QuantEv: quantifying the spatial distribution of intracellular events

Thierry Pécot, Liu Zengzhen, Jérôme Boulanger, Jean Salamero, Charles Kervrann

QuantEv: quantifying the spatial distribution of intracellular events

Thierry Pécot[1], Liu Zengzhen[2], Jérôme Boulanger[2], Jean Salamero[2,3], and Charles Kervrann[1]

[1]Inria, Centre Rennes-Bretagne Atlantique, SERPICO Project Team, F-35042 Rennes

[2]Institut Curie, PSL Research University, CNRS UMR 144, Space Time Imaging of Endomembranes Dynamics Team, F-75248, Paris

[3]Cell and Tissue Imaging Facility - IBiSA, Institut Curie, F-75248 Paris

Analysis of the spatial distribution of endomembrane trafficking is fundamental to understand the mechanisms controlling cellular dynamics, cell homeostasy, and cell interaction with its external environment in normal and pathological situations. The development of automated methods to visualize and quantify the spatial distribution of intracellular events is essential to process the ever-increasing amount of data generated with modern light microscopy. We present a generic and non-parametric framework to quantitatively analyze and visualize the spatio-temporal distribution of intracellular events from different conditions in fluorescence microscopy. From the spatial coordinates of intracellular features such as segmented subcellular structures or dynamic processes like vesicle trajectories, QuantEv automatically estimates weighted densities for each dimension of the 3D cylindrical coordinate system and performs a comprehensive statistical analysis from distribution distances. We apply this approach to study the spatio-temporal distribution of moving Rab6 fluorescently labeled membranes with respect to their direction of movement in cells constrained in crossbow- and disk-shaped fibronectin patterns. We also investigate the position of the

generating hub of Rab11 positive membranes and the effect of actin disruption on Rab11 trafficking in coordination with cell shape. An Icy plugin and a tutorial are available at http://icy.bioimageanalysis.org/plugin/QuantEv.

Modern light microscopy associated with fluorescence molecule tagging allows studying the spatial distribution of intracellular events. Unfortunately, fluorescent images are complex to analyze and additional software is needed to evaluate statistical differences between different conditions [1]. Automatic methods have the obvious advantage of being quicker and reproducible. However, most computational methods are based on the complex combination of heterogeneous features such as statistical, geometrical, morphological and frequency properties [2], which makes difficult to draw definitive biological conclusions. Additionally, most experimental designs, especially at single cell level, pool together data coming from replicated experiments of a given condition [3–5], neglecting the biological variability between individual cells.

Micro-patterning is now a well established strategy to reduce morphological variability by imposing constraints on adhesion sites, which has been shown to influence the cytoskeleton geometry and transport carrier localization [3,6]. This technique opened the way to pairwise comparisons of conditions with a two-sample kernel density-based test by pooling together all data from each condition [7]. Unfortunately, it does not consider the sample-to-sample variability because all replicated experiments from a given condition are simply merged together. Additionally, the visualization of the kernel density maps enables to average several experiments but fails to identify specific locations of interest in the cell (*e.g.* docking areas). Finally, assessing the dynamical behavior of

labeled membrane structures, a fundamental task for trafficking analysis, remains out of scope in this framework.

In this paper, we describe a method that we call QuantEv dedicated to the analysis of the spatial distribution of intracellular events represented by any static or dynamical descriptor (*e.g.* detected points, segmented regions, trajectories...) provided that the descriptors are associated with spatial coordinates. QuantEv offers a unifying frame to decipher complex trafficking experiments at the scale of the whole cell. It is typically able to detect subtle global molecular mechanisms when trajectory clustering fails. An overview of the approach is presented in Fig. 1. Our approach first computes 3D histograms of descriptors in a cylindrical coordinate system (parameterized by radius r, angle θ and depth z) with computational cell shape normalization, enabling comparisons between cells of different shape. Densities are obtained via adaptive kernel density estimation [8,9]. Visualization through histograms and densities allows giving a clear biological interpretation of the experiments. We use the Earth Mover's Distance [10] and the Circular Earth Mover's Distance [11] to measure the dissimilarity between densities associated to different experimental conditions. A statistical analysis on these distances reliably takes into account the biological variability over replicated experiments. By computing weighted densities for each point in the cell as the reference center, QuantEv identifies the point that gives the most uniform angular distribution. This point may coincide with a biological structure of interest that would act as the events emitter or attracter.

In the section *Results*, we describe the application of QuantEv to detect significant differences between molecular trafficking and phenotypes observed in cells with various shapes. The

first application is concerned with the distribution of membranes labeled by GFP-Rab6 as a hallmark of vesicular carriers in crossbow- and disk-shaped cells. Rab6 proteins are transiently anchored to moving transport carriers from the Golgi apparatus located at the cell center to Endoplasmic Reticulum entry sites or to plasma membrane [12–17], both assumed to be located at the cell periphery. Micro-patterns impose constraints on the cytoskeleton and consequently influence the spatial distribution of Rab6 transport carriers, as confirmed with kernel density maps [3]. We apply QuantEv to visualize and quantify this influence and to localize regions in the cell associated with Rab6 trafficking stages. In addition, Rab6 positive membranes were reported to move from and towards the Golgi in apparent close proportions [16,18], and yet these membrane associated proteins are believed to traffic in majority from the Golgi located at the cell center to the cell periphery [12–16] where they should dissociate from membranes and recycle back to the cytosol. To investigate these apparently antagonist statements, we apply QuantEv on Rab6 trajectories to characterize the dynamical behaviors of these transport carriers.

The second application focuses on the dynamics of mCherry-Rab11 positive membranes. Rab11 is known to primarily localized to the Endosomal Recycling Compartment and it organizes spatially and temporally recycling from this compartment [19–22]. Here, we confirm by using QuantEv the hypothesis that the labeled transport intermediates are uniformly distributed around the ERC and close to the plasma membrane plane. Furthermore, we also investigate the progressive effect of actin disruption induced by Latrunculin A injection on the Endosomal Recycling Compartment localization with respect to time. We finally apply QuantEv to analyze the joined influence of actin disruption and cell shape on the rad ial distribution of Rab11 vesicles trafficking.

Results

Visualizing and quantifying the influence of micropatterns on the spatial distribution of Rab6 positive membranes Quantification is unavoidable for nowadays cell biology experiments to assess differences between populations or to evaluate the effects of certain drugs or genes. Easy to interpret visualization coming with quantification is a real plus for biologists to understand what are the features explaining statistical significance. We applied the QuantEv approach to visualize the spatial distribution of Rab6 positive membranes in crossbow- and disk-shaped cells (see Fig. S1 **a-b**) and quantify their differences. We compared the results obtained with QuantEv to those obtained with the more conventional kernel density (KD) maps [3,4]. The KD approach concludes that the distribution of Rab6 positive membranes are clearly different between micro-patterns (see Fig. S2 **a-b**, p-value $= 0$) as expected, but also leads to a significant difference between distributions from a same micro-pattern (see Fig. S2 **c**). Instead, QuantEv selectively identifies a significant difference between the angular distributions of Rab6 positive membranes from crossbow- and disk-shaped cells (p-value $= 0.0006$) while the radial and in-depth distributions are similar for both micro-patterns (see Fig. 2 **a-b**). QuantEv also highlights a distribution peak for a radius (resp. a depth) at the two-thirds the distance between the Golgi region border and the cell periphery (resp. between cell top and cell bottom), for both micro-patterns (see Fig. 2 **a;c-d**). This peak corresponds to an accumulation of Rab6 positive membranes and identifies the area where they enter a docking phase before switching to a tethering phase. The KD maps do not show such subtle mechanisms in their primary formulation.

Inwards and outwards Rab6 positive membranes show two distinctive dynamical behaviors

Rab6 positive membranes are trafficking from the Golgi located at the cell center to the cell periphery [12–16] and at the same time move from and towards the Golgi in comparable proportions [16,18]. To reconcile these two antagonist statements, we applied QuantEv as follows. Rab6 trajectories were classified into two categories (Fig. 3 **a-b**): i) vesicles moving towards the cell periphery; ii) vesicles moving towards the Golgi. As shown in Fig. 3 **c-d**, the proportion of Rab6 positive membranes moving towards the cell periphery and towards the Golgi are close for both micro-patterns (0.501 *versus* 0.499 for crossbow-shaped cells, 0.515 *versus* 0.485 for disk-shaped cells). The radial distributions shown in Fig. 3 **c-d** display two distinctive modes for vesicles moving towards the cell periphery and those moving towards the Golgi (p-value $= 0.0304$ for crossbow-shaped cells, p-value $= 0.0047$ for disk-shaped cells). Between the Golgi and the distribution peak shown in Fig. 2 **a**, Rab6 vesicles are predominantly moving towards the cell periphery. Between this peak and the cell periphery, they are in majority moving towards the Golgi, indicating that during their docking-tethering phase, the vesicles are predominantly moving towards the cell center. These two distinctive dynamical behaviors are consistent with the aforementioned antagonist statements. To go further in the analysis, we pooled together the trajectories of Rab6 positive membranes from crossbow- and disk-shaped cells as they show a similar radial distribution. We then applied QuantEv without any weight (Fig. 3 **f**) and by considering the confinement ratio (see Fig. 3 **g**) and the total path length (see Fig. 3 **h**) to normalize the radial distribution. Rab6 positive membranes moving towards the cell periphery have a much more direct path than the ones moving towards the Golgi (0.54 *versus* 0.46, p-value $= 6 \times 10^{-8}$, see Fig. 3 **g**) while Rab6 positive membranes moving

towards the Golgi follow longer trajectories than the ones moving towards the cell periphery (0.54 *versus* 0.46, p-value $= 0.0001$, Fig. 3 **h**). In summary, this analysis clearly demonstrated that Rab6 positive membranes move predominantly and quite directly from the Golgi to the cell periphery until they enter a docking phase. Then, they mostly go back towards the cell center by following long and indirect trajectories.

The Endosomal Recycling Compartment (ERC) organizes Rab11 angular distribution Rab11 positive recycling membranes originate their journey from the so-called Endosomal Recycling Compartment (ERC). We formulate the assumption that Rab11 positive membranes are uniformly distributed at the membrane plane around the ERC position with cell, whatever the cell shape is. To test this hypothesis, we used images acquired at the membrane with TIRF microscopy showing Rab11 proteins (see Fig. S1 **c-d**). Most labeled membranes of the ERC are not located near the cell surface. However, for each TIRF sequence, one HiLo (Highly Inclined TIRFM) or pseudo-wide field image was acquired, giving the biology expert the capacity to manually define its location (red disks in Fig. 4 **a**). To test our assumption, the QuantEv uniformity analysis is applied by considering intensity on segmented regions. The results are shown in Fig. 4 **a** (blue disks). To have a line of comparison, we also plot the cell centers as green disks in Fig. 4 **a**. Interestingly, the blue disk is close to the red disk for all image sequences except one (second line, middle image in Fig. 4 **a**). The blue disk is also closer to the red disk than the green disk in 7 out of 8 image sequences (see Fig. 4 **a-b**). Although the point that gives the most uniform angular distribution does not strictly coincide with the manually identified ERC, it is sufficiently close to indicate that the Rab11 positive membranes are quite uniformly distributed around the ERC position at the membrane plane

whatever the cell shape is. This indicates that the ERC corresponds to the organizing hub of the Rab11 carrier vesicles.

Joint actin disruption and cell shape influence on Rab11 radial distribution Applying the QuantEv uniformity analysis at each time step of a sequence allows studying the location stability of the particle emitter or attractor. To test if the estimated ERC location is stationary over time, we computed the Euclidean distance between the reference point estimated at time $t = 0$ and the points estimated for the next frames. On the image sequences considered in the previous section (see Fig. 4 **a**), this distance remains stable (see Fig. S3 **a**). We analyzed cells treated with Latrunculin A, which inhibits actin polymerization and disrupts actin filaments (see Fig. S1 **e-f**). We show that the ERC location is moving away as the drug is affecting the cell (see Fig. S3 **b**), enlightening the role of cytoskeleton in stabilizing the localization of the ERC. We then acquired image sequences at 10, 15, 20 and 25 minutes after Latrunculin A addition and we extracted Rab11 trajectories. The confinement ratio of Rab11 tracks is decreasing with time (see Fig. S4), which is consistent with actin cytoskeleton being involved in Rab11 vesicle trafficking. The radial distribution of Rab11 vesicles is constantly shifting from the cell periphery to the cell center for both micro-patterns (see Fig. 5). However, we first observe significant differences between the radial distrubutions for the two tested micropatterns controlling the actin organization at injection time (p-value = 0.0129). After Latrunculin A treatment, we progressively observe no difference between the radial distributions, suggesting that the actin organization is drastically perturbed. All these quantifications allow us to conclude that exocytosis/recycling vesicle trafficking is cell shape-actin organization dependent.

Discussion

This article presents a robust computational framework taking into account cell variability to quantify the distribution of fluorescently labeled proteins. Using dynamical descriptors, detailed insight into dynamical processes is also unraveled and the uniformity analysis allows to localizing an organizing region for the observed biological objects.

QuantEv enables us to state that the angular distribution of Rab6 positive membranes from crossbow- and disk-shaped cells are statistically different. It also enables to visualize the Rab6 positive membranes distribution along the depth and between the Golgi and the cell periphery, identifying the locations where they enter their docking phase. By considering the directions of the moving Rab6 positive membranes, QuantEv allows demonstrating that these membranes first move predominantly and directly towards the cell periphery before reaching their docking phase. They then go back to the cell center in an undirected and long fashion. This intriguing result showing statistically bi-directional movements of Rab6 was reported before. The Rab6 positive vesicles generated at the Golgi membranes are predestined to the cell periphery, in order to deliver their exocytic cargo [16,18], which should favor a centrifuge directionality. Our data reconciles this two apparently opposed observations and show for the first time, that a majority of Rab6 vesicles reverses their movement only towards close docking-fusion sites and only during this ultimate phase of docking-fusion.

We demonstrate with the help of QuantEv that Rab11 positive membranes are uniformly distributed around the Endosomal Recycling Compartment (ERC) at the plasma membrane plane.

This shows that the ERC represents an organizing hub for the Rab11 carrier vesicles. By applying the QuantEv uniformity analysis along time, we exhibit how the ERC location is affected by actin disruption caused by Latrunculin A injection. The radial distribution analysis of Rab11 positive membranes in crossbow- and disk-shaped cells combined with Latrunculin A injection reveals the interconnection between cell shape and actin organization on Rab11 trafficking at the plasma membrane, and more generally on the exocytosis/recycling vesicle distribution.

In conclusion, QuantEv has the potential to become a very popular analysis method for dynamics and intracellular event analysis as i) it is publicly available; ii) it is fully automated and non-parametric; iii) it provides results that are easy to biologically interpret; iv) it performs a statistical analysis that takes into account the biological variability over the replicated experiments of a same condition and is efficient with small and large amounts of data. QuantEv is quite flexible since the user can specify any distance, including reference to a unique point and to membrane borders [23]. The QuantEv plugin can be applied with Cartesian, cylindrical and spherical coordinate systems to analyze flat or rounded cells. It is worth noting that this approach can be applied to constrained and non-constrained cell shapes. An Icy plugin and a tutorial are available at http://icy.bioimageanalysis.org/plugin/QuantEv.

Materials and methods

Sample preparation In the first dataset, we use cell lines stably expressing fluorescently tagged proteins in order to minimize the cell-to-cell variability in fluorescence signal. HeLa cells stably

transfected with GFP-tagged Rab6 proteins were maintained in DMEM supplemented with 10% fetal bovine serum. Cells were spread onto fibronectin Cytoo chips (Cytoo Cell Architect) 4 to 5 hours before imaging. Cell adhesion on micropatterns both constrains the cells in terms of lateral movement and averages their size and shape. Two types of micropatterns were used (disk-shaped and crossbow-shaped, Cytoo Cell Architect, $1100\mu m^2$)[6]. For a second set of experiments, wild type RPE1 cells were grown in Dulbecco's Modified Eagle Medium, Nutrient Mixture F-12 (DMEM/F12) supplemented with 10% (vol/vol) FCS in 6 well plates. RPE1 cells were transiently transfected with plasmids coding for Rab11a-GFP, and Langerin-mCherry using the following protocol: 2 μg of each DNAs, completed to 100 μL with DMEM/F12 (FCS free) were incubated for 5 min at room temperature. 6 μL of X-tremeGENE 9 DNA Transfection Reagent (Roche) completed to 100 μL with DMEM/F12 (FCS free) were added to the mix and incubated for further 15 min at room temperature. The transfection mix was then added to RPE1 cells grown one day before and incubated further at 37^{o}C overnight. Cells were then spread onto fibronectin Cytoo chips (Cytoo Cell Architect) for 4h at 37^{o}C with F-12 (with 10% (vol/vol) FCS, 10mM Hepes, 100 units/ml of penicillin and 100ug/ml of Strep) before imaging. When specified, 2 mM Latrunculin A (Sigma) was dissolved to 0.02 mM in F-12 DMEM. 300 μL of culture medium with Latrunculin A (600 nM) was added to establish a final Latrunculin A concentration of 3 μM.

Data acquisition For the Rab6 dataset, the 488 nm laser of a spinning-disk confocal microscope (Ti Eclipse, Nikon, S.A, France equipped with spinning disk system and CoolSnap HQ2 CCD, from Roper Scientific S.A.R.L, France) was used to acquire 3D 380 $\times$ 380 $\times$ 8 stacks (the voxel resolution is 64.5 nm $\times$ 64.5 nm $\times$ 300 nm) at a rate of one stack per second. 18 image sequences

with crossbow-shaped cells and 22 image sequences with disk-shaped cells were acquired. The system was driven by the Metamorph software (Molecular Devices). The volume rendering of two images from this dataset are shown in Fig. S1 **a-b**.

For the Rab11 dataset, live-cell imaging was performed using simultaneous dual color Total Internal Reflection Fluorescence (TIRF) microscopy. All imaging was performed in full conditioned medium at 37^{o}C and 5% CO2 unless otherwise indicated. Simultaneous dual color TIRF microscopy sequences were acquired on a Nikon TE2000 inverted microscope equipped with a x100 TIRF objective (NA=1.49), an Azymuthal TIRF module (Ilas2, Roper Scientifc), an image splitter (DV, Roper Scientific) installed in front of an EMCCD camera (Evolve, Photometrics) that can be bypassed or not, depending on the experimental conditions, as indicated in the text, and a temperature controller (LIS). GFP and m-Cherry were excited with a 488 nm and a 561 nm laser, respectively (100mW). The system was driven by the Metamorph software (Molecular Devices). Four selected image projections from this data set are shown in Fig. S1 **c-f**.

Data availability We use two datasets in this study that are publicly available on the *iMANAGE* database at https://cid-curie.fr/iManage/standard/login.html with username *public* and password *Welcome!1* in the project entitled QuantEv-Data.

Event detection and localization Before applying QuantEv, the intracellular events have to be identified and localized. The Rab6 proteins are extracted for each image sequence by using the C-CRAFT method [24] with default parameters, except the p-value that ranges from 0.0025 to 0.0125 depending on the noise level, available on *Icy* [25]. The Rab11 positive membranes are segmented

at each time point with the ATLAS algorithm [26] with default parameters, except the *p*-value that ranges from 0.05 to 0.045 depending on the noise level. In both cases, a variance stabilization transform [27] is performed to take into account the Poisson-Gaussian nature of the noise in the CCD sensors. Finally, the Rab6 and Rab11 trajectories are estimated with the multiple hypothesis tracking method [28] with default parameters, available on *Icy* [25].

Weighted density estimation The localization of events needs to be defined on a common coordinate system to compare the experiments. We propose to use the cylindrical coordinate system where only a reference point such as the event emitter or attracter and a reference direction have to be specified by the user. To fairly compare experiments with different cell shapes, we define appropriate distances to obtain normalized densities, *i.e.* independent from the cell shape. We illustrate the importance of shape normalization in the *Supporting Information* and in Fig. S5 **a-b**.

More formally, let us define Ω the 3D cell support and $\partial\Omega$ the 3D cell surface. Let us consider a set of N sample points associated with intracellular events $\mathcal{S} = \{(r_i, \theta_i, z_i, w_i, d_{\theta_i}, d_{z_i}), i \in [1, N]\}$, where (r_i, θ_i, z_i) denote the spatial cylindrical coordinates. The weight w_i enables to take into account features associated to events such as intensity, track length, confinement ratio... w_i can typically be a function of fluorescence intensity, proportional to the number of molecules observed at a given location. The distance d_{θ_i} is equal to the Euclidean distance between the coordinate system origin $O \in \Omega$ projected on plane z_i (O_{z_i}) and the point $P_{\theta_i,z_i} \in \partial\Omega$ with angle θ_i at plane z_i, such that $d_{\theta_i} = \|P_{\theta_i,z_i} - O_{z_i}\|_2^2$. The distance d_{z_i} is equal to the Euclidean distance between the coordinate system origin O and the point $P_{r_i,\theta_i} \in \partial\Omega$ with radius r_i and angle θ_i

such that $d_{z_i} = ||P_{r_i,\theta_i} - O||$. These two distances allow estimating normalized densities that are independent from cell shapes. All these coordinates are illustrated on a simple example for the event i in Fig. S7. We propose to estimate three densities defined as follows:

$$\begin{aligned} f(r) &= \frac{1}{Z_{r,\theta}} \sum_{i=1}^{N} G_{\hat{\sigma}_r}(r_i - r) \frac{w_i}{d_{\theta_i}}, \\ f(\theta) &= \frac{1}{Z_{r,\theta}} \sum_{i=1}^{N} H_{\hat{\kappa}}(\theta_i - \theta) \frac{w_i}{d_{\theta_i}}, \\ f(z) &= \frac{1}{Z_z} \sum_{i=1}^{N} G_{\hat{\sigma}_z}(z_i - z) \frac{w_i}{d_{z_i}}, \end{aligned} \tag{1}$$

where $G_{\hat{\sigma}}(\cdot)$ is a Gaussian kernel with bandwidth $\hat{\sigma}$, $H_{\hat{\kappa}}$ is a von Mises kernel with concentration $\hat{\kappa}$ such that $H_{\hat{\kappa}}(\theta) = \frac{e^{\hat{\kappa}\cos\theta}}{2\pi I_0(\hat{\kappa})}$ and $I_0(\cdot)$ is the Bessel function of order 0. The bandwidths $\hat{\sigma}_r$ and $\hat{\sigma}_z$ are estimated with the Silverman's rule of thumb [8] and $\hat{\kappa}$ is estimated using the robust rule of thumb proposed by Taylor *et al.* [9]. The normalization constants are defined as follows:

$$Z_{r,\theta} = N \sum_{i=1}^{N} \frac{w_i}{d_{\theta_i}}, \qquad Z_z = N \sum_{i=1}^{N} \frac{w_i}{d_{z_i}}. \tag{2}$$

Statistical procedure Quantitative comparison between different conditions is mandatory to analyze biological data. In most computational biology studies, data from different experiments corresponding to the same condition are pooled together [3,4]. This usual procedure enables to add statistical power when comparing two conditions. Therefore, it is especially useful when few data are available. Unfortunately, pooling data together presents two main drawbacks. First, if large amounts of data are available, the opposite problem arises and the statistical tests may become significant for every comparison [29]. One solution is to downsample the data, but the amount of downsampling becomes another issue. Second, pooling data together for one condition partially

hides the variability between the replicated experiments for this condition. As an example, let us consider a study aimed at analyzing the effects of a drug on a sample of normal individuals. To evaluate the drug efficiency, a comparison between normal individuals and individuals that were administered the drug is conducted. Let us assume that the drug is effective on half the individuals. Consequently, normal individuals are compared to a mix of normal individuals and individuals with the drug effects. This comparison should not be statistically significant as the drug is not efficient on all individuals. However, the effects on the individuals for which the drug is efficient might hide the fact that it is not efficient on all individuals if all the data are pooled together. In what follows, we propose to compute a distance between all experiments instead of a distance between conditions. The idea is demonstrated in the *Supporting Information* and validated on synthetic image sequences (see Fig. S6 and Fig. S5 **c-d**).

Distance between densities We propose to compute the earth mover's distance (also known as the Kantorovich-Rubinstein or the first order Wasserstein distance) between every replicate of every condition to apply a statistical test. This transport-based distance demonstrated its efficiency for other studies on cell phenotypes [30]. The discrete Earth Mover's Distance (EMD) between two unidimensional distributions is simply defined as the sum of the absolute differences between their cumulated distribution functions [10]:

$$EMD(f^1, f^2) \quad = \quad \sum_{i=1}^{K} |F^1(i) - F^2(i)|, \tag{3}$$

where F^1 and F^2 are the cumulated distribution functions of f^1 and f^2. Although the EMD depends on the number of bins K, EMD proportions are kept intact when the number of bins is high enough as shown in Fig. S5 **e-f**. For the angular distribution, the Circular Earth Mover's

Distance (CEMD) [11] is defined as:

$$\text{CEMD}(f^1, f^2) = \min_{k \in \{1,\dots,K\}} \sum_{i=1}^{K} |Q_k^1(i) - Q_k^2(i)|, \tag{4}$$

with

$$Q_k(i) \quad = \quad \begin{cases} \sum_{j=k}^{i} f(j) & \text{if} \quad i \geq k, \\ \sum_{j=k}^{K} f(j) + \sum_{j=1}^{i} f(j) & \text{if} \quad i < k. \end{cases} \tag{5}$$

Difference between conditions The EMD and CEMD enable to compute a distance between two single experiments for the radial, angular and in-depth densities. The distances between the replicates of one condition and the replicates of another condition give an idea about the difference between the two conditions. However, a baseline distance is also needed to state if the difference is random or significant. Therefore, if we consider two conditions, two distances are defined for each experiment and each density:

i) the *intra-condition distance*: average distance between the density and all the other densities for the same condition;

ii) the *inter-condition distance*: average distance between the density and all the other densities from the other condition.

We define as the *condition difference* the difference between the inter-condition distance and the intra-condition distance. If the *condition difference* is high, the two conditions are different.

Statistical test A statistical test is applied on the difference distance to state if the observed conditions are significantly different. A non-parametric statistical test is better suited as there is no underlying model for the condition difference. In addition, a negative condition difference implies that the current experiment is closer to the replicated experiments of the other condition than the replicated experiments of the same condition. Consequently, the condition difference has to be positive if the conditions are different. For those two reasons, we propose to use the one-sided non-parametric Wilcoxon signed-rank test on the condition differences for all experiments to state if two conditions are statistically different.

Analysis of uniform distribution of events In case we focus on the intracellular events assumed to be uniformly distributed around a given biological object, *e.g.* the events emitter, QuantEv allows us to estimate a location for this traffic-organizing component. This source location is then defined as the reference point with the most uniform angular distribution. It is established that the maximum entropy corresponds to the most uniform distribution. Consequently, the reference point O^* is defined as the location that maximizes the entropy:

$$O^* = \max_{O \in \Omega} -\sum_{i=1}^{N} f(\theta_i) \log f(\theta_i). \tag{6}$$

The most straightforward way to find this point is to estimate the entropy map that gives, for each point in Ω, the entropy value computed with the current point used as the reference center. We also propose to use the bisection method to speed up the computation (about ten time faster than the entropy map computation). A uniformity analysis conducted on simulations is presented in the *Supporting Information* and in Fig. S5 **g-h**. The entropy criterion can be extended to detect multiple organizing components if needed.

Code availability The jar file of the QuantEv Icy plugin is available at http://icy.bioimageanalysis.org/plugin/QuantEv. The source code can be extracted from this jar file.

ACKNOWLEDGMENTS This work was supported by the France-BioImaging infrastructure (ANR-10-INBS-04). The Cell and Tissue Imaging Facility at Institut Curie is a member of France-BioImaging.

1. Murphy R, et al. (2012) Focus on Bioimage Informatics. *Nat. Methods* 9(7):629–742.

2. Peng H (2008) Bioimage informatics: a new area of engineering biology. *Bioinformatics* 24(17):1827–1836.

3. Schauer K et al. (2010) Probabilistic density maps to study global endomembrane organization. *Nat. Methods* 7(7):560–566.

4. Merouane A et al. (2015) Automated profiling of individual cell-cell interactions from high-throughput time-lapse imaging microscopy in nanowell grids (TIMING). *Bioinformatics* 31(19):3189–3197.

5. Biot E et al. (2016) Strategy and software for the statistical spatial analysis of 3D intracellular distributions. *Plant J.* 87(2):230–242.

6. Thery M et al. (2005) The extracellular matrix guides the orientation of the cell division axis. *Nat. Cell Biol.* 7:947–953.

7. Duong T, Goud B, Schauer K (2012) Closed-form density-based framework for automatic detection of cellular morphology changes. *Proc. Natl. Acad. Sci. U.S.A.* 109(22):8382–8387.

8. Silverman B (1986) *Density estimation for statistics and data analysis*. (CRC press) Vol. 26.

9. Taylor C (2008) Automatic bandwidth selection for circular density estimation. *Computational statistics and data analysis* 52(7):3493–3500.

10. Rubner Y, Tomasi C, Guibas L (2000) The earth mover's distance as a metric for image retrieval. *Int. J. Comp. Vision* 40(2):99–121.

11. Rabin J, Delon J, Gousseau Y (2011) Transportation distances on the circle. *J. Math. Imaging Vision* 41(1-2):147–167.

12. White J et al. (1999) Rab6 coordinates a novel golgi to ER retrograde transport pathway in live cells. *J. Cell Biol.* 147(4):743–760.

13. Chavrier P, Goud B (1999) The role of ARF and Rab GTPases in membrane transport. *Curr. Opin. Cell Biol.* 11(4):466–475.

14. Echard A et al. (2000) Alternative splicing of the human Rab6a gene generates two close but functionally different isoforms. *Mol. Biol. Cell* 11(11):3819–3833.

15. Opdam F et al. (2000) The small GTPase Rab6B, a novel Rab6 subfamily member, is cell-type specifically expressed and localised to the Golgi apparatus. *J. Cell. Sci.* 113(Pt 15):2725–2735.

16. Grigoriev I et al. (2007) Rab6 regulates transport and targeting of exocytotic carriers. *Dev. Cell* 13(2):305–314.

17. Bardin S et al. (2015) Phenotypic characterisation of RAB6A knockout mouse embryonic fibroblasts. *Biol. Cell* 107(12):427–439.

18. Grigoriev I et al. (2011) Rab6, Rab8, and MICAL3 cooperate in controlling docking and fusion of exocytotic carriers. *Curr. Biol.* 21(11):967–974.

19. Ullrich O, Reinsch S, Urbe S, Zerial M, Parton RG (1996) Rab11 regulates recycling through the pericentriolar recycling endosome. *J. Cell Biol.* 135(4):913–924.

20. Boulanger J et al. (2014) Fast high-resolution 3D total internal reflection fluorescence microscopy by incidence angle scanning and azimuthal averaging. *Proc. Natl. Acad. Sci. U.S.A.* 111(48):17164–17169.

21. Gidon A et al. (2012) A Rab11A/myosin Vb/Rab11-FIP2 complex frames two late recycling steps of langerin from the ERC to the plasma membrane. *Traffic* 13(6):815–833.

22. Baetz NW, Goldenring JR (2013) Rab11-family interacting proteins define spatially and temporally distinct regions within the dynamic Rab11a-dependent recycling system. *Mol. Biol. Cell* 24(5):643–658.

23. Heride C et al. (2010) Distance between homologous chromosomes results from chromosome positioning constraints. *J. Cell. Sci.* 123(Pt 23):4063–4075.

24. Pécot T et al. (2015) Background fluorescence estimation and vesicle segmentation in live cell imaging with conditional random fields. *IEEE Trans. Image Process.* 24(2):667–680.

25. de Chaumont F et al. (2012) Icy: an open bioimage informatics platform for extended reproducible research. *Nat. Methods* 9(7):690–696.

26. Basset A, Boulanger J, Salamero J, Bouthemy P, Kervrann C (2015) Adaptive Spot Detection With Optimal Scale Selection in Fluorescence Microscopy Images. *IEEE Trans. Image Process.* 24(11):4512–4527.

27. Boulanger J, Kervrann C, Bouthemy, P. Elbau P, Sibarita JB, Salamero J (2009) Patch-based non-local functional for denoising fluorescence microscopy image sequences. *IEEE Trans. Med. Imaging* 28(12).

28. Chenouard N, Bloch I, Olivo-Marin JC (2013) Multiple hypothesis tracking for cluttered biological image sequences. *IEEE Trans. Pattern Anal. Mach. Intell.* 35(11):2736–2750.

29. Olivier J, Walter SR (2015) Too much statistical power can lead to false conclusions: a response to 'Unsuitability of the epidemiological approach to bicycle transportation injuries and traffic engineering problems' by Kary. *Inj. Prev.* 21(4):289.

30. Basu S, Kolouri S, Rohde GK (2014) Detecting and visualizing cell phenotype differences from microscopy images using transport-based morphometry. *Proc. Natl. Acad. Sci. U.S.A.* 111(9):3448–3453.

Figure 1: **Overview of QuantEv approach.**

Figure 2: **Spatial distribution analysis of Rab6 proteins for crossbow- and disk-shaped cells. a** Histograms (bar plots) and densities (lines) of the spatial distribution of Rab6 positive membranes with respect to radius r, angle θ and depth z. These distributions come from 18 (resp. 22) image sequences with a crossbow-shaped cell (blue bar plots and lines) (resp. a disk-shaped cell (orange bar plots and lines)). **b** Box and whisker plots of the condition differences with respect to radius r, angle θ and depth z over the 40 image sequences. p-values under conditions of one-sided Wilcoxon signed-rank test when considering the condition differences are indicated below the box and whisker plots. A star (*) indicates that the p-value is smaller than 0.05. **c-d** Overlay of the average intensity projection map of an image sequence with a crossbow- **c** (resp. disk- **d**) shaped cell and the radial levels at 0, 0.6 and 0.8. The scale bar corresponds to 5 μm.

Figure 3: **Spatial distribution analysis of moving Rab6 proteins. a-b** Overlay of the average intensity projection map of an image sequence with a crossbow- **a** (resp. disk- **b**) shaped cell and the Rab6 positive membrane trajectories moving towards the cell periphery (red trajectories) and towards Golgi (green trajectories). The scale bar corresponds to 5μm. **c-d** Histograms (bar plots) and densities (lines) of the spatial distribution of Rab6 positive membranes moving towards the cell periphery (green bar plots and lines) or towards the Golgi (pink bar plots and lines) with respect to radius r, the θ and the z. These distributions come from 18 image sequences with a crossbow-shaped cell **c** and 22 image sequences with a disk-shaped cell **d**. The box and whisker plots of the condition differences of the spatial distribution of moving Rab6 positive membranes with respect to radius r, angle θ and depth z over the 40 image sequences for crossbow- **c** and disk-shaped cells **d** are next to the histograms and densities. **e** Illustration of the displacement distance and the total path length of a trajectory. The confinement ratio is defined as the ratio between the displacement distance and the total path length. **f-h** Histograms (bar plots) and densities (lines) of the radial distributions of Rab6 positive membranes trajectories with no weights **f**, and weighted by the confinement ratio **g** and total path length **h**. These distributions come from 40 image sequences with crossbow-shaped cells and disk-shaped cells. **h** Box and whisker plots of the condition differences of the radial distribution with no weights, weighted by the confinement ratio and by the total path length over the 40 image sequences. p-values under conditions of one-sided Wilcoxon signed-rank test when considering the condition differences are indicated below the box and whisker plots. A star (*) indicates that the p-value is smaller than 0.05.

Figure 4: **Estimation of the Endosomal Recycling Center (ERC) location from the angular distribution of Rab11 positive membranes.** **a** The red disks correspond to the manual annotation, the blue disks to the point defining the most uniform angular distribution of Rab11 positive membranes and the green disks correspond to the cell centers. These disks are displayed over the average intensity projections of the image sequences showing Rab11 positive membranes. The scale bar corresponds to 5μm. **b** Euclidean distances between the manually annotated ERC and the cell centers (green disks) or the points giving the most uniform angular distribution (blue disks).

Figure 5: **Latrunculin A influence on Rab11 radial distribution.** **a** Histograms (bar plots) and densities (lines) of the radial distribution of Rab11 positive membranes on crossbow- and disk-shaped cells weighted by intensity at Latrunculin A Injection Time, 10, 15, 20 and 25 minutes after injection. **b** Box and whisker plots of the condition differences of the radial distribution between crossbow- and disk-shaped cells at Latrunculin A Injection Time, 10, 15, 20 and 25 minutes after injection. p-values under conditions of one-sided Wilcoxon signed-rank test when comparing crossbow- and disk-shaped cells (9 images sequences for each micro-pattern at each injection time) are indicated below the box and whisker plots.

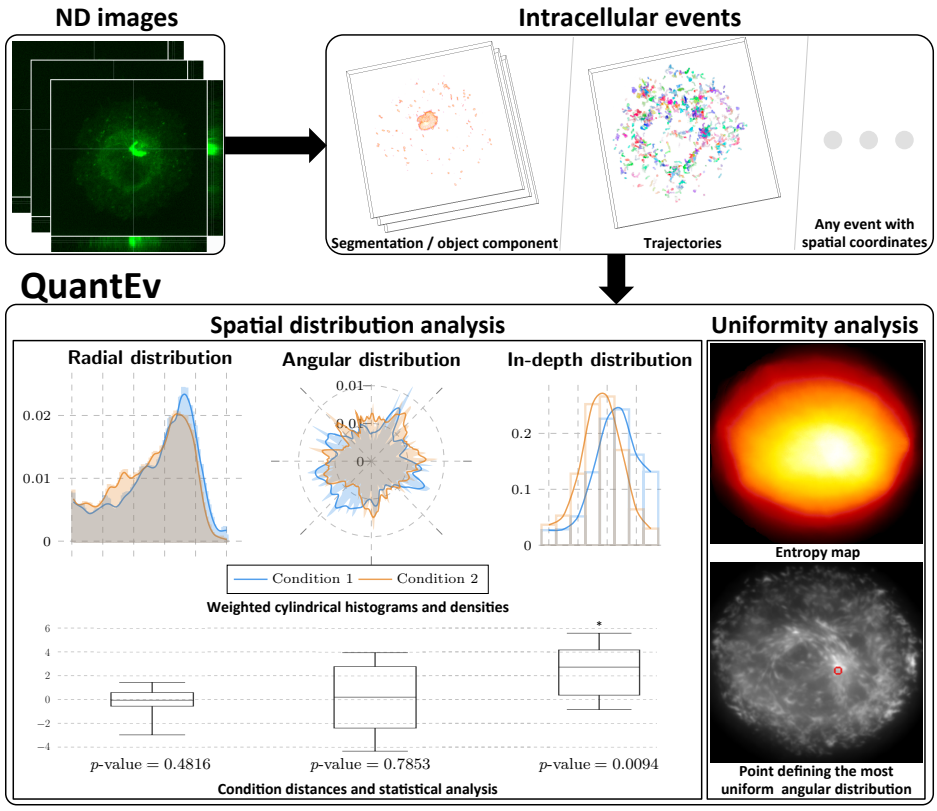

Fig. 1. Overview of QuantEv approach.

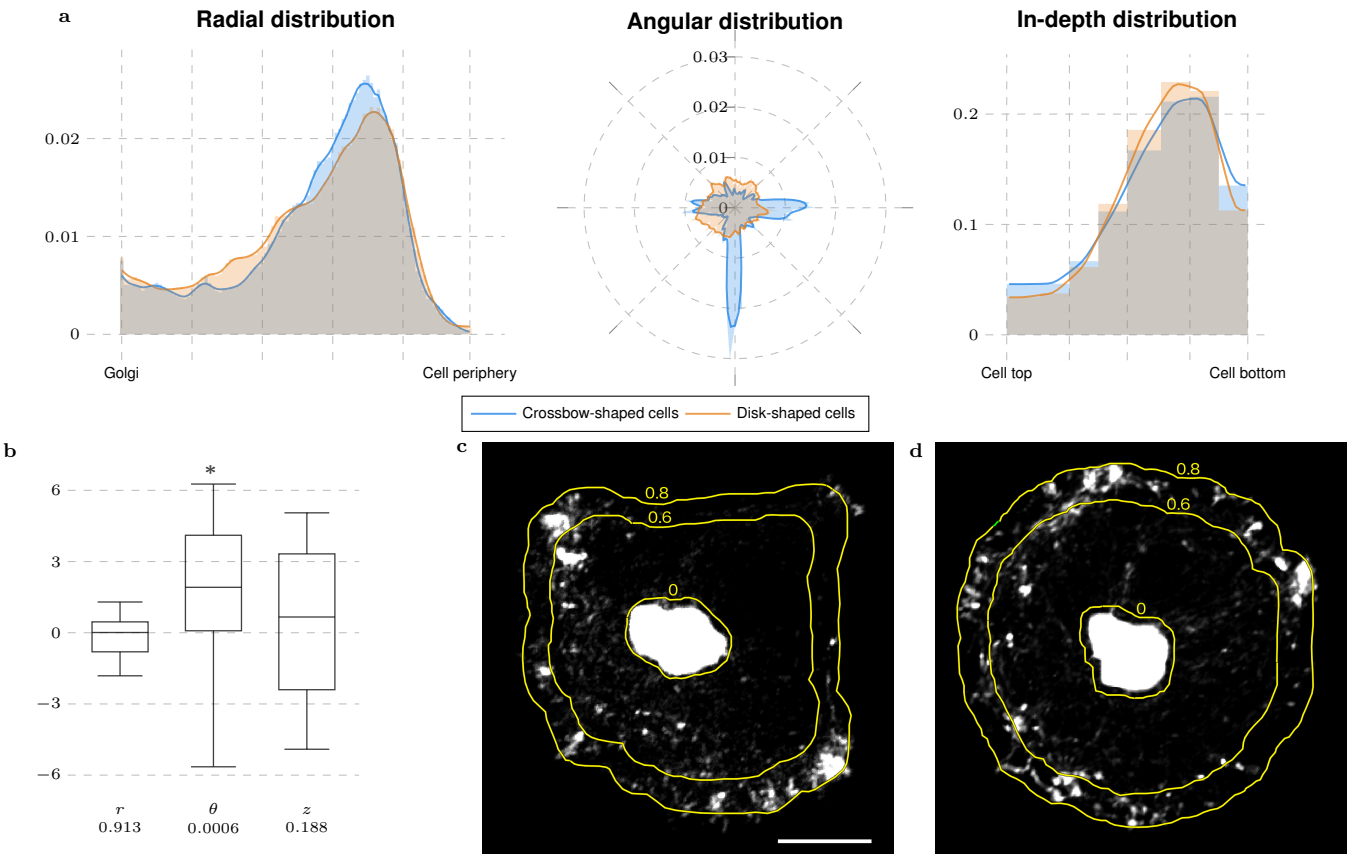

Fig. 2. Spatial distribution analysis of Rab6 proteins for crossbow- and disk-shaped cells.

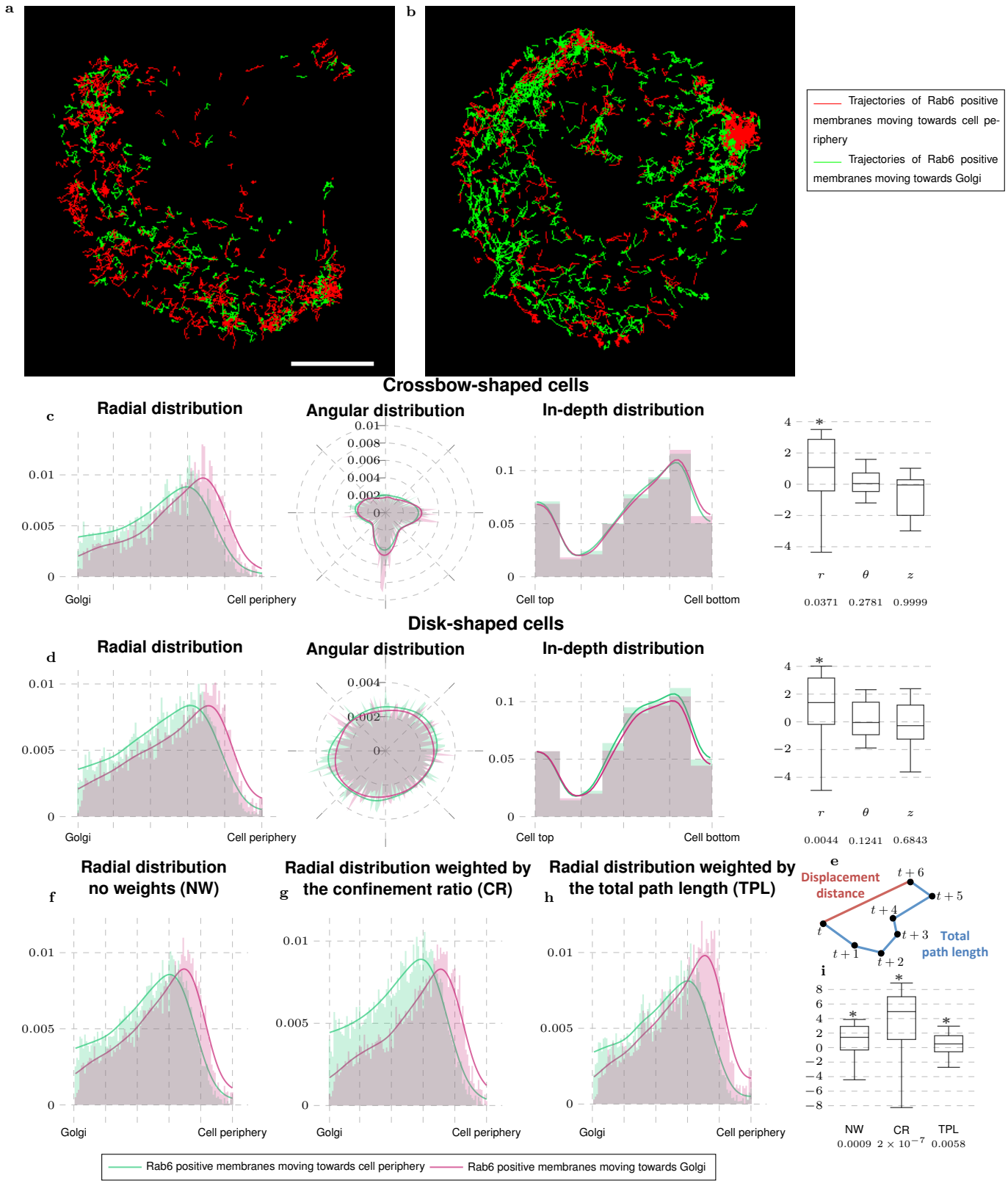

Fig. 3. Spatial distribution analysis of moving Rab6 proteins.

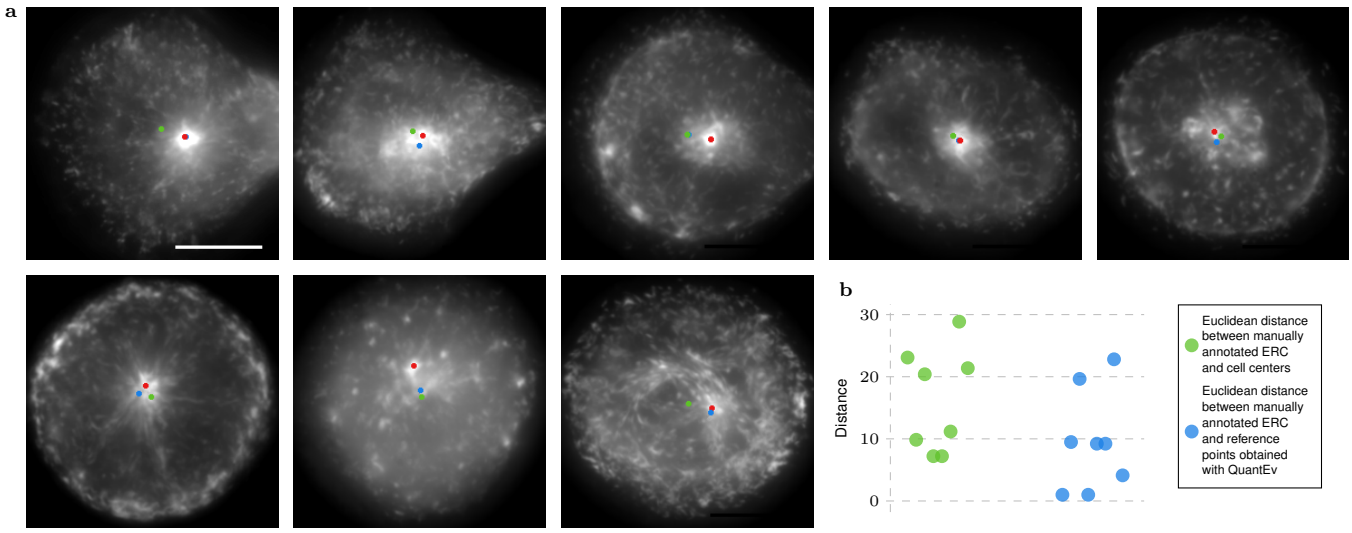

Fig. 4. Estimation of the Endosomal Recycling Center (ERC) location from the angular distribution of Rab11 positive membranes.

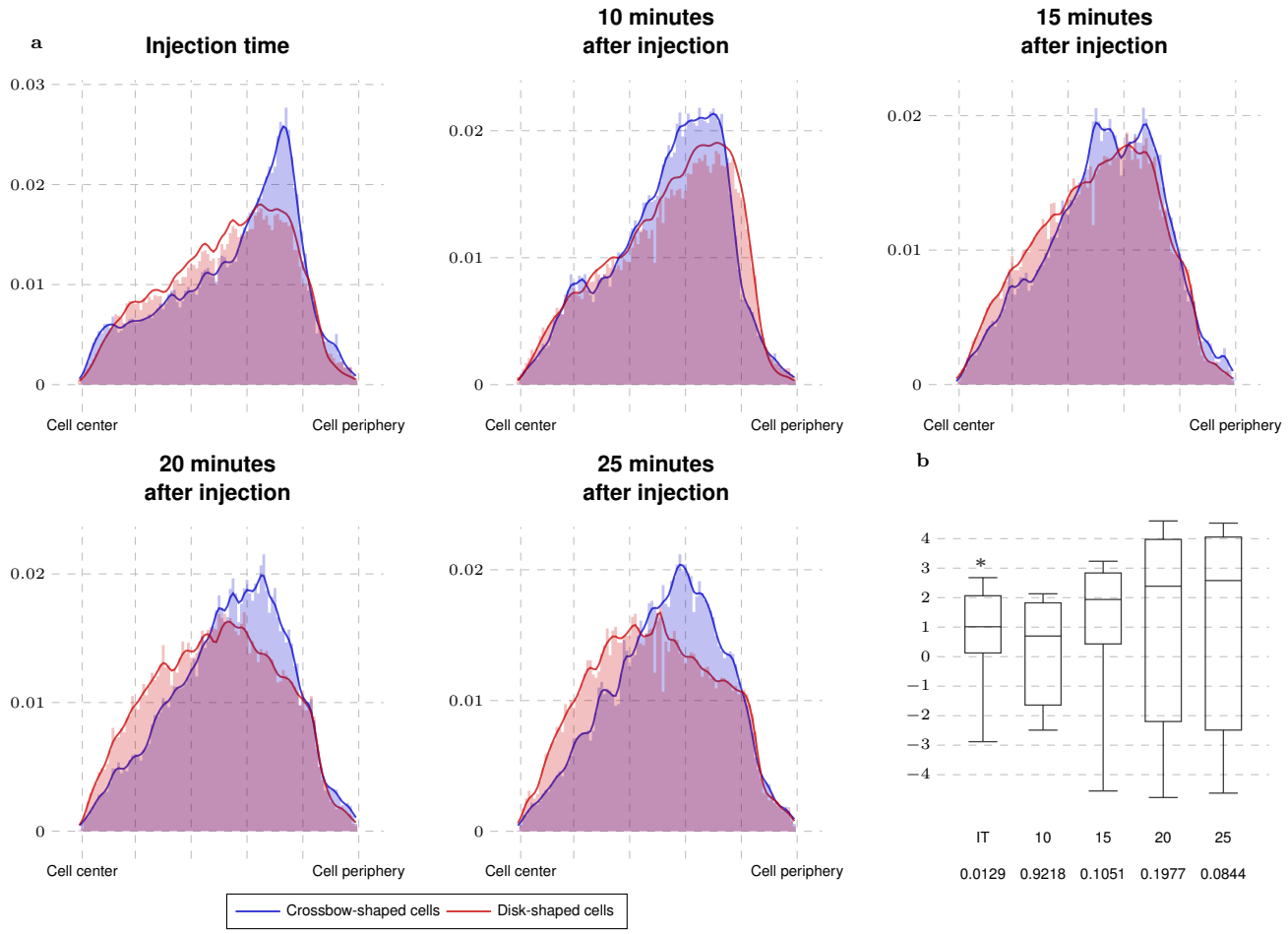

Fig. 5. Latrunculin A influence on Rab11 radial distribution.

Supporting Information

Datasets Fig. S1 shows fluorescence images taken from the different datasets used in the study.

Sensitivity to cell shape The cell shape influences the spatial distribution of intracellular events. The distances d_θ and d_z were introduced to compute a distribution that is invariant from the cell shape (see Section *Weighted density estimation*). To quantify the cell shape influence and to evaluate the pertinence of the normalization with distances, we generate image sequences with vesicles trafficking on a square-shaped region. In these simulations, vesicles are uniformly distributed over 16 different paths and are moving from the cell center to the cell periphery (see Fig. S6 **a**). As the cell is square-shaped, the vesicles moving to the cell corners travel a longer distance than the other vesicles so the number of vesicles on these paths is higher. Consequently, the spatial distribution of vesicles is not uniform over the radius and angle ranges (see purple histograms in Fig. S5 **a-b**). Nevertheless, the vesicles are generated over the paths with an equal probability in the simulations, meaning that the distribution over the different paths is uniform. By weighting the distribution of spatial coordinates with the distance between the cell center and the cell periphery, the shape dependence is accurately corrected as shown in the green histograms of Fig. S5 **a-b**.

Statistical analysis To evaluate the effect of pooling data together on the statistical analysis, 20 image sequences with uniform distribution over the paths (Fig. S6 **b**) and 10 image sequences with isotropic distribution over the paths (Fig. S6 **c**) are generated. 4 groups are then defined from these simulations:

- group #1: 10 image sequences with uniform distribution;
- group #2: 10 other image sequences with uniform distribution;
- group #3: 10 image sequences with isotropic distribution (6 paths with a probability equal to 0.1 and 10 paths with a probability equal to 0.04);
- group #4: 5 image sequences with uniform distribution and 5 image sequences with isotropic distribution.

The analysis of variance (ANOVA), usual method for biological studies, is compared to the QuantEv statistical approach. For the ANOVA analysis, the vesicle mass centers are extracted from the simulations and the pair (r, θ) is used to compare two groups. The intensity observed in the segmented vesicles is used for the QuantEv approach. For both methods, several amounts of data are considered: from 1% to 100% data for the ANOVA analysis; from 2 vs. 2 to 10 vs. 10 image sequences for QuantEv.

With the ANOVA analysis on pooled data, the *p*-values are low with a small amount of data when comparing groups #1 and #3 (see Fig. S5 **c**). But they also start to be low when comparing groups #1 and #2 for an amount of data that reaches about 50% (see Fig. S5 **c**). These results indicate that there is a gradient of *p*-values consistent with actual differences between the spatial distributions. However, the values lead to a significant difference between all groups (see Fig. S5 **c**). It demonstrates that it is difficult to deal with pooled data when the amount of data is high. When comparing groups #1 and #4, there should not be any statistical difference, as group #4 is

constituted of particles with different distributions. But the ANOVA analysis on the pooled data is not able to grasp this variability between replicated experiments of a same condition and the *p*-values are low with a small amount of data (about 5%, see Fig. S5 **c**).

The QuantEv statistical approach does not lead to any statistical difference for radius for the three comparisons (see Fig. S5 **d**), a result that is consistent with the data. By using QuantEv, it turns out that angular distributions are statistically different when comparing groups #1 and #3 while they are not for the two other comparisons (see Fig. S5 **d**). These experiments demonstrate that the QuantEv statistical approach is not disturbed by large amounts of data because it considers the distributions over the sequences. They also demonstrate that QuantEv takes into account the variability between replicated experiments of a same condition as the comparison involving groups #1 and #4 does not conclude to any statistical difference.

Uniform distribution of events To evaluate the QuantEv uniformity analysis, we simulate 10 image sequences with particles uniformly distributed over the different paths on a network for which the origin is not centered in the image (Fig. S6 **d**). Fig. S5 **g** shows the entropy map obtained for one simulation. Fig. S5 **h** shows the different reference points estimated over the ten simulations as green disks. These results are not perfect, as the reference centers are not estimated to be located at the exact particle emitter location. However, if the particles are distributed with equal probability on all paths, this does not imply that the actual number of generated particles is the same on all paths so the estimation cannot be perfect. The estimated reference points for these simulations are close to the particle emitter location, which demonstrates the potential of this approach.

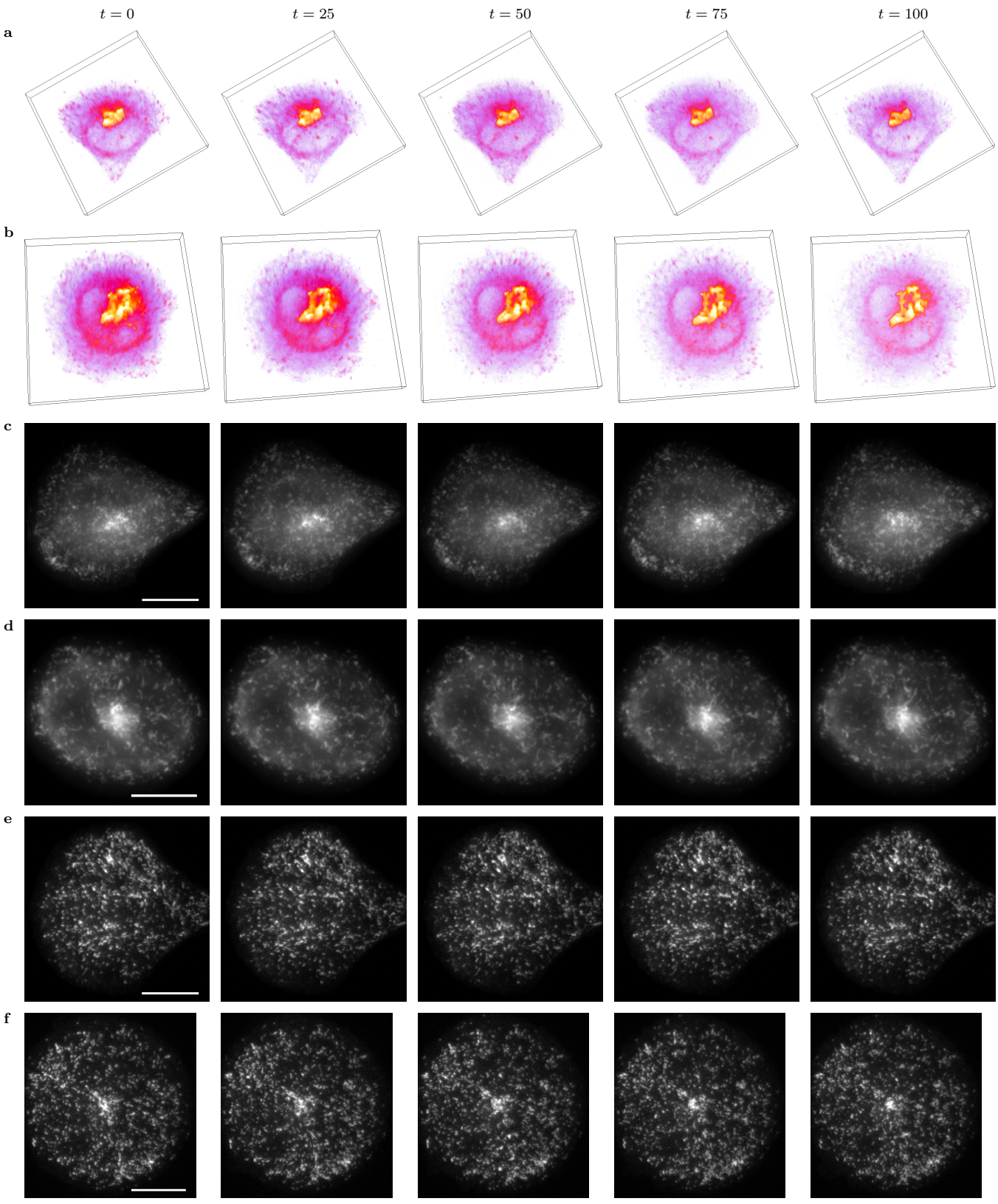

Fig. S1. Example images from the datasets.

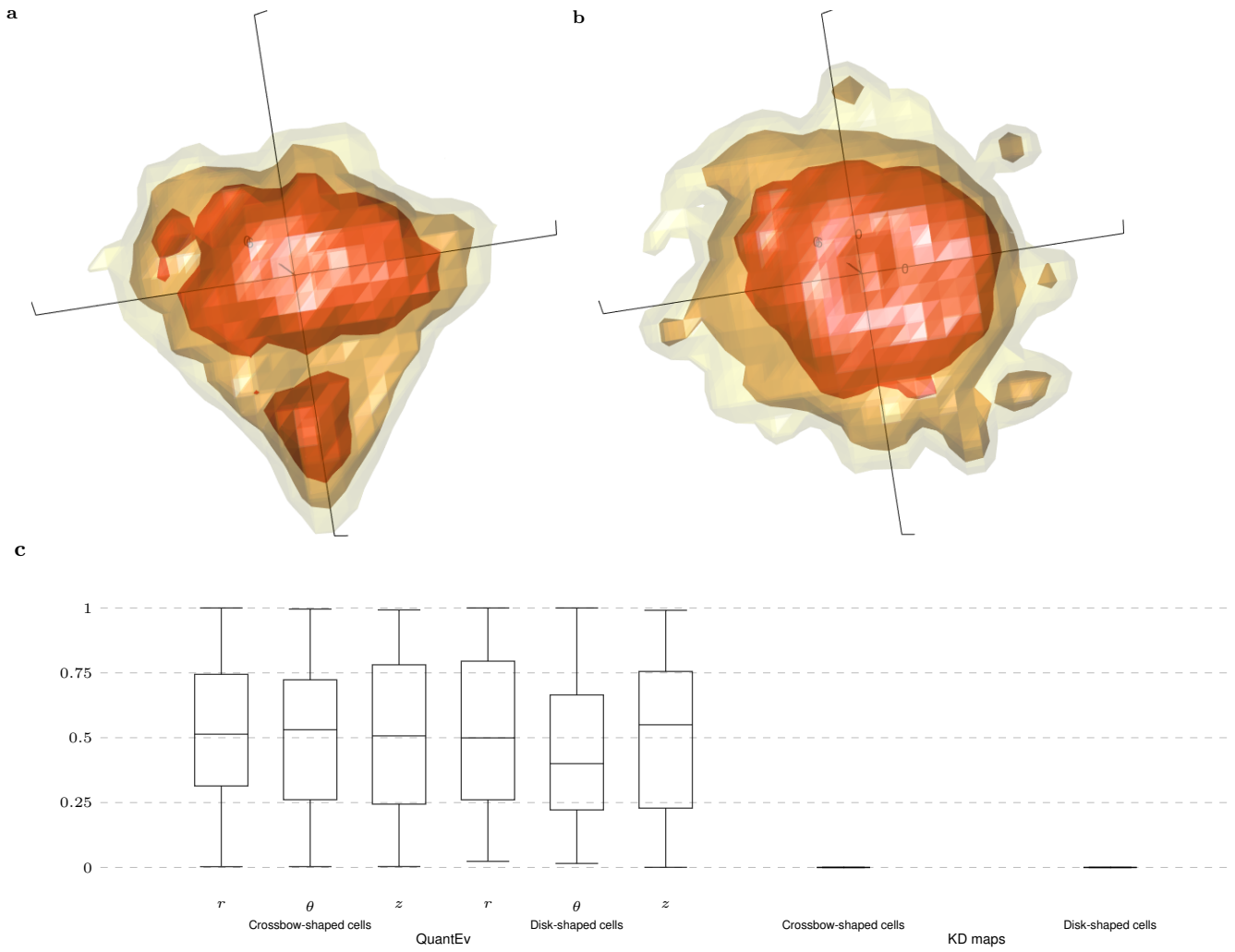

Fig. S2. Spatial distribution analysis of Rab6 proteins for crossbow- and disk-shaped cells with kernel density (KD) maps and QuantEv.

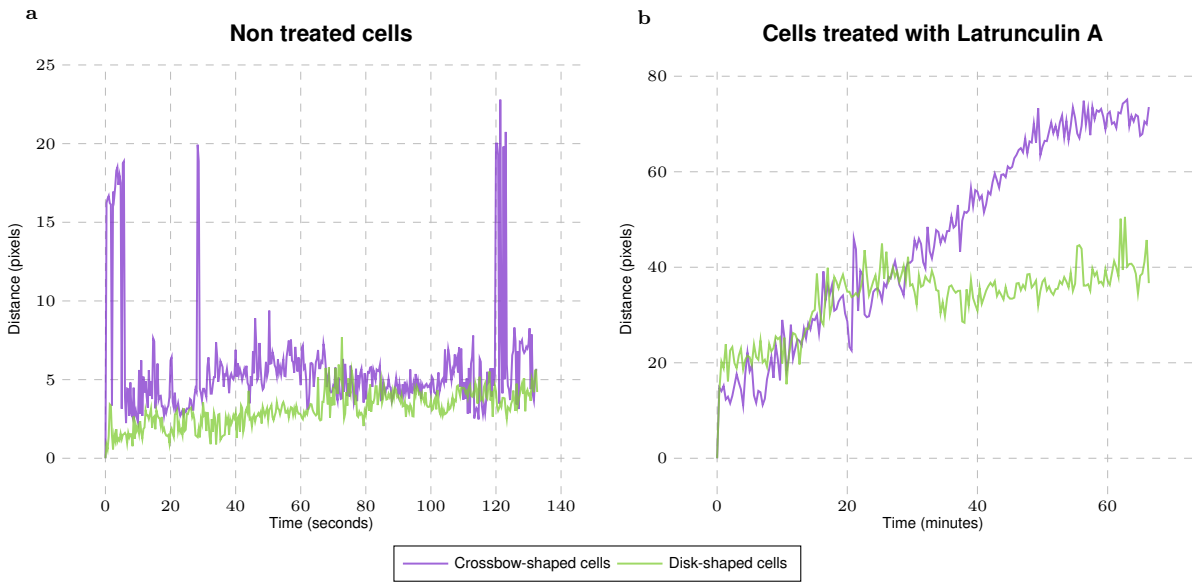

Fig. S3. Evolution of the point giving the most uniform angular distribution over time.

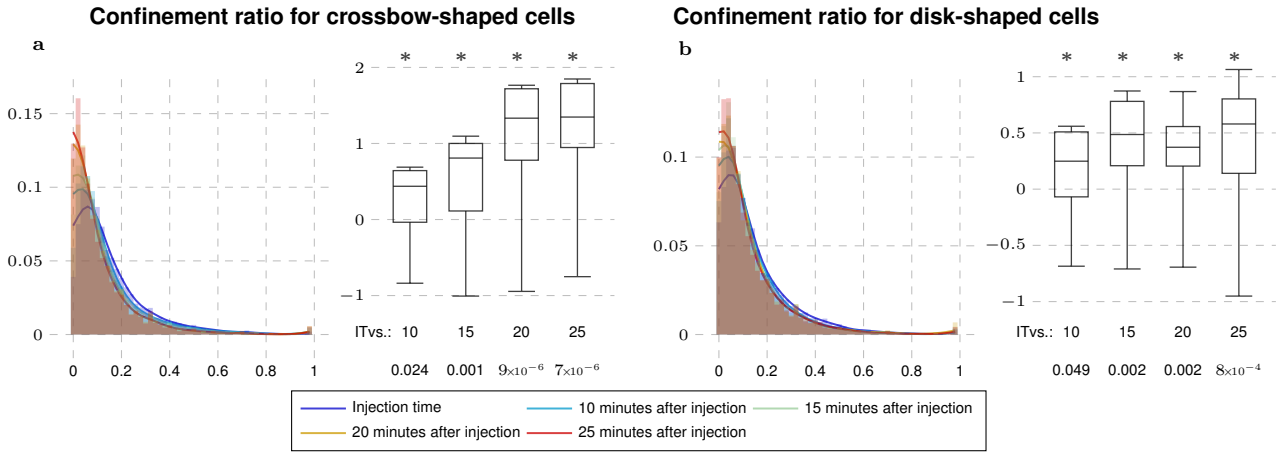

Fig. S4. Confinement ratio of Rab11 positive membranes with Latrunculin A injection.

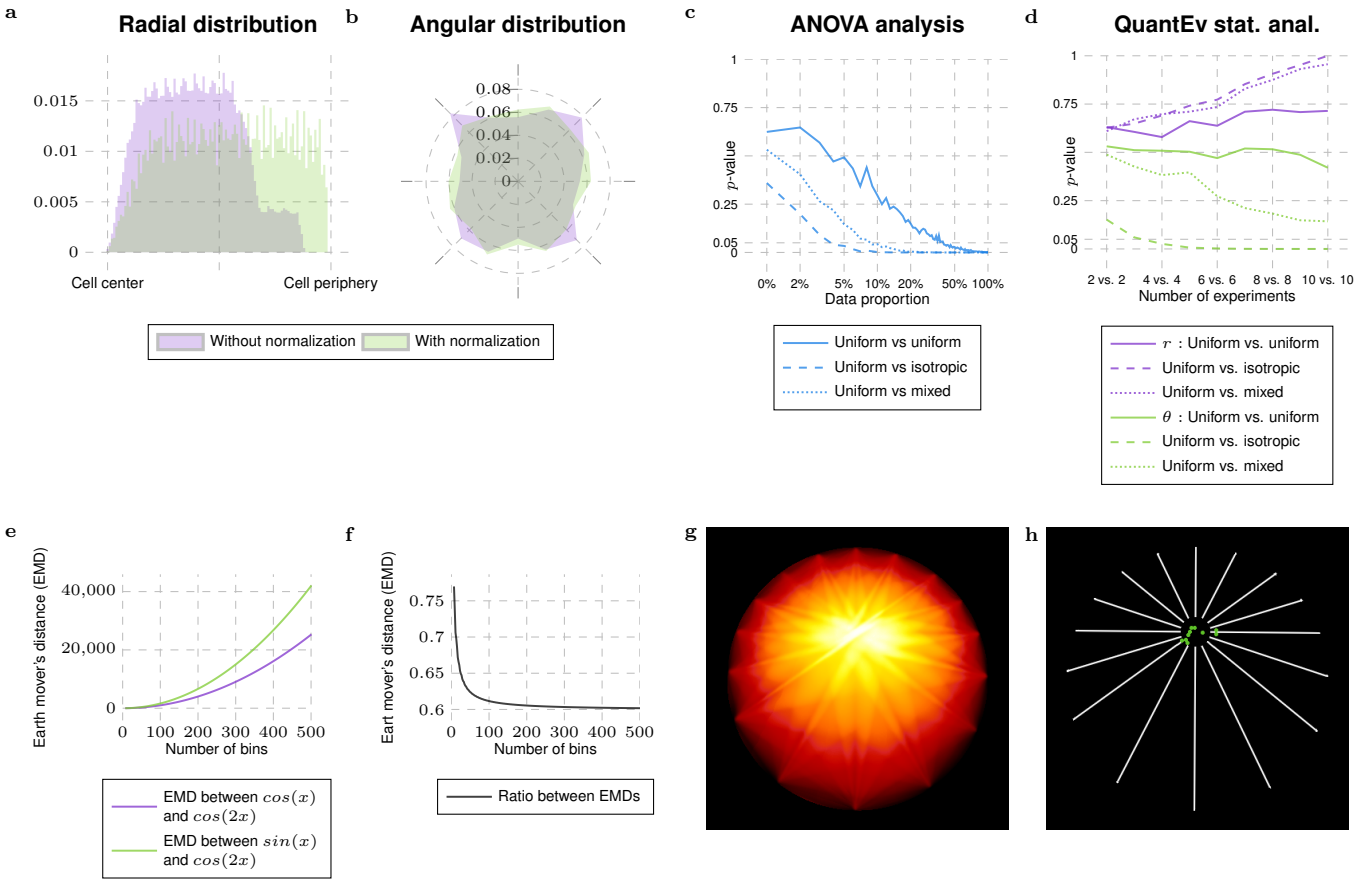

Fig. S5. Evaluation and tests on simulated data.

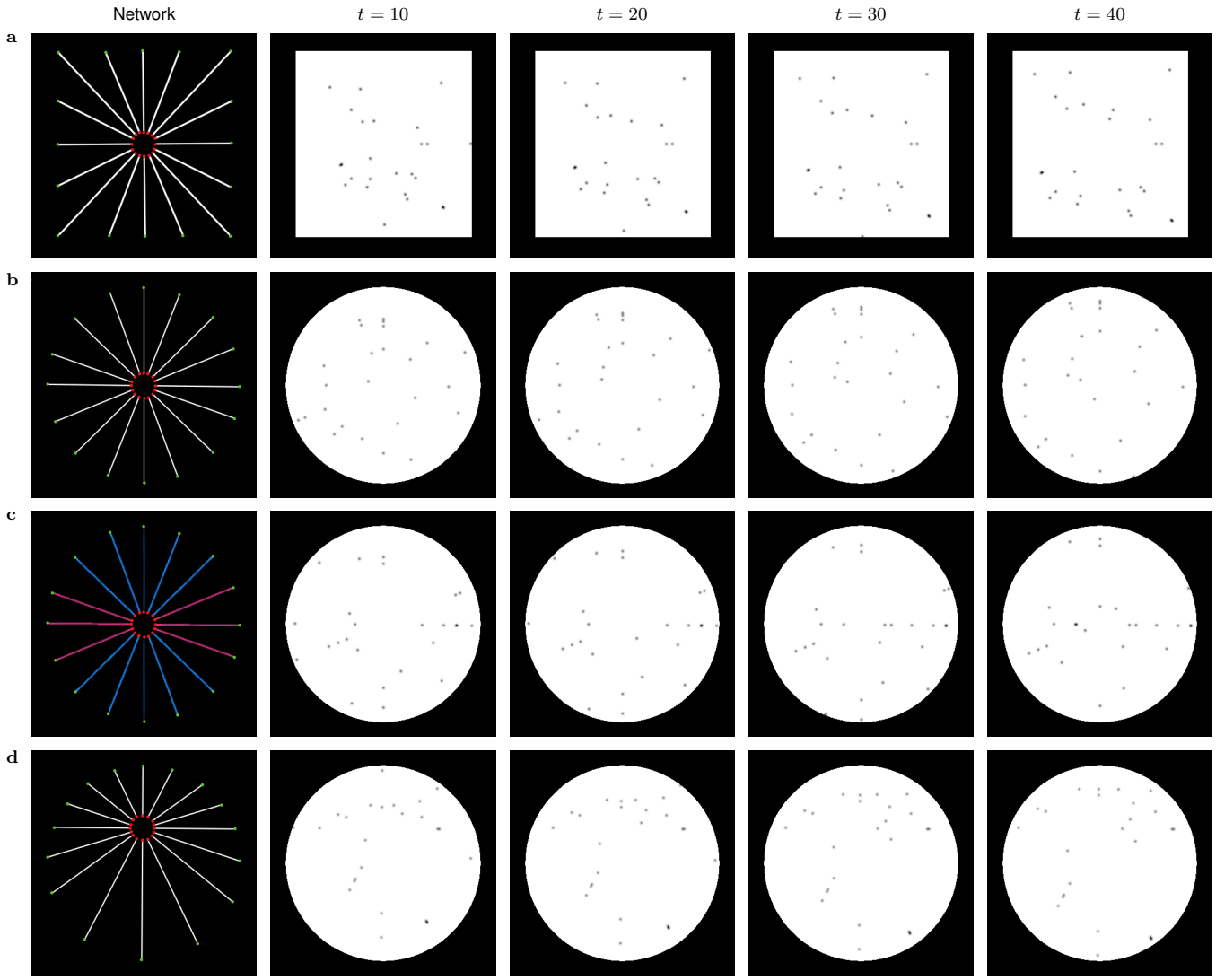

Fig. S6. Image sequences simulated to evaluate QuantEv performance.

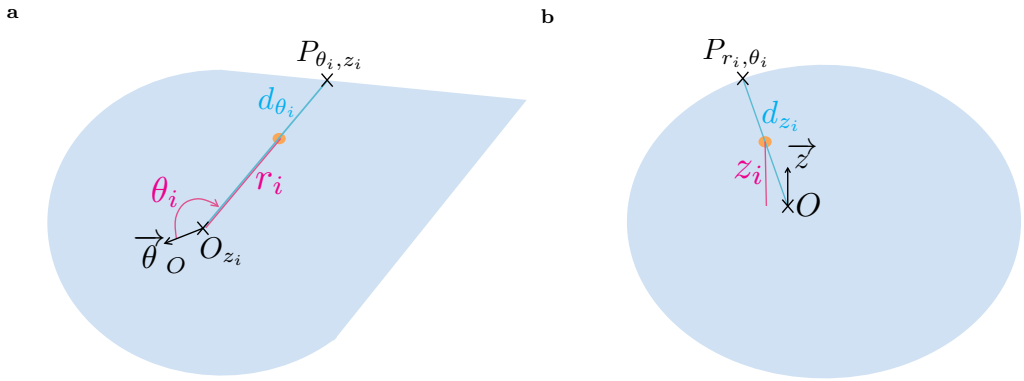

Fig. S7. Representation of the coordinates of an event.

Fig. S1. Example images from the datasets. a-b Volume renderings of fluorescent images taken from 2 sequences showing Rab6 proteins in a crossbow-shaped (**a**) and a disk-shaped cell (**b**). **c-d** Fluorescent images taken from 2 sequences showing Rab11 proteins in a crossbow-shaped (**c**) and a disk-shaped cell (**d**). **e-f** Fluorescent images taken from 2 sequences showing Rab11 proteins treated with Latrunculin in a crossbow-shaped (**e**) and a disk-shaped cell (**f**). In Figs. **c-f**, the intensity over the planes is averaged, a gamma correction is applied for a better visualization and the scale bars correspond to 5μm.

Fig. S2. Spatial distribution analysis of Rab6 proteins for crossbow- and disk-shaped cells with kernel density (KD) maps and QuantEv. a-b 3D KD maps obtained with kernel density maps[3] when considering all image sequences with crossbow- **a** and disk-shaped cells **b**. **c** Box and whisker plots of the p-values obtained when comparing randomly 100 times 2 groups of crossbow-shaped cells or disk-shaped cells with QuantEv and KD maps[3].

Fig. S3. Evolution of the point giving the most uniform angular distribution over time. Average Euclidean distance between the point giving the most uniform distibution at time $t = 0$ and the point estimated at further frames for normal cells **a** (3 image sequences for crossbow-shaped cells and 5 image sequences for disk-shaped cells) and cells treated with Latrunculin A **b** (6 image sequences for each micro-pattern).

Fig. S4. Confinement ratio of Rab11 positive membranes with Latrunculin A injection. Histograms (bar plots) and densities (lines) of the confinement ratio of Rab11 positive membranes on crossbow- (**a**) and disk-shaped cells (**b**) at Latrunculin A Injection Time (IT), 10, 15, 20 and 25 minutes after injection. Box and whisker plots of the corresponding condition differences (9 image sequences for each micro-pattern) are displayed on the side of the graphs. p-values under conditions of one-sided Wilcoxon signed-rank test when considering the condition differences are indicated below the box and whisker plots. A star (*) indicates that the p-value is smaller than 0.05.

Fig. S5. Evaluation and tests on simulated data. a-b p-values obtained with ANOVA statistical analysis (**a**) and QuantEv statistical analysis (**b**) when considering the spatial distribution of particles uniformly distributed (Fig. S6 **b**), isotropically distributed (Fig. S6 **c**) and a mix of uniformly and isotropically distributed particles over 16 paths. **c-d** Radial and angular distributions of particles moving on square-shaped cells from 10 simulated image sequences (Fig. S6 **a**) with (green histograms) and without normalization (purple histograms) with respect to the distance between the cell center and the cell periphery. **e** Eath Mover's Distance (EMD) between $cos(x)$ and $cos(2x)$ (purple curve) and between $sin(x)$ and $cos(2x)$ (green curve). **f** Ratio between the two EMDs shown on left plot. **g** Entropy map obtained on a simulated image sequence (see Fig. S6 **d**) showing at each point the angular distribution entropy obtained when considering this point as the reference. **h** Estimated reference points (green disks) obtained for 10 different simulated image sequences.

Fig. S6. Image sequences simulated to evaluate QuantEv performance. First column: networks used to generate 4 image sequences. The particle origins are labeled as red disks while particle destinations appear as green disks. Particles for the sequences **a**, **c** and **d** are uniformly distributed over the different paths. Particles for the sequence **c** are distributed with a probability equal to 0.1 over the pink paths and with a probability equal to 0.04 over the blue paths. Images corresponding to time $t = 10$, $t = 20$, $t = 30$ and $t = 40$ taken from one simulated image sequence for each network are illustrated in columns 2 to 5.

Fig. S7. Representation of the coordinates of an event. Representation of the coordinates of an event i depicted as an orange disk in the cylindrical coordinate system with origin O and reference direction $\vec{\theta}_O$ for z fixed (**a**) and along z axis (**b**).