bioRxiv preprint doi: https://doi.org/10.1101/188060; this version posted September 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Obtaining 3D Super-resolution Information from 2D Super-resolution Images through a 2D-to-3D Transformation Algorithm

Andrew Ruba, Joseph Kelich, Wangxi Luo and Weidong Yang*
Department of Biology, Temple University, Philadelphia, Pennsylvania, USA

Correspondence to: Dr. Weidong Yang, Department of Biology, Temple University, Philadelphia, Pennsylvania, USA. weidong.yang@temple.edu

Abstract

Currently, it is highly desirable but still challenging to obtain three-dimensional (3D) super-resolution information in studies of structures in fixed specimens as well as dynamic pictures in live cells with a high spatiotemporal resolution. Here we introduce an approach, without using 3D super-resolution microscopy or real-time 3D particle tracking, to achieve 3D sub-diffraction-limited information with a spatial resolution down to 1 nm. This is achieved by converting 2D super-resolution images through a 2D-to-3D transformation algorithm. The method has been successfully applied to obtain structural and functional information for 25-300 nm sub-cellular organelles that have rotational symmetry. In this article, we will provide a comprehensive analysis of this new method by using experimental data and computational simulations.

Introduction

Since stated by Ernst Abbe in 1873, the resolution of conventional light microscopy has been limited to approximately 200 nanometers laterally (x,y) and 600 nanometers axially (z) due to light diffraction from the microscope objective[1,2]. Currently, super-resolution light microscopy techniques break this limitation and allow for the capture of static or dynamic images with sub-diffraction resolution (< 200 nm) in all three axes. The techniques generally fall into two broad categories: optical based approaches such as stimulated emission depletion (STED) microscopy, which generate a sub-diffraction illumination volume due to the nonlinear optical response of fluorophores in samples through laser modifications; and single-molecule based mathematical approaches such as photoactivated light microscopy (PALM) and stochastic optical reconstruction microscopy (STORM). PALM and STROM utilize mathematical functions to localize the centroids of fluorophores and then reconstitute these centroids to form super-resolution images[1-4]. Although these super-resolution techniques have revolutionized imaging of biological samples via unprecedented spatial resolution, they are still limited in acquisition time (seconds to hours) and axial spatial resolution (typically > 50-100 nm). Meanwhile, fast, three-dimensional (3D) super-resolution imaging is critical for obtaining structural or dynamic information in live cells, which are inherently 3D objects. Moreover, many biological functions in sub-cellular organelles are near or below the spatiotemporal resolution limit of current 3D super-resolution imaging techniques, such as nucleocytoplasmic transport through 50-nm nuclear pore channels with millisecond transport times[7-9].

Typically, 3D super-resolution imaging is inherently more technical demanding than 2D super-resolution imaging. This is due to the fact that the point spread function (PSF) of the emitting fluorescent probe in the axial dimension is much larger than in the lateral dimension at the focal

bioRxiv preprint doi: https://doi.org/10.1101/188060; this version posted September 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

plane of the light microscopy objective[2]. Several methods have been developed to improve axial resolution in fluorescence microscopy. One category is to alter the shape of the PSF of the fluorescent probe along the optical axial position and then determine the probe's axial information by referencing a pre-determined relationship (setup through control experiments) between the shape of the PSF and the corresponding locations in the z dimension. The other is to use two objectives to improve the axial resolution after comparing fluorescent signals of the probes from these objectives with or without interference. Typically the above approaches involve expensive and complex optical implements. Here we introduce an alternative approach of overcoming this axial dimensional resolution limitation that does not require unique optics. The method is to transform 2D super-resolution images or single-molecule localization distributions into their corresponding 3D probability information. This can be done by utilizing biological structures that are inherently symmetrical in one of the three dimensions. The whole process includes three major steps: 1) determine the structure of a sub-cellular organelle by electron or expansion microscopy; 2) conduct single-particle tracking and/or single-molecule localization to get 2D super-resolution images or 2D single-molecule localization distributions in the structure; and 3) transform these experimentally obtained 2D data to their corresponding computational 3D probability information based on the previously obtained structure. Here, we will generally introduce the method but mainly focus on a detailed analysis of the 2D-to-3D transformation process and demonstrate its applications in determining 3D structural and functional information in sub-cellular sub-micrometer organelles that have rotational symmetry[10,11], such as the nuclear pore complex (NPC) and the primary cilia. Also, we further suggest that the 2D-to-3D transformation process can be extended to convert 2D super-resolution images obtained from 2D super-resolution light microscopy techniques, by using STORM-based 2D data of microtubules as an example.

Results

Mathematical concept and detailed process for the 2D-to-3D transformation algorithm As shown in Fig. 1 and Online Methods, the main idea behind the 2D-to-3D transformation algorithm is that, for any radially symmetric biological structure (typically determined by using electron microscopy) such as NPCs, primary cilium and microtubules seen in Figure 1A-D, an area matrix in the radial dimension can be developed (y and z dimensions in Figure 1E). As molecules locate in or travel through these structures, their locations can be projected into the xy or the yz plane, depending on whether microscopy imaging of the structure is conducted at the lateral or the axial dimension respectively (Fig. 1F). Typically, in practice it is much easier to obtain 2D super-resolution images of these targeted molecules in the xy plane than the yz plane because of the much smaller radial dimension than the axial dimension in these structures. In the end however, the same information on the structure can be obtained. After further projecting the 2D molecular spatial locations into the y dimension from either the xy or the yz plane, the obtained two y-dimension histograms in principle are identical as demonstrated in Fig. 1G. Then, based on the two identical y-dimension histograms, each column in the y-dimension histogram projected from the xy plane will be equal to the areas times the densities for each radial bin of the yz plane (Fig. 1 G-H). Finally, as detailed in the mathematical calculation shown in Online Methods, the densities in the radial dimension can be obtained by solving the matrix equations, which eventually reconstitute the corresponding 3D super-resolution information for the structure (Fig. 1 I-K).

Monte Carlo simulation demonstrates the parameters required for accurate 3D density map

bioRxiv preprint doi: https://doi.org/10.1101/188060; this version posted September 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

reconstruction Next, we use Monte Carlo simulation to demonstrate that two critical parameters, the single-molecule localization error of targeted molecules and the number of single-molecule locations, will determine the reproducibility of obtaining accurate 3D super-resolution information in the biological structures with rotational symmetry. Typically, the spatial localization of individual targeted molecules labeled with fluorophores or fluorescent proteins were determined with non-zero localization errors because of background noise in real experiments. To mimic typical transport routes of proteins in the NPC or primary cilia, our Monte Carlo simulations were performed where varying numbers of single-molecule locations were randomly simulated on an ideal radius (R_I) (Supplemental Figure 3A-B). Then, single-molecule localization error (σ_{LE}) was added to R_I by sampling an error value from a normal distribution with a standard deviation of σ_{LE} (Supplemental Figure 3C). Subsequently, the 2D-to-3D transformation algorithm was performed on only the y dimensional data of the simulated single-molecule localization distribution around R_I to model the loss of z dimensional information during the 2D microscopy projection process. The peak position of the transformed 3D density histogram was then determined by Gaussian fitting to produce a measured radius (R_M) which may deviate from the R_I due to limited number of simulated locations and non-zero single-molecule localization errors (Supplemental Figure 3D). We conducted 10,000 iterations of this process and obtained 10,000 R_M values, in which the mean of the R_M values converges on R_I as expected (Supplemental Figure 3E). To quantify how reproducible a single experimental data set is, we set out to determine how many individual R_M values from the whole distribution of R_M fell within an acceptable range of R_I. The acceptable range was defined as the $R_I \pm \sigma_{LE}$ because, in principle, any single R_M value can only be accurately localized within the range of two standard deviations of its Gaussian fitting according to the Rayleigh Criterion (Supplemental Figure 3F). We expect that a high number of simulated single-molecule locations or low simulated localization error would increase the number of iterations that fall within the acceptable range, thus resulting in a high reproducibility rate. Two critical steps in the process are to, first, optimize the bin size for each set of simulation parameters. This is accomplished by determining the smallest bin size that produces no statistical difference by Chi square analysis between the original 2D histogram and the back-calculated 2D histogram obtained by multiplying the 3D density histogram by the corresponding area matrix (Supplemental Figure 4). The second step is to account for the sensitivity of the inner bins of the area matrix when determining the accurate R_M peak fitting (Supplemental Figure 5).

To test the effects of single-molecule localization error on the final R dimensional peak fitting obtained for the 3D transformed density histogram, Monte Carlo simulations were performed with an R_I of 25 nm, data point number of 1,000,000, and σ_{LE} ranging from 0 to 30 nm (Figure 2A-D). In principle, as the single molecule localization error becomes excessively large, the peaks in the 3D density map will become heavily overlapped on the y dimensional axis and, subsequently, the radial axis after the transformation algorithm. This will obscure the peak at R_I and make it indistinguishable. A ratio between the error of a bimodal Gaussian fitting and the error of a single Gaussian fitting is used to determine the indistinguishable overlap (Figure 2E-H). As shown in Figure 2I, the bimodal fitting error becomes much larger than the single peak fitting error beyond 21-nm localization error. This indicates that the experimental localization error cannot exceed 21 nm for any structure or transport route with a radius of 25 nm This is smaller than the theoretical single-molecule localization error of 25 nm predicted by the Rayleigh Criterion mainly due to the aforementioned sensitivity of the inner bins of the area matrix in this 2D-to-3D transformation process (Supplemental Figure 5).

bioRxiv preprint doi: https://doi.org/10.1101/188060; this version posted September 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

To determine the effects based on the quantity of single molecule locations constituting 3D transformed density map, Monte Carlo simulations were performed with an R_I of 25 nm, σ_{LE} of 10 nm, and data point numbers ranging from 50 to 1,000. Representative simulations are shown for 50, 100, and 500 points, and the corresponding reproducibility percentage was calculated for each point number after 10,000 iterations (Figure 3A-F). Remarkably, only 100 and 350 points are needed to achieve 90% and 100% reproducibility respectively (Figure 3G). It is highly feasible to obtain 100-500 points experimentally, although the number of points is higher than the initial Nyquist Sampling theorem estimation of the minimum 38 single-molecule locations (Supplementary Information) because of a non-uniform distribution of locations through the area matrix.

A general formula developed for estimating the accuracy of final 3D information Through the above simulations, we noticed that the parameters for determining σ_{TR}, the standard deviation of the R_M distribution (Figure 4 A-B), are analogous to the parameters for determining single-molecule localization error (σ_{LE}) experimentally[19]. Thus, we set out to develop a general formula for estimating σ_{TR}, the accuracy of the final 3D density map. The development firstly starts with the standard error on the mean, $\frac{s^2}{N}$, where s is the standard deviation of the individual 3D density map peak fitting, and N is the number of single-molecule locations collected to determine the 3D density map. Next, binning of the single-molecule locations can introduce extra noise to the final accuracy, which can be estimated by the variance of a continuous distribution as $a^2/12$, where a is the bin size. The above two steps generate an accuracy of $\frac{s^2 + a^2/12}{N}$. Another source of noise is the fact that, unlike single fluorescent molecules, which emit photons from single spots, the locations that constitute the 3D density map come from a distribution that is spread by the convolution of two sources as mentioned before: the localization error of single molecules and the physical width of the biological transport route. Therefore, this noise is accounted for by dividing the number of single-molecule locations collected by 4s that is the segment of the radius over which 95% of the 3D density map is spread, resulting in the estimated accuracy of $\frac{s^2 + a^2/12}{N/(4s)}$. Taking the square root of the variance equation yields that standard deviation, which gives the final version of formula to estimate the error σ_{TR}, $\sigma_{TR} = \sqrt{\frac{s^2 + a^2/12}{N/(4s)}}$. Finally, we tested the formula by comparing the predicated results for σ_{TR} with the simulated results by using the same parameters outlined in Figure 4D-F. As shown, the formula works very well in estimating the accuracy σ_{TR} for 3D density maps. It is noteworthy that the accuracy curve approaches 1-nm or ≤ 1 nm resolution in 3D mapping with 1,000 or more single-molecule locations, which is also feasible to obtain experimentally (Fig. 3C). Thus in theory, given enough single-molecule data, the provided resolution can approach that of electron microscopy.

SPEED Microscopy Our lab has previously developed SPEED microscopy to fill the technique niche of capturing single molecules transporting through sub-diffraction-limit bio-channels at high spatial (< 10 nm) and temporal (< 1ms) resolution[10,11]. We achieve this through four main technical modifications on traditional epifluorescence or confocal light microscopy. (1) A small inclined or

bioRxiv preprint doi: https://doi.org/10.1101/188060; this version posted September 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

vertical illumination PSF is used for the excitation of single transiting molecules through bio-channels in the focal plane. This greatly increases the allowable detection speed (up to 0.2 ms per frame for the CCD camera we used) by reducing the number of camera pixels required for detection. Also, it significantly avoids out-of-focus fluorescence. (2) A high optical density (100–500 kW/cm^2) in the small illumination volume causes a high number of photons from the fluorophores be emitted. (3) A collection of high-number of photos in a short time period increases spatial resolution by reducing the effects of molecular diffusion on the single-molecule localization precision of moving molecules (Online Methods). (4) Pinpointed illumination in live cells causes negligible photo-induced toxicity. Thus, SPEED microscopy meets the needs of high 2D spatiotemporal resolution for *in vivo* single-molecule tracking in dynamic bio-channels such as single NPCs. However, SPEED does not directly obtain any 3D information. As demonstrated below, we have employed SPEED microcopy to obtain 2D single-molecule data in glass nanocapillary tubes, the NPC, and primary cilia. However, other single-particle tracking or super-resolution microscopy approaches may be employed to obtain similar 2D data sets for processing by the 2D-to-3D transformations.

Experimental validation of 2D-to-3D transformation process in several systems: glass nanocapillary tube, nuclear pore complex, primary cilia, and microtubules Since the 2D-to-3D transformation algorithm requires radial symmetry, we first used an ideal artificial glass nanocapillary to test the algorithm's accuracy when coupled with SPEED microscopy for data acquisition. The glass nanocapillaries (GNCs) were fabricated using laser-assisted capillary-pulling of quartz micropipettes which can generate pore diameters ranging from 20 nm to 300 nm. The dimensions of GNCs used in this study were determined by Helium scanning transmission ion microscopy to have an inner radius ~35 nm. With that parameter in mind, the dimensions of the GNC were re-measured by determining the 3D density map of 1-nM Alexa Fluor 647 that was pumped into the inner lumen of the GNC. After thousands of 2D spatial locations for Alexa Fluor 647 were collected with a single-molecule localization precision of ≤ 5 nm, the final 3D density map revealed a radius of 37±2 nm, agreeing well with the 35-nm inner radius of the GNC imaged by helium ion microscopy (Fig. 5 A-C).

After the inner diameter of the GNC was confirmed by the application of the 2D-to-3D transformation algorithm, we moved to two live cell systems: Importin β1 (a major transport receptor in nucleocytoplasmic transport) tracking through the NPC and externally-labeled SSTR3 (a major transmembrane protein in primary cilia[20]) on the surface of primary cilia on the ciliary shaft. Previously, our lab has revealed that Importin β1 assists the movement of protein cargo via interactions at the periphery of the NPC, a selective gate between the cytoplasm and nucleus[7,8]. In this analysis, we present a total of 450 2D spatial locations for Importin β1 within the NPC. The corresponding 3D density map clearly shows a high density region for Importin β1 at 23 nm along the radius of the NPC with a 100% reproducibility rate (Fig. 5 D-F). Similarly in primary cilia, SSTR3, was externally-labeled with Alexa Fluor 647[21] and tracked using SPEED microscopy along the length of primary cilia, ~125 nm in radius[17,18,22]. Agreeing well with the diameter reported by electron microscopy, the externally-labeled SSTR3 was shown to have a high density region at 127 nm along the radius of primary cilia, with a 97% reproducibility rate (Fig. 5 G-I).

Lastly, to test whether the 2D-to-3D density transformation algorithm could be applied beyond SPEED microscopy, we measured the diameter of microtubules, in which tubulins were labeled

bioRxiv preprint doi: https://doi.org/10.1101/188060; this version posted September 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

by a primary and secondary antibody conjugated to Alexa Fluor 647 and then imaged by 2D-STORM[23]. By converting the published 2D super-resolution image for a microtubule into its 3D structure by our 2D-to-3D transformation algorithm, we determined the diameter of microtubules in this specific sample to be 64±1 nm with a reproducibility rate of 90% (Fig. 5 J-L). This result agrees well with previous determinations by using EM[24] and 3D super-resolution microscopy[5].

Discussion

In this paper, we presented a detailed analysis of the 2D-to-3D transformation algorithm that enabled us to obtain 3D super-resolution information from 2D super-resolution images or 2D single-molecule localization data without using 3D light microscopy setups. The roles that several critical factors played in reproducing the accurate 3D super-resolution information, such as the single molecule localization error, the binning effects, and the number of single molecule locations, have been fully discussed. The successful applications in various systems, including the GNC, the NPC and primary cilia in live cells, and microtubules in fixed samples, prove the robustness of achieving accurate 3D super-resolution information by combining 2D experimental data and the 2D-to-3D transformation algorithm.

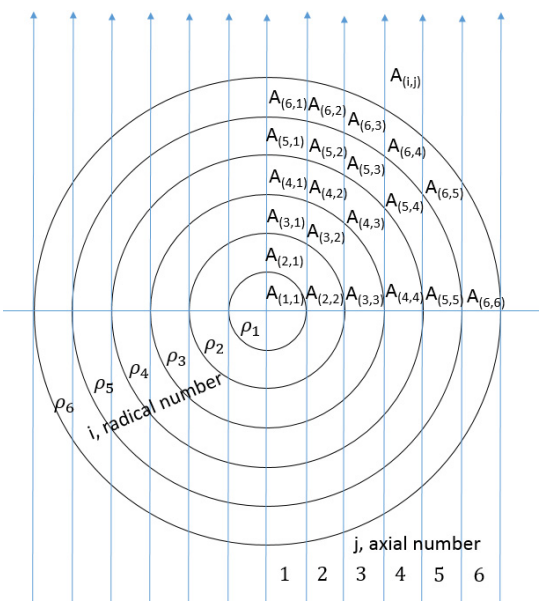
It is noteworthy that one prerequisite of the algorithm is the radial symmetry of the underlying biological structure. Normally, transmission electron microscopy is appropriate for determining the structure and its rotational symmetry with high spatial resolution. Another alternative approach is expansion microscopy, by which the size of a specimen can be enlarged 4.5 to 20 times without significant distortion and the enlarged structure can be labeled and imaged by epi-fluorescence or confocal light microscopy[25-26]. As demonstrated, with the known structures of the NPC, primary cilia, and microtubules, area matrices could be developed for the cylindrical structures based on their rotational symmetry for 2D-to-3D transformation algorithm. Furthermore, area matrices could also be developed for other regular or even irregularly shaped, radially symmetric structures. The only requisition is that the density of the molecules of interest is constant along a given radial bin.

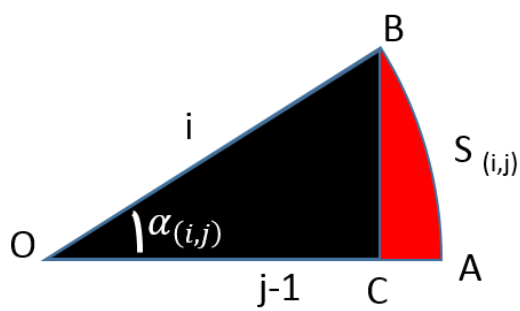
bioRxiv preprint doi: https://doi.org/10.1101/188060; this version posted September 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Online Methods

Calculating the regions of the area matrix

In the 2D-to-3D density transformation process, we use an area matrix to reflect the contribution of each ring to the 2D distribution. As shown below, we define i as the radial number and j as the axial number. The density of single molecule locations in the same radius i is supposed to be uniform given the radial symmetry and the density is labeled as ρ_i here. The cross-ection of radial number i and axial number j is defined as $A_{(i,j)}$. In the following equations, Δr is the bin size, ρ_i is the density of molecules in the *i*th ring, h is the half of illumination depth and ρ is the constant background density outside the region of interest.





To determine the area of each sub-region ($A_{(i,j)}$), we always start to calculate the area of the fan-shape area at $j=i$, in which:

$$\cos\alpha_{(i,j)} = \frac{j-1}{i}, \quad \sin\alpha_{(i,j)} = \frac{\sqrt{i^2-(j-1)^2}}{i}, \quad \alpha_{(i,j)} = \cos^{-1}\frac{j-1}{i}$$

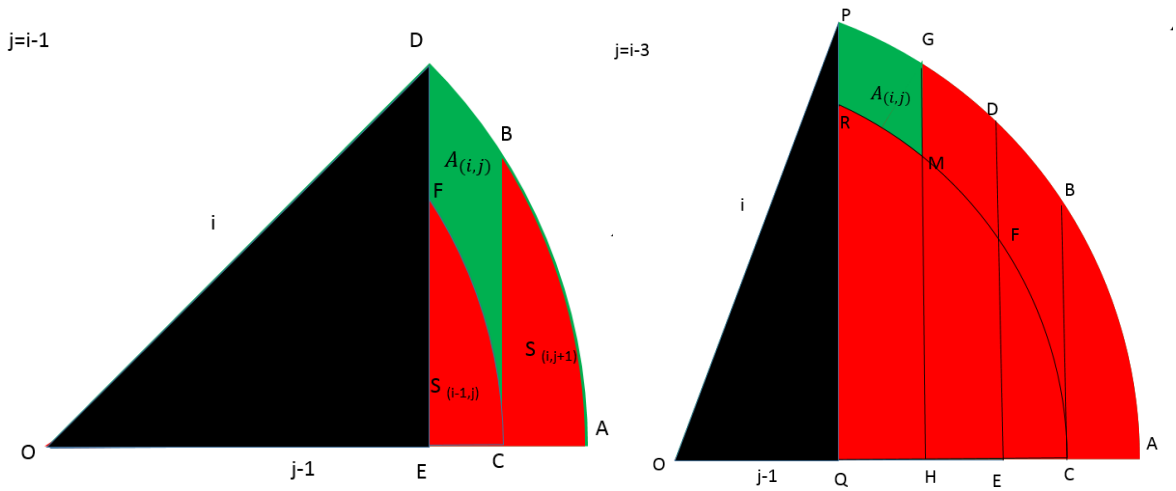
The red-shaded area is defined as $S_{(i,j)}$:

$S_{(i,j)}=S_{ABC}=S_{AOB}-S_{BOC}$

$$S_{(i,j)} = \frac{1}{2}\alpha_{(i,j)} i^2\Delta r^2 - \frac{1}{2}\sin\alpha_{(i,j)}\cos\alpha_{(i,j)}\, i^2\Delta r^2$$

$$S_{(i,j)} = \frac{1}{2}[\cos^{-1}(\frac{j-1}{i})]\, i^2\Delta r^2 - \frac{1}{2}[\frac{\sqrt{i^2-(j-1)^2}}{i}]\,(\frac{j-1}{i})\, i^2\Delta r^2 \qquad \text{(Eq. 1)}$$

bioRxiv preprint doi: https://doi.org/10.1101/188060; this version posted September 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



When $j \neq i$ *and* $i > j$, we need to calculate the area of the green-shaded region as follows:

$S_{CBDF} = S_{ADE} - S_{CFE} - S_{ABC}$ $S_{GMRP} = S_{AQP} - S_{CQR} - S_{AHG} + S_{CMH}$

$$A_{(i,j)} = S_{(i,j)} - S_{(i-1,j)} - S_{(i,j+1)} + S_{(i-1,j+1)} \quad \text{(Eq. 2)}$$

When $i < j$, $A_{(i,j)} = 0$ (Eq. 3)

Following the above equations, all $A_{(i,j)}$ can be precisely calculated.

Finally, N_j, the number of events detected in the jth column can be calculated with the following equation:

$$N_j = 2 * \{\sum_{i=j}^{n} \rho_i A_{(i,j)} + [h\Delta r - \sum_{i=j}^{n} A_{(i,j)}]\rho\} \quad \text{(Eq. 4)}$$

As soon as N_j and $A_{(i,j)}$ are known, ρ_i will be calculated.

Single molecule localization precision

For immobile molecules or fluorescent nuclear pores, the fluorescent spot was fitted to a 2D symmetrical or an elliptical Gaussian function, respectively, and the localization precision was determined by the standard deviation (s.d) of multiple measurements of the central point. So, the precision is presented as mean ± s.d. However, for moving molecules, the influence of particle motion during image acquisition should be considered in the determination of localization precision. In detail, the localization precision for moving substrates (σ) was determined by the formula $\sigma = \sqrt{F[\frac{16(s^2 + a^2/12)}{9N} + \frac{8\pi b^2 (s^2 + a^2/12)^2}{a^2 N^2}]}$, where F is equal to 2, N is the number of collected photons, a is the effective pixel size of the detector, b is the standard deviation of the

bioRxiv preprint doi: https://doi.org/10.1101/188060; this version posted September 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

background in photons per pixel, and $s = \sqrt{s_0^2 + \frac{1}{3} D\Delta t}$, s_0 is the standard deviation of the point spread function in the focal plane, D is the diffusion coefficient of substrate and Δt is the image acquisition time[2-5]. In our experiments, more than 1000 signal photons were collected from each targeted moving molecule. As a result, the localization precision for the moving molecules is calculated to be <10 nm based on the above equations while the parameters were determined experimentally (N >1000, a=240 nm, b≈2, s_0=150±50 nm, Δt=*0.4 ms,* D is < 0.1mm^2/s for the tested substrates in the central channel of nuclear pores[6-7]). Therefore, the localization precision for moving molecules in our previous measurements is a mixture of precisions smaller than 10 nm[27-30].

bioRxiv preprint doi: https://doi.org/10.1101/188060; this version posted September 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figures

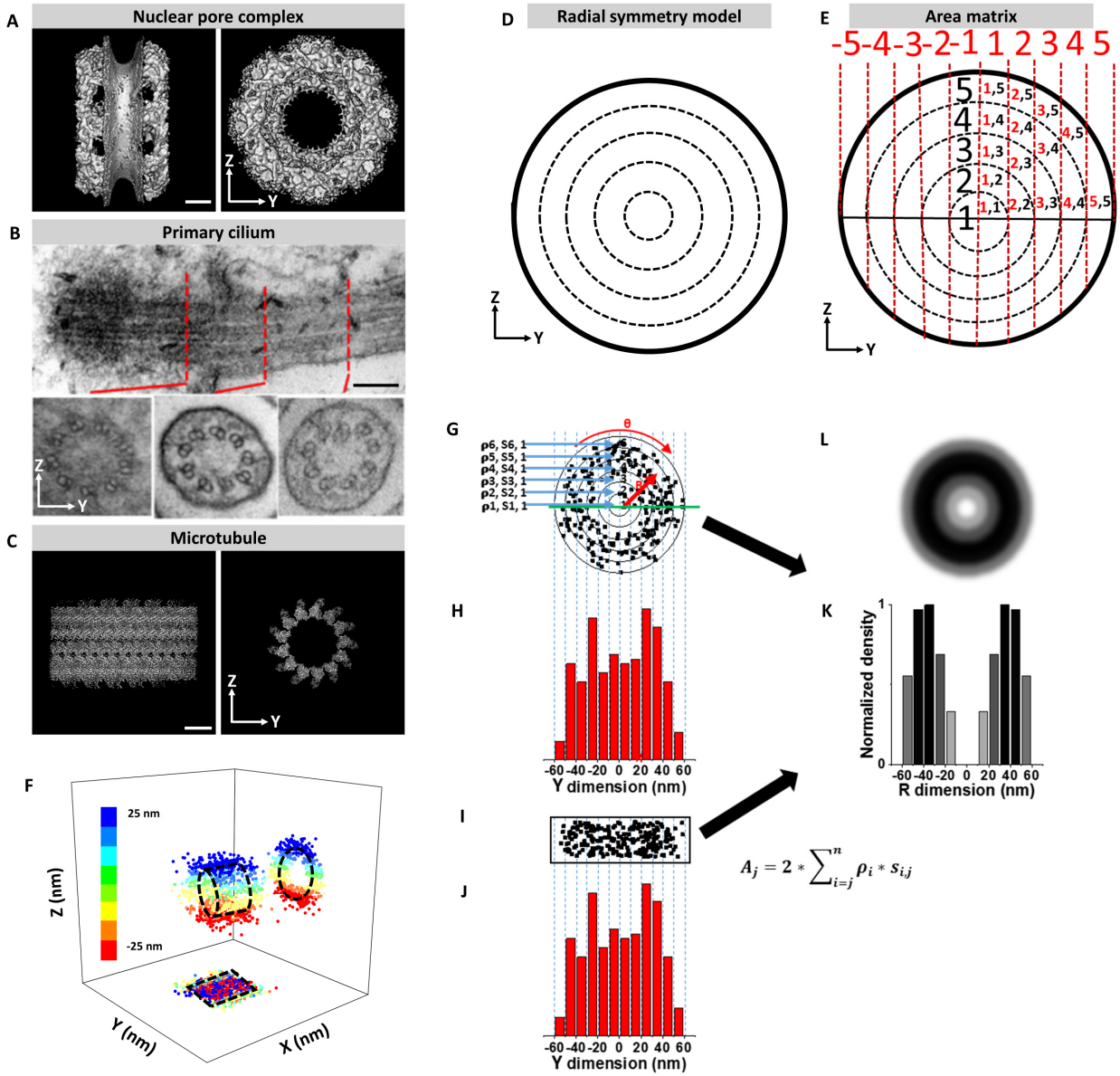


Figure 1. An area matrix can be developed for radially symmetric biological structures that reflects 2D single molecule data. **(A)** 3D electron tomography image of the nuclear pore complex averaged rotationally with 8-fold symmetry[31,32]. Scale bar = 20 nm. **(B)** Transverse slice of the primary cilia using transmission electron microscopy[33,34]. From left to right, the red lines indicate cross-sections of the basal body, transition zone, and ciliary shaft of the primary cilium. Scale bar = 100 nm. **(C)** 3D electron tomography image of a microtubule averaged rotationally with 13-fold symmetry[32,35]. Scale bar = 10 nm. **(D)** Structures from (A), (B), and (C) may be simplified as a radially symmetric circle in the y and z dimensions. **(E)** An area matrix may be designed where the radially symmetrical simplified model is divided along the y dimension due to the fact that the projection of 2D single molecule data is along the y dimension even though it came from the y and

bioRxiv preprint doi: https://doi.org/10.1101/188060; this version posted September 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

z dimensions. **(F)** Simulated single molecule data in the x, y, and z dimensions from a cylinder with ideal radius of 25 nm and localization error of 5 nm for demonstration of the 2D-to-3D density transformation. Dotted lines represent the ideal structure from which the simulated originated. **(G)** Single molecule data from (F) in the y and z dimensions superimposed with the area matrix from (E). **(H)** Histogram of the single molecule data from (G) along the y dimension to model the effects of 2D microscopy, from which a y dimensional histogram can also be obtained. **(I)** Single molecule data from (F) in the x and y dimensions. **(J)** Histogram of the single molecule data from (I) along the y dimension. **(K)** The resultant radial density map of the single molecule data from (G) determined by dividing the number of points in each radial bin by the area of each radial bin as well as the resultant radial density map of only the y dimensional histogram from (I) determined by using the 2D-to-3D density transformation. **(L)** Radial density map from (K) expressed as a cartoon reconstruction of the density of single molecules through the original 3D structure. The intensity of rings, from black to white, indicates highest to lowest normalized density.

bioRxiv preprint doi: https://doi.org/10.1101/188060; this version posted September 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

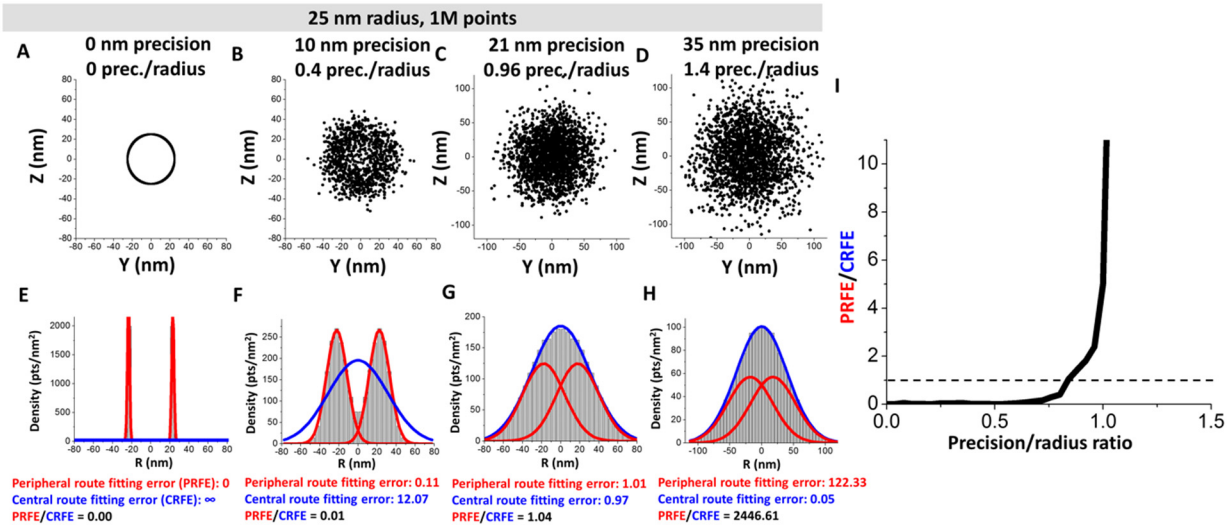


Figure 2. **Varying the simulated precision shows the resolution limit of the 3D transformation algorithm. (A), (B), (C), and (D)** 1,000,000 locations were simulated on an ideal radius (R_I) of 25 nm with localization errors (σ_{LE}) of 0, 10, 21, and 35 nm respectively. Scatter plots were down-sampled to 2,000 locations for visualization. **(E), (F), (G), and (H)** The corresponding 3D transformed density histogram of each simulation from (A), (C), (E), and (G) respectively. To determine whether the density peaks could be distinguished after a given localization error had been added to the ideal data, the ratio between the fitting error of a bimodal Gaussian distribution (PRFE) and a single Gaussian distribution was used (CRFE). If the bimodal fitting error was less, then the peaks in the 3D transformed distribution can likely be distinguished. If the single peak fitting error was less, the localization error was too great and the two peaks were indistinguishable. **(I)** Up to 21 nm, the PRFE was much less than the CRFE. Above 21 nm, the CRFE was much less than the PRFE. Therefore, the maximum localization error allowed to distinguish a 25 nm ideal radius is ~21 nm.

bioRxiv preprint doi: https://doi.org/10.1101/188060; this version posted September 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

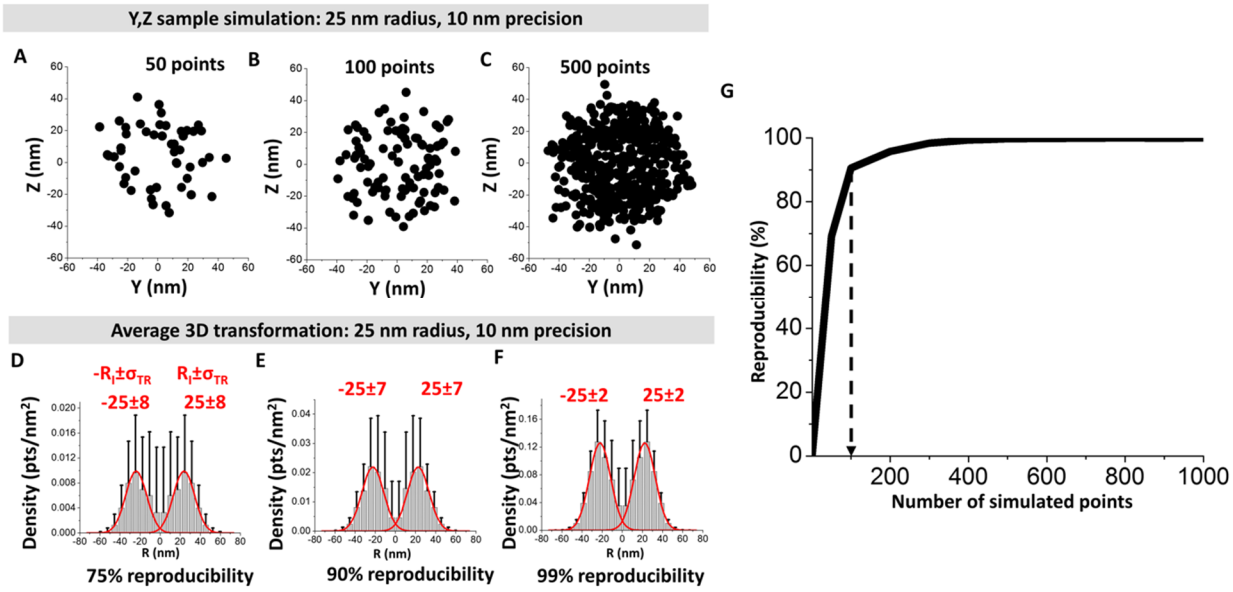


Figure 3. **Varying the number of simulated points shows the effect of sampling error on the peak fitting of the 3D transformed data. (A), (B), and (C)** 50, 100, and 500 locations were simulated on an ideal radius (R_I) of 25 nm with localization error (σ_{LE}) of 10 nm. **(D), (E), and (F)** The average 3D transformation of 10,000 simulated data sets for each number of simulated locations. Error bars indicate standard deviation of each histogram bin value while $R_I \pm \sigma_{TR}$ indicates the average peak fitting ± the standard deviation of the peak fittings. That is, the standard deviation (σ_{TR}) of the R_M distribution outlined in Figure 1E. The reproducibility percentage, the number of peak fittings that fell within the localization error, is shown beneath each 3D transformed density histogram. **(G)** The plot of reproducibility percentage for 10,000 iterations of each number of points up to 1,000. The simulation parameters were the same as (A), (B), and (C) except for the varying point number.

bioRxiv preprint doi: https://doi.org/10.1101/188060; this version posted September 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

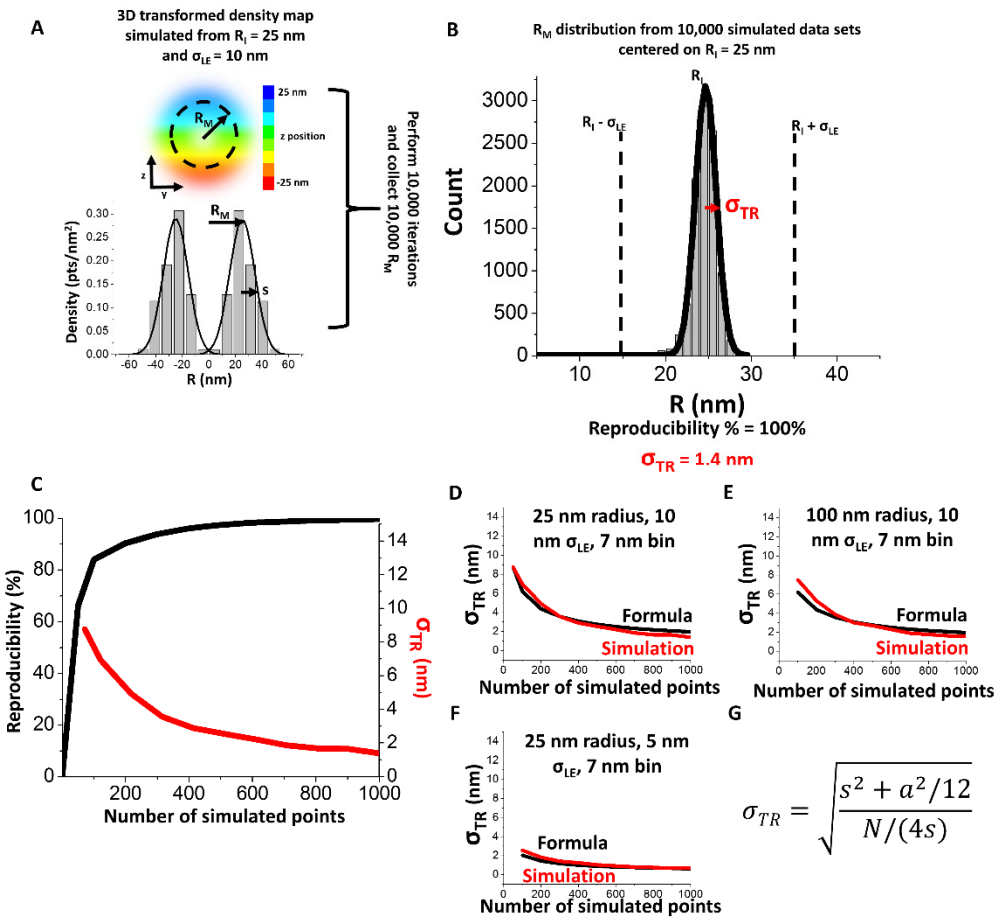


Figure 4. **A general formula developed for estimating the accuracy in the 3D transformed density map. (A)** 10,000 data sets were simulated with an ideal 25-nm radius (R_I) and a localization error (σ_{LE}) of 10 nm. The resultant 3D histograms were then each fitted with a Gaussian function to localize the mean position of each peak, which is designated as the mean radius R_M. **(B)** The histogram for all the R_M values was determined and the number of simulated data sets that fell within the acceptable range of $R_I \pm \sigma_{LE}$ were counted. After 10,000 simulations of data sets with 1,000 points, R_I of 25 nm, and σ_{LE} of 10 nm, the reproducibility percentage was 100% while the standard deviation of the R_M distribution (σ_{TR}) was 1.4 nm. **(C)** The reproducibility percentage and the corresponding σ_{TR} were plotted across all point values up to 1,000 points for R_I of 25 nm and σ_{LE} of 10 nm. **(D)** σ_{TR} may be determined by simulation, as has been done up to this point. However, a formula to estimate σ_{TR} may be derived based on the formula for single molecule localization error. The average error for the formula compared to the simulation with the parameters from (A) and (B) yields 13±17%.

bioRxiv preprint doi: https://doi.org/10.1101/188060; this version posted September 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

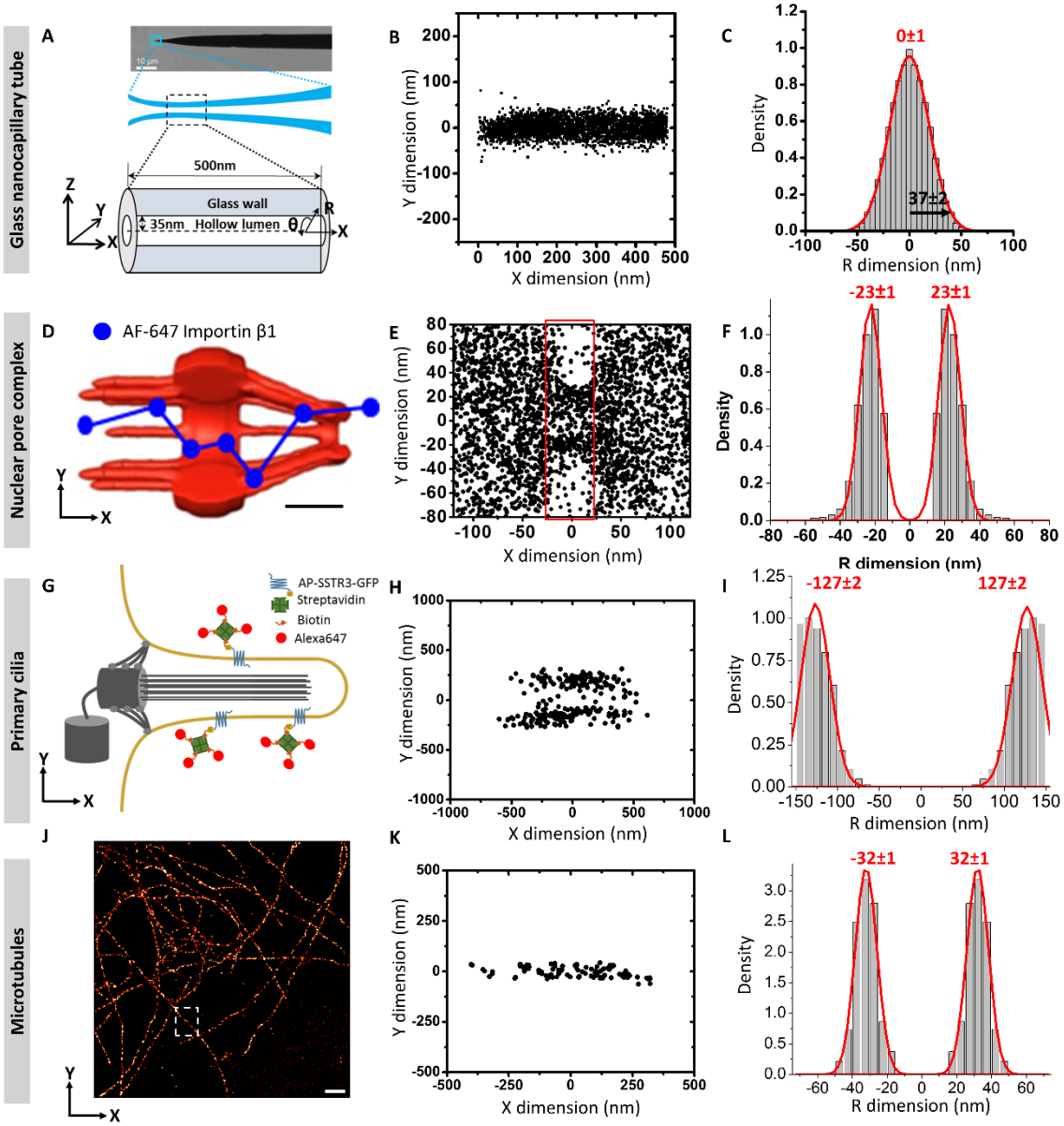
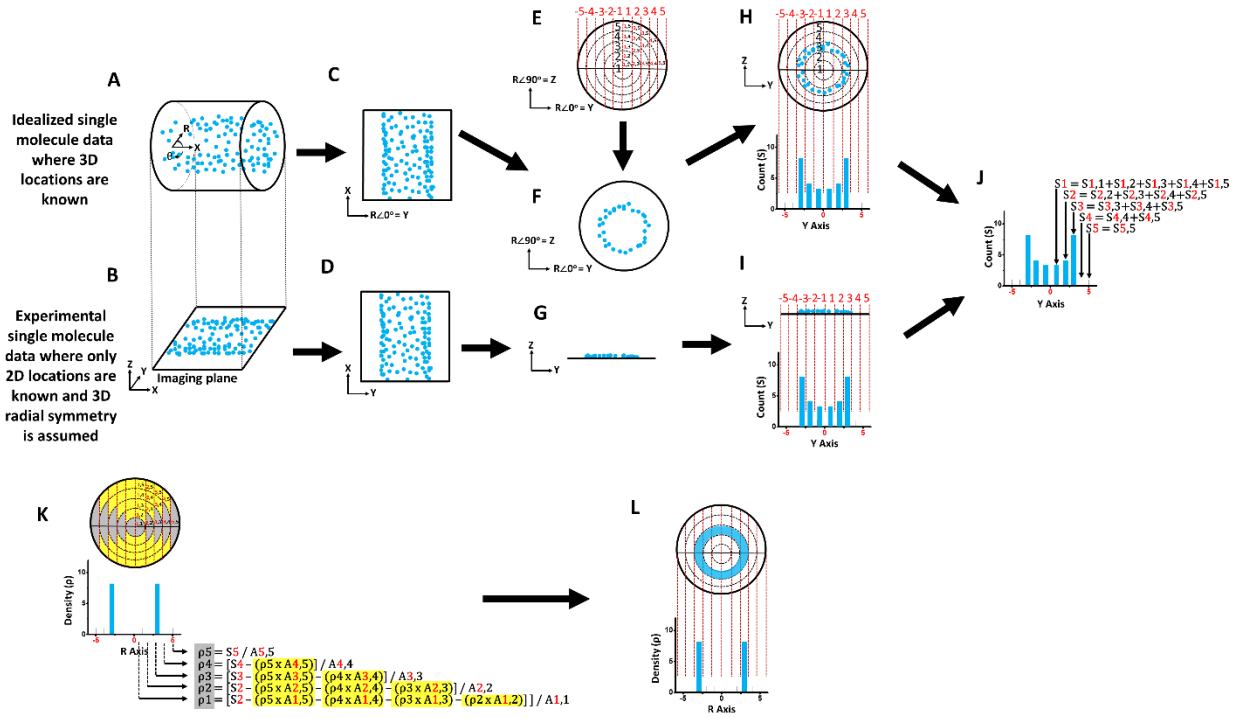


Figure 5. **3D density maps accurately reconstruct the locations of labeled molecules in various structures. (A)** Schematic of the glass nanocapillary tube with dimensions of 35 nm inner radius. **(B)** x and y dimensional single molecule data from the tracking of Alexa Fluor 647 inside the glass nanocapillary tube. **(C)** 3D density map of (B) showing the width at 2 standard deviations from the mean is 37 nm. **(D)** Schematic of the nuclear pore complex with a model trajectory for Alexa Fluor 647-labeled Importin β1. Scale bar = 50 nm. **(E)** x and y dimensional single molecule data from the tracking of Alexa Fluor 647-labeled Importin β1 in the nuclear pore complex. **(F)** 3D density map of (E) showing the peak fitting of the Importin β1 transport route is 23 nm. **(G)** Schematic showing the Alexa Fluor 647 externally-labeled SSTR3 in the shaft of a primary cilium. **(H)** x and y dimensional single molecule data from the tracking of externally-labeled SSTR3 in the primary cilium. **(I)** 3D density map of (H) showing the peak fitting of the SSTR3 transport route is 127 nm. **(J)** 2D STORM data showing the reconstructed super-resolution image of tubulin

bioRxiv preprint doi: https://doi.org/10.1101/188060; this version posted September 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

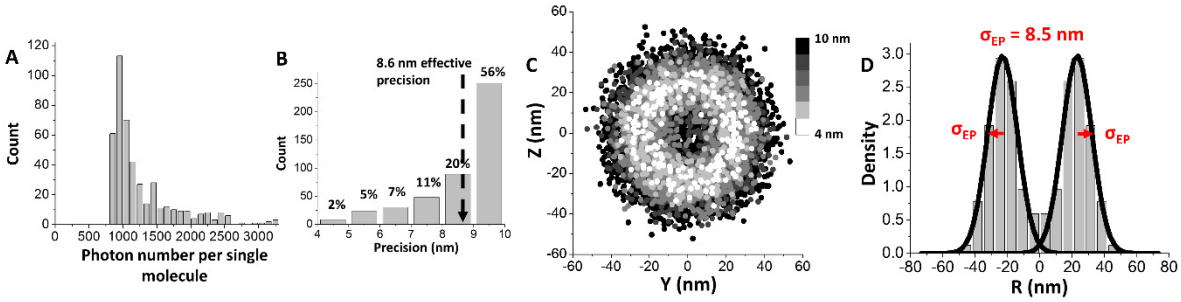
labeled with a primary and secondary antibody attached to Alexa Fluor 647. Scale bar = 1 μm. **(K)** x and y dimensional single molecule data from section of 2D STORM data in the dashed white box from (J). **(L)** 3D density map of (K) showing the peak fitting of the labeled tubulin from (K) to be 32 nm.

bioRxiv preprint doi: https://doi.org/10.1101/188060; this version posted September 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



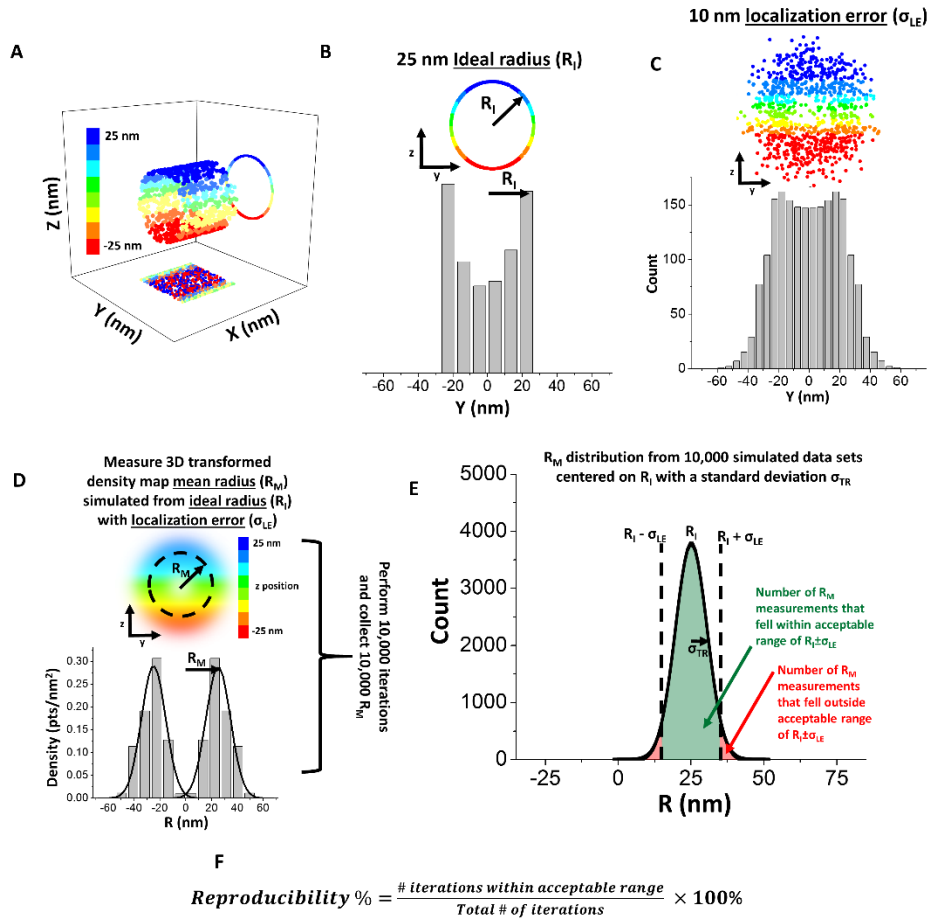
Supplementary Figure 1. **(A)** Idealized simulated single molecule data in x, y, and z dimensions. **(B)** Single molecule data from (A) projected to just the x and y dimensions, as it would be during the microscopy imaging process. **(C) and (D)** In the x and y dimensions, both the 3D single molecule data and the 2D projected data are identical. **(E)** An area matrix subdivided radially and axially to account for the radially symmetric structure from which the single molecule data originated. **(F)** Single molecule data from (A) viewed in the y and z dimensions. **(G)** Single molecule data from (B) viewed in the y and z dimensions. Since the data is projected, the z dimension is uniform. **(H) and (I)** y dimensional histograms from (F) and (G) are identical due to the fact that the z dimension is projected during the histogram process. **(J)** Each column in the y dimensional histogram in both (H) and (I) originate from the sum of the corresponding axial bins in the area matrix. **(K)** The 3D density histogram is calculated by using the outermost radial bin as a reference, calculating its density, then subtracting its contribution from the inner axial bins. The process is then repeated for the next most outer radial bin until the inner most radial bin is reached. **(L)** The final 3D density map gives a map of single molecule radial density rather than information about the 3D location of each single molecule.

bioRxiv preprint doi: https://doi.org/10.1101/188060; this version posted September 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



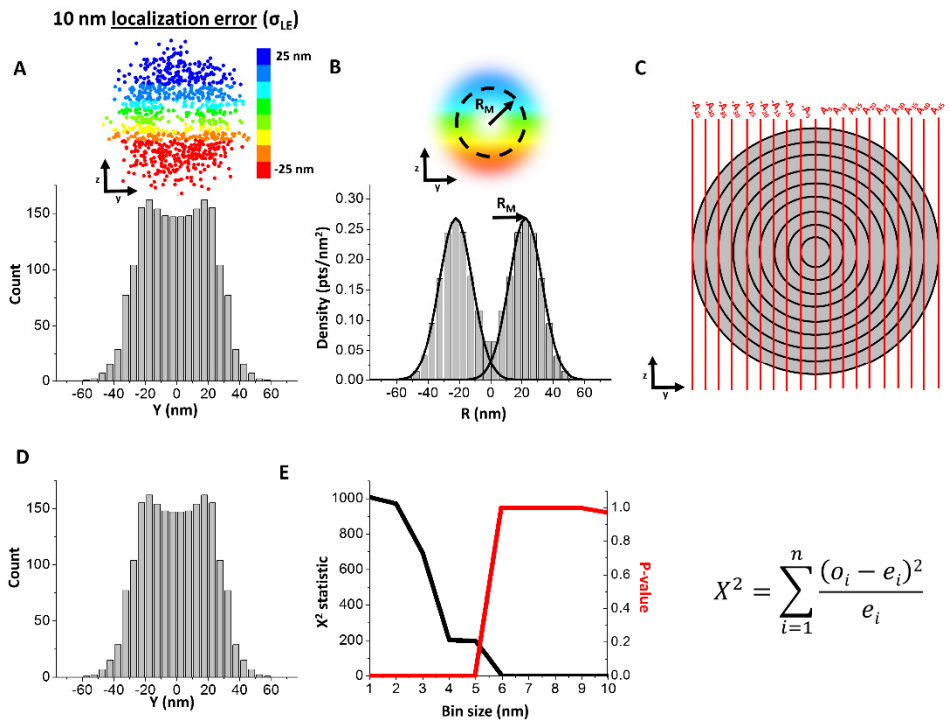
Supplementary Figure 2. **Minimum amount of experimental data points needed for a reliable 3D distribution of Imp β1 in the NPC**. **(A)** During the data selection process, single molecules with less than ~800-1000 photons are removed before the final 3D transformation analysis. **(B)** In the precision distribution, this corresponds to a cutoff precision of 10 nm. Each whole number precision value below 10 nm threshold contributes a certain percentage to the overall weighted average precision of 8.6 nm. **(C)** As an example, 10,000 single molecule points were simulated on an ideal radius (R_I) of 25 nm and given a localization error based on the precision distribution from (E). Grayscale values indicate precision assigned to each single molecule. **(D)** The 3D transformation from (F) shows that the spread of each peak is approximately equal to the weighted average or "effective" precision (σ_{EP}) of the distribution rather than the cutoff precision.

bioRxiv preprint doi: https://doi.org/10.1101/188060; this version posted September 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



Supplementary Figure 3. **SPEED microscopy and 3D transformation reproducibility percentage using a simulation-based approach.** For any given set of simulated data, the bin size is varied from 1 nm to the precision that is 10 nm in this example. **(A)** Data sets were simulated in three dimensions. Color bar indicates z position of the simulated points. **(B)** Each data set was simulated first with an ideal 25-nm radius (RI). (C) Subsequently, a localization error (σ_{LE}) of 10 nm was added to each point. **(C)** Using a 5-nm bin size for demonstration, the 2D histogram of the simulated data set with a 25-nm radius and 10-nm localization precision was determined. **(D)** 10,000 data sets were simulated with an ideal 25-nm radius (R_I) and a localization error (σ_{LE}) of 10 nm. The resultant 3D histograms were then each fitted with a Gaussian function to localize the mean position of each peak, which is designated as the mean radius R_M. **(E)** The histogram for all the R_M values was determined and the number of simulated data sets that fell within the acceptable range of $R_I \pm \sigma_{LE}$ were counted. The acceptable range of $R_I \pm \sigma_{LE}$ was chosen because, in principle, the Rayleigh criterion limited the resolution of any single 3D histogram to the spread of that distribution, which was due to the simulated localization error (σ_{LE}). After 10,000 simulations, the histogram for R_M values converges on the mean (R_I) from which they were originally sampled, while the spread of the R_M histogram (σ_{TR}) converges on a value that is due to the number of simulated points in each distribution and simulated localization error. **(F)** Reproducibility percentage was defined as the number of R_M values that fell within the acceptable range of $R_I \pm \sigma_{LE}$ divided by the total number of simulated data sets and multiplied by 100%.

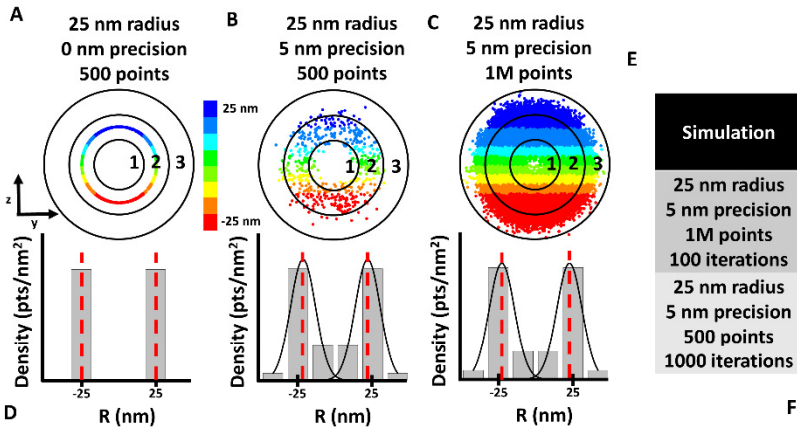
bioRxiv preprint doi: https://doi.org/10.1101/188060; this version posted September 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



$$X^2 = \sum_{i=1}^{n} \frac{(o_i - e_i)^2}{e_i}$$

Supplementary Figure 4. **Optimal bin size determination using Chi-square error analysis.** For any given set of simulated data, the bin size is varied from 1 nm to the precision that is 10 nm in this example. **(A)** Data sets were simulated in three dimensions. Color bar indicates z position of the simulated points. Each data set was simulated first with an ideal 25-nm radius (R_I). Subsequently, a localization error (σ_{LE}) of 10 nm was added to each point. Using a 5-nm bin size for demonstration, the 2D histogram of a simulated data set with a 25-nm radius and 10-nm localization precision was determined. **(B)** The 3D density histogram was then obtained via the 2D to 3D transformation algorithm and the peaks were fit with Gaussian distributions. **(C)** The 5-nm bin size area matrix was calculated and multiplied by the 3D density distribution in (C) to reconstruct the 2D distribution. **(D)** The values of the reconstructed 2D distribution were then compared bin-by-bin to the original 2D distribution (as shown in A) using the Chi-square analysis equation where 'o' refers to the observed histogram values in (D), 'e' refers to the expected histogram values in (A), 'i' refers to the bin, and 'n' refers to the total number of bins with histogram values in them. **(E)** The Chi-square statistic and p-value were then plotted across the potential bin size values. A p-value ≤ 0.05 indicates that the 2D histograms in (A) and (D) are different from each other, suggesting the lack of enough data to allow sufficient sampling from each bin. A p-value > 0.05 indicates that the 2D histograms are not statistically different and likely have enough data points to accurately measure the value in each bin. Chi-square analysis was performed 10 times for each set of simulation parameters.

bioRxiv preprint doi: https://doi.org/10.1101/188060; this version posted September 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

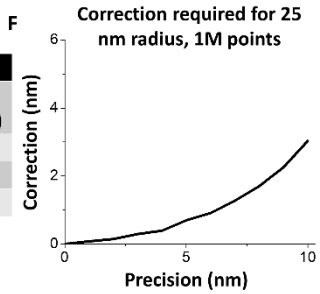


E

Simulation	Peak radius fit (nm)	Correction (nm)	Final radius (nm)
25 nm radius 5 nm precision 1M points 100 iterations	24.30 ± 0.02	+0.7	25.00
25 nm radius 5 nm precision 500 points 1000 iterations	24.30 ± 1.41	+0.7	25.00

D

Ring #	Ring area (nm²)	0 nm prec. – 500 pts		5 nm prec. – 500 pts		5 nm prec. – 1M pts	
	$(\pi r_n^2 - \pi r_{n-1}^2)$	Number of points	Density (pts/nm²)	Number of points	Density (pts/nm²)	Number of points	Density (pts/nm²)
1	842.66	0	0	43	0.49	68680	78.70
2	2617.99	500	0.19	414	0.16	819690	313.10
3	4363.32	0	0	43	0.01	111580	25.57



Supplementary Figure 5. **Sensitivity of inner bins necessitates slight correction of peak position during simulation. (A), (B), and (C)** Simulated single molecule data and corresponding 3D density histogram for simulation with 25 nm radius, 0 nm localization error, and 500 points. Red dashed lines indicate mean peak fitting. **(B)** Simulated single molecule data and corresponding 3D density histogram for simulation with 25 nm radius, 5 nm localization error, and 500 points. Red dashed lines indicate mean peak fitting. **(C)** Simulated single molecule data and corresponding 3D density histogram for simulation with 25 nm radius, 5 nm localization error, and 1 million points. Red dashed lines indicate mean peak fitting. **(D)** Table showing the calculation to obtain each bin of the 3D density histogram. **(E)** Table showing that even one million points does reconstruct a precise 25 nm peak fitting due to the fact that the inner radial bins have smaller area and are slightly more sensitive to changes in density. **(F)** Correction required for each precision up to 10 nm for a 25 nm radius. This correction process was performed before each simulation to accurately localize the R_M density peak and correlate it to the ideal R_I from which the data was simulated.

bioRxiv preprint doi: https://doi.org/10.1101/188060; this version posted September 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Acknowledgements
The project was supported by grants from the National Institutes of Health (NIH GM097037, GM116204 and RGM122552A to W.Y.).

Author contributions
A.R., J.K., W.L., and W.Y. designed experiments; A.R., J.K., W.L., and W.Y. performed single-molecule tracking and super-resolution SPEED microscopy experiments; A.R., J.K., and W.L. prepared plasmids and established cell lines; A.R., J.K., W.L., and W.Y. conducted data analysis and wrote the manuscript.

References

1 Leung, B. O. & Chou, K. C. Review of super-resolution fluorescence microscopy for biology. *Applied spectroscopy* 65, 967-980 (2011).
2 Huang, B., Bates, M. & Zhuang, X. Super-resolution fluorescence microscopy. *Annual review of biochemistry* 78, 993-1016 (2009).
3 Betzig, E. *et al.* Imaging intracellular fluorescent proteins at nanometer resolution. *Science* 313, 1642-1645 (2006).
4 Hell, S. W. & Wichmann, J. Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy. *Optics letters* 19, 780-782 (1994).
5 Huang, B., Wang, W., Bates, M. & Zhuang, X. Three-dimensional super-resolution imaging by stochastic optical reconstruction microscopy. *Science* 319, 810-813 (2008).
6 Pavani, S. R. P. *et al.* Three-dimensional, single-molecule fluorescence imaging beyond the diffraction limit by using a double-helix point spread function. *Proceedings of the National Academy of Sciences* 106, 2995-2999 (2009).
7 Akey, C. W. & Radermacher, M. Architecture of the Xenopus nuclear pore complex revealed by three-dimensional cryo-electron microscopy. *The Journal of Cell Biology* 122, 1-19 (1993).
8 Akey, C. W. Interactions and structure of the nuclear pore complex revealed by cryo-electron microscopy. *The Journal of Cell Biology* 109, 955-970 (1989).
9 Yang, W. & Musser, S. M. Nuclear import time and transport efficiency depend on importin β concentration. *The Journal of cell biology* 174, 951-961 (2006).
10 Ma, J., Goryaynov, A., Sarma, A. & Yang, W. Self-regulated viscous channel in the nuclear pore complex. *Proceedings of the National Academy of Sciences* 109, 7326-7331 (2012).
11 Ma, J. & Yang, W. Three-dimensional distribution of transient interactions in the nuclear pore complex obtained from single-molecule snapshots. *Proceedings of the National Academy of Sciences* 107, 7305-7310 (2010).
12 Czarnecki, P. G. & Shah, J. V. The ciliary transition zone: from morphology and molecules to medicine. *Trends in cell biology* 22, 201-210 (2012).
13 Moores, C. in *Methods in Cell Biology* Vol. 88 (ed Thomas Müller-Reichert and Paul Verkade) Ch. 18, 299-317 (2008).
14 Weis, K. Regulating access to the genome: nucleocytoplasmic transport throughout the cell cycle. *Cell* 112, 441-451 (2003).
15 Fried, H. & Kutay, U. Nucleocytoplasmic transport: taking an inventory. *Cellular and molecular life sciences* 60, 1659-1688 (2003).

bioRxiv preprint doi: https://doi.org/10.1101/188060; this version posted September 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

16 Fahrenkrog, B. & Aebi, U. The nuclear pore complex: nucleocytoplasmic transport and beyond. *Nature reviews. Molecular cell biology* 4, 757 (2003).
17 Marshall, W. F. & Nonaka, S. Cilia: tuning in to the cell's antenna. *Current Biology* 16, R604-R614 (2006).
18 Scholey, J. M. & Anderson, K. V. Intraflagellar transport and cilium-based signaling. *Cell* 125, 439-442 (2006).
19 Thompson, R. E., Larson, D. R. & Webb, W. W. Precise nanometer localization analysis for individual fluorescent probes. *Biophysical journal* 82, 2775-2783 (2002).
20 Händel, M. *et al.* Selective targeting of somatostatin receptor 3 to neuronal cilia. *Neuroscience* 89, 909-926 (1999).
21 Ye, F. *et al.* Single molecule imaging reveals a major role for diffusion in the exploration of ciliary space by signaling receptors. *Elife* 2, e00654 (2013).
22 Yang, T. T. *et al.* Superresolution pattern recognition reveals the architectural map of the ciliary transition zone. *Scientific reports* 5, 14096 (2015).
23 Sage, D. *et al.* Quantitative evaluation of software packages for single-molecule localization microscopy. *Nature methods* 12, 717 (2015).
24 Weber, K., Rathke, P. C. & Osborn, M. Cytoplasmic microtubular images in glutaraldehyde-fixed tissue culture cells by electron microscopy and by immunofluorescence microscopy. *Proceedings of the National Academy of Sciences* 75, 1820-1824 (1978).
25 Chen, F., Tillberg, P. W. & Boyden, E. S. Expansion microscopy. *Science* 347, 543-548 (2015).
26 Chang, J.-B. *et al.* Iterative expansion microscopy. *Nature* 201, 7 (2017).
27 Mortensen, K. I., Churchman, L. S., Spudich, J. A. & Flyvbjerg, H. Optimized localization analysis for single-molecule tracking and super-resolution microscopy. *Nature methods* 7, 377-381 (2010).
28 Quan, T., Zeng, S. & Huang, Z.-L. Localization capability and limitation of electron-multiplying charge-coupled, scientific complementary metal-oxide semiconductor, and charge-coupled devices for superresolution imaging. *Journal of biomedical optics* 15, 066005-066005-066006 (2010).
29 Robbins, M. S. & Hadwen, B. J. The noise performance of electron multiplying charge-coupled devices. *IEEE transactions on Electron Devices* 50, 1227-1232 (2003).
30 Deschout, H., Neyts, K. & Braeckmans, K. The influence of movement on the localization precision of sub-resolution particles in fluorescence microscopy. *Journal of biophotonics* 5, 97-109 (2012).
31 von Appen, A. *et al.* In situ structural analysis of the human nuclear pore complex. *Nature* 526, 140 (2015).
32 Protein Data Bank Japan (PDBj). Date accessed: 8/3/2017
33 Blair, H. J. *et al.* Evc2 is a positive modulator of Hedgehog signalling that interacts with Evc at the cilia membrane and is also found in the nucleus. *BMC biology* 9, 14 (2011).
34 https://openi.nlm.nih.gov. Date accessed: 8/3/2017. This was modified from original version. License: https://creativecommons.org/licenses/by/2.0/
35 Ti, S.-C. *et al.* Mutations in human tubulin proximal to the kinesin-binding site alter dynamic instability at microtubule plus-and minus-ends. *Developmental cell* 37, 72-84 (2016).

bioRxiv preprint doi: https://doi.org/10.1101/188060; this version posted September 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.