

# A Novel Organelle Map Framework for High-Content Cell Morphology Analysis in High Throughput

Journal of Biomolecular Screening  
2014, Vol. 19(2) 317–324  
© 2013 Society for Laboratory  
Automation and Screening  
DOI: 10.1177/1087057113497399  
jbx.sagepub.com  


Kristine Schauer<sup>1</sup>, Jean-Philippe Grossier<sup>1</sup>, Tarn Duong<sup>1,\*</sup>, Violaine Chapuis<sup>2</sup>, Sébastien Degot<sup>2</sup>, Aurianne Lescure<sup>3</sup>, Elaine Del Nery<sup>3</sup>, and Bruno Goud<sup>1</sup>

## Abstract

A screening procedure was developed that takes advantage of the cellular normalization by micropatterning and a novel quantitative organelle mapping approach that allows unbiased and automated cell morphology comparison using black-box statistical testing. Micropatterns of extracellular matrix proteins force cells to adopt a reproducible shape and distribution of intracellular compartments avoiding strong cell-to-cell variation that is a major limitation of classical culture conditions. To detect changes in cell morphology induced by compound treatment, fluorescently labeled intracellular structures from several tens of micropatterned cells were transformed into probabilistic density maps. Then, the similarity or difference between two given density maps was quantified using statistical testing that evaluates differences directly from the data without additional analysis or any subjective decision. The versatility of this organelle mapping approach for different magnifications and its performance for different cell shapes has been assessed. Density-based analysis detected changes in cell morphology due to compound treatment in a small-scale proof-of-principle screen demonstrating its compatibility with high-throughput screening. This novel tool for high-content and high-throughput cellular phenotyping can potentially be used for a wide range of applications from drug screening to careful characterization of cellular processes.

## Keywords

density mapping, cell biology, lysosomes, Rab7

## Introduction

Many studies have highlighted the complexity of eukaryotic cells. To cope with the enormous number of assays that need to be performed to study the function of cellular proteins and to characterize cellular regulatory circuits, high-throughput approaches have been pushed forward in the past decade. High-throughput approaches monitor many cellular phenotypes under similar assay conditions and thus should allow a better understanding of cell behavior. However, because cultured cells often display strong morphological cell-to-cell variation,<sup>1</sup> high-throughput approaches have experienced some limitations. It has been shown that population size, local cell density, and cell position within a cellular islet determine cellular morphology and responsiveness.

A primary method to study cellular function is to examine cell morphology after a given manipulation such as treatment with chemical compounds or specific gene silencing by RNA interference. These techniques in conjunction with fluorescent microscopy are frequently exploited for functional analyses at the scale of high-throughput studies.<sup>2</sup> Currently, a variety of numerical features are extracted from images that are further analyzed with classification strategies.<sup>3,4</sup> Features represent

any measured property derived from the image, such as total/mean/standard deviation of fluorescence intensity, texture, Zernike shape descriptions, and so forth. Unfortunately, they often suffer from a lack of biologically meaningful, human

<sup>1</sup>Molecular Mechanisms of Intracellular Transport, Unité Mixte de Recherche 144 Centre National de la Recherche Scientifique/Institut Curie, Paris, France

<sup>2</sup>Cytoo Cell Architects Minatoc, Grenoble, France

<sup>3</sup>BioPhenics Platform, Institut Curie–Translational Research Department, Hôpital Saint Louis, Paris, France

\*Current address: Theoretical and Applied Statistics Laboratory (LSTA), University of Paris, Paris, France

Received Feb 8, 2013, and in revised form Apr 14, 2013; May 24, 2013; Jun 19, 2013. Accepted for publication Jun 20, 2013.

Supplementary material for this article is available on the *Journal of Biomolecular Screening* Web site at <http://jbx.sagepub.com/supplemental>.

## Corresponding Author:

Kristine Schauer, Molecular Mechanisms of Intracellular Transport, Unité Mixte de Recherche, 144 Centre National de la Recherche Scientifique/ Institut Curie, 75005 Paris, France.  
Email: [kristine.schauer@curie.fr](mailto:kristine.schauer@curie.fr)

interpretable measurements (e.g., texture or Zernike shape descriptions).

We have recently developed a computational approach that takes advantage of the cellular normalization by micropatterning and a novel quantitative organelle mapping approach that allows automated cell morphology comparison. Briefly, morphological cell-to-cell variations are first reduced by plating cells on adhesive microfabricated patterns of extracellular matrix proteins. This *in vitro*<sup>5</sup> confinement of cell shape mimics the restriction of space that cells experience in body tissues.<sup>6</sup> Then, images of fluorescently marked intracellular proteins are transformed into a cloud of coordinate points, and their spatial organization is rigorously quantified by probabilistic density mapping,<sup>7</sup> an applied mathematical method. Using this approach, we found that intracellular compartments are spatially well defined at the single-cell level in conditions of confined cell adhesion.<sup>7</sup> Because the organization of each tested trafficking compartment was highly reproducible, density mapping was employed to quantify subtle alterations of cellular morphologies.<sup>8</sup> To compare density maps that summarize multivariate data, such as spatial organization of intracellular structures, we have developed a fully automated test that estimates all parameters needed to compute the *p* value (i.e., the mean and variance parameters of the asymptotically normal null distribution) directly from the segmented data. This test does not require any subjective decisions and allowed the detection of subtle changes in cell organization induced by disruption of the cellular cytoskeleton.<sup>8</sup>

In the present work, the density-mapping approach was adjusted to high-throughput screening by implementing cell normalization, density map calculation, and comparison into a 96-well format. The framework was tested for its performance at different magnifications and for different cell shapes. Given that cell morphology can be monitored by probabilistic density maps, statistical significance of the similarity between two cellular morphologies was assessed in an automated and unbiased manner. Thus, density mapping is a novel tool for high-content and high-throughput cellular phenotyping that can be employed without a high level of statistical training.

## Materials and Methods

### Cells and Sample Preparation

An EGFP-Rab7-expressing stable cell line was generated by transfection of the plasmid pEGFP-Rab7 into hTERT RPE-1 cells (Invitrogen, Carlsbad, CA) and selection with 500 µg/mL geneticin. Cells were grown in DMEM/F12 (for RPE-1) or DMEM (for HeLa) medium supplemented with 10% fetal bovine serum and 1% penicillin/streptomycin. Seventy percent confluent cells were trypsinized and seeded on micropattern-printed 96-well CYTOOplates (CYTOO)

at a concentration of 4000 cells/well. To depolymerize microtubules, nocodazole was added in decreasing concentrations (10 µM, 1 µM, 0.1 µM, 10 nM, 1 nM, 0.1 nM), and cells were incubated at 37 °C for 1 h. Taxol, cytochalasin D, carbonyl cyanide *m*-chlorophenyl hydrazone (CCCP), and ibuprofen were used following the same procedure. All drugs were purchased from Sigma (St. Louis, MO). Cells were fixed with 4% (w/v) formaldehyde for 15 min and washed three times with phosphate-buffered saline (PBS). In the case of an immunofluorescence staining, formaldehyde-fixed cells were quenched with 0.05 M NH<sub>4</sub>Cl, washed three times with PBS, and permeabilized in PBS/0.2% bovine serum albumin/0.05% saponin. Cells were then incubated with primary antibody anti-Lamp1 (BD Pharmingen, Franklin Lakes, NJ) for 1 h, washed in PBS, and incubated with Alexa Fluor 488- or Cy3-coupled secondary antibodies (Jackson ImmunoResearch, West Grove, PA). After three washes with PBS, nuclei were stained with 0.2 µg/mL Dapi.

### Immunofluorescence Image Acquisition and Analysis

Images were acquired using the INCell2000 automated wide-field microscope (GE Healthcare) equipped for image deconvolution. Acquisition was performed using either 10× or 20× dry objectives. Images were sorted in ImageJ<sup>9</sup> using the micropatterned single-cell sorting (MSCS) plugin and segmented with the multidimensional image analysis interface running under MetaMorph (Universal Imaging Corporation, Bedford Hills, NY) based on wavelet decomposition.<sup>10</sup> The watershed function was routinely applied to detect precisely individual structures in dense regions.

### Kernel Density Estimator

The coordinates of the segmented structures were processed for density estimation programmed in the *ks* library<sup>11</sup> in the R programming language<sup>12</sup>: the probability density function *f* for each data sample of *n* coordinates  $X_1, X_2, \dots, X_n$  was estimated. A nonparametric, unbinned kernel density estimator was used. At each of the data points, a kernel function *K* was centered. The kernel functions were then summed to form the kernel density estimator  $\hat{f}$ :

$$\hat{f}_H(\mathbf{x}) = \frac{1}{n} \sum_{i=1}^n K_H(\mathbf{x} - \mathbf{X}_i)$$

in which  $K_H$  is the Gaussian kernel with mean zero and variance matrix *H*. To estimate *H* (also known as the bandwidth), the plug-in selector in the *ks* library was used, which has been shown to be reliable for two-dimensional and three-dimensional spatial distributions.<sup>13</sup> For visualizing kernel density estimates, probability contours were

employed.<sup>14,15</sup> For a 96-well plate, the analysis was adapted to compute the series of density maps in a batch mode.

### Statistical Error of Kernel Density Estimation

The statistical error for kernel density estimates was determined by calculating the estimated mean integrated squared error (MISE) as a function of the number of cells analyzed. For the combined sample of the  $n$  endomembrane coordinates from  $m$  cells, the average number of endomembranes per cell ( $n/m$ ) was computed. To compute the estimated MISE for  $j$  cells,  $j = 1, 2, \dots, m$ ,  $j \times (n/m)$  coordinates were resampled from the total of  $n$  coordinates. This resampling was performed 100 times for each  $j$  to ensure statistical robustness.

### Statistical Testing of Density Maps

The pairwise comparison of endosomal spatial distributions was carried out using the two-sample kernel density-based test introduced by Duong et al.,<sup>8</sup> in which  $X_1, X_2, \dots, X_{n_1}$  was the sample from the control density map  $f_1$  and  $Y_1, Y_2, \dots, Y_{n_2}$  was the sample from the treatment density map  $f_2$ . If the usual null hypothesis  $H_0: f_1 = f_2$  holds, the two density maps will be statistically the same. The test statistic was  $\hat{T} = \hat{\psi}_1 + \hat{\psi}_2 - (\hat{\psi}_{1,2} + \hat{\psi}_{2,1})$ , where

$$\hat{\psi}_1 = \frac{1}{n_1^2} \sum_{i=1}^{n_1} \sum_{i_2=1}^{n_1} K_{H_1}(X_i - X_{i_2}), \quad \hat{\psi}_2 = \frac{1}{n_2^2} \sum_{j=1}^{n_2} \sum_{j_2=1}^{n_2} K_{H_2}(Y_j - Y_{j_2})$$

$$\hat{\psi}_{1,2} = \frac{1}{n_1 n_2} \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} K_{H_1}(X_i - Y_j), \quad \hat{\psi}_{2,1} = \frac{1}{n_1 n_2} \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} K_{H_2}(X_i - Y_j).$$

where  $K$  is the kernel function and the bandwidth matrix parameters are  $H_1$  and  $H_2$ . Assuming that the null hypothesis holds  $H_0: f_1 = f_2 = f$ , then, as  $n_1, n_2 \rightarrow \infty$ , the test statistic  $\hat{T}$  follows an asymptotic normal distribution. More precisely,  $\frac{\hat{T} - \mu_T}{\sigma_T \sqrt{1/n_1 + 1/n_2}} \xrightarrow{d} N(0,1)$ , where  $\mu_T = [n_1^{-1} |H_1|^{-1/2} + n_2^{-1} |H_2|^{-1/2}] K(0)$  and  $\sigma_T^2 = 3 \left[ \int f(x)^3 dx - \left( \int f(x)^2 dx \right)^2 \right]$ . (Additional mathematical details are presented by Duong et al.<sup>8</sup>). This test statistic was interpreted as comparing intrasample pairwise differences  $X_{i_1} - X_{i_2}$  and  $Y_{j_1} - Y_{j_2}$  to the intersample pairwise differences  $X_{i_1} - X_{j_1}$ . If the latter are larger than the former, this indicates that the samples are different. The asymptotic normality allows for the easy computation of a  $p$  value. Note that this setup is for a single pair of density maps. To potentially compare all 96 wells against each other, the statistical tests were adopted to a batch mode.

## Results and Discussion

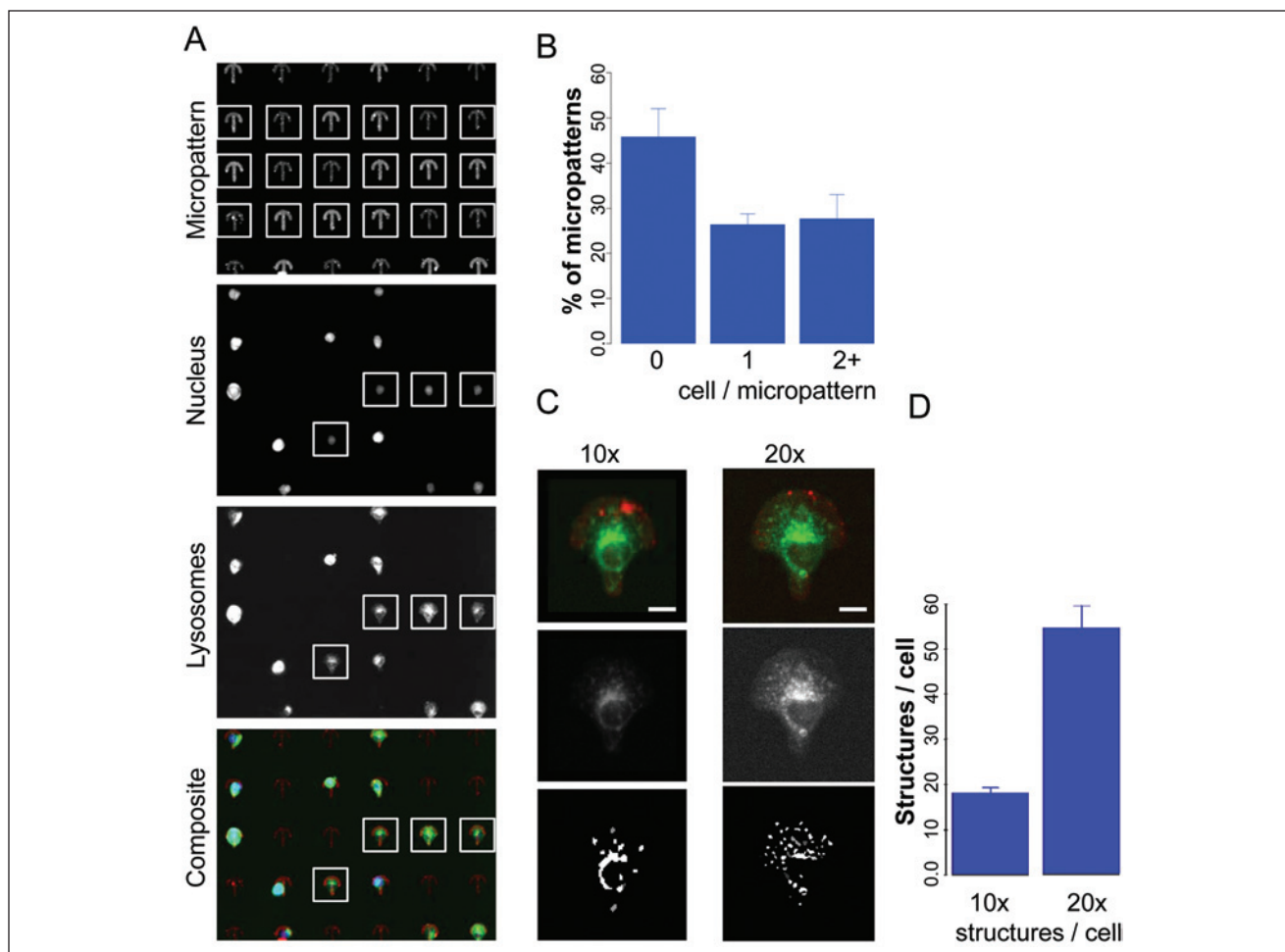
### Cell Plating on 96-Well Micropatterned Substrates and Single-Cell Selection

Cell seeding on micropatterned substrates does not provide complete coverage of all micropatterns, with one single cell required for analysis. Thus, the number of single

micropatterned cells in a 96 well was first estimated. Commercially available micropatterned 96-well plates (CYTOO) were obtained that contained an average of 4000 micropatterns/well according to the manufacturer's estimation. To determine the correlation between cell number and micropattern occupation per single cells, different numbers of cells per well were seeded (3200, 3500, 4000, 4500, or 4800 cells per well). RPE-1 cells that stably expressed EGFP-Rab7, a marker of lysosomes, were used to visualize intracellular compartments. Lysosomes are involved in several important cellular functions, including processing of nutrients, ligands, and receptors during endocytosis that are transported along microtubules. Images were acquired at 10 $\times$  and 20 $\times$  for up to four fluorescence channels, including fluorescence staining of micropatterns (Cy5), fluorescence staining of the nucleus (Dapi), and fluorescence stainings for intracellular structures of interest (green fluorescent protein [GFP]; **Fig. 1A**). A plugin was programmed in ImageJ (MSCS) to select for patterns that contained only one nucleus and thus one single cell. Briefly, after the detection of each full pattern (**Fig. 1A**, upper panel), the same region was cropped for the Dapi channel. Segmentation analysis of the nucleus sorted the images into three categories: no cell per pattern, one cell per pattern, and more than one cell per pattern. Significant differences in the average number of single cells were not detected when varying cell numbers (data not shown). Thus, 4000 cells per well were used for all following experiments. Based on the analysis of 1,199,973 single patterns detected in more than 20 plates, approximately 25% of micropatterns were occupied by a single cell in independent seeding assays. **Figure 1B** shows the statistics from one typical experiment using nine plates on which a total of 514,571 patterns were detected containing 135,916 single cells (26.4% single-cell occupancy). After the identification of single cells on patterns, the corresponding images visualizing the intracellular compartments were cropped using the pattern as reference (**Fig. 1A**). As a result, stacks of images of single aligned cells were obtained for a given intracellular structure (e.g., lysosomes). Positional information of all structures of interest was extracted by segmentation analysis in Metamorph (Universal Imaging Corporation). Approximately 16 structures per cell were detected at 10 $\times$  magnification and approximately 55 structures per cell at 20 $\times$  magnification (**Fig. 1C, D**).

### Implementation of Cell Normalization and Density Map Calculation into a 96-Well Setup

To set up the overall workflow for phenotypic screening using the organelle mapping approach, the 96-well CYTOOplate starters was first employed, which contains five different micropattern shapes (disc-, crossbow-, I-, Y-, and L-shaped micropatterns) in two different sizes of 700  $\mu\text{m}^2$



**Figure 1.** Single-cell selection on the 96-well micropatterned support. **(A)** Fluorescence images of crossbow-shaped micropatterns as well as nuclei and lysosomes from corresponding crossbow-shaped micropatterned cells imaged at 20 $\times$ . The composite image shows micropatterns in red, the nuclei in blue, and lysosomes in green. Micropatterns/cells that are framed were selected for further analysis: all micropatterns were selected whose images were complete; cells were selected only if they contained one nucleus. **(B)** Analysis of cell distribution on micropatterns. Results are shown in percentages and bars represent the SD from nine independent plates on which a total of 514,571 patterns were detected containing 135,916 single cells. **(C)** Composite images of one crossbow-shaped micropattern in red and lysosomes from one single cell in green and corresponding fluorescence images and segmented images of endosomes at 10 $\times$  and 20 $\times$  image magnification. Scale bars are 10  $\mu$ m. **(D)** Analysis of the number of structures detected by segmentation in single cells at 10 $\times$  and 20 $\times$  magnification. Bars represent SD from 2687 (10 $\times$ ) or 5332 (20 $\times$ ) single cells.

and 1100  $\mu$ m<sup>2</sup> (Suppl. Fig. 1A). Columns 1 and 12 of the 96-well plate corresponded to control wells (homogeneously coated/no micropattern) and were not analyzed. To simplify the overall process, RPE-1 cells that stably expressed EGFP-Rab7 were added to 96 wells (4000 cells per well), fixed, imaged at 20 $\times$ , sorted, and segmented. All coordinates of Rab7-marked lysosomes from all single cells in one well were aligned and plotted on the 96-well plate format (Suppl. Fig. 1B). The aligned coordinates were used to calculate the average density maps in each well according to Schauer et al.<sup>7</sup> in the R software<sup>12</sup> in a batch mode. For each well, Gaussian functions (kernels) with mean zero and an optimized variance were centered at each of the data points for all cells and summed, revealing the underlying density throughout the cell that was visualized by the 25%, 50%, and 75%

probability contours (Suppl. Fig. 1C). Contours represent the smallest area in which a given percentage of endomembranes are found: for example, the 50% probability contour defines the smallest area in which 50% of endomembranes are located. A 96-well plate map for visualization was programmed that reveals for each well information about the well number, the treatment condition, and an additional parameter such as micropattern shapes and the statistics of the analyzed well, including identified single-cell and intracellular structure numbers (Suppl. Fig. 1D). This interface allows direct and intuitive visualization of subcellular changes of the analyzed intracellular compartment. Then, wells were pairwise compared using a density-based test.<sup>8</sup> This test computes a *p* value to evaluate the statistical difference of two density maps in a robust, automatic, and unbiased manner



that allows its usage in high-throughput approaches in which user intervention needs to be minimal. For a 96-well plate, there are potentially many possible pairwise comparisons; thus, the statistical tests were adopted to a batch mode.

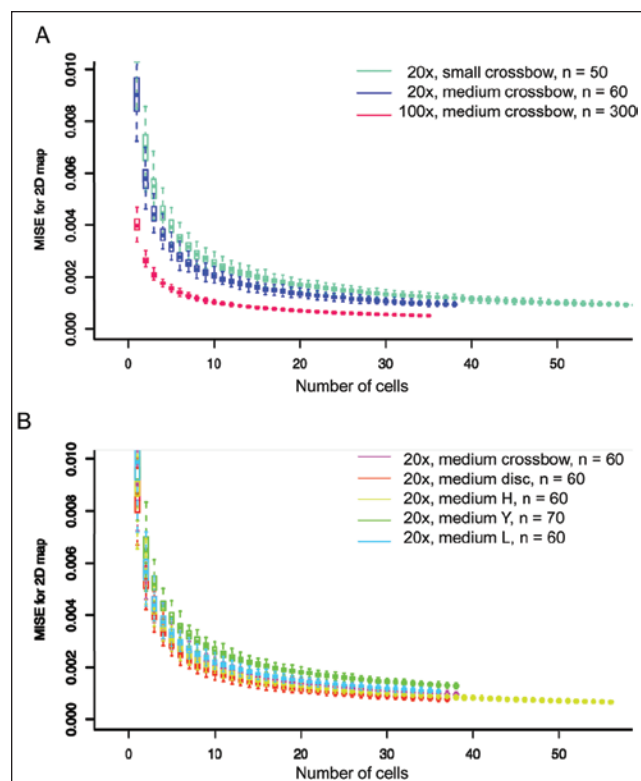
Together, the organelle map framework was adjusted to a high-throughput analysis that allows biological scientists even without a high level of statistical training to identify hits following a user-friendly workflow. It requires cell alignment, segmentation, density calculation, and pairwise statistical testing that are all programmed into a batch mode for the 96-well format to streamline the statistical analysis pipeline.

### Cell Number Estimation for Stable Density Maps at Low Magnification and Analysis of Different Pattern Shapes

We have previously shown that 30 to 40 cells were sufficient to obtain stable density maps of different intracellular compartments imaged at 100 $\times$  (Fig. 2A, pink).<sup>7</sup> As the total number of structures is important for the reliability of density maps and thus determines the number of cells required for analysis, the reliability of density maps was estimated for an imaging magnification of 20 $\times$ , which is typically used for high-throughput screening. The MISE of the density estimation of late endosomes was computed in cells grown on small- (700  $\mu\text{m}^2$ , light blue) and medium- (1100  $\mu\text{m}^2$ , dark blue) sized micropatterns as a function of the number of cells analyzed (Fig. 2A) and compared with the MISE obtained for cells imaged at 100 $\times$ , as shown previously.<sup>7</sup> For less than 10 cell, the box plots of estimated MISEs showed a large variability, with significant improvement in estimated MISE when adding data for each subsequent cell. The estimated MISE only marginally diminished between 20 and 30 cells and did virtually not change thereafter. Density maps from data acquired at 20 $\times$  were more heterogeneous than those from 100 $\times$  acquisition (pink) for small numbers of cells; however, all density maps became stable for more than 30 cells. This demonstrated that using density estimation with micropatterned cells at a magnification of 20 $\times$  is robust from only several tens of cells for both sizes of micropattern.

In addition, the impact of cell shape on density mapping was analyzed. The MISE was computed for the density maps of lysosomes in cells adhering to the five different micropattern shapes present on the CYTOO plate starter (Fig. 2B). Although the MISE behaved very similarly for different shapes, round patterns gave the most stable density maps (red), followed by H-shaped (yellow), crossbow-shaped (pink), L-shaped, and Y-shaped patterns. This analysis showed that the choice of the shape is not critical for density-based analysis.

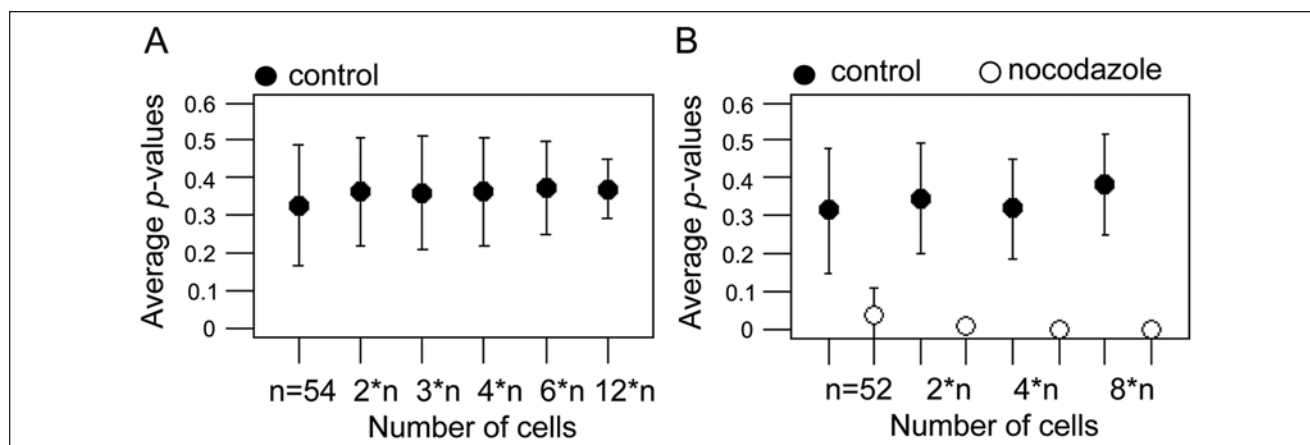
Together, density-based image analysis performed comparably well between high- and low-magnification images and is adaptable to different micropattern shapes.



**Figure 2.** Estimated mean integrated squared error (MISE) for density estimation as a function of the number of cells analyzed. (A) MISE of density maps of lysosomes as a function of the number of cells analyzed with  $n$  average structures per cell for cells grown on crossbow-shaped patterns of different sizes and different imaging magnification. Small micropatterned cells were 700  $\mu\text{m}^2$  (pattern diameter 26.5  $\mu\text{m}$ ), and medium-sized micropatterned cells were 1100  $\mu\text{m}^2$  (pattern diameter 33  $\mu\text{m}$ ); both were imaged at 20 $\times$ . Big micropatterned cells were 1300  $\mu\text{m}^2$  (pattern diameter 36  $\mu\text{m}$ ) and were imaged at 100 $\times$ .<sup>7</sup> (B) MISE of density maps of lysosomes as a function of the number of cells analyzed for cells grown on medium-sized micropatterns of different shapes.

### Unbiased Comparison of Cell Morphology

To determine assay robustness, a series of quality control analyses was performed in 96-well plates that contained the same micropattern shape (L-form). Again, 4000 EGFP-Rab7-expressing RPE-1 cells were seeded per well. The first control plate contained DMSO in one half of the plate and no treatment in the other half of the plate (Suppl. Fig. 2A). To assess the reproducibility of density maps between DMSO-treated and nontreated conditions, 20 fields per well at 20 $\times$  were acquired, resulting in the detection of 23,038 micropatterns, with a total of 5184 single cells detected (about 54 cells per well). Density maps for each condition were obtained (Suppl. Fig. 2B), each DMSO-treated well was compared with all nontreated wells of the same row using our statistical test,<sup>8</sup> and the average  $p$  value was calculated (Fig. 3A). To study how the  $p$  value behaved for



**Figure 3.** Control analysis of cell morphology comparisons. **(A)** Average  $p$  values of the comparisons between nontreated cells and DMSO-treated cells grown on L-shaped micropatterns for increasing numbers of analyzed cells (see also **Suppl. Fig. 2**). Nontreated cells from 1, 2, 3, 4, 6, or 12 wells were compared to corresponding DMSO-treated cells from 1, 2, 3, 4, 6, or 12 wells. Bars represent the SD of at least 16 pairwise comparisons (see also **Suppl. Fig. 2A**). **(B)** Average  $p$  values of the comparison between nontreated cells and DMSO-treated cells (closed circle) as well as nontreated cells and nocodazole-treated cells (open circle) grown on L-shaped micropatterns for increasing numbers of analyzed cells (see also **Suppl. Fig. 3**). Nontreated cells from 1, 2, 4, or 8 wells were compared with either corresponding DMSO-treated cells from 1, 2, 4, or 8 wells or corresponding nocodazole-treated cells from 1, 2, 4, or 8 wells. Bars represent the SD of at least four pairwise comparisons (see also **Suppl. Fig. 3**).

increasing numbers of cells, 2, 3, 4, 6, or 12 wells were pooled (**Suppl. Fig. 2B**). Then, DMSO-treated wells were compared pairwise with all corresponding pooled nontreated wells, and the average  $p$  values were calculated. As expected, the average  $p$  values were  $p > 0.05$  for all well numbers (and cell numbers) analyzed, indicating that there was no significant difference in lysosomal positioning between nontreated and DMSO-treated cells, even for as little as several tens of cells (**Suppl. Fig. 2C**). The corresponding box plots are shown in **Supplementary Figure 2D**.

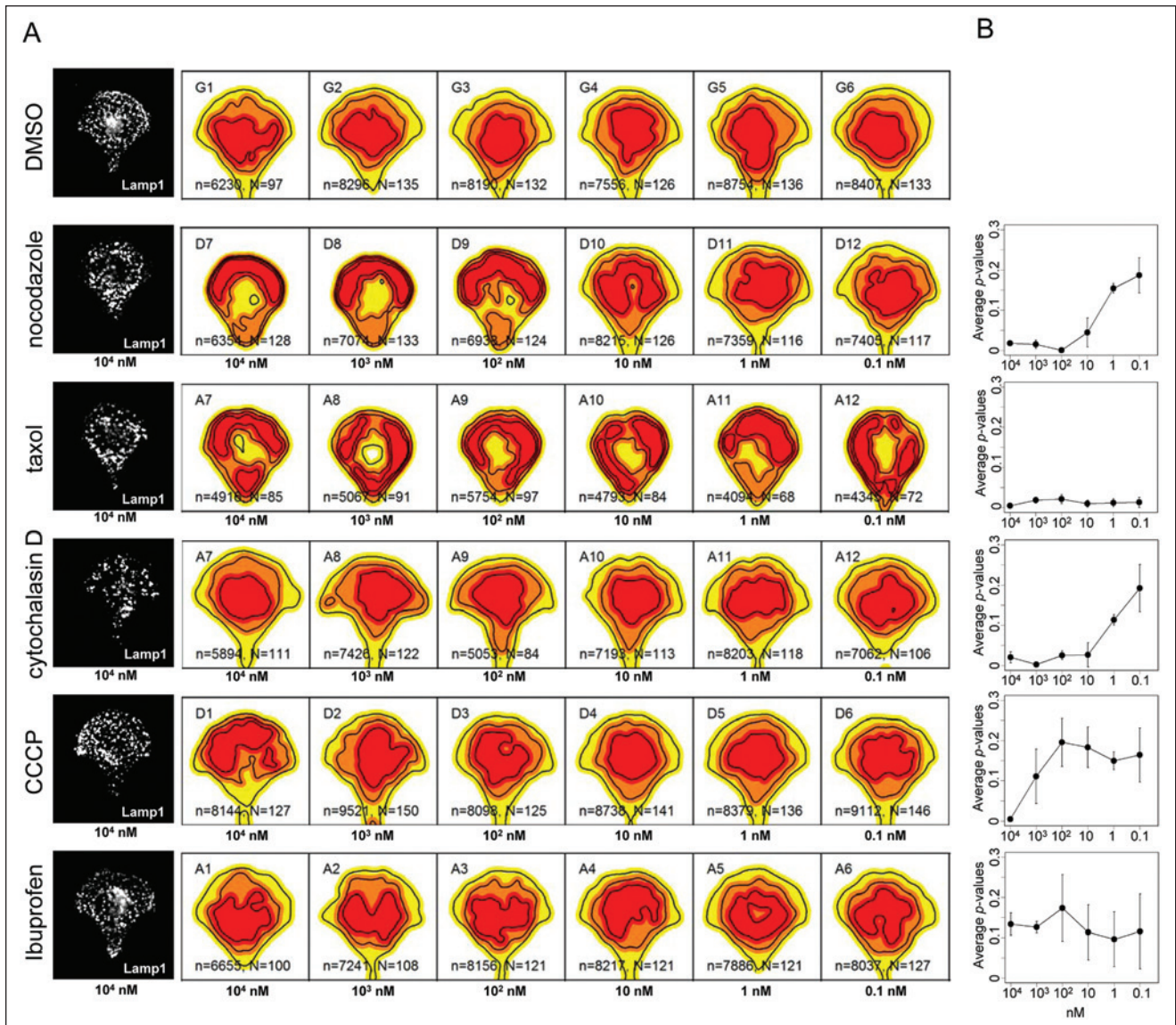
Next, the method's potential to detect changes in a screen setup was tested by monitoring lysosome positioning upon depolymerization of the microtubule cytoskeleton. A second control plate was prepared containing no treatment, DMSO treatment, and nocodazole treatment in two separated columns (**Suppl. Fig. 3A**). Cells were treated with 10  $\mu$ M nocodazole, a concentration that did not alter cellular shape as judged by representative images of single cells and average intensity projections of all cells from one nontreated well or one nocodazole-treated well (**Suppl. Fig. 3B**). Using the same image-processing settings as for the first plate, we obtained about 52 cells per well. Again, cells were either analyzed from single wells or from pooled wells (2, 4, and 8 wells), which increased the number of analyzed cells. Density maps were obtained for each condition (**Suppl. Fig. 3C**), and the average  $p$  values were calculated for all possible combinations of comparison (**Suppl. Fig. 3D**). When nontreated and DMSO-treated cells were compared, the average  $p$  values were  $p > 0.05$ , indicating that there is no significant difference in lysosomal positioning even for as little as several tens of cells (**Fig. 3B** and **Suppl. Fig. 3D**).

When nontreated and nocodazole-treated cells were compared, average  $p$  values of  $p < 0.05$  were obtained in all tested conditions (**Fig. 3B** and **Suppl. Fig. 3D**). The corresponding box plots are shown in **Supplementary Figure 3E**. Increasing cell numbers led to lower  $p$  values for nocodazole-treated cells, indicating that increasing cell numbers was advantageous. Sources of false-positives and false-negatives at low cell numbers may be noncomplete spreading of cells on micropatterns, as cell spreading was not accessed under our experimental conditions.

Together, the results of this analysis demonstrated that changes in intracellular distribution of lysosomes due to microtubule disruption could be detected with statistical significance in a screening setup with several tens of cells.

### Detection of Morphological Changes Induced by Different Drug Treatments

To further demonstrate the capability of this framework to detect changes induced by different chemical compounds, several well-characterized drugs were tested in different dilutions. For this analysis, HeLa cells were used that were seeded into a 96-well containing crossbow-shaped micropatterns and were immunostained with  $\alpha$ -Lamp1 antibody to detect lysosomes. Forty-two fields per well were acquired, giving rise to about 150 single cells per well. A decision was made to oversample, because HeLa cell morphology is more heterogeneous than that of RPE-1 cells<sup>7</sup> (e.g., the organization of the Golgi apparatus can be fragmented or not), potentially requiring more cells for a robust analysis. In addition to nocodazole, other drugs were tested, including taxol, which stabilizes microtubules; cytochalasin D,



**Figure 4.** Detection of morphological changes after drug treatment in a 96-well setup. **(A)** Fluorescence images and corresponding density maps of lysosomes at the indicated drug concentrations. Fluorescence images show a representative image of lysosomes at 10  $\mu\text{M}$  drug concentrations that did not alter cell shape. Density maps were calculated using all single cells in one well containing  $N$  cells and  $n$  structures. The color of probability contours represents the smallest regions in which 25% (red), 50% (orange), and 75% (yellow) of lysosomes are found. Scale bars are 10  $\mu\text{m}$ . **(B)** Average  $p$  values of the comparison between DMSO-treated cells and drug-treated cells at different drug concentrations. Bars represent the SD of well triplicates (see also **Suppl. Fig. 4**).

which inhibits actin polymerization; and CCCP, which inhibits the proton motive force as well as ibuprofen as negative control. Each drug was tested in dose response with concentrations ranging from 10  $\mu\text{M}$  to 0.1 nM (1:10 serial dilutions), and each drug concentration was tested in triplicate (**Suppl. Fig. 4**). Each of the wells containing drugs was compared with 12 DMSO-treated control wells, and the corresponding  $p$  values were calculated. Changes in lysosomal positioning were clearly detected by density maps (**Fig. 4A**) and the statistics of the black box test (**Fig. 4B**). This was observed for all tested drugs except ibuprofen, for

which lysosomal distribution was not altered even at the highest concentration (10  $\mu\text{M}$ ). The average  $p$  values of triplicates demonstrated that taxol treatment was detected at all tested concentrations ( $p < 0.05$ ) and that treatment with nocodazole and cytochalasin D was detected with  $p < 0.05$  at as low as 1 nM of drug concentration. Treatment with CCCP gave rise to significant changes in lysosomal positioning only for 10  $\mu\text{M}$ . Thus, a concentration-dependent increase in  $p$  values was found for several drugs tested. Furthermore, density maps revealed that changes in lysosomal positioning were similar upon microtubule disruption



or stabilization (treatment with nocodazole or taxol, respectively) but differed upon interfering with the actin cytoskeleton or the energy metabolism of the cell (treatment with cytochalasin D or CCCP, respectively). This direct and intuitive visualization of changes in the distribution of intracellular structures is useful to assign functional roles to tested molecules. To further validate our approach for screening, small known compound libraries will be tested to assess the robustness of the framework and/or the hit rate.

One limitation in applying our approach for very high-throughput screening (e.g., genomewide screening or big drug libraries) is the large amount of acquired images that is required to detect sufficient single cells as they represent only a quarter of all imaged micropatterns. To solve this problem, smart microscopes will be required that perform a preimaging at low resolution and specifically acquire regions with many single cells. Optimally, only well-spread cells should be selected that reduce the number of analyzed cells to about 30, as indicated by the MISE estimation. Currently, our approach is suitable for screens with a lower amount of conditions to test, such as well-defined drug libraries, siRNA libraries targeting specific gene families, or in secondary screens.

Together, density-based analysis was tested in a screening setup, in which changes in lysosomal positioning were detected for several drugs tested. Density-based screening may inherit several advantages of our probabilistic image analysis method<sup>7,8</sup>: (1) Unlike state-of-the-art high-content/high-throughput analysis on unrestricted cells having massive cell-to-cell variation, statistically significant results are obtained with only several tens of cells per condition. (2) Density-based screening estimates all required test parameters directly from the data and do not require computationally intensive analytical techniques such as classification. As image acquisition facilities are developed at an accelerated speed and advanced microscopes acquire thousands of images daily, there is a rising need for automated image analysis tools. (3) The organelle maps for each well allow direct visualization of cellular phenotypes. Thus, our organelle mapping framework in micropatterned cells may represent a major step toward quantitative, high-content, and fast analysis for automated detection of cell morphology changes in high-throughput screens. This framework has the potential to become a gold standard for cell phenotyping in high-throughput screening.

### Acknowledgments

We acknowledge Hesna Kara for the construction of stably GFP-expressing cell lines. We would like to thank Jean-Baptiste Brault and Lena Oesterlin for critical reading of the article.

### Declaration of Conflicting Interests

The authors declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: A patent (PCT No. EP2012/064239) has been filed on the density-based two-sample test.

### Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: K.S. received funding from the Fondation pour la Recherche Médicale en France and Association pour la Recherche sur le Cancer. J.-P.G. is supported by the Association pour la Recherche sur le Cancer. This project was further supported by grants from Agence Nationale de la Recherche (#2010 BLAN 122902), the Centre National de la Recherche Scientifique and Institut Curie, and the OSEO ISI program to CYTOO.

### References

1. Snijder, B.; Sacher, R.; Ramo, P.; et al. Population Context Determines Cell-to-Cell Variability in Endocytosis and Virus Infection. *Nature* **2009**, *461*, 520–523.
2. Sacher, R.; Stergiou, L.; Pelkmans, L. Lessons from Genetics: Interpreting Complex Phenotypes in RNAi Screens. *Curr. Opin. Cell Biol.* **2008**, *20*, 483–489.
3. Carpenter, A. E.; Jones, T. R.; Lamprecht, M. R.; et al. CellProfiler: Image Analysis Software for Identifying and Quantifying Cell Phenotypes. *Genome Biol.* **2006**, *7*, R100.
4. Loo, L. H.; Lin, H. J.; Steininger, R. J., III.; et al. An Approach for Extensively Profiling the Molecular States of Cellular Subpopulations. *Nat. Methods* **2009**, *6*, 759–765.
5. Thery, M.; Racine, V.; Piel, M.; et al. Anisotropy of Cell Adhesive Microenvironment Governs Cell Internal Organization and Orientation of Polarity. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 19771–19776.
6. Gumbiner, B. M. Cell Adhesion: The Molecular Basis of Tissue Architecture and Morphogenesis. *Cell* **1996**, *84*, 345–357.
7. Schauer, K.; Duong, T.; Bleakley, K.; et al. Probabilistic Density Maps to Study Global Endomembrane Organization. *Nat. Methods* **2010**, *7*, 560–566.
8. Duong, T.; Goud, B.; Schauer, K. Closed-Form Density-Based Framework for Automatic Detection of Cellular Morphology Changes. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 8382–8387.
9. Schneider, C. A.; Rasband, W. S.; Eliceiri, K. W. NIH Image to ImageJ: 25 Years of Image Analysis. *Nat. Methods* **2012**, *9*, 671–675.
10. Racine, V.; Sachse, M.; Salamero, J.; et al. Visualization and Quantification of Vesicle Trafficking on a Three-Dimensional Cytoskeleton Network in Living Cells. *J. Microsc.* **2007**, *225*, 214–228.
11. Duong, T. ks: Kernel Density Estimation and Kernel Discriminant Analysis for Multivariate Data in R. *J. Stat. Software* **2007**, *21*, (7).
12. R\_Development\_Core\_Team. *R: A Language and Environment for Statistical Computing*. Vienna, **2013**.
13. Duong, T., Koch, I., Wand, M. P. Highest Density Difference Region Estimation with Application to flow Cytometric Data. *Biom. J.* **2009**, *51*, 504–521.
14. Bowman, A. W.; Foster, P. Density Based Exploration of Bivariate Data. *Stat. Comput.* **1993**, *3*, 171–177.
15. Hyndman, R. Computing and Graphing Highest Density Regions. *Am. Stat.* **1996**, *50*, 120–126.