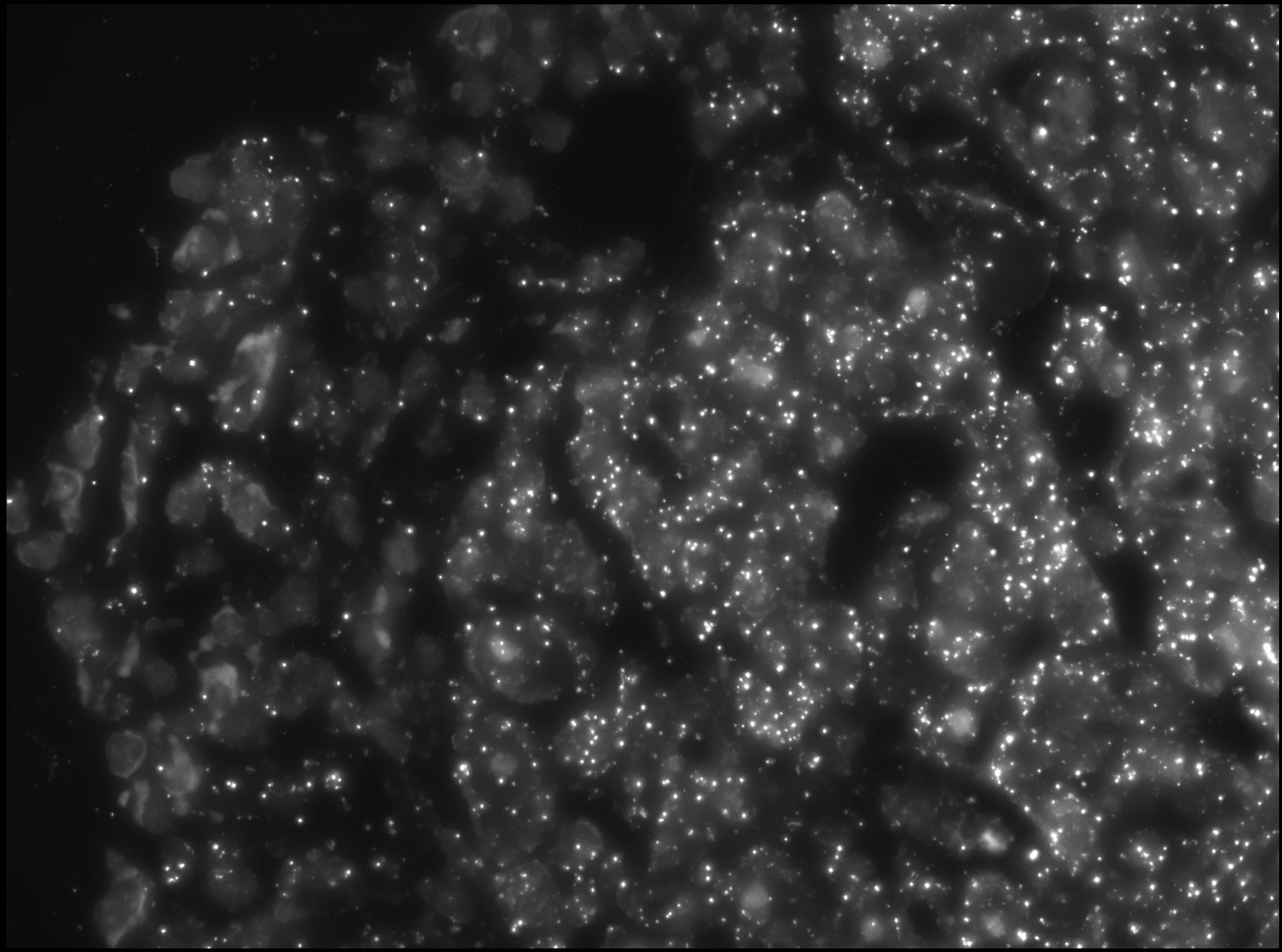
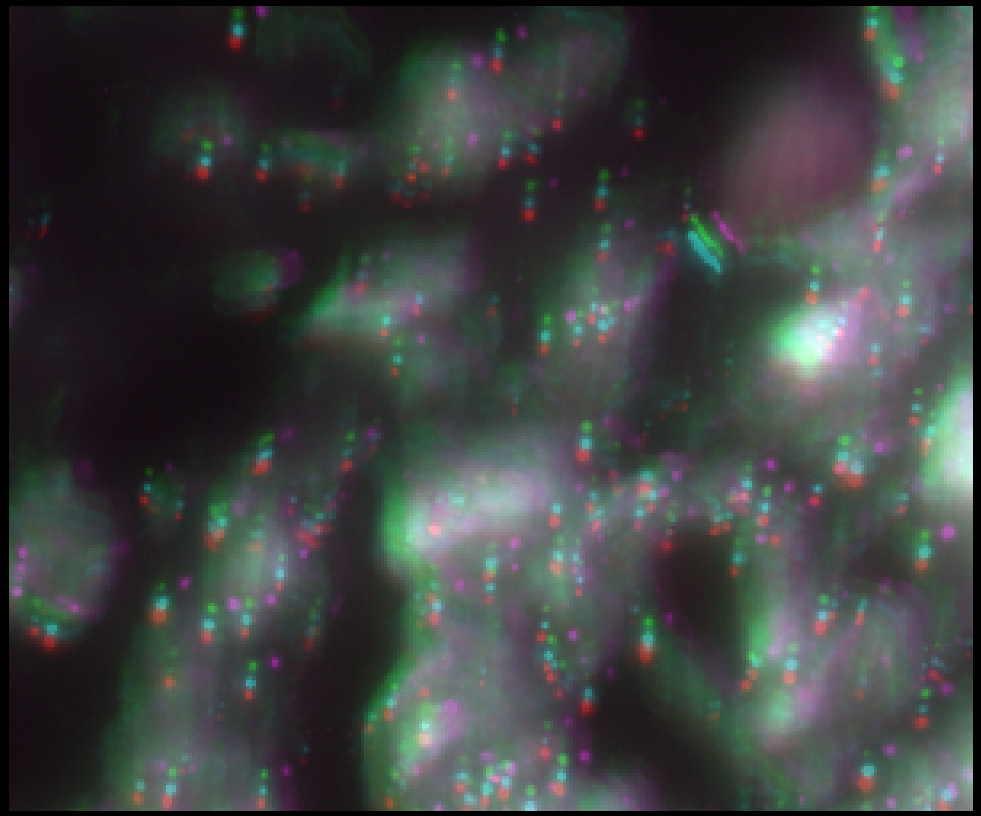
Per-pixel sequencing

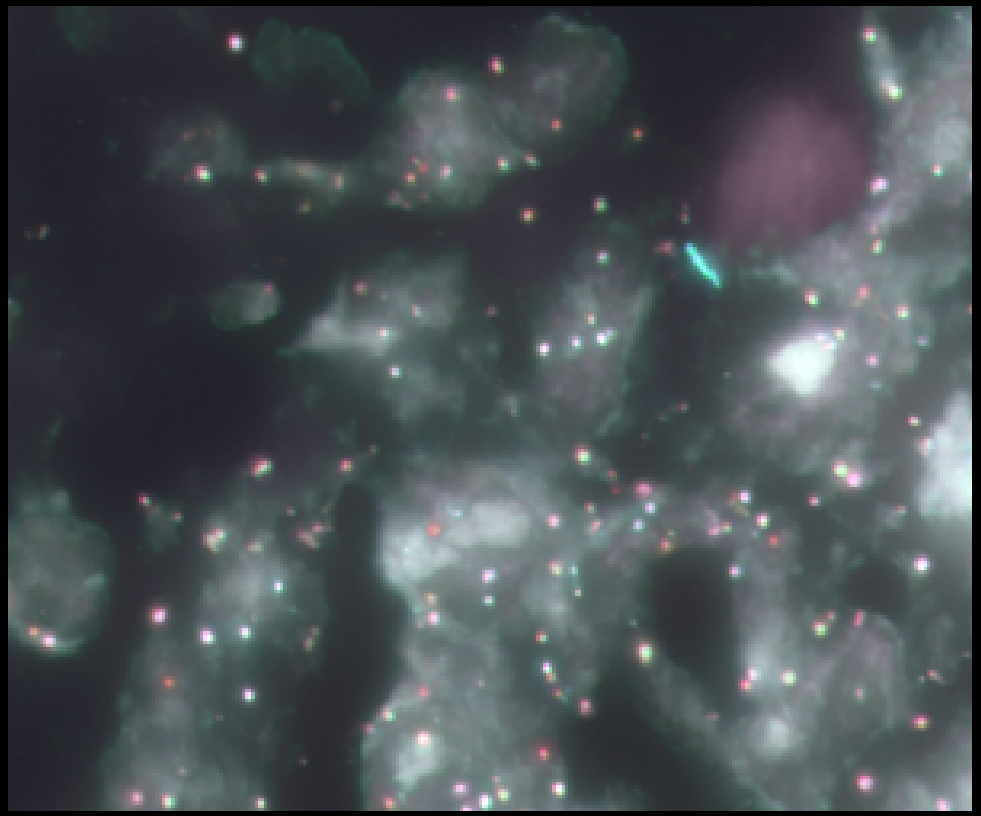
Kellen Dye

Four hybridization cycles

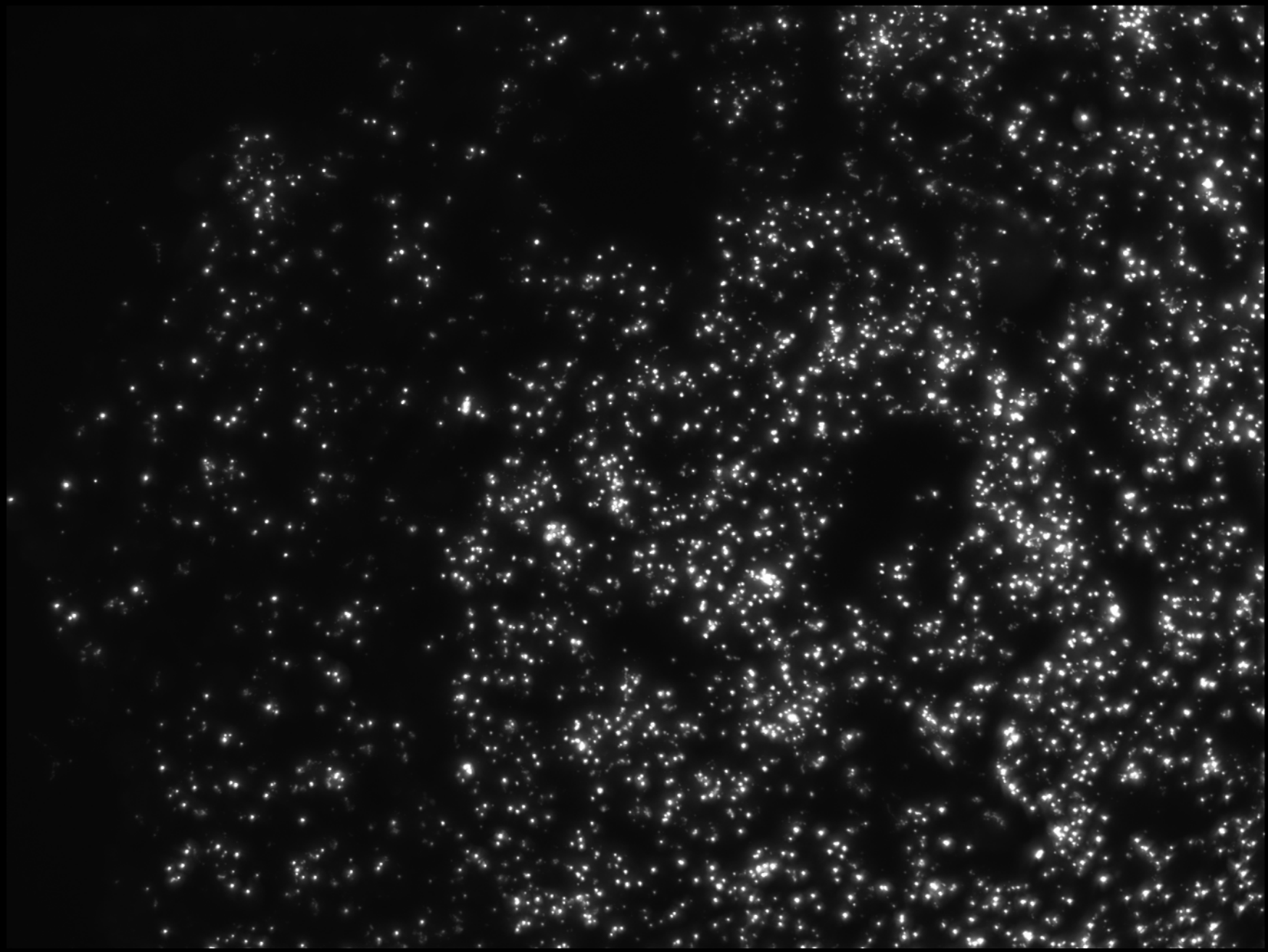


Registration: RANSAC with SURF

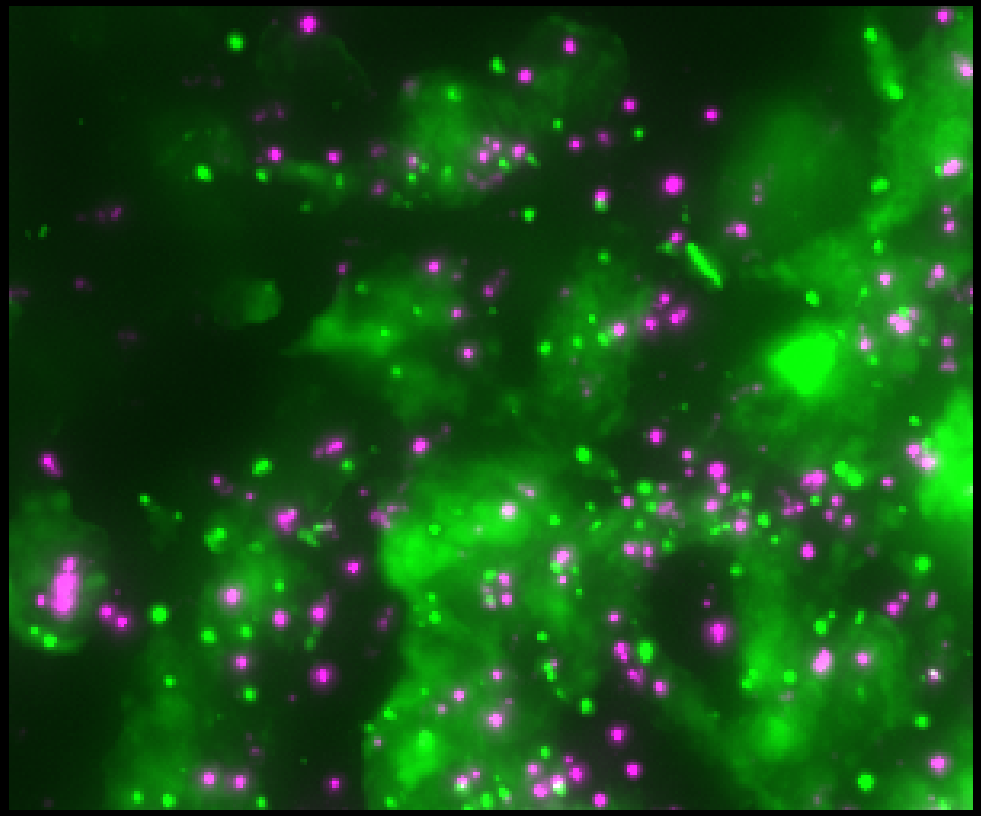


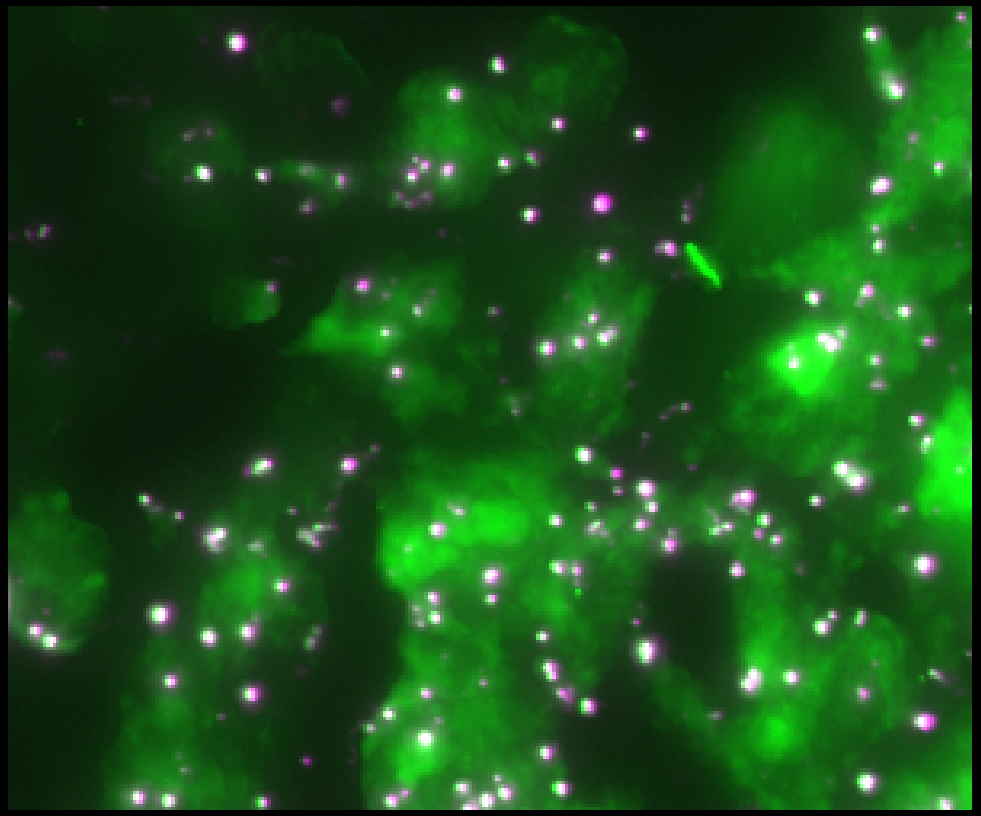


General stain

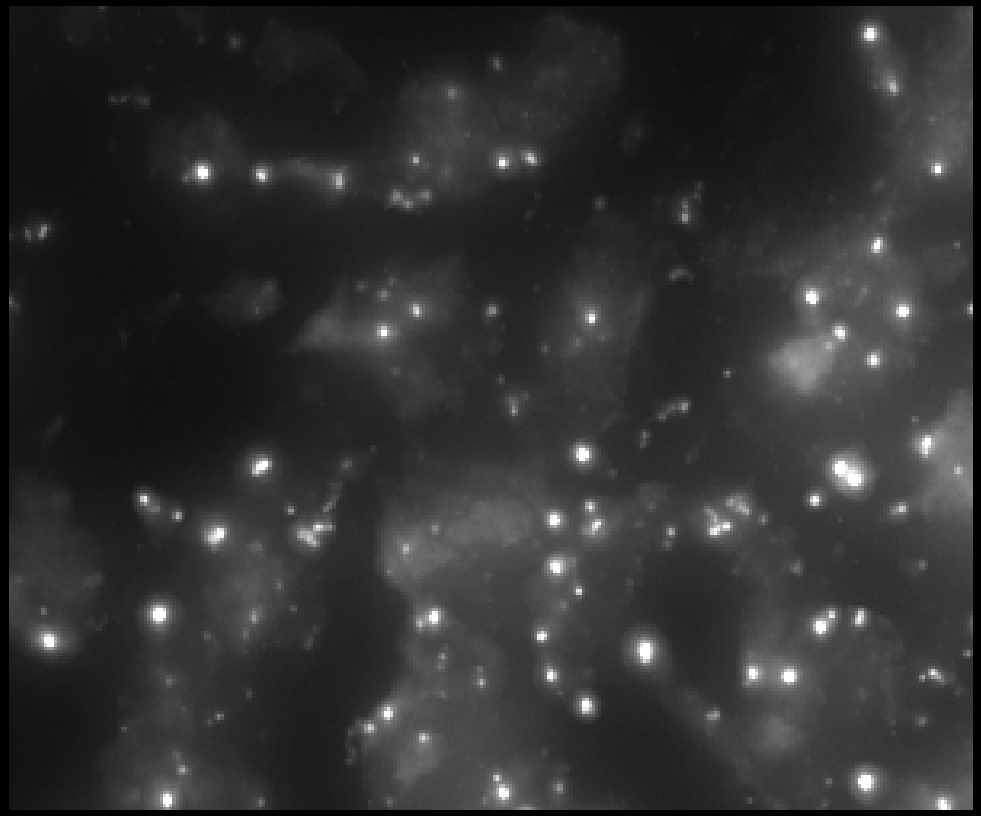


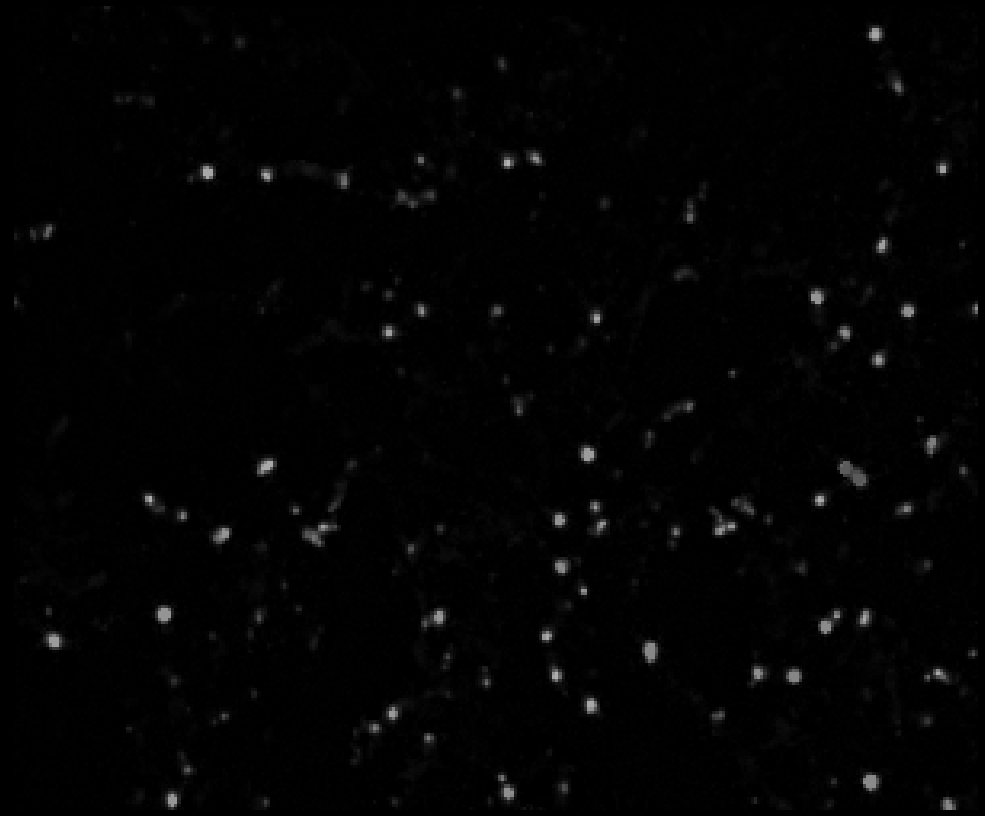
Registration: Coherent Point Drift





Preprocessing: top-hat filter

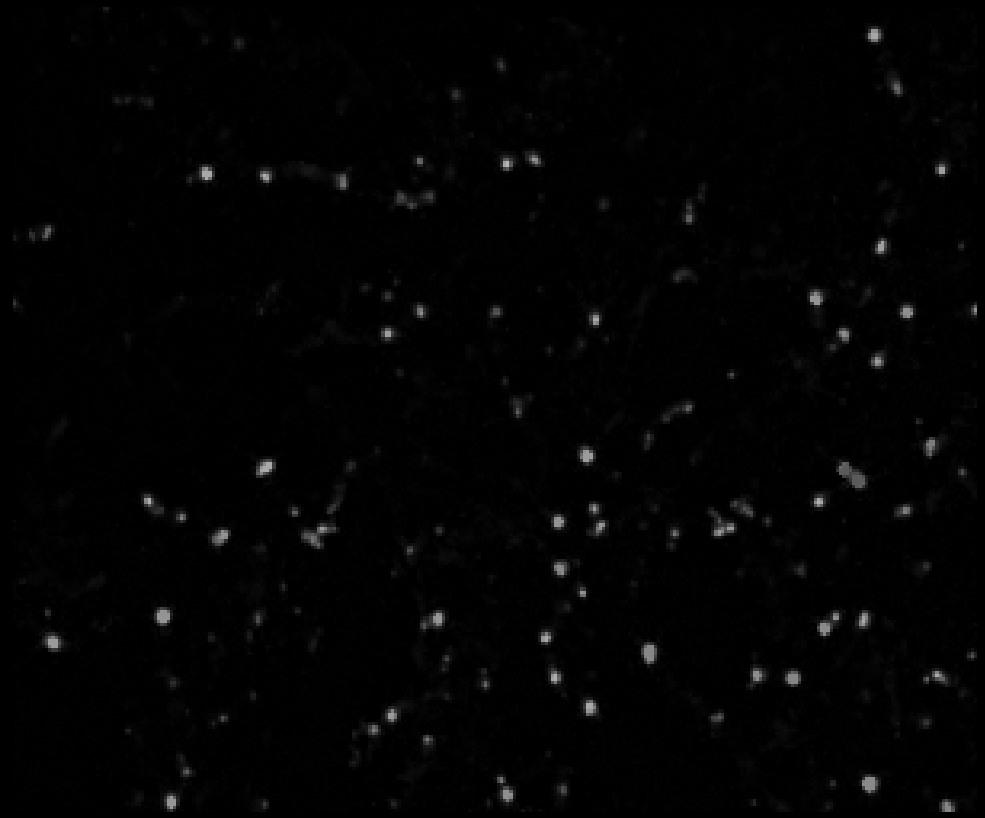


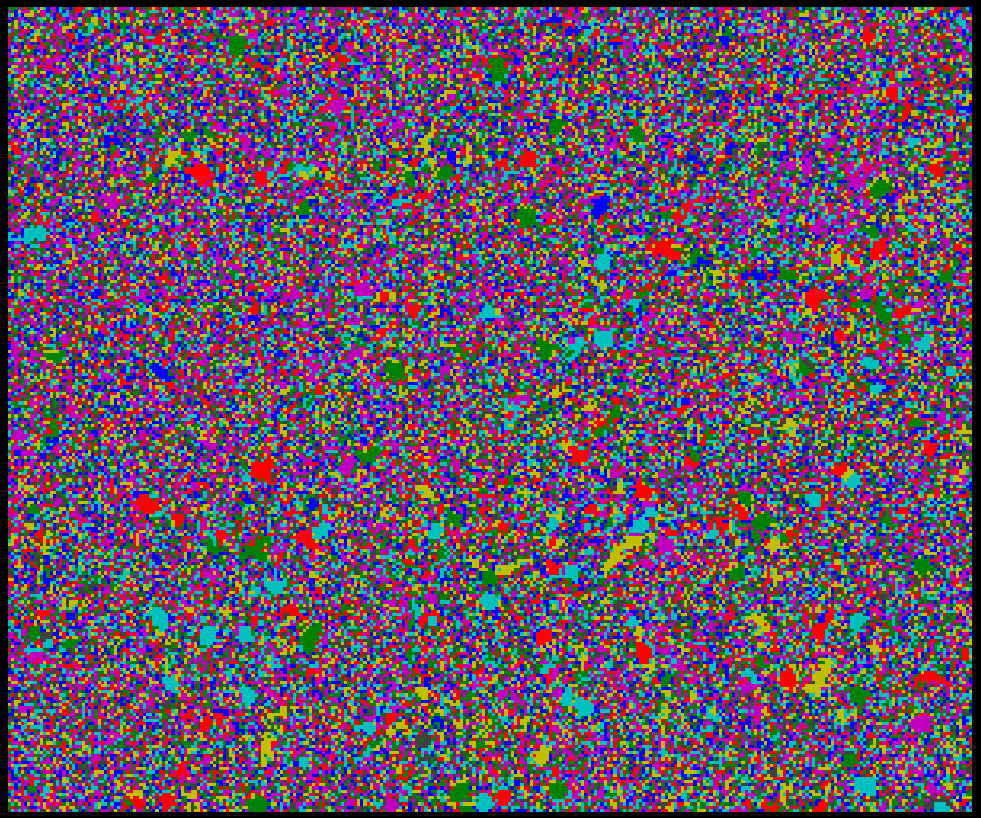


Per-pixel sequencing

$$seq(x,y) = \mathop{\Big\|}_{c \in \{1,2,3,4\}} \underset{b_c \in \{A,C,T,G\}}{argmax} \; I_{b_c}(x,y)$$

Per-pixel sequencing

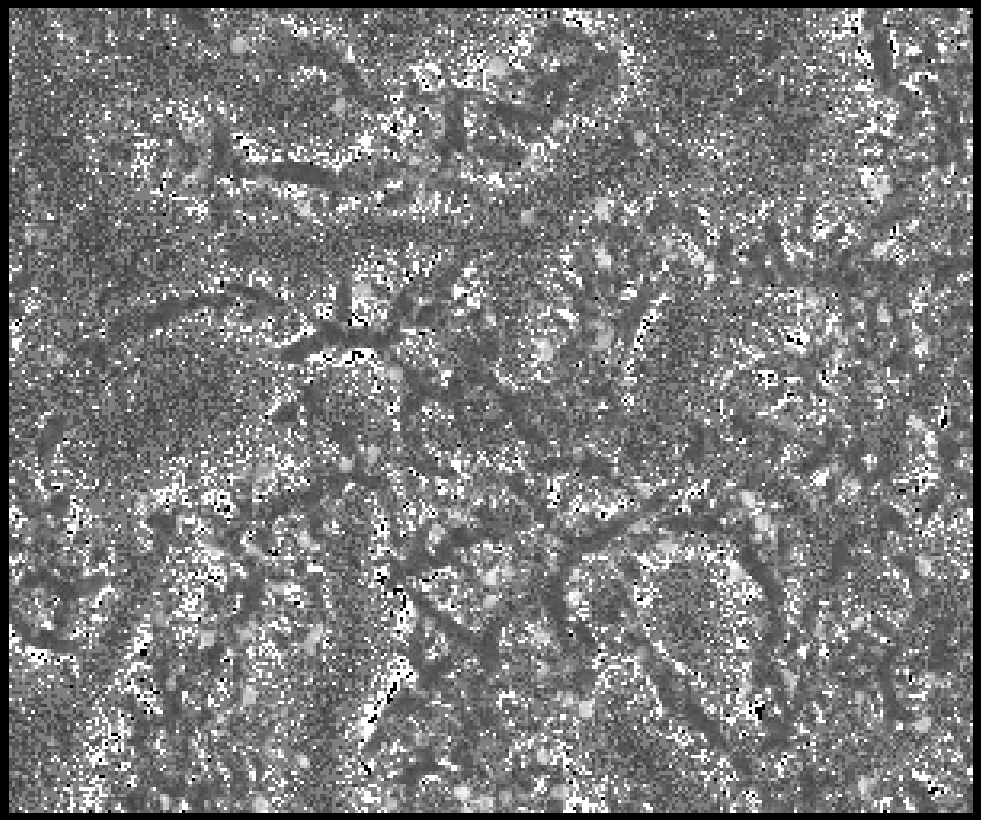


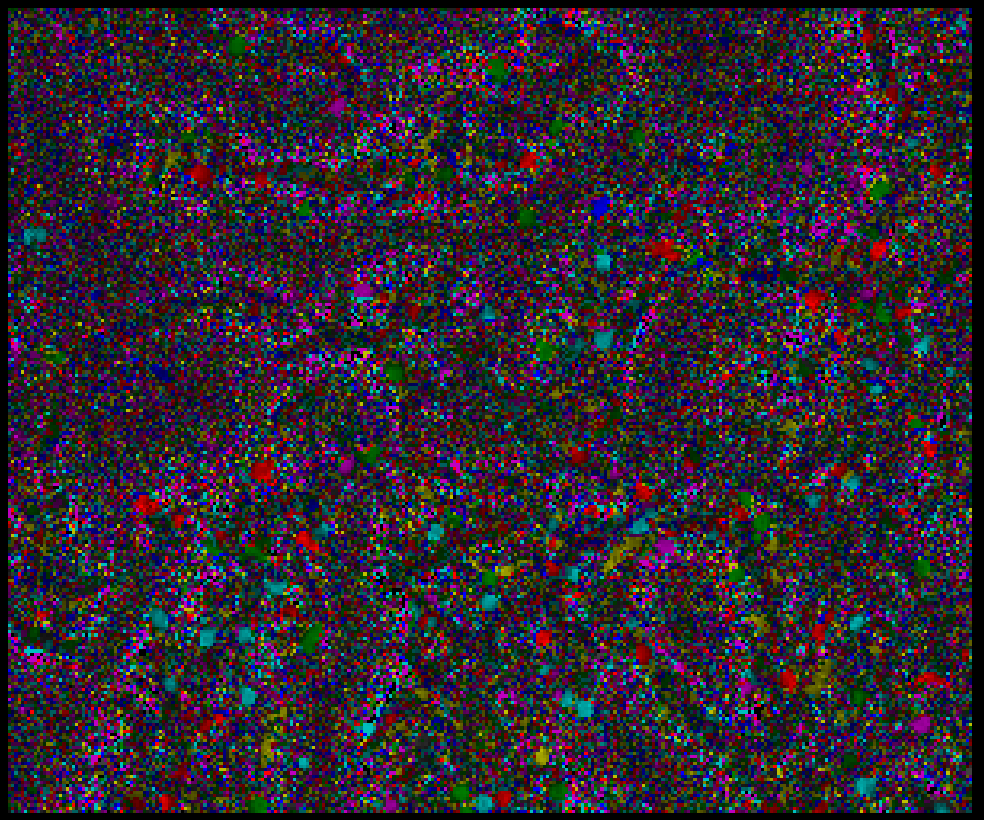


Quality for sequence

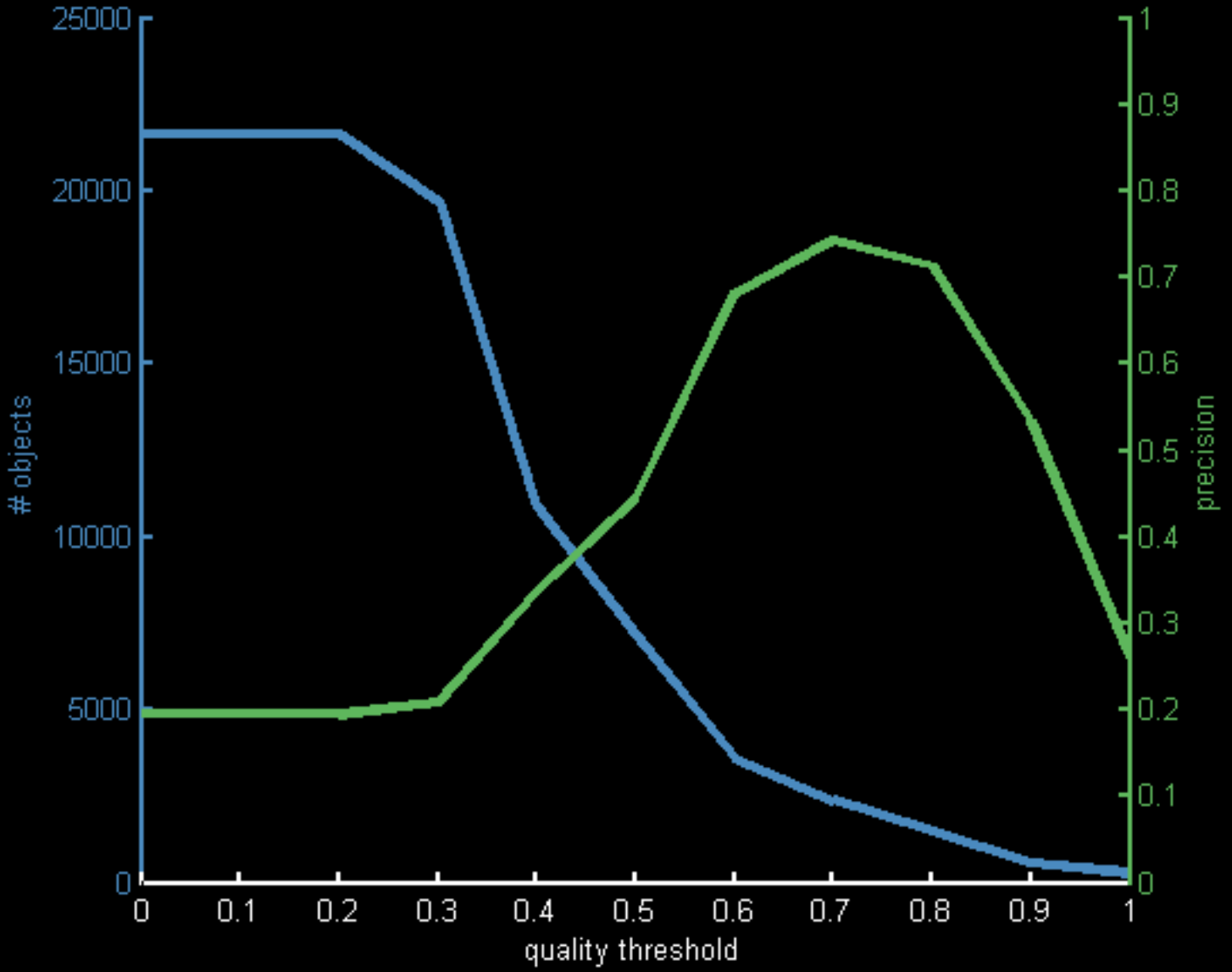
$$quality(x,y) = \min_{c \in \{1,2,3,4\}} \frac{\max_{b_c \in \{A,C,T,G\}} I_{b_c}(x,y)}{\sum_{b_c \in \{A,C,T,G\}} I_{b_c}(x,y)}$$

Quality for sequence





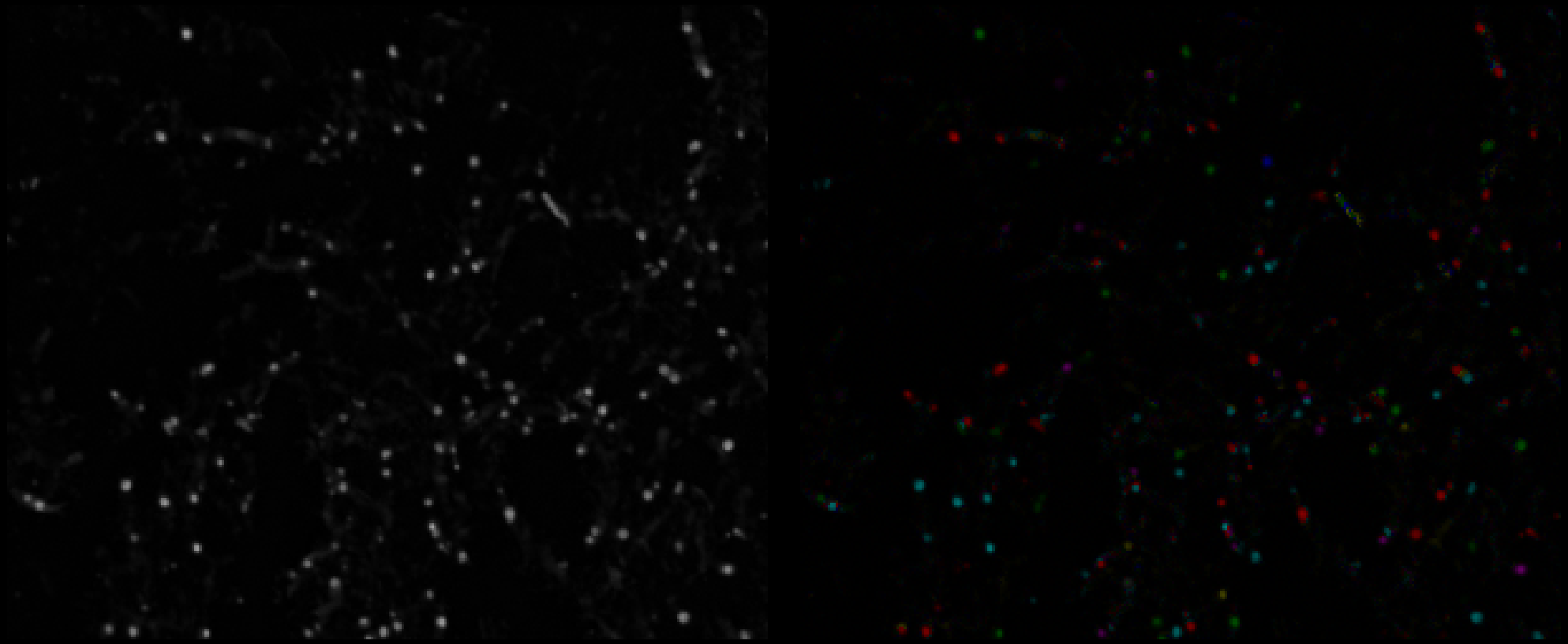
Quality for sequence



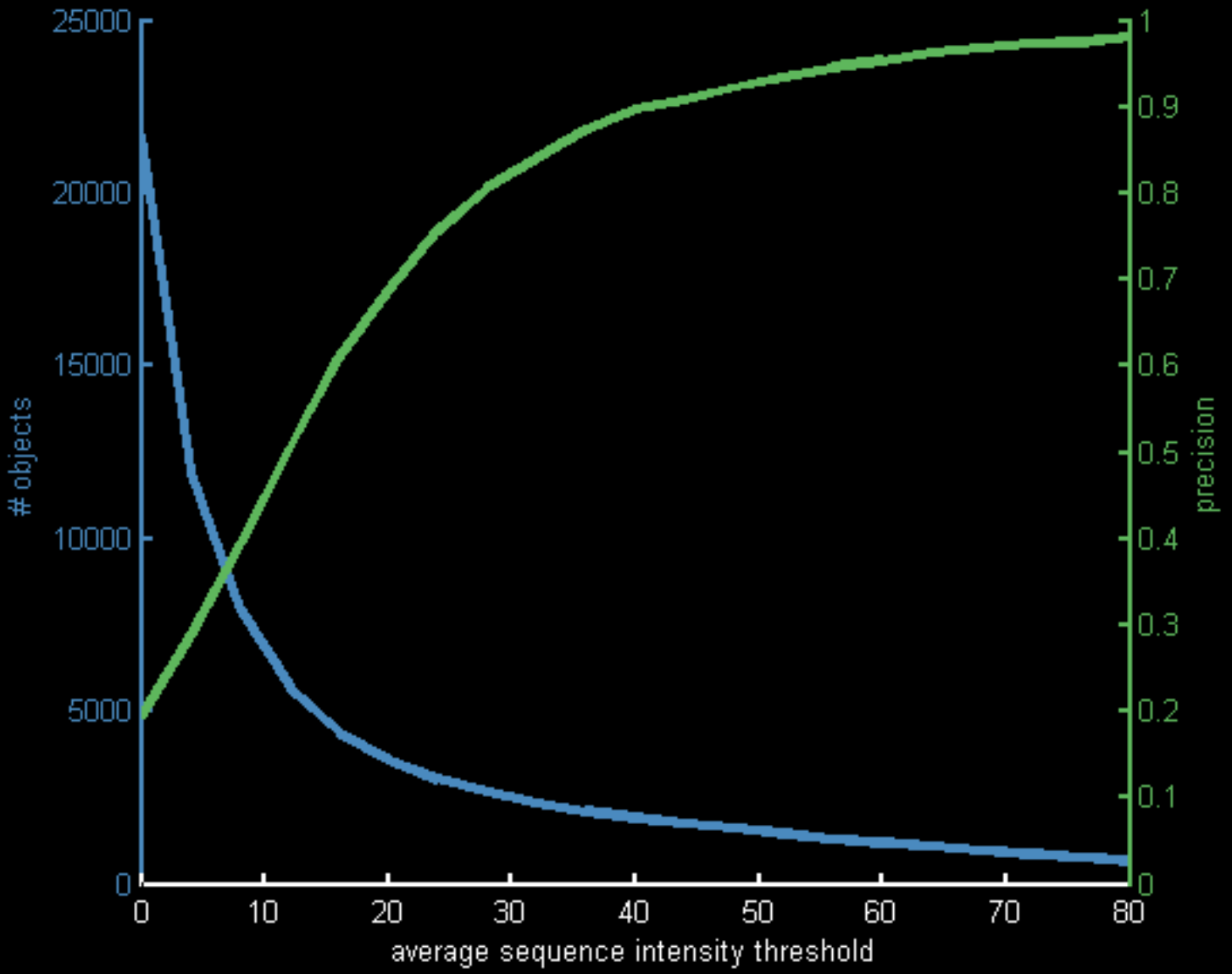
Average intensity for sequence

$$avg(x,y) = \frac{1}{|c|} \sum_{\substack{b_C \in \\ seq(x,y)}} I_{b_C}(x,y)$$

Average intensity for sequence



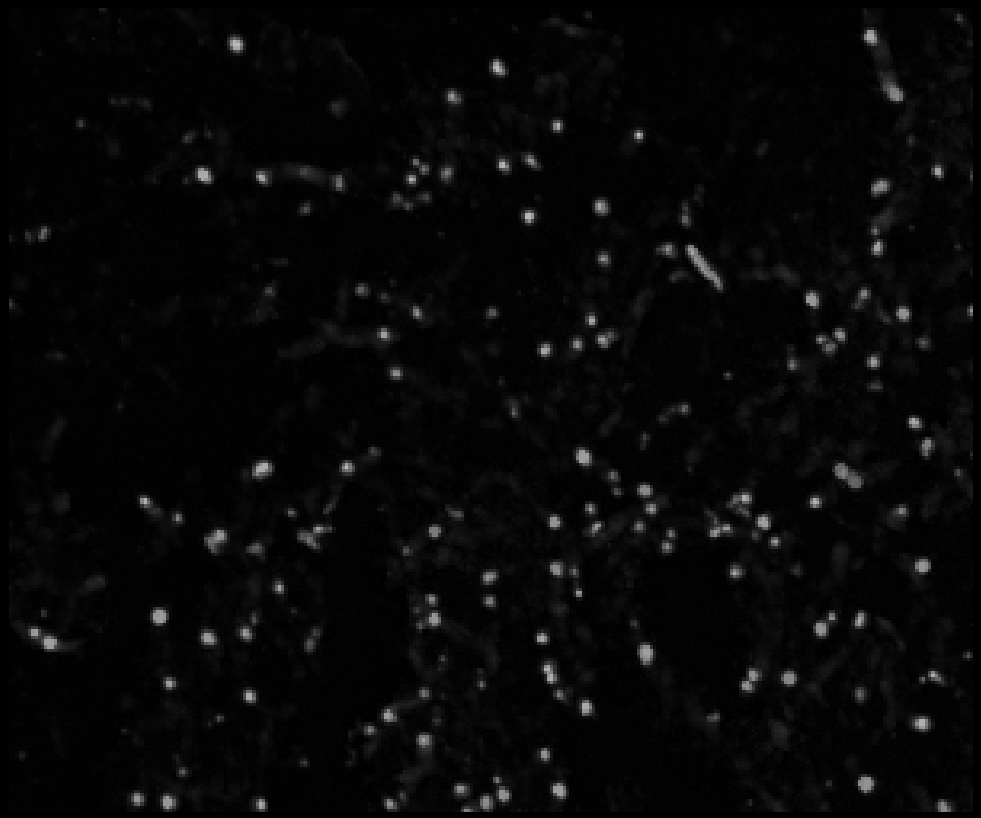
Average intensity for sequence

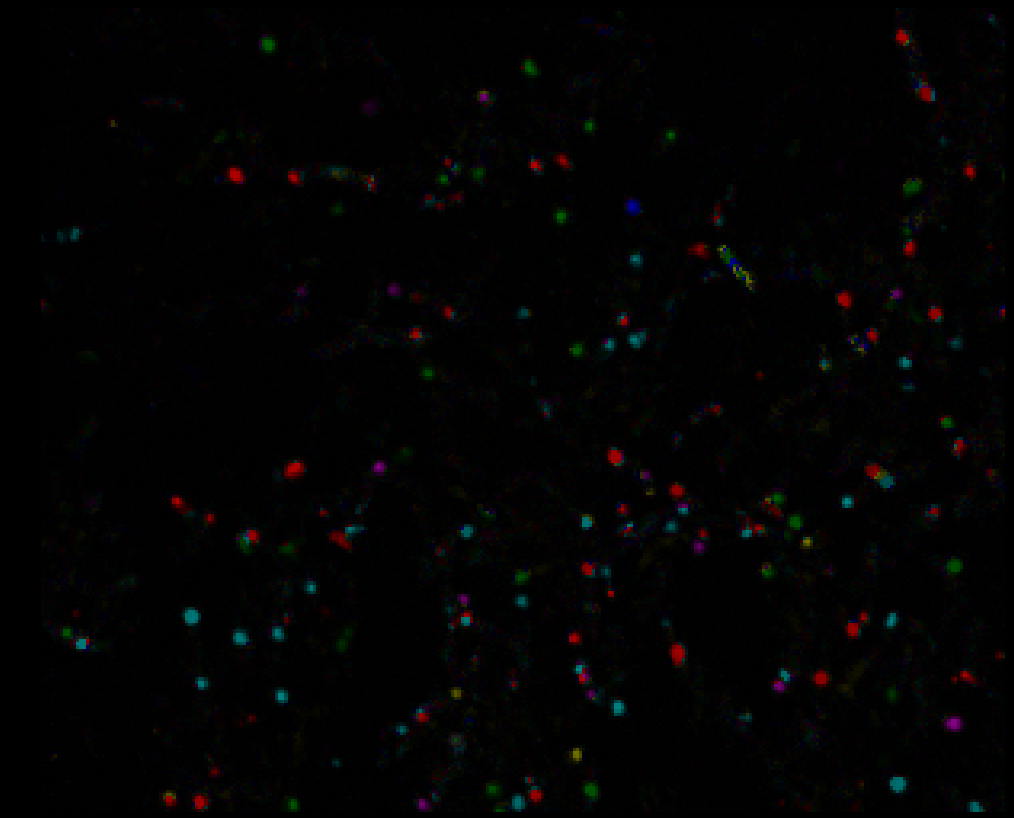


Maximum intensity for sequence

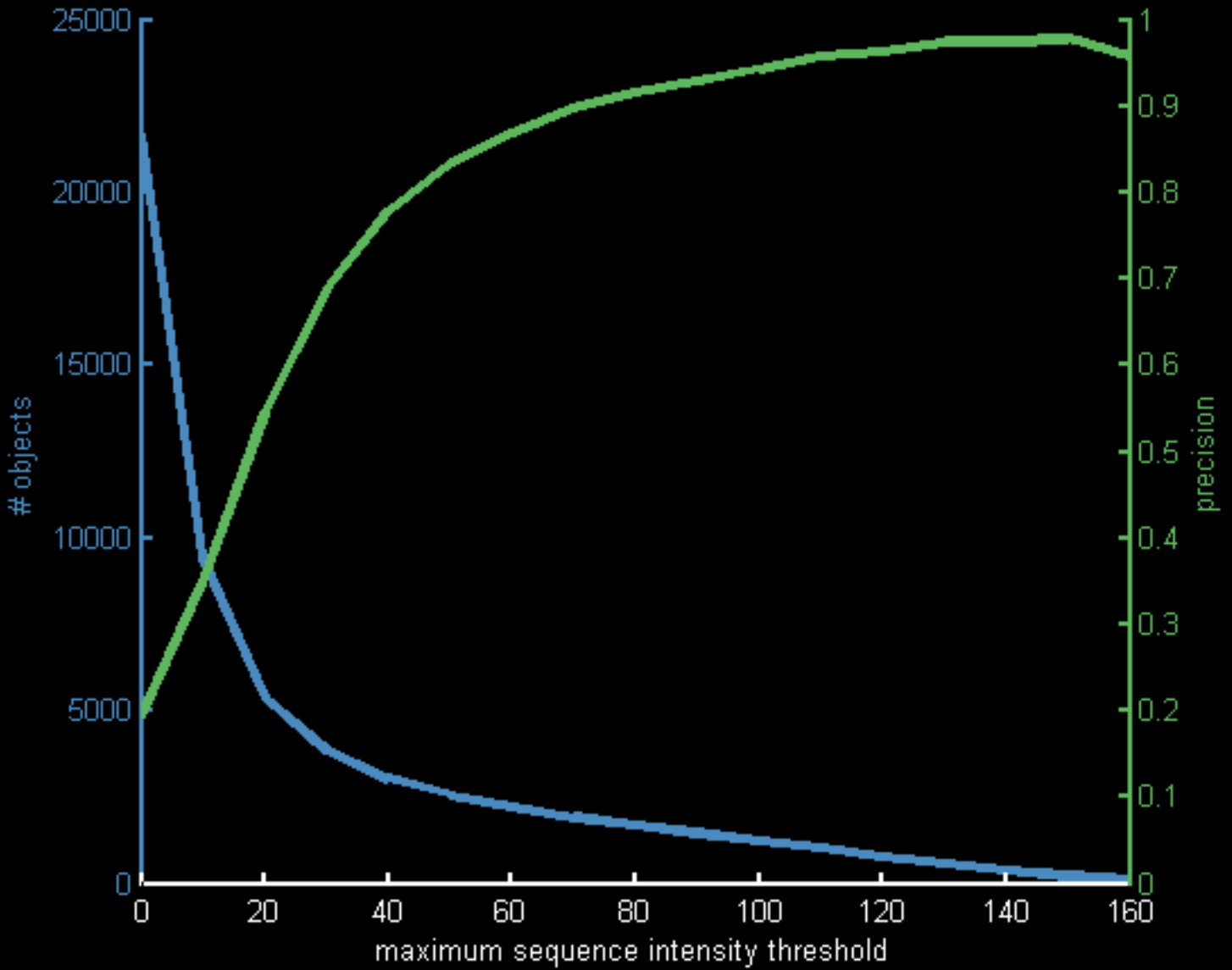
$$maxIntensity(x,y) = \max_{\substack{b_C \in \\ seq(x,y)}} I_{b_C}(x,y)$$

Maximum intensity for sequence

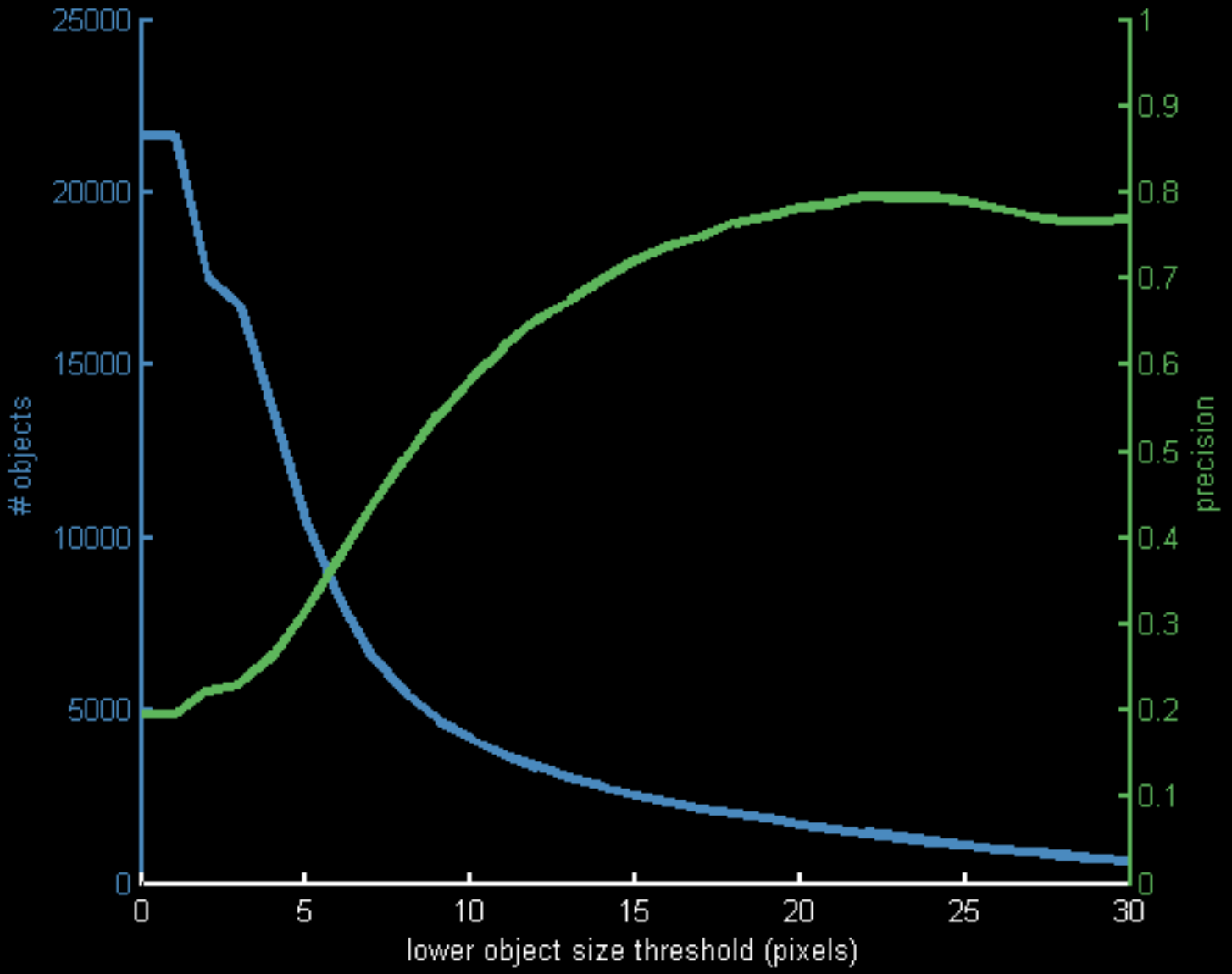




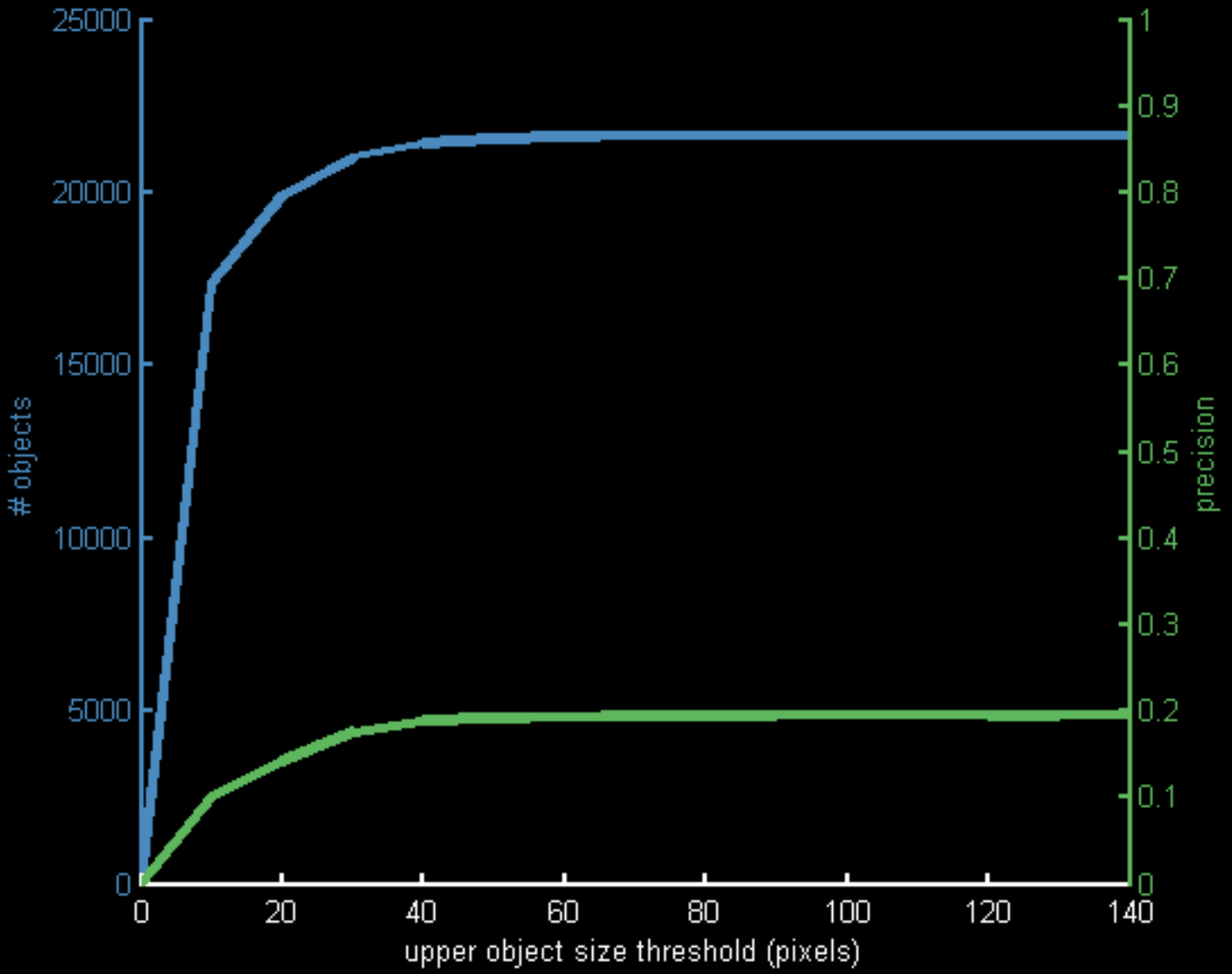
Maximum intensity for sequence



Region size: lower threshold



Region size: upper threshold



Exclude

A priori known error sequences: AAAA, CCCC, GGGG, TTTT

Regions outside an inclusive binary threshold of the general stain

$$avg(x, y) < 25$$

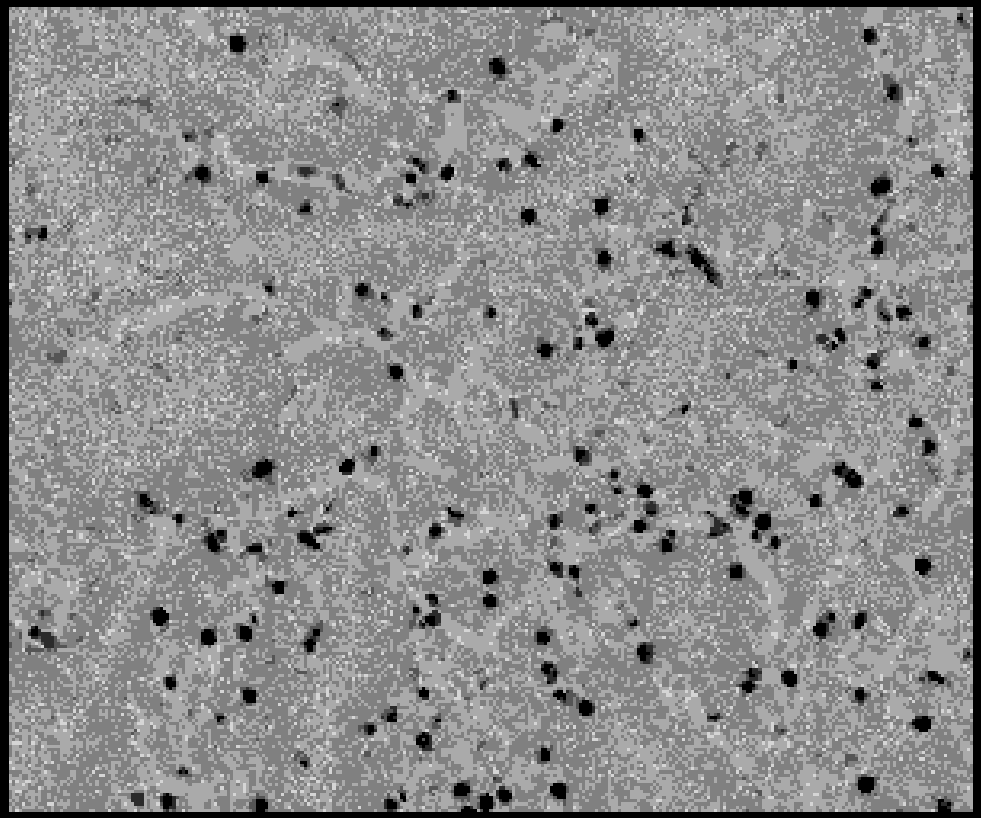
$$maxIntensity(x, y) < 40$$

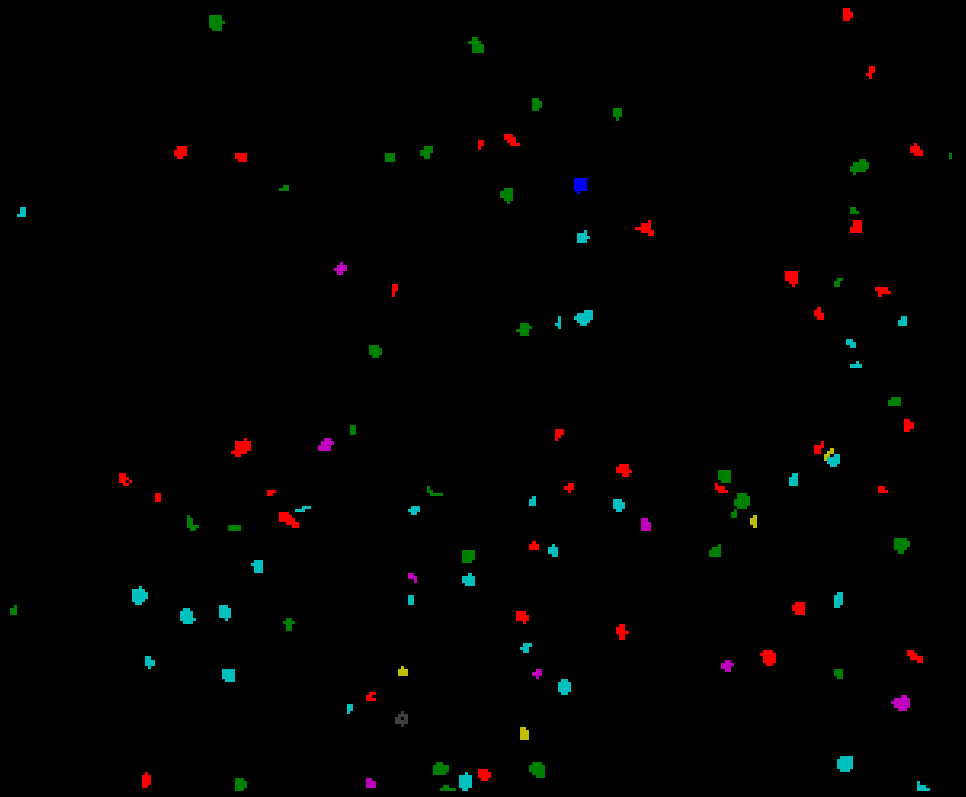
$$quality(x, y) < 0.475$$

$$size > 80$$

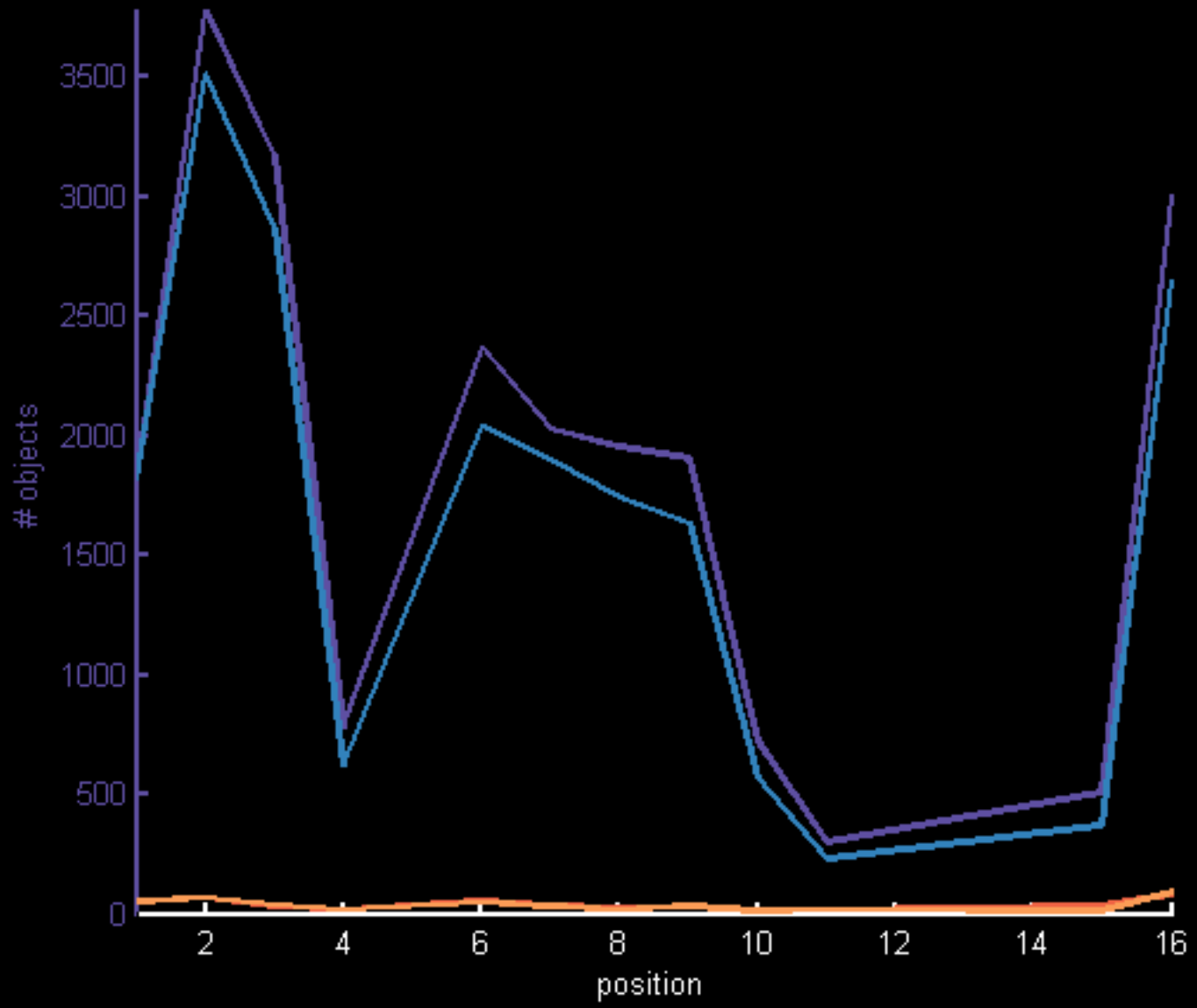
$$size < 5$$

Exclude

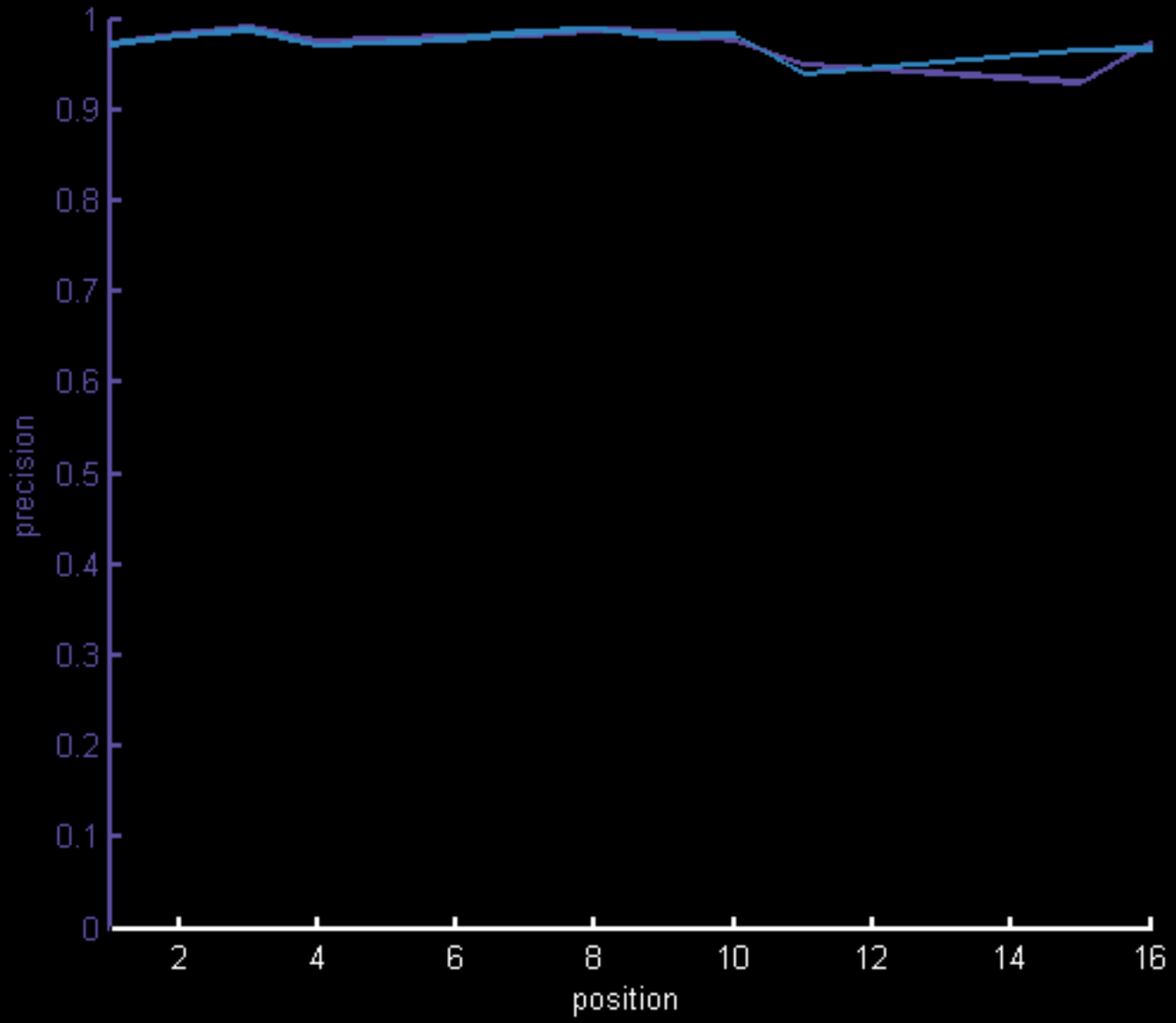


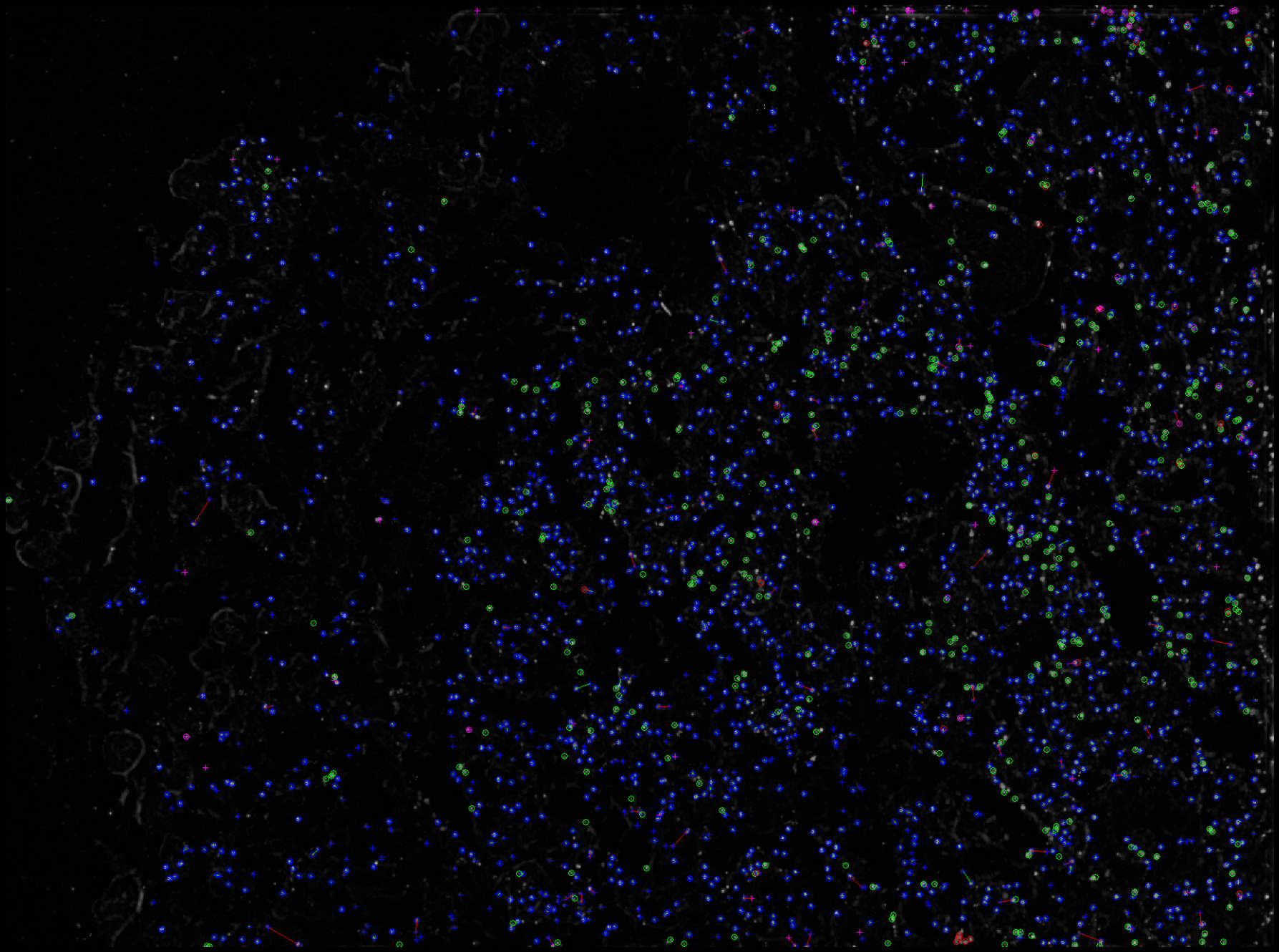


Comparison with existing blob-based approach



Comparison with existing blob-based approach





Potential speedup

Just use connected components

Not *that many* touching blobs of the same class

Potential improvement

Larger overlap for each imaging position

Use regional information for registration of the general stain

Potential improvement

Fully rigid registration

Both RANSAC and Coherent Point Drift are affine

Scaling and shear mostly introduce error

Thanks!