# A non-parametric procedure for co-localization studies in fluorescence microscopy

Frédéric Lavancier[1,2], Thierry Pécot[1] and Charles Kervrann[1]

[1] Team SERPICO; Inria Rennes - Bretagne Atlantique, F-35042 Rennes
[2] Laboratoire de Mathématiques Jean Leray, University of Nantes, F-44322 Nantes

# Outline

Motivation

Previous work

Our approach

Evaluation

Conclusion

# Outline

Motivation

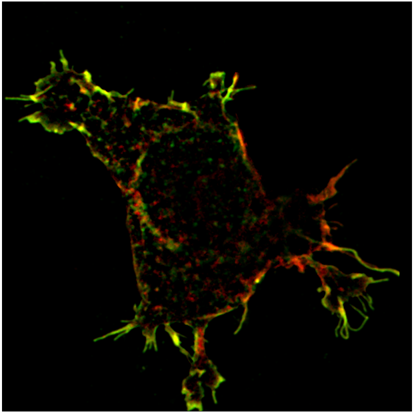
Previous work

Our approach

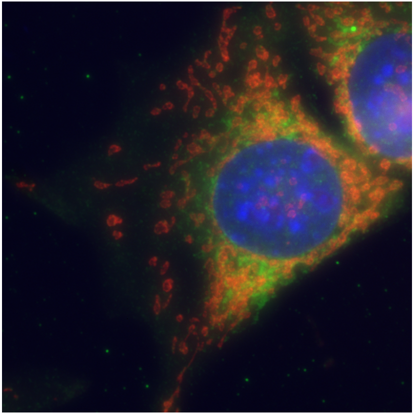
Evaluation

Conclusion

# Co-localization

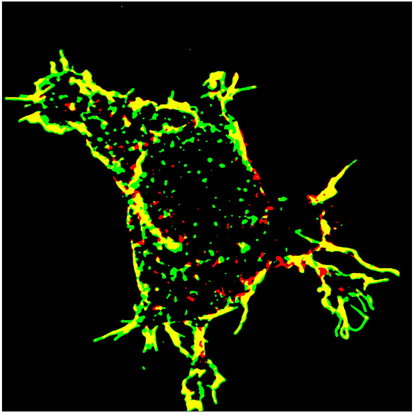


c-Src tyrosine kinase
serotonin receptor

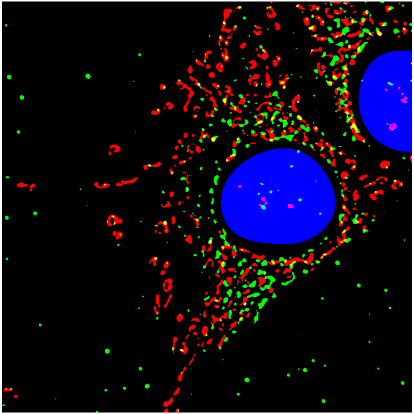


H1B mitochondria
EB1 (microtubules plus ends)
DNA

# Co-localization



c-Src tyrosine kinase
serotonin receptor



H1B mitochondria
EB1 (microtubules plus ends)
DNA

# Co-localization

# Outline

Motivation

Previous work

Our approach

Evaluation

Conclusion

# Common co-localization approaches

- $\Omega$: **image domain**
- $I_c(x)$: **intensity** observed at site $x \in \Omega$ in channel $c$
- $\overline{I}_c$: **mean intensity** for channel $c$

# Common co-localization approaches

- $\Omega$: **image domain**
- $I_c(x)$: **intensity** observed at site $x \in \Omega$ in channel $c$
- $\overline{I}_c$: **mean intensity** for channel $c$

- **Pearson's correlation coefficient:**

$$PCC = \frac{\sum_{x \in \Omega} (I_1(x) - \overline{I}_1)(I_2(x) - \overline{I}_2)}{\sqrt{\sum_{x \in \Omega} (I_1(x) - \overline{I}_1)^2 (I_2(x) - \overline{I}_2)^2}}$$

# Common co-localization approaches

- $\Omega$: **image domain**
- $I_c(x)$: **intensity** observed at site $x \in \Omega$ in channel $c$
- $\overline{I}_c$: **mean intensity** for channel $c$

- **Pearson's correlation coefficient:**

$$PCC = \frac{\sum_{x\in\Omega}(I_1(x)-\overline{I}_1)(I_2(x)-\overline{I}_2)}{\sqrt{\sum_{x\in\Omega}(I_1(x)-\overline{I}_1)^2(I_2(x)-\overline{I}_2)^2}}$$

- **Manders' co-localization coefficients:**

$$MCC_1 = \frac{\sum_{x\in\Omega} I_1(x)\mathbf{1}_{I_2(x)>0}}{\sum_{x\in\Omega} I_1(x)}, \quad MCC_2 = \frac{\sum_{x\in\Omega} I_2(x)\mathbf{1}_{I_1(x)>0}}{\sum_{x\in\Omega} I_2(x)}$$

# Costes method

**Permutation test** applied to the **Pearson's** correlation coefficient (PCC)

# Costes method

**Permutation test** applied to the **Pearson's** correlation coefficient (PCC)

- The **PCC** between the two channels is computed

# Costes method

**Permutation test** applied to the **Pearson's** correlation coefficient (PCC)

- The **PCC** between the two channels is computed
- The pixels of one of the two images are **randomly permuted** $n$ times
- For **each permutation**, the **PCC** is computed

# Costes method

**Permutation test** applied to the **Pearson's** correlation coefficient (PCC)

- The **PCC** between the two channels is computed
- The pixels of one of the two images are **randomly permuted** $n$ times
- For **each permutation**, the **PCC** is computed
- The **proportion of PCCs** obtained with permuted data that are higher than the PCC obtained for the two channels define the **p-value**

# Costes method

**Permutation test** applied to the **Pearson's** correlation coefficient (PCC)

- The **PCC** between the two channels is computed
- The pixels of one of the two images are **randomly permuted** $n$ times
- For **each permutation**, the **PCC** is computed
- The **proportion of PCCs** obtained with permuted data that are higher than the PCC obtained for the two channels define the **p-value**

- As neighbor pixels are correlated, **blocks** are preferred to pixels for **permutations**

# Lagache method

**Object-based** method in the framework of **marked point process**

# Lagache method

**Object-based** method in the framework of **marked point process**

- $A_1 = \{x_1, \ldots, x_{n_1}\}$: **mass centers** of segmented objects in channel 1
- $A_2 = \{y_1, \ldots, y_{n_2}\}$: **mass centers** of segmented objects in channel 2

# Lagache method

**Object-based** method in the framework of **marked point process**

- $A_1 = \{x_1, \ldots, x_{n_1}\}$: **mass centers** of segmented objects in channel 1
- $A_2 = \{y_1, \ldots, y_{n_2}\}$: **mass centers** of segmented objects in channel 2
- **Asymmetric** Ripley's K function

$$K_{12}(r) = \frac{|\Omega|}{n_1 n_2} \sum_{x \in A_1} \sum_{y \in A_2} \mathbf{1}_{d_{xy} < r} b(x, y, r)$$

where

| | |
|---|---|
| $d_{xy}$ | **Euclidean distance** between $x$ and $y$ |
| $b(x, y, r) = \dfrac{\lvert c(x, d_{xy}) \rvert}{\lvert c(x, d_{xy}) \rvert \cap \Omega}$ | **boundary correction** term |
| $c(x, d_{xy})$ | disk (sphere) centered at $x$ with radius $d_{xy}$ |

# Lagache method

- **Reduced** statistics

$$\tilde{K}_{12}(r) = \frac{K_{12}(r) - \mathbb{E}\{K_{12}(r)\}}{\sqrt{Var\{K_{12}(r)\}}}$$

# Lagache method

- **Reduced** statistics

$$\tilde{K}_{12}(r) = \frac{K_{12}(r) - \mathbb{E}\{K_{12}(r)\}}{\sqrt{Var\{K_{12}(r)\}}}$$

- **Asymptotic normality** of $\tilde{K}_{12}(r)$

$$\tilde{K}_{12}(r) \rightarrow \mathcal{N}(0,1) \text{ as } n_2 \rightarrow \infty$$

# Lagache method

- **Reduced** statistics

$$\tilde{K}_{12}(r) = \frac{K_{12}(r) - \mathbb{E}\{K_{12}(r)\}}{\sqrt{Var\{K_{12}(r)\}}}$$

- **Asymptotic normality** of $\tilde{K}_{12}(r)$

$$\tilde{K}_{12}(r) \to \mathcal{N}(0,1) \text{ as } n_2 \to \infty$$

- The **null hypothesis** of $A_2$ uniform distribution with a confidence level of $1-\gamma$ is rejected if:

$$\tilde{K}_{12}(r) > z_\gamma$$

# Outline

Motivation

Previous work

Our approach

Evaluation

Conclusion

- $\Gamma_1$: **random set** in $\mathbb{R}^d$
- $\Gamma_2$: **random set** in $\mathbb{R}^d$

$$p_1 = P(o \in \Gamma_1), \;\; p_2 = P(o \in \Gamma_2), \;\; p_{12} = P(o \in \Gamma_1 \cap \Gamma_2), \;\; \forall o \in \mathbb{R}^d$$

- $\Gamma_1$: **random set** in $\mathbb{R}^d$
- $\Gamma_2$: **random set** in $\mathbb{R}^d$

$$p_1 = P(o \in \Gamma_1), \;\; p_2 = P(o \in \Gamma_2), \;\; p_{12} = P(o \in \Gamma_1 \cap \Gamma_2), \;\; \forall o \in \mathbb{R}^d$$

- $\Omega$: **image domain**

$$\hat{p}_1 = |\Omega|^{-1} \sum_{x \in \Omega} \mathbf{1}_{\Gamma_1}(x), \;\; \hat{p}_2 = |\Omega|^{-1} \sum_{x \in \Omega} \mathbf{1}_{\Gamma_2}(x), \;\; \hat{p}_{12} = |\Omega|^{-1} \sum_{x \in \Omega} \mathbf{1}_{\Gamma_1}(x) \mathbf{1}_{\Gamma_2}(x)$$

- $\Gamma_1$: **random set** in $\mathbb{R}^d$
- $\Gamma_2$: **random set** in $\mathbb{R}^d$

$$p_1 = P(o \in \Gamma_1), \;\; p_2 = P(o \in \Gamma_2), \;\; p_{12} = P(o \in \Gamma_1 \cap \Gamma_2), \;\; \forall o \in \mathbb{R}^d$$

- $\Omega$: **image domain**

$$\hat{p}_1 = |\Omega|^{-1} \sum_{x \in \Omega} \mathbf{1}_{\Gamma_1}(x), \;\; \hat{p}_2 = |\Omega|^{-1} \sum_{x \in \Omega} \mathbf{1}_{\Gamma_2}(x), \;\; \hat{p}_{12} = |\Omega|^{-1} \sum_{x \in \Omega} \mathbf{1}_{\Gamma_1}(x) \mathbf{1}_{\Gamma_2}(x)$$

- **Pearson correlation**:

$$\hat{\rho} = \frac{\hat{p}_{12} - \hat{p}_1 \hat{p}_2}{\sqrt{\hat{p}_1(1-\hat{p}_1)\hat{p}_2(1-\hat{p}_2)}}$$

- $\Gamma_1$: **random set** in $\mathbb{R}^d$
- $\Gamma_2$: **random set** in $\mathbb{R}^d$

$$p_1 = P(o \in \Gamma_1), \;\; p_2 = P(o \in \Gamma_2), \;\; p_{12} = P(o \in \Gamma_1 \cap \Gamma_2), \;\; \forall o \in \mathbb{R}^d$$

- $\Omega$: **image domain**

$$\hat{p}_1 = |\Omega|^{-1} \sum_{x \in \Omega} \mathbf{1}_{\Gamma_1}(x), \;\; \hat{p}_2 = |\Omega|^{-1} \sum_{x \in \Omega} \mathbf{1}_{\Gamma_2}(x), \;\; \hat{p}_{12} = |\Omega|^{-1} \sum_{x \in \Omega} \mathbf{1}_{\Gamma_1}(x) \mathbf{1}_{\Gamma_2}(x)$$

- **Pearson correlation**:

$$\hat{\rho} = \frac{\hat{p}_{12} - \hat{p}_1 \hat{p}_2}{\sqrt{\hat{p}_1 (1 - \hat{p}_1) \hat{p}_2 (1 - \hat{p}_2)}}$$

- Let $D = \hat{p}_{12} - \hat{p}_1 \hat{p}_2$

$$
\begin{aligned}
E(D) &= 0 \\
Var(D) &= |\Omega|^{-2} E\left( \sum_{x\in\Omega} \mathbf{1}_{\Gamma_1}(x)\mathbf{1}_{\Gamma_2}(x) - |\Omega|^{-1} \sum_{x\in\Omega} \mathbf{1}_{\Gamma_1}(x) \sum_{x\in\Omega} \mathbf{1}_{\Gamma_2}(x) \right)^2 \\
&= S1 + S2 + S3
\end{aligned}
$$

where

$$
\begin{aligned}
S_1 &= |\Omega|^{-2} \sum_{x\in\Omega} \sum_{y\in\Omega} C_1(x-y) C_2(x-y) \\
S_2 &= -2|\Omega|^{-3} \sum_{x\in\Omega} \left( \sum_{y\in\Omega} C_1(x-y) \right) \left( \sum_{y\in\Omega} C_2(x-y) \right) \\
S_3 &= |\Omega|^{-4} \sum_{x\in\Omega} \sum_{y\in\Omega} C_1(x-y) \sum_{x\in\Omega} \sum_{y\in\Omega} C_2(x-y)
\end{aligned}
$$

and $C_1$ and $C_2$ are **auto-covariance** functions

$$
\begin{aligned}
E(D) &= 0 \\
Var(D) &= |\Omega|^{-2} E\left(\sum_{x\in\Omega} \mathbf{1}_{\Gamma_1}(x)\mathbf{1}_{\Gamma_2}(x) - |\Omega|^{-1}\sum_{x\in\Omega}\mathbf{1}_{\Gamma_1}(x)\sum_{x\in\Omega}\mathbf{1}_{\Gamma_2}(x)\right)^2 \\
&= S1 + S2 + S3
\end{aligned}
$$

where

$$
\begin{aligned}
S_1 &= |\Omega|^{-2}\sum_{x\in\Omega}\sum_{y\in\Omega} C_1(x-y)C_2(x-y) \\
S_2 &= -2|\Omega|^{-3}\sum_{x\in\Omega}\left(\sum_{y\in\Omega} C_1(x-y)\right)\left(\sum_{y\in\Omega} C_2(x-y)\right) \\
S_3 &= |\Omega|^{-4}\sum_{x\in\Omega}\sum_{y\in\Omega} C_1(x-y)\sum_{x\in\Omega}\sum_{y\in\Omega} C_2(x-y)
\end{aligned}
$$

and $C_1$ and $C_2$ are **auto-covariance** functions

$$Var(D) \sim S_1 \text{ if } |\Omega| \to \infty$$

# Hypothesis testing

- Following [1], it can be proved that:

$$\frac{D}{\sqrt{S_1}} \to \mathcal{N}(0,1) \text{ as } |\Omega| \to \infty$$

[1] Shigeru Mase. Asymptotic properties of stereological estimators of volume fraction for stationary random sets. *Journal of Applied Probability*

# Hypothesis testing

- Following [1], it can be proved that:

$$\frac{D}{\sqrt{S_1}} \to \mathcal{N}(0,1) \text{ as } |\Omega| \to \infty$$

- If $\Gamma_1$ and $\Gamma_2$ are $m$-**dependent stationary random sets** and if $\Gamma_1$ is **independent** of $\Gamma_2$, then

$$T := \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)\hat{p}_2(1-\hat{p}_2)}{\hat{S}_1}}\hat{\rho} \to \mathcal{N}(0,1) \text{ as } |\Omega| \to \infty$$

[1] Shigeru Mase. Asymptotic properties of stereological estimators of volume fraction for stationary random sets. *Journal of Applied Probability*

# Hypothesis testing

- Following [1], it can be proved that:

$$\frac{D}{\sqrt{S_1}} \to \mathcal{N}(0,1) \text{ as } |\Omega| \to \infty$$

- If $\Gamma_1$ and $\Gamma_2$ are $m$-**dependent stationary random sets** and if $\Gamma_1$ is **independent** of $\Gamma_2$, then

$$T := \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)\hat{p}_2(1-\hat{p}_2)}{\hat{S}_1}}\hat{\rho} \to \mathcal{N}(0,1) \text{ as } |\Omega| \to \infty$$

- The **null hypothesis** of independence is **rejected** at the asymptotic level $\alpha \in (0,1)$ if $T > q(\alpha)$ corresponding to

$$p\text{-value} \quad = \quad 1 - \Phi(T)$$

[1] Shigeru Mase. Asymptotic properties of stereological estimators of volume fraction for stationary random sets. *Journal of Applied Probability*

# Outline


Motivation

Previous work

Our approach

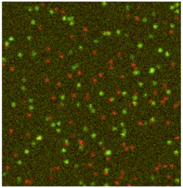
Evaluation

Conclusion

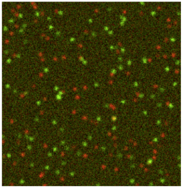
# Computation time

| | 2D image<br>256x256<br>50 objects | 2D image<br>256x256<br>200 objects | 2D image<br>256x256<br>3500 objects | 2D+t image<br>256x256x1000<br>100 objects | 3D image<br>256x256x60<br>100 objects |
|---|---|---|---|---|---|
| **Costes method**<br>**with 1000 permutations**<br>ImageJ plugin | 6.1 s | 6.2 s | 6.1 s | 38 min 20 s | 3 min 3 s |
| **Lagache method**<br>Icy plugin | 1 s | 1.96 s | 12.38 s | 12 min 39 s | 4.37 s |
| **Our method**<br>C++ implementation | 0.18 s | 0.2 s | 0.19 s | 29.5 s | 10 s |

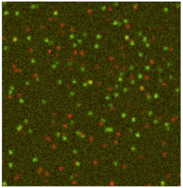
0% colocalization

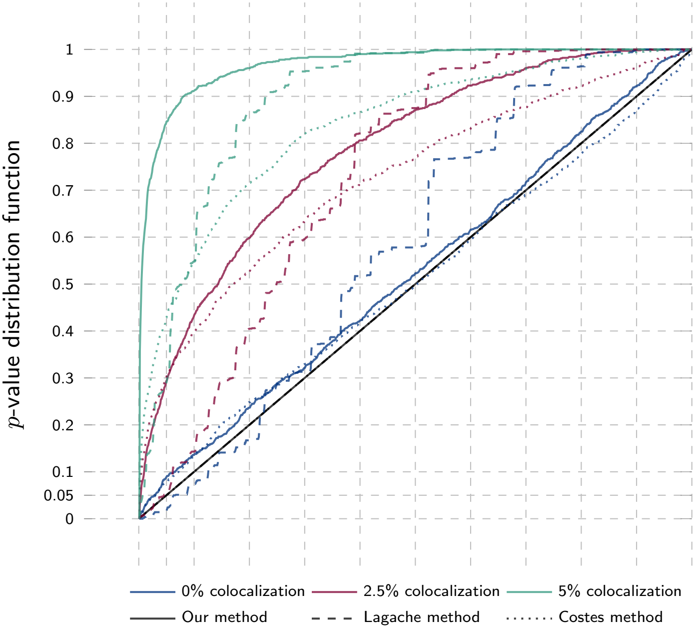


2.5% colocalization

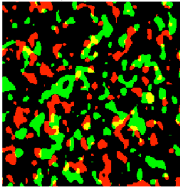


5% colocalization

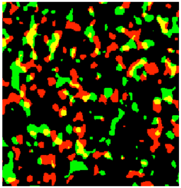




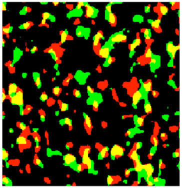
$\rho = 0$

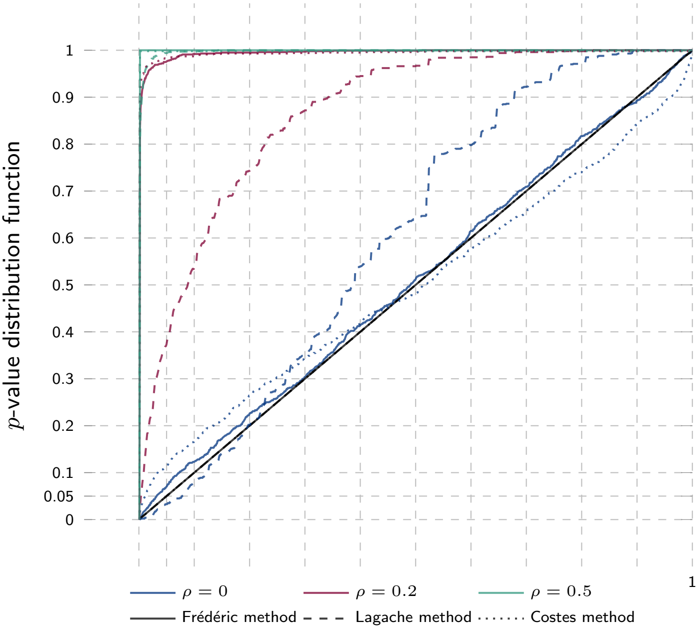


$\rho = 0.1$

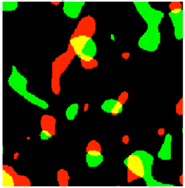


$\rho = 0.2$

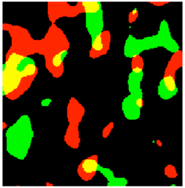




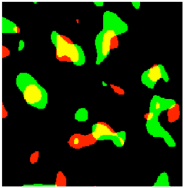
$\rho = 0$

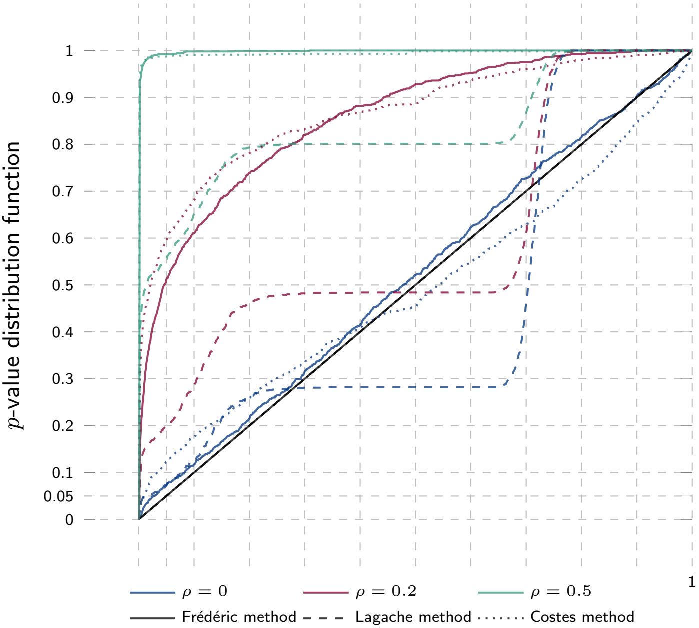


$\rho = 0.1$



$\rho = 0.2$

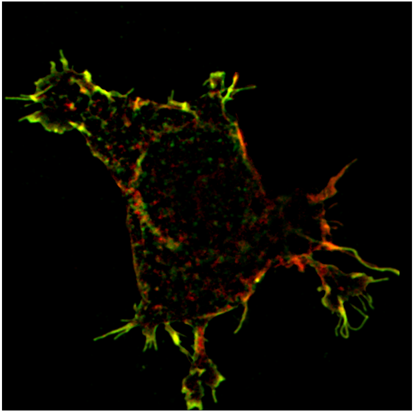




# Other evaluations on synthetic images

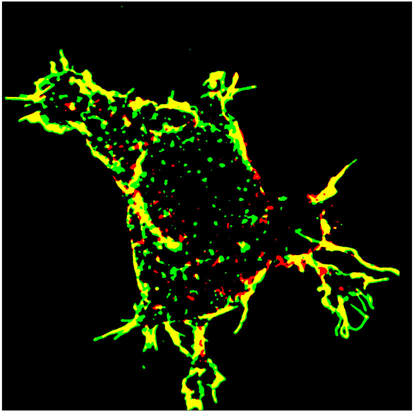
- 3D images
- Different number of objects in each channel
- Different size for objects in channel 1 and objects in channel 2

# Co-localization



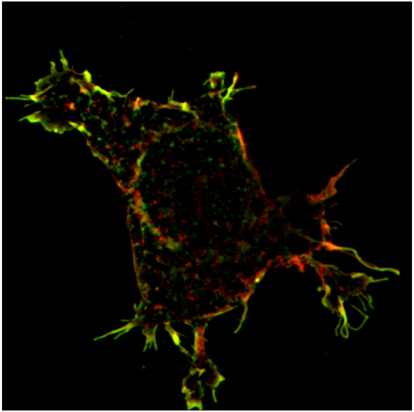
c-Src tyrosine kinase

serotonin receptor

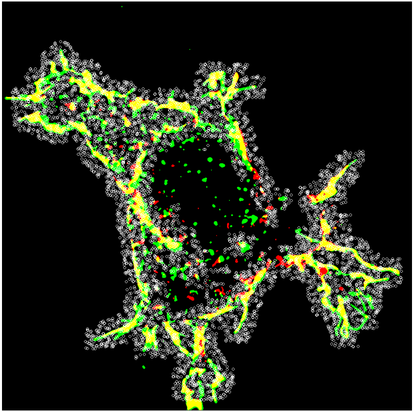


$p$-value $= 0$

# Co-localization

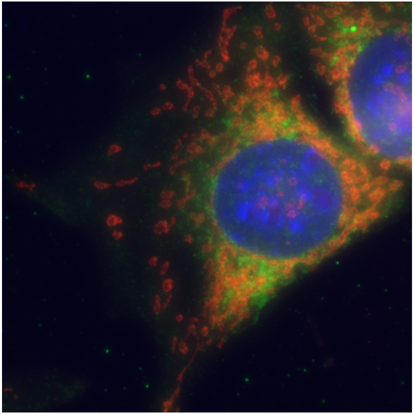


c-Src tyrosine kinase
serotonin receptor

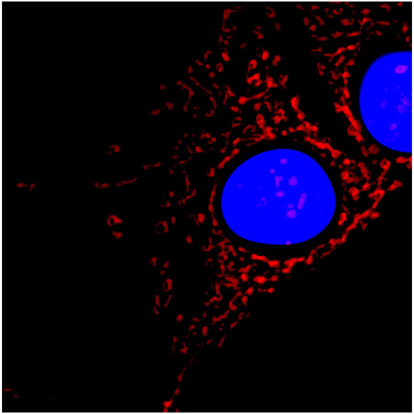


White circles: co-localization hits

# Co-localization

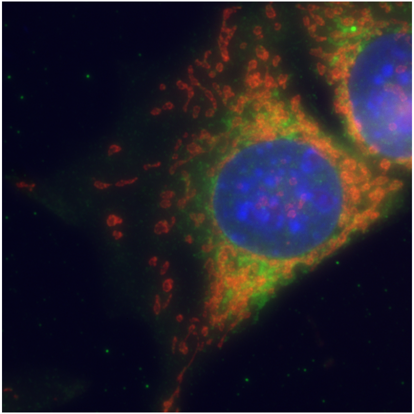


H1B mitochondria
EB1 (microtubules plus ends)
DNA

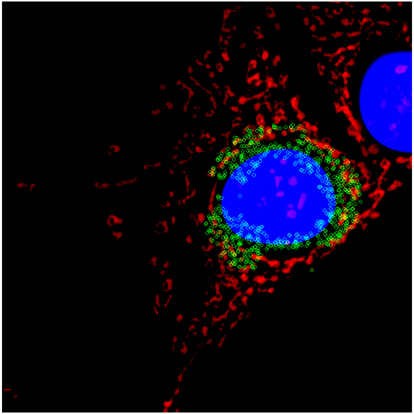


$p$-value $= 0$

# Co-localization

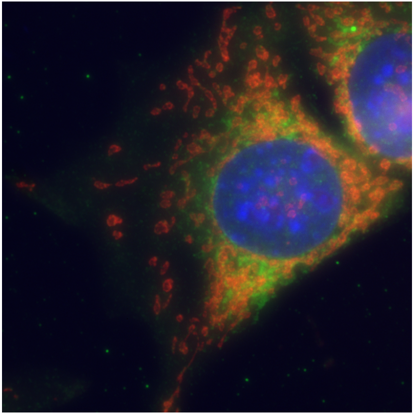


H1B mitochondria
EB1 (microtubules plus ends)
DNA

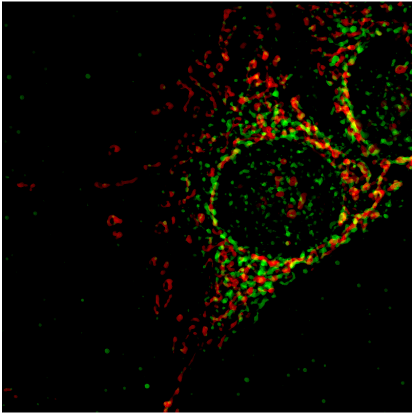


Green circles: anti co-localization hits

# Co-localization

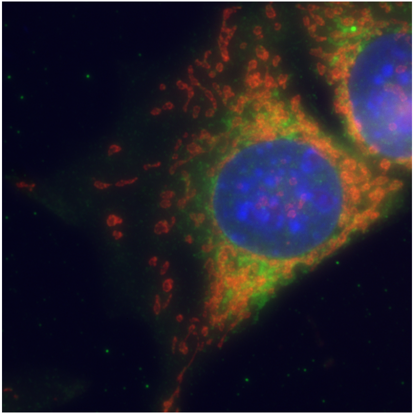


H1B mitochondria
EB1 (microtubules plus ends)
DNA

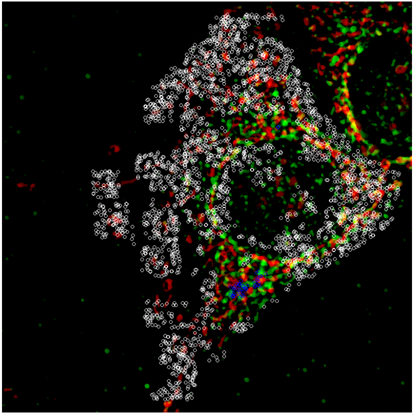


$p$-value $= 0$

# Co-localization



H1B mitochondria
EB1 (microtubules plus ends)
DNA



White circles: co-localization hits
Blue circles: anti co-localization hits

# Outline

Motivation

Previous work

Our approach

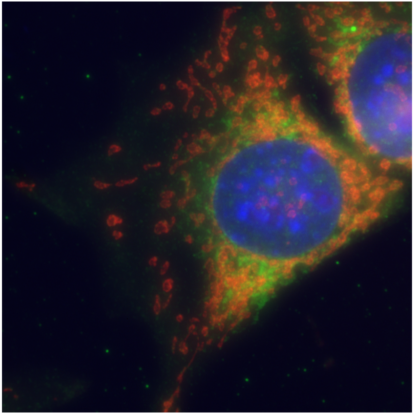
Evaluation

Conclusion

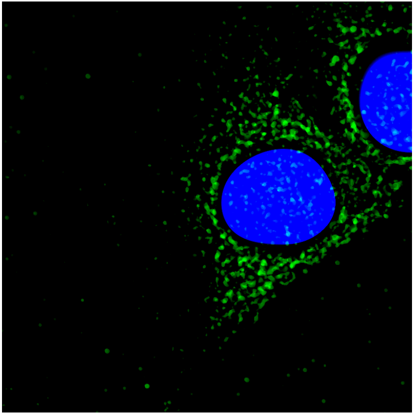
# Conclusion

- **Non-parametric** procedure to test **co-localization** and **anti-colocalization**
- **Fast** and **reliable** approach
- Method adapted to **any size** of 2D and 3D objects, enabling **localized co-localization**

# Co-localization

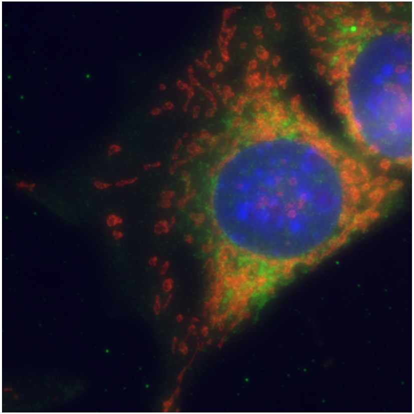


H1B mitochondria
EB1 (microtubules plus ends)
DNA

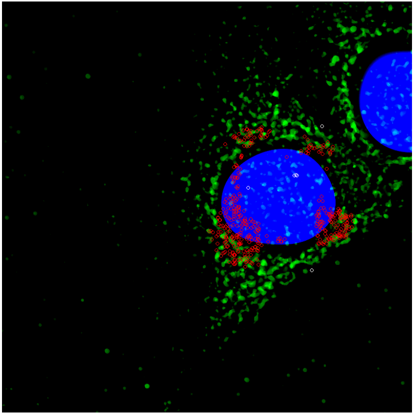


$p$-value $= 0.0799$

# Co-localization



H1B mitochondria
EB1 (microtubules plus ends)
DNA



White circles: co-localization hits
Red circles: anti co-localization hits