Published in final edited form as:

Nat Protoc. 2013 February ; 8(2): 345–354. doi:10.1038/nprot.2013.005.

Quantifying Spatial Organization in Point Localization Superresolution Images Using Pair Correlation Analysis

Prabuddha Sengupta¹, Tijana Jovanovic-Taliman^{1,2}, and Jennifer Lippincott-Schwartz^{1,3}

¹The Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, Maryland 20892, USA

Abstract

The distinctive distributions of proteins within subcellular compartments both at steady state and during signaling events play an essential role in cell function. Here, we describe a method for delineating the complex arrangement of proteins within subcellular structures visualized using point localization superresolution imaging. The approach, called pair correlation photoactivated localization microscopy (PC-PALM), employs a pair-correlation algorithm to precisely identify single molecules in point localization superresolution imaging datasets, and is used to decipher quantitative features of protein organization within subcellular compartments, including the existence of protein clusters, and the size, density and number of proteins in these clusters. We provide a step-by-step protocol for PC-PALM, illustrating its analysis capability for four plasma membrane proteins tagged with photoactivatable green fluorescent protein. The experimental steps for PC-PALM can be carried out in three days, and the analysis can be done in ~6–8 hours. Researchers need to have substantial experience in single molecule imaging and statistical analysis to conduct the experiments and carry out this analysis.

Keywords

PALM; pair correlation; spatial organization; quantitative analysis

INTRODUCTION

Super-resolution imaging techniques using fluorescent probes offer the capability of unravelling the spatial organization of biological structures at nanoscopic length scales^{1–5}. Among the techniques reaching the highest subdiffraction resolution are those using single molecule point localization approaches, including: photoactivated localization microscopy (PALM)⁶, fluorescence photoactivation localization microscopy (fPALM)⁷, stochastic optical reconstruction microscopy (STORM)⁸, direct STORM (dSTORM)⁹,

³Corresponding author, jlippin@helix.nih.gov, Phone: 301-402-1010; Fax: 301-402-0078, <https://science.nichd.nih.gov/confluence/display/sob/Home>.

²Current address: Department of Chemistry, University of Hawaii at Manoa, Honolulu, Hawaii 96822, USA

AUTHOR CONTRIBUTIONS

P.S. and T.J.-T. designed and performed experiments, analysed data, and wrote the paper; J.L.-S. designed the experiments, and wrote the paper.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

SUPPLEMENTARY INFORMATION

Supplementary Software. Compute_Paircorrelation

Matlab code for computation of autocorrelation and cross-correlation functions from binary images of single molecule spatial distribution.

ground state depletion microscopy followed by individual molecule return (GDSDIM)¹⁰, bleaching/blinking assisted localization microscopy (BaLM)¹¹, and generalized SHRIMP (gSHRIMP)¹². These methods produce superresolution images through a pointillist strategy, in which individual fluorescent molecules are localized through multiple cycles of photoactivation, photoswitching, blinking or bleaching to temporally isolate their fluorescence from other spatially proximal molecules. The process is performed thousands of times, with individual frames in the image series containing only a sparsely distributed subset of fluorescing molecules. The precise location of individual molecules in each frame is estimated by mathematical fitting^{13, 14}. Subsequently, the fitted localizations of all molecules in all frames are summed into a super-resolution image, which contains both nanometric structural context and molecular detail.

The point-localization superresolution (PL-SR) imaging techniques described above have produced spectacular images of biological structures at dramatically improved spatial resolution^{15–17}. Nevertheless, extracting precise information from these images regarding single molecule locations and their complex nanoscale organization has proved challenging. This is because rigorous, statistical analysis methods are required both to precisely identify single molecules in PL-SR images and to elucidate spatial scales of molecular organization. Without such statistical methods, it is impossible to interpret molecular distributions in PL-SR images.

Here, we discuss a new approach for achieving such meticulous analysis of PL-SR image datasets. The method, termed pair correlation PALM (PC-PALM), both precisely identifies single molecules in an image and allows the spatial parameters of protein organization to be quantified (i.e., size and protein packing density of higher-order biological structures). It was initially described as a technique in *Nature Methods*¹⁸. We provide a detailed protocol for PC-PALM, give troubleshooting tips and describe its benefits for quantifying different parameters of single molecule spatial organization, including the presence of molecule clusters, their size, density and abundance in a particular site. The workflow for PC-PALM analysis is shown in Fig. 1. The protocol should additionally serve as a useful supplement to previously published protocols describing experimental steps for performing PL-SR microscopy^{19, 20}.

Difficulties in precisely identifying single molecules in PL-SR images

A prerequisite for quantifying protein distributions in PL-SR images is the precise identification of single molecules. This capability is essential for interpreting whether proteins within a particular subcellular compartment are randomly distributed or clustered. But precisely identifying a single molecule in PL-SR images is not easy. A single fluorescent molecule fluoresces across multiple frames before being irreversibly photobleached^{18, 21–23}. Because the calculated position of the molecule in each frame varies a small amount (due to differences in the number of photons collected in each frame), the single molecule usually appears as a cluster of “peaks” (localized fluorescent spots) in the summed frames of the superresolution image. If the cluster of peaks all appear in successive frames, then it is possible to group them and assign them to a single protein. Specifically, peaks spread over a certain area with a radius termed group radius are assigned to a single protein. Using such grouping of peaks, statistical methods like Ripley’s spatial analysis²⁴ have attempted to interpret characteristics of protein distributions in PL-SR images^{25, 26}. The pitfall of these approaches, however, is that most photoactivatable fluorescent proteins and fluorescent dyes blink off for a variable time period before becoming fluorescent again^{27–29}, making grouping inaccurate. An additional problem occurs for these analysis methods based on grouping when images are collected using fast acquisition rates. This often results in more than one fluorescent spot in a diffraction limited area per frame^{30–35},

requiring the researcher to either discard the data or use alternative methods to resolve the overlapping spots.

As a consequence of these problems, spatial analysis algorithm programs like Ripley's (that use grouping) frequently misassign peaks due to the long and variable blinking times of fluorescent molecules. This can introduce severe artifacts in the final analyses, since spatial errors acquired using these methods at short spatial scales become integrated and propagated at longer distances. Alternative strategies to grouping, therefore, are needed to precisely identify single molecules in PL-SR images if quantitative assessment of protein organization is desired.

Pair-correlation PALM and its applications

A recently developed statistical approach that avoids the drawbacks of grouping peaks is PC-PALM. Instead of grouping peak radii or blink time intervals, PC-PALM employs a statistical algorithm that uses pair-correlation analysis to quantify the spatial organization of point patterns present in a pointillist image^{18, 21}. The algorithm takes advantage of the specific spatial signature of the peaks belonging to the same fluorescent molecule to calculate the contribution of multiple appearances of this molecule relative to the spatial organization of all point patterns. In other words, it uses the unique spatial characteristics of peak clusters to quantify and remove the effect of multiple appearances, thereby avoiding the difficult (and sometimes intractable) problem of finding an optimal time window for grouping of peaks. Thus, by carrying out the entire analysis in the spatial domain, PC-PALM circumvents the problems associated with grouping peaks (carried out in both spatial and temporal dimensions).

PC-PALM is adaptable to all point localization based super-resolution imaging for quantitative analysis^{21, 36}. The approach is particularly useful when single probe molecules or single proteins-of-interest are detected multiple times during an experiment. This often confounds quantitative analysis of data in various superresolution optical microscopy and electron microscopy experiments. As long as the multiple signals from a single molecule has a defined spatial distribution, the PC-PALM approach can be used to subtract out the contribution of the multiple appearances, and provide an accurate measurement of the actual spatial organization of fluorescent proteins in a sample.

In STORM, multiple appearances of organic dyes pose an especially difficult problem for spatial analysis of protein populations since these dyes can blink and reappear across a time-scale of seconds to minutes. Conventional grouping and spatial analysis methods (i.e., Ripley's) will often provide inaccurate measurements in high-density images obtained by STORM. However, the multiple peaks of a single organic dye have a definite spatial distribution determined by the average positional uncertainty of the STORM experiment. PC-PALM offers a robust solution for identifying individual molecules in this case, and can be used for quantitative spatial analysis of STORM images³⁶.

PC-PALM can also be implemented for analysis of gold point-patterns in electron microscopy images^{36,37}. Detection of spurious clustering due to labelling of single proteins with multiple gold nanoparticles is a serious problem in immunoelectron microscopy datasets. The gold nanoparticles are normally attached to antibodies bound to proteins-of-interest, and are expected to spatially sample a defined area centred on the protein. PC-PALM can be used to remove the contribution of multiple labelling for measuring the spatial clustering features in a specimen.

PC-PALM theory

The mathematical derivation for the correlation functions has been described in detail previously in separate publications^{18, 36, 37}. In this derivation, the autocorrelation function, $g(r)$, of a point pattern measures the average probability density of points at a distance 'r' around each point. Protein domains are characterized by changes in density across the domain boundary. Consequently, a pair-correlation function can be used to detect the presence of protein clusters and to characterize the spatial scales of protein organization in such clusters.

A key concept in PC-PALM is the recognition that even though a single fluorescent protein makes multiple appearances across many frames in an image acquisition series, the fluorescent peaks that arise from a single protein are spatially organized in a defined way. These peaks are scattered across a two dimensional (2D) Gaussian surface, which represents the point spread function (PSF) (i.e. the localization uncertainty) of the imaging technique. The average localization uncertainty of the proteins can be directly estimated from the mathematical fitting of the peaks during the peak-localization process. This localization uncertainty represents the standard deviation of the 2D Gaussian surface associated with the peaks from a single protein. Consequently, the magnitude of the autocorrelation function contributing to the measured total correlation function can be estimated from the autocorrelation function of the 2D Gaussian function of the cluster of peaks from a single protein, and is represented as the following¹⁸:

$$g(r)^{stoch} = \frac{1}{4\pi\sigma_s^2\rho^{protein}} \exp\left(\frac{-r^2}{4\sigma_s^2}\right) \quad \text{Eqn. 1}$$

where σ_s is the average uncertainty in single molecule localization, and $\rho^{protein}$ is the average protein density.

The autocorrelation function of the protein molecules is equal to one if they are distributed randomly across the image. In such a case, the computed autocorrelation function is:

$$g(r)^{peaks} = g(r)^{stoch} + 1 \quad \text{Eqn. 2}$$

The autocorrelation function describing the relative distribution of proteins will have a value greater than one across spatial scales over which the proteins are clustered when proteins are non randomly distributed. An exponential function can be used to approximate the correlation function of proteins if they are present in randomly distributed clusters of no definite shape. Then, the calculated autocorrelation function is given by¹⁸:

$$g(r)^{peaks} = \frac{1}{4\pi\sigma_s^2\rho^{protein}} \exp\left(\frac{-r^2}{4\sigma_s^2}\right) + \left(A \exp\left(\frac{-r}{\xi}\right) + 1\right) * g(r)^{PSF} \quad \text{Eqn. 3}$$

Here, the pre-exponential factor A is proportional to the density of proteins in the clusters, and the correlation length, ξ , is proportional to the cluster size. The PSF (related to uncertainty of position determination) of the imaging method is represented by $g(r)^{PSF}$, and

is equal to $\frac{1}{4\pi\sigma_s^2} \exp\left(\frac{-r^2}{4\sigma_s^2}\right)$.

The cluster parameters ξ and A are obtained by fitting the computed autocorrelation to Eqn. 3.

$N^{cluster}$, the average occupancy of a cluster, is equal to the following:

$$N^{cluster} \approx 2A\pi\xi^2\rho^{protein} \quad \text{Eqn. 4a}$$

The ratio of density of proteins in clusters to the average density across the entire image (increased density of proteins) is given by:

$$\psi^{cluster} = 2A \quad \text{Eqn. 4b}$$

The PC-PALM analysis method is a powerful approach to detect the presence of morphologically featureless, clusters of protein scattered across a two dimensional surface. Moreover, it can be used to determine the physical characteristics of such clusters. The autocorrelation function calculates the radially averaged correlation function for all the depicted points (fluorescent spots) in the entire image. Using computational modeling, we have confirmed that this analysis can provide an accurate estimation of clustering parameters for radially-symmetric or randomly-shaped protein domains, which are not polarized in any specific direction. Because the PC-PALM algorithm calculates average values of protein parameters over an entire image, it is only suitable for characterizing protein domains with no preferred direction of organization. This is generally not a problem when clusters of proteins are not expected to have any preferred direction in the way they are distributed across a two dimensional (2D) surfaces like the plasma membrane. But for asymmetric clusters with elongated morphology it is a problem, as averaging decreases the degree of clustering detected in such structures. Other commonly used statistical algorithms, including Ripley's analysis, share this drawback.

Another limitation of PC-PALM analysis (and all other statistical analysis) is that it provides a description of clusters only in terms of the observed number of proteins. In most PL-SR imaging experiments, however, not all photoactivatable molecules become fluorescent during the imaging period. This can occur for a number of reasons: a fraction of the fluorescent molecules may never undergo photoactivation during an experiment; some fluorophores may undergo premature photobleaching; there may be an inadequate number of photons to fit a molecule; and a high local background signal can preclude detection and localization of a fraction of the fluorescent molecules. Thus, the number of molecules detected in a cluster during the quantitative analysis is not necessarily equal to the absolute number of proteins present in such clusters. Nonetheless, PC-PALM can still provide significant quantitative information about nanoscale changes in protein organization during various physiological processes. This is because it can provide key information regarding the size, density and protein abundance in protein clusters, even in the absence of knowing absolute number of proteins. Additionally, by performing PC-PALM analysis on a reference structure containing a known number of PA-FP tagged proteins, it is possible to estimate the fraction of PA-FP molecules that are detected, on average, during the PALM experiment. This can be used to calibrate of the protein number obtained from PC-PALM analysis, and to estimate the absolute number of PA-FP tagged proteins present in a previously uncharacterized structure or protein domain. However, appropriate reference structures are not always available to implement such a calibration scheme. Common advantages and limitations of PC-PALM approach are listed in Table 1.

Pair cross-correlation of superresolution images

In addition to performing autocorrelation analysis to detect the clustering of a single kind of protein molecule, pair-correlation analysis can also be used to delineate the relative spatial organization of two different proteins with respect to each other. Cross-correlation analysis of two-colour superresolution images provide estimates of spatial scales of co-clustering of

two proteins at length scales down to the localization uncertainty of the experiments (~50–60 nm for PALM). Unlike the autocorrelation function, the cross-correlation function is not affected by multiple appearances. Instead, the cross-correlation quantifies the actual correlation between two different photoactivatable proteins, convolved with the PSFs of positional uncertainty. Thus, if the cross-correlation function is modelled by an exponential function, then the measured cross-correlation, $c(r)^{peaks}$, can be fit to the following equation:

$$c(r)^{peaks} = \left(A \exp\left(\frac{-r}{\xi}\right) + 1 \right) * c(r)^{PSF} \quad \text{Eqn. 5}$$

In Eqn. 5, $c(r)^{PSF}$ is the cross-correlation of the PSF of the two photoactivatable probes. It should be noted that the problem of under-estimation of protein numbers can be more severe during two-color imaging, where photoactivation and imaging of one PA-FP can lead to simultaneous photoactivation of the PA-FP, which is not being imaged. Thus, imaging conditions need to be carefully optimized to minimize the loss in protein localization for two-color superresolution experiments.

EXPERIMENTAL DESIGN

Choice of fixative

Use a combination of formaldehyde and glutaraldehyde for proper fixation and immobilization of proteins^{18, 38}. To perform quantitative spatial analysis of PALM data sets for characterizing protein organization at spatial scales of ~50–250 nm, it is paramount to immobilize the proteins and block their lateral diffusion. At the very least, the fixation conditions should confine lateral movement of proteins during the experimental window to spatial scales considerably smaller than the spatial scales of interrogation. We found that short-term incubation with 4% (wt/vol) formaldehyde does not immobilize plasma membrane proteins effectively. Addition of 0.2% (wt/vol) glutaraldehyde, in contrast, largely immobilizes plasma membrane proteins. Glutaraldehyde quenches the fluorescence of photoactivatable proteins to a certain extent, so do not use a higher percentage of glutaraldehyde or the sampling density will be significantly decreased. Additionally, prolonged exposure to fixative can result in increased quenching of fluorescent proteins, decreasing the sampling density. Following incubation with fixatives, the sample should be promptly quenched with glycine (or BSA) to minimize subsequent fluorescence quenching. Image cells within 4–5 hours of fixation since the photoactivatable proteins lose their fluorescence with time.

Choice of PA-FPs for two-color experiments

Choose PA-FPs for cross-correlation experiments carefully so that the two PA-FPs can be differentially activated and imaged with minimal cross talk between the two probes. Clear separation of the spectral profile and differential sensitivities to the activation laser will minimize incorrect assignment of probe localizations. Readout laser induced activation of some PA-FPs like photoactivatable green fluorescent protein (PAGFP)³⁹ helps to minimize cross-activation induced loss of protein localization during two-color PALM experiments. Photoactivatable mCherry1 (PAmCh)⁴⁰ has spectrally separated emission profile from that of PAGFP, and is not efficiently activated by 488 nm laser and low power 405 nm laser. Consequently, PAGFP and PAmCh are suitable for two-color PALM experiments^{18, 40}.

MATERIALS

REAGENTS

- Cultured cells

- Plasmid encoding the protein of interest, tagged with an appropriate PA-FP
- Acetone (Sigma, #179124)
- Hydrogen peroxide (Sigma, #216763),
- Ammonium hydroxide (Sigma, #221228)
- Hellmanex III (Fisher, # 14-385-864)
- Fibronectin (Sigma, # F2006)
- Phosphate Buffered Saline (Invitrogen, 3AM9625)
- Phenol red free DMEM (Cellgro, #17-205-CV) or other growth medium
- Fetal bovine serum (Cellgro, #35-016-CV)
- L-Glutamine (Cellgro, #25-005-CI)
- Trypsin, without phenol red (Cellgro, #25-054-CI)
- FuGene 6 Transfection Agent (Promega, #E2692)
- 16% (wt/vol) Paraformaldehyde (Electron Microscopy Sciences, #15710)
- 25% (wt/vol) Glutaraldehyde (Electron Microscopy Sciences, #16220)
- Glycine (Sigma, #G8898)
- Bovine Serum Albumin (Sigma, #A2153)
- Tetraspeck beads (Invitrogen, #T-7279)
- 3-aminopropyl trimethoxysilane (Sigma, #281778)
- Purified photoactivatable fluorescent proteins (PA-FPs) (See Reagent Setup)

EQUIPMENT

- #1.5 coverslips (Warner Instruments, #CS-25R15) see Equipment Setup
- Olympus IX81 microscope modified to image in total internal reflection mode, and equipped with a 60×1.45 numerical objective (Olympus, PlanApoN) and an EMCCD camera (Andor Technology, DV887ECS-BV).

EQUIPMENT SETUP

Coverslips for imaging—Coverslips need to be cleaned thoroughly to remove dirt and all other sources of background fluorescence. Clean high refractive indexed coverslips (#1.5) by boiling in acetone for 10 min at 70°C, and then incubating in a mixture of 1:1:5 hydrogen peroxide/ammonium hydroxide/water for 1 hour at 70°C. Alternatively, incubate the coverslips with 1% (vol/vol) Hellmanex III for 12 hours. Wash the cleaned coverslips thoroughly (at least 5 times) with distilled water. Next sterilize the coverslips by washing with 200-proof ethanol. Flame the coverslips in a tissue-culture hood, and place them in sterile 35 mm culture dishes. The coverslips can be stored in a dry place and cells can be directly plated on these coverslips before the experiment.

REAGENT SETUP

Cells—Any cell line that adheres well to fibronectin (or other extracellular matrix protein) coated coverslips and can be efficiently transfected with plasmid constructs (corresponding to proteins of interest) can be used. Cells should be grown and maintained in phenol red free medium.

Purified photoactivatable proteins—Express PA-FPs to be used as fluorescent probes for superresolution experiments in bacterial culture, and obtain purified PA-FPs using an appropriate protein purification technique. Store the purified protein at -80°C . The purified PA-FP will be used in Steps 21–23 of the Procedure to estimate the average number of appearances of single PA-FP molecules.

Fixative—Prepare the fixative by diluting paraformaldehyde and glutaraldehyde to make a solution with final concentration of 4% (wt/vol) paraformaldehyde and 0.2% (wt/vol) glutaraldehyde in phosphate buffer saline (PBS). Make 1 ml aliquots of the fixative and store them at -80°C . Thaw the fixative just before fixing the cells.

PROCEDURE

Sample Preparation TIMING 2.5 days

- 1| Incubate clean coverslips (see Equipment Setup) with $10\text{ }\mu\text{g/mL}$ fibronectin in PBS (pH 7.4) for 20 min at 37°C .

CRITICAL STEP PALM involves imaging of single molecules with low signal to noise ratio; thus, any residual fluorescence from dirt on the coverslip will add to the background noise and decrease the contrast. This, in turn, will affect the efficiency of single molecule detection. Furthermore, dirty coverslips increase the chances of detection of spurious signal. Commercially available tissue culture dishes with coverslips mounted at the bottom are usually not suitable for single molecule imaging due to high background signal.

?TROUBLESHOOTING

- 2| Plate cells on fibronectin-coated coverslips at a density such that cells will have ~50% confluency after 24 hours.

CRITICAL STEP Fibronectin coating is crucial for achieving flat, adherent tracts of ventral cell surface. The number of detected photons and the background signal determine the localization precision of a single fluorescent molecule.

CRITICAL STEP Non-fluorescent (phenol-red free) media should be used for culturing cells throughout the Procedure to minimize background fluorescence during single molecule experiments.

- 3| Using FuGENE 6 transfection reagent or any other transfection method, transfect cells with a plasmid encoding the protein of interest tagged with PA-FP 24 hours after plating. To avoid increased auto-fluorescence of cells, wash away the residual transfection agent 6 hours after transfection by replacing the medium with fresh medium.

CRITICAL STEP Culture dishes containing transfected cells should be kept wrapped in aluminium foil or any other opaque material during incubation and transferring from one place to another to minimize the photoactivation of fluorescent proteins by ambient light.

- 4| Wash cells twice with PBS pre-warmed at 37°C when the cells are expressing optimal amount of labelled protein (usually 18–24 hours post transfection). Immediately fix the cells with fixative reagent for 30 min at room temperature.
- 5| Next, quench unreacted fixative with 25 mM glycine in PBS or 10 mg mL^{-1} BSA in PBS (filtered) for 10 minutes.

- 6| Wash 3–4 times with PBS. Incubate cells with Tetraspeck beads or other fiducial markers in PBS for 10 min. During this time the beads will settle down and adhere to the glass coverslip. Remove the non-adherent, floating beads by washing the cells with PBS. Image the cells immediately after fixation.

CRITICAL STEP Calibrate the concentration of fiducial markers and time of incubation so that only a few of the fiducial markers (4–10) are present in any random field-of-view. In order to achieve accurate drift correction, try to image cells that have fiducial markers evenly distributed around the perimeter of the cell.

?TROUBLESHOOTING

Imaging TIMING 1-2 hours

- 7| Choose highly spread cells from sub-confluent areas of the coverslip using transmitted light or DIC. This helps to avoid 3-dimensional geometries in the TIRF zone, ensuring that membrane topography does not affect the quantitative analysis of protein organization. It is important to ensure cells are plated at adequate density (i.e., ~50% confluent during transfection) since having either too sparse or almost confluent cells may impact cell physiology (i.e. adhesion, metabolism) and experimental reproducibility.
- 8| Choose the magnification such that the pixel size is similar to the standard deviation of the point spread function, which is optimal for accurate single molecule localization. In the presence of non-negligible background fluorescence, this gives the optimal localization¹³.
- 9| Choose maximal linear EM gain on the EMCCD camera, and use frame transfer and constant baseline level options during acquisition.
- 10| Choose appropriate filters and laser lines for selected PA-FPs. Some PA-FPs like PAGFP undergo efficient activation when illuminated with the readout (i.e. excitation) laser, and single colour PALM experiments can be performed with these PA-FPs using a single laser.
- 11| Choose the activation and excitation laser power and the exposure time such that enough photons per fluorophore are detected in each frame to ensure good localization (~ 25 nm), and that the density of single molecule peaks is low enough for detection of single, isolated peaks. Use the same excitation laser power for the entire experiment so that the photophysical parameters do not change during the course of the experiment. Image cells until all detectable proteins are recorded. For two-colour experiments, design the activation and imaging protocol to ensure that the localized single molecules can be unambiguously assigned to one or the other probe. Carefully calibrate the laser intensities in each step to ensure proper optimal detection of single molecules and minimize signal loss due to cross-activation. For example, use the following protocol for a two-colour experiment with PAGFP and PAmCh¹⁸: Image PAGFP exhaustively using the activation/readout 488 nm laser; Next, use low power 405 nm laser in combination with 488 nm laser to record the residual PAGFP signal; Finally, use a combination of 405 nm activation laser and 561 nm readout laser to image PAmCh.

?TROUBLESHOOTING

Analysis TIMING 6–8 hours

- 12| Identify single molecule peaks in each frame and determine the positional coordinates using a least-squared fit to a two-dimensional Gaussian point spread function.
- 13| Estimate the localization uncertainty (σ_{xy}^2) of each localized single molecule peak using the following equation¹⁴:

$$\sigma_{xy}^2 = \frac{\sigma_s^2}{N} \left(\frac{16}{9} + \frac{8\pi\sigma_s^2 b^2}{Na^2} \right) \quad \text{Eqn. 6}$$

where $\sigma_s^2 = s^2 + a^2/12$, s is the standard deviation of the point spread function, a is pixel dimension in the image, N is the number of photons detected from the molecule, b is the number of background photons.

?TROUBLESHOOTING

- 14| Use the positional information of the fiducials to perform drift-correction using a suitable fitting algorithm. For two-color experiments, align the two channels by registering multiple fiducial positions from the two data sets.

CRITICAL STEP The errors in drift correction and channel registration contribute to decreased resolution of the PALM image. Accurate drift-correction and image registration are critical for accurate characterization of spatial scales of organization. Typically, the positional error contributed by drift can be reduced to a few nanometers, so that drift has a minor contribution to overall positional uncertainty.

- 15| Generate a histogram of the localization precision (σ_{xy}) of all peaks that have been localized in the experiment. Determine the average positional uncertainty (σ^{raw}) by fitting the distribution to a skewed Gaussian or using an appropriate function.
- 16| Replace peaks that appear in successive frames and within an area of radius $2.5 \times \sigma^{raw}$ with one peak. Such peaks correspond to a single image of a molecule stretched across multiple frames because of the interruptions due to the acquisition frame rate. Calculate positional coordinates (x_m and y_m) of each resultant peaks from the positional coordinates of the individual peaks of the corresponding group weighted with corresponding the number of photons in each peak as follows:

$$x_m = \frac{\sum_{i=1}^{npeaks} x_i N_i}{\sum_{i=1}^{npeaks} N_i} \quad \text{Eqn. 7a}$$

where x_i and N_i are the positional coordinate and number of photons detected for the individual peaks of the group, and $npeaks$ is the number of peaks in the group. The correlation function has been derived such that its stochastic part measures the correlation due to multiple appearances of a single PA-FP arising from blinking events with at least one intervening dark frame. The grouping procedure in this step substitutes peaks representing a single appearance of a PA-FP (with no detectable blinking) with a single peak. This is required to ensure that the measured correlation of peaks can be correctly described by the model correlation function.

- 17| Next, calculate the localization precision (weighted over number of photons) of each grouped peak in X- (σ_x^m) and Y-direction (σ_y^m) as follows⁶:

$$\sigma_x^m = \left[\sqrt{\frac{\sum_{i=1}^{npeaks} \{(x_i - x_m)^2 \times N_i\}}{\sum_{i=1}^{npeaks} N_i \times npeaks} + \frac{\sum_{i=1}^{npeaks} \{\sigma_i^2 \times N_i\}}{\sum_{i=1}^{npeaks} N_i \times npeaks}} \right] / \sqrt{2} \quad \text{Eqn. 7b}$$

Finally, determine the localization precision of the grouped peak (σ^{peak}) as the maximum of σ_x^m and σ_y^m .

- 18| Obtain an estimate of the mean localization of the detected single molecules (σ_s) from the distribution of sigma of the peaks (σ^{peak}) by fitting the distribution to a skewed Gaussian or other appropriate function. Discard peaks with σ^{peak} -values greater than $3 \times \sigma^s$.
- 19| Use the position coordinates of the peaks to generate a binary PALM image where all detected peaks are represented as pixels with a value of 1, while the rest of the pixels have a value of 0. The pixel size of the binary image should be small enough so that spatially adjacent molecules do not overlap, and they are individually represented in the image. Thus, the pixel size should be smaller than the minimum distance between two adjacent molecules. This binary PALM image is used for the analysis of spatial organization of protein molecules. For two-color experiments, generate two separate binary images representing the spatial distribution of the two different proteins.
- 20| Perform Option A for autocorrelation analysis in order to investigate spatial organization of a particular protein (use binary images generated from single channel, i.e. one-color, superresolution data set). Perform Option B for cross-correlation analysis to investigate relative spatial organization (spatial interaction) of two different proteins (use binary images generated from two-color superresolution data sets).

Option A: Autocorrelation Analysis **TIMING several hours**

- i. First, assess the average number of times that a PA-FP appears in the binary image (due to blinking) before it is irreversibly photobleached. Incubate clean coverslips for 5 min with 5% (wt/vol) solution of 3-aminopropyl trimethoxysilane in acetone at 25°C. Wash the coverslips successively with acetone followed by PBS. Next, treat the coverslips with 0.25% (wt/vol) glutaraldehyde in PBS for 30 minutes.
- ii. Ultracentrifuge a freshly thawed solution of purified PA-FP. Collect the supernatant (containing monomeric protein) and measure the concentration of PA-FP. Finally, immobilize PA-FP on the functionalized coverslip by incubating the coverslips with a mixture of 1 nM PA-FP and 5 μ M bovine serum albumin in PBS for 30 minutes at 25°C inside a humid container. Remove unreacted protein by washing the coverslips thoroughly with PBS.

CRITICAL STEP The exact concentration of PA-FP molecules and incubation time should be calibrated so that the coverslips are modified with a sparse concentration of PA-FPs.

- iii. Image sparsely distributed, immobilized PA-FPs immobilized on the coverslip using identical imaging conditions as those used for imaging cells (Steps 8–11), and estimate the average number of appearances of peaks.

?TROUBLESHOOTING

- iv. Next, calculate the average density of PA-FP modified proteins ($\rho^{protein}$) from the density of detected peaks in the image (ρ^{peaks}) and the average number of appearances.
- v. Randomly choose regions ($4\text{--}16\ \mu\text{m}^2$) from the binary image for correlation analysis.

CRITICAL STEP: Avoid visible holes in the zoomed out binary images when selecting sections for correlation analysis. This will ensure that membrane topological effects do not contribute significantly to the spatial analysis.

- vi. Compute two dimensional (2D) autocorrelation functions for the selected ROI using the following equation:

$$g(r) = \frac{1}{\rho_{peaks}^2} \left\{ \frac{FFT^{-1}(|FFT(Im)|^2)}{FFT^{-1}(|FFT(W)|^2)} \right\} \quad \text{Eqn. 8}$$

$FFT()$ and $FFT^{-1}()$ represent the computation of the 2D discrete Fourier and 2D discrete inverse Fourier transforms, respectively, using an FFT algorithm. Use either Matlab's in-built functions *fft2* and *ifft2* for these FFT computations (see, accompanying Matlab code in Supplementary Data), or any other suitable computing software. Im is the binary PALM image appended with zeros in all directions. The added stretch of pixels (with zero-value) corresponds to length scale larger than that being interrogated by correlation function, i.e. the dimension of the largest structure being investigated. This padding of image with zero-pixels helps in avoiding artifactual correlations arising from the periodicity of the discrete Fourier functions. A window function, W , is used to ensure that only the correlation in the interval corresponding to the original binary image contributes to the analysis. W is a matrix with same dimensions as the image Im , and it is similarly appended with a stretch of zero-valued pixels. The pixels in the central measurement area have value 1.

- vii. Use non-linear least squares analysis to fit the calculated autocorrelation values ($g(r)$, $r > 0$) to Eq. 2. Adjust the model parameters σ_s and $\rho^{protein}$ during the fitting procedure. If the parameter values extracted from the fit are within $\pm 25\%$ of the measured values, it suggests that the protein molecules are randomly organized. This completes the autocorrelation analysis of this particular data set. However, if the measured correlation function cannot be fit with estimated values of parameters ($\pm 25\%$ of the measured values), it suggests that the protein molecules are organized in a non-random fashion. To assess the clustering parameters of protein organization, proceed to step 20A (viii)
- viii. If the measured correlation function cannot be adequately described by a random model (Eqn. 2), then the proteins are organized non-randomly. Use an exponential function described in Eqn. 3 to approximate the protein correlation. Fit the measured correlation function to Eq. 3 to obtain an estimate of the model parameters ξ and A . Constrain σ_s and $\rho^{protein}$ to be consistent with the calculated values during the fitting procedure. Alternatively, choose any other physically meaningful model that can describe the protein correlation function.
- ix. Calculate the mean size of clusters (ξ), average number of proteins in each cluster ($N^{cluster}$), and the enhanced density of proteins in the cluster ($\psi^{cluster}$) using Eqns. 4a,b.

Option B: Cross-correlation Analysis TIMING 0.5–1 hours

- i. Choose regions (4–16 μm^2) for cross-correlation analysis from either one of the binary images. Avoid visible holes in zoomed out image to minimize artifacts arising from membrane topography.
- ii. Calculate the cross-correlation function for the two proteins for chosen regions using the following equation:

$$c(r) = \text{Real}\left(\frac{\text{FFT}^{-1}(\text{FFT}(Im^{ch1}) \times \text{conj}(\text{FFT}(Im^{ch2})))}{\rho^{ch1}\rho^{ch2}N(r)}\right) \quad \text{Eqn. 9}$$

Real() represents the real part, *conj()* indicates complex conjugate, Im^{ch1} and Im^{ch2} are the binary PALM images of the two PA-FPs respectively, and ρ^{ch1} and ρ^{ch2} are the mean peak densities of the two PA-FPs across the image, respectively.

- iii. Fit the cross-correlation values computed as a function of radius in step 20B(ii) to Eqn. 5 to get an estimate of the spatial scales of correlation.

TROUBLESHOOTING

Troubleshooting advice can be found in Table 2.

TIMING

Coverslip cleaning, approximately 12 h

Steps 1–6, Cell plating, growth, transfection, and fixation, approximately 2.5 d

Steps 7–10, Microscope setup, several minutes to 1 h

Step 11, Data acquisition, 10 min–1 h for one color; depends on expression density and imaging conditions

Steps 12–19, Data analysis, approximately several hours; depends on number of detected peaks

Step 20A Auto-correlation analysis, approximately several hours; depends on number of detected peaks

Step 20B, Cross-correlation analysis, 0.5–1 h

ANTICIPATED RESULTS

PC-PALM provides detailed description of organization of proteins on a nano-scale level. It can give insight into clustering and/or oligomerization of proteins on the plasma membrane, which is especially important for understanding biological mechanisms such as signaling. Typical results are shown in Fig. 2, in which two lipid anchored proteins (PAGFP-GPI and Lyn-PAGFP) and two transmembrane proteins (Lat-PAGFP and VSVG-PAGFP) are imaged and analyzed using pair-correlation function in an expressing COS-7 cells¹⁸. The measured correlation functions were fit to a clustered model. The fit parameters are used to quantify the physical characteristics of the nanoscale spatial distribution of the four PAGFP-labeled proteins across the plasma membrane. PAGFP-GPI and VSVG-PAGFP are present in small clusters of uniform size (<60 nm). These clusters are densely packed, containing three or less proteins. Lyn-PAGFP and Lat-PAGFP appeared organized into clusters with radii of variable size (radius between 60–180 nm). These clusters contained three or more proteins.

The local densities of Lyn and Lat-PAGFP were 2–4 times higher than that expected for a randomly distributed protein.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We would like to thank George Patterson and Yan Fu (National Institute of Biomedical Imaging and Bioengineering) for DNA constructs and purified proteins, Harald Hess and Gleb Stengel (Howard Hughes Medical Institute, Janelia Farm Research Campus) for analysis software and valuable discussion.

REFERENCES

1. Patterson G, Davidson M, Manley S, Lippincott-Schwartz J. Superresolution imaging using single-molecule localization. *Annu Rev Phys Chem.* 2010; 61:345–367. [PubMed: 20055680]
2. Toomre D, Bewersdorf J. A new wave of cellular imaging. *Annu Rev Cell Dev Biol.* 2010; 26:285–314. [PubMed: 20929313]
3. Patterson GH. Fluorescence microscopy below the diffraction limit. *Semin Cell Dev Biol.* 2009; 20:886–893. [PubMed: 19698798]
4. Flors C, Earnshaw WC. Super-resolution fluorescence microscopy as a tool to study the nanoscale organization of chromosomes. *Curr Opin Chem Biol.* 2011; 15:838–844. [PubMed: 22098720]
5. Sigrist SJ, Sabatini BL. Optical super-resolution microscopy in neurobiology. *Curr Opin Neurobiol.* 2012; 22:86–93. [PubMed: 22051692]
6. Betzig E, et al. Imaging intracellular fluorescent proteins at nanometer resolution. *Science.* 2006; 313:1642–1645. [PubMed: 16902090]
7. Hess ST, Girirajan TPK, Mason MD. Ultra-high resolution imaging by fluorescence photoactivation localization microscopy. *Biophysical journal.* 2006; 91:4258–4272. [PubMed: 16980368]
8. Rust MJ, Bates M, Zhuang XW. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nature methods.* 2006; 3:793–795. [PubMed: 16896339]
9. Wombacher R, et al. Live-cell super-resolution imaging with trimethoprim conjugates. *Nature methods.* 2010; 7:717–719. [PubMed: 20693998]
10. Folling J, et al. Fluorescence nanoscopy by ground-state depletion and single-molecule return. *Nature methods.* 2008; 5:943–945. [PubMed: 18794861]
11. Burnette DT, Sengupta P, Dai YH, Lippincott-Schwartz J, Kachar B. Bleaching/blinking assisted localization microscopy for superresolution imaging using standard fluorescent molecules. *Proceedings of the National Academy of Sciences of the United States of America.* 2011; 108:21081–21086. [PubMed: 22167805]
12. Simonson PD, Rothenberg E, Selvin PR. Single-molecule-based super-resolution images in the presence of multiple fluorophores. *Nano Lett.* 2011; 11:5090–5096. [PubMed: 22003850]
13. Thompson RE, Larson DR, Webb WW. Precise nanometer localization analysis for individual fluorescent probes. *Biophysical journal.* 2002; 82:2775–2783. [PubMed: 11964263]
14. Mortensen KI, Churchman LS, Spudich JA, Flyvbjerg H. Optimized localization analysis for single-molecule tracking and super-resolution microscopy. *Nature methods.* 2010; 7:377–381. [PubMed: 20364147]
15. Kanchanawong P, et al. Nanoscale architecture of integrin-based cell adhesions. *Nature.* 2010; 468:580–584. [PubMed: 21107430]
16. Loschberger A, et al. Super-resolution imaging visualizes the eightfold symmetry of gp210 proteins around the nuclear pore complex and resolves the central channel with nanometer resolution. *Journal of cell science.* 2012; 125:570–575. [PubMed: 22389396]
17. Kopek BG, Shtengel G, Xu CS, Clayton DA, Hess HF. Correlative 3D superresolution fluorescence and electron microscopy reveal the relationship of mitochondrial nucleoids to

- membranes. *Proceedings of the National Academy of Sciences of the United States of America*. 2012; 109:6136–6141. [PubMed: 22474357]
18. Sengupta P, et al. Probing protein heterogeneity in the plasma membrane using PALM and pair correlation analysis. *Nature methods*. 2011; 8:969–975. [PubMed: 21926998]
 19. Gould TJ, Verkhusha VV, Hess ST. Imaging biological structures with fluorescence photoactivation localization microscopy. *Nature protocols*. 2009; 4:291–308.
 20. van de Linde S, et al. Direct stochastic optical reconstruction microscopy with standard fluorescent probes. *Nature protocols*. 2011; 6:991–1009.
 21. Sengupta P, Lippincott-Schwartz J. Quantitative analysis of photoactivated localization microscopy (PALM) datasets using pair-correlation analysis. *Bioessays*. 2012; 34:396–405. [PubMed: 22447653]
 22. Annibale P, Vanni S, Scarselli M, Rothlisberger U, Radenovic A. Identification of clustering artifacts in photoactivated localization microscopy. *Nature methods*. 2011; 8:527–528. [PubMed: 21666669]
 23. Annibale P, Vanni S, Scarselli M, Rothlisberger U, Radenovic A. Quantitative photo activated localization microscopy: unraveling the effects of photoblinking. *PloS one*. 2011; 6:e22678. [PubMed: 21818365]
 24. Kiskowski MA, Hancock JF, Kenworthy AK. On the use of Ripley's K-function and its derivatives to analyze domain size. *Biophysical journal*. 2009; 97:1095–1103. [PubMed: 19686657]
 25. Owen DM, et al. PALM imaging and cluster analysis of protein heterogeneity at the cell surface. *J. Biophotonics*. 2010; 3:446–454. [PubMed: 20148419]
 26. Williamson DJ, et al. Pre-existing clusters of the adaptor Lat do not participate in early T cell signaling events. *Nature immunology*. 2011; 12:655–662. [PubMed: 21642986]
 27. Dickson RM, Cubitt AB, Tsien RY, Moerner WE. On/off blinking and switching behaviour of single molecules of green fluorescent protein. *Nature*. 1997; 388:355–358. [PubMed: 9237752]
 28. Schwille P, Kummer S, Heikal AA, Moerner WE, Webb WW. Fluorescence correlation spectroscopy reveals fast optical excitation-driven intramolecular dynamics of yellow fluorescent proteins. *Proceedings of the National Academy of Sciences of the United States of America*. 2000; 97:151–156. [PubMed: 10618386]
 29. Ha T, Tinnefeld P. Photophysics of fluorescent probes for single-molecule biophysics and super-resolution imaging. *Annu Rev Phys Chem*. 2012; 63:595–617. [PubMed: 22404588]
 30. Smith CS, Joseph N, Rieger B, Lidke KA. Fast, single-molecule localization that achieves theoretically minimum uncertainty. *Nature methods*. 2010; 7:373–375. [PubMed: 20364146]
 31. Holden SJ, Uphoff S, Kapanidis AN. DAOSTORM: an algorithm for high-density super-resolution microscopy. *Nature methods*. 2011; 8:279–280. [PubMed: 21451515]
 32. Huang F, Schwartz SL, Byars JM, Lidke KA. Simultaneous multiple-emitter fitting for single molecule super-resolution imaging. *Biomed Opt Express*. 2011; 2:1377–1393. [PubMed: 21559149]
 33. Quan T, et al. High-density localization of active molecules using Structured Sparse Model and Bayesian Information Criterion. *Opt Express*. 2011; 19:16963–16974. [PubMed: 21935056]
 34. Cox S, et al. Bayesian localization microscopy reveals nanoscale podosome dynamics. *Nature methods*. 2012; 9:195–200. [PubMed: 22138825]
 35. Wang Y, Quan T, Zeng S, Huang ZL. PALMER: a method capable of parallel localization of multiple emitters for high-density localization microscopy. *Opt Express*. 2012; 20:16039–16049. [PubMed: 22772294]
 36. Veatch SL, et al. Correlation Functions Quantify Super-Resolution Images and Estimate Apparent Clustering Due to Over-Counting. *PloS one*. 2012; 7
 37. Veatch SL, Chiang EN, Sengupta P, Holowka DA, Baird BA. Quantitative Nanoscale Analysis of IgE-FcεpsilonRI Clustering and Coupling to Early Signaling Proteins. *J Phys Chem B*. 2012; 116:6923–6935. [PubMed: 22397623]
 38. Tanaka KA, et al. Membrane molecules mobile even after chemical fixation. *Nature methods*. 2010; 7:865–866. [PubMed: 20881966]

39. Testa I, Mazza D, Barozzi S, Faretta M, Diaspro A. Blue-light (488 nm)-irradiation-induced photoactivation of the photoactivatable green fluorescent protein. *Appl Phys Lett*. 2007; 91:133902–133904.
40. Subach FV, et al. Photoactivatable mCherry for high-resolution two-color fluorescence microscopy. *Nature methods*. 2009; 6:153–159. [PubMed: 19169259]

Figure 1.

Work flow for PC-PALM analysis. A pointillist, super-resolution image is constructed by combining the localizations from all the frames of the PALM image series. Next, a region of interest (ROI, red box) ($4\text{--}16\ \mu\text{m}^2$) is chosen for correlation analysis. Finally, the autocorrelation function of peaks in the ROI is measured (black circles), and fit (red line) to an appropriate model to obtain estimates of cluster size, number of proteins in clusters, and the density of proteins in clusters. The protein correlation function and the correlation function corresponding to multiple appearances of a protein are extracted from the fitting procedure, and are represented as green and blue circles, respectively. Images adapted from reference (18).

Figure 2. Spatial organization of plasma membrane proteins expressed in COS-7 cells evaluated by PC-PALM analysis. Distribution of cluster radius (a), distribution of proteins per cluster (b), and the increased local density (c) for PAGFP-GPI, VSVG-PAGFP, Lyn-PAGFP and Lat-PAGFP estimated by PC-PALM analysis of composite PALM images. (d) Cartoon of steady state lateral distribution of four plasma membrane proteins generated on the basis of insights obtained from PC-PALM analysis. Images adapted from reference (18).

Table 1**Common Advantages and Limitations of PC-PALM Approach**

Advantages	Limitations
Uses robust statistical algorithm to circumvent artifacts associated with multiple appearances of a single molecule	Cluster parameters are estimated based on detected (not absolute) number of proteins
Provides precise estimation of size, density and protein abundance in protein domains	Involves radial averaging, and as such, is not optimal for accurate characterization of asymmetric, elongated protein domains
Can quantify spatial organization over a wide spatial scale (25nm – microns)	Errors in measurement of average number of appearances can lead to inaccurate estimation of protein abundance in clusters
Can be used to characterize spatial organization of high density of proteins	Time consuming
Employs genetically encoded fluorescent protein reporters with precise 1:1 labeling ratio	Requires skilled user with substantial knowledge of statistics, single molecule photophysics and computer programming

Table 2

Troubleshooting

Step	Problem	Possible reason	Solution
1	High background fluorescence	Coverslips not cleaned properly, Phenol red was included in the medium	Follow the cleaning protocol stringently and avoid additives such as Phenol red in the medium
6	Fiducial density too low	Cells are too dense	Decrease amount of plated cells
		Number of fiducial markers is too low	Add additional buffer with fiducials and/or incubate for longer times
9–11	Well dispersed cells hard to find	High plating density of the cells	Decrease amount of plated cells
	Many holes appear in the image of the cell	Cells did not adhere well to the coverslip	Optimize fibronectin coating, try different agents for cell adhesion
	Little or no signal in the image	Transfection efficiency not optimal	Optimize transfection by varying various conditions (transfection reagent, DNA amount, incubation time etc.)
		Imaging parameters are not optimal	Adjust imaging parameters such as laser power, EM gain, etc.
	Lateral or axial drift	Air flow going onto the objective	Prevent the air flowing onto objective
		Immersion oil contaminated	Replace the immersion oil
13	Many molecules are localized with precision >25 nm	High fluorophore density in one frame	Adjust imaging parameters to reduce number of activated molecules
		Molecules do not emit enough photons	Increase the laser power, adjust the focus
20A (iii)	Difficult to determine the average number of appearances of peaks	Selected region has high fluorophore density	Reduce the amount of PA-FP on coverslip or select cell with lower expression level