



HAL
open science

COMBINING GLOBAL TRACKING WITH STATISTICAL CLASSIFICATION TO ANALYZE ENDOCYTOSIS DYNAMICS USING TIRF MICROSCOPY

Thibault Lagache, Giacomo Nardi, Laetitia Bertot, Alexandre Grassart,
Nathalie Sauvonnet, Jean-Christophe Olivo-Marin

► **To cite this version:**

Thibault Lagache, Giacomo Nardi, Laetitia Bertot, Alexandre Grassart, Nathalie Sauvonnet, et al.. COMBINING GLOBAL TRACKING WITH STATISTICAL CLASSIFICATION TO ANALYZE ENDOCYTOSIS DYNAMICS USING TIRF MICROSCOPY. IEEE International Symposium on Biomedical Imaging, Apr 2017, Melbourne (AUS), France. 10.1109/ISBI.2017.7950706 . hal-04778865v1

HAL Id: hal-04778865

<https://hal.science/hal-04778865v1>

Submitted on 19 Nov 2024 (v1), last revised 27 Nov 2024 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

COMBINING GLOBAL TRACKING WITH STATISTICAL CLASSIFICATION TO ANALYZE ENDOCYTOSIS DYNAMICS USING TIRF MICROSCOPY

Thibault Lagache^{◦,§,*} Giacomo Nardi^{◦,§} Laetitia Bertot[‡]
Alexandre Grassart[‡] Nathalie Sauvonnet[‡] Jean-Christophe Olivo-Marin^{◦,†}

[◦] Institut Pasteur, BioImage Analysis Unit, CNRS UMR 3691. Paris, France
CNRS URA 2582, Paris, France

[‡] Institut Pasteur, Molecular Microbial Pathogenesis Unit, INSERM U1202. Paris, France
CNRS URA 2581, Paris, France

[§] Equally contributing authors

ABSTRACT

In this work we set several mathematical tools to study the role of Dynamin and Endophilin in the Clathrin-mediated endocytosis process. Their different dynamics and co-localizations are observed by using TIRF microscopy. We define in particular a novel tracking method in order to track the Clathrin-coated pits and quantify their co-localization with the proteins involved in the process. Thereafter, we perform a statistical analysis of the lifetimes of Clathrin-coated pits tracks by a suitable mixture model. This allows to define different populations of interest and point out the role of Endophilin and Dynamin in the process.

Index Terms— Global tracking, Hybrid model selection, Endocytosis, Clathrin, Fitting, Mixture model

1. INTRODUCTION

Endocytosis is a key biological process that enables cell to internalize molecular cargos from extracellular medium. The most studied pathway involves Clathrin-coated pits at cell surface where cargos arrival to specific receptors activates pit maturation, inward budding and final pinching of Clathrin-coated vesicles containing cargos [1]. Effective cargo endocytosis by Clathrin-coated pits is conditioned by the precise orchestration of multiple molecular players [2], and the understanding of the key actors of endocytosis success and kinetics is a very active field of investigation.

A central factor essential for the completion of Clathrin vesicle is Dynamin (Dnm). This pinching enzyme directs the fission of the vesicle from the plasma membrane. The role of the membrane curvature sensitive protein Endophilin (Endo) is more debated. Indeed, Endophilin is an actor of Clathrin dependent endocytosis that is proposed to trigger Dynamin recruitment and to act on the vesicle scission (Sundborger A 2014). However, Endophilin seems dispensable for Clathrin vesicle completion thereby raising the question of its exact role.

Here, we investigate the role of Dnm and Endo in Clathrin mediated endocytosis. For this, we engineered genome-edited cells and obtained fluorescent Endo, Dnm and Clathrin light chain (Clc) in physiological conditions. We then monitored during 300s live cell

endocytosis dynamics with dual color, total internal reflexion fluorescence (TIRF) microscopy. TIRF microscope uses an evanescent wave to selectively illuminate and excite fluorophores close to cell membrane (< 200 nanometers depth), removing fluorescence background coming from cell cytosol and allowing the live monitoring of endocytosis at time frequency above 1 Hz. Thus, TIRF imaging is particularly well-suited for analyzing the kinetics of endocytosis, as the disappearance of Clathrin spot at specific pits marks either the pinching off the Clathrin coated vesicle inside the cell, or the disassembly of the coated pit before effective vesicle internalization. Dual-color TIRF allowed us to monitor either Endo or Dnm in addition to Clc and analyzed Clc dynamics when associated or not with one of its partner. Finally, we also used cells depleted for either Dnm or Endo to test the effect of each protein on endocytosis.

The two main technical issues that arise when analyzing live TIRF images of genome-edited Clc are the robust tracking of Clc tracks over time and the statistical interpretation of track lifetimes. Indeed, the high density of Clathrin-coated pits and the poor signal to noise ratio in genome-edited cells impairs the robust tracking of endocytosis events which may last several minutes. Moreover, the pinching-off and internalization of Clathrin vesicles is indistinguishable from pits disassembly before effective entry as both events are marked by spot disappearance from TIRF field of view.

To tackle this two issues, we first propose a new global tracking method that fully uses the high confinement of Clc spots at pits' positions during endocytosis (Section 2). Second, as unsuccessful endocytosis shall present a lifetime shorter than true internalizations [3], we propose a statistical analysis of tracks lifetimes by a suitable mixture model selected with Bayesian information (Section 3). Altogether, these new tracking and statistical methods offer an insight into the dynamical role of Endophilin and Dynamin in the Clathrin-mediated endocytosis.

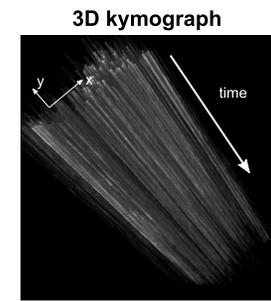
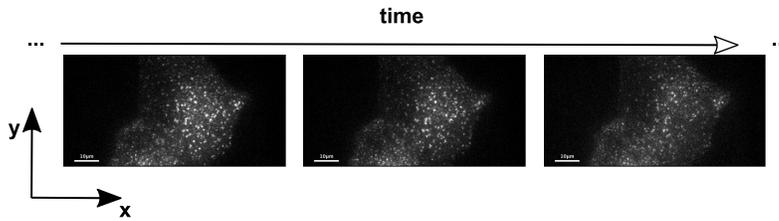
2. TRACKING

High density of Clathrin spots and their weak signal to noise ratio (SNR) in genome-edited conditions make their robust tracking over time difficult. This hinders the use of standard tracking algorithms that use detections at times $t - k, \dots, t$, with $k \sim 1 - 5$ typically, to construct spot trajectories iteratively. Indeed due to many false and missing detections, endocytic tracks are usually shortened, and step by step false associations lead to artifactual displacement of spots while they are supposed to stay nearly static over time (Fig. 1-A).

*T.L. is funded by a fellowship from the Fondation pour la Recherche Médicale and a Grant from the Philippe Foundation. Present Address: the Neuro-Technology Center at Columbia University, 10027 New York

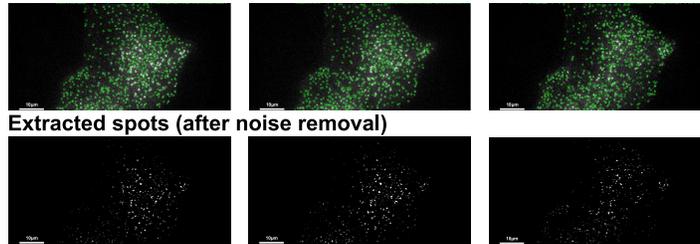
[†]Corresponding author: jcolivo@pasteur.fr.

A - Endocytic tracks are highly confined

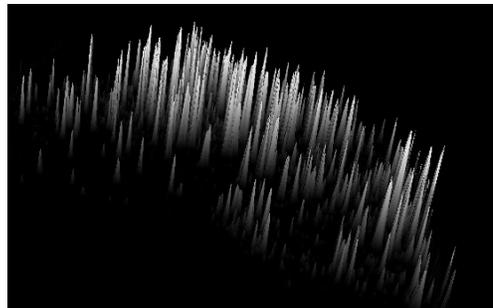


B - Principles of non-Bayesian, global tracking

1- Automatic detection of clathrin spots with à trou wavelet transform

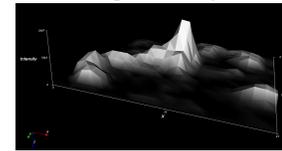


Elevation map of spots' sum

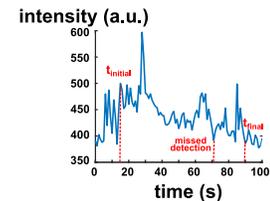
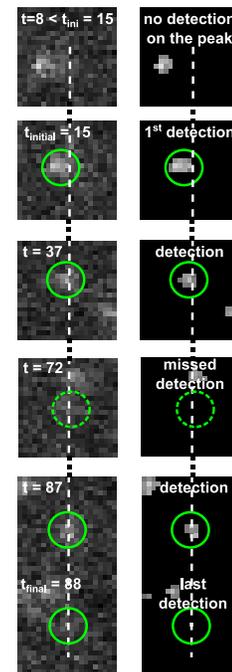


One Gaussian peak
= One track position

Back tracking for
each gaussian peak



Gaussian peak



C - Comparison with Bayesian algorithm

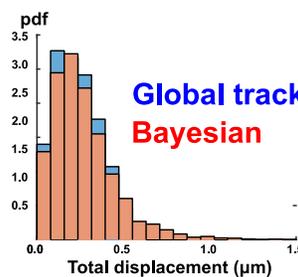
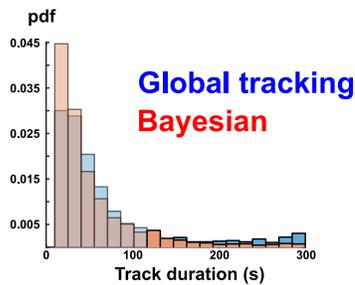


Fig. 1. A - Endocytic tracks are highly confined Each endocytosis spot begins by the appearance of a Clathrin fluorescent spot, its stationarity during the whole track and its final disappearance from the TIRF field of view, that marks either the internalization of the Clathrin-coated vesicle or its disassembly before effective cargo internalization. The 3D (x,y,t) kymograph of tracks highlights the highly confined motion of endocytic spots. Indeed, trajectories appear as lines parallel to the time dimension, which indicates a quasi-constant (x,y) position over time. **B - Principles of global tracking 1-** We first use wavelet-based denoising and Clathrin spots' detection. We then sum the detected spots over time and obtain a global elevation map (x,y,intensity) of the whole process. Gaussian peaks indicate putative endocytic sites with multiple spot appearances. **2-** We then detect each peak with local maximum algorithm and reconstruct endocytic tracks by associating detected spots to closest peaks. The unambiguous localization of endocytic spots allows us to robustly monitor the intensity of each track, even when detections are missing. **C - Comparison with enhanced Multiple Hypothesis Tracking (eMHT) algorithm.** We plotted tracks' durations and total displacements obtained with our global tracking method (red) and eMHT tracking algorithm [4], (blue).

Recently, Aguet et al. proposed a *master-slave* approach where an over-expressed, bright molecule that is very stable all along the endocytosis process is used as a *master* channel and helps to determine the positions of endocytic events unambiguously [5]. Then, the genome-edited *slaved* molecule of interest, which is in physiological concentrations and much less bright than over-expressed *master* signal, can be robustly tracked at pit position along the *master* track. The main issue with Aguet’s method is that it requires a specific fluorescent channel for Master’s tracking, which complicates experimental settings and diminishes the number of fluorescent channels that can be used for the multi-color analysis of different molecules implicated in the endocytosis process.

Here we take advantage of spots’ stationarity to determine the positions of putative endocytic sites very precisely, and then use these *Master* positions to back-track fluorescent spots robustly.

To determine the positions of endocytic sites, we first denoise live TIRF images and detect bright fluorescent spots by applying a *à trou* wavelet transform to images before thresholding statistically wavelet coefficients [6]. For this, we used *Spot detector* plugin (plugin ID R3M2Y2) in open-source image analysis software Icy [7] (<http://icy.bioimageanalysis.org>) (Fig. 1-B-1).

Then, we sum denoised images and obtain an elevation map $(x, y, \text{intensity})$ of the whole endocytic process (Fig. 1-B-1). We model each denoised single spot with a two-dimensional Gaussian profile $I(x, y) = A_i e^{-\frac{1}{2\sigma^2}((x-x_i)^2 + (y-y_i)^2)}$, where σ is the single spot spread due to diffraction, (x_i, y_i) the position of the spot i and A_i its amplitude. The sum of the k (detected) spots along a specific track confined around position (x_0, y_0) with deviation σ_0 (that accounts for both localization error and confined motion) gives rise to a Gaussian peak with mean position (x_0, y_0) , amplitude $A_0 = \sum_{i=1..k} A_i \approx k\bar{A}$, with \bar{A} the mean spot amplitude, and standard deviation $\sqrt{\sigma_0^2 + \sigma^2}$. Consequently, Gaussian peaks with high amplitude indicate long-lasting endocytic events with multiple spot detections over time.

We localize endocytic spots by detecting peak positions with standard local maxima algorithm. Then, using these *Master* positions of endocytic spots, we construct global Clathrin tracks by associating detected spots to closest Gaussian peak (Fig. 1-B-2). The unambiguous localization of endocytosis positions allows us to reconstruct the full intensity profile of each endocytic track even for frames where detections are missing (Fig. 1-B-2).

Finally, we compared our global method to state of the art tracking algorithm (enhanced Multiple Hypothesis Tracking (eMHT) method [8, 4], implemented in *Spot tracking* plugin in Icy (plugin ID L5S9M5) (Figure 1-C). We observe that the number of shortest tracks (<40 seconds) was diminished by 30% with global tracking (42% of the tracks) compared to eMHT (54.8% of tracks). We also observed that the number of tracks with a large total displacement ($> 0.5\mu\text{m}$), which shall correspond to tracks with false step-by-step associations leading to artifactual spot motion, was drastically diminished with our method (3.0% of tracks) compared to eMHT (9.8%).

We thus conclude that tracks obtained with global tracking are much confined and longer than tracks obtained with standard multiple-hypothesis tracking methods, and better corresponds to confined and long-lasting endocytose process.

3. STATISTICAL ANALYSIS

This section is devoted to the statistical analysis of the Clc-lifetimes in order to define different populations and compare their behaviour

depending on the depletion of each of the proteins. It has been previously reported that clathrin-coated pits lifetimes are very heterogeneous because many CLC tracks failed to form vesicles [3, 9, 10]. Then, the goal of this study is first to distinguish active Clathrin tracks from non-internalized ones (abortive or static). Although these questions have already been addressed in [3, 5], this work allows to identify the different populations more precisely in order to deeply describe the endocytosis process (Section 4).

By using the dual-color TIRF microscopy we monitor during 300s the co-localization between Clc and Dnm or Endo. In the control cases (siC) we observe both the Clc tracks presenting a co-localization with Endo or Dnm, and the Clc tracks with no co-localization. In the depleted cases (siEndo, siDnm) we observe the Clc-tracks presenting a co-localization with the remaining protein. In the control cases we have 4353 tracks that are co-localized with Clc (Clc-siC-coloc) and 6931 tracks that are non-co-localized (Clc-siC-non-coloc). The data sets of co-localized tracks when Dnm or Endo is depleted (Endo-Clc-siDnm-coloc and Dnm-Clc-siEndo-coloc, respectively) contain 3761 and 2353, respectively.

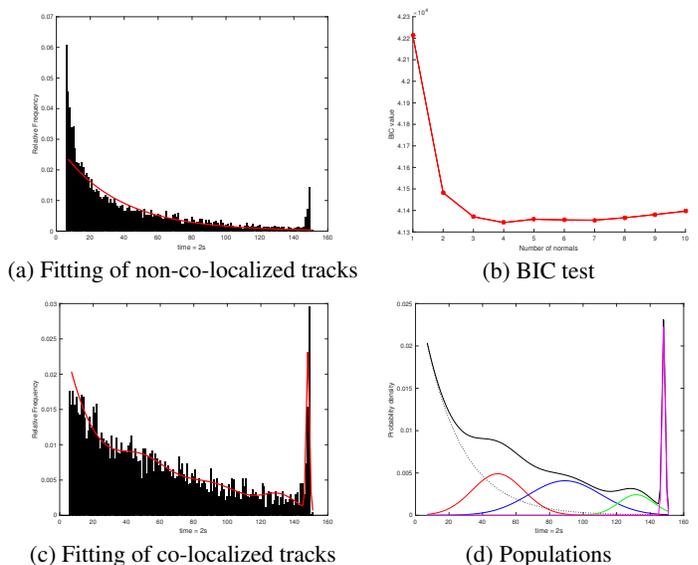


Fig. 2. Fitting of lifetimes distribution for co-localized and non-co-localized tracks in the control cases

As shown in Fig. 2(a) the Clc-siC-non-coloc data set follows an exponential distribution $\text{Exp}(x, \lambda) = \lambda e^{-\lambda x}$ whose parameter can be easily estimated by the *expfit* function of Matlab 2016a ($\lambda \approx 0.03$). This reveals short-lived CLC tracks at the plasma membrane that should account for aberrant detection and/or non-productive clathrin assembly [3, 9, 10]. The life-times distribution of co-localized tracks is different from the non-co-localized ones but part of the distribution fits with an exponential one, like in the non-co-localized case. These short-lived tracks should represent the previously defined early/late abortive tracks defining non-productive clathrin pit assembly [3, 9, 10].

Then, concerning the fitting for lifetimes of Clc-siC-coloc, we use the following mixture model

$$F(x, \theta) = p_0 \text{Exp}(x, \gamma) + \sum_{i=1}^N p_i \mathcal{N}(x, \mu_i, \sigma_i^2),$$

where $\sum_{i=0}^N p_i = 1$ and θ denotes the set of parameters of the model: $\theta = (\gamma, \mu_1, \sigma_1, \dots, \mu_N, \sigma_N, p_0, p_1, \dots, p_N)$.

The exponential distribution corresponds to the set of abortive tracks and $\mathcal{N}(x, \mu_i, \sigma_i^2) = \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{-\frac{(x-\mu_i)^2}{2\sigma_i^2}}$ represents the normal distribution corresponding to the i -th population of interest.

The $(3N + 1)$ parameters are estimated by maximizing the likelihood function via the expectation-maximization algorithm. The number of normal distributions is chosen by the *Bayesian information criterion* (BIC) (see Fig.2(b)). The BIC function is minimized by $N = 4$ (Fig. 2(b)) meaning that, for the lifetimes of co-localized tracks, over the exponential population, we have four Gaussian populations of interest identified by the intervals $\mu_i \pm \sigma_i$ ($i = 1, \dots, 4$): Pop1 = 98s (± 22 s, 30% of tracks), Pop2 = 178s (± 50 s, 20% of tracks), Pop3 = 264s (± 18 s, 7% of tracks) and Pop4 = 296s (± 13 , 7% of tracks). The other parameters of the selected model are the following: $(p_0, p_i) = (0.4759, 0.1885, 0.2190, 0.0643, 0.0522)$ (for $i = 1, \dots, 4$), and $\gamma \approx 0.04$.

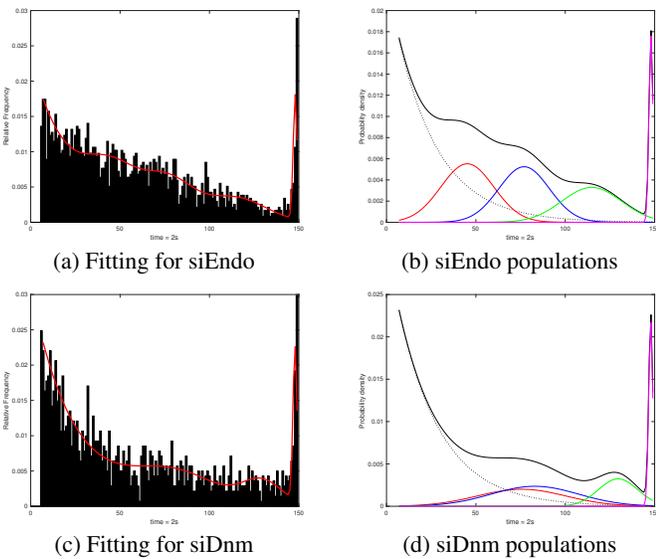


Fig. 3. Fitting of lifetimes distribution for co-localized tracks in the cases siEndo and siDnm

4. BIOLOGICAL RESULTS

In order to quantify how the depletions of Endophilin and Dynamin influence the endocytosis process and affect the different populations, we fit the model F with $N = 4$ to the data sets Endo-Clc-siDnm-coloc and Dnm-Clc-siEndo-coloc (Fig. 3). As shown in Fig. 3(a-b), Endophilin depletion does not affect the distribution of Clc-tracks co-localized with Dynamin and there is no significant difference concerning the distribution of populations.

Strikingly, the depletion of Dynamin in the cells leads to a different Clc lifetimes distribution (Fig. 3(c-d)) and we have the following populations: Pop1 \approx Pop2 = 156s (± 60 , 20%), Pop3 = 258s (± 21 , 8%), and Pop4 = 294s (± 14 , 8%). The fact that the first population overlaps almost completely with the second one suggests that the model has been overfit. This is actually corroborated by the result of the BIC test (for the model F fitted to the Endo-Clc-siDnm-coloc

data set) which gives $N = 3$. This essentially means that this distribution can be described by three normal populations, instead of four like in the control condition (Fig. 2(d)), and that the first population of the control case disappeared.

Since Dynamin is absolutely required in the process, this result reveals that the first population represents the active endocytic Clc tracks. In contrast, populations 3 and 4 are static Clc tracks that are not internalized during almost all the movies time (300 sec). These populations would be particularly enriched in Clathrin lattice objects. Population 2 might represent a mixture of late productive Clc tracks in which the process would be delayed with some non-productive tracks that would be finally abortive.

We finally conclude that the methods above described define a new methodology to study the endocytosis process. As said in Section 2, the tracking model strongly improves the results with respect to other algorithms. Moreover, the robust statistical analysis by a mixture model and the BIC test for selection allow to precise the role of different proteins by comparing several populations of interest.

5. REFERENCES

- [1] Harvey T McMahon and Emmanuel Boucrot, “Molecular mechanism and physiological functions of clathrin-mediated endocytosis,” *Nat Rev Mol Cell Biol*, vol. 12, no. 8, pp. 517–33, Jul 2011.
- [2] Marcus J Taylor, David Perrais, and Christien J Merrifield, “A high precision survey of the molecular dynamics of mammalian clathrin-mediated endocytosis,” *PLoS Biol*, vol. 9, no. 3, pp. e1000604, Mar 2011.
- [3] Dinah Loerke et al., “Cargo and dynamin regulate clathrin-coated pit maturation,” *PLoS Biol*, vol. 7, no. 3, pp. e57, Mar 2009.
- [4] Nicolas Chenouard, Isabelle Bloch, and Jean-Christophe Olivo-Marin, “Multiple hypothesis tracking for cluttered biological image sequences,” *IEEE Trans Pattern Anal Mach Intell*, vol. 35, no. 11, pp. 2736–3750, Nov 2013.
- [5] François Aguet et al., “Advances in analysis of low signal-to-noise images link dynamin and ap2 to the functions of an endocytic checkpoint,” *Dev Cell*, vol. 26, no. 3, pp. 279–91, Aug 2013.
- [6] Jean-Christophe Olivo-Marin, “Extraction of spots in biological images using multiscale products,” *Pattern Recognition*, vol. 35, no. 9, pp. 1989–1996, 2002.
- [7] Fabrice de Chaumont et al., “Icy: an open bioimage informatics platform for extended reproducible research,” *Nat Methods*, vol. 9, no. 7, pp. 690–6, Jun 2012.
- [8] Nicolas Chenouard et al., “Objective comparison of particle tracking methods,” *Nat Methods*, vol. 11, no. 3, pp. 281–9, Mar 2014.
- [9] Marcelo Ehrlich et al., “Endocytosis by random initiation and stabilization of clathrin-coated pits,” *Cell*, vol. 118, no. 5, pp. 591–605, 2004 Sep 3.
- [10] Marcel Mettlen et al., “Cargo- and adaptor-specific mechanisms regulate clathrin-mediated endocytosis,” *J. Cell Biol.*, vol. 188, no. 6, pp. 919–33, 2010 Mar 22.