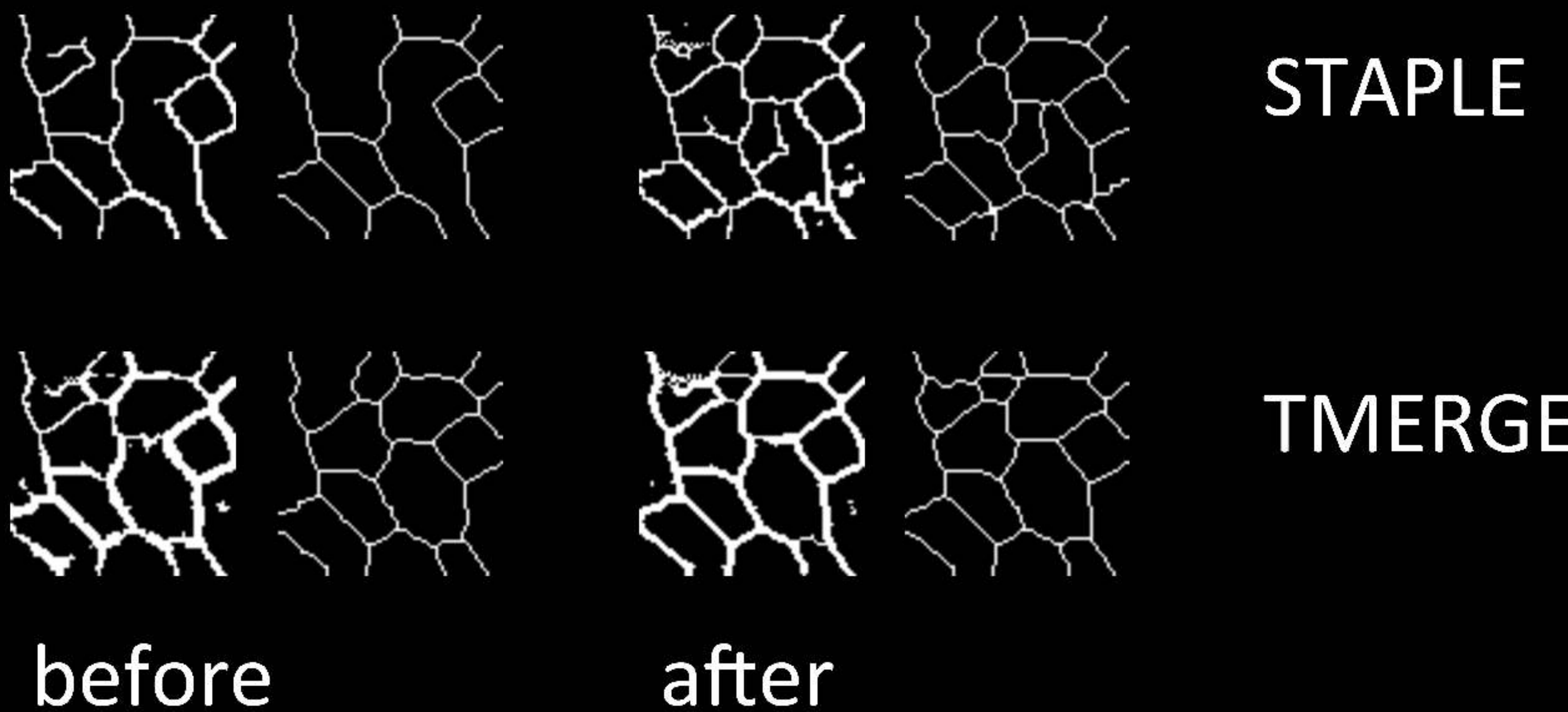


Meristem data – segmentations per tile

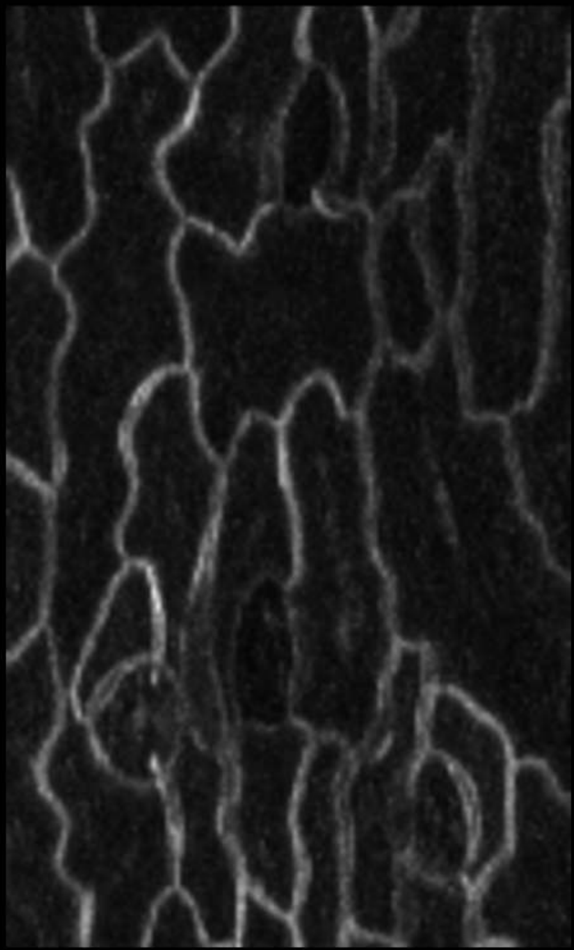


Modeling Consensus – without outliers

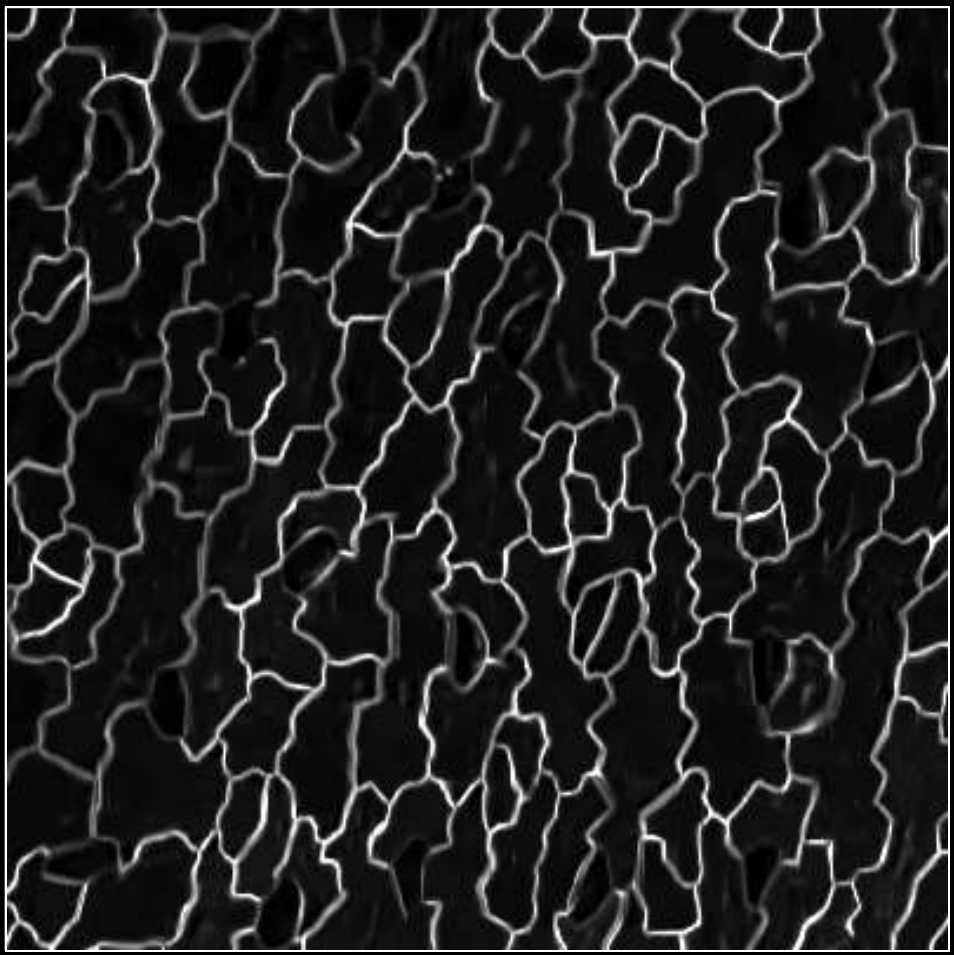


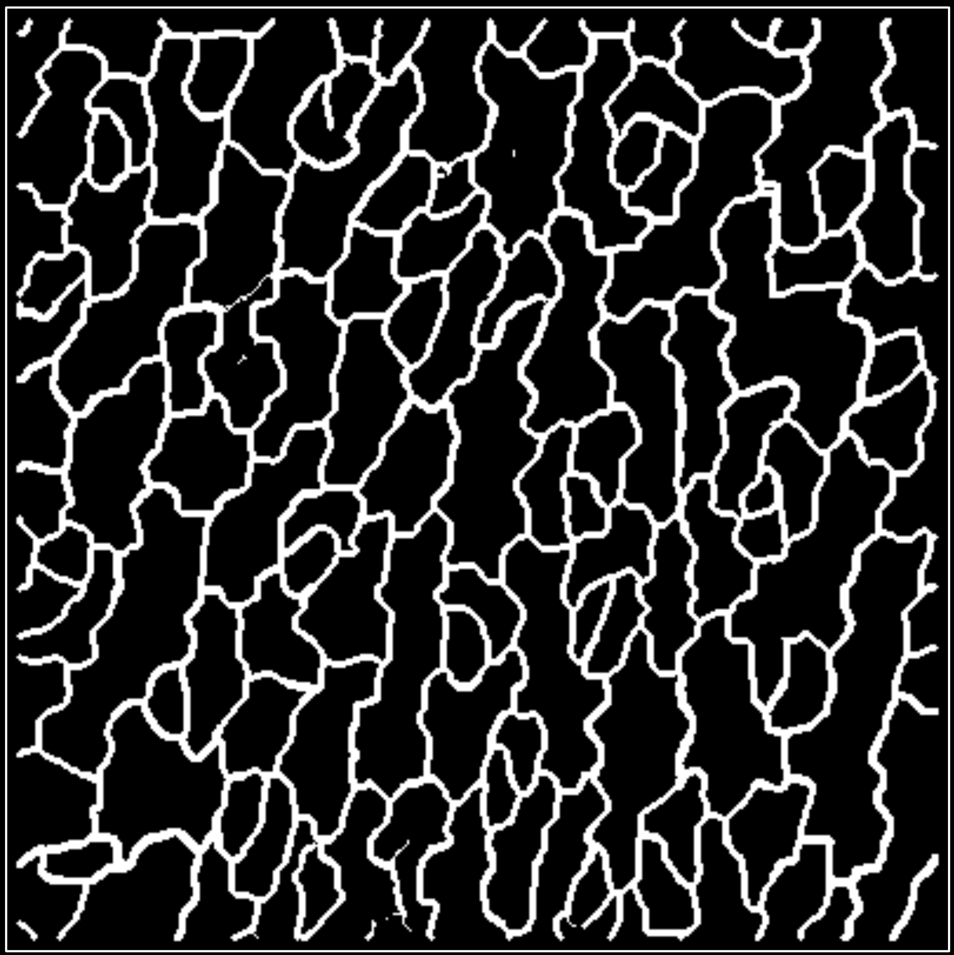


Modeling Consensus – results for sepal segmentation



Modeling Consensus – combined results for sepal





Modeling Consensus

Combining algorithms & model parameters & data

Let m_i be the $i-$th segmentation model $whose\ parameters$ θ_k can assume m distinct values

$$S_i = \{s_i^k | \mathcal{M}_i(\theta_k, x) \rightarrow s_i^k, k = 1..m\}$$

$$|S_i| = m_i, \quad i = 1..n$$

$$\theta_k \neq \theta_j, k \neq j \quad \rightarrow \quad possible\ that\ s_i^k = s_i^j, k \neq j$$

Modeling Consensus

Combining algorithms & model parameters & data

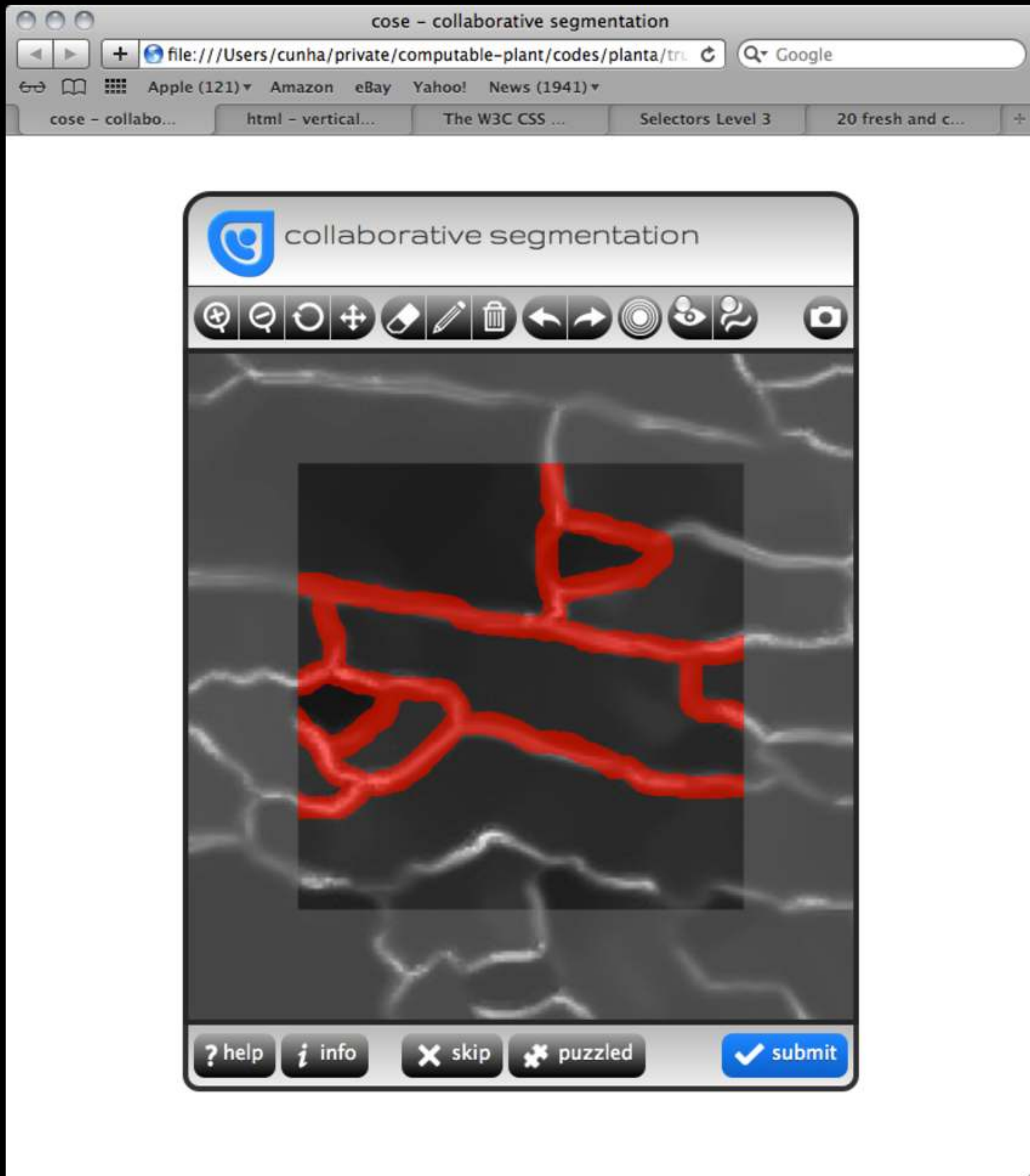
$$S = \bigcup S_i, \qquad |S| = m_1 + ... + m_n$$

$$S = \{s_j | s_j \in S_i\}$$

$$\begin{array}{l} \quad optimization\ problem: \\ find \quad s^\star = \ argmin_s \quad d(s, s_j) \end{array}$$

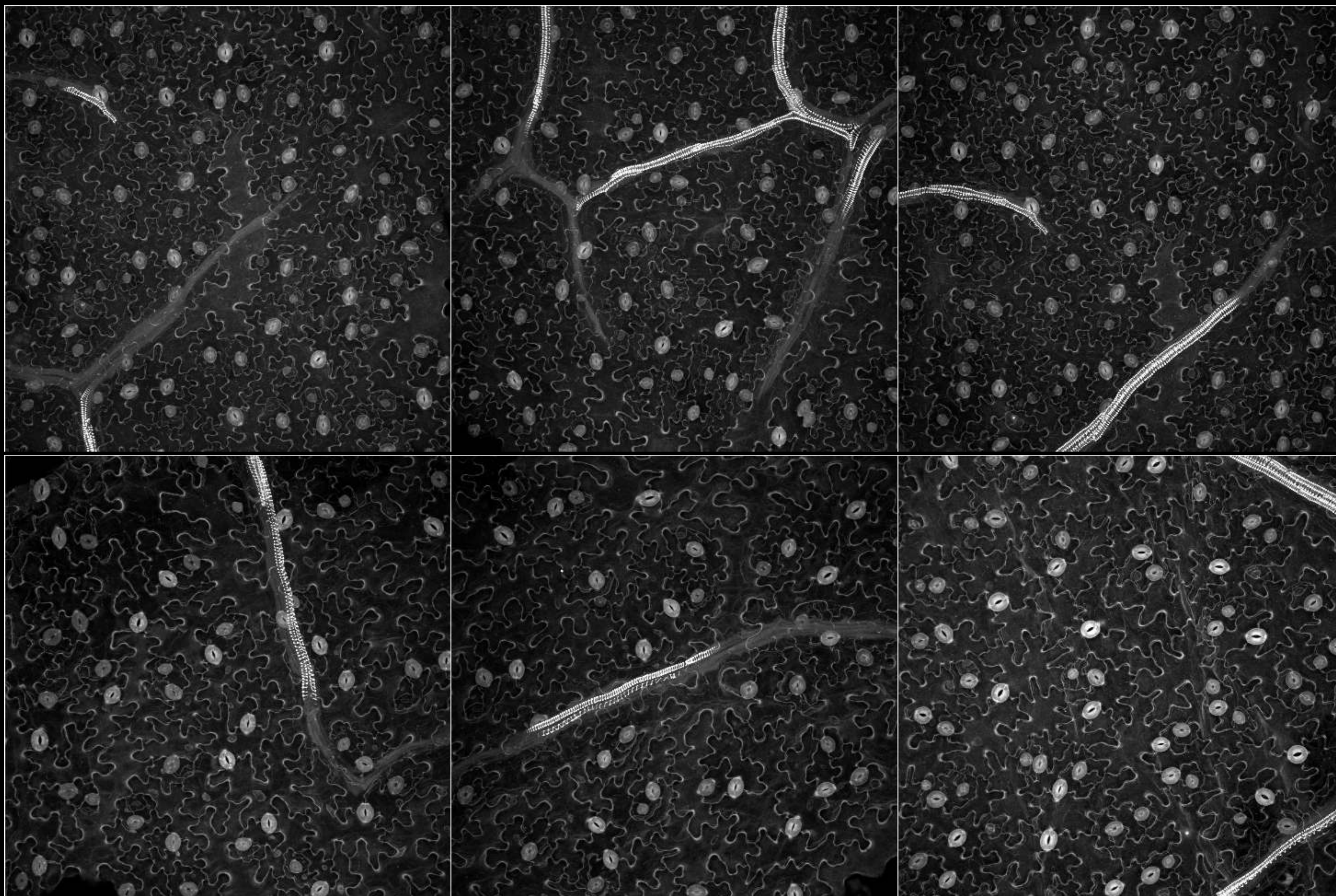
$$d \ = measure\ of\ distance$$

Collaborative segmentation http://cose.cacr.caltech.edu



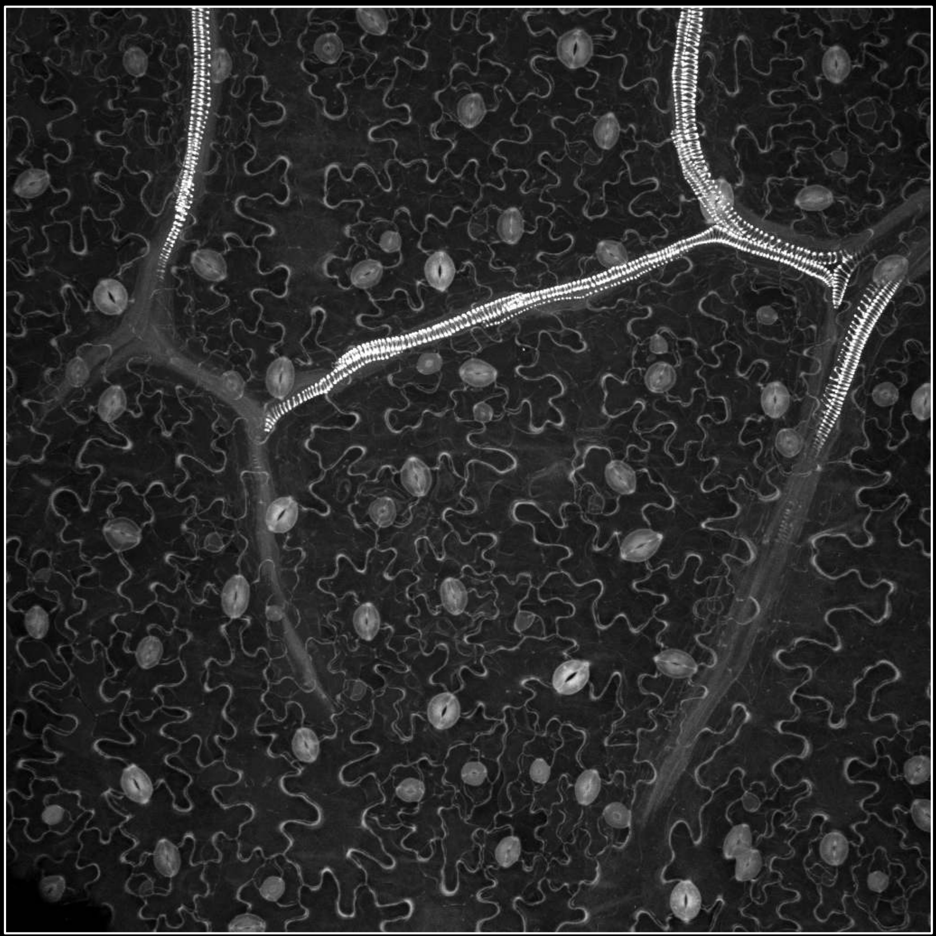
Better max intensity projection (Meyerowitz lab) - Arun

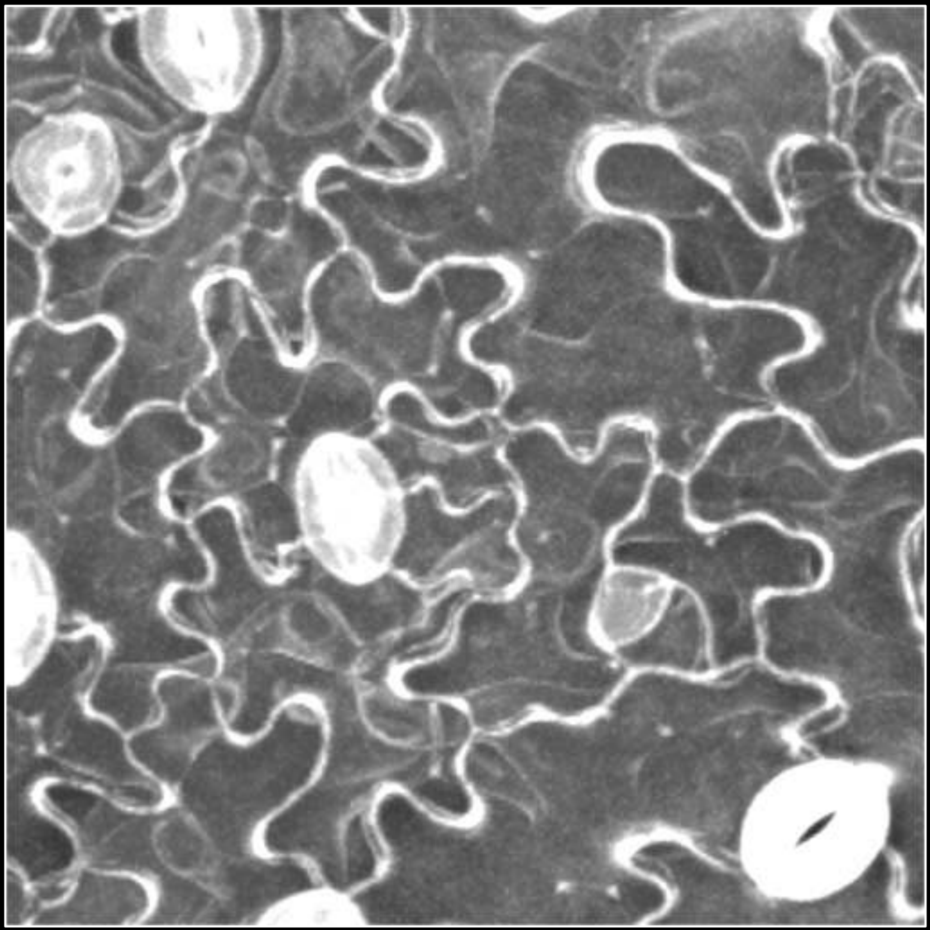
Projections have defects which reduce the number of segmented cells.



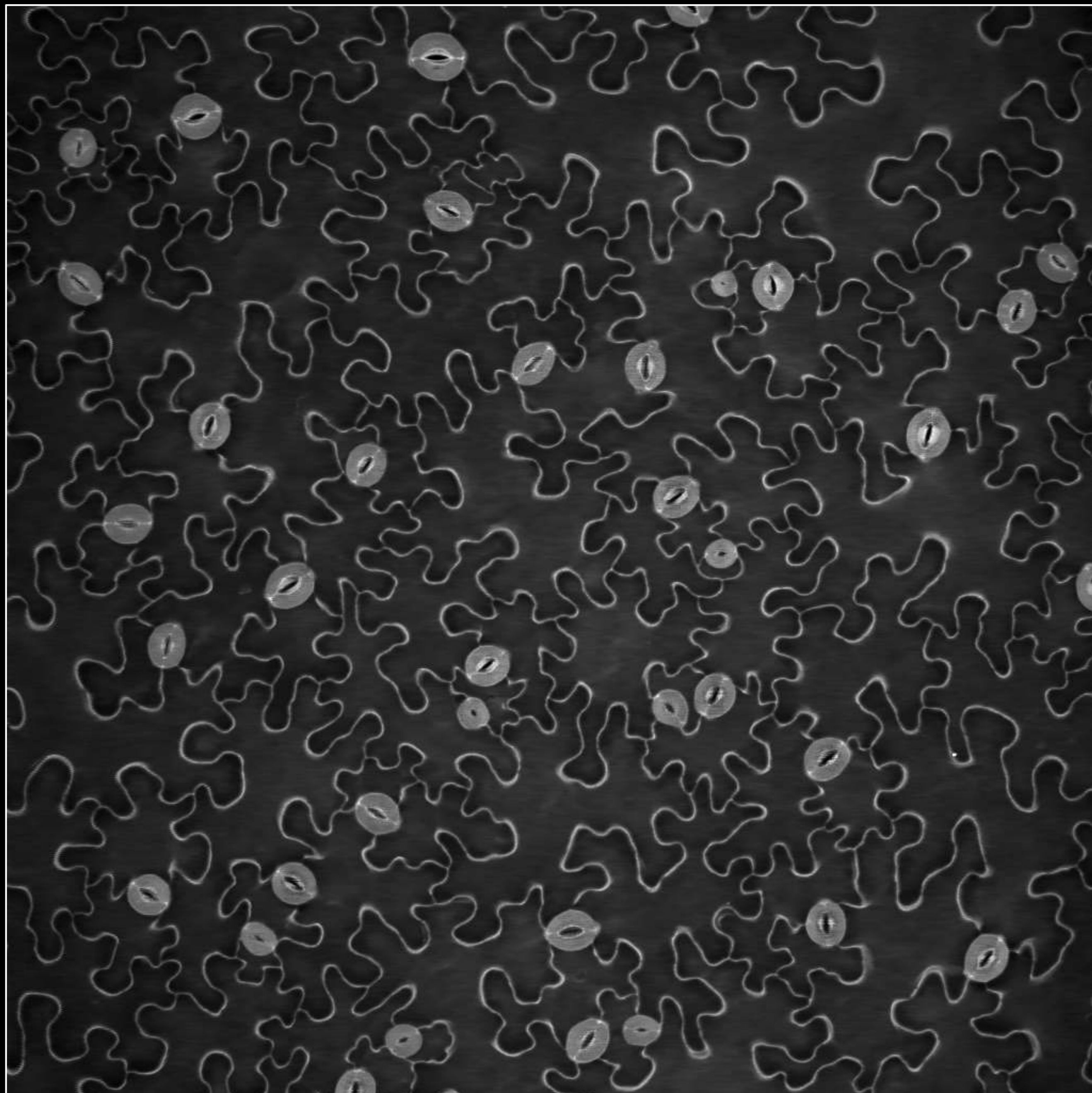
Better max intensity projection (Meyerowitz lab) - Arun

Diverse patterns and plenty of spurious data:



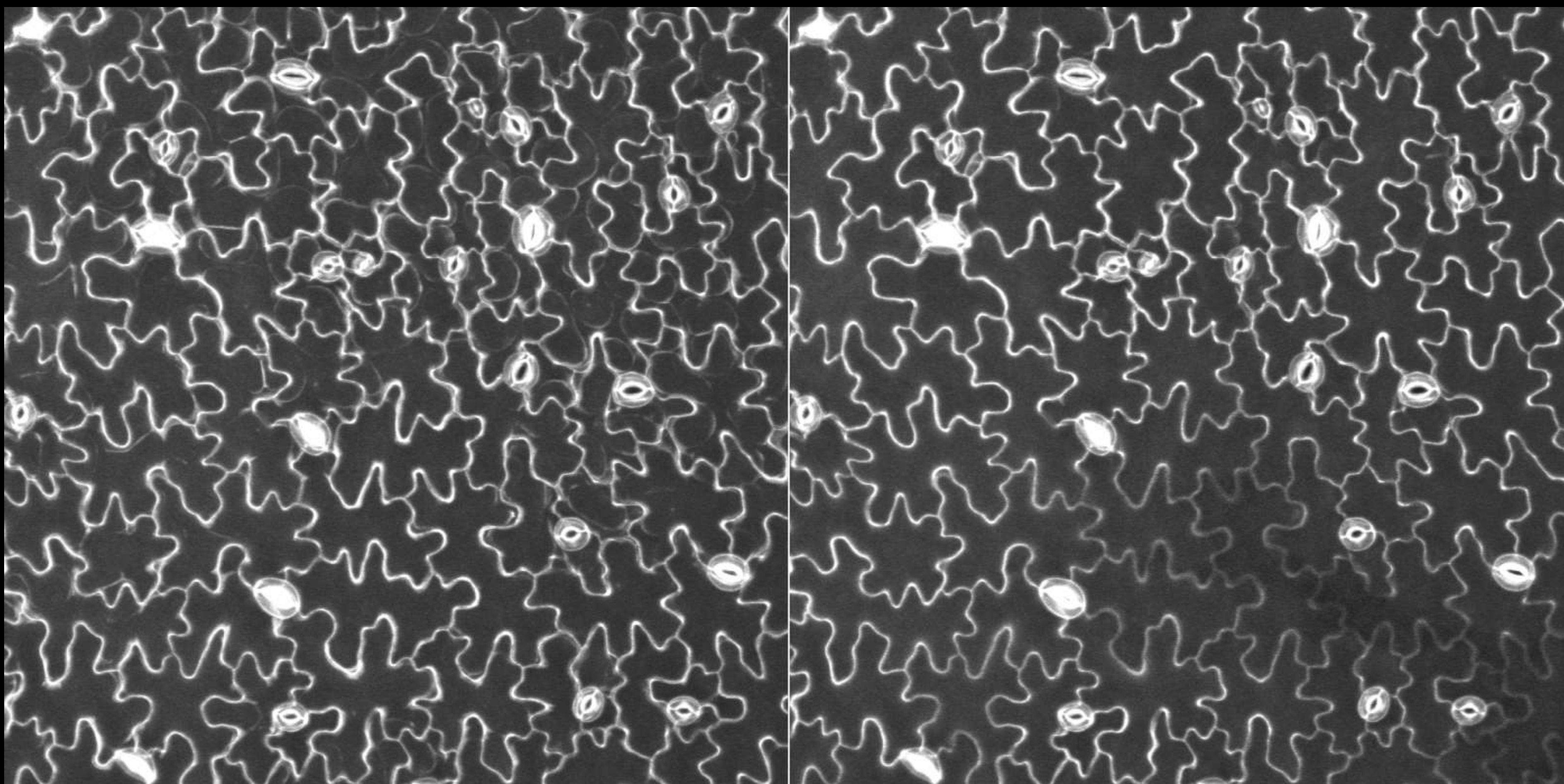


Better max intensity projection (Meyerowitz lab) - Arun



Better max intensity projection (Meyerowitz lab) - Arun

Segmenting pavement cells after improving projection



Better max intensity projection – segmented cells

Classify DNA spots (Stathopoulos lab) - Leslie

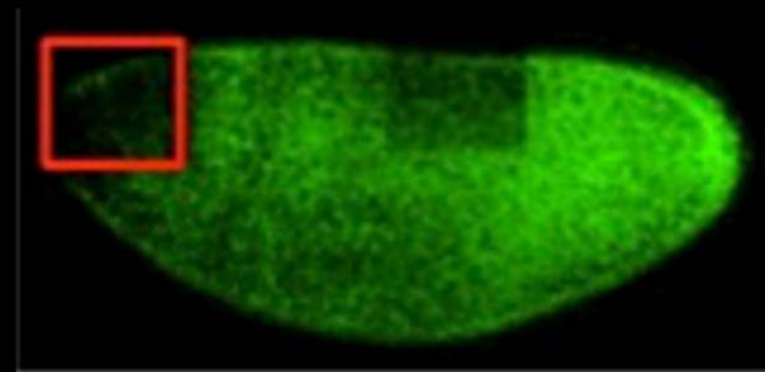
Chromatin is a complex of DNA, RNA, and proteins that looks like beads on a string when imaged on an optical microscope.

We study how a certain gene in early *Drosophila* embryo is controlled by different chromatin conformations which is analyzed based on the relative location of three segments of a single DNA.

This triplet, (3,5,P), is identified with the help of fluorescently labeled DNA probes and *fluorescence in-situ hybridization*, FISH, and images acquired with a widefield microscope.

We perform spot detection to find the location of triplets and classify their spatial configuration.

Drosophila embryo



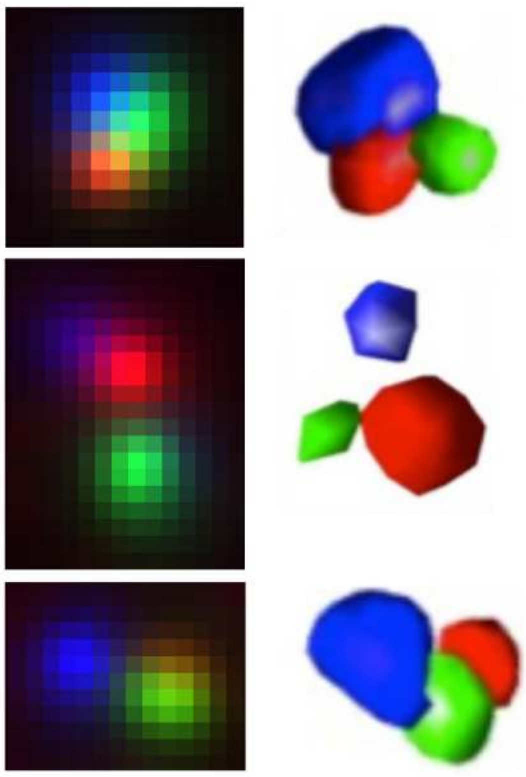
Drosophila melanogaster.
We acquire multiple image stacks with 2048 x 2048 x 101 voxels (red square above) at different positions on the embryo. Each stack covers one layer of cells and it is as parallel as possible to the embryo surface thus allowing the usage of projections for validation of results. Stack resolution, in microns, is 0.065 x 0.065 x 0.100.

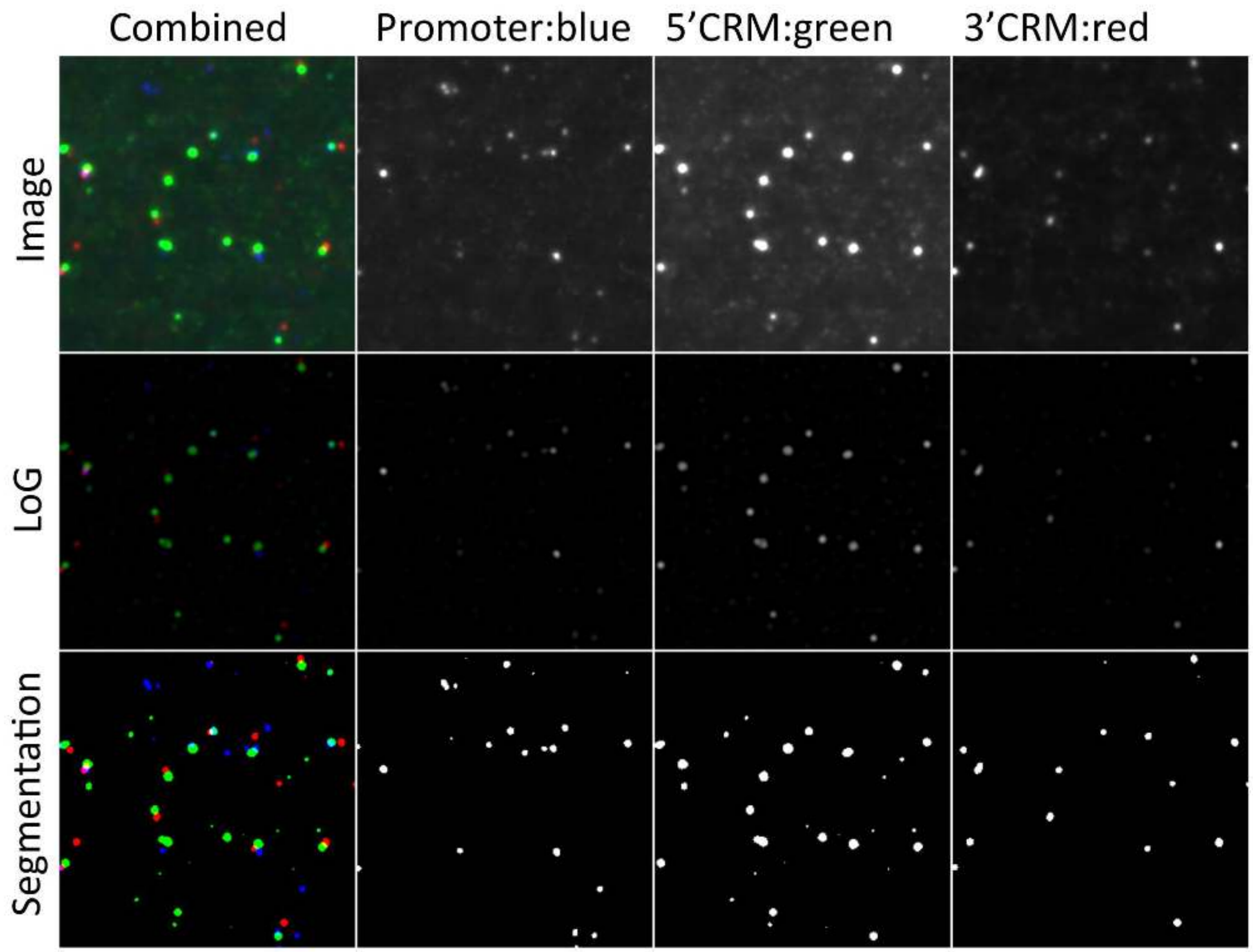
Spots as Gaussian in 3D

Spot detection

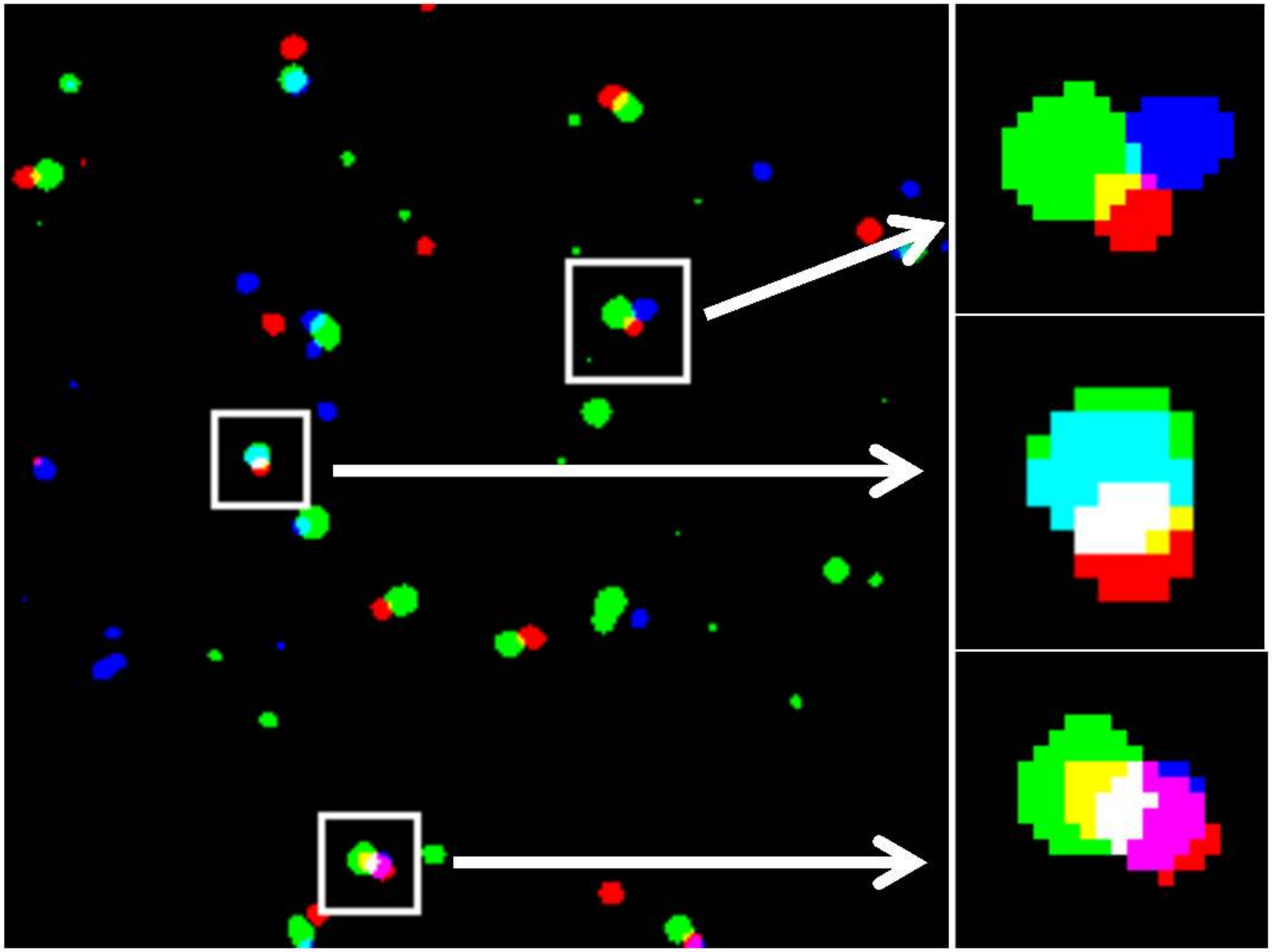
Computed separately for each channel corresponding to 3′CRM, 5′CRM, and Promoter probes and then assembled back to determine the formation of valid triplets (3,5,P) based on Euclidean distances in 3D.

Spots as Gaussian blobs in 3D





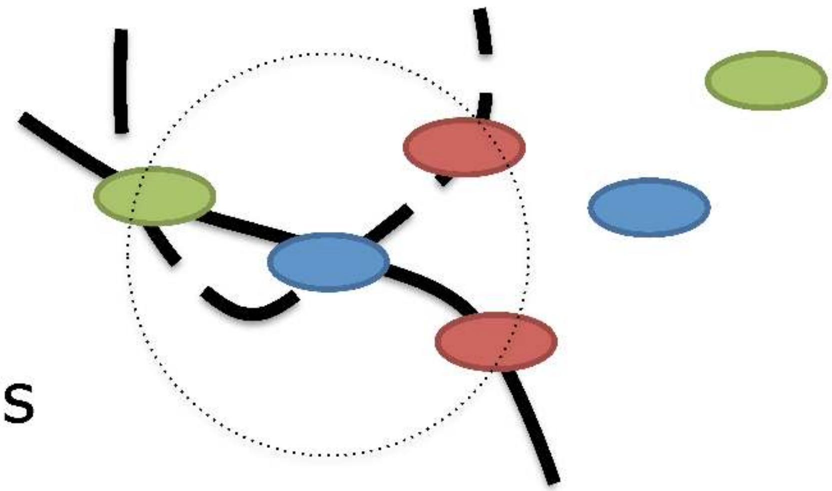
Triplet formation



Whenever three spots of different color are close enough they form a triplet (3,5,P) provided no other spot makes the triplet ambiguous (ill resolved). A maximum intensity projection might help validate triplet formation.

Beads on a string

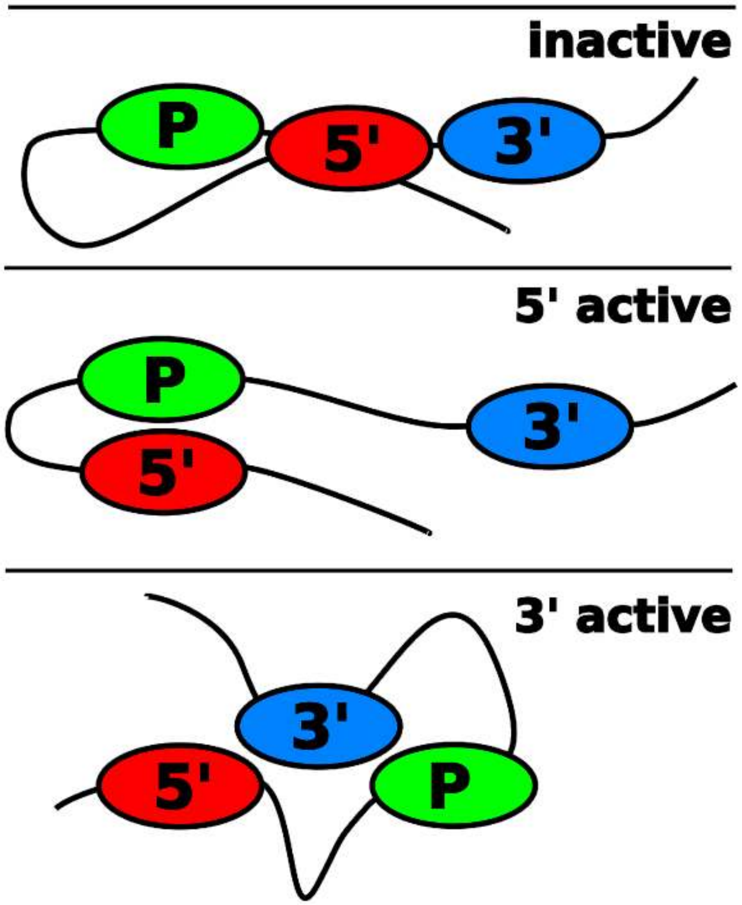
We image the beads – 3,5,and P – but not the strings. We can't resolve which string (DNA) a bead (DNA segment) belongs to thus we can't always tell triplets apart: two beads of the same type might encroach on the neighborhood of a different bead type.



Which of the two black strings (solid and dashed) is the correct DNA for the blue bead on the left?

Triplet classification

Conformations



Triplets are classified based on the pairwise proximity of its components compared to a **touching threshold *t***, chosen according to experiments, e.g. *t = 0.2um*. Each triplet is then labeled with a three digits binary word:

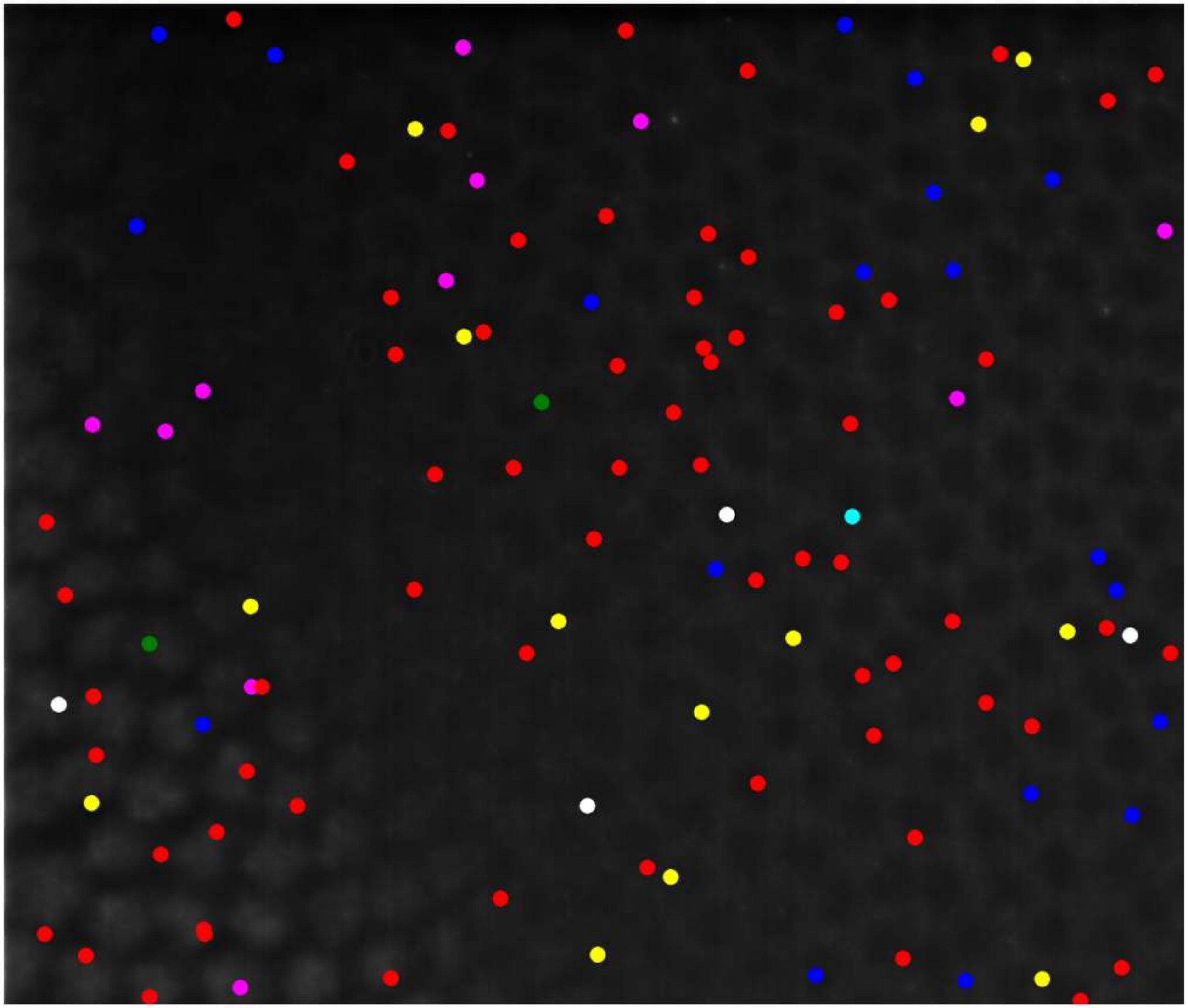
bit_1 = if ***dist(3,5) > t*** then ***0*** else ***1***
bit_2 = if ***dist(3,P) > t*** then ***0*** else ***1***
bit_3 = if ***dist(5,P) > t*** then ***0*** else ***1***

The eight possible configurations are color coded according and triplets are painted back in the image with their respective colors for a visual inspection and analysis.

Triplet classification

Color coded classification for triplets on image:

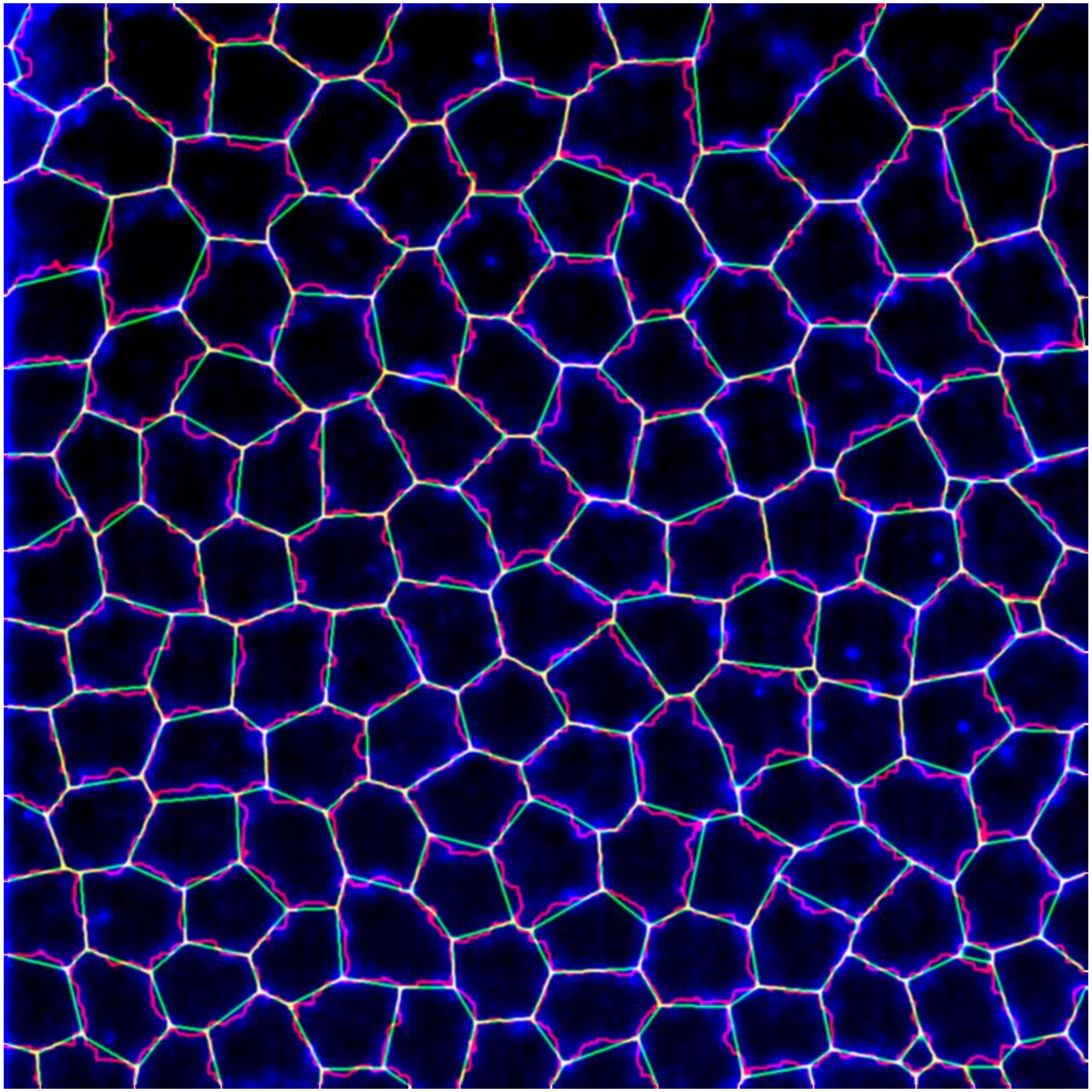
Bits	Count	Color
000	4	White
001	64	Red
010	2	Green
100	1	Cyan
011	19	Blue
101	11	Magenta
110	0	Gray
111	13	Yellow



Processing steps

1. **Deconvolution** with theoretical point spread function (Huygens)
2. **Filtering with Laplacian of Gaussian**, LoG, and tuned parameters (kernel size and variance) per slice
3. **Local adaptive threshold** (foreground decision is based on distance to neighborhood mean and jump size) can account for different amplitudes of Gaussians
4. **Connected components in 3D**, where spots sufficiently close at adjacent slices are merged
5. **Find centroid of spots** in each channel to execute range search on complementary channels
6. **Form triplets** discarding ambiguous cases
7. **Classify triplets** according to the proximity between its pair components 3-5, 3-P, 5-P as compared to a target touching distance.

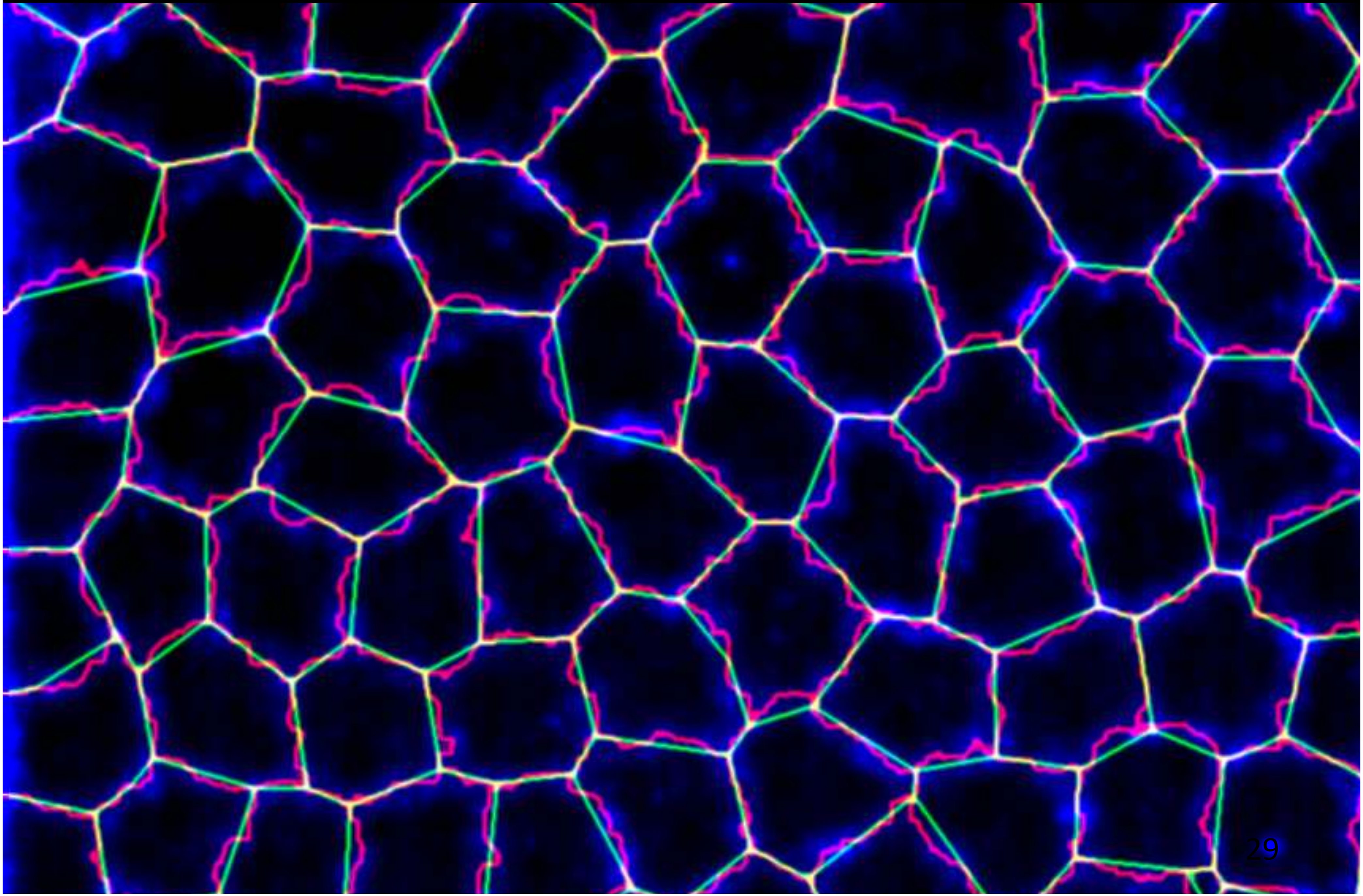
Cell segmentation to improve results



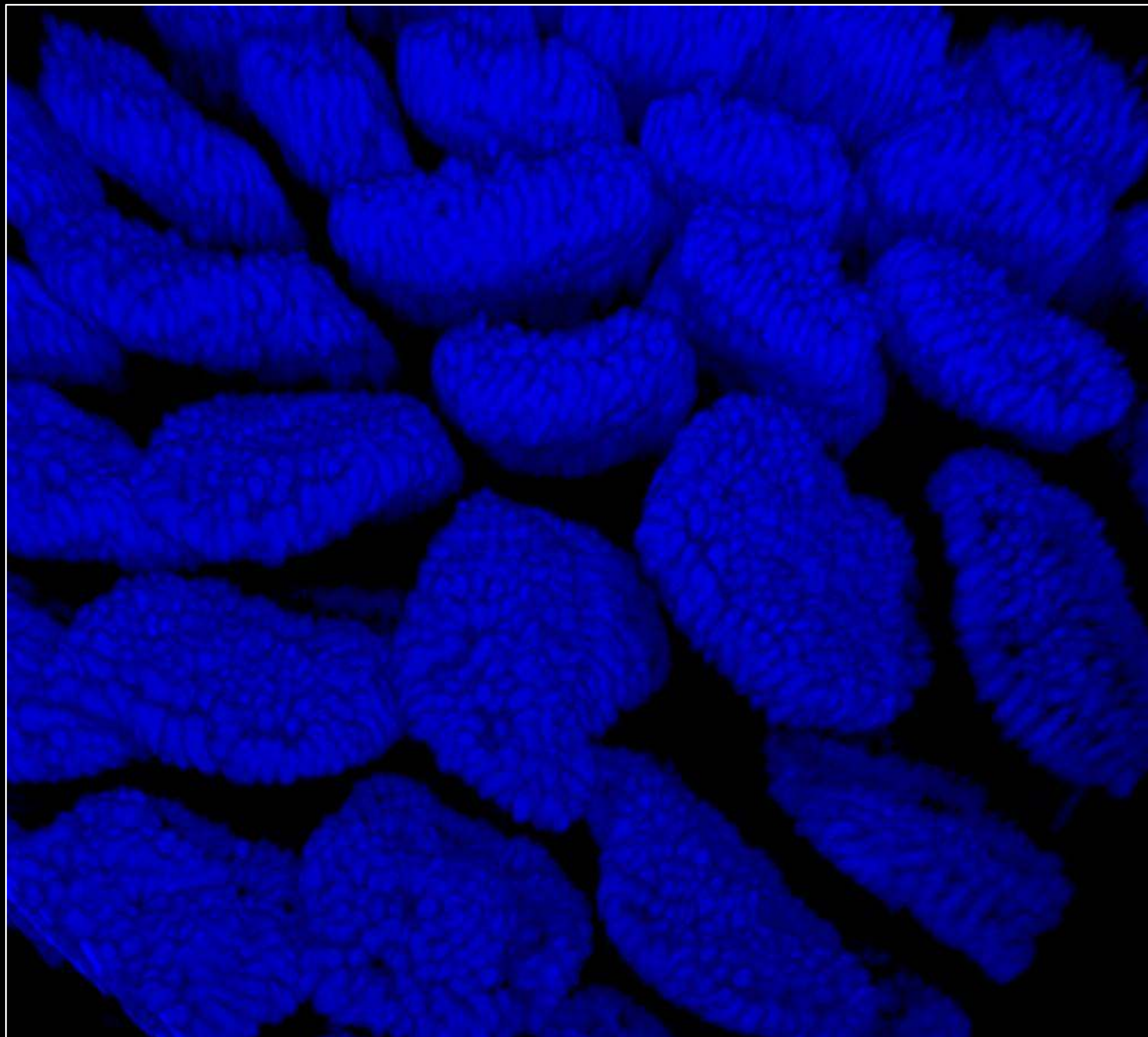
Improvements

The idea is to isolate each cell and then do a local search for triplets which might occur at most twice per cell. For that, we segment cells using their plasma membrane signal, as shown in the picture on the left.

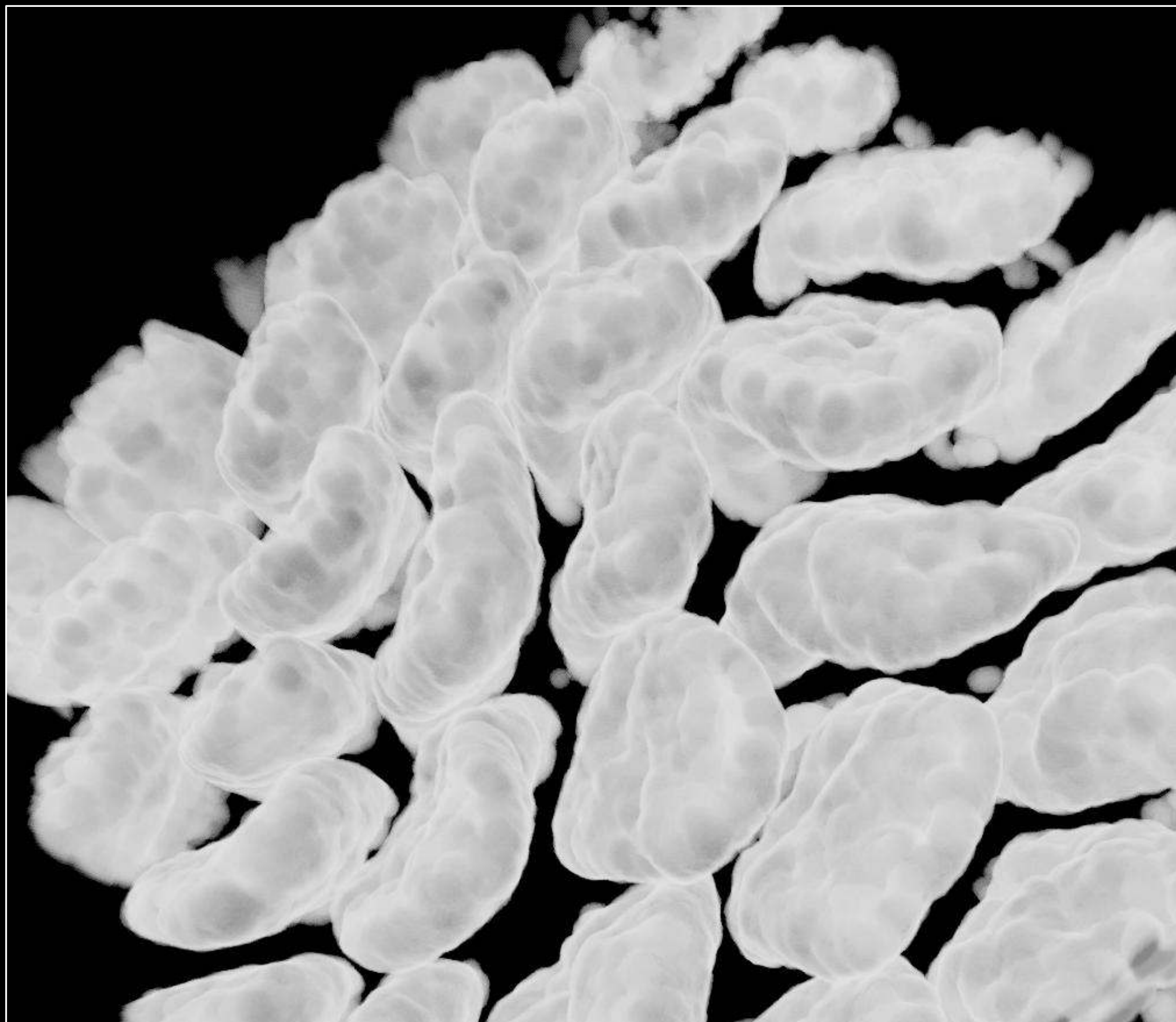
Cell segmentation to improve results



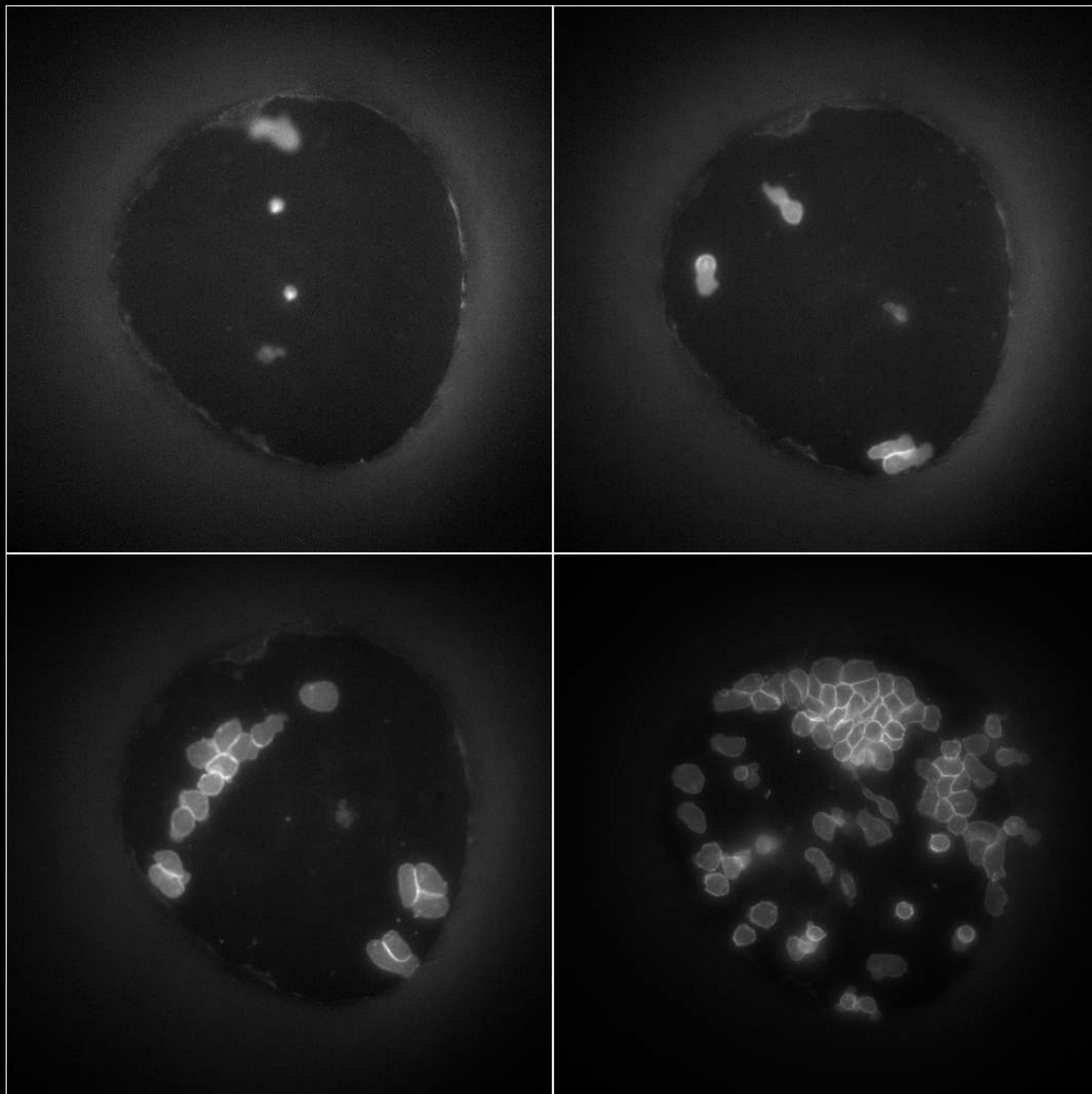
Villi segmentation (Ismagilov lab) - Octavio



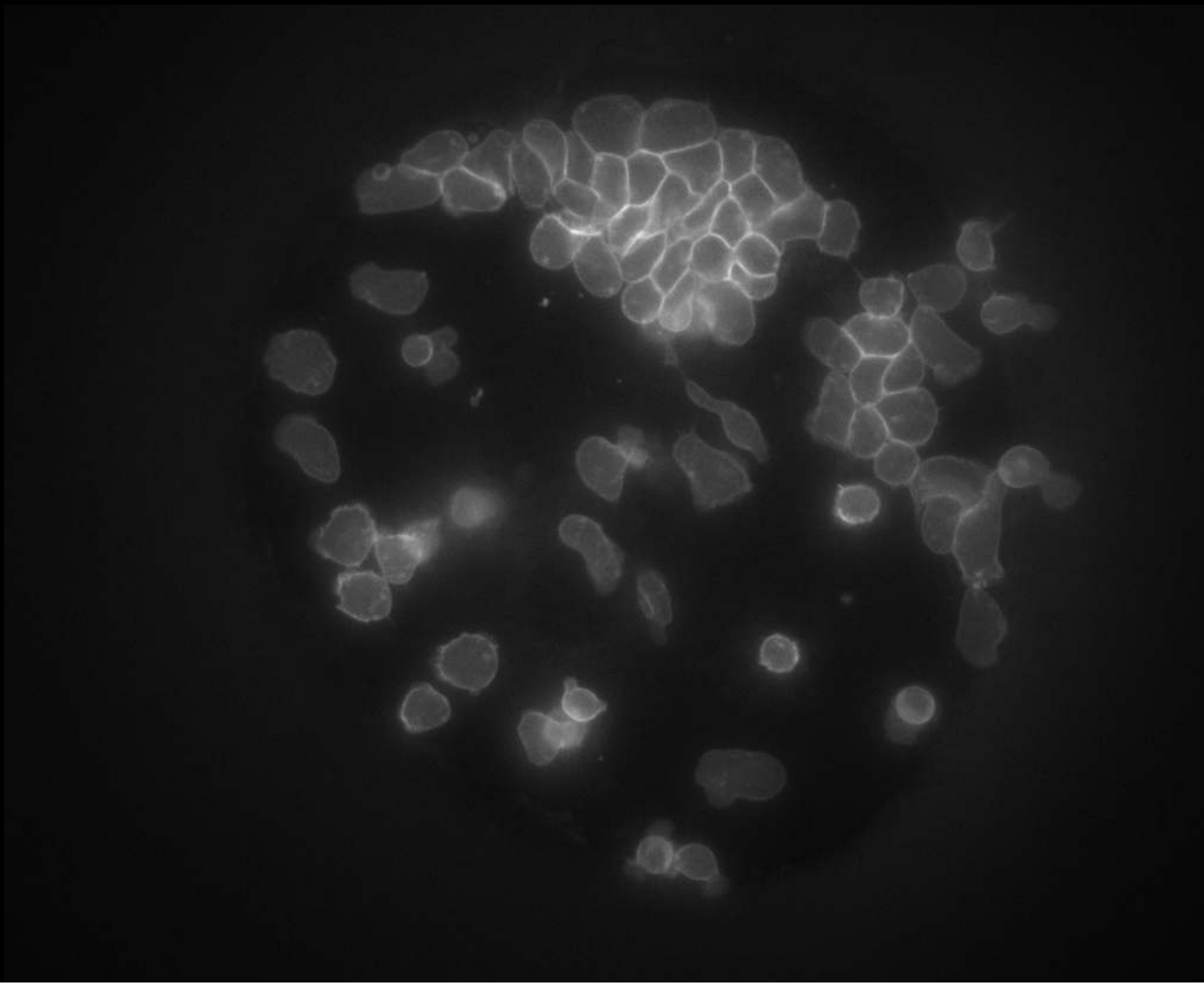
Villi segmentation (Ismagilov lab) - Octavio



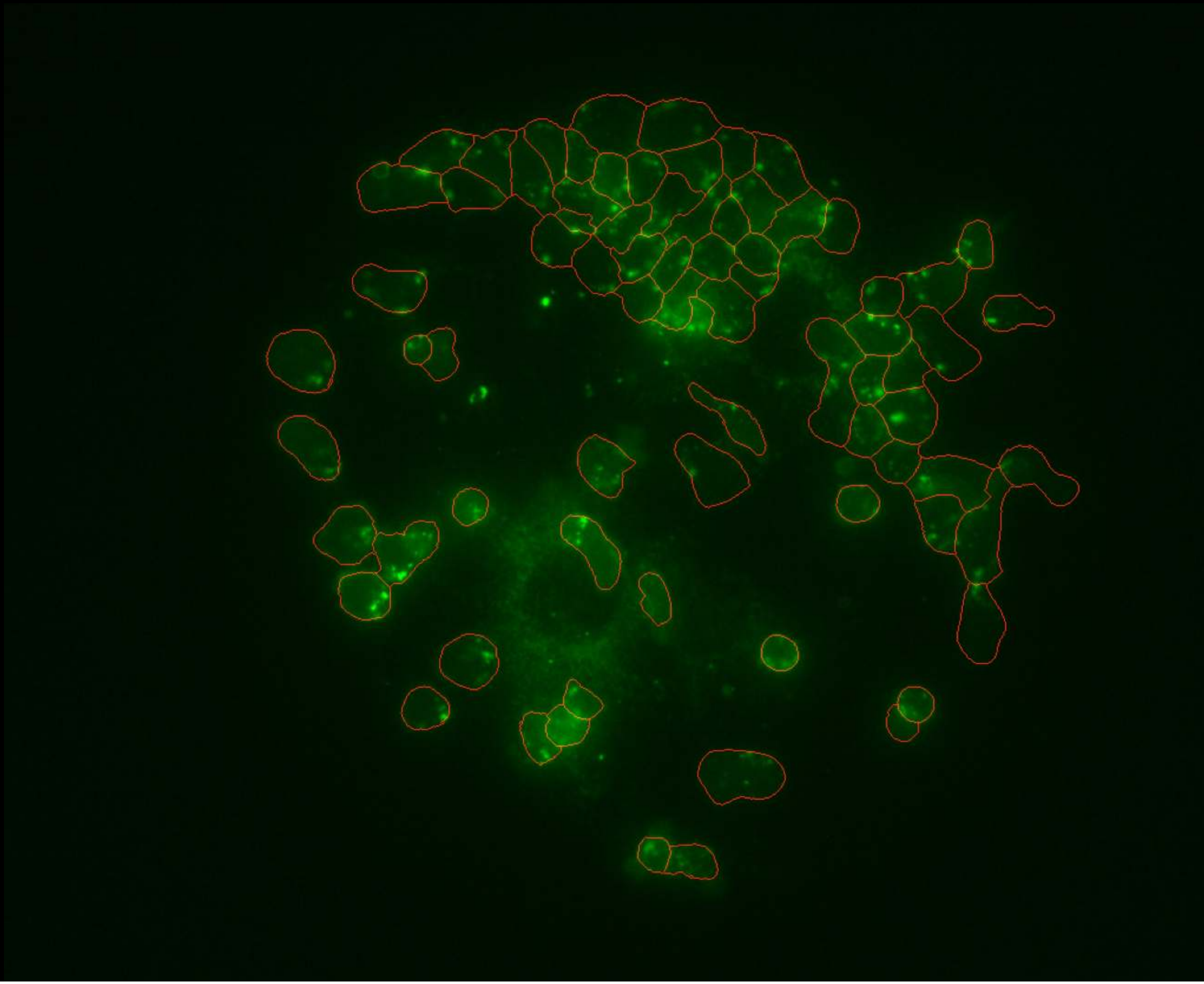
Analysis of T-cells (Rothenberg lab) – Mary Yui



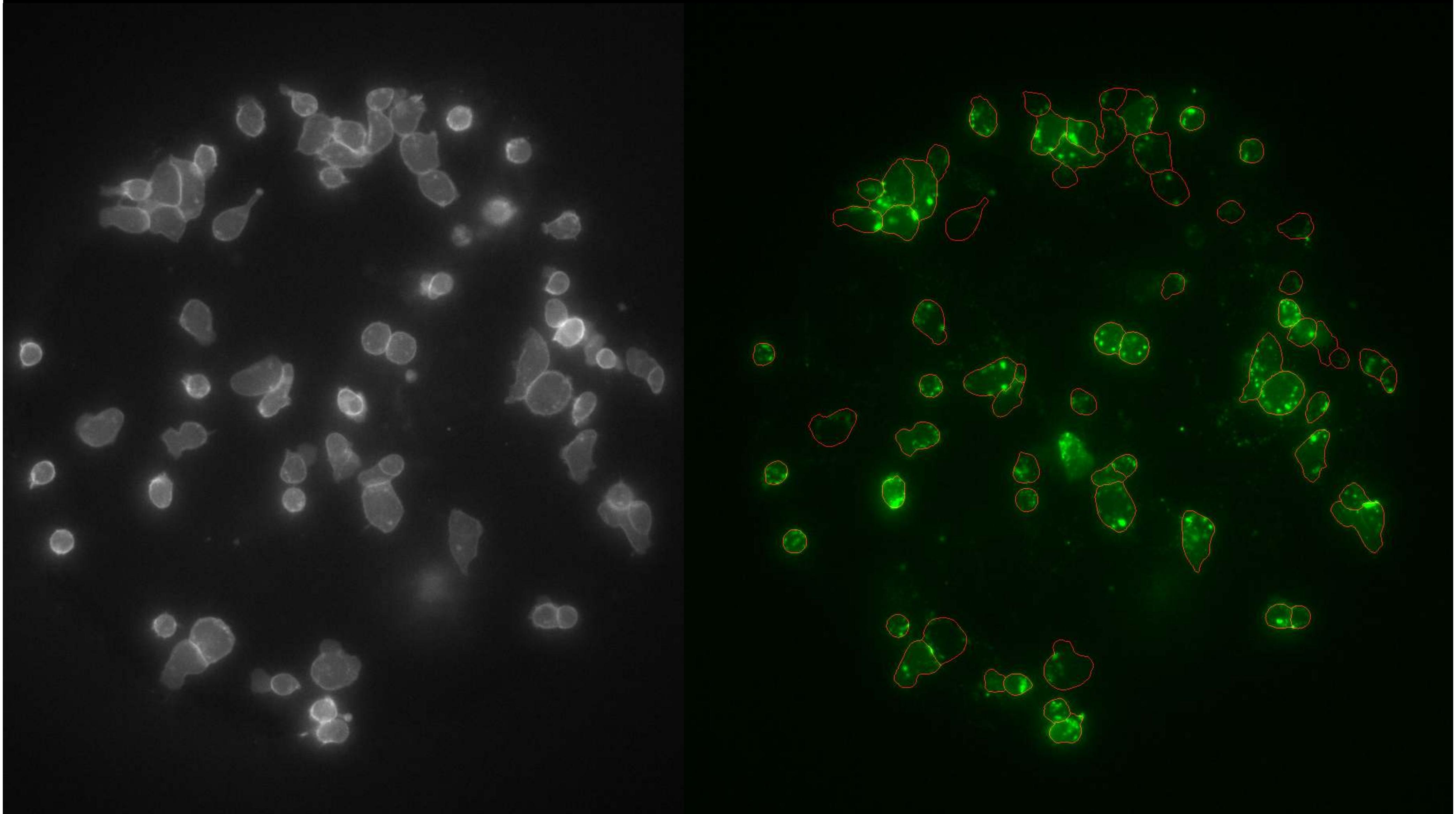
Analysis of T-cells (Rothenberg lab) – Mary Yui



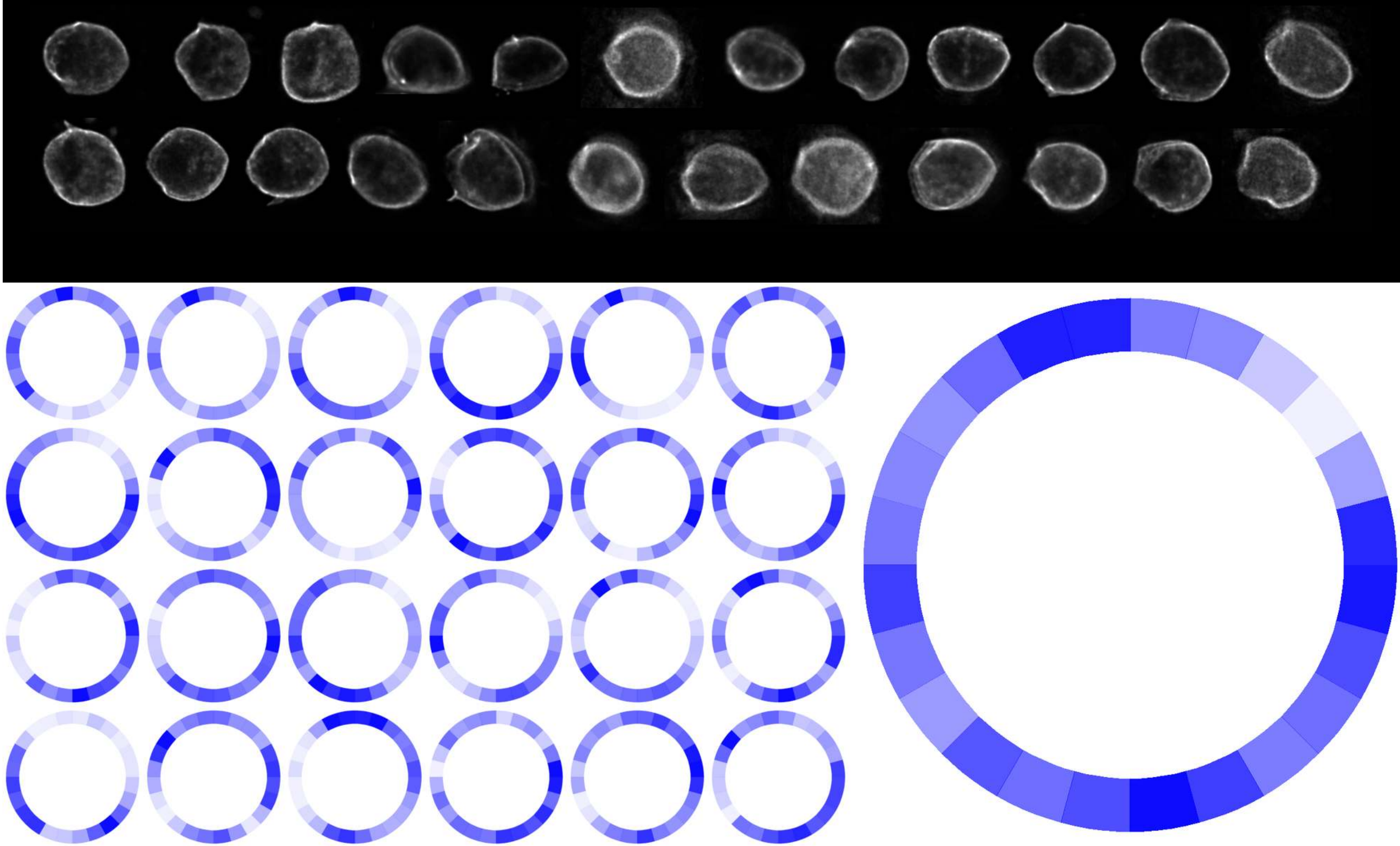
Analysis of T-cells (Rothenberg lab) – Mary Yui



Analysis of T-cells (Rothenberg lab) – Mary Yui



Signal on cell membrane (Sternberg lab) - Mihoko

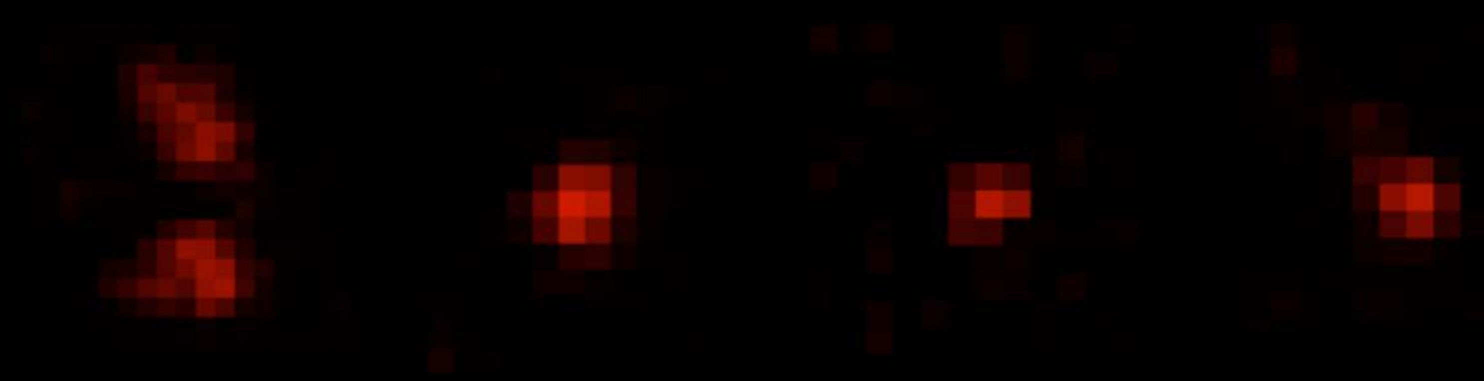


Spot detection, RNA (Pierce lab) – Harry, Maayan

Amplification time for smHCR in chicken embryos

EphA4 (B1: 27 pairs & B2: 27 pairs)

63x objective, pixel size (x,y) = 264 nm, z = 400 nm



90 min 60 min 45 min 30 min

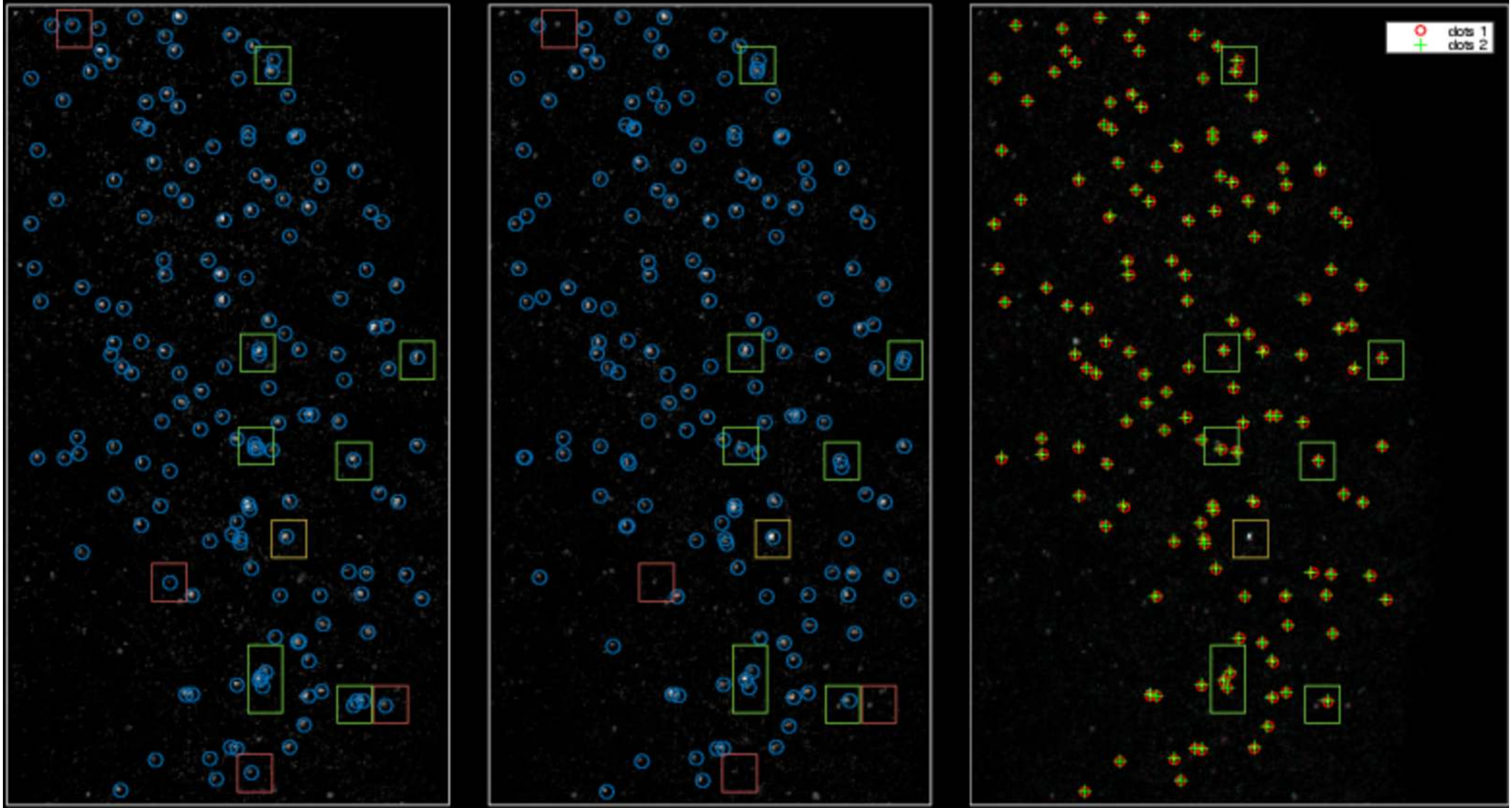
Spot detection, RNA (Pierce lab) – Harry, Maayan

Dots in Channel 1 ROI

Dots in Channel 2 ROI

Colocalized dots

Spot detection, RNA (Pierce lab) – Harry, Maayan

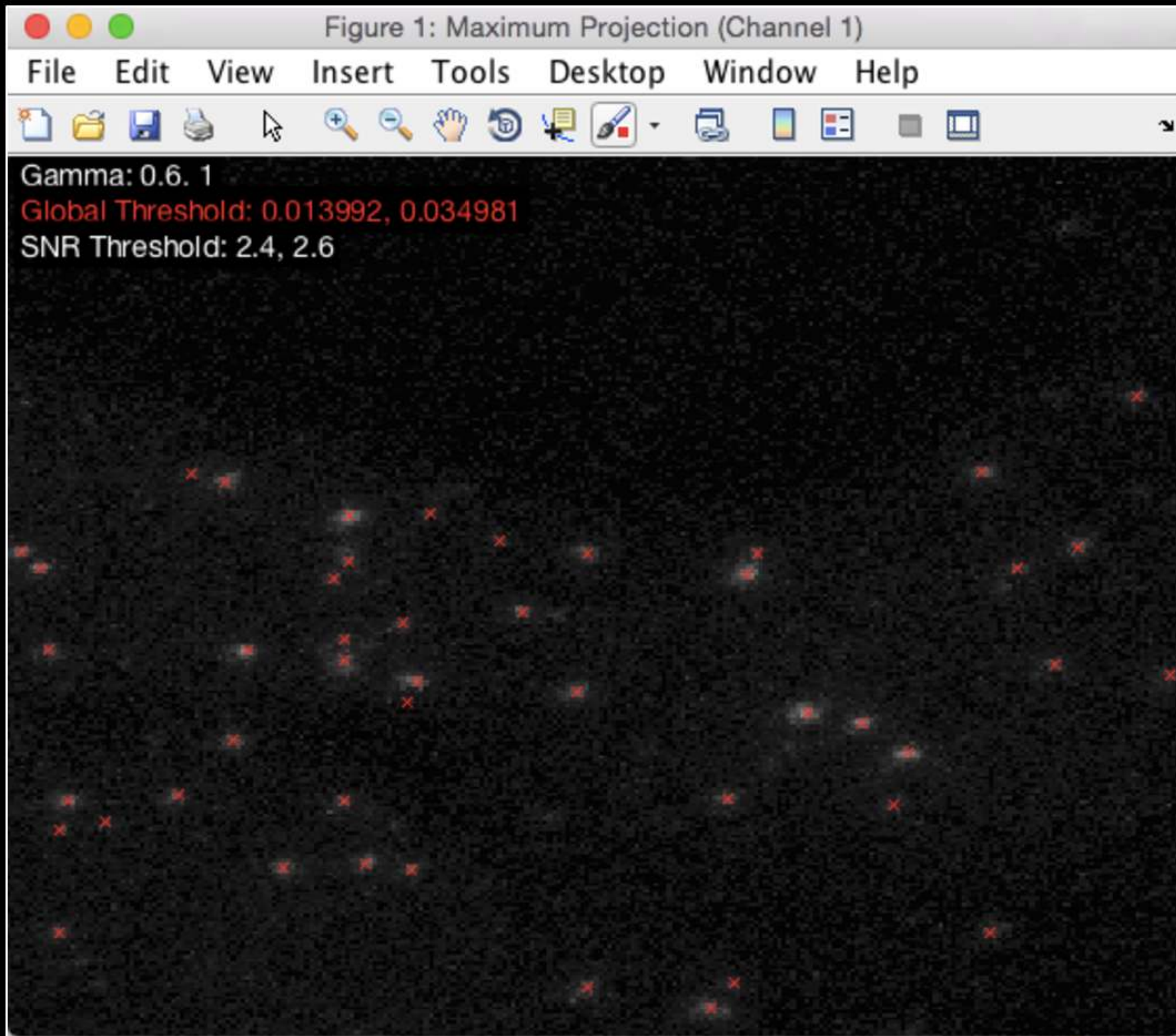


Dots in Channel 1 ROI
160 dots
88% colocalization

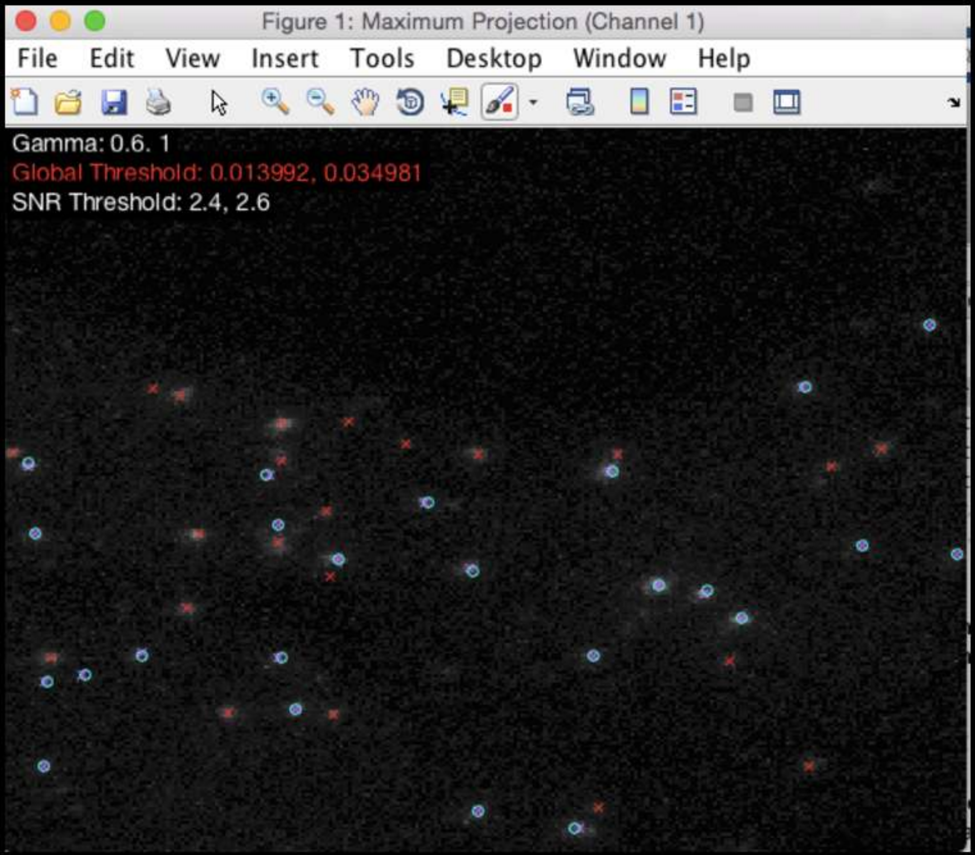
Dots in Channel 2 ROI
154 dots
91% colocalization

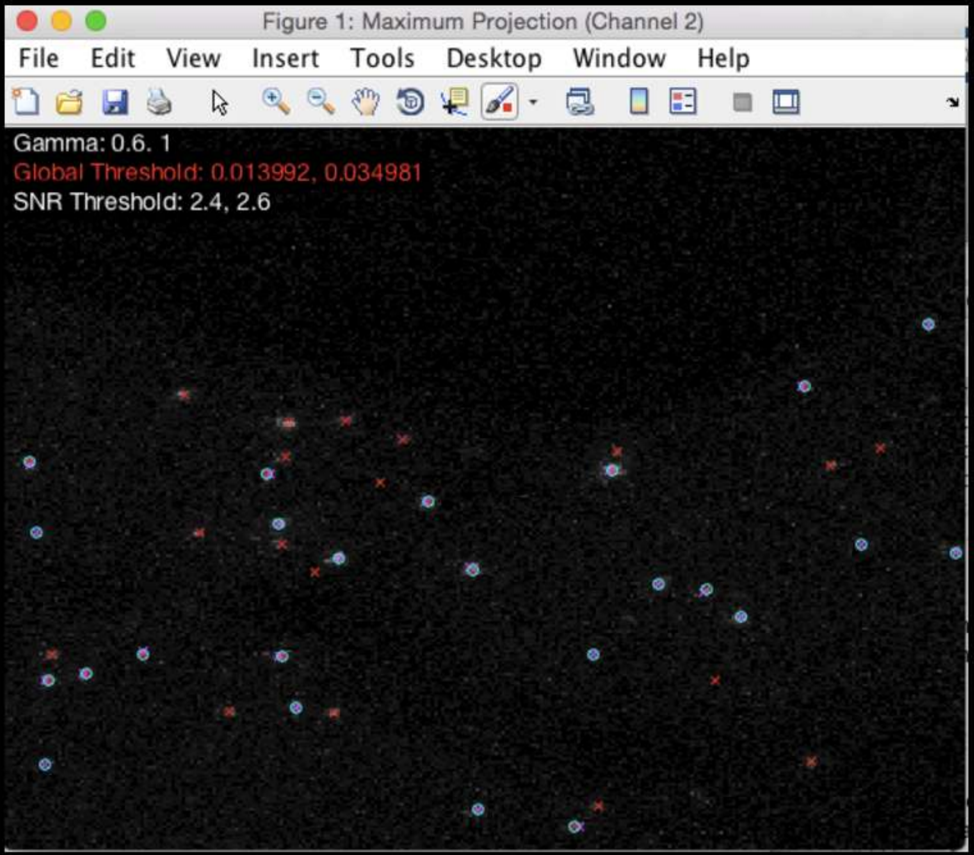
Colocalized dots
140 dots

Spot detection, RNA (Pierce lab) – Harry, Maayan



Spot detection, RNA (Pierce lab) – Harry, Maayan





Spot detection, RNA (Pierce lab) – Harry, Maayan

----------------------- Files -----------------------
File Channel 1: C1-BRAF-signal1
File Channel 2: C2-BRAF-signal1

--------------------- Parameters ---------------------
Gamma: 0.600000, 1.000000
Global Threshold: 0.013992, 0.034981
SNR Threshold: 2.400000, 2.600000

---------------------- Results ----------------------
Total Detections Channel 1: 146195
Total Detections Channel 2: 155348

Thresholded Detections Channel 1: 137
Thresholded Detections Channel 2: 105

Colocalized Detections Channel 1: 59
Colocalized Detections Channel 2: 59

Percentage Colocalized Detections Channel 1: 43.07%
Percentage Colocalized Detections Channel 2: 56.19%

Spot detection, RNA (Pierce lab) – Harry, Maayan

C1-BRAF-signal1_colocalizations.csv

id	scale	xpos	ypos	zpos	intensity	meanWindowIntensity	snrCriterion	matchDistance	radius	intensityRatio
44055	1.4142	1	92	10	0.0096745	0.015387	8.0585	2.1635	2	0.067741
45500	1.4142	311	343	10	0.041121	0.069639	10.02	2.9463	2	0.94889
46477	1.4142	378	511	10	0.043416	0.054831	13.345	2.3834	2	0.77409
46696	1.4142	366	554	10	0.038746	0.0607	14.221	2.3834	2	1.5889
46856	1.4142	202	583	10	0.044696	0.07907	10.965	2.3834	2	1.4316
47237	1.4142	568	650	10	0.04843	0.07097	15.246	2.1635	2	0.66297
47728	1.4142	413	729	10	0.016464	0.026296	9.264	2.3834	2	0.64287
52111	1.4142	380	535	11	0.026077	0.041464	10.57	0	2	1.0708
52200	1.4142	545	553	11	0.08073	0.13497	13.369	2.1635	2	1.4713
52319	1.4142	687	580	11	0.026956	0.042311	11.31	2.3834	2	0.6823
52471	1.4142	637	612	11	0.011645	0.019247	7.8353	2.1635	2	0.3622
52645	1.4142	335	649	11	0.027315	0.041293	11.818	3.1114	2	0.48376
56134	1.4142	634	429	12	0.011881	0.019201	7.7936	0	2	0.30622
56366	1.4142	583	483	12	0.030279	0.047161	11.011	1	2	1.0084
56531	1.4142	318	528	12	0.028856	0.067621	8.2684	2.3834	2	0.49423
56548	1.4142	444	532	12	0.012949	0.026204	2.9092	2.1635	2	0.21042
56579	1.4142	599	539	12	0.019233	0.032663	9.6375	0	2	0.86798
60207	1.4142	617	461	13	0.015382	0.025574	10.029	2.1635	2	0.34075
60397	1.4142	264	506	13	0.043624	0.064017	12.11	2.1635	2	1.223
60468	1.4142	483	524	13	0.028874	0.046593	11.288	2.3834	2	0.54214
60525	1.4142	290	539	13	0.019712	0.030142	7.6397	2.3834	2	0.55123