RSC Advances

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/advances

Journal Name

ARTICLE

Super-resolution imaging of STAT3 cellular clustering during the nuclear transport

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

www.rsc.org/

Jing Gao,[a] Feng Wang,[c] Junling Chen,[a,b] Jianzhong Wang,[d] Mingjun Cai,[a] Haijiao Xu,[a] Junguang Jiang,[a] and Hongda Wang*[a]

Signal transducer and activator of transcription 3 (STAT3) mediates signals from the plasma membrane to the nucleus, and participate in gene expression in response to cytokines and growth factors. STAT3 translocation to the cell nucleus and subsequent recycling to the cytoplasm is the critical step for multiple functions. However, the distribution and molecular composition of STAT3 during the nuclear transport upon activating by different stimuli remains largely elusive. By direct stochastic optical reconstruction microscopy (dSTORM), we visualized the nucleocytoplasmic distribution of STAT3 in HeLa cells following different stimuli. Our work reveals that STAT3 is preclustered in the cytoplasm before activation, and that interleukin-6 promotes the formation of more and larger clusters. Further analysis indicates that STAT3 clusters vary with the nuclear translocation of STAT3. Moreover, dual-color dSTORM imaging shows the colocalization of STAT3 and STAT1 clusters, and suggests that the ratios of their colocalization are associated with STAT functions.

Introduction

Various stimuli can induce the activation of transcription factors that lead to gene expression in the nucleus. Signal transducers and activators of transcription (STATs) are latent in the cytoplasm until they are activated by cytokines and growth factors [1-2]. In response to stimulation, phosphorylated STATs undergo dimerization, subsequently move to the nucleus, bind to DNA, and activate the transcription of their target genes [2-3]. There are seven mammalian STAT proteins (STATs 1, 2, 3, 4, 5A, 5B, and 6) that mediate distinct activating signals and elicit a wide spectrum of physiological outcomes [4-5]. As a key representative, STAT3 is activated by cytokines such as interleukin-6 (IL-6), a cytokine that is well-studied in the JAK/STAT signaling pathway [6-7], as well as ligands that bind to receptor-tyrosine kinases e.g. epidermal growth factor (EGF) [8]. STAT3 plays pivotal roles in embryo implantation [9], development [10], acute phase reactions [11], and immune responses [12]. Because of its diverse functions, aberrant activities of STAT3 can lead to many diseases, even certain neoplasias [13].

As a critical mediator for transmitting signals between the plasma membrane and nucleus, STATs need to be translocated to the nucleus from cytoplasm and then return to the cytoplasm. Thus, trafficking behavior is crucial for the functions of STAT proteins and might offer a new type of therapy to cure STAT-induced diseases [14]. Some earlier studies proposed that inactive STATs reside in the cytoplasm as isolated monomers and then dimerize until they are phosphorylated [1]. However, the recent statosome model predicts that STATs exist as dimers or high-molecular-weight complexes (statosome) with other proteins whose sizes are in the range of 200–400 kDa and 1–2 MDa [15-16]. Although the activation mechanism and functional role of STATs have been intensely investigated by biological experiments, the molecular composition and distribution of STATs during the nuclear transport after activating by stimuli is still unknown.

Recently, a number of technologies have emerged to solve similar issues [17-19]. One of them is super-resolution fluorescence microscopy, which has surpassed the light diffraction limit of ~200 nm, allowing the observation of biological processes and cellular structures at nanometer scales [20-21]. It can be divided into two categories: one is based on spatially patterned illumination, such as stimulated emission depletion microscopy [22-23] or saturated structured-illumination microscopy [24]; the other is based on single-molecule localization of individual fluorescent molecules, including (direct) stochastic optical reconstruction microscopy ((d)STORM) [25-27] and photoactivation localization microscopy [28-29]. Because of the high maneuverability of these instruments and the emergence of many commercial fluorescent dyes, (d)STORM has been widely used in many fields [30-32]. The improvement of spatial resolution enables high-precision localization of proteins and the elucidation of many detailed structures or features of multi-protein complexes [33-36].

Therefore, based on our earlier study about the relationship between the cellular localization of STAT1 and the cell cycle [37], we further located STAT3 during the nuclear trafficking under IL-6 or EGF stimulation by the super-resolution imaging technique

[a] State Key Laboratory of Electroanalytical Chemistry, Changchun Institute of Applied Chemistry, Chinese Academy of Sciences, Changchun, China. E-mail: hdwang@ciac.ac.cn
[b] University of Chinese Academy of Sciences, Beijing, China.
[c] Institute of Immunology, The First Bethune Hospital Academy of Translational Medicine, Jilin University, Changchun, China.
[d] School of Computer Science and Information Technology, Northeast Normal University, Changchun, China.
Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 10.1039/x0xx00000x

(dSTORM) in this work. Our data directly verified the statosome model that unactivated STAT3 existed as clusters in the cytoplasm, and these clusters became larger and were transported into the nucleus after IL-6 stimulation. Further statistical analysis indicated that these clusters changed as STAT3 was translocated into and out of the nucleus. Moreover, the dual-color dSTORM imaging revealed the colocalization of EGF-activated STAT3 and STAT1 clusters in the whole cell.

Experimental section

Cell culture

Human HeLa cells (purchased from Shanghai Institute of Biological Sciences) were grown in Dulbecco's modified Eagle's medium (DMEM, Hyclone) supplemented with 10% fetal calf serum (Biochrom AG, Germany) and 1% penicillin/streptomycin (Invitrogen) in a humidified 5% CO_2 atmosphere at 37℃. Cells were passaged every two or three days.

Antibody labeling

Anti-STAT1 antibody (a mouse monoclonal antibody epitope mapping between amino acids 613-739 of STAT1α p91 of human origin, C-111) and anti-STAT3 antibody (a mouse monoclonal antibody raised against amino acids 50-240 mapping at the N-terminus of STAT3 p92 of human origin, F-2) were purchased from Santa Cruz Biotechnology, Inc. (Santa Cruz, CA). At first, antibodies were respectively labeled with Alexa Fluor 647 or Alexa Fluor 532 (Invitrogen) in an appropriate concentration. 100 μl anti-STAT3 or STAT1 (100 μg/ml) antibodies were stained by 2 μl Alexa Fluor 647 or Alexa Fluor 532 (1 mg/ml, dissolved in DMSO) and shook for 2 hours in dark at room temperature. To remove excess dyes, the solution was filtered out by gel filtration using illustra NAP-5 columns (GE Healthcare). The absorbance of each tube at 280 nm (antibodies) and 650 nm (Alexa Fluor 647) or 532 nm (Alexa Fluor 532) was measured by absorption spectroscopy assay. Tubes in which the labeling ratio of Alexa Fluor 647 (Alexa Fluor 532) and antibody was between 0.7 and 1 dye/protein were pooled for use.

Cytokine treatment and sample preparation

Prior to cell plating, standard microscope slides were sonicated for ~15 min in 1% Micro-90 concentrated cleaning solution (Sigma-Aldrich), rinsed thoroughly in Milli-Q water and sterilized in the biosafety cabinet under ultraviolet light for 30 min. HeLa cells were seeded onto the pre-cleaned slides for approximate 24 h to achieve a ~60% confluence. To observe nuclear transport of STAT3, cells were treated with 50 ng/ml IL-6 (Millipore) for 0, 20, 40, 60, 120 and 240 min, respectively. Activation of both STAT3 and STAT1 was done by adding 50 ng/ml EGF (Peprotech) for 30 min. Then cells following different stimulation time were fixed in 4% paraformaldehyde (Fisher) for 10 min at the room temperature, washed by PBS for three times, permeablized with 0.1% Triton X-100 (Roche) for 10 min, and washed by PBS for three times again. Prepared cell samples were blocked by incubating in 1% BSA for 30 min. After washing out the blocking buffer by PBS for three times, cells were stained with 50 μl Alexa Fluor 647-conjugated STAT3 antibodies (and 50 μl Alexa Fluor 532-conjugated STAT1 antibodies) for 40 min in dark at room temperature. After that, cells were incubated with 1: 500 dilution of Hoechst 33342 (Sigma-Aldrich) for 10 min, and washed out the staining solutions three times with PBS.

Super-resolution imaging

STORM imaging was performed on a home-built microscope equipped with appropriate lasers. Samples were covered with the STORM imaging buffer containing Tris (50 mM, pH 8.0), NaCl (10 mM), glucose (10% w/v), glucose oxidase (500 μg/ml; Sigma), catalase (40 μg/ml; Sigma), β-ME (1% v/v; Sigma) for single color imaging or MEA (10% v/v, 1 M; Sigma) for dual-color imaging. Samples were observed on a Nikon Ti-E microscope equipped with an oil-immersion objective (100×1.49NA lens; Nikon, Japan) and imaged on an electron-multiplying charge-coupled device (EMCCD) camera (Photometrics, Cascade II). The illuminated area was 256×256 pixels with the pixel size of 160 nm. Excitation was provided by a 640 nm laser line (100 mW) and a 532 nm laser line (100 mW) with AOTF-based intensity control. To produce the photoswitching necessary for dSTORM imaging, moderately high powers were used, typically between 20 and 30 mW on the sample. In the range of 2-4 μm above the bottom of the cell, a single fluorescence image of the nucleus was firstly obtained by 405 nm laser, and then at the same imaging depth, dSTORM images of STAT3 or STAT1 were captured with oblique illumination to reduce as much background as possible. 640 nm laser was used to acquire the Alexa Fluor 647 signal for imaging STAT3 and 532 nm laser was used to acquire the Alexa Fluor 532 signal for imaging STAT1 if needed. For Alexa Fluor 647 and Hoechst 33342 signals, an excitation filter (ZET405/488/561/647x, Chroma), a dichromic mirror (ZT405/488/561/647rpc, Chroma) and an emission filter (ZET405/488/561/640m, Chroma) were set in a beam path. For Alexa Fluor 647 and Alexa Fluor 532 signals, an excitation filter (ZET532/647x, Chroma) and the corresponding dichromic mirror and emission filter were set in a beam path. Typically, 5000 images were captured for each cell with 40 ms integration time per frame. X-Y drift and alignment differences between different channels were corrected by localizing 100 nm TetraSpeck microspheres (Invitrogen) immobilized on the sample coverslip. The average deviation between Alexa Fluor 647 and Alexa Fluor 532 channels was 8-10 nm.

Image reconstruction

A freely available plug-in for Image J named quickPALM [38] was used to analyze raw images. Raw data were first preprocessed to detect single-molecule events via background subtraction with a minimum SNR of 2–4. Then fluorescence peaks were identified in each frame and fitted a least square fit with an elliptical Gaussian function. Individual least-squares fit estimates were performed by a threshold of the peak height and the peak widths in the two lateral dimensions. After rejecting the poor fit and asymmetric PSFs, the coordinates of detected molecules were determined by the centers of gravity of their PSFs. A super-resolution image was generated as a density map using the precise localization data of single fluorescent molecules. The merging images of the two or three channels were established via Image-Pro Plus 6.0 (Media Cybernetics, Inc.).

Measurement of localization precision

Single molecule localization precision was analyzed to detect the resolution on our home-built instrument. Adequately diluted Alexa Fluor 647-conjugated anti-STAT3 antibodies (about 10 nM) were added onto the prepared cell surface, incubated for 40 min, and imaged. Similarly, 5000 images were recorded with 40 ms integration time per frame to generate a spot localization map. From repetitively switching fluorophores, the localization precision was 29 nm for Alexa Fluor 647-STAT3 antibody by measuring the full-width at half-maximum (FWHM) of the localization distribution of individual fluorophores (Fig. S1). Meanwhile, the average number of localized spots in a single Alexa Fluor 647-labeled STAT3 antibody was calculated, which was 27.8±5.1 (s.d.) (100 dyes analyzed).

Cluster analysis by Ripley's K-function

To analyze the spatial distribution of STAT3 during the entry to the nucleus, Ripley's K-function was applied to characterize STAT3 clustering based on the localization data. 4×4 μm^2 examined regions in the reconstructed dSTORM images were randomly selected. Ripley's K-function [39] is then calculated as:

$$K(r) = \frac{A}{N^2}\sum_{i=1}^{N}\sum_{j=1, j\neq i}^{N}\delta_{ij} \quad (1)$$

Where A is the image area, N is the number of total localizations in the area, r is the spatial scale (radius) for the K-function calculation and δ_{ij} is the distance between the i-th and the j-th points. Here, if δ_{ij} is less than r, the value will be one, otherwise δ_{ij} =0. The linear transformation of K(r), namely H-function [40], is used to interpret the spatial randomness:

$$H(r) = L(r) - r = \sqrt{K(r)/\pi} - r \quad (2)$$

The amplitude of H(r) will be zero for particles with a random distribution, and positive for clustering particles. Edge-effects were eliminated by weighting edge points and cropping image edges after the calculation. The values of L(r) generated by each particle were used to produce a cluster map by interpolating a surface plot with L(r) as the z-axis. Then a binary cluster map was generated through an appropriate L(r) threshold, that is particles with their values of L(r)-r>0 are selected. Finally the information of clustering could be extracted from the binary map, such as the number and the size of clusters. All calculations and image processing were performed in Matlab.

Co-cluster analysis using a combined univariate and bivariate G&F method

To investigate the localization relationship between STAT3 and STAT1 during nuclear transport, a combined univariate and bivariate Getis and Franklin's local point pattern analysis method was used to quantify the co-clustering of the two species. The value of L(r) or L(r)cross is given by:

$$L(\mathrm{r})_j = \sqrt{A\sum_{i=1}^{N}\left(\frac{\delta_{ij}}{N}\right)/\pi} \quad (3)$$

where A is the image area, N is the number of total localizations in the area, r is the spatial scale (r = 100 nm in this experiment), and δ_{ij} is the distance between the i-th and the j-th points. If δ_{ij} is less than r, the value will be one, otherwise δ_{ij} =0. Values of L(r) are calculated for one species by counting the number of its own species within the circle of radius r and values of L(r)cross are calculated for one species by counting the number of the different species within the circle of radius r.

The L(100) and L(100)cross values for each molecule were plotted as a scatter plot and the color code denoted the number of molecules at the position. Linear trend lines were generated using linear least-squares fitting in Matlab. Cluster thresholds of 100 for nuclear areas and 150 for cytoplasmic areas were applied to divide the plot into four quadrants which indicated the percentage of co-clustering, individual clustering and random distribution of the two species.

Results and discussion

Super-resolution imaging of IL-6-activated STAT3 nuclear transport

To know the nuclear transport of STAT3, we located STAT3 in HeLa cells during IL-6 stimulation with time. Fig. 1 shows the localization of STAT3 in the whole cell before and after IL-6 stimulation for 20, 40, 60, 120 and 240 min. The reconstructed dSTORM images of STAT3 labeled with Alexa Fluor 647-conjugated STAT3 antibodies (Fig. 1, the first column) and the single fluorescent images of the corresponding nucleus stained with Hoechst 33342 (Fig. 1, the second column) were independently acquired; the merged channels of STAT3 and the nucleus are shown in the third column. The serial images clearly show the translocation of STAT3 in and out of the nucleus following IL-6 stimulation. Nuclear STAT3 increased significantly from 0 to 20 min after IL-6 addition, remained stable during 20–40 min, and gradually decreased after 60 min up to 240 min (Fig. 1, the last column). This result indicates that tyrosine phosphorylated STAT3 is rapidly transported into the nucleus and then slowly exports from the nucleus following continued stimulation. Interestingly, there were always some STAT3 in the cytoplasm during the stimulation, demonstrating that IL-6 only induced some but not all STAT3 to enter the nucleus (Fig. 1, the fourth column). This result agrees with previous studies that STAT3 is not only present in the nucleus, but is continuously shuttled between the nucleus and cytoplasm as well, which is independent of tyrosine phosphorylation [41-42]. Therefore, we showed that STAT3 had still existed in the cytoplasm even after IL-6 stimulation for a relative long time.

To further accurately analyze the localization of STAT3 during the stimulation, we quantitatively measured STAT3 expression level in the cytoplasm and the nucleus using reconstructed dSTORM images at different time points. In general, when fluorescent dyes attach to a type of antibodies, the number of localizations per dye is a constant value under the same experimental conditions, such as excitation, image capture, and data processing. Therefore, the number of localizations, to a large extent, can represent the amount of labelled proteins when the label ratio of dye and antibody is close to one (the ratio in our experiment ranged from 0.7 to 1). We determined the average number of localizations with

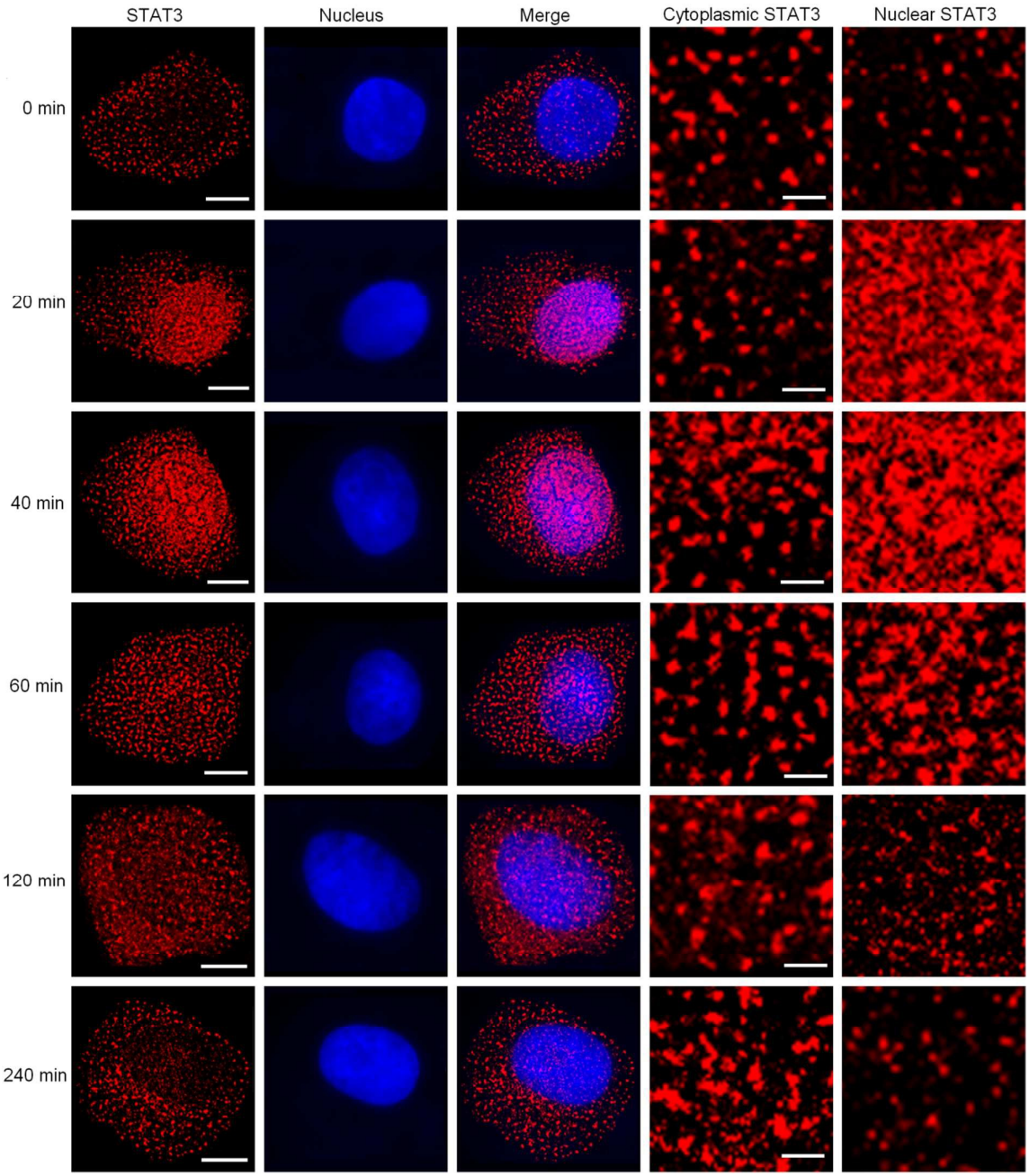


Fig. 1 IL-6-induced STAT3 translocation from the cytoplasm to the nucleus. STAT3 was labeled with Alexa Fluor 647-conjugated STAT3 antibodies and the nucleus with Hoechst 33342. The reconstructed dSTORM images of STAT3 in HeLa cells without IL-6 stimulation and with 20, 40, 60, 120, and 240 min IL-6 stimulation are shown in the first column. The single fluorescent images of the corresponding nucleus are displayed in the second column, and the merged images of STAT3 and the nucleus in the third column. The typical zoom-in views of cytoplasmic and nuclear STAT3 are listed in the fourth and fifth columns respectively. Each scale bar represents 10 μm in the first three columns, and 1 μm in the last two columns.

a single Alexa Fluor 647-conjugated STAT3 antibody to be 27.8±5.1 (mean±s.d.) by quantitative 100 dyes (see Experimental section and Fig. S1 for detail), which could be used to estimate STAT3 density in different locations of the cell as described [37]. The normalized total localization number of STAT3 at different IL-6 stimulation time was essentially steady with a slight increase from 0 to 20 min and a slow decrease after 40 min (Fig. 2A). However, the ratio of nuclear: cytoplasmic STAT3 localizations fluctuated sharply (Fig. 2B). In 0-20

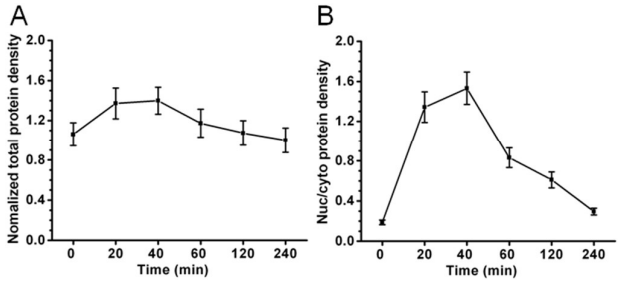


Fig. 2 Time course analysis of STAT3 localization in HeLa cells after IL-6 stimulation. (A) Normalized total localization number of STAT3 proteins in HeLa cells with different IL-6 stimulation times. (B) The ratio of nuclear to cytoplasmic (nuc/cyto) localizations of STAT3 proteins in HeLa cells with different stimulation times. Each data set was obtained from 20 cell samples in five independent experiments. Values are means±s.d.

min, the proportion of nuclear STAT3 localizations went up dramatically from 0.2 to 1.3. In 20–40 min, it continued to increase but relatively slow, and reached a peak 1.5 at approximately 40 min. After this point, the ratio began to decline until it returned to the unstimulated state. The statistic data together with the dSTORM images reveal that STAT3 could rapidly respond to IL-6 activation and enter the nucleus, and then gradually export from the nucleus. The total quantity of STAT3 in the whole cell did not change after IL-6 stimulation, suggesting that the fluctuation of nuclear STAT3 was due to its translocation.

Characterization of IL-6-induced STAT3 clustering

We next attempted to determine what forms of STAT3 exist during IL-6-activated STAT3 nuclear trafficking. Previous studies have indicated that phosphorylation of STAT3 by cytokines or growth factors can generate induce STAT3 homodimers or heterodimers with other STATs, and these STAT3 complexes move from the cytoplasm to the nucleus. Thus, we hypothesized that activated STAT3 may form clusters just as STAT1 which was described in our earlier study [37]. To confirm the validity of this hypothesis, we determined the distribution characteristics of STAT3 using Ripley's K-function analysis (see Experimental section, Eqn 1), which is a widely used spatial statistic method to test randomness or aggregation for spatial point patterns [39-40] and is routinely applied to analyze protein heterogeneity for super-resolution data [20, 43-44]. Here, we used the transformation of H-function (see Experimental section, Eqn 2), where the larger value of H corresponds to the region of more aggregation, and the value of r corresponding to the maximum of H, is considered to approximate to the cluster size in the region [45]. As shown in Fig. S2, a 4×4 μm^2 region of the reconstructed dSTORM image in the cytoplasm of a HeLa cell without stimulation was taken to exemplify the use of Ripley's K-function. According to Eqn 1 and 2, the H-function plot was obtained, which illustrated that the r value of maximal aggregation was approximately 190 nm with a clustering range stretched to 500 nm (above the level of a random distribution). Simultaneously, the

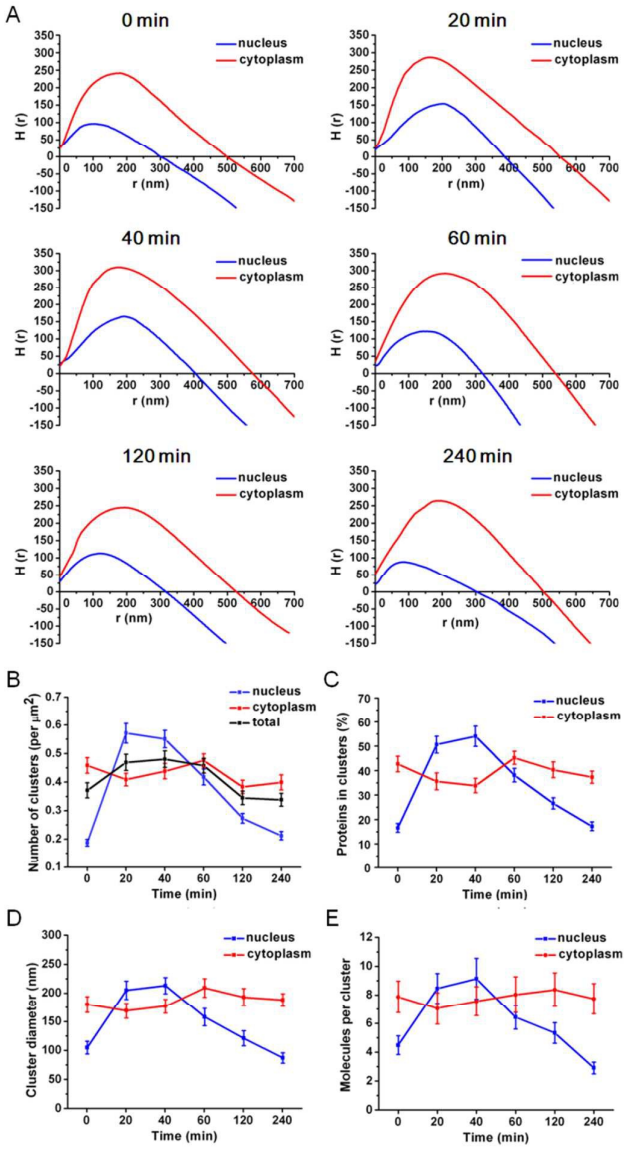


Fig. 3 Cluster analysis of nuclear transport of IL-6-induced STAT3 using Ripley's K-function. (A) Ripley's K-function analysis of STAT3 protein clustering. Regions (4×4 μm^2) of reconstructed dSTORM images of STAT3 in both the nucleus and cytoplasm of HeLa cells with the different stimulation times were calculated. Plots shown are representatives of 120 analyzed regions from 20 cells. (B) The number of STAT3 clusters per μm^2 in the nucleus, cytoplasm and whole cell at different stimulation times. (C) The percentage of STAT3 proteins participating in clusters in the nucleus and cytoplasm at different stimulation times. (D) Changes in the average STAT3 cluster diameter in the nucleus and cytoplasm. (E) The average amount of STAT3 protein per cluster in the nucleus and cytoplasm. Each data set in (B)–(E) were from 20 cells in five independent experiments. Every cell sample was chosen 6 regions (3 in the cytoplasm and 3 in the nucleus). Values are means±s.d.

pseudo-colored cluster map was also established by interpolating the surface plot with L(r) of every point as the z value. Finally, the binary cluster map was obtained through an appropriate L(r) threshold, from which the amount, diameter, shape and other parameters of clusters could be extracted.

We then analyzed different regions of STAT3 in both the cytoplasm and nucleus of HeLa cells following different IL-6 stimulation time. Fig. 3A shows the representative plots of Ripley's K-function analysis, indicating that STAT3 formed clusters in the whole cell during the stimulation. As mentioned above, we extracted the information of clusters including cluster amount per μm^2 (Fig. 3B), the percentage of proteins participating in clusters (Fig. 3C), and cluster diameter (Fig. 3D) at the different stimulation times according to binary cluster maps. We found that the cluster density in the whole cell was initially high and then reduced with stimulation time (Fig. 3B, the black line), which indicated that IL-6 activation promoted the formation of STAT3 clusters. We next observed the change of this parameter in the nucleus and cytoplasm, respectively. In the nucleus, cluster density increased sharply during 0–20 min stimulation, remained steady from 20 to 40 min, and gradually decreased to unstimulated levels from 40 to 240 min. In contrast to the nucleus, cluster density in the cytoplasm firstly decreased and then increased. However, the increase of nuclear cluster number during 0–20 min stimulation was more than the decrease of cytoplasmic cluster number, indicating that not only cytoplasmic STAT3 was translocated to the nucleus, but also new STAT3 clusters were generated in the nucleus after stimulating by IL-6. As stimulation continued, nuclear STAT3 gradually exported out of the nucleus, and the distribution of STAT3 clusters returned to the initial state. Variations in the percentage of proteins participating in clusters were consistent with changes in cluster number, confirming the above view. Meanwhile, cluster size also changed significantly during this process. For nuclear areas, the average cluster diameter was merely 100 nm before IL-6 activation, increased to more than double (approximately 210 nm) after stimulating for 20–40 min, and finally reduced to 100 nm. However, cytoplasmic cluster diameter only fluctuated moderately. Before IL-6 addition, there was a marked difference between the cytoplasmic and nuclear cluster size because STAT3 proteins have not been translocated to the nucleus. With IL-6 stimulation, the cluster diameter in the cytoplasm was smaller than that in the nucleus at 20 and 40 min, suggesting that IL-6 induced the formation of large aggregations in the nucleus and that some cytoplasmic clusters might enter the nucleus. After that, STAT3 began to export out of the nucleus, causing cluster size in the cytoplasm to become larger than that in the nucleus again. These results suggest that the changes in cluster size are associated with the nuclear translocation of STAT3 after IL-6 stimulation. Meanwhile, the parameters of Ripley's K-function plots listed in Table S1, r_{ave} (the radius of maximal clustering), $H(r)_{max}$ (clustering degree) and r_{max} (clustering range), were in accord with this trend change.

Based on the localization number of each cluster, we semi-quantitatively estimated the mean number of proteins in both cytoplasmic and nuclear clusters during the stimulation (Fig. 3E). This value was calculated by dividing the total localization number inside a cluster by the average localization number of single Alexa Fluor 647-conjugated STAT3 antibody. Similar to other parameters, the average protein amount of cytoplasmic clusters remained steady at approximately 8 (whether stimulated or not); however, the value of nuclear clusters increased from 4 to over 8 after IL-6 activation. An earlier study using fluorescence correlation spectroscopy in live Hep3B cells expressing STAT3-green fluorescent protein reported that IL-6 stimulation induced the generation of high-molecular-weight complexes of STAT3 (>1 MDa) in the cytoplasm [16]. Our findings on the formation of STAT3 clusters further verify this view and suggest that these clusters may be oligomers consisting of approximately two to four dimers.

Overall, STAT3 distributes in clusters in HeLa cells. Without stimulation, STAT3 mainly aggregates into clusters in the cytoplasm, which may facilitate the rapid formation of dimers because of the closer distance between STAT3. After IL-6 addition, there is a dramatic increase in cluster number and size in the nucleus, indicating that STAT3 begins to enter the nucleus to form new and larger clusters. Over time, the cluster number in the nucleus decreases gradually, indicating the export of nuclear STAT3. Many intracellular processes are mediated by vesicles [46-47], and it has reported that STAT3 is targeted to endosomes and associated with other vesicular elements in response to IL-6 or growth factor stimulation [48-49]. Hence, it is possible, although untested here, that the nuclear transport of STAT3 clusters also takes place along the vesicular pathway.

Visualization of the distribution of STAT3 and STAT1 after EGF stimulation

In addition to cytokines such as IL-6, STAT3 can be activated by growth factors to form homodimers and heterodimers simultaneously. EGF, as a key member of growth factors, can exert a variety of effects by activating STAT pathway including cell growth, differentiation and survival. Therefore, we further investigated the distribution of two STAT family members, STAT3 and STAT1, following EGF stimulation. Fig. 4 shows dual-color dSTORM images of STAT3 and STAT1 in HeLa cells without (Ctrl) and with 30 min-EGF stimulation (EGF). A significant amount of nuclear accumulation of both STAT3 and STAT1 was observed after EGF stimulation. Preliminary observations of the merged images revealed that STAT3 and STAT1 have varying degrees of colocalization whether stimulated or not. In control HeLa cells, a portion of STAT3 was associated with STAT1 in the cytoplasm, while there was scarcely colocalization in the nucleus. In EGF-stimulated cells, nuclear colocalization was obvious, whereas the change in the cytoplasm could not be quantitatively determined. Despite this, we could observe that there was a colocalization relationship between STAT3 and STAT1 in the whole cell, particularly in the nucleus following EGF stimulation, which is consistent with previous studies [2]. The imaging mode of STAT3 and STAT1 used in dSTORM was oblique illumination and 3D fluorescence imaging was performed to rule out the possibility that there was a certain overlap of nuclear and cytoplasmic regions during the acquisition of dSTORM images. The 3D fluorescence data (Fig.S3) was similar to the results of dSTORM images as expected, although they could not provide a clear enough organization of STAT3 and STAT1 due to the resolution of fluorescence microscopy.

The interesting observation is that the distribution pattern of EGF-

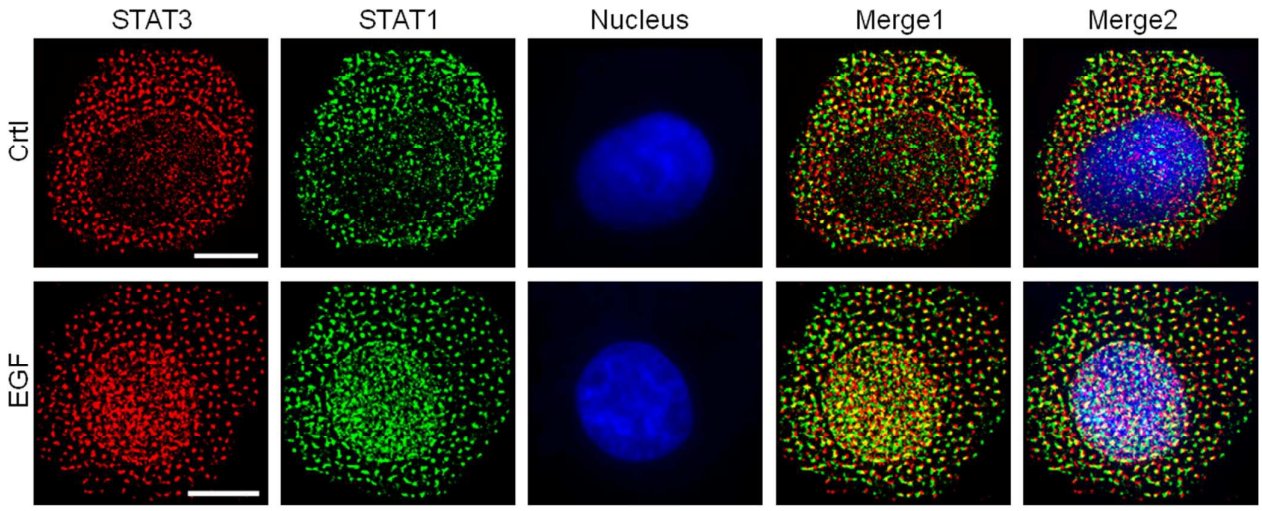


Fig. 4 dSTORM imaging showing colocalization of STAT3 and STAT1 in HeLa cells before and after EGF stimulation. STAT3 was labeled with Alexa Fluor 647-conjugated STAT3 antibodies, STAT1 with Alexa Fluor 532-conjugated STAT1 antibodies, and the nucleus with Hoechst 33342. dSTORM images of STAT3 and STAT1 are shown in the first two columns and the corresponding nucleus image is shown in the third column. Merged panels of STAT3 and STAT1 are shown in the fourth column (Merge 1). Merged panels of STAT3, STAT1 and the nucleus are shown in the last column (Merge 2). Each scale bar represents 10 μm.

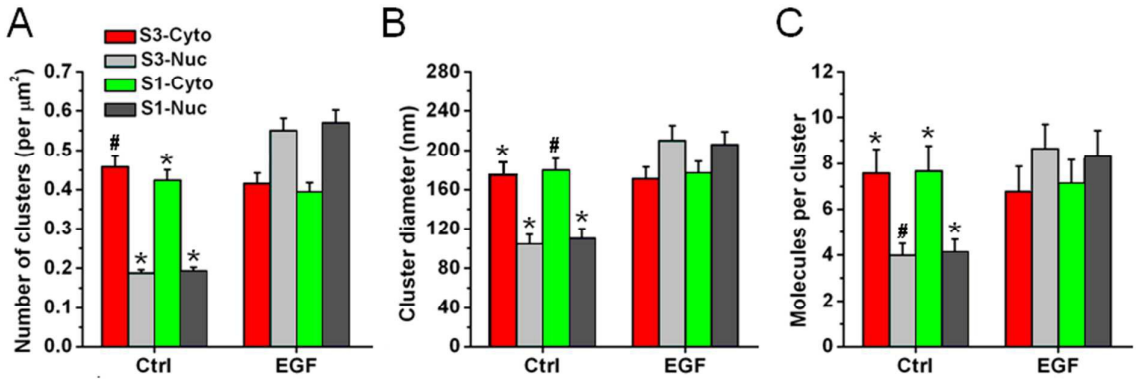


Fig. 5 Clustering properties of EGF-activated STAT3 and STAT1. (A) The number of STAT3 and STAT1 clusters per μm^2 in the nucleus and cytoplasm of control and EGF-stimulated HeLa cells. (B) The average cluster diameter of STAT3 and STAT1 in the nucleus and cytoplasm with and without EGF stimulation. (C) The average number of molecules in one STAT3 or STAT1 cluster in the nucleus and cytoplasm before and after EGF addition. Each data set were from 60 regions of 20 cells in five independent experiments. Every cell sample was chosen 6 regions (3 in the cytoplasm and 3 in the nucleus). Values are means±s.d. **P*<0.05, #*P*<0.01, Ctrl vs EGF, two-tailed unpaired *t*-test.

activated STAT3 and STAT1 is similar to that of IL-6-induced STAT3. Therefore, we examined the clusters of STAT3 and STAT1 using Ripley's K-function. The H-function parameters (Table S2) of both STAT members confirmed the clustering of STAT3 and STAT1 in the cytoplasm and nucleus either before or after EGF stimulation. This finding implies that clustering during the nuclear transport may be universal to all members of the STAT family. On the basis of Ripley's K-function analysis, the amount and size of clusters were extracted as shown in Fig. 5A and B. Without EGF activation, cluster number in the cytoplasm was two times more than that of the nucleus for both STAT3 and STAT1. However, the results were the opposite after EGF activation, with more nuclear cluster amount than cytoplasmic one. When focusing on the cytoplasmic cluster amount, the values of both STAT3 and STAT1 only modestly reduced after EGF treatment. However, with respect to the nuclear cluster number, a dramatic increase was observed following EGF stimulation. Moreover, the cluster size of STAT3 and STAT1 had a similar trend with cluster number. Finally, we estimated the

number of molecules in single clusters using the aforementioned method (Fig. 5C). There was only a slight decline in the value of cytoplasmic clusters for both STAT3 and STAT1 after EGF treatment, but the value of nuclear clusters increased from four to over eight. Taken together, similar with IL-6 stimulation, the statistical results illustrate that STAT3 and STAT1 can form clusters before activation, which facilitates oligomerization to enable fast signal transduction, and that EGF promotes nuclear import of STAT3 and STAT1 and the formation of new and large clusters in the nucleus. In addition, the cluster distribution and size of STAT3 and STAT1 are very similar during the whole activation process possibly because the two proteins share approximately 70% sequence homology and have a similar crystal structure as tyrosine phosphorylated dimers in the nucleus [2].

Co-clustering analysis of STAT3 and STAT1 in the whole cell

Because STAT3 and STAT1 can form clusters individually, we questioned whether the overlapping areas (Fig. 4, yellow regions of merged images) were co-clusters consisting of both species. Therefore, to explicitly detect the colocalization of STAT3 and STAT1 under EGF stimulation, we used a combined univariate and bivariate version of Getis and Franklin's analysis method [50] to quantify the degree of co-clustering in single-molecule localization data. In addition to testing the clustering of each species, it is possible to acquire intuitive co-cluster information, for example, the number of clusters of one species overlapping with clusters or random distributive molecules of the other species. This method is more advantageous and reliable than Pearson's correlation coefficient for single-molecule localization data [50].

As shown in Fig. 6, the dual-color dSTORM image of STAT3 and STAT1 in the nucleus and cytoplasm (4×4 μm^2 regions), without and with EGF stimulation, were selected for analysis (Fig. 6, left panel). Considering that the average cluster diameters of STAT3 and STAT1 are greater than 100 nm in both nucleus and cytoplasm, with and without stimulation (Fig. 5B), the value of r was set at 100 nm, allowing for the search of cluster colocalization on a 100-nm scale. The L(100) and L(100)cross values for each molecule were calculated according to Eqn 3 (see Experimental section) and scatter diagrams were plotted (Fig. 6, middle panel). The scatter plots show the data of STAT3 molecules only. The x-axis (L(100)) denotes the self-clustering degree of STAT3 molecules and the y-axis (L(100)cross) represents to what extent STAT3 molecules are within STAT1 clusters.. Each molecule with the same L(100) and L(100)cross values was marked by one color. Based on the positions of all the points, linear trend lines (green) were generated using linear least-squares fitting in Matlab. To determine whether molecules were in or out of clusters, an appropriate clustering threshold was defined based on informed biological assumptions. Molecules whose local point density was over the threshold, were identified as being inside clusters. As described above, the average STAT3 or STAT1 numbers of cytoplasmic and nuclear clusters are different in control and IL-6-stimulated cells (Fig. 5C). For example, in nuclear areas of STAT3, the molecular number of clusters is almost more than four and the average number of localized points in a single Alexa Fluor 647-labeled STAT3 antibody is about 28, so the local point density within a nuclear cluster is approximately 112 (obtained by 28×4). As for cytoplasmic areas, the value is approximately 168 (28×6). Thus, we set two clustering threshold values, 100 for nuclear areas and 150 for cytoplasmic areas, according to their individual molecular density in the clusters. The thresholds drawn by red lines divided the plots into four quadrants. The proportion of each quadrant occupied by STAT3 molecules is displayed as a histogram (Fig. 6, right panel), from which the localization relationship between STAT3 and STAT1 was obtained. Quadrant I denotes that STAT3 molecules are not in any cluster and are only randomly distributed. Quadrant II represents that STAT3 molecules form clusters with their own species. Quadrant III indicates that random distributive STAT3 molecules are in STAT1 clusters. Quadrant IV indicates that STAT3 molecules exist simultaneously in both STAT3 and STAT1 clusters.

We first examined data derived from control cells. For nuclear areas, 23% of STAT3 molecules were in STAT3 clusters (Quadrants II and IV) whereas 20% of molecules existed within STAT1 clusters (Quadrants III and IV). The overlap was only less than 5% (IV) and the remaining 60% were in Quadrant I. The results reveal that the majority of STAT3 molecules are randomly distributed and there is no apparent co-clustering of STAT3 and STAT1. The reason may be that the two STAT species have not entered the nucleus. For cytoplasmic areas, we observed that 20% of STAT3 molecules were self-clustered (II) and 19% of the randomly distributed STAT3 molecules overlapped with STAT1 clusters (III). These two values are essentially in agreement with those of nuclear areas (19% and 16%, respectively). However, it was found that 27% of STAT3 molecules resided simultaneously in both STAT3 and STAT1 clusters (IV), which was almost six times greater when compared with the nuclear areas. This resulted in a noticeable decrease in Quadrant I (34%). The degree of co-clustering is also illustrated by the trend line showing the positive correlation between L(100) and L(100)cross. These results demonstrate that before EGF stimulation, STAT3 and STAT1 clusters are colocalized to some extent in the cytoplasm. This distribution pattern may shorten the distance between STAT3 and STAT1 proteins and benefit the formation of heterodimers when being activated.

We next applied this method to data from EGF-stimulated cells. The proportion of Quadrant II and III increased in the nucleus and decreased in the cytoplasm slightly compared with control cells, indicating that self-clustering of STAT3 and STAT1 increased in the nucleus and decreased in the cytoplasm at a slow rate. In contrast, the percentage of molecules in Quadrant I and IV markedly changed in the nucleus and cytoplasm. The co-clustering of STAT3 and STAT1 (IV) increased from approximately 5% to 37% in the nucleus, and the value also increased from 27% to 46% in the cytoplasm. Thus, the proportion of only randomly distributed STAT3 molecules (I) reduced accordingly, from 60% to 26% in the nucleus and from 34% to 22% in the cytoplasm. The colocalization of STAT3 and STAT1 clusters in stimulated cells was much higher than that in control cells, particularly for nuclear areas, revealing that EGF promotes the generation of more co-clusters, with a considerable number of these co-clusters being transported to the nucleus. Moreover, compared with the increase of self-clustering and co-clustering after stimulation, we know that nuclear accumulation of STAT3 and STAT1 clusters mainly contributes to the formation of co-clusters.

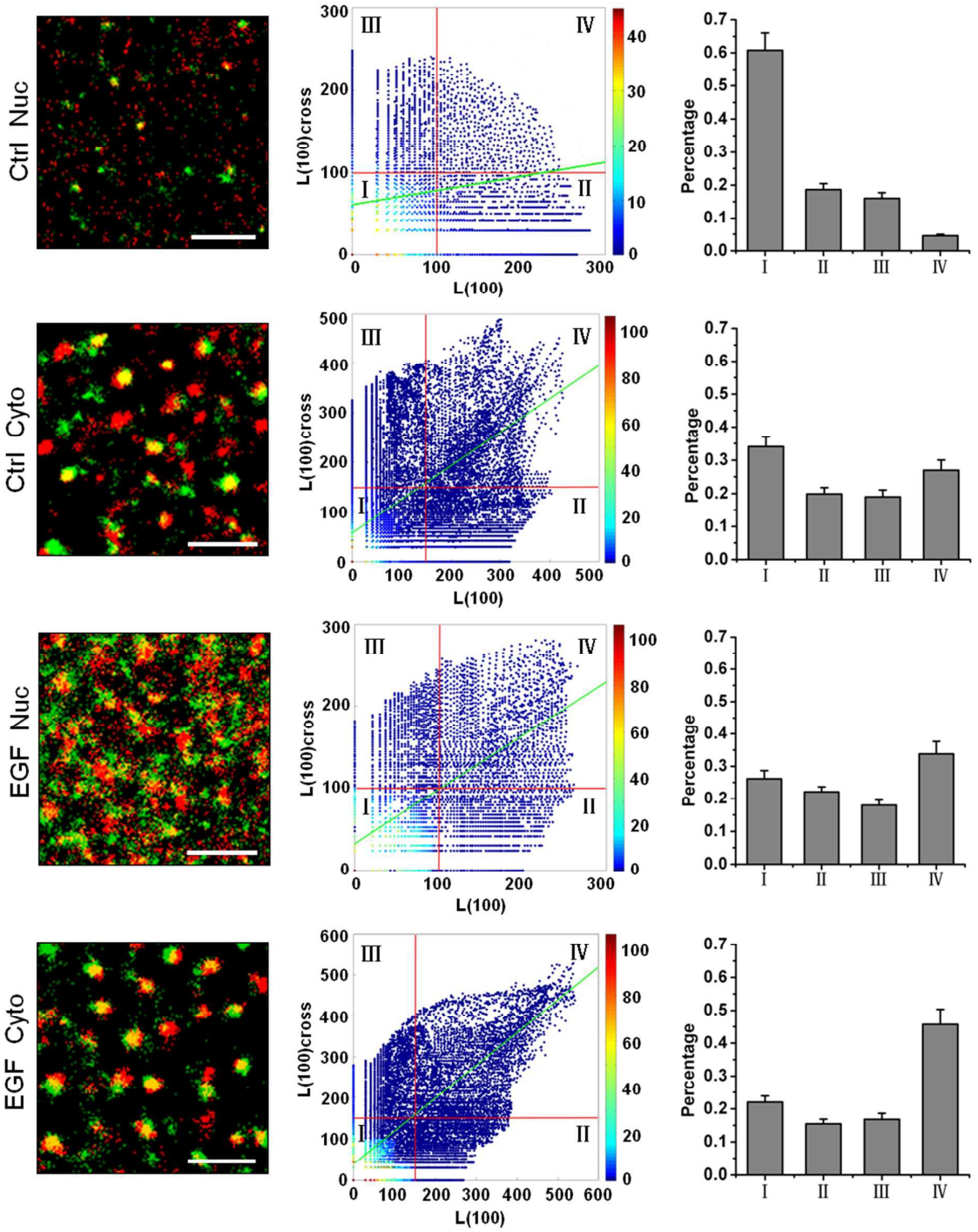


Fig. 6 Co-cluster analysis of STAT3 and STAT1 before and after EGF stimulation using a combined univariate and bivariate G&F analysis method. Regions (4×4 μm^2) of reconstructed dual-color dSTORM images of STAT3 and STAT1 in both the nucleus and cytoplasm of HeLa cells with and without EGF stimulation for 30 min were measured, as shown in the left panel. The scatter plots of combined univariate and bivariate G&F analyses for the four representative regions at a spatial scale of r = 100 nm together with trend lines (green) are shown in the middle panel. Clustering thresholds of 100 for nuclear areas and 150 for cytoplasmic areas are indicated by red lines, which divide the plots into four quadrants. The proportion of STAT3 in every quadrant is displayed as histograms in the right panel. Each data set were from 60 regions of 20 cells in five independent experiments. Every cell sample was chosen 6 regions (3 in the cytoplasm and 3 in the nucleus). Values are means±s.d. Scale bars represent 1 µm.

What are the components of the co-clusters? There are two possible hypotheses. One is that two clusters concatenate without molecular mixing, while the other is that two clusters completely merge to form a new cluster. Till now, we can not decide which one stands or both stand. Although the two species can generate homodimers and heterodimers, the proportion of each type of dimers in co-clusters remains unclear and needs further study. It is of interest here that the co-clustering in the cytoplasm is more striking than that in the nucleus after EGF stimulation. This may be because heterodimers and homodimers of STAT3 and STAT1 need to separate to bind individual target DNA motifs and activate transcription [2, 8]. Therefore, we infer that the components of co-clusters are different in the nucleus and cytoplasm. The nuclear co-clusters may be more prone to consist of heterodimers, whereas the cytoplasmic co-clusters may include not only the merging of heterodimers but also the concatenation of STAT3 and STAT1 homodimers, which leads to more co-clusters in the cytoplasm.

Conclusions

In summary, we have successfully used dSTORM to characterize the cellular clustering of STAT3 along with STAT1 in HeLa cells under different stimuli. Either STAT3 or STAT1 forms clusters in the cytoplasm whether it is activated or not, which may facilitate fast response to stimuli. The cluster characteristics of STAT3 are the same in both IL-6 and EGF stimulations, i.e., more and larger STAT3 clusters form and enter the nucleus after stimulation. Moreover, dual-color dSTORM imaging and co-clustering analysis reveal that STAT3 clusters colocalize with STAT1 clusters and the composition of the co-clusters may be related to their functions. Our study reveals the characteristics of STAT3 self-clustering and co-clustering with STAT1 in mediating signals from the cytoplasm to the nucleus, which can be generalized to other STAT members. Moreover, this method, combined with further developed algorithms, can pave the way to study the accurate stoichiometry of STAT co-clusters in the near future.

Acknowledgements

This work is supported by MOST (grant no. 2011CB933600), NSFC (Grant no. 21373200), and the “100 talent Program” of CAS.

References

1 J. E. Darnell, I. M. Kerr and G. R. Stark, *Science*, 1994, **264**, 1415-1421.
2 J. E. Darnell, *Science*, 1997, **277**, 1630-1635.
3 M. Karin and T. Hunter, *Curr. Biol.*, 1995, **5**, 747-757.
4 J. N. Ihle, *Curr. Opin.Cell Biol.*, 2001, **13**, 211-217.
5 D. E. Levy and J. Darnell, *Nat. Rev. Mol. Cell Biol.*, 2002, **3**, 651-662.
6 M. Minami, M. Inoue, S. Wei, K. Takeda, M. Matsumoto, T. Kishimoto and S. Akira, *Proc. Natl. Acad. Sci. U. S. A.*, 1996, **93**, 3963-3966.
7 M. Berishaj, S. P. Gao, S. Ahmed, K. Leslie, H. Al-Ahmadie, W. L. Gerald, W. Bornmann and J. F. Bromberg, *Breast Cancer Res.*, 2007, **9**.
8 C. Schindler and J. Darnell Jr, *Annu. Rev. Biochem.*, 1995, **64**, 621-652.
9 L. Robb, R. Li, L. Hartley, H. H. Nandurkar, F. Koentgen and C. G. Begley, *Nat. Med.*, 1998, **4**, 303-308.
10 X. Xie, K. S. Chan, F. Cao, M. Huang, Z. Li, A. Lee, I. L. Weissman and J. C. Wu, *Stem Cells Dev.*, 2009, **18**, 205-214.
11 U. M. Wegenka, J. Buschmann, C. Lütticken, P. C. Heinrich and F. Horn, *Mol. Cell. Biol.*, 1993, **13**, 276-288.
12 K. Takeda, B. E. Clausen, T. Kaisho, T. Tsujimura, N. Terada, I. Förster and S. Akira, *Immunity*, 1999, **10**, 39-49.
13 H. Yu, D. Pardoll and R. Jove, *Nat. Rev. Cancer*, 2009, **9**, 798-809.
14 P. A. Johnston and J. R. Grandis, *Mol. Interventions*, 2011, **11**, 18-26.
15 M. I. Ndubuisi, G. G. Guo, V. A. Fried, J. D. Etlinger and P. B. Sehgal, *J. Biol. Chem.*, 1999, **274**, 25499-25509.
16 K. Watanabe, K. Saito, M. Kinjo, T. Matsuda, M. Tamura, S. Kon, T. Miyazaki and T. Uede, *Biochem. Biophys. Res. Commun.*, 2004, **324**, 1264-1273.
17 L. Xu, S. Zhao, W. Ma, X. Wu, S. Li, H. Kuang, L. Wang and C. Xu, *Adv. Funct. Mater.*, 2016, **26**, 1602-1608.
18 S. Li, L. Xu, W. Ma, X. Wu, M. Sun, H. Kuang, L. Wang, N. A. Kotov and C. Xu, *J. Am. Chem. Soc.*, 2016, **138**, 306-312.
19 X. Zhao, L. Xu, M. Sun, W. Ma, X. Wu, H. Kuang, L. Wang and C. Xu, *Small*, 2016, DOI: 10.1002/smll.201503629.
20 J. Gao, Y. Wang, M. Cai, Y. Pan, H. Xu, J. Jiang, H. Ji and H. Wang, *Nanoscale*, 2015, **7**, 2511-2519.
21 J. Chen, T. Tong and H. Wang, *J. Innov. Opt. Health Sci.*, 2016, **9**, 1630007.
22 S. W. Hell and J. Wichmann, *Opt. Lett.*, 1994, **19**, 780-782.
23 T. A. Klar and S. W. Hell, *Opt. Lett.*, 1999, **24**, 954-956.
24 M. G. Gustafsson, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 13081-13086.
25 M. J. Rust, M. Bates and X. Zhuang, *Nat. Methods*, 2006, **3**, 793-795.
26 M. Heilemann, S. van de Linde, M. Schuttpelz, R. Kasper, B. Seefeldt, A. Mukherjee, P. Tinnefeld and M. Sauer, *Angew. Chem., Int. Ed.*, 2008, **47**, 6172-6176.
27 S. van de Linde, A. Loeschberger, T. Klein, M. Heidbreder, S. Wolter, M. Heilemann and M. Sauer, *Nat. Protoc.*, 2011, **6**, 991-1009.
28 E. Betzig, G. H. Patterson, R. Sougrat, O. W. Lindwasser, S. Olenych, J. S. Bonifacino, M. W. Davidson, J. Lippincott-Schwartz and H. F. Hess, *Science*, 2006, **313**, 1642-1645.
29 S. T. Hess, T. P. K. Girirajan and M. D. Mason, *Biophys. J.*, 2006, **91**, 4258-4272.
30 T. Klein, S. Proppert and M. Sauer, *Histochem. Cell Biol.*, 2014, **141**, 561-575.
31 J. Wu, J. Gao, M. Qi, J. Wang, M. Cai, S. Liu, X. Hao, J. Jiang and H. Wang, *Nanoscale*, 2013, **5**, 11582-11586.
32 M. Zhang, J. Gao, J. Chen, M. Cai, J. Jiang, Z. Tian and H. Wang, *Sci. China Chem.*, 2016, DOI: 10.1007/s11426-015-0535-0.

33 N. Ehmann, S. van de Linde, A. Alon, D. Ljaschenko, X. Z. Keung, T. Holm, A. Rings, A. DiAntonio, S. Hallermann, U. Ashery, M. Heckmann, M. Sauer and R. J. Kittel, *Nat. Commun.*, 2014, **5**.

34 Y. Wang, J. Gao, X. Guo, T. Tong, X. Shi, L. Li, M. Qi, Y. Wang, M. Cai, J. Jiang, C. Xu, H. Ji and H. Wang, *Cell Res.*, 2014, **24**, 959-976.

35 B. Dudok, L. Barna, M. Ledri, S. I. Szabo, E. Szabadits, B. Pinter, S. G. Woodhams, C. M. Henstridge, G. Y. Balla, R. Nyilas, C. Varga, S.-H. Lee, M. Matolcsi, J. Cervenak, I. Kacskovics, M. Watanabe, C. Sagheddu, M. Melis, M. Pistis, I. Soltesz and I. Katona, *Nat. Neurosci.*, 2015, **18**, 75-86.

36 K. Schuecker, T. Holm, C. Franke, M. Sauer and R. Benavente, *Proc. Natl. Acad. Sci. U. S. A.*, 2015, **112**, 2029-2033.

37 J. Gao, F. Wang, Y. Liu, M. Cai, H. Xu, J. Jiang and H. Wang, *Sci. Rep.*, 2015, **5**, 9045-9045.

38 R. Henriques, M. Lelek, E. F. Fornasiero, F. Valtorta, C. Zimmer and M. M. Mhlanga, *Nat. Methods*, 2010, **7**, 339-340.

39 B. D. Ripley, Journal of the Royal Statistical Society Series B-Methodological, 1979, **41**, 368-374.

40 D. M. Owen, C. Rentero, J. Rossy, A. Magenau, D. Williamson, M. Rodriguez and K. Gaus, *J. Biophotonics*, 2008, **3**, 446-454.

41 A. L. Pranada, S. Metz, A. Herrmann, P. C. Heinrich and G. Muller-Newen, *J. Biol. Chem.*, 2004, **279**, 15114-15123.

42 L. Liu, K. M. McBride and N. C. Reich, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 8150-8155.

43 B. F. Lillemeier, M. A. Moertelmaier, M. B. Forstner, J. B. Huppa, J. T. Groves and M. M. Davis, *Nat. Immunol.*, 2010, **11**, 90-U106.

44 J. S. Aaron, B. D. Carson and J. A. Timlin, *Small*, 2012, **8**, 3041-3049.

45 M. A. Kiskowski, J. F. Hancock and A. K. Kenworthy, *Biophys. J.*, 2009, **97**, 1095-1103.

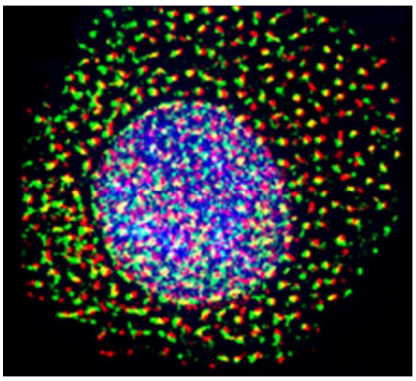
46 C. Pangarkar, D. Anh-Tuan and S. Mitragotri, *J. Controlled Release*, 2012, **162**, 76-83.

47 M. P. Brenner, *Curr. Biol.*, 2012, **22**, R597-R598.

48 A. H. Bild, J. Turkson and R. Jove, *Embo J.*, 2002, **21**, 3255-3263.

49 M. Shah, K. Patel, S. Mukhopadhyay, F. Xu, G. Guo and P. B. Sehgal, *J. Biol. Chem.*, 2006, **281**, 7302-7308.

50 J. Rossy, E. Cohen, K. Gaus and D. M. Owen, *Histochem. Cell Biol.*, 2014, **141**, 605-612.



STAT3 cellular clustering revealed by super-resolution fluorescence microscopy