HAL
open science

Development of advanced methods for super-resolution microscopy data analysis and segmentation

Leonid Andronov

To cite this version:

Leonid Andronov. Development of advanced methods for super-resolution microscopy data analysis and segmentation. Biophysics. Université de Strasbourg, 2018. English. NNT : 2018STRAJ001 . tel-02918100

HAL Id: tel-02918100

https://theses.hal.science/tel-02918100v1

Submitted on 20 Aug 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ DE STRASBOURG

ÉCOLE DOCTORALE des Sciences de la Vie et de la Santé

Centre de Biologie Intégrative, IGBMC, UMR7104, Illkirch

THÈSE présentée par :

Leonid ANDRONOV

soutenue le : **09 janvier 2018**

pour obtenir le grade de : **Docteur de l'université de Strasbourg**

Discipline/ Spécialité : Biophysique et biologie structurale

Development of advanced methods for super-resolution microscopy data analysis and segmentation

Développement de méthodes avancées pour l'analyse et la segmentation de données de microscopie à super-résolution

THÈSE dirigée par :

Dr. KLAHOLZ Bruno P.	DR, CNRS, IGBMC, Illkirch

RAPPORTEURS :

Dr. SALAMERO Jean	DR, CNRS, Institut Curie, Paris
Dr. SIBARITA Jean-Baptiste	IR, CNRS, Institut Interdisciplinaire de Neurosciences, Bordeaux

---

AUTRES MEMBRES DU JURY :

Dr. DAVIDSON Irwin	DR, Inserm, IGBMC, Illkirch

Acknowledgments

First of all, I would like to thank my thesis advisor Bruno Klaholz for giving me the opportunity to work in an excellent team, providing exceptional support and motivation.

I thank the committee members Jean-Baptiste Sibarita, Jean Salamero and Irwin Davidson, who agreed to evaluate my work.

I also thank all current and past members of our team for interesting discussions, particularly during lab meeting. I would like to specially acknowledge Alexander Myasnikov for his help with Imagic and other cryo-EM image processing software, Igor Orlov for his ideas on Voronoi diagrams and Jonathan Michalon in regards to visualization software for 3D Voronoi diagrams and IT support. Also thanks to Ottilie, Nadia, Kundhavai, Evelina, Christine and other people, who were around for chatting

I thank the engineers at the Imaging Center of the IGBMC for their support and trainings in light microscopy. I especially thank the head of the Imaging Platform Jean-Luc Vonesch, who provided organizational and scientific suggestions. Also, I am grateful to Yves Lutz, who taught me handling of cells and helped a lot with immunofluorescence and other aspects of sample preparation for most of my projects. Of course I cannot forget Pascal Kessler, who was always ready to help on different aspects and with whom we continue super-resolution experiments now. And as well Didier Hentsch, who took care of many organizational and technical questions.

I thank Ali Hamiche and post-docs from his team, Khalid Ouararhni and Muhammad Shuaib, together with whom we worked on the CENP-A project. I thank as well Evi Soutoglou and Charlene Lemaitre who was a PhD student in Evi's team, with whom we did super-resolution imaging in the framework of their DNA repair project.

I thank my friends and colleagues from the Faculty of Pharmacy and other Ukrainians in Strasbourg, with whom I spent a lot of time together and who brought much fun into my life in France.

Finally, I am grateful to my family who was always involved and supportive despite the distance that separates us.

Summary

Since the 19th century the resolution of microscopy is known to be limited by diffraction to the order of magnitude of the light wavelength. For the visible light this limit is around 200 nm, which makes precise fluorescent imaging of protein complexes non feasible. This limitation seemed to be fundamental until the invention of super-resolution microscopy in the beginning of the 21st century. Super-resolution microscopy allows to improve the resolution up to 10 times in nowadays routine experiments, and even up to 100 times in some special cases. Among numerous super-resolution techniques, single-molecule localization microscopy (SMLM) is remarkable, first of all, for best practically achievable resolution but also for the direct access to properties of individual molecules. Huge efforts have been made in technical developments of SMLM, and now it is a method commonly used in biological research with a wide choice of markers and experimental procedures, suitable for 2D and 3D imaging of fixed and live cells and tissues.

The primary data format of SMLM is a list of the coordinates (and other properties) of individual fluorophores, which is not the most common data type for microscopy. Therefore, specially adapted methods for processing of this data had to be developed in order to get a maximum of information from the single-molecule coordinates. Some methods for SMLM data processing were developed or adapted from other fields, but they lacked automatization and integration. The first objective of my project was therefore development of a software, which would combine the most useful post-processing steps in a single tool. We developed such a tool, SharpViSu (Andronov *et al.*, 2016a), with features such as: a graphical user interface, correction of drift and chromatic aberrations, selection of localization events based on their intensity and frame number, reconstruction of data into 2D images or 3D volumes with different visualization techniques, estimation of resolution with Fourier ring correlation, possibility of extension with plugins. The software works with different popular formats of localization tables and produces images and graphs in formats useful for analysis and publications.

For drift correction, we implemented a cross-correlation-based method which calculates drift directly from data, without the need for fiducial markers. We have shown that this method produces best results when it is applied in an iterative way. Also, we introduced a new visualization method, based on Voronoi tessellations and we showed that it preserves resolution equally well or better than the histogram or Gaussian display, which are also available in SharpViSu.

One of the common applications for fluorescence microscopy is cluster analysis. With SMLM this becomes even more important because of the possibility to study clustering of proteins at the subdiffraction scale. Methods for image segmentation developed for conventional microscopy are not optimal for SMLM because the image is a secondary data type in SMLM, while the primary data are the localizations and they should be used directly in order to get the most out of the experiments. We therefore developed a new method, ClusterViSu (Andronov *et al.*, 2016b) that is based on the elegant geometry and properties of Voronoi tessellations. The method compares the distribution of the sizes of the Voronoi polygons of the experimental points with that of the noise model (randomly distributed points). This noise modeling allows to determine the segmentation threshold, the minimal local density of molecules in the clusters, in an automated manner. We proposed to estimate the local density in the neighborhood of a molecule as the inverse value of the surface area of the Voronoi polygon that contains this molecule. This provides a simple and unambiguous estimation, unlike other methods that count molecules inside circles of a given radius. By interpolating the local densities, associated with every molecule, to a regular grid, an image called a density map, can be calculated. This map should be binarized using the threshold determined as described above.

We applied this method for cluster analysis of simulated data, as well as nuclear pore complexes, chromatin and RNA Polymerase II SMLM data. The characteristics of clusters obtained with ClusterViSu compare well with results obtained with other methods and reported in the literature. Besides segmentation, this technique also allows for colocalization estimation between distributions of two proteins, as we demonstrated on an example of RNA Polymerase II and histone H2B.

ClusterViSu, as well as most of the other previously published segmentation methods for SMLM, were implemented only in two dimensions. While 2D imaging is common in SMLM, the most challenging biological objects, such as chromatin, require three dimensional information for drawing correct conclusions about their structure. As I show in this work, processing of 3D data only in two dimensions can strongly distort the result of segmentation, produce meaningless output or make the segmentation completely impossible, especially for dense structures with high levels of background noise (Andronov *et al.*, 2017). This is why SMLM data processing methods have to be implemented also in 3D. While many methods can be extended for the third dimension, in fact only few ready to use 3D tools exist so far. This is also because of the extended complexity of 3D data that may strongly increase the processing time.

We therefore explored the segmentation of 3D localization data and extended ClusterViSu for the third dimension (3DClusterViSu, Andronov *et al.*, 2017). We have shown that the 2D concept works as well for 3D. We validated 3DClusterViSu on simulated data and then applied it to cluster analysis of histones H2B and CENP-A. Our data shows that H2B forms heterogeneous clusters in the cell nuclei, while CENP-A forms cluster of relatively homogeneous size with a diameter of 260 nm ± 54 nm in the centromeric regions of chromatin in the middle of the G1 phase of the cell cycle. We have also shown on an example of microtubules that the Voronoi segmentation method can be used for noise reduction if only the localizations with the local density higher than the threshold are kept. Since our 3D data was obtained with the most widely used astigmatism method, the density of localizations was not isotropic, it drops while going out of focus. Because of that we used a modified noise model, the complete spatially random distribution in the X and Y dimensions and a Gaussian distribution in the Z dimension. 3DClusterViSu is written in Matlab and available as a stand-alone application. Visualization of 3D Voronoi diagrams is made possible with another, Python-based tool.

The last part of my thesis project was super-resolution imaging of centromeric chromatin throughout the cell cycle. The centromere is typically defined epigenetically with the CENP-A protein that replaces the canonical histone H3 in some of nucleosomes of the centromeric chromatin. A peculiarity of these nucleosomes is that CENP-A, unlike other histones, is not deposited during DNA synthesis, but is deposited later, in early G1 phase. The detailed structure of chromatin, including the centromeric one, is still poorly understood; we therefore aimed at finding patterns in the subdiffraction-scale structure of centromeric chromatin during the CENP-A deposition in G1 phase. Using SMLM and advanced data processing methods, we demonstrated for the first time that CENP-A forms clusters of about the same size for different chromosomes, the shape of the clusters transforms upon the CENP-A deposition in early G1 from a hollow to a filled shape, the CENP-A chaperone HJURP is found approximately in the middle of CENP-A clusters in early G1 phase. Even though the exact mechanism of this transformation remains unknown, this study sheds light on the structural transformations of the centromeric chromatin during the CENP-A deposition.

Since we found other shapes of CENP-A clusters in non-synchronized cells, it could be interesting in the future to study the centromeric chromatin by 3D imaging with a high temporal resolution over the whole cell cycle. Additionally, in order to solve the chromatin structure at the range of ~10–30 nm, studies with electron microscopy, focused ion beam and correlative

light-electron microscopy will need to be performed, a promising area of research where SMLM will play a central role.

Résumé de thèse

Depuis le XIXe siècle il est connu que la résolution de la microscopie est limité par la diffraction jusqu'à l'ordre de magnitude le la longueur d'onde de la lumière. Pour la lumière visible cette limite est d'environ 200 nm, ce qui rend l'imagerie de complexes protéiques par la fluorescence non réalisable. Cette limitation semblait d'être insurmontable jusqu'à l'invention de la microscopie à super-résolution au début du XXIe siècle. La microscopie à super-résolution a permis d'améliorer la résolution d'un facteur 10 dans les expériences de routine, et même jusqu'à 100 fois dans certains cas spécifiques. Parmi les nombreuses techniques de super-résolution, la microscopie par localisation de molécules individuelles (single-molecule localization microscopy, SMLM, **Fig. I**) se distingue principalement par sa meilleure résolution réalisable en pratique mais aussi pour l'accès direct aux propriétés des molécules individuelles. Des grands progrès ont été réalisés dans des développements techniques de la SMLM, ce qui en fait une méthode couramment utilisée dans la recherche biologique avec un vaste choix des sondes et des procédures expérimentales, qui convient pour l'imagerie en 2D et en 3D des cellules et des tissus fixés et vivants.

Les données primaires de la SMLM sont une liste des coordonnées (et d'autres propriétés) des fluorochromes individuels, qui ne sont pas des données couramment utilisées en microscopie. Des méthodes spécialement adaptées pour le traitement de ces données devaient donc être développées afin que l'on puisse obtenir le plus d'information des coordonnées des molécules individuelles. Quelques méthodes de traitement des données de la SMLM ont été développé ou adaptées à partir d'autres domaines, mais elles n'étaient pas automatisées ni intégrées. Le premier objectif de mon projet était donc le développement d'un logiciel combinant les étapes du post-traitement les plus importantes selon notre expérience, dans un seul outil. Nous avons développé le logiciel SharpViSu (Andronov *et al.*, 2016a), avec les fonctionnalités suivantes : une interface graphique, une correction des dérives et des aberrations chromatiques, une sélection des événements de localisations basée sur leurs intensité et numéro d'image, une reconstruction des données dans des images 2D ou dans des volumes 3D par le moyen de différentes techniques de visualisation (**Fig. II**), une estimation de la résolution à l'aide de la corrélation des anneaux de Fourier, une possibilité d'extension avec des plugins. Le logiciel fonctionne avec différents formats populaires des tableaux de localisation et crée des images et des graphiques dans les formats les plus utiles pour l'analyse et la présentation de résultats.

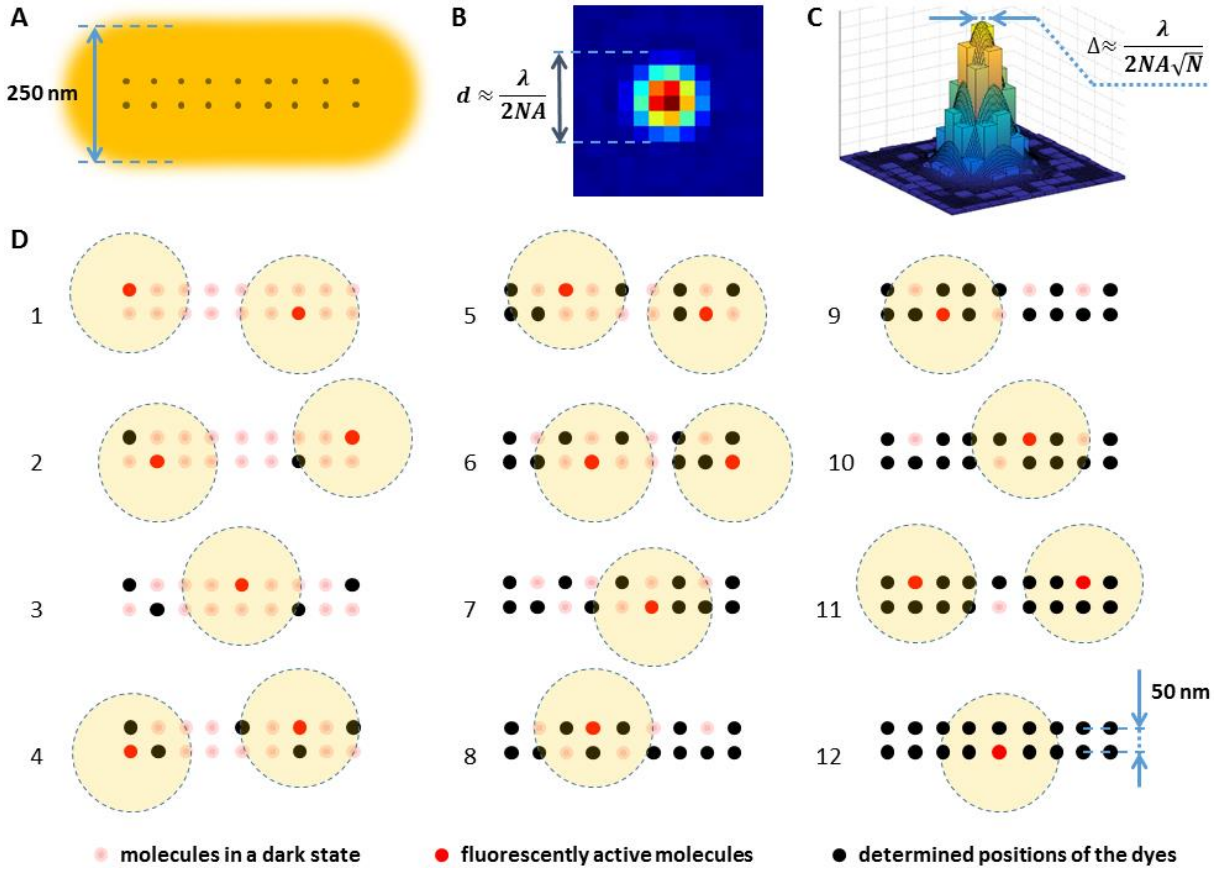


Figure I. Le principe de la microscopie par localisation de molécules individuelles. (**A**) Dans la microscopie conventionnelle à fluorescence tous les fluorochromes émettent la lumière en simultané. Par conséquent, les détails de l'échantillon plus petits que l'ordre de la longueur d'onde de la lumière sont cachés à cause de la diffraction. (**B**) Une seule molécule de colorant est projetée sur une caméra dans une forme d'un disque de diamètre *d*. (**C**) Le centre de l'image de la molécule individuelle, qui correspond à la position de la molécule, peut être déterminé avec la précision Δ, qui est plus élevée que *d* et qui dépend du nombre de photons *N* détecté de cette molécule (9). (**D**) Pendant une expérience SMLM, à chaque instant la plupart des colorants se trouve dans un état sombre non-fluorescent (points rouges clairs). Un petit nombre de fluorochromes aléatoires, moins qu'un par la région limitée par la diffraction, est activé et imagé séparément en tant que des molécules individuelles (points rouges). Les positions de ces molécules peuvent donc être déterminées avec une haute précision Δ (points noirs). Une fois ces colorants sont blanchis ou renvoyés dans l'état sombre, une autre portion aléatoire de fluorochromes peut être imagée et localisée sur des clichés suivants. L'image complète à super-résolution peut être reconstruite depuis les coordonnés après que la plupart des colorants est passée le cycle d'activation – localisation – désactivation (**D, 12**). **1–12** signifient le numéro du cliché. Les cercles jaunes représentent la taille latérale de la PSF.

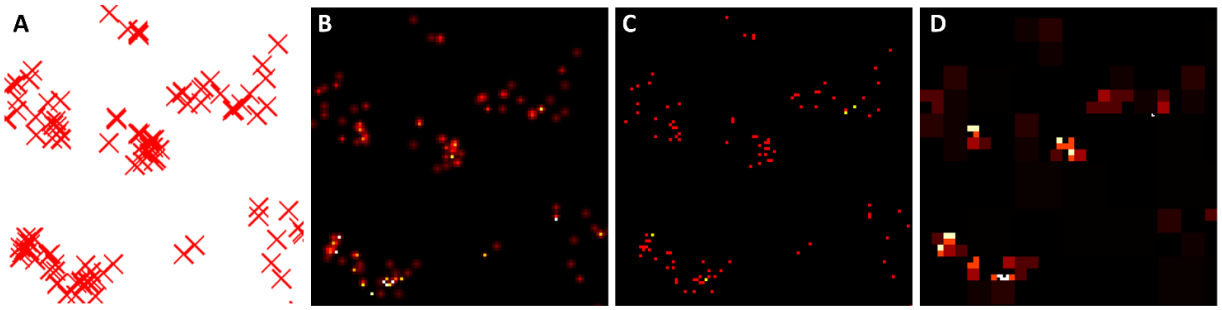


Figure II. Des différentes méthodes de visualisation de données de la SMLM. (**A**) Un nuage de points. (**B**) La visualisation de chaque molécule dans une forme d'une fonction gaussienne. (**C**) Un histogramme 2D. (**D**) Un histogramme adaptative avec la capacité maximale de la classe de 2 localisations. Les données: la protéine TPR dans une cellule U2OS, marquée avec les anticorps secondaires conjugué Alexa 647 (Lemaître et al., 2014). 1 µm x 1 µm FOV.

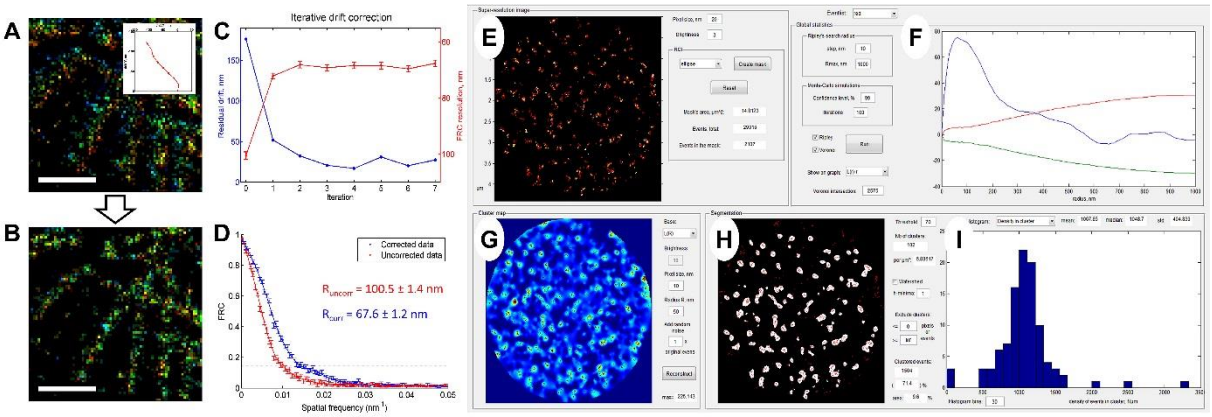


Figure III. Les fonctionnalités de SharpViSu. (**A, B**) Un fragment de 1,5 µm x 1,5 µm d'une image super-résolutive de la β-tubuline dans une cellule HeLa, reconstruite avec l'encodage du temps dans la teinte, avant (A) et après 7 itérations de la correction de dérive (B). Le tracé de la dérive obtenu par SharpViSu est indiqué dans l'encart. La barre d'échelle, 500 nm (**C**) La réduction de la dérive résiduelle (en bleu) et l'amélioration de la résolution estimée par FRC (en rouge) en itérant la correction de dérive. (**D**) FRCs des données avant et après la correction démontrent une amélioration significative de la résolution. (**E-I**) Interface de ClusterViSu, un plugin pour la segmentation de données de la SMLM. (**E**) Une région d'intérêt est sélectionnée. (**F**) La statistique sur les localisations avec la fonction L(r)-r de Ripley des données expérimentales (en bleu) et l'intervalle de confiance de 99% pour les points distribués de façon aléatoire (en rouge et en vert) démontrent un clustering. (**G**) La carte de densités calculée sur la base de la fonction L de Ripley. (**H**) La carte de densités binarisée. (**I**) L'histogramme de la densité des localisations dans les clusters. La figure est adaptée de (Andronov et al., 2016a).

Pour la correction de la dérive, nous avons implémenté une méthode basée sur la cross-corrélation, qui calcule la dérive directement à partir des données, sans nécessité de marqueurs fiduciaires. Nous avons démontré que cette méthode produit de meilleurs résultats quand elle est employée de façon itérative (**Fig. III**). En outre, nous avons introduit une nouvelle méthode de visualisation, basée sur les diagrammes de Voronoï. Nous avons démontré qu'elle conserve la résolution aussi bien voire mieux que l'affichage par un histogramme ou par des fonctions gaussiennes, qui sont aussi disponibles dans SharpViSu.

Une des applications ordinaires de la microscopie à fluorescence est le partitionnement des données en clusters. Elle est devenue encore plus importante avec l'invention de la SMLM en raison de la possibilité d'étudier le clustering de protéines à l'échelle sous-diffractionnelle. Les méthodes de segmentation d'images développées pour la microscopie conventionnelle ne sont pas optimales pour le SMLM parce que les images super-résolutives sont des données secondaires pour la SMLM. Les données primaires sont les localisations qui doivent être utilisées directement afin d'obtenir le maximum d'informations à partir des expériences. Nous avons donc développé une telle méthode, ClusterViSu (Andronov *et al.*, 2016b) qui est basé sur la géométrie élégante des diagrammes de Voronoï (**Fig. IV**). La méthode compare la distribution des tailles des polygones de Voronoï des points expérimentaux avec celle du modèle de bruit (les points distribués d'une façon aléatoire). Cette modélisation de bruits permet de déterminer le seuil de segmentation, c'est-à-dire la valeur minimale de la densité locale de molécules dans les clusters, par une procédure automatique. Nous avons proposé d'estimer la densité locale dans l'environnement d'une molécule comme la valeur inverse de la surface du polygone de Voronoï qui contient cette molécule. Ceci donne une estimation simple et non ambiguë, contrairement à d'autres méthodes qui comptent les molécules dans des cercles d'un certain rayon. En interpolant les densités locales, associées avec chaque molécule, sur une grille régulière, une image (qui s'appelle la carte de densité) peut être calculée. Cette carte doit être binarisée en utilisant le seuil déterminé par la méthode décrite ci-dessus.

Nous avons appliqué cette méthode au partitionnement de données simulées, du complexe du pore nucléaire, de la chromatine et de l'ADN polymérase II. Les caractéristiques des clusters obtenues avec ClusterViSu correspondent aux résultats obtenus en utilisant d'autres méthodes et rapportés dans la littérature. En plus de la segmentation, cette technique permet aussi d'estimer la colocalisation entre des distributions de deux protéines, comme nous avons démontré sur un exemple d'ADN polymérase II avec l'histone H2B.

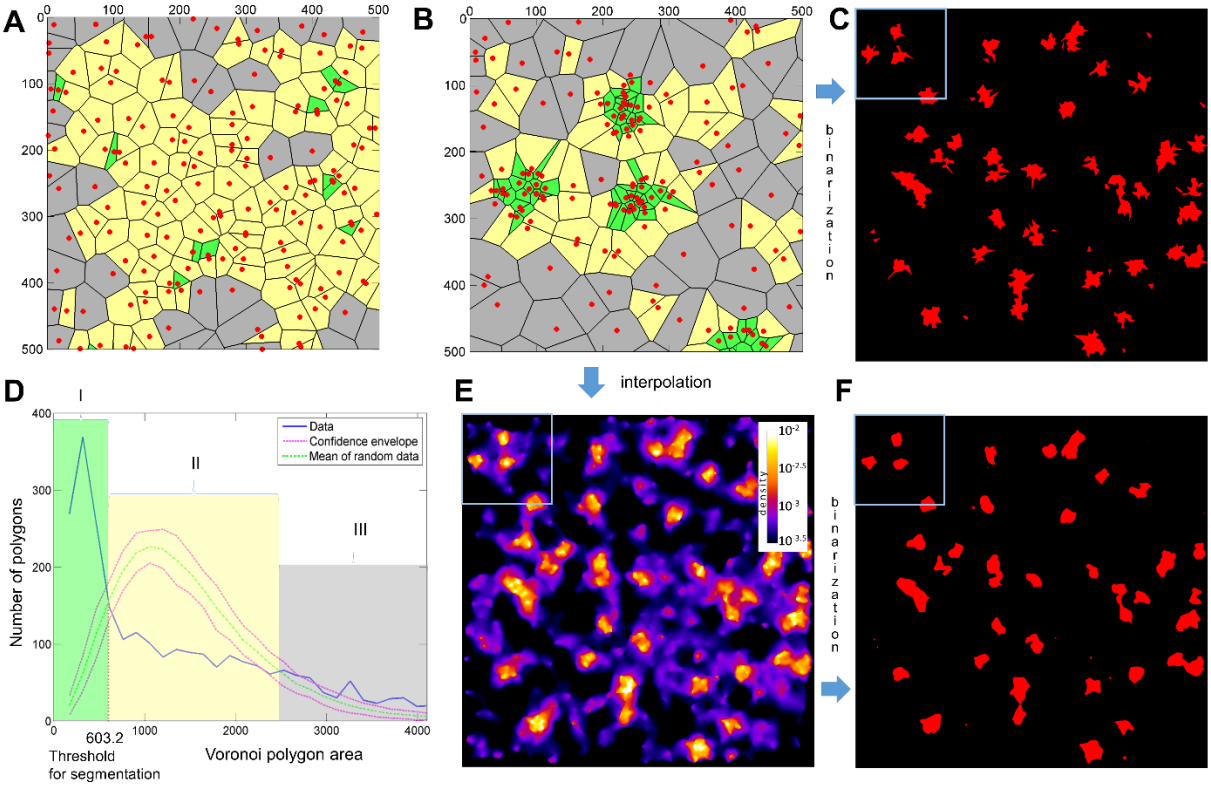


Figure IV. Le principe de segmentation basé sur les diagrammes de Voronoï. (**A**) Une région de données synthétiques avec les points distribués aléatoirement. (**B**) Un jeu de données clustérisées synthétique avec le même nombre de points que dans les données distribuées aléatoirement. (**C**) Le partitionnement obtenu par la segmentation directe du diagramme de Voronoï permet la visualisation des clusters. (**D**) La distribution des surfaces des polygones de Voronoï des données clustérisées (en bleu) et la distribution moyenne des polygones de Voronoï des données aléatoires (en vert) avec l'enveloppe de confiance obtenu par les simulations Monte-Carlo, ce qui permet à définir la valeur du seuil pour la segmentation automatisée. Les trois régions caractéristiques : les petits polygones clustérisés (en vert), les polygones intermédiaires correspondants à la distribution aléatoire (en jaune), les grands polygones se trouvant au fond dans la distribution clustérisée. (**E**) L'interpolation des densités locales à une grille de pixels produit une carte de densités. (**F**) Le partitionnement obtenu par le seuillage de la carte de densités. Les cadres bleus dans (C, E, F) correspondent à la région indiquée en (B). La figure est adaptée de (Andronov et al., 2016b).

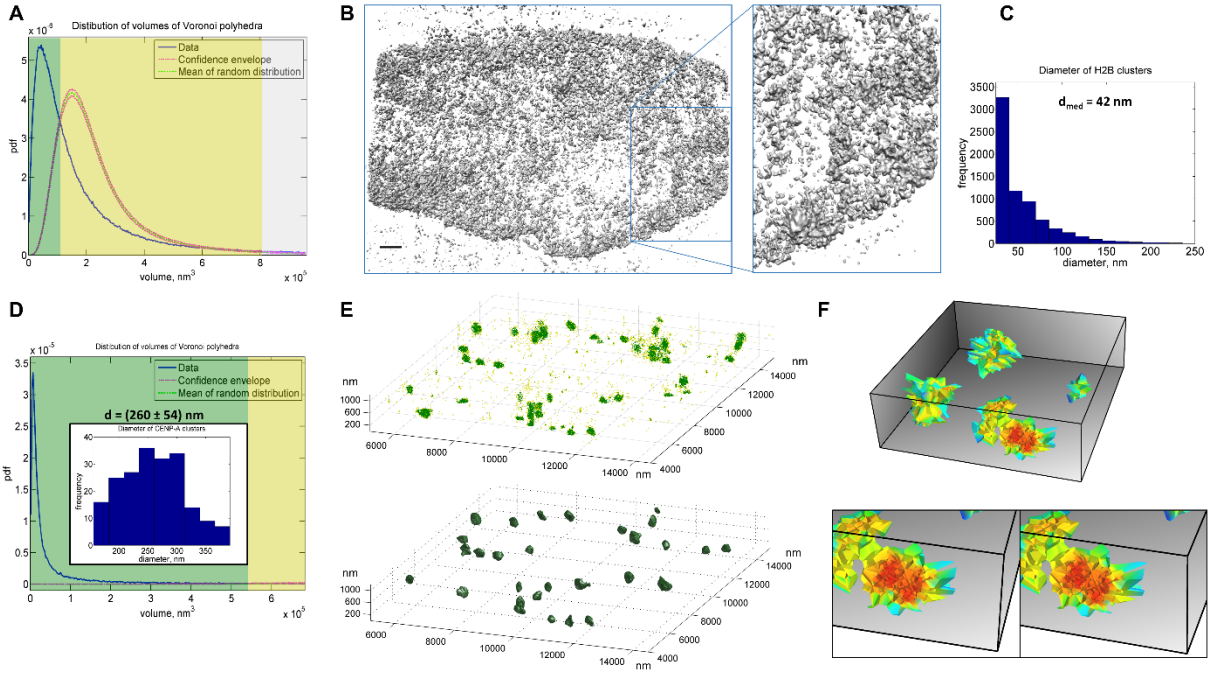


Figure V. Le partitionnement de données des histones H2B et CENP-A par les diagrammes 3D de Voronoï. (**A-C**) L'histone H2B détectée avec les anticorps secondaires marqués par Alexa-647 dans une cellule HeLa. (**A**) La distribution des volumes des polyèdres de Voronoï indique un clustering de l'histone H2B. (**B**) La carte de densité en 3D représentée comme une isosurface au niveau du seuil dans le logiciel Chimera (Pettersen et al., 2004); barre d'échelle, 1µm. (**C**) L'histogramme des diamètres équivalents des clusters d'H2B avec la valeur médiane de 42 nm. (D-F) CENP-A détecté avec les anticorps secondaires marqués par Alexa-647 dans des cellules U2OS dans une phase G1 avancée. (**D**) La distribution des volumes des polyèdres de Voronoï démontre un fort clustering de CENP-A qui permet la quantification de ses propriétés. L'histogramme des diamètres des centromères est affiché dans l'encart. (**E**) Un nuage de molécules colorées selon la densité locale (en haut) ; la carte de densité segmentée montre les clusters de CENP-A dans les régions centromèriques de la cellule (en bas). (**F**) La représentation 3D des clusters de CENP-A affichés comme une tesselation de Voronoï avec notre outil de visualisation en 3D (en bas : la région zoomée dans une représentation stéréo ; le diamètre moyen est de 260 nm contenant en moyenne 418 localisations dans un cluster 3D donné). La figure est adaptée de (Andronov et al., 2017).

Nous avons donc exploré la segmentation de données de localisation en 3D et élargi ClusterViSu sur la troisième dimension (3DClusterViSu, Andronov *et al.*, 2017). Nous avons démontré que le concept 2D fonctionne aussi pour la 3D. Puisque nos données 3D étaient obtenues en utilisant la méthode d'astigmatisme qui est la plus populaire, la densité de localisations n'était pas isotrope, car elle était plus basse pour les localisations hors focus. Pour cette raison, nous avons utilisé un modèle de bruit modifié: une distribution uniforme aléatoire dans le plan (X,Y) et une distribution Gaussienne aléatoire le long de la direction Z. Nous avons validé 3DClusterViSu sur des données simulées puis nous l'avons utilisé sur le partitionnement de localisations des histones H2B et CENP-A (**Fig. V**). Nos données ont montré que H2B forme des clusters hétérogènes dans les noyaux cellulaires, tandis que CENP-A forme des clusters d'une taille relativement homogène avec le diamètre de 260±54 nm dans les régions centromériques de la chromatine au milieu de la phase G1 du cycle cellulaire. Nous avons aussi démontré avec un exemple de microtubules que la méthode de segmentation de Voronoï peut être utilisée pour la réduction du bruit si l'on garde uniquement les localisations avec la densité locale plus importante que le seuil. 3DClusterViSu est écrit en Matlab et est disponible comme une application standalone. La visualisation des diagrammes de Voronoï en 3D est rendue possible à l'aide d'un autre outil écrit en Python.

La dernière partie de mon projet de thèse était l'imagerie à super-résolution de la chromatine centromérique tout au long du cycle cellulaire. Le centromère est défini épigénétiquement par la protéine CENP-A, qui remplace une histone canonique, H3, dans certains nucléosomes de la chromatine centromérique. Une particularité de ces nucléosomes est que CENP-A, à la différence des autres histones, n'est pas déposé pendant la synthèse de l'ADN, mais elle est déposée plus tard, au début de la phase G1. La structure détaillée de la chromatine, et notamment de la chromatine centromérique, est toujours très peu connue; nous avons donc cherché à identifier des régularités dans la structure de la chromatine centromérique à l'échelle sous-diffractionnelle tout au long de la déposition de CENP-A dans la phase G1.

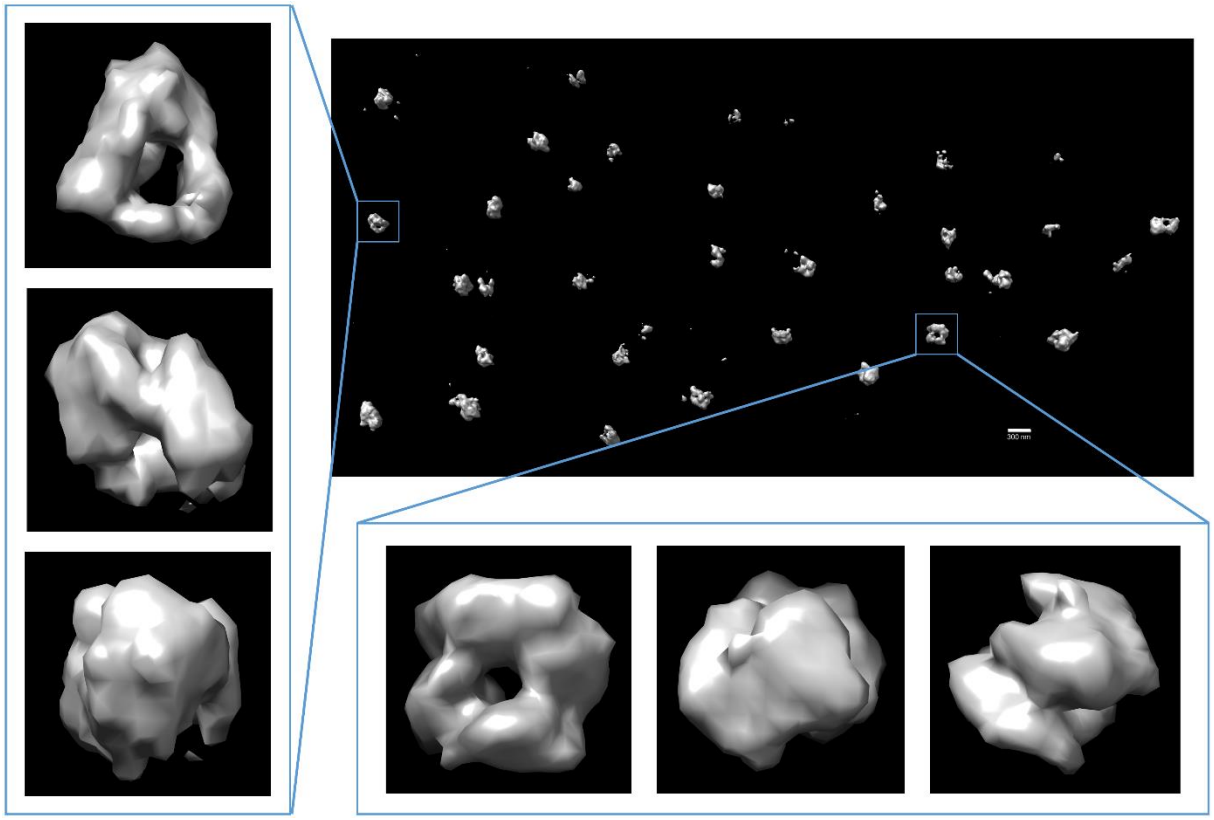


Figure VI. L'imagerie en SMLM 3D de CENP-A dans une cellule U2OS fixée à 1.5 heures après la mitose. Les clusters creux sont visibles. L'image est créée comme une carte de densité en 3D (Andronov et al., 2017) et affichée avec le logiciel Chimera (Pettersen et al., 2004). Les images agrandies montrent des vues pivotées d'une particule. La barre d'échelle, 300 nm; la taille des boxes, 600 nm.

En utilisant SMLM et des méthodes avancées de traitement de données, nous avons démontré pour la première fois que CENP-A s'assemble dans des clusters de la taille similaire pour des chromosomes différents; la forme des clusters change d'une forme creuse (**Fig. VI**) à une forme remplie pendant la déposition de CENP-A au début de la phase G1; le chaperon de CENP-A HJURP se trouve à peu près au centre des clusters de CENP-A au début de la phase G1 (**Fig. VII**). Même si le mécanisme précis de telles transformations reste inconnu, cette étude fait la lumière sur les transformations structurelles de la chromatine centromérique au cours de la déposition de CENP-A.

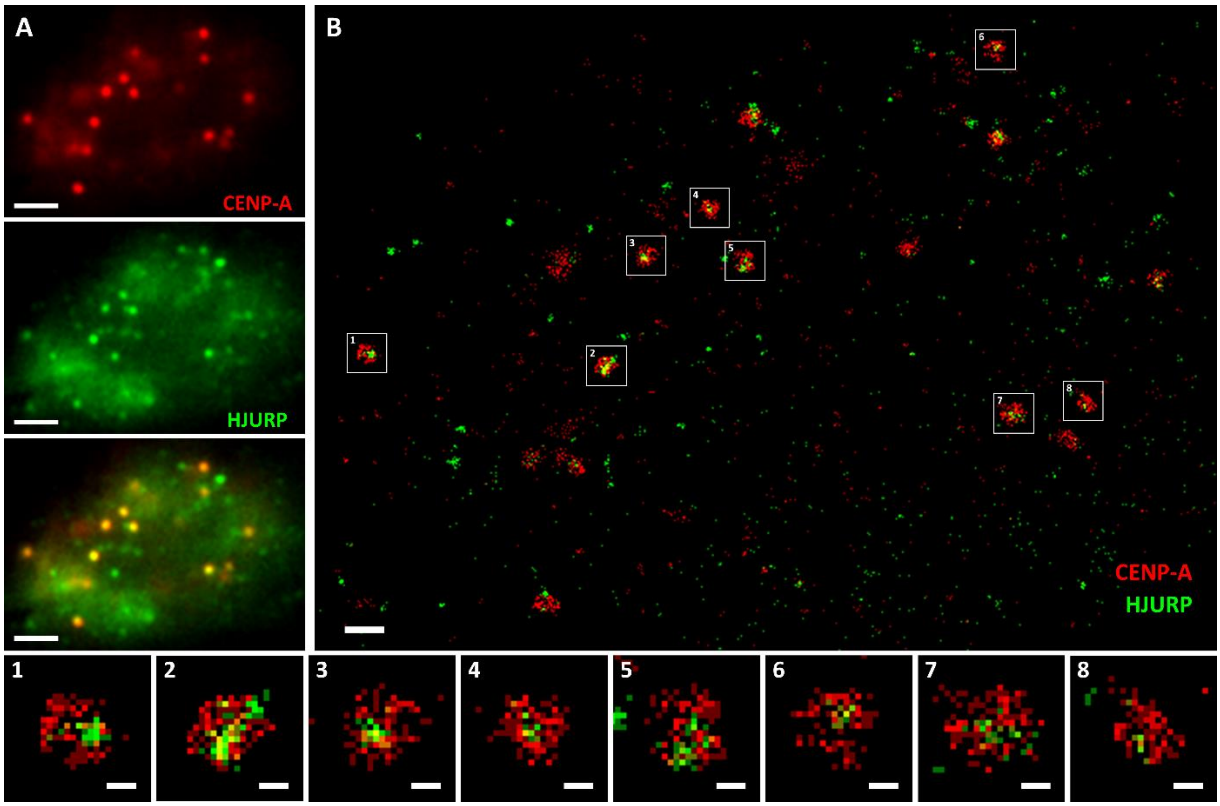


Figure VII. L'imagerie en SMLM de CENP-A avec son chaperon HJURP. (**A**) Les images à fluorescence en champs large de CENP-A et HJURP démontrent leur colocalisation au niveau des centromères. (**B**) La SMLM démontre que CENP-A (en rouge) et HJURP (en vert) forment des clusters de tailles différentes qui ne colocalisent pas complétement. (Panels en bas **1**–**8**) Les centromères agrandis de l'image B mettent en évidence qu'un cluster d'HJURP se situe souvent dans le centre vide d'un cluster de CENP-A. CENP-A était marqué avec les anticorps secondaires conjugués Alexa 647 et HJURP avec ceux conjugués Alexa 555. Les images en super-résolution (B, 1–8) sont reconstruites dans le mode d'histogramme avec la taille de classe de 20 nm. Les barres d'échelle, 2 µm (A), 500 nm (B) et 100 nm (1–8).

Puisque nous avons trouvé d'autres formes des clusters de CENP-A dans des cellules non-synchronisées, dans l'avenir les études de la chromatine centromérique peuvent être complétées avec une imagerie 3D avec une plus haute résolution temporelle sur l'ensemble du cycle cellulaire. De plus, afin de résoudre la structure chromatinienne à l'échelle de ~10–30 nm, des études par microscopie électronique, par sonde ionique focalisée et par microscopie corrélative devront être effectuées.

Articles

1. Andronov, L., Michalon, J., Ouararhni, K., Orlov, I., Hamiche, A., Vonesch, J.-L., Klaholz, B. (2017). 3DClusterViSu: 3D clustering analysis of super-resolution microscopy data by 3D Voronoi tessellations. *bioRxiv* 146456. *Sous revue dans Bioinformatics.*

2. Orlov, I., Myasnikov, A.G., Andronov, L., Natchiar, S.K., Khatter, H., Beinsteiner, B., Ménétret, J.-F., Hazemann, I., Mohideen, K., Tazibt, K., , Tabaroni, R., Kratzat, H., Djabeur, N., Bruxelles, T., Raivoniaina, F., Pompeo, L., Torchy, M., Billas, I., Urzhumtsev, A., Klaholz, BP. (2017). The integrative role of cryo electron microscopy in molecular and cellular structural biology. *Biol. Cell* **109**, 81–93.

3. Andronov, L., Orlov, I., Lutz, Y., Vonesch, J.-L., and Klaholz, B.P. (2016b). ClusterViSu, a method for clustering of protein complexes by Voronoi tessellation in super-resolution microscopy. *Sci. Rep.* **6**, 24084.

4. Andronov, L., Lutz, Y., Vonesch, J.-L., and Klaholz, B.P. (2016a). SharpViSu: integrated analysis and segmentation of super-resolution microscopy data. *Bioinformatics* **32**, 2239–2241.

5. Lemaître, C., Grabarz, A., Tsouroula, K., Andronov, L., Furst, A., Pankotai, T., Heyer, V., Rogier, M., Attwood, K.M., Kessler, P., Dellaire, G., Klaholz, B., Reina-San-Martin, B., Soutoglou, E. (2014). Nuclear position dictates DNA repair pathway choice. *Genes Dev*. **28**, 2450–2463.

Communications

1. Leonid Andronov. 2D and 3D segmentation of single-molecule localization microscopy data using Voronoi diagrams. *Imagerie quantitative: Journées thématiques super-résolution.* Le 13 novembre 2017. ESPCI Paris, France. *Communication orale.*

2. Leonid Andronov, Jonathan Michalon, Khalid Ouararhni, Igor Orlov, Yves Lutz, Jean-Luc Vonesch, Bruno Klaholz. Segmentation and cluster analysis of protein complexes by Voronoi tessellation in 3D localization microscopy. *Focus on Microscopy-2017*. Les 09–12 avril 2017, Bordeaux, France. *Communication par affiche.*

3. Leonid Andronov. Visualization and segmentation of protein complexes in super-resolution microscopy. *Mini-symposium: Novel bio-imaging tools for understanding Life.* Le 27 janvier 2017. Faculté de médecine, Strasbourg, France. *Communication orale.*

4. Leonid Andronov. Localizing protein complexes with super-resolution microscopy. *Symposium on Protein-Protein Interactions and Complexes.* Le 15 novembre 2016, IGBMC, Illkirch, France. *Communication orale.*

5. Leonid Andronov. Developments in super-resolution microscopy: the GSDIM setup and processing with the SharpViSu software. *IGBMC Internal seminar.* Le 15 septembre 2016. *Communication orale.*

6. Leonid Andronov, Igor Orlov, Yves Lutz, Jean-Luc Vonesch, Bruno Klaholz. ClusterViSu, a method for clustering and colocalization analysis of protein complexes by Voronoi tessellation in single-molecule microscopy. *ICON Europe 2016*. Les 07–10 juin 2016, Bâle, Suisse. *Communication par affiche.*

7. Leonid Andronov, Yves Lutz, Jean-Luc Vonesch, Bruno Klaholz. SharpGSDIM, a software for processing of super-resolution microscopy data. *Focus on Microscopy-2015*. Les 29 mars – 1 avril 2015, Göttingen, Allemagne. *Communication par affiche.*

8. Leonid Andronov. Super-resolution GSDIM microscopy. *Imagerie au Campus d'Illkirch.* Le 13 décembre 2013, IGBMC, France. *Communication orale.*

Table of contents

Acknowledgments 3
Summary 4
Résumé de thèse 8
1. Introduction 25
1.1. Fluorescence microscopy 27
1.2. Super-resolution microscopy 32
1.2.1. Total internal reflection fluorescence microscopy 33
1.2.2. Structured illumination microscopy 35
1.2.3. Stimulated emission depletion microscopy 40
1.2.4. Single-molecule localization microscopy 44
1.2.4.1. Methods for three-dimensional SMLM 51
PSF modifications: astigmatism 51
PSF modifications: double-helix PSF 54
Detection path splitting 55
4Pi microscopy 56
1.2.5. Comparison of super-resolution techniques 58
1.3. SMLM: data processing 60
1.3.1. Single-molecule localization 60
1.3.2. Fitting of highly dense fluorophores 61
1.3.3. Effects of molecular orientation on localization accuracy 62
1.3.4. Size of labeling system and localization accuracy 64
1.3.5. Drift correction 65
1.3.6. Correction of chromatic aberrations 66
1.3.7. Super-resolution image reconstruction 68
1.3.8. Estimation of real resolution 69
1.3.8.1. Fourier ring correlation 70
1.3.9. Cluster analysis of SMLM data 72
1.3.9.1. Nearest-neighbor analysis (Clark-Evans statistics) 72
1.3.9.2. Ripley's K and L functions 73
1.3.9.3. Pair correlation functions 77
1.3.9.4. DBSCAN 78
1.3.10. Colocalization analysis 78
1.3.10.1. Image-based analysis 79
1.3.10.2. Coordinate-based analysis 80

1.4. Chromatin, histones and centromeres 82
1.4.1. Super-resolution imaging of chromatin 88
2. Objectives of the research project 90
3. Results and discussion 93
3.1. SharpViSu, a software for SMLM data processing 94
3.1.1. Publication 1 “SharpViSu: integrated analysis and segmentation of super-resolution microscopy data” 95
3.2. ClusterViSu, a method for 2D cluster analysis of SMLM data 96
3.2.1. Publication 2 “ClusterViSu, a method for clustering of protein complexes by Voronoi tessellation in super-resolution microscopy” 97
3.3. 3DClusterViSu, a method for 3D cluster analysis of SMLM data 98
3.3.1. Publication 3 “3DClusterViSu: 3D clustering analysis of super-resolution microscopy data by 3D Voronoi tessellations” 99
3.4. SR imaging of CENP-A throughout the cell cycle 100
3.4.1. Methods 108
4. Conclusions and perspectives 110
4.1. General discussion 111
4.2. Perspectives 115
5. References 117
6. Appendix 140
6.1. Publication 4 “Nuclear position dictates DNA repair pathway choice” 141
6.2. Review “The integrative role of cryo electron microscopy in molecular and cellular structural biology” 142
6.3. Posters 143

List of abbreviations

AU	Airy unit
CCAN	Constitutive centromere-associated network
CenH3	Centromeric H3 variant
CENP-A	Centromere protein A
COPI	Coat protein
DBSCAN	Density-based spatial clustering of applications with noise
dSTORM	Direct stochastic optical reconstruction microscopy
EGFP	Enhanced green fluorescent protein
EM	Electron microscopy; Electron multiplication
EMCCD	Electron-multiplying charge-coupled device
ET	Electron tomography
EYFP	Enhanced yellow fluorescent protein
Fab	Antigen-binding fragment
FCS	Fluorescence correlation spectroscopy
FIB	Focused ion beam
FITC	Fluorescein isothiocyanate
FLIM	Fluorescence lifetime microscopy
FOV	Field of view
FP	Fluorescent protein
FPALM	Fluorescence photoactivation localization microscopy
FRC	Fourier ring correlation
FWHM	Full width at half maximum
GFP	Green fluorescent protein
GSD	Ground state depletion
GSDIM	Ground state depletion followed by individual molecule return
H3K9me3	Histone H3 modified by trimethylation of lysine 9
HJURP	Holliday junction recognition protein
HOH	Higher-order harmonics
IF	Immunofluorescence
iPALM	Interferometric photoactivated localization microscopy
LP	Localization precision
LS	Least-squares

MEA	β-mercaptoethylamine
MLE	Maximum-likelihood estimation
MT	Microtubule
NA	Numerical aperture
NL	Nonlinear
NMR	Nuclear magnetic resonance spectroscopy
PA	Photoactivated
PALM	Photoactivated localization microscopy
PALMIRA	Photoactivated localization microscopy with independently running acquisition
PSF	Point spread function
RESOLFT	Reversible saturable optical linear fluorescence transitions
ROI	Region of interest
SAXS	Small-angle X-ray scattering
SANS	Small-angle neutron scattering
SD	Standard deviation
SEM	Scanning electron microscopy
S_i	i-th singlet state
SIM	Structured illumination microscopy
SM	Single-molecule
SMLM	Single-molecule localization microscopy
SNR	Signal-to-noise ratio
SPT	Single particle tracking
SR	Super-resolution
SSIM	Saturated structural illumination microscopy
STED	Stimulated emission depletion
STORM	Stochastic optical reconstruction microscopy
T_i	i-th triplet state
TIRF	Total internal reflection microscopy
uPAINT	Universal points accumulation for imaging in nanoscale topography
VH	Variable domain of the immunoglobulin heavy chain
VL	Variable domain of the immunoglobulin light chain
VVSRM	Virtual volume super-resolution microscopy
W-4PiSMSN	Whole-cell 4Pi single-molecule switching nanoscopy

List of figures

Figure 1. Resolution of the optical system. 26
Figure 2. Jablonski diagram of a dye molecule. 28
Figure 3. Scheme of the epifluorescence microscope. 29
Figure 4. Principle of confocal microscopy. 30
Figure 5. Comparison of confocal with wide-field fluorescence images. 31
Figure 6. Scale of things studied by integrated structural biology. 33
Figure 7. Principle of total internal reflection fluorescence microscopy. 34
Figure 8. Comparison of TIRF with conventional epifluorescence imaging. 35
Figure 9. Principle of structured illumination microscopy. 36
Figure 10. Nonlinear structured illumination microscopy. 39
Figure 11. Stimulated emission depletion microscopy. 41
Figure 12. Principle of single-molecule localization microscopy. 47
Figure 13. Some of the first images obtained with SMLM. 49
Figure 14. Jablonski diagram of a typical fluorophore suitable for GSDIM microscopy. 50
Figure 15. GSDIM imaging of β-tubulin in HeLa cells labeled with Alexa 647-conjugated secondary antibodies. 51
Figure 16. 3D SMLM based on astigmatism. 52
Figure 17. 3D SMLM of β-tubulin labeled with Alexa 647-conjugated secondary antibodies in a HeLa cell. 53
Figure 18. 3D SMLM with engineered double-helix PSF. 54
Figure 19. Scheme of the 3D SMLM with dual-focal-plane imaging (detection path). 55
Figure 20. Multiphase interferometric microscope. 57
Figure 21. Approximation of the lateral profile of the PSF (the Airy Pattern) with the Gaussian function. 61
Figure 22. Comparison of the performance of sparse localization algorithms with a high-density fitting algorithm DAOSTORM. 62
Figure 23. Imaging of individual dipoles with different orientations. 63
Figure 24. Lateral chromatic aberrations in SMLM. 67
Figure 25. Different methods for SMLM data visualization. 69
Figure 26. FRC curve for the dataset shown in **Fig. 15**. 71
Figure 27. Example of Ripley's statistics for a SMLM dataset. 75
Figure 28. Detailed cluster analysis using the Ripley's statistics. 76

Figure 29. Coordinate-based colocalization analysis of SMLM data. 81
Figure 30. Schematic of the eukaryotic cell cycle..... 83
Figure 31. Nucleosome in complex with linker histone H1..... 84
Figure 32. Structure of the DNA entrance and exit of the human CENP-A nucleosome..... 85
Figure 33. Function of CENP-A..... 87
Figure 34. Model of centromeric chromatin organization based on SMLM imaging of unfolded centromeres. 89
Figure 35. CENH3 chromatin is organized in ring structures throughout the cell cycle of rye cells..... 89
Figure 36. SMLM imaging of CENP-A in one single HeLa cell..... 100
Figure 37. SMLM images of centromeres in different non-synchronized HeLa cells..... 101
Figure 38. SMLM imaging of centromeres in synchronized U2OS cells. 102
Figure 39. 3D reconstructions of the centromere at 1.5 hours after mitosis based on the common line approach. 103
Figure 40. 3D SMLM imaging of CENP-A in a U2OS cell fixed at 1.5 hours after mitosis. 104
Figure 41. Statistical analysis of CENP-A clusters in fixed and synchronized U2OS cells. 105
Figure 42. SMLM imaging of CENP-A with its chaperone HJURP. 106
Figure 43. STED imaging of CENP-A and HJURP proteins..... 107

List of tables

Table 1. Comparison of commonly used super-resolution techniques: SMLM, STED and SIM..... 59

1. Introduction

The ability to see small objects is indispensable for proper understanding of biological systems. This is why many discoveries in biology are tightly bound with developments of microscopy technologies. The invention of the first microscopes in the 17th century made it possible to observe previously unseen microscopic structure of life, which became the starting point of modern biology. Improvement in optics made possible observations of increasingly smaller objects; however later it became evident that there is a fundamental limit to resolution, which cannot be avoided by any improvement in optics. This limit, also known as the diffraction limit, was described in 1873 by Ernst Abbe (Abbe, 1873):

$$d \approx \frac{0.61\lambda}{n sin\Theta} \approx \frac{\lambda}{2NA} \tag{1}$$

where d is the radius of the first minimum of the diffraction spot from a point source, λ is the wavelength of light, n is the refractive index of the medium between the objective and the object, θ is the half-angle of light collection, NA = n·sinθ is the numerical aperture of the objective. NA is limited to ≲ 1.5 by the refractive index of immersion oil and glass. Thus, whatever the size of an object, its image in a transmission microscope cannot reveal details smaller than the order of the wavelength of the observation light.

The response of an imaging system to a point light source is called the point spread function (PSF). For an ideal optical system this function is known as the Airy pattern (Airy, 1835), it can be calculated from light diffraction on a circular aperture (**Fig. 1A**). According to the Rayleigh resolution criterion, two points can be resolved when the first diffraction minimum of the image of the first source coincides with the maximum of another (Lord Rayleigh, 1879), that is why the formula (1) is the resolution limit for microscopy (**Fig. 1B-C**).

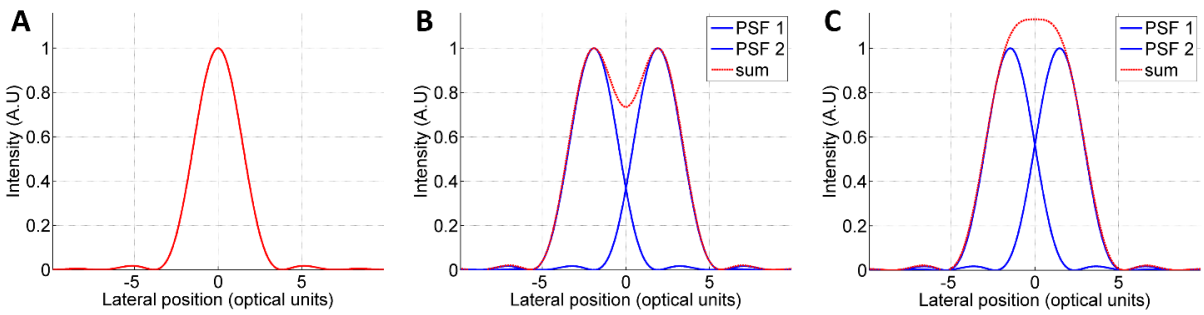


Figure 1. Resolution of the optical system. (**A**) Lateral profile of the PSF of an ideal optical system. (**B**) Images of two point sources are resolved according to the Rayleigh criterion: the two main diffraction maxima coincide with the first diffraction minima. (**C**) Two point sources separated with a slightly smaller distance are not resolved.

1.1. Fluorescence microscopy

Photoluminescence is the property of atoms and molecules to emit light following excitation by an outside source of photons. Fluorescence is a particular case of luminescence, resulting from singlet–singlet electronic relaxation, with a lifetime in the order of nanoseconds. This phenomenon originates from the electron properties of the matter, which can be represented as a Jablonski diagram for the case of molecules (Jabłoński, 1933) (**Fig. 2**). As a result of these properties, the absorption and emission spectra are usually separated from each other and the emission usually occurs at longer wavelengths (Stokes shift (Stokes, 1852)). Since the spectra reflect the chemical structure of dyes, they are also different for different chemical compounds. Thanks to these, fluorescent molecules started to be used as markers for microscopic observations of different biological components, such as proteins, membranes, organelles, etc. The technique which allows to image fluorescently labeled objects is called fluorescence microscopy (Brumberg and Krylova, 1953; Ellinger and Hirt, 1929; Ploem, 1967) (**Fig. 3**).

Thanks to the properties of fluorescence, it became possible to image, *e.g.* different proteins in the same cell or tissue with high contrast and specificity. Even though many biological objects are fluorophores by themselves, i.e. they demonstrate so-called autofluorescence, this signal in most cases is not specific nor strong enough to be used in practice. This is why molecules of interest are usually need to be labeled with a high-quality fluorophore. There are several ways to label a molecule or a structure, among the most used methods are DNA- and membrane-binding fluorescent agents (*e.g.* DAPI (Kapuscinski, 1995), DiI (Shiraishi et al., 1992)), fluorescently labeled phalloidin (Shiraishi et al., 1992), fluorescently labeled antibodies and genetically encoded systems, such as fluorescent proteins (FPs, *e.g.* Green Fluorescent Protein-GFP (Shimomura et al., 1962), mCherry (Shaner et al., 2004), etc.).

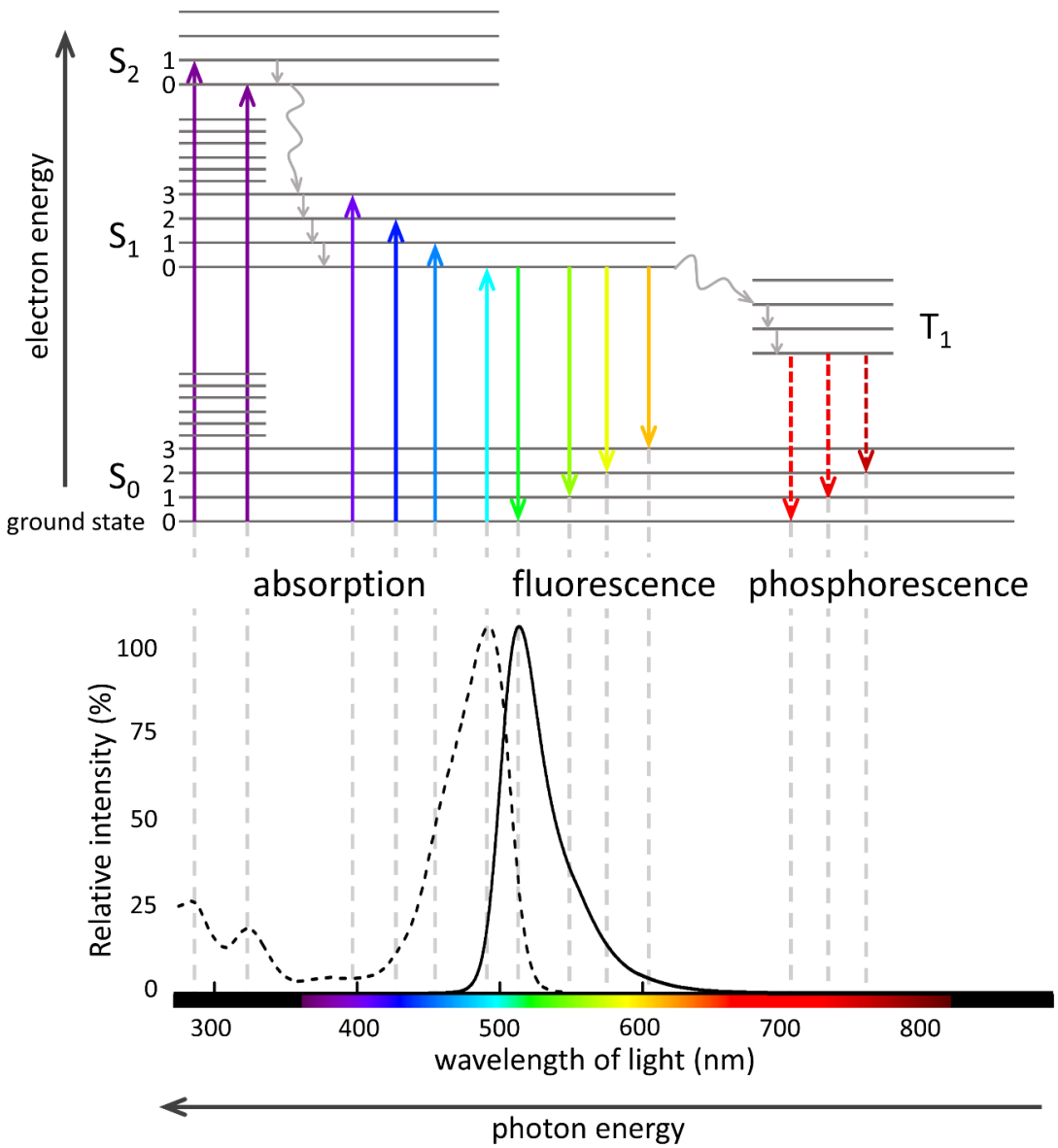


Figure 2. Jablonski diagram of a dye molecule. S_0-S_2 are the singlet electron states. At the room temperature and in the absence of excitation almost all molecules are in the ground state. A photon can be absorbed with some probability, bringing an electron to one of the excited singlet states S_1, S_2 or S_n. To relax back to the ground state, the electron can follow different paths, from which the transitions from S_1 to S_0 can result in an emission of a photon (fluorescence). The lifetime of the S_1 state is in the order of 10^{-7} to 10^{-9} s. T_1 is a triplet state (typical lifetime: 10^{-4} to 10 s). The probability of transition from a singlet to a triplet state (intersystem crossing) is low compared to a de-excitation through the singlet states, that is why the T_1 state is normally not populated and phosphorescence (radiative T_1-S_0 transitions) is very weak. The spectra of the Fluorescein dye is represented in the bottom: excitation in dashed and fluorescence in solid line (author: Zadelrob, attributed under a CC BY-SA 3.0 license https://en.wikipedia.org/wiki/File:Fluorescein-spectra3.svg.).

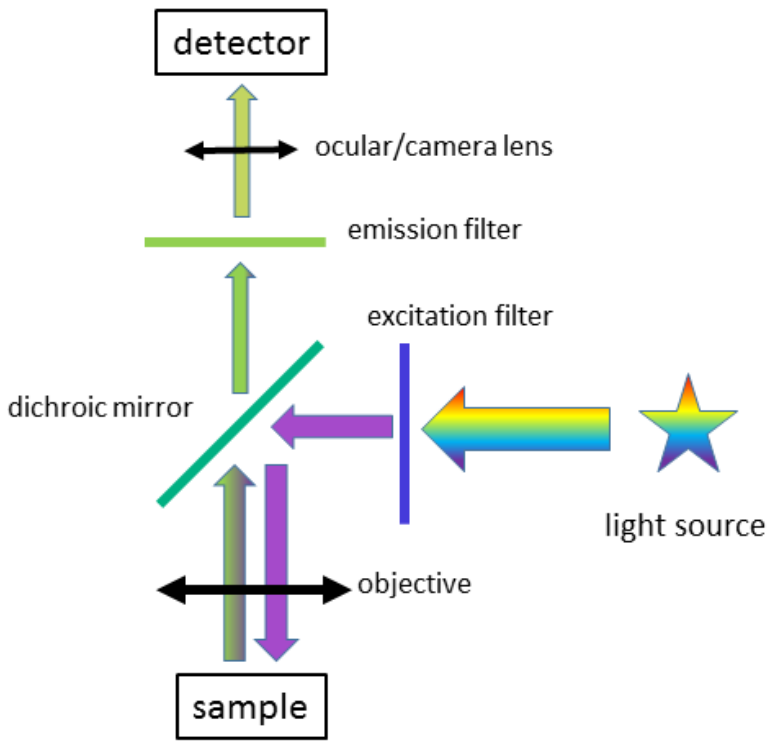


Figure 3. Scheme of the epifluorescence microscope. The excitation light is selected by an excitation filter from a broad-spectrum light source (*e.g.*, a gas discharge lamp). A dichroic mirror reflects the excitation light to the sample. After passing through the objective, this light reaches the specimen with a fluorescent labeling. The fluorescence, together with scattered and reflected light, is collected back by the same objective. The dichroic mirror reflects excitation light but passes the longer-wavelength fluorescence (Brumberg and Krylova, 1953; Ploem, 1967). Emission filter additionally cuts out any light outside the fluorescence spectrum. The image of the fluorescently labeled object can then be seen in the ocular or registered with a camera. The excitation filter can be omitted if a single-mode laser or a light-emitting diode is used as a light source.

A major breakthrough in microscopy occurred with the invention and commercialization of the confocal laser scanning microscope (Davidovits and Egger, 1971; Minsky, 1961; Sheppard and Choudhury, 1977). While in the conventional wide-field fluorescence microscope the image contains out-of-focus light from the whole depth of the sample, which strongly limits image contrast and thus observation of thick specimens, in the confocal microscope most of the out-of-focus light is cut off by a pinhole (**Fig. 4**) that allows effective optical sectioning (**Fig. 5**) (White et al., 1987). The point scanning system of a confocal microscope also allows a very flexible spectral separation of dyes, because excitation and emission filters can be replaced with

monochromators. In order to image different fluorophores at the same time, several detectors can be used simultaneously.

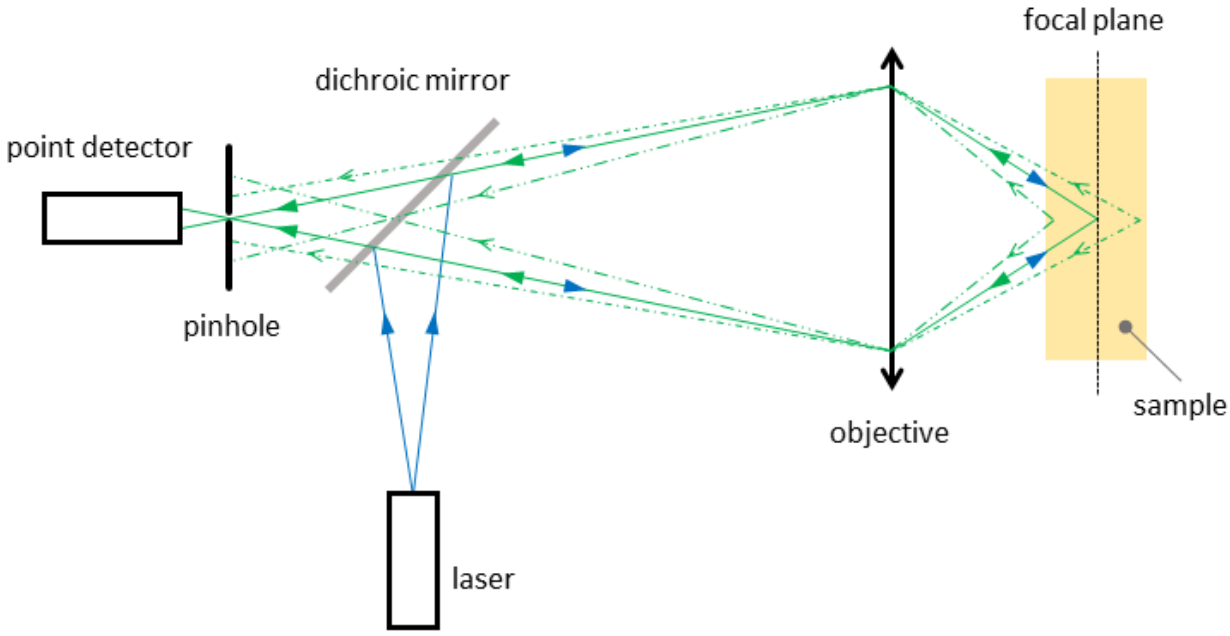


Figure 4. Principle of confocal microscopy. Excitation light after reflection from a dichroic mirror is focused by the objective into a point in the focal plane inside the sample (blue arrows). Fluorescence from this point of the sample (green arrows) is collected back by the objective, passes through the dichroic mirror and through a diaphragm in the conjugated focal plane (so-called pinhole) and is detected by a point light detector (*e.g.* by a photomultiplier or by an avalanche photodiode). Fluorescence from any out-of-focus regions of the sample is cut off by the pinhole (green dashed lines).

Confocal imaging can also improve resolution of fluorescence microscopy roughly 1.4 times in both lateral and axial directions as compared to the wide-field fluorescence microscope (Wilson, 2011). The lateral and the axial full widths at half minima of wide-field and confocal microscopes can be calculated accordingly to:

$$LFWHM_{wf} = 0.51\frac{\lambda}{NA} \qquad LFWHM_{conf} = 0.37\frac{\lambda}{NA} \tag{2}$$

$$AFWHM_{wf} = 0.89\frac{\lambda}{n-\sqrt{n^2-NA^2}} \qquad AFWHM_{conf} = 0.64\frac{\lambda}{n-\sqrt{n^2-NA^2}} \tag{3}$$

where n is the refractive index of the immersion medium.

The width of the optical section is limited by diffraction for small pinhole sizes and grows linearly with the increasing the pinhole size (Wilson, 2011):

$$FWHM_{conf} \approx 0.67 \frac{\lambda}{n-\sqrt{n^2-NA^2}} \sqrt{1+AU^2} \quad (4)$$

The formulas (2) – (3) are valid for a small pinhole diameter which passes only a minor part of the fluorescent light. To have a satisfying signal-to-noise ratio (SNR), in practice pinhole diameters of at least 1 airy units (1 AU) are used (the size of the main peak of the Airy pattern), which brings resolution closely to the wide-field case while retaining good optical sectioning (4) and high intensity (Wilson, 2011).

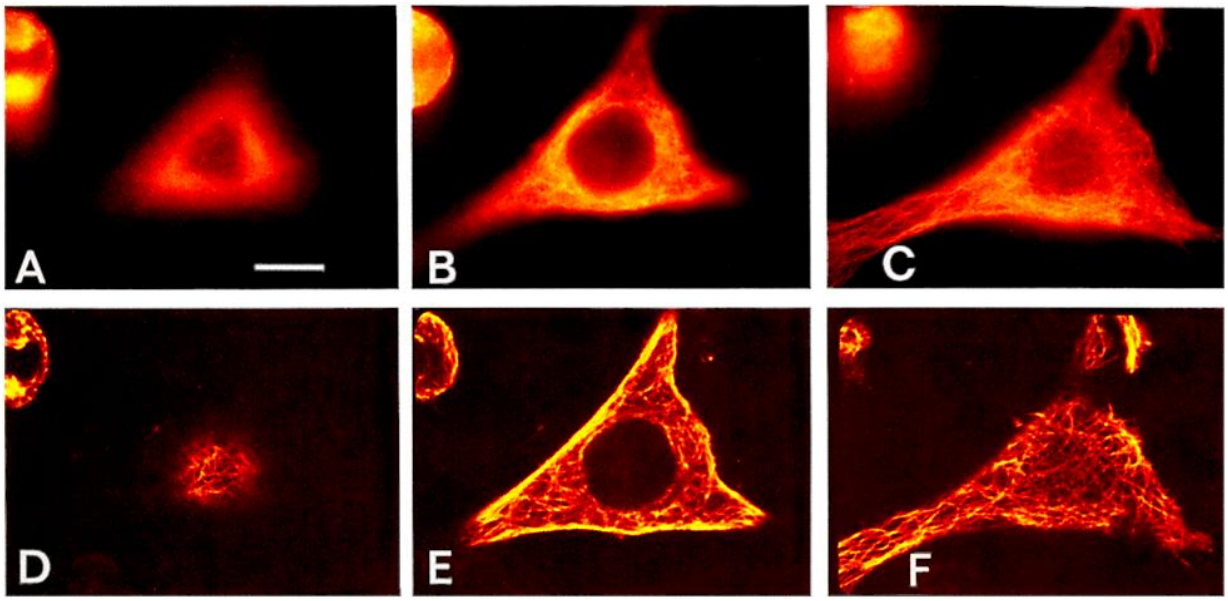


Figure 5. Comparison of confocal with wide-field fluorescence images. Example of tubulin labeled with fluorescein isothiocyanate (FITC)-conjugated secondary antibodies in a HeLa cell. Confocal images (**D-F**) as compared to wide-field images (**A-C**) allow to see much more details in the same region due to removal of the out-of-focus fluorescence. Scale bar is 10 µm. Image is adapted from (White et al., 1987).

A recent improvement to the confocal microscope was the introduction of Airyscan detectors by Zeiss (Huff, 2015). Instead of a pinhole and a point detector it uses a 32 component 2D detector on which the conjugated focal plane is projected. Each detector element acts as a small 0.2 AU pinhole, while the whole array has the diameter of 1.25 AU. The signals of all the units are reassigned in order to get an image with high SNR and resolution.

1.2. Super-resolution microscopy

According to the diffraction limit (1), it is impossible to obtain an image in a classical way with resolution significantly better than the wavelength of light. For the visible light (λ = 400…700 nm) this limit can be around 150 nm at best, which is obviously not enough for precise studies on molecular structures. However, recently it turned out that this limit can be overcome in fluorescence microscopy using special imaging schemes or special properties of fluorophores. Such techniques, named "super-resolution microscopy", include a number of different methods, among them the most popular nowadays are: total internal reflection microscopy, structured illumination microscopy, stimulated emission depletion microscopy and single-molecule localization microscopy. These techniques will be discussed in detail below.

Many biological structures, big protein complexes or viruses, have a size below the resolution limit (**Fig. 6**), so resolution improvement of highly contrast, specific and multi-color fluorescence imaging down to this scale was highly beneficial for research in biology. This is why the invention and development of super-resolution fluorescence microscopy techniques which happened in practice at the beginning of the 21st century was quickly marked by a Nobel Prize in Chemistry (2014). Super-resolution microscopy has rapidly become a technique which bridges the resolution gap between traditional light microscopy and highly resolutive techniques, such as X-ray crystallography and cryo-electron microscopy (**Fig. 6**), and thus allows to link atomic structures of proteins with their functions in the cellular context and eventually in the organism.

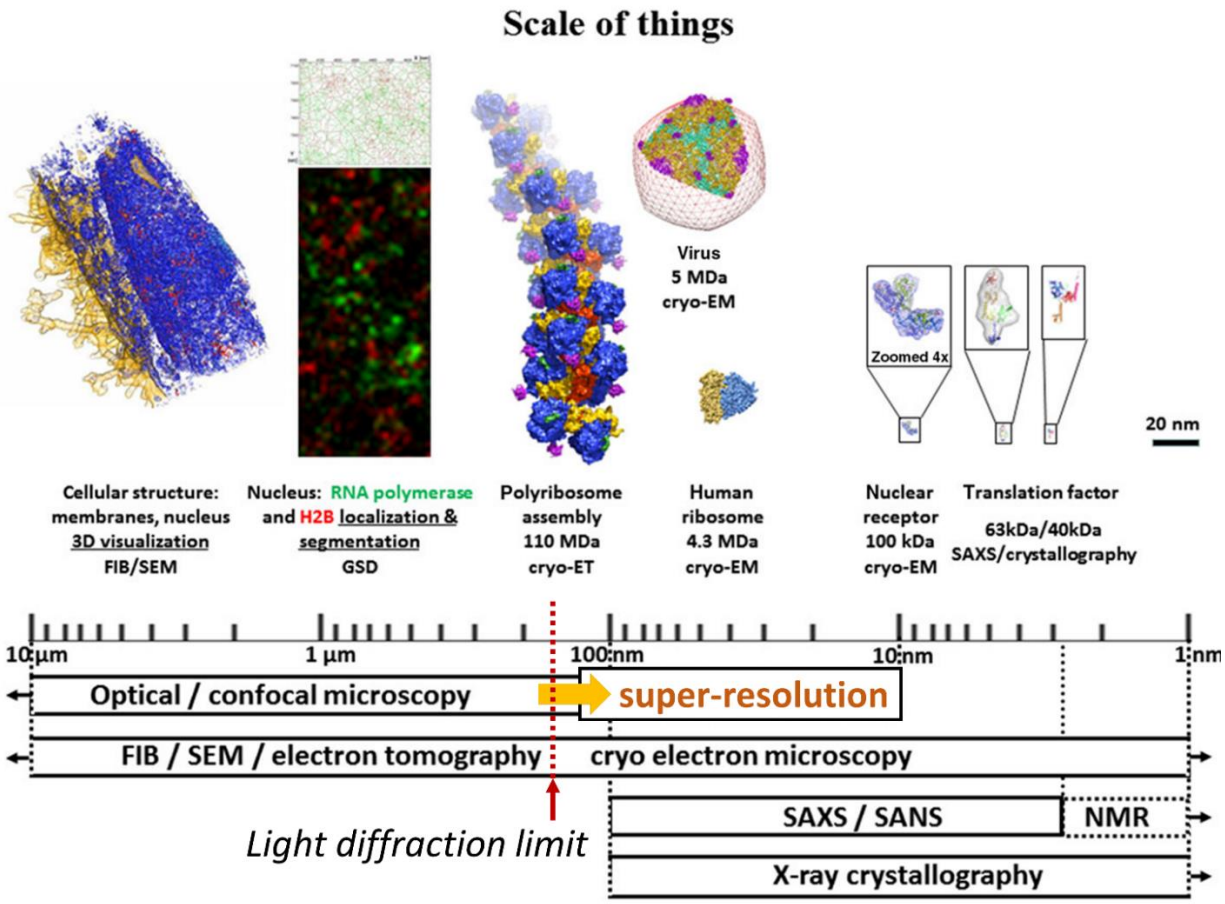


Figure 6. Scale of things studied by integrated structural biology. Image is adapted from (Orlov et al., 2017) and demonstrates examples of macromolecular complexes studied in our lab.

1.2.1. Total internal reflection fluorescence microscopy

When an electromagnetic wave strikes a boundary between two media where the refractive index of the first medium is greater than that of the second one, at an angle larger than a certain critical angle, the wave cannot pass through the boundary and completely reflects back to the first medium. This phenomenon is known as total internal reflection. Even though the wave is completely reflected, it penetrates to the second medium to a distance in the order of the wavelength in the form of an evanescent wave, which has an exponential decay:

$$I(z) = I_0 \cdot \exp(-z/d), \tag{5}$$

where I is the intensity of the evanescent wave, I_0 is the intensity of the incident wave, z is the perpendicular distance from the boundary, d is the characteristic depth of decay (the depth at which the excitation intensity decreases *e* times):

$$d = \frac{\lambda}{4\pi n_2}\left(\frac{sin^2\theta}{sin^2\theta_c} - 1\right)^{-\frac{1}{2}}, \quad (6)$$

where λ is the wavelength of light, n_2 is the refractive index of the second medium, θ_c = arcsin (n_2/n_1) is the critical angle and θ is the angle of incidence ($\theta > \theta_c$).

The idea to use evanescent waves for transmission light microscopy appeared back in 1956 (Ambrose, 1956) and was developed later for fluorescence microscopy (Axelrod, 1981). If a fluorescent molecule is found in a close proximity to the boundary, it will be excited. At the same time, any molecule further than the characteristic depth (6) will get almost no light. A very thin excitation plane improves axial (along Z axis) resolution and removes background signal. In practice the thickness of the excited region can be as small as 50–100 nm (**Fig. 7–8**) in total internal reflection (TIRF) microscopy.

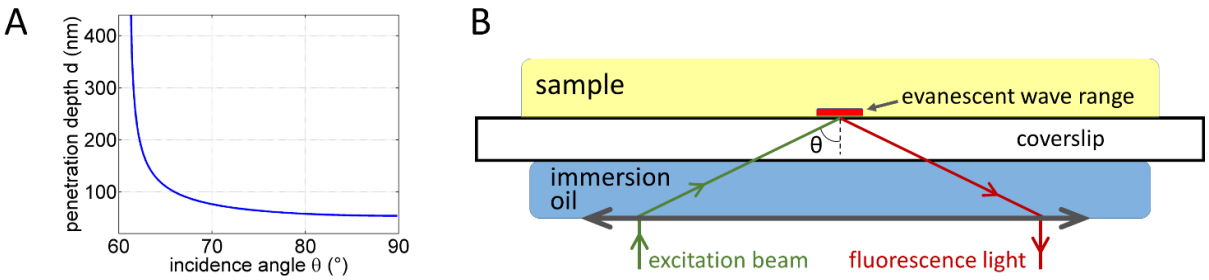


Figure 7. Principle of total internal reflection fluorescence microscopy. (**A**) Penetration depth *d* as the function of incidence angle *θ*. The curve is calculated using formula (5) with the following parameters: $\lambda = 500\ nm$ (green light); $n_1 = 1.52$ (glass, immersion oil); $n_2 = 1.33$ (water). (**B**) Geometry of TIRF microscopy with through the objective excitation. Since the angle θ should be larger than the critical angle, only high-NA oil immersion objectives are suitable for this geometry. In other geometries the sample can be illuminated *e.g.* through a prism and light can be collected also with a lower NA objective (Ambrose et al., 1999).

TIRF microscopy became widely adopted for imaging of cell-substrate contacts (Axelrod, 1981; Burmeister et al., 1998), motor proteins (Vale et al., 1996), endo- and exocytosis (Oheim et al., 1998), adhesions (Choi et al., 2008) and cytoskeleton (Manneville et al., 2003; Vale et al., 1996). Since single-molecule (SM) imaging needs very low background intensity and confocal scanning techniques are not practical for detection of individual molecules, TIRF microscopy became indispensable for such experiments (Horn et al., 2016; Vale et al., 1996). The evident shortcomings of TIRF microscopy are illumination of only one pre-surface layer and improvement of resolution only in the axial direction.

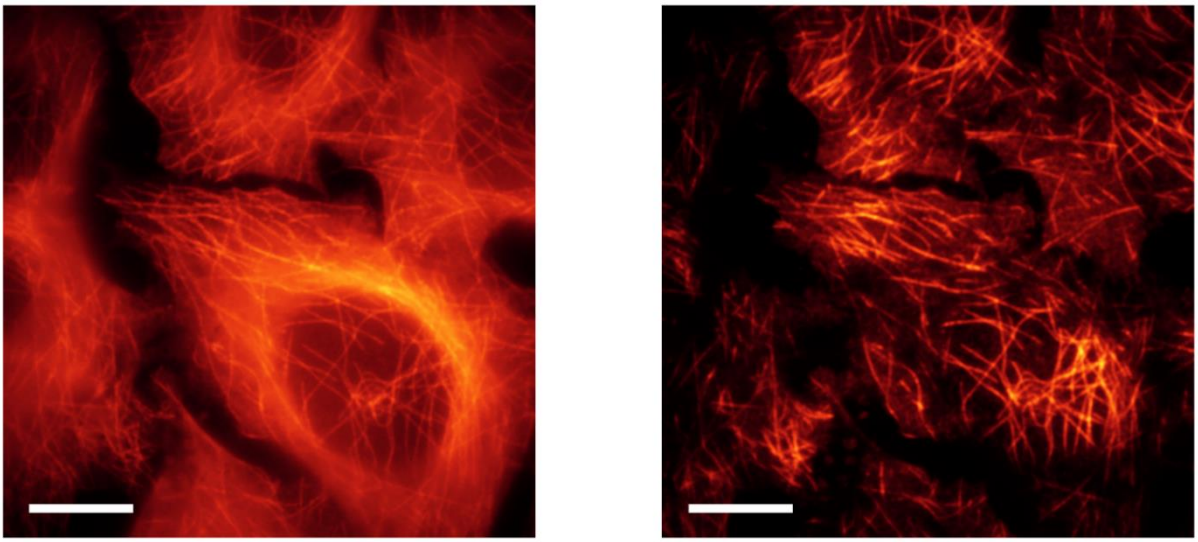


Figure 8. Comparison of TIRF with conventional epifluorescence imaging. β-tubulin detected with Alexa 647 conjugated secondary antibodies imaged with conventional epifluorescence illumination (left) and with through-the-objective TIRF (right). Note the absence of background fluorescence and the presence of a stronger signal from microtubules close to the coverslip in the TIRF image. Scale bars, 10 µm. The images were acquired on the in-house Leica SR GSD system.

1.2.2. Structured illumination microscopy

The maximum spatial frequency that can be detected through a microscope is limited by diffraction (1): $k_0=2NA/\lambda$, which can be regarded as a circle of the radius k_0 in the 2D Fourier space (**Fig. 9**). However, higher frequencies of the sample can also be accessed through a microscope if one reduces their apparent frequency using patterned illumination. Indeed, two overlaid periodic patterns produce an interference pattern which can have lower frequencies than present in the original patterns. This phenomenon is known as Moiré fringes and can produce undesirable effects, *e.g.* in television, when a periodic pattern of an object interferes with the periodic lattice of the digital image detector. As a result, low-frequency fringes become apparent even though the high-frequency pattern of the object is not resolved (**Fig. 9A**).

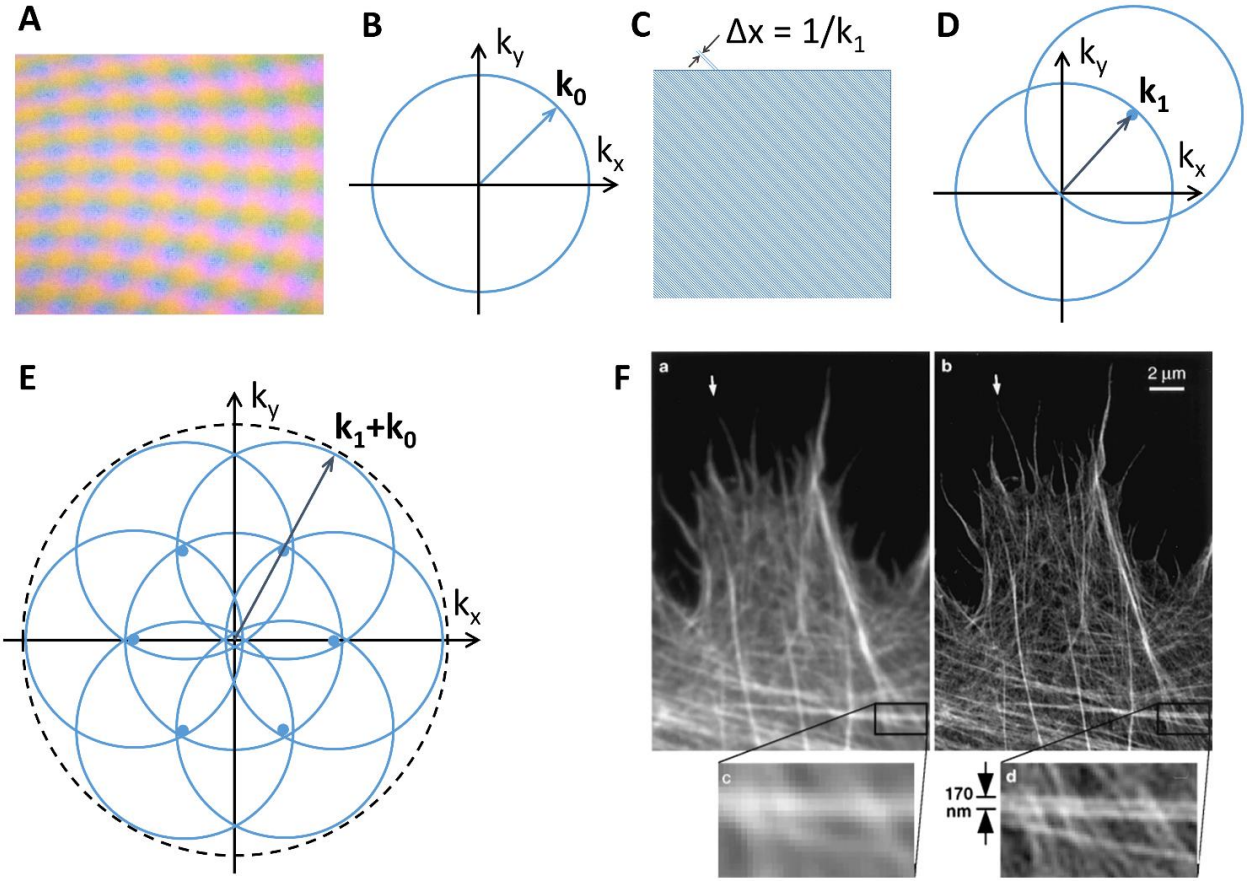


Figure 9. Principle of structured illumination microscopy. (**A**) Moiré pattern observed when a computed display is imaged with a digital camera: even though the underlying pixels are too small to be seen, the Moiré fringes reflect the periodic structure of the display and the camera. The observed pattern has much lower spatial frequency than the underlying lattices of pixels. (**B**) The region observable through a conventional wide-field microscope corresponds to a circle of the radius $k_0 = 2NA/\lambda$ in Fourier space. (**C**) Illumination pattern with period Δx for SIM. (**D**) Such an illumination allows to observe higher frequencies inside the circle with the center in $\mathbf{k_1}$. (**E**) Patterns with different phases and angles allow to fill the whole Fourier space extending the resolution to $|\mathbf{k_1} + \mathbf{k_0}| = 4NA/\lambda$. (**F**) Images of the actin cytoskeleton in a HeLa cell taken with a conventional (left) and a SIM (right) microscope (Gustafsson, 2000).

Since the observed emission intensity of fluorophores is the product of the density of the dyes with the intensity of the excitation light, if the sample is illuminated with a patterned illumination with spatial frequency $\mathbf{k_1}$ (**Fig. 9C**), the resulting moiré pattern will contain the difference frequency $\mathbf{k} - \mathbf{k_1}$, where $\mathbf{k}$ is a frequency present in the sample. The Moiré fringes will be observable if $|\mathbf{k} - \mathbf{k_1}| < k_0$ that can be seen as a shift of the observable region to a higher frequency range, up to the frequency $k_1 + k_0$ (**Fig. 9D-F**). The k_1 frequency is, however, limited

by the same diffraction limit to $k_{1,max} = k_0 = 2NA/\lambda$. The maximum observable frequency with this method is therefore $k_1 + k_0 = 4NA/\lambda$, *i.e.* the resolution is only improved twice in the best case (Gustafsson, 2000, 2005).

Another application of structured illumination is the optical sectioning. When the sample is illuminated with a fine pattern, the shape of the pattern will affect only the in-focus part of the specimen. The out of focus blur is similar for any position of the pattern, that is why it can be removed by subtracting images acquired with different positions of the illumination grid (Neil et al., 1997). The resulting strength of the optical sectioning is comparable with that of confocal microscopy (Karadaglić and Wilson, 2008). In order to extend structured illumination microscopy (SIM) for an axial resolution improvement, a 3D structural illumination scheme was developed, which – besides a lateral periodic pattern – has also an axial one, thus achieving a twofold gain in resolution in all three dimensions (Gustafsson et al., 2008).

SIM allows to image whole cells in multiple colors using conventional dyes with lateral resolution approaching 100 nm (Schermelleh et al., 2008). Despite the fact that SIM requires at least 9 raw images for reconstitution of one 2D image with extended resolution, it is a wide-field technique and its speed is still quite high as compared with other super-resolution techniques or even with confocal scanning microscopy. Moreover, unlike other techniques, it does not require particularly powerful excitation which helps avoiding bleaching and phototoxicily. This is why besides only a twofold improvement in resolution SIM is becoming more popular for live cell imaging. Recently, endocytic and cytoskeletal dynamics were imaged at 84 nm lateral resolution with SIM using a 1.7-NA objective and TIRF illumination. The authors have been able to depict multiple proteins with a speed below 1s per color per frame, which gave insights into *e.g.* clathrin-mediated endocytosis, dynamics of focal adhesion and interactions of filamentous actin with myosin (Li et al., 2015).

The considerations about the twofold improvement of resolution are valid in case of linear response of fluorophores on excitation (*i.e.*, the intensity of fluorescence is directly proportional to the intensity of illumination). If the response in not linear, the shape of the emission pattern will no more be proportional to the illumination pattern, giving rise to harmonics on multiple of k_1 frequencies. This is equivalent to the extension of the observable region in Fourier space to nk_1+k_0, where n is the number of the maximal observed harmonic. Since the observed image is the sum of all components with different harmonics, in order to retrieve the whole image, at least n measurements with different phase shifts are needed. Also, more grid orientations are needed for complete filling of the large Fourier space (Gustafsson, 2005). The SIM techniques

that use a nonlinear phenomenon to achieve high resolution are called nonlinear SIM (NL-SIM) or saturated SIM (SSIM).

A particular nonlinear process first proposed for super-resolution SIM is the saturation of the excited S_1 state (Gustafsson, 2005; Heintzmann et al., 2002). Fluorescent molecules are transferred to the excited state and stay there for some average time, so-called lifetime (**Fig. 2**). During this time they are not accessible for excitation, meaning that the absorption decreases with increasing excitation intensity. If the sample is illuminated with a high power sinusoidal pattern, the response of the fluorophores will be almost linear near the valleys of the pattern and the saturation will occur near the peaks. As a result, the emission pattern will have a deformed shape with flattened peaks. This will lead to the appearance of higher harmonics in the Fourier space, and their strength will be higher with increasing excitation intensity. This effect can produce theoretically unlimited resolution, which is limited in practice by SNR and by the maximum power that the sample can withstand (Gustafsson, 2005). The resolution of 49 nm was demonstrated by applying this method to fluorescent beads (Gustafsson, 2005). For cell imaging, however, this is difficult to achieve because the strong excitation would bleach conventional fluorophores faster than the required number of images can be acquired. Special labels such as nanoparticles should be employed to make this method feasible. For live cell imaging another challenge is sample motion, because for one super-resolution image hundreds of raw frames with different phase and orientation of the illumination pattern should be acquired (**Fig. 10**), and during this time the sample should not move for distances larger than the resolution value.

Microscopy techniques, which use nonlinear fluorescence response, got the name RESOLFT (REversible Saturable Optical Linear Fluorescence Transitions). This applies not only to NL-SIM, but also to other techniques, such as stimulated emission depletion (STED) and ground state depletion (GSD) (see below). One of the saturable transitions that does not require high excitation power is the photoswitching of organic fluorophores (Heilemann et al., 2005; Irie et al., 2002) and fluorescent proteins (Dickson et al., 1997; Gurskaya et al., 2006). Such fluorophores can be switched on and off (between a fluorescent and a non-fluorescent state, or between two spectrally distinct fluorescent states) using light of a particular wavelength. For example, the Dronpa fluorescent protein can be switched on with a 405 nm wavelength, and it fluoresces and switches off under a 488 nm illumination (**Fig. 10A**). With this FP a resolution around 50 nm was demonstrated on proteins expressed inside cells (**Fig. 10D**) (Hirvonen et al., 2008; Rego et al., 2012). The intensity of the photoswitching light used in these experiments

was in the order of moderate 1–12 W/cm^2, as compared with ~8 MW/cm^2 peak power and ~30 W/cm^2 average power used for the excitation saturation (Gustafsson, 2005). Recently, new photoswitchable FPs with properties optimized for RESOLFT were developed: *e.g.*, Kohinoor (Lu-Walther et al., 2016; Tiwari et al., 2015) and Skylan-NS (Zhang et al., 2016). They withstand a high number of on/off cycles, have improved photostability, brightness and on/off contrast ratio, which allowed *e.g.* to image live cells at a resolution of 60 nm (Zhang et al., 2016).

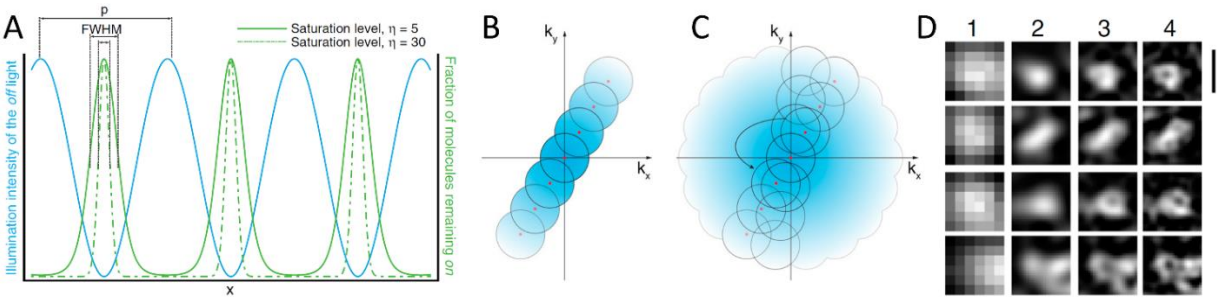


Figure 10. Nonlinear structured illumination microscopy. (**A**) Patterned light (λ=488 nm, blue) is used for switching off the fluorescence of Dronpa. A fraction of molecules at the minima of the pattern will remain fluorescent (green). With increasing exposure time (proportional to η) the emission pattern becomes narrower (green dashed line), its width can be much smaller than the diffraction limit. The emission pattern is no more sinusoidal which gives rise to higher-order harmonics (HOH). (**B**) Spatial frequencies observable with SIM extended to higher harmonics. (**C**) The entire Fourier space is filled by rotating the illumination pattern. (**D**) Clusters of a protein associated with the nuclear core complex POM121-Dronpa as imaged with (1) conventional microscopy, (2) linear SIM, (3) NL-SIM with one HOH, (4) NL-SIM with two HOH. Scale bar is 200 nm. The figure is adapted from (Rego et al., 2012).

A particular problem of SIM is the sensitivity to minor changes in parameters of experiments and image processing, which leads to appearance of artifacts in the final super-resolution images. Factors such as instability of the excitation intensity, imprecision in the shifts of the illumination pattern, bleaching of the sample, spherical aberrations due to the mismatch of the refractive index through the light path from the specimen to the objective, sample motion, etc. all influence the result of the image reconstruction (Choi and Kim, 2013; Förster et al., 2016; Schaefer et al., 2004). The artifacts typically look like periodic stripes or dots, which can sometimes be confused with real appearance of the labeled object. To reduce such artifacts,

different methods of their detection and correction have been proposed (Choi and Kim, 2013; Förster et al., 2016; Schaefer et al., 2004), but the intrinsic resolution limitation remains.

1.2.3. Stimulated emission depletion microscopy

In the confocal scanning microscope the size of the excited region at a given time is limited by diffraction (formula 1), and the detected fluorescence originates therefore from the same diffraction-limited region. S. Hell and J. Wichmann back in 1994 proposed to inhibit the fluorescence in the periphery of the confocal PSF using a phenomenon of stimulated emission, such that the detected light will originate only from the very center of the PSF, where the depletion intensity equals zero. This would effectively reduce the size of the PSF and improve resolution (Hell and Wichmann, 1994). During the stimulated emission, an incoming photon interacts with an excited electron and stimulates it to emit a photon with properties identical to the incoming photon, such as the wavelength. With this phenomenon an excited state of a molecule, which is the source of fluorescence, can be depleted using a higher-wavelength light outside the detection range (**Fig. 11**). The depletion beam can be made to have a shape of a torus with the zero intensity coinciding with the center of the excitation beam. Because the remaining fluorescence intensity depends on the depletion intensity in a nonlinear way the resulting PSF can be much smaller than the diffraction limit. The resolution of the STED microscope can be estimated by:

$$LFWHM_{STED} \approx 0.45 \frac{\lambda}{NA\sqrt{1+{}^{I_0}/_{I_{sat}}}} \tag{7}$$

Where I_0 is the intensity of the peak of the depletion beam and I_{sat} is the saturation intensity, *i.e.* the intensity at which the emission is reduced twice (Westphal and Hell, 2005). The resolution of STED microscopy is therefore not principally limited and it improves with increasing depletion intensity.

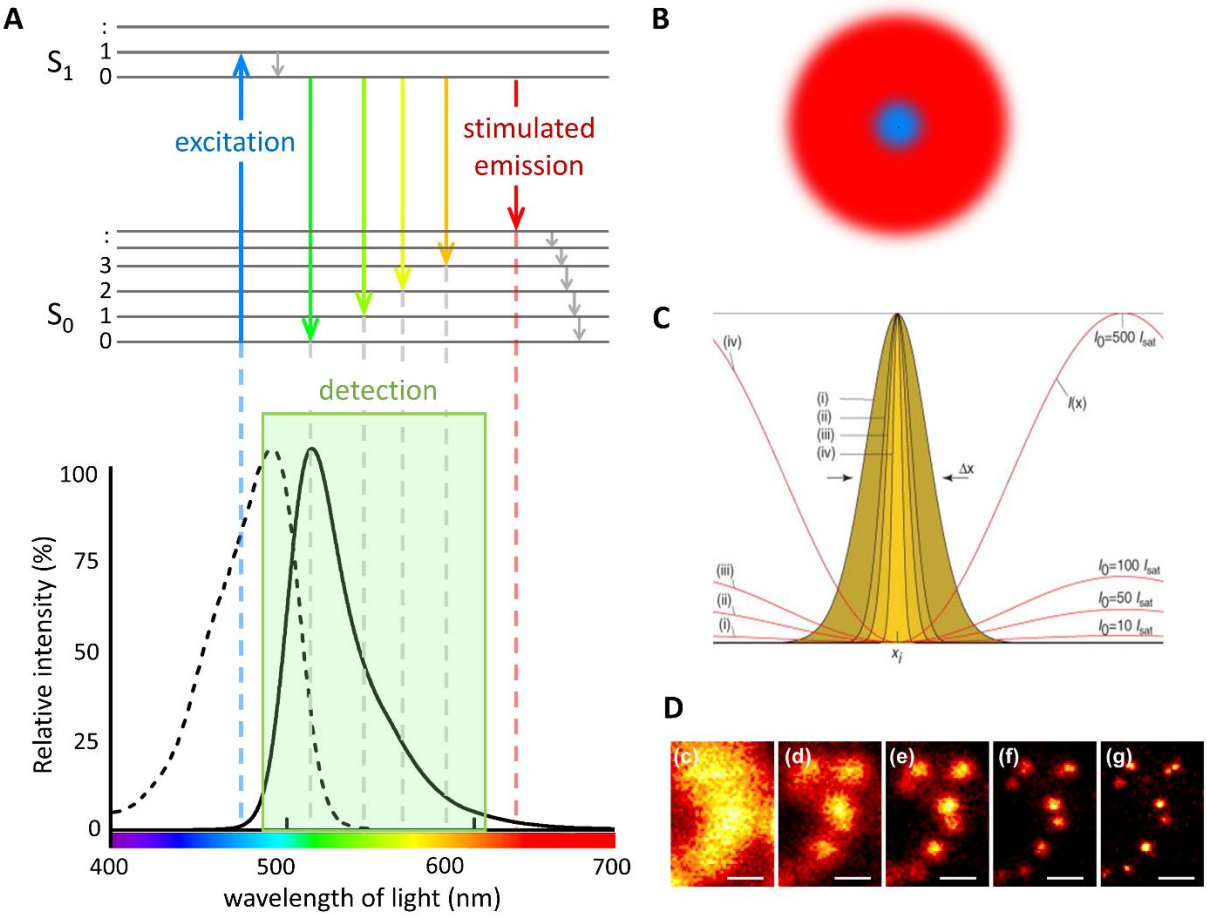


Figure 11. Stimulated emission depletion microscopy. (**A**) Jablonski diagram of a dye molecule with its excitation (dashed line) and emission (solid line) spectra. For STED microscopy, the dye is excited by a laser line (blue), the periphery of the PSF is depleted with a donut-shaped STED beam (red), most of its emission spectra between the two laser beams is detected (green). (**B**) Overlay of the excitation (blue) and the depletion (red) beams as used in STED microscopy, XY view. (**C**) Even though both excitation and depletion patterns are diffraction-limited, the resulting emission PSF (yellow) can be made smaller than the diffraction limit by increasing the intensity I_0 of the STED beam (Hell et al., 2004). (**D**) Experimentally observed improvement of resolution with increasing STED beam intensity, from 0 (c) to 1.7 GW/cm^2 (g) on example of 24 nm fluorescent beads. Scale bars, 200 nm. **C-D** are adapted from (Harke et al., 2008; Hell et al., 2004); spectra by Zadelrob, attributed under a CC BY-SA 3.0 license https://en.wikipedia.org/wiki/File:Fluorescein-spectra3.svg.

In practice, however, the reachable resolution is limited by the STED beam intensity which the sample can withstand. This is why STED microscopy has demonstrated a very high resolution (6 nm) on bright and photostable nanoparticles (Rittweger et al., 2009) and a more moderate

resolution of 30–70 nm with more practical for labeling organic dyes (Schermelleh et al., 2010), reaching ~70–90 nm practically on biological samples.

For the improvement of axial resolution additionally to the lateral one, the STED beam can be specially designed to have maxima above and below the focal plane, together with the ring lateral shape within the focal plane. This allows to achieve, *e.g.* isotropic resolution in the order of 100 nm in all three directions (Klar et al., 2000). Nowadays, commercial systems, such as Leica STED 3X or Abberior Easy3D allow for adjustment of the intensity for the lateral and the axial "donuts" which results in tuning of the lateral and axial resolutions or "PSF engineering".

Since STED is a point scanning technique similar to confocal laser scanning microscopy, one of the principal disadvantages of STED is the low imaging speed. To solve this problem, parallelized imaging schemes have been developed, e.g. with the wide-field excitation, a depletion pattern with 2000 minima and an on-camera detection (Bergermann et al., 2015). However, a point scanning technique has a number of advantages, such as flexible spectral separation of dyes, which allows for easy multicolor imaging. Other techniques previously available on the base of the confocal setup have been extended for STED microscopy. For example, two photon excitation STED microscopy allows for super-resolution imaging rather deep into tissues, *e.g.* in the brain tissue (Bethge et al., 2013; Moneron and Hell, 2009; Takasaki et al., 2013); STED-fluorescence correlation spectroscopy (STED-FCS) sheds light on molecular mobility of lipids in the plasma membrane (Honigmann et al., 2014); STED-fluorescence lifetime imaging (STED-FLIM) permits precise dye separation and colocalization studies (Bückers et al., 2011; Lesoine et al., 2012) and improvement of resolution using gated lifetime detection (Vicidomini et al., 2011).

Despite impossibility of one-photon fluorescence excitation with the STED wavelength, the very high intensity of depletion can be sufficient for two-photon excitation of fluorophores (Danzl et al., 2016). Additionally, STED light can excite higher-energy S_n states from the S_1 state (Danzl et al., 2016). Both processes lead to an increased background and to an early bleaching of fluorophores. Recent studies sought for decreasing the undesirable effects of powerful STED excitation. In one of them it was proposed to use reversibly switchable FPs, such as rsEGFP (Grotjohann et al., 2011), rsCherryRev (Stiel et al., 2008) or Dronpa (Ando et al., 2004). The off-state of such fluorophores was found not to be accessible for a simultaneous excitation and depletion illumination (Danzl et al., 2016). To avoid the STED-induced bleaching, the sample has to be illuminated by a "switching-off" donut-shaped beam prior to

depletion with a beam of the same shape. This combination allows not only for strongly reduced bleaching, but also for resolution improvement because of the synergetic action of the deactivation and the depletion beams. This technique is called “Protected STED” (Danzl et al., 2016).

Another demonstrated improvement of the STED technique is a restriction of the scanned field of view (FOV) of the sample. By doing so, the ROI does not suffer from the strong intensity of the crest of the donut and as a consequence, the fluorophores inside the ROI do not bleach. In fact, the very high intensity of the donut’s crest is not needed by itself for resolution improvement, it is only needed to approach the saturation intensity close to the center of the PSF. The saturation intensity (STED intensity at which the half of fluorophores are depleted) is orders of magnitude smaller than the peak STED beam intensity (Göttfert et al., 2017). By restricting the ROI to a square of around 100 x 100 nm^2, the MINFIELD-STED approach allows to get around 100 times more signal from the same ROI as compared to a wide field STED (Göttfert et al., 2017). One of the imaging modalities proposed by the authors is a low-resolution scanning for the search of the ROI with a consecutive high-power STED scanning of small ROIs in order to get best signal and resolution. This approach can be especially useful for 3D STED where the laser intensities are much higher and shaped beams are spread over large areas which provokes increased exposure for the depletion light during conventional wide-field scanning (Göttfert et al., 2017).

STED microscopy has found a wide range of applications in biological science. In neurobiology, for example, it was used for fixed and live brain imaging of pre- and post-synaptic structures (Bär et al., 2016; Nishimune et al., 2016; Tønnesen et al., 2011) as well as axons in the brain (Chéreau et al., 2017) and in the peripheral nervous system (D’Este et al., 2017). Other applications of STED include studies of phagocytosis (Baranov et al., 2016), chromatin organization at double strand brakes (Reindl et al., 2017), nanoscale dynamics of plasma membrane lipids (Eggeling et al., 2009), transforming growth factor receptors at post-Golgi vesicles (Ruan et al., 2016) and dynamics of human immunodeficiency virus maturation (Hanne et al., 2016).

1.2.4. Single-molecule localization microscopy

When a fluorescently labeled sample is excited by light with the wavelenght inside the absorption band of the fluorophores, the dyes usually emit fluorescence all in the same time continuously until the excitation is switched off or the dyes are bleached. Since the microscope image of every molecule is the PSF with the size *d* (1), during the simultaneous observation of all the dyes all the corresponding PSFs overlay, which produces a diffraction-limited image and therefore the observation of details smaller than the order of *d* becomes impossible (**Fig. 1B-C**, **Fig. 12A**). However, if there is no more than one molecule per diffraction-limited volume, the PSFs of neighboring dyes do not overlap. Since the PSF is usually axially symmetric around the position of the source, the position of the molecule can be determined as the center of the Gaussian-fitted PSF. This is routinely used for single-particle tracking (SPT) or SM tracking with up to a nanometer precision, which allowed to see *e.g.* the hand-over-hand walking of myosin V on actin (Yildiz et al., 2003), kinesin-driven movements (Gelles et al., 1988) or diffusion of single lipid molecules through a membrane (Schmidt et al., 1996).

The precision of SM localization is given by (Thompson et al., 2002):

$$\langle(\Delta x^2)\rangle = \frac{\sigma^2}{N} + \frac{a^2}{12N} + \frac{8\pi\sigma^4 b^2}{a^2 N^2}, \tag{8}$$

where Δx is the standard deviation (SD) of the localization error, σ is the SD of the PSF, N is the number of detected photons, a is the pixel size and b is the SD of the background noise. In this expression the first term corresponds to the photon-counting noise, the second one corresponds to the pixelation noise originating from a non-zero pixel size and the third term stands for the background noise contribution.

A transition point in the number of photons can be introduced (Thompson et al., 2002):

$$N_t = \frac{8\pi\sigma^4 b^2}{a^2(\sigma^2 + {a^2}/{12})} \tag{9}$$

The localizations of spots with $N < N_t$ will be dominated by background noise, the localizations with $N > N_t$ will be dominated by shot (photon-counting) noise.

In the ideal case, the pixel size can be made arbitrarily small and the background noise can be removed. In this case the localization precision (LP) is limited only by photon counting and can be expressed simply:

$$\Delta x \approx \frac{\sigma}{\sqrt{N}} \tag{10}$$

The number of photons that can be collected from a single fluorophore is in the order of 10^3–10^6, depending on fluorophore and imaging conditions (Qu et al., 2004; Yildiz et al., 2003), which generates a LP in the nanometer range.

Back in 1994 Eric Betzig proposed to perform super-resolution imaging by separation of the PSFs in a multidimensional space using distinguishing optical characteristics (Betzig, 1995; Burns et al., 1985). For example, if different dyes have distinct spectral properties, they should be separated based on these characteristics up to the level where there is less than one molecule per the multidimensional PSF volume. With this separation, the coordinates of the individual molecules could be assigned as the centers of the corresponding PSFs. The principle was first proposed (Betzig, 1995) and later demonstrated (Oijen et al., 1999) for super-resolution imaging of single molecules isolated in a host crystal matrix under cryo-conditions (Ambrose et al., 1991; Talon et al., 1992). In practice, however, only a few markers inside the diffraction-limited region could be resolved (Churchman et al., 2005; Lacoste et al., 2000), orders of magnitude less than it would be necessary for resolving features in a dense cellular context. The use of the spectral properties for SM separation was not practical because the absorption and the fluorescence spectra of dye molecules are broad under normal conditions and even if they were narrow enough, their fine spectral separation would require a very complex microscopy setup.

Super-resolution microscopy by exploiting the time dimension proved to be more practical. First attempts to separate dyes in time relied on bleaching of fluorophores. Despite simultaneous light emission during excitation, individual fluorophores bleach at different time points. In other words, diminution of fluorescence of a labeled object under constant excitation occurs in discrete steps corresponding to SM bleaching. By analyzing the intensity profiles of diffraction-limited regions over time, a number of states with different numbers of active fluorophores can be distinguished. The position of the longest-lasting molecule can be determined as a simple fitting of the SM image of the last state before the complete bleaching of the region. The positions of the second-longest-lasting dye can be determined by fitting the difference image of the two last states before the complete bleaching, and so on (Qu et al., 2004). With this method, again only a few fluorophores could be resolved (Gordon et al., 2004; Qu et al., 2004). Additionally, the technique can only be applied for studies of processes stationary in time and

space, because for proper subtraction of the PSFs the fluorophores should not move throughout acquisition (Qu et al., 2004).

Practically, molecule separation by photoactivation turned out to be the most successful technique. Fluorescent proteins, such as mutants of GFP (Yang et al., 1996), were found to have a non-fluorescent "dark" form where the protein passes after excitation of the fluorescent form. The protein could transfer back from the dark to the fluorescent form upon excitation with near-UV light (Dickson et al., 1997). At the beginning of 2000s several photoactivatable and photoswitchable FPs with improved characteristics have been reported (Ando et al., 2002; Chudakov et al., 2003; Patterson and Lippincott-Schwartz, 2002). Betzig *et al.* used them, fused with the proteins of interest, for super-resolution imaging by photoactivating only a small amount of FPs at a time (Betzig et al., 2006). The sample was illuminated continuously at a wavelength near the excitation maximum of the "on" form of FPs. At the beginning of the experiment, a small fraction of PFs was in the activated state and they could be resolved as single molecules. After some time, these dyes bleached and the number of observed molecules decreased. Then, a brief 405 nm light was applied for photoconversion of the next portion of FPs from the dark state. This procedure continued until most of dyes have been photoconverted and bleached (**Fig. 12D**). In total, ~10^4…10^5 images were acquired producing ~10^5…10^6 localized single molecules. The sample was illuminated most often in a TIRF geometry (**Fig. 7**) for background reduction; the images were acquired with a highly sensitive electron-multiplying charge coupled device (EMCCD) camera for photon counting. In all other respects the imaging system was a common epifluorescence microscope (**Fig. 3**). This technique could resolve individual fluorophores at such densities as the densities of proteins in cells (**Fig. 13 A-B**). This method was termed "photoactivated localization microscopy" or PALM. Similar concept, using PA-GFP (Patterson and Lippincott-Schwartz, 2002), was proposed by another group under the name "fluorescence photoactivation localization microscopy" (FPALM) (Hess et al., 2006).

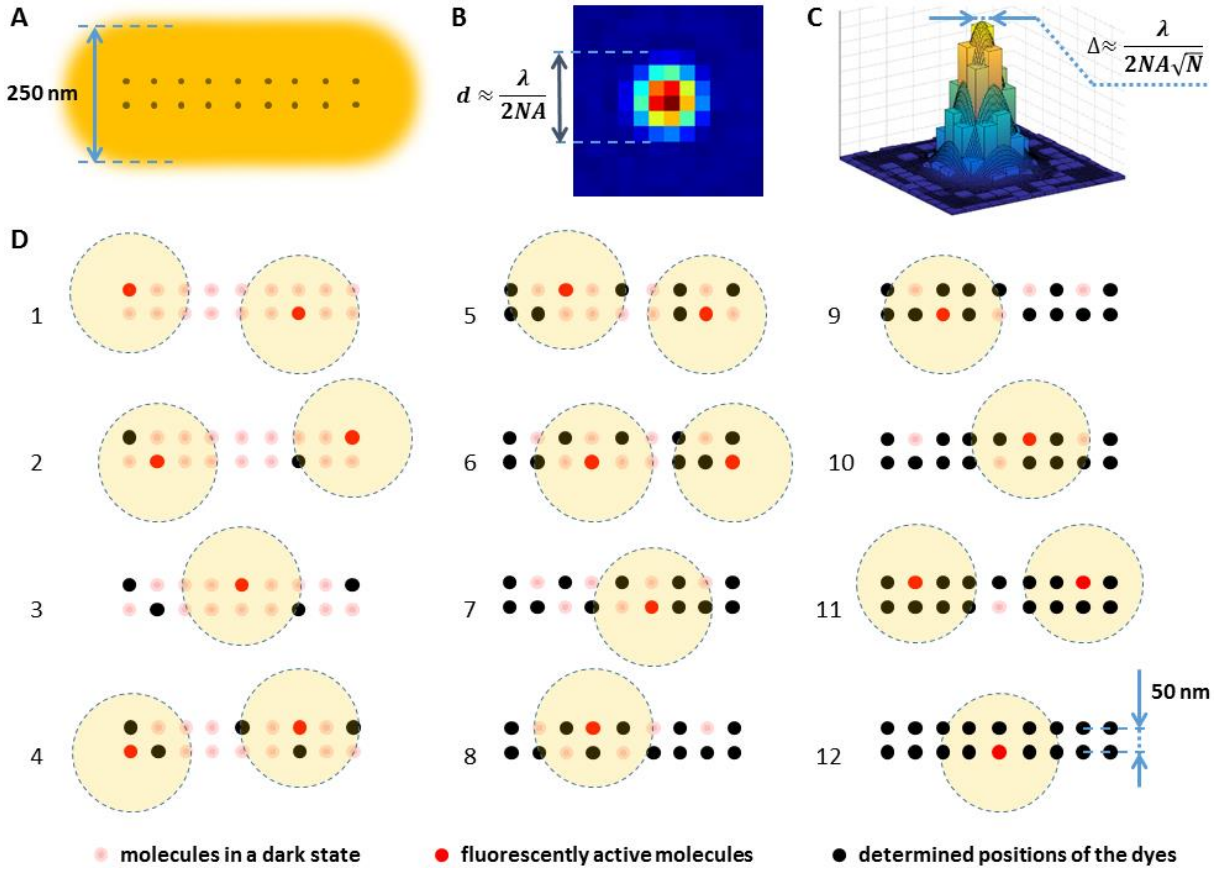


Figure 12. Principle of single-molecule localization microscopy. (**A**) In conventional fluorescence microscopy all fluorophores emit light simultaneously, therefore any details of the sample smaller than the order of the wavelength are obscured due to light diffraction. (**B**) A single dye molecule is imaged on a camera as a disk with diameter *d*. (**C**) The center of the SM image, which corresponds to the position of the molecule, can be determined with precision Δ, higher than *d* and dependent on the number of photons *N* detected from this molecule (9). (**D**) During a SMLM experiment, in every moment most of the dyes are found in a non-fluorescent dark state (light red dots). A small number of random fluorophores, less than one per the diffraction-limited region, is activated and imaged separately as single molecules (red dots). The positions of these molecules can be therefore determined with a high precision Δ (black dots). After these dyes are bleached or brought back to the dark state, another random portion of fluorophores can be imaged and localized in the next frames. The complete super-resolution image can be reconstructed from the SM coordinates after most of the dyes have passed the activation – localization – deactivation cycle (**D, 12**). **1–12** stand for the frame number. The yellow circles represent the lateral size of the PSF.

Organic fluorophore Cy5 in an oxygen scavenging buffer with β-mercaptoethanol or β-mercaptoethylamine (MEA) was found to exhibit a switching behavior between a fluorescent and a dark state upon illumination with light of different wavelengths (Bates et al., 2005; Heilemann et al., 2005). As a mechanism responsible for the switching, a transition to a long-living dark state from the triplet states was proposed (Bates et al., 2005), however the exact mechanism remains unknown (Heilemann et al., 2005). The dark state has a lifetime in the order of hours (Bates et al., 2005). Additionally, Cy3 dye in a very close proximity (< 1 nm) of Cy5 molecule largely facilitated the photoactivation of Cy5. It was therefore proposed to create a SM optical switch, where a Cy5 (reporter) and a Cy3 (activator) molecule is attached to a DNA molecule. The switch can be converted to the dark state by excitation with red light (a 638 nm laser) and activated to the fluorescent state with green light (532 nm) (Bates et al., 2005).

At the same time as (Betzig et al., 2006; Hess et al., 2006), Rust *et al.* published a technique which provides super-resolution imaging by activating a sparse subset of Cy5/Cy3 fluorophores at a time (Rust et al., 2006), a method they called "stochastic optical reconstruction microscopy" (STORM). They have also demonstrated subdiffraction imaging by immunolabelling of proteins with this optical switch (**Fig. 13 C-D**). The approach was quickly extended to a broad range of fluorophores, including the Cy5.5/Cy3, Cy7/Cy3, Alexa 405/Cy5, Cy2/Cy5, Alexa 647/Cy3 reporter/activator pairs (Bates et al., 2007). Besides the dyes and the laser lines, both methods, (F)PALM and STORM, use the same principle, the photoactivation of a small number of fluorophores at a time with following localization of SM PSFs (**Figs. 12, 13**). To avoid confusion, further in the text super-resolution imaging techniques based on localization of individual fluorophores will be called "Single-Molecule Localization Microscopy" (SMLM).

By using fast spontaneously blinking FPs and high camera frame rates (frame exposure time being approximately equal to the on-time of fluorophores), it became possible to significantly improve the acquisition speed and SNR. This implementation also required only one laser (488 nm) because the used FP rsFastLime (Stiel et al., 2007) could be excited and photoactivated at the same wavelength. This method was called PALM with independently running acquisition (PALMIRA) (Egner et al., 2007; Geisler et al., 2007). A short time later, rsFastLime and Cy5 were used together for double-color experiments. The Cy5 channel was acquired first, using a 633 nm laser excitation, and then the second channel (rsFastLime) was imaged with a 488 nm excitation (Bock et al., 2007).

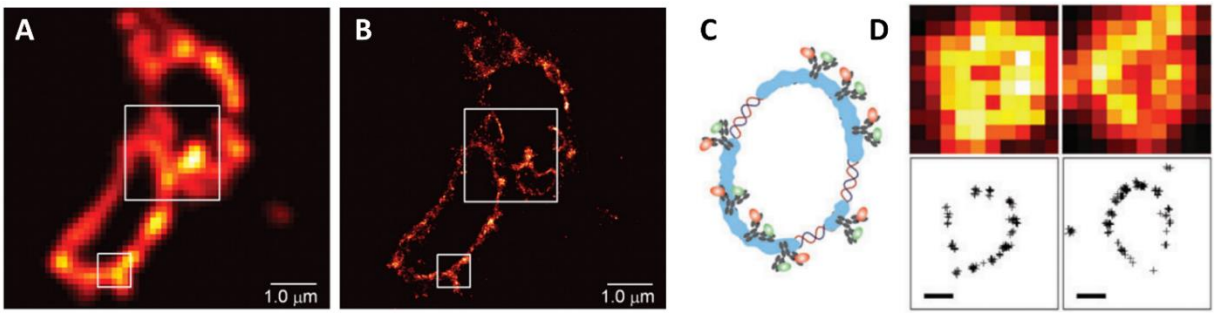


Figure 13. Some of the first images obtained with SMLM. Conventional TIRF (**A**) and PALM (**B**) images of a lysosomal transmembrane protein CD63 tagged with the Kaede FP (Betzig et al., 2006). Indirect immunofluorescence labeling of RecA-coated circular plasmid DNA: secondary antibodies are labeled with both Cy3 and Cy5 dyes (**C**); conventional TIRF image (**D**, top) and reconstructed STORM images of the same objects (**D**, bottom) (Rust et al., 2006). Scale bars: 1 µm (**A**, **B**), 300 nm (**D**). The diffraction-limited image **A** was obtained by summing all the frames of the SM experiment. The image **B** is obtained by rendering every localization as a Gaussian function with the width equivalent to the LP of the corresponding molecule. Images **A** and **B** are adapted from (Betzig et al., 2006); images **C** and **D** are adapted from (Rust et al., 2006).

Soon it was demonstrated that common cyanine dyes, such as Cy5, Alexa 647, Cy7 or Alexa 680, can be used directly, without coupling with an activator fluorophore, for super-resolution imaging, when switched on and off under simultaneous excitation with appropriate intensity at two wavelengths (Heilemann et al., 2005, 2008). It was an important finding because it allowed to use for super-resolution microscopy commercially available dye-conjugated antibodies with little or no modification of common immunofluorescence (IF) protocols. This technique was called “direct STORM” or dSTORM (Heilemann et al., 2008).

At the same time, Fölling *et al.* proposed to exploit long-living electronic states intrinsic for any dye molecule, such as the triplet state T_1 (Fölling et al., 2008), a concept previously proposed for STED-like ground-state depletion (GSD) microscopy (Bretschneider et al., 2007; Hell and Kroug, 1995). This would extend the localization super-resolution techniques beyond photoconvertible fluorophores. Long-living molecular states, accessible from the S_1 state, allow to push there most of the dye molecules, which depletes the singlet states where very few molecules remain. A continuous illumination with intensity $I > I_s = h\upsilon/(\Phi_{isc}\sigma\tau) \approx 1$ kW/cm² minimizes the fraction of molecules in the ground state to $\varepsilon \approx \tau_{fl}/(\Phi_{isc}\tau)$, where $h\upsilon$ is the excitation photon energy, Φ_{isc} is the probability of the intersystem conversion (from the singlet

to the triplet states), τ_{fl} is the fluorescence lifetime, τ is the triplet state lifetime and σ is the absorption cross-section of the molecules in the ground state (Fölling et al., 2008) (**Fig. 14**). Typically, $\Phi_{isc} \leq 10^{-3}$ meaning that a molecule with high quantum yield emits at least 10^3 photons before going to the dark triplet state. The lifetimes of the triplet and the other dark states are in the order of 10^{-3}–10^3 s, which is 10^5–10^{11} times longer than the fluorescence lifetimes ($\tau_{fl} \sim 10^{-9}$–10^{-8}). This means that under sufficiently powerful excitation ($I > I_s$) the fraction of molecules in the bright singlet state can be reduced to $\varepsilon << 10^{-2}$ (Fölling et al., 2008) that can be sufficient for spatial separation of individual molecules even in densely labeled samples (**Fig. 15**).

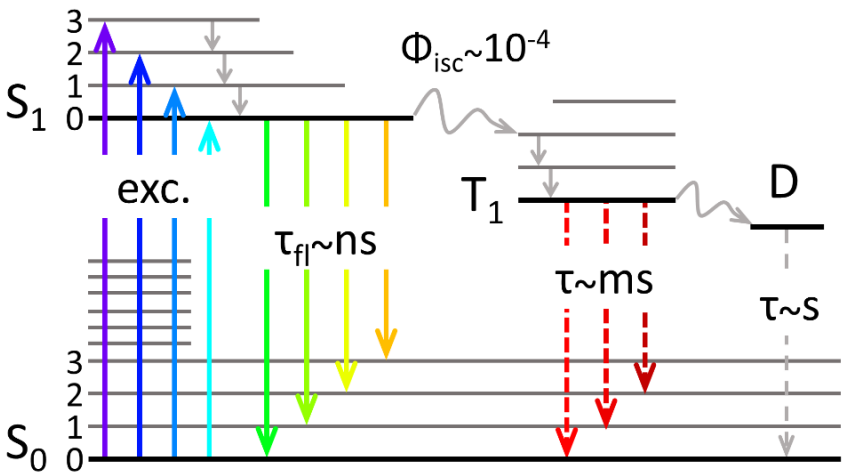


Figure 14. Jablonski diagram of a typical fluorophore suitable for GSDIM microscopy. The molecule is excited from the ground state S_0 to one of the S_1 states, from where it can relax back to S_0 with fluorescence emission, or it can pass to the triplet state T_1 (and then possibly to the dark state D) with probability Φ_{isc}. In the T_1 and D states molecules are not accessible for excitation at the excitation wavelength of the ground state. These states are long-living, therefore a high population of molecules can be pushed and kept there with powerful excitation. This can dramatically deplete the singlet system providing observation of individual molecules (**Fig. 15**) (Fölling et al., 2008).

This technique was termed "Ground-State Depletion followed by Individual Molecule return" (GSDIM). It was found that, exploiting the GSDIM principle, many conventional fluorophores can be used for super-resolution (SR) imaging. Among them are Atto 532, Atto 655, Alexa 488, FITC, Oregon Green, Alexa 647 (**Fig. 15**), Texas Red; FPs EGFP, EYFP, Citrine and many others (Dempsey et al., 2011; Fölling et al., 2008; Heilemann et al., 2009; Linde et al., 2008).

The method in principle requires only one laser, because dye activation (the conversion from the long-living states to the ground state) occurs spontaneously. However, for some fluorophores this activation can be optionally accelerated with a shorter-wavelength excitation (Fölling et al., 2008). The wide choice of dyes and the simple setup allow easy multicolor imaging, including with excitation of a single wavelength and spectral unmixing of fluorophores (Fölling et al., 2008; Testa et al., 2010).

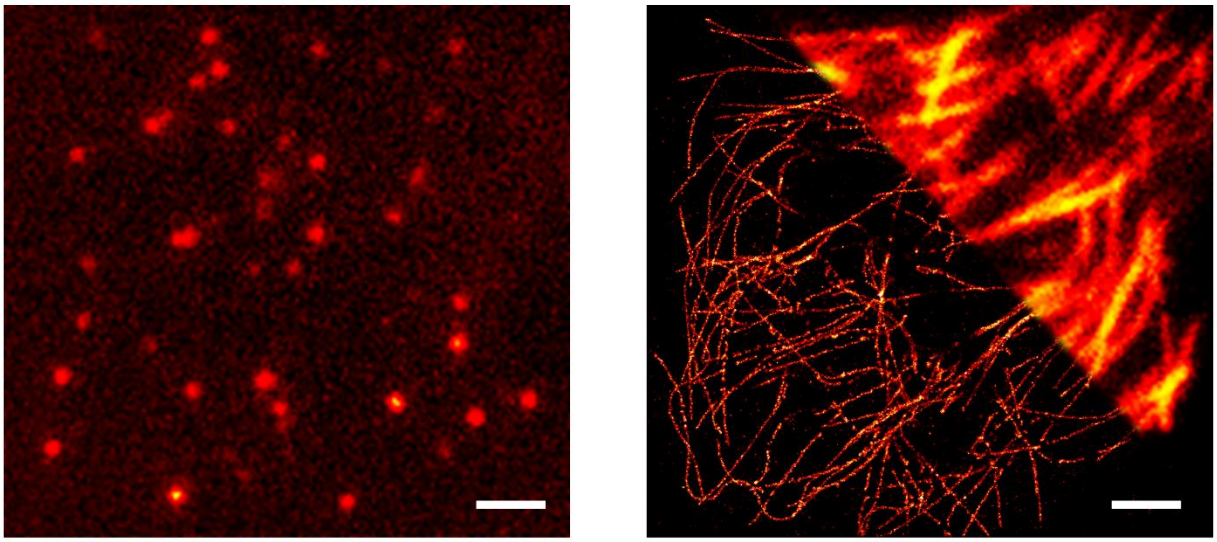


Figure 15. GSDIM imaging of β-tubulin in HeLa cells labeled with Alexa 647-conjugated secondary antibodies. One of the frames acquired during the experiment (left). Reconstructed SR image with the conventional TIRF image in the top right corner (right). The data was acquired on the in-house Leica SR GSD system (TIRF illumination, excitation with the 642 nm laser at 50% intensity, 6.34 ms frame exposure time and 300x EM gain) and processed with SharpViSu (Andronov et al., 2016a) for drift correction and visualization in the histogram mode with a pixel size of 16 nm.

1.2.4.1. Methods for three-dimensional SMLM

PSF modifications: astigmatism

Two-dimensional (x,y) coordinates of individual dyes can be obtained straightforward as the coordinates of the centroid of the PSFs providing improved lateral resolution in SMLM. However, this does not provide any improvements in the axial (z) direction and the axial resolution remains the same as for the conventional wide-field microscopy case (3). Moreover,

imaging of multiple planes in the Z direction is limited by fluorophore bleaching. Since most of cellular and tissual structures are 3D objects, their studies would highly benefit from super-resolution imaging in all three dimensions (see also section 3.3).

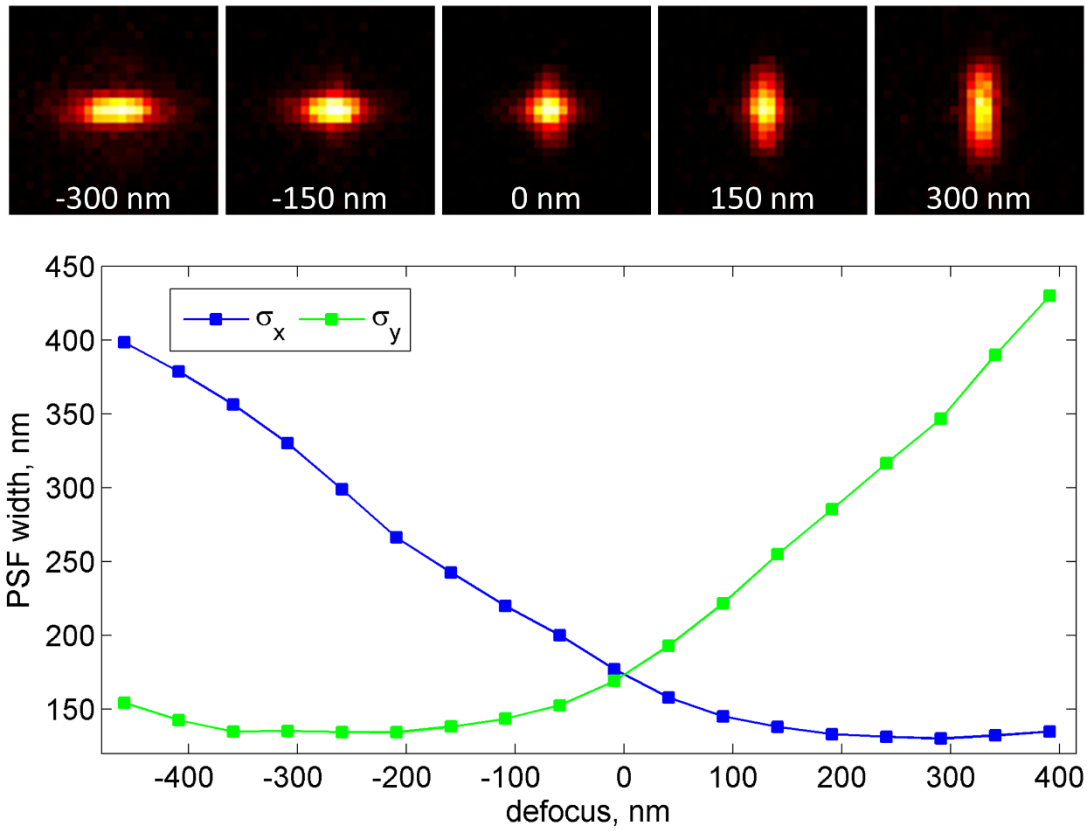


Figure 16. 3D SMLM based on astigmatism. (top) Images of a subdiffraction-sized fluorescent bead through a microscope with astigmatism. Numbers indicate the defocus value. (bottom) Change of the lateral size of the SM image with defocusing. σ_x, σ_y is the 2D Gaussian spread in the x and y direction, correspondingly. The experiment was acquired on the in-house Leica SR GSD system within the framework of its adaptation for astigmatism-based 3D SMLM.

The PSF of an objective without aberrations is axially symmetric, and while the PSF width increases with defocusing, it did not provide a sufficiently robust method for the axial position determination, especially in low SNR conditions of bioimaging. Also, the expansion of the PSF occurs in both directions of defocusing that makes the determination of the exact position ambiguous. One of the ways to determine this position in SPT applications was breaking the axial symmetry of the PSF with an artificially induced astigmatic aberration (Holtzer et al., 2007; Kao and Verkman, 1994). With astigmatism, the PSF becomes elongated with defocusing, and the shapes of the PSF for under- and overfocusing are perpendicular to each other (**Fig. 16**), which allows a robust and unambiguous determination of the axial coordinate

of individual particles. The reported precision of the Z-position determination is in the order of 12–30 nm (Holtzer et al., 2007; Kao and Verkman, 1994).

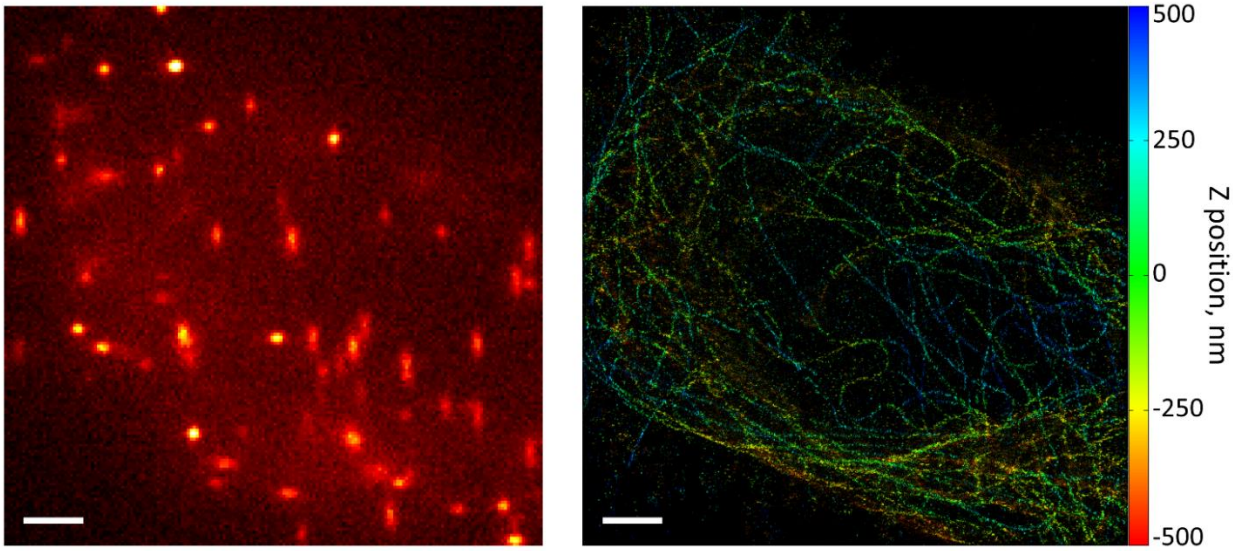


Figure 17. 3D SMLM of β-tubulin labeled with Alexa 647-conjugated secondary antibodies in a HeLa cell. (left) A frame of an acquisition with elongated by astigmatism PSFs corresponding to individual molecules situated at different depths. (right) Reconstructed 3D image with depth encoded in hue. Scale bars, 2 µm. Acquired on a Leica GSD 3D instrument (Wetzlar, Germany).

This technique has been naturally extended for SMLM (Huang et al., 2008a). For creation of astigmatism, a weak cylindrical lens can be introduced in the imaging path. Alternatively, astigmatism can be created with adaptive optics which also allows for flexible correction of other optical aberrations (Izeddin et al., 2012). Fitting of PSFs with a 2D Gaussian function

$$f(x,y) = A \cdot \exp\left(-\left(\frac{(x-x_0)^2}{2\sigma_x^2} + \frac{(y-y_0)^2}{2\sigma_y^2}\right)\right) \quad (11)$$

provides not only the (x_0, y_0) centroid coordinates, but also the widths (σ_x, σ_y) in both X and Y directions. The functions $\sigma_x = f(z)$ and $\sigma_y = f(z)$ can be calibrated using fluorescent subdiffraction objects and known objective defocus (**Fig. 16**). From this calibration the Z coordinates of fluorophores in an experimental dataset can be determined (Huang et al., 2008a) (**Fig. 17**). Using the Cy3/Alexa 647 activator/reporter pair, a LP of ~24 nm in the lateral and of ~52 nm in the axial direction was reached (Huang et al., 2008a). The concept was quickly extended for multicolor imaging, *e.g.* for studies of interactions between mitochondria and microtubules (Huang et al., 2008b). Disadvantages of 3D SMLM with astigmatism are non-

isotropic LP, degradation of resolution when going out of focus and imaging depth limited to 600-1000 nm around the focal plane.

PSF modifications: double-helix PSF

Another reported method for breaking the axial symmetry of the PSF was giving it a double helical shape with a liquid crystal spatial light modulator (Pavani and Piestun, 2008). When applied to SMLM, every individual fluorophore is depicted as two close diffraction-limited spots, rotated around their common center with an angle, proportional to the axial distance between the fluorophore and the focal plane (**Fig. 18**) (Pavani et al., 2009). With this method, the SM LP of ~12 nm in the lateral and of ~20 nm in the axial direction was reported. The advantage of this method is an extended to ~2 µm depth of field, an LP closer to isotropy and less dependent on the axial position (Lew et al., 2010). Double-helix PSF was applied *e.g.* for life-cell two-color 3D super-resolution imaging of cytoskeletal proteins and cell membrane in bacteria (Lew et al., 2011).

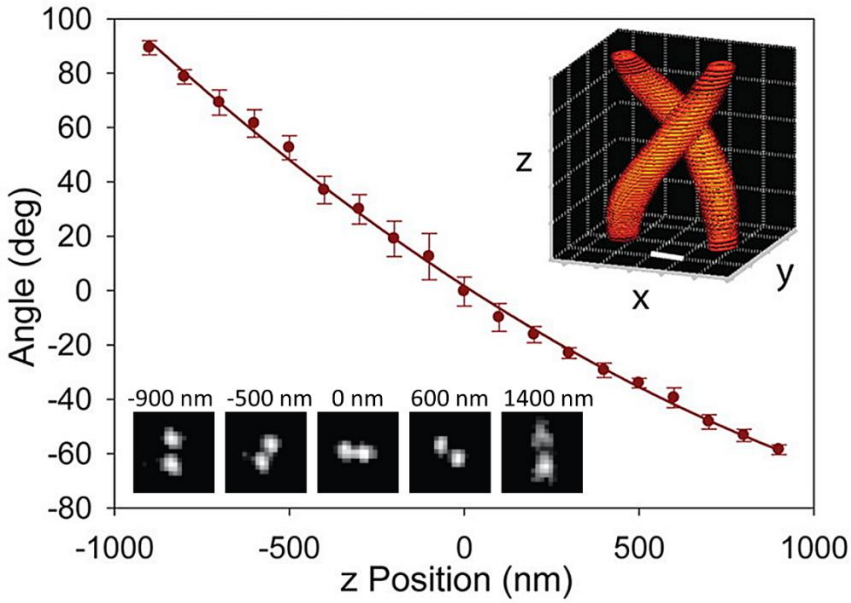


Figure 18. 3D SMLM with engineered double-helix PSF. (top right) 3D view of the double helix PSF, scale bar is 400 nm. (middle) Calibration curve, the dependence of the rotation angle with respect to the horizontal versus the axial position. (bottom left) Images of a fluorescent bead at different Z positions. Figure is adapted from (Pavani et al., 2009).

Detection path splitting

By modifying the detection path of a fluorescence microscope, it is possible to simultaneously detect several different focal planes in the sample (**Fig. 19**). This detection scheme was first used for particle tracking in living cells (Prabhat et al., 2004) and then was extended for SMLM (Juette et al., 2008). By separating two focal planes on ~500 nm in the sample, the fluorophores situated between them are captured with two cameras in different ways in terms of spot intensity and shape. This gives access to their lateral (axial) position with a precision of 30 nm (75 nm) over a ≲1 µm-axial detection range (Juette et al., 2008). An advantage of the biplane scheme is that the axial and the lateral resolution is virtually independent on the axial position of the fluorophore (Juette et al., 2008).

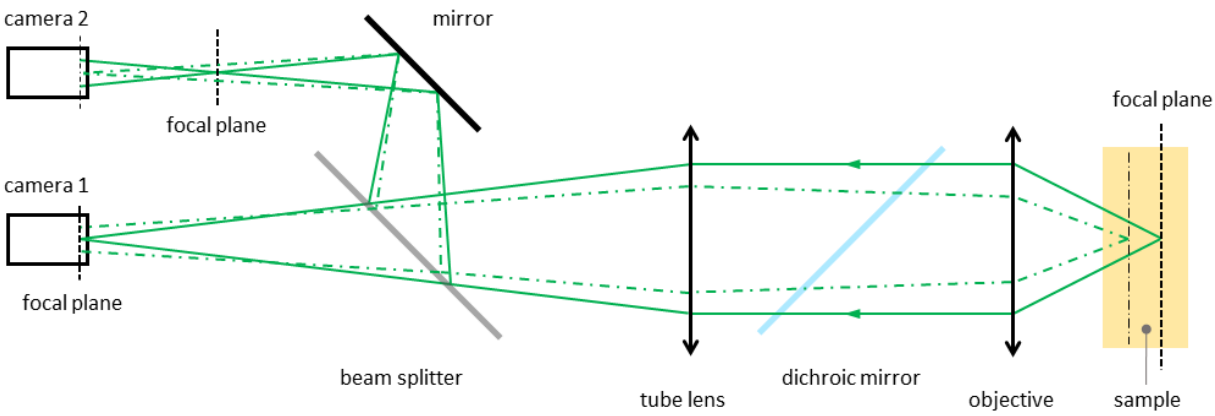


Figure 19. Scheme of 3D SMLM with dual focal plane imaging (detection path). Using a 50:50 beam splitter, half of the fluorescence is diverted on a second camera (or on the second half of a single camera (Juette et al., 2008)) via a longer light path resulting in a shifted focus position.

Other interesting examples of 3D SM imaging with detection path splitting include "Parallax" where the fluorescence beam is split into two parts corresponding to the views of the sample by the objective from the two sides. When the fluorophore goes away from the focal plane, the two images move towards or away from each other, reflecting the axial position (Sun et al., 2009). Another example is the "virtual volume super resolution microscopy" (VVSRM) which allows, using a mirror behind the sample, to image a side view of fluorophores at the same time with the front view. The side view provides straightforwardly the axial position of the dyes with a precision, equivalent to the lateral one (Tang et al., 2010).

4Pi microscopy

In the classical epifluorescence geometry the illumination and the collection of light happen from only one side of the sample. Fluorescence is emitted in all directions around the fluorophore, therefore the light spread in the direction opposite of the objective is lost. This geometry also leads to a PSF elongated in the axial direction and as the consequence to non-isotropic resolution. If a second objective is placed from the other side of the sample and it is used simultaneously with the first one for a coherent illumination and detection, the PSF of such a microscope can approach the isotropic one as a result of a constructive interference (Hell and Stelzer, 1992). This technique was first implemented by S. Hell and E. Stelzer for laser scanning confocal microscopy (Hell and Stelzer, 1992). In the ideal case, one objective can collect light from a solid angle of 2π, thus two objectives could collect it from the complete sphere of 4π. This fact gave the name to this technique, "4Pi microscopy". The best results from the 4Pi confocal microscope could be obtained when both the excitation beam interfered in the focal spot and the fluorescence light interfered in the detector spot: an axial resolution of 75 nm was reported (Hell et al., 1994).

A similar two-objective concept was developed for widefield microscopy with simple lamp illumination by M. Gustafsson *et al.* (Gustafsson et al., 1995) and termed "I^5M" (Gustafsson et al., 1999, 1996). The technique was able to achieve an axial resolution of 70 nm as well (Gustafsson et al., 1999). The biggest problem of both 4Pi and I^5M techniques is the presence of intense side lobes in the axial direction on both sides of the central peak of the PSF, which leads to appearance of "ghost images". To reduce artifacts, a deconvolution is necessary, which is not always feasible in practice because it requires an exact PSF shape. For 4Pi microscopy, however, the side lobes can be strongly reduced using two-photon excitation (Hänninen et al., 1995) due to quadratic dependence on the illumination intensity (Bewersdorf et al., 2006).

The double-objective detection scheme quickly found its applications for 3D SMLM. Same photons emitted by a fluorophore can interfere with themselves after having passed through the two objectives. By letting the two beams interfere in *e g.* three different ways and imaging the interference intensity with three cameras (**Fig. 20**), the axial position of the fluorophore could be determined with sub-10 nm precision. This technique was called "interferometric PALM" (iPALM) (Shtengel et al., 2009) and allowed *e.g.* to reveal nanoscale protein organization in focal adhesions (Kanchanawong et al., 2010).

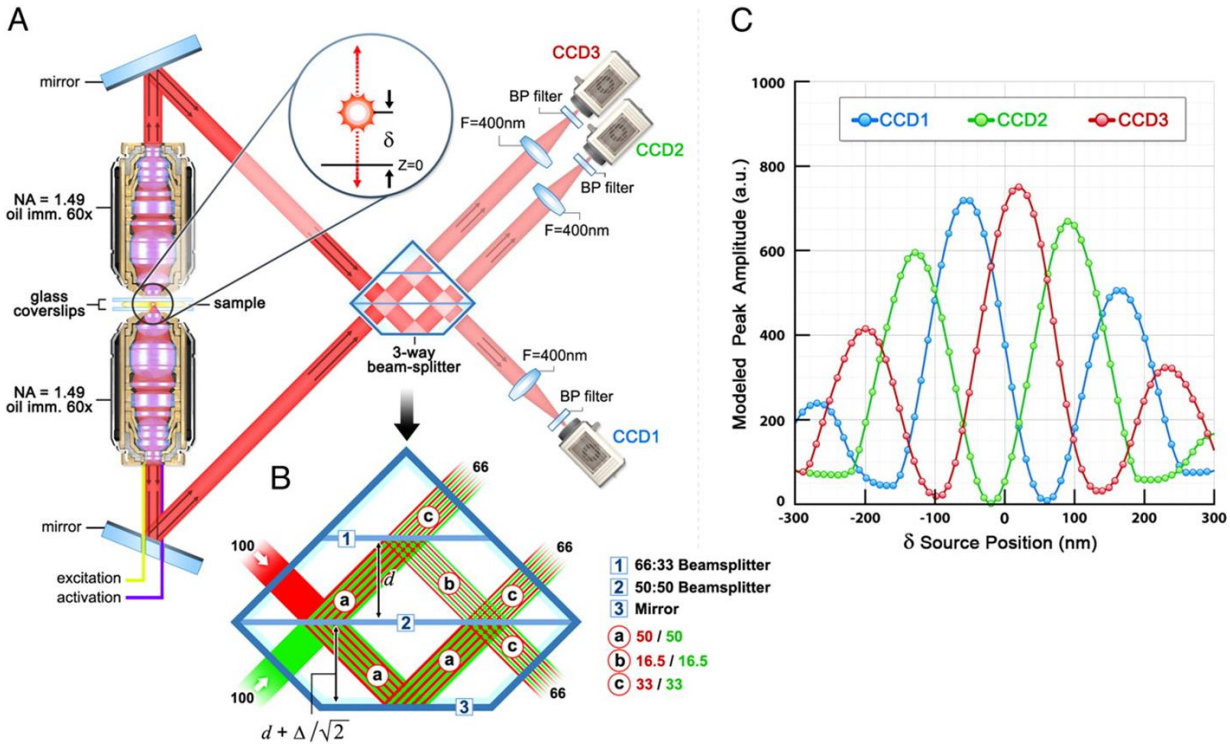


Figure 20. Multiphase interferometric microscope. (**A**) Simplified lightpath of the microscope. (**B**) Scheme of a 3-way beam splitter where the two emission beams interfere. (**C**) The fluorophore axial position is encoded unambiguously into the intensities of its image in the three camera images. Figure adapted from (Shtengel et al., 2009).

Different variations of the double-objective scheme were developed. For example, double-objective detection with an astigmatic PSF proved to achieve ~9 nm/19 nm of lateral/axial LP (Xu et al., 2012). Four-channel detection provided impressing 2.3-3.5 nm/3.5-17.5 nm of lateral/axial LP over up to 1 µm-thick axial layers in two colors (Aquino et al., 2011). Recently, "whole-cell 4Pi single-molecule switching nanoscopy" (W-4PiSMSN) using both the four-channel detection and astigmatism could resolve numerous structures (endoplasmic reticulum, bacteriophages, mitochondria, nuclear pore complexes, primary cilia, Golgi-apparatus-associated COPI vesicles, mouse spermatocyte synaptonemal complexes) with nearly isotropic resolution of 10–20 nm, including for two-color and whole-cell imaging (Huang et al., 2016).

Advantages of the 4Pi geometry over PSF modifications and biplane techniques also include: possibility to collect twice more photons than it is possible using only one objective, which results in √2-fold improvement in LP (9); no widening of PSF and therefore weak dependence of LP on axial position and higher tolerable density of active molecules at a time (Hell et al.,

2009). The obvious disadvantages are complex optical setup and unconventional sample mounting.

1.2.5. Comparison of super-resolution techniques

Many different super-resolution techniques have been developed in the beginning of the 21st century, but now there are three main families of techniques which have found the widest applications in biology: SMLM, STED and SIM. Every technique has its advantages and drawbacks (**Tab. 1**) and thus a particular technique should be chosen depending on the biological question. In terms of resolution, the best results can be obtained with SMLM while linear SIM provides only two-fold improvement over the conventional microscopy. However, to achieve high resolution, lots of data have to be acquired, from where the inverse dependence between the resolution and the experiment duration. In practice, despite the different nominal resolution, images obtained with SMLM and STED may look similarly (Wegel et al., 2016). Surprisingly, very dense and fine structures as the actin network can look better when imaged with SIM (Wegel et al., 2016). Long acquisitions of SMLM make it almost unusable for live imaging, except for studies of molecular trajectories with single particle tracking PALM (sptPALM (Manley et al., 2008)) or universal points accumulation for imaging in nanoscale topography (uPAINT (Giannone et al., 2010)). SIM is obviously best suited for imaging of fast processes with a moderate improvement of resolution. The clear advantages of STED are a direct optical image formation without data processing and a low propensity to artifacts. SMLM, besides best nominal resolution, is also unique because it gives access to the properties of individual molecules.

	SMLM	STED/RESOLFT	SIM/NLSIM
Principle	localization of stochastically switched-on fluorophores	PSF sharpening by depleting its periphery	observation of high frequencies through Moiré interference
Lateral resolution, nm, typical/record	30/10	50/20	120/50
Typical acquisition time	1–20min	0.1-1 min	0.5-2 s
Optical setup	conventional wide-field	confocal with modifications	wide-field with modifications
Data processing required	yes	no	yes
Fluorophore choice	wide but limited	limited to photostable fluorophores	almost not limited
Special mounting medium	necessary in most cases, except PALM	common antifade is advisable	not indispensable
Excitation intensity	strong	strong	moderate
Output data	point data and image	image	image
Access to SM properties	intrinsic	limited	no
Examples of dedicated commercial systems	Nikon N-STORM Leica SR GSD Bruker Vutara 352 ZEISS Elyra DeltaVision OMX	Abberior ExpertLine Leica STED 3X MicroTime 200	DeltaVision OMX ZEISS Elyra Nikon N-SIM

Table 1. Comparison of commonly used super-resolution techniques: SMLM, STED and SIM. The typical resolution and acquisition time are based on performance of commercial systems and not of custom-built setups.

1.3. SMLM: data processing

SMLM experiments substantially differ from conventional fluorescence microscopy. SMLM is the only super-resolution fluorescence microscopy technique, which allows to get some information from every individual fluorescent molecule. It is therefore not surprising that the SMLM data, which have to incorporate SM properties, are different from mere microscopy images. After a single molecule image is depicted into a sensor, its individual properties (at least the coordinates, the spread of the PSF and the detected number of photons) should be determined. After that, the raw acquisition frames are no more needed and all the subsequent processing can be performed on the determined molecule properties. This SM information is nothing else than point data and should be treated as such. Many methods have been developed for processing of microscopy images, but they cannot be directly applied to SMLM data. On the other hand, point data can be found in many scientific fields, and a lot of processing techniques exist for them. However, in the microscopy field it is a relatively rare type of data, and therefore new, specially adapted for SMLM, data processing method had to be developed.

1.3.1. Single-molecule localization

After capturing the SM "blinks" into a camera, the first data processing step is the determination of SM properties from these camera frames. The precision of SM localization is fundamentally limited by photon counting, because a single photon can be detected only in one of pixels with a probability given by the PSF shape. To retrieve the PSF shape precisely, many photons have to be collected, and the origin of the PSF can be determined at best with the localization precision given by formula 8.

The simplest way to determine the SM position is just to take the centroid of the diffraction spot:

$$C_x = \frac{\sum_{i=1}^{n}\sum_{j=1}^{m}(x_i \cdot I_{ij})}{\sum_{i=1}^{n}\sum_{j=1}^{m} I_{ij}} \ , \tag{12}$$

where C_x is the x (y) centroid coordinate, I is an n-by-m matrix (input image), x_i is the coordinate of the pixel on x (y) axis (Cheezum et al., 2001). This is a fast algorithm (Deschout et al., 2014), which does not require a PSF or noise model and is used in many popular software, such as QuickPALM (Henriques et al., 2010) or Leica LAS AF. In case of a non-uniform

background, this estimation will be falsely shifted towards the brighter background region. For a correct estimation, the background correction should be performed beforehand, *e.g.* by subtracting the non-blinking component of image frames (Hoogendoorn et al., 2014). Obviously, the center of mass method fails to correctly detect molecules in high-density conditions, where the PSFs partially overlap (Sage et al., 2015).

Since the lateral PSF can be conveniently approximated with a Gaussian function (11) (**Fig. 21**), a widespread localization method is the least-squares (LS) fitting of this function to the diffraction spots (Betzig et al., 2006). Another possibility is a maximum-likelihood estimation (MLE) of SM parameters (Aguet et al., 2005; Smith et al., 2010). For best results, however, this technique requires a good model for the PSF and the noise distribution (Small and Stahlheber, 2014). MLE was shown to outperform LS fitting, especially at low photon counts (Abraham et al., 2009; Small and Stahlheber, 2014).

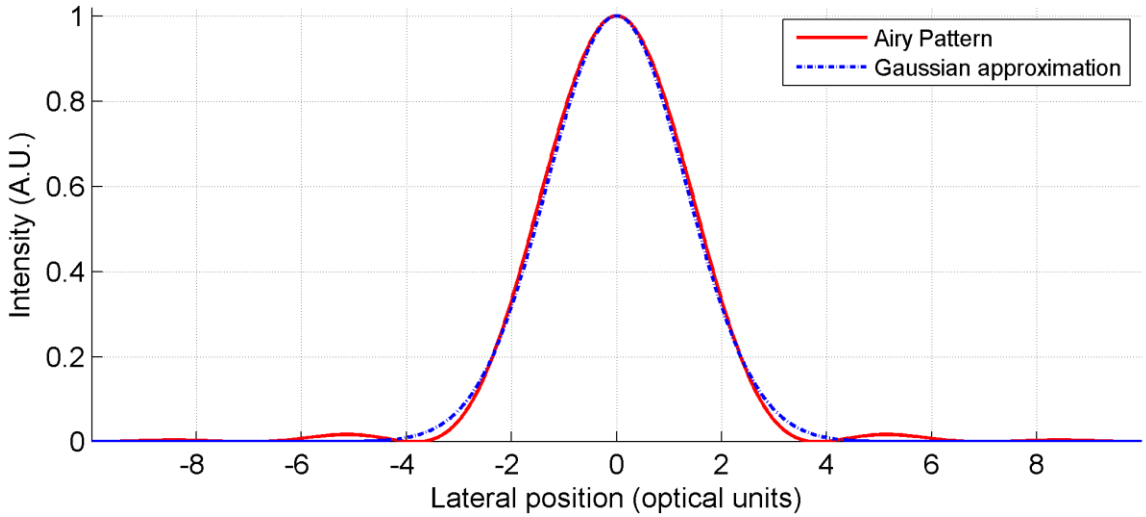


Figure 21. Approximation of the lateral profile of the PSF (the Airy Pattern) with the Gaussian function (10).

1.3.2. Fitting of highly dense fluorophores

The density of activated dyes depends on the properties of the fluorophores and experimental parameters and it cannot be always insured to be less than one per the diffraction-limited surface ($\lesssim$ 1 μm^{-2} (Huang et al., 2011)) at a time. In the case of dense emitters, the simple algorithms of the center of mass or fitting with a single Gaussian would give meaningless positions (Small and Stahlheber, 2014), so such algorithms usually discard spots which cannot be precisely fitted

with a single PSF. However, it would be beneficial to be able to find positions of several fluorophores inside overlapping PSFs, this would reduce the acquisition time for the equal number of localizations and expand the SMLM for poorly blinking fluorophores or densely labeled structures. Such algorithms, based either on LS fitting (Holden et al., 2011) or MLE (Huang et al., 2011, 2013; Quan et al., 2011), have been developed (**Fig. 22**). They allow to increase the density of resolvable molecules up to 10 μm^{-2} (Huang et al., 2011), which is a one order of magnitude improvement over the single-emitter fitting methods.

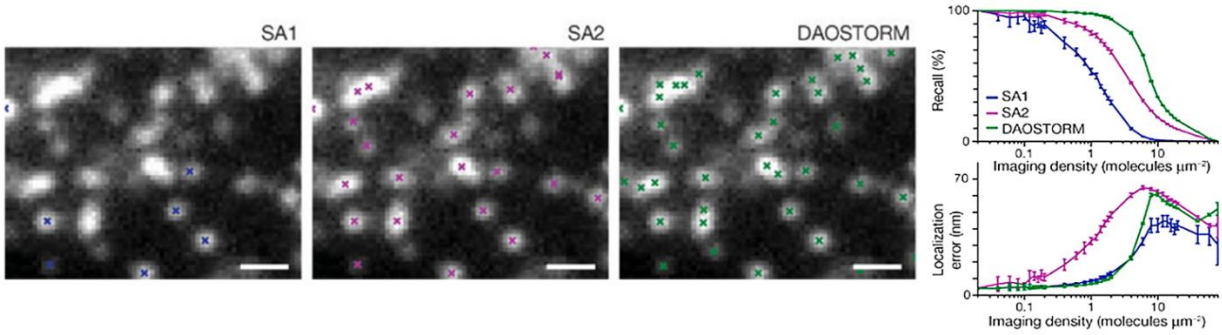


Figure 22. Comparison of the performance of sparse localization algorithms with a high-density fitting algorithm DAOSTORM (Holden et al., 2011). SA1 fits spots with a Gaussian of variable size and ellipticity; if the shape of the PSF is too elliptical, the localization is rejected. SA2 fits spots with a Gaussian of fixed shape and size, without size- or shape- based filtering. DAOSTORM fits spots with a model PSF in an iterative way, including for overlapping spots. The model PSF is generated from low-density images. (right) Recall and localization error of the algorithms on randomly distributed molecules of different densities. The figure is adapted from (Holden et al., 2011).

1.3.3. Effects of molecular orientation on localization accuracy

The light emission by a fluorophore can be represented as the radiation of a simple oscillating electric dipole (Selényi, 1939). As a consequence, the single-molecule fluorescence is polarized and its intensity is not isotropic in the space around the molecule: the intensity is zero along the dipole momentum direction and is maximal in the perpendicular direction (Lakowicz, 2006). When an individual molecule is imaged with a lens, the non-isotropic emission of a tilted dipole illuminates the lens asymmetrically, which leads to a PSF tilted with respect to the optical axis. If the molecule is out of exact focus, the centroid of the PSF does not correspond the SM

position anymore (Backlund et al., 2014; Engelhardt et al., 2011; Stallinga and Rieger, 2010) (**Fig. 23**). It was shown that, *e.g.* at defocuses over a ± 200 nm range where the brightness diminishes to 20% of its maximum, the localization error would be ± 65 nm. For molecules with random orientations, the average error is around 32 nm (Engelhardt et al., 2011). The effect is less pronounced for TIRF microscopy because of a limited axial range (Enderlein et al., 2006; Engelhardt et al., 2011). At the same time, a similar phenomenon practically does not affect STED microscopy, the shift being < 2 nm (Engelhardt et al., 2011). For SMLM, one of possibilities to avoid the inaccuracy of localization is to assure free rotation of fluorophores, meaning that hardening mounting media are poorly compatible with high localization accuracy (Stallinga and Rieger, 2010). Since the SM fluorescence is polarized, it is also possible to determine the dipole orientation by splitting the detection path of a microscope and imaging several different polarization states (Foreman et al., 2008; Stallinga and Rieger, 2012), however, this strongly complicates the optical setup and image analysis.

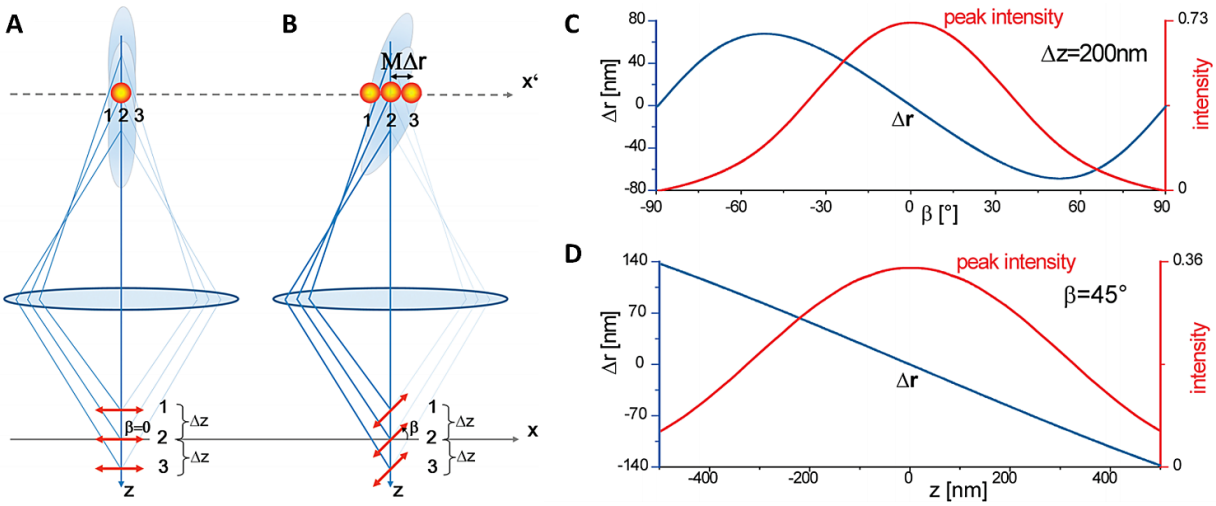


Figure 23. Imaging of individual dipoles with different orientations. (A) Molecules with the dipole moment parallel to the focal plane ($\beta = 0°$) are represented as PSFs parallel to the optical axis z. Defocusing Δz does not affect the centroid position of the PSF. (B) Molecules with the dipole moment tilted on an angle β with respect to the focal plane are imaged as tilted PSFs. (C) The lateral shift of the centroid position Δr and its peak intensity as functions of the tilt angle β, the maximal shift is at $\beta \approx 52°$. (D) The lateral shift of the centroid position Δr and its peak intensity as functions of the axial position *z*. The calculations were made for an oil immersion objective with NA = 1.4, excitation and emission wavelengths of 532 nm and 600 nm, respectively. Adapted from (Engelhardt et al., 2011).

1.3.4. Size of labeling system and localization accuracy

Labeling of proteins of interest with specific dye-stained antibodies (immunofluorescence, IF) is a technique commonly used in biology. The discovery of blinking of common organic dyes conveniently extended IF for high-resolution SMLM studies. However, at the resolution level offered by SMLM the size of the labeling system between the molecule of interest and the fluorophores should be considered. In typical IF protocols a primary site-specific antibody is used with secondary fluorophore-conjugated antibodies. Every antibody is a heavy (~150 kDa) protein with maximal dimensions around 13 nm (Harris et al., 1998). Thus, in the case of indirect immunolabelling, the separation between the antigen and fluorophores will be around 10–20 nm. Also, every secondary antibody usually contains several fluorophores, and several secondary antibodies can bind to the primary one. Consequently, in a super-resolution image a protein molecule will be displayed as a blob of 20–30 nm in diameter, situated 10–20 nm aside from the epitope, which distorts imaging of fine structures (Ries et al., 2012).

FPs are much smaller (< 4 nm (Yang et al., 1996)) and thus they provide a good localization accuracy, but their photon budget is much smaller than that of organic dyes, and FPs do not allow to label wild-type proteins without modifying them (Allen et al., 2013). To conveniently increase the quality of SMLM with FP-targeted proteins, their targeting with fluorophore-coupled nanobodies was proposed (Ries et al., 2012).

Nanobodies are variable fragments of heavy-chain antibodies produced in camelids (Harmsen and Haard, 2007). While in conventional antibodies, the variable fragments of both the heavy (VH) and the light (VL) chains are required for nominal affinity (Sundberg and Mariuzza, 2002; Ward et al., 1989), the antibodies of camelids possess only heavy chains (Hamers-Casterman et al., 1993) and thus their variable fragments (VHH or Nanobody®) are sufficient for target binding. These VHH are small proteins (~15 kDa), around 3 nm in size (Spinelli et al., 1996), and can be efficiently produced in bacteria (Olichon and Surrey, 2007; Rahbarizadeh et al., 2005). Therefore, labeled with fluorophores, they can be used as specific markers for super-resolution applications, *e.g.* for detection of GFP-fused proteins (Ries et al., 2012; Szymborska et al., 2013) or for labeling of native proteins (Mikhaylova et al., 2015; Pleiner et al., 2015). Other possibilities for reduction of the localization inaccuracy include: primary antibodies directly conjugated with fluorophores, fluorescently labeled Fab or $F(ab')_2$ antibody fragments, aptamers (Castro et al., 2017; Opazo et al., 2012), fluorescently labeled phalloidin for actin imaging, photoswitching DNA-intercalating dyes (Flors et al., 2009), genetically encoded

labels, such as SNAP- (Jones et al., 2011; Keppler et al., 2003) and Halo-tags (Barlag et al., 2016; Los et al., 2008).

1.3.5. Drift correction

As noted previously, SMLM is the slowest among the common super-resolution techniques and the typical acquisition time is in the order of minutes (**Tab. 1**). At this scale of time and resolution spatial drift becomes a major issue. The origins of drift are multiple: temperature variations, airflows, vibrations, mechanical instability, sample mounting, etc. Many hardware solutions were proposed for drift reduction. This includes the compensation of the axial drift using a reflection from the coverslip (Nikon Perfect Focus System, Olympus Zero Drift system, Leica Adaptive Focus Control, etc.), a dedicated microscope stage with objective fixed on the stage (Leica SuMo stage), special Petri dishes (Nahidiazar et al., 2016) or an active stabilization of the microscope using a high-precision detection of a fiducial marker position (Carter et al., 2007).

Drift can also be corrected through image processing. In conventional time-lapse microscopy or in cryo-electron microscopy consecutive frames will appear shifted because of drift, and this shift can be detected via cross-correlation (Guizar-Sicairos et al., 2008) and then subtracted from the raw frames in order to have a drift-free image or movie (Wade and Fitzke, 1998). Frames of SMLM do not directly provide this possibility, because the fluorophores are activated at random positions so the drift detection by cross-correlation is not possible. In order to have fiducial marks for cross-correlation calculation, fluorescent polystyrene (Rust et al., 2006) or gold beads (Betzig et al., 2006) can be introduced inside the sample. In most cases, instead of finding the shift between the raw images, it is more efficient to localize the beads at the same time as the blinking fluorophores and then subtract time-smoothed bead trajectories from the SM coordinates (Bates et al., 2007).

Alternatively, the sample can be imaged with bright field transmission microscopy during the SMLM acquisition and the shift between the bright field images can be calculated, which can be used either for the real-time stage repositioning or for drift subtraction from the SM coordinates in the post-processing (Tang et al., 2014). This method requires modifications of the optical system, but provides sub-nanometer precise drift correction without introducing beads in the sample (Tang et al., 2014).

Even though individual frames contain too few information for reliable shift estimation, their subsets consisting of several tens or hundreds of images contain enough information to partially reconstruct an image. The images reconstructed from different subsets are similar except the high-resolution details, and this similarity can be sufficient for a robust determination of the shift between them. Therefore, a drift correction procedure specific for SMLM consists in: 1) division of the whole acquisition localization list into consecutive subsets; 2) reconstruction of an SR image from each subset; 3) determination of the shift between the images with sub-pixel precision using cross-correlation; 4) interpolation of the shifts to the whole acquisition; 5) subtraction of the interpolated drift from the coordinates of every localization, taking into account their time of appearance (frame number). This method was used since first SMLM experiments (Huang et al., 2008a, 2008b) and was shown to be precise to a sub-5 nm level for suitable structures (Geisler et al., 2012; Mlodzianoski et al., 2011). It uses present in any SMLM dataset information and does require neither fiducial markers in the sample nor modifications of the optics. See also section 3.1.

1.3.6. Correction of chromatic aberrations

The refractive index of any medium, including the optical glass, depends on the wavelength of light. This phenomenon is known in optics as dispersion. As the consequence, the optical parameters of the lenses become dependent on the wavelength as well. This leads to different focal lengths and magnifications for different colors and to misalignment of structures labeled with different fluorophores and imaged through the microscope. Such misalignments are termed chromatic aberrations, which can be divided into axial chromatic aberrations (different focus positions for different wavelengths) and lateral chromatic aberrations (lateral shift between images acquired with different wavelengths, the shift typically increases from the center to the corner of the image). Such aberrations can be partially corrected using combinations of multiple lenses made of different sorts of glass, but even the best objectives on the market, called plan-apochromats, have some residual chromatic aberrations, up to several hundred nanometers in both axial and lateral directions (Erdelyi et al., 2013; Juškaitis and Wilson, 1999; Kozubek and Matula, 2000). The shift is often under the diffraction limit and is hardly seen in conventional microscopy, but in SMLM the aberrations are obvious (**Fig. 24**).

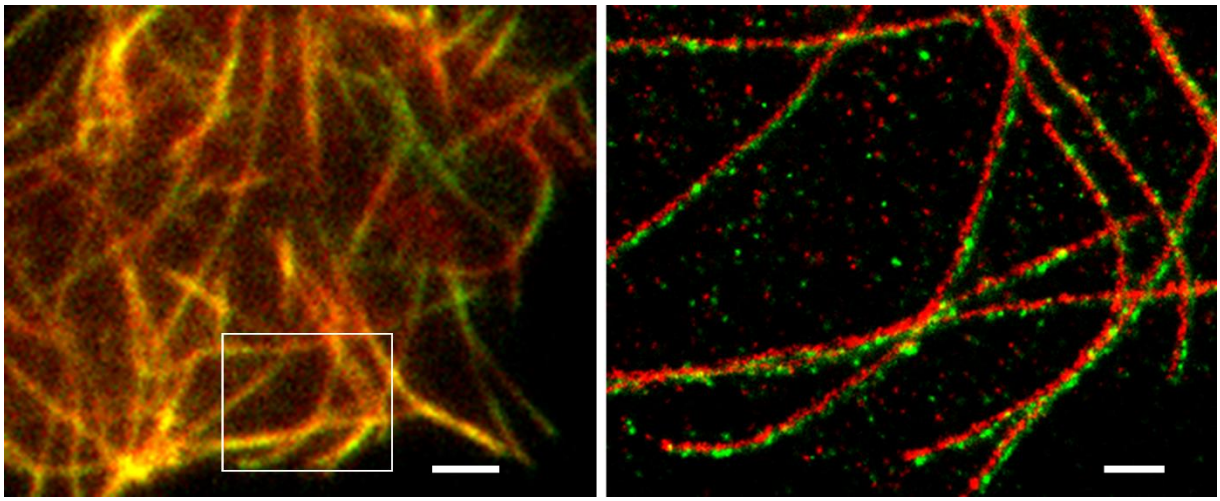


Figure 24. Lateral chromatic aberrations in SMLM. (left) Wide-field image of β-tubulin co-labeled with Alexa 647- (red) and Alexa 488- (green) conjugated secondary antibodies in a HeLa cell. Both labelings overlay when imaged with the diffraction-limited resolution. (right) SMLM image of the same cell zoomed in the white rectangle. The images of the two labelings do not overlay at high resolution. Note that the shift is visible when its direction is perpendicular to the microtubules. The SR dataset was corrected for drift, but not for chromatic aberrations. Scale bars: 2 µm (left) and 500 nm (right). Data was acquired on the in-house Leica SR GSD system and processed with LAS AF and SharpViSu (Andronov et al., 2016a).

In early works on multicolor SMLM the chromatic aberrations were corrected by simple shift of reconstructed images (Bock et al., 2007; Linde et al., 2009); however it quickly became evident that this phenomenon needs more precise corrections because the lateral shift between images of different colors is not constant in the FOV (Erdelyi et al., 2013). It was shown that the lateral chromatic aberrations of an objective can be calibrated by imaging multicolor fluorescent beads and the shift between the channels can be efficiently approximated with a second-order polynomial in order to determine the displacement at an arbitrary point of the FOV (Erdelyi et al., 2013). After subtracting this fit from the SM coordinates, the residual shift between channels dropped from ~45 nm for the uncorrected data to ~16 nm (Erdelyi et al., 2013; Erdélyi et al., 2015) – see section 3.1 – which is smaller or in the order of the value of the other errors in SMLM and thus this correction can be sufficient for most applications. This method successfully found its application for SMLM-based colocalization studies (Andronov et al., 2016a; Grove et al., 2014).

1.3.7. Super-resolution image reconstruction

Unlike in other microscopy methods, in SMLM the final image is not obtained in a direct optoelectronic way. To get the SR image, the SM coordinates have to be somehow translated in a raster image, *i.e.* in a matrix consisting of picture elements – pixels. This is not an unambiguous procedure and several different methods have been used for SR data rendering (**Fig. 25**).

If the density of molecules is low, their positions can be simply indicated by markers in the coordinate space (Rust et al., 2006). The result can be exported in a vector or raster graphic format. This method is used for indicating of positions of sparse molecules and it is not suitable for dense data because of overlap of the markers (Baddeley et al., 2010). For dense data it is more efficient to translate the localizations into an image where the pixel gray values would reflect the local density of fluorophores like in conventional fluorescence microscopy. One of such possibilities is to build a 2D histogram on the localizations and use its bins as pixels and the counts inside the bins as the pixels' gray values (Bossi et al., 2008). For low-dense data this method produces very noisy image and thus requires increasing the pixel size.

Another widely used possibility is to depict every localization as a 2D Gaussian kernel, which width would be equivalent to the localization precision of the given molecule, calculated from the number of the detected photons (Bates et al., 2007; Betzig et al., 2006). This approach can be useful if one molecule activates very few times, then the Gaussian-blurred image can give an idea about the real position of the molecule. In the case of multiple "blinks" of the same fluorophore this Gaussian blur will additionally broaden the distribution of the localizations, which will lead to deterioration of the image resolution up to $\sqrt{2}$ times (Baddeley et al., 2010; Coltharp et al., 2014).

The problem with the histogram visualization is that it samples data with a constant step (the bin or the pixel size). As an improvement to this method it was proposed to adapt the histogram bin size to the local density of the molecules and recalculate the local density (used as the gray values of pixels) as N/S for every bin, where N is the number of detected molecules within the bin's area and S is the surface area of the bin (Baddeley et al., 2010).

Also, it was proposed to use a Delaunay triangulation built on the localization coordinates (Baddeley et al., 2010). In this case the surface is tiled with triangles and the brightness of the triangles is inversely proportional to their area. This is a method similar to the adapted histogram, the difference is mostly in the shape of the sampling: square in the histogram and

triangles in the triangulation. Since an image composed of triangles looks unnatural, the images were additionally smoothened, either by randomly displacing the points around their positions or by averaging images built from different subsets of data (Baddeley et al., 2010). The triangulation-based methods demonstrated best results in terms of preservation of the image resolution, SNR and image continuity, but they are more computationally demanding and include an ambiguous smoothening step.

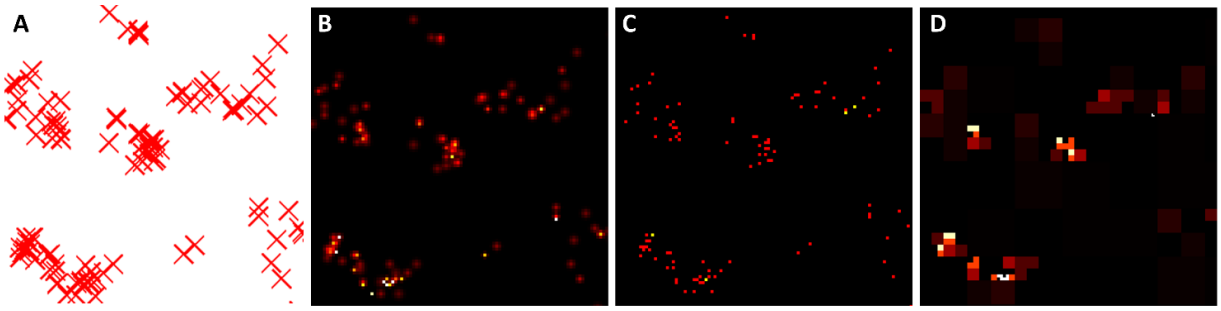


Figure 25. Different methods for SMLM data visualization. (**A**) Scatter plot. (**B**) Visualization of each molecule as a Gaussian function. (**C**) 2D histogram. (**D**) Adaptive histogram with maximal bin capacity of 2 localizations. Data: TPR protein in a U2OS cell, labeled with Alexa 647 conjugated secondary antibodies (Lemaître et al., 2014). 1 µm x 1 µm FOV.

1.3.8. Estimation of real resolution

While in conventional microscopy the resolution of a microscope is commonly defined by Rayleigh's criterion (**Fig. 1**) and by Abbe's formula (1), in SMLM the definition of resolution is less evident. On the one hand, the precision to determine the molecule position is defined at least by the PSF and the number of detected photons and this precision should obviously affect the resolution. On the other hand, in order to resolve details in the sample, the positions of a sufficient number of fluorophores should be determined, *i.e.* in order to resolve two features spaced by a distance *x*, at least two molecules over this distance should be localized. Also, according to the Nyquist-Shannon sampling theorem, to resolve a certain distance *x*, fluorescent labels should be available at least each *x/2*. Because during a SMLM acquisition the molecules are localized gradually, the resolution should also improve gradually. However, if there is a non-corrected sample drift, the resolution will on the contrary deteriorate with increasing acquisition time. Therefore, the SMLM resolution depends on both localization precision and density of localized fluorophores (Betzig et al., 2006; Cordes et al., 2010; Fitzgerald et al.,

2012; Shroff et al., 2008). The exact resolution value depends on many parameters (nature of fluorophores, mounting medium, concentration of labels, laser intensity, mechanical stability of the system, algorithms of data processing, etc.) and is difficult to estimate *a priori*. This is why the actual resolution is usually measured on final data, after all the processing.

A widespread way to estimate the resolution was the measurement of the size of structures in the SR image and its comparison with their real size. For this purpose tubulin fibers, non-specifically bound antibodies, DNA origami with defined separation between several fluorophores and other objects can be used (Cordes et al., 2010; Mikhaylova et al., 2015; Schmied et al., 2014; Szymborska et al., 2013). The obvious disadvantage of this method is the necessity for objects with known dimensions, which can be rarely satisfied because the size of the imaged structures is usually not known, it has to be determined from the SR experiment.

1.3.8.1. Fourier ring correlation

In order to have an unambiguous resolution criterion which does not require *a priori* known structures, it was proposed to employ the Fourier ring correlation (FRC) method (Banterle et al., 2013; van Heel et al., 1982; Nieuwenhuizen et al., 2013; Saxton and Baumeister, 1982). The method operates on two independently built images of the same object, which can be reconstructed from two randomly selected subsets of SM localizations. It calculates the correlation coefficient between the two images at different frequencies, which in Fourier space is a correlation between pixels along two rings (Saxton and Baumeister, 1982), giving rise to the term FRC:

$$FRC(R) = \frac{\sum_{\vec{r}\in R} \widehat{f_1}(\vec{r})\widehat{f_2}(\vec{r})^*}{\sqrt{\sum_{\vec{r}\in R} |\widehat{f_1}(\vec{r})|^2}\sqrt{\sum_{\vec{r}\in R} |\widehat{f_2}(\vec{r})|^2}}, \quad (13)$$

where *FRC(R)* is the FRC value at the spatial frequency *R*, $\widehat{f_1}(\vec{r})$ and $\widehat{f_2}(\vec{r})$ are the Fourier transforms of the two images. At low frequencies, since the two images are similar, the correlation is high, and it decreases while going to higher frequencies and sampling fine details and noise, which are less similar in the images (**Fig. 26**). The inverse value of the resolution is defined as is the frequency at which the FRC curve drops below the threshold 1/7 (0.143) (Nieuwenhuizen et al., 2013) or where the FRC curve intersects with the 2σ curve, determined as: $F_{2\sigma}(R) = \frac{2}{\sqrt{N_p(R)/2}}$, where $N_p(R)$ is the number of pixels in the ring R (Banterle et al., 2013).

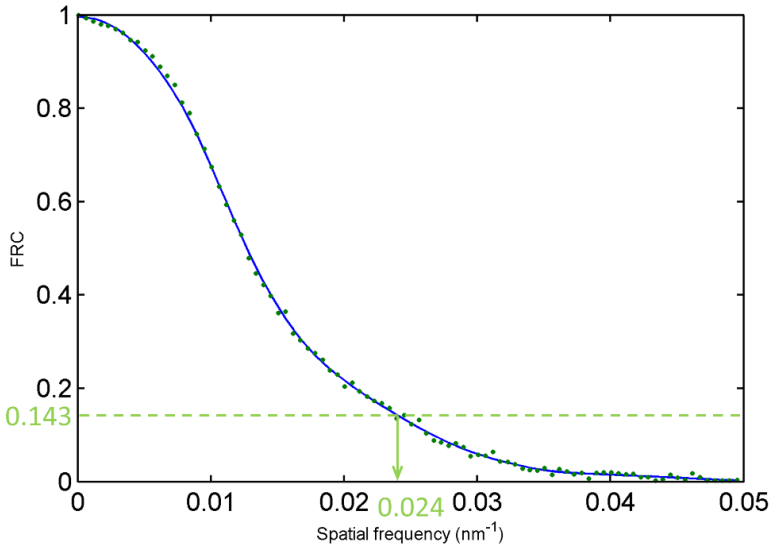


Figure 26. FRC curve for the dataset shown in **Fig. 15**. Calculated in SharpViSu (Andronov et al., 2016a), after merging consecutive localizations. The cut-off frequency by the 1/7th criterion is 0.024 nm^{-1} that corresponds to the resolution of 42 nm. The maximal frequency in the graph is 0.05 nm^{-1}, which is the Nyquist frequency for the pixel size of 10 nm.

Since a single fluorophore may appear on several consecutive frames, this may create artificially high correlation between datasets from randomly selected localizations and lead to an overestimation of the resolution. To account for this, the localizations of the same molecule in the consecutive frames should be combined. Alternatively, the dataset can be first split into many consecutive blocks of localizations and then these blocks can be split randomly into two subsets (Nieuwenhuizen et al., 2013). Indeed, the FRC criterion was shown to be sensitive to parameters that affect the resolution of SMLM data, *e.g.* to the localization precision, the labeling density, the drift and the data processing (Banterle et al., 2013; Nieuwenhuizen et al., 2013). Also, it can be applied for both 2D and 3D data (Nieuwenhuizen et al., 2013). In the 3D case this method is named Fourier shell correlation (FSC) and is employed routinely for the estimation of resolution in the cryo-EM field (Liao and Frank, 2010).

1.3.9. Cluster analysis of SMLM data

Cluster analysis in general means grouping of objects into sets according to similarity in their properties. In classical bioimaging this task was usually performed by segmentation of micrographs, in order to detect *e.g.* cells or organelles and automatize studies of their shapes or spatial distributions. With the invention of super-resolution microscopy novel applications for cluster analysis appeared, such as studies of spatial-temporal distributions of individual proteins and complexes. Because for the coordinate-targeted super-resolution methods (STED, SIM) the primary data type is the image, cluster analysis of such data can still rely on common image segmentation techniques used for wide-field and confocal imaging. In SMLM, the image is secondary after the localization data, and in order to get best results SMLM data processing should use the raw point data instead of reconstructed images. The localization data format was relatively rare in the microscopy field and as the consequence, little or no comprehensive processing methods, such as cluster or colocalization analysis were available until recently. Similar data types, however, are common in such fields, as geography, astronomy, botany, epidemiology, ecology, computational geometry, etc. Adaptation of methods from these fields as well as developments of new processing techniques dedicated for bioimage data would therefore help to bring new insights into biology at the level of unprecedented resolution offered by SMLM.

1.3.9.1. Nearest-neighbor analysis (Clark-Evans statistics)

This method aims at characterizing the deviation of the actual dataset from the random distribution of points (Clark and Evans, 1954). It measures for each data point the distance *r* to its nearest neighbor. The mean observed density for a population of *N* points is given by $\overline{r_A} = \frac{\sum r}{N}$. For an infinitely large completely random distribution of points with the mean density of ρ this value was shown to be $\overline{r_E} = \frac{1}{2\sqrt{\rho}}$. The ratio is $R = \frac{\overline{r_A}}{\overline{r_E}}$ is introduced in order to measure the degree of deviation of the observed dataset from the random distribution. For randomly distributed points this ratio *R = 1*, for the maximally aggregated (clustered) dataset *R = 0*, and in the case of the maximal spacing (the molecules are distributed on a hexagonal pattern) it was shown that *R* equals *2.1491*. Thus, *R* takes values from *0* to *2.1491* and reflects the overall geometrical properties of the distribution. To test if the experimental distribution is statistically different from the random distribution, a z-test can be performed with the z-score given by $c =$

$\frac{\bar{r}_A - \bar{r}_E}{\sigma_{\bar{r}_E}}$, where $\sigma_{\bar{r}_E} = \frac{0.26136}{\sqrt{N\rho}}$ is the standard error of the mean distance to the nearest neighbor in a randomly distributed set of N points (Clark and Evans, 1954).

This method was mostly used for studies of distributions of trees in forests (Clark and Evans, 1954) or of animal behavior (Boettcher et al., 1994; Leising and Yen, 1997), but its usage for SMLM experiments is yet to be demonstrated in practice (Nicovich et al., 2017).

1.3.9.2. Ripley's K and L functions

Ripley's K function (Ripley, 1976) for a set of *n* points can be defined as:

$$K(r) = \lambda^{-1} \sum_{i \neq j} \frac{I(d_{ij} < r)}{n} \tag{14}$$

where λ is the average density of points, d_{ij} is the distance between the ith and the jth points and *I* is the indicator function (*1* if the argument is true and *0* otherwise). In other words, *K(r)* is proportional to the average number of localizations within the distance r around a randomly chosen point. Unlike the nearest neighbor analysis, this function describes properties of a dataset at many distance scales (Dixon, 2001).

For many particular distributions the analytical expression of *K(r)* is known. For example, for the complete spatial randomness (a homogeneous Poisson process) the expression is very simple, $K(r) = \pi r^2$ (Dixon, 2001). For the evaluation of experimental data, *K(r)* can be estimated directly by (14), but because the experimental ROI is always limited, the K-function will be underestimated for the points close to the borders with increasing *r*. To correct for this, a weighting scheme can be used (Ripley, 1976; Tilke, 1992).

To evaluate an experimental dataset, the experimental *K(r)* has to be compared with *K(r)* for a hypothetical distribution of points, taking into accounts its deviation around the mean value due to fluctuations. In practice, however, for evaluation of experimental data it is more convenient to use the L-function (Ripley, 1979):

$$L(r) = \sqrt{K(r)/\pi} \tag{15}$$

because for a homogeneous distribution of points $L(r) = r$ and the variance of *L(r)* is approximately constant. It is therefore easy to test a dataset for randomness by evaluating the deviation from zero of the function $H(r) = L(r) - r$, or alternatively, by evaluating just one

value: $L_m = sup_r|L(r) - r|$ or $L_s = \sum_r|L(r) - r|$ (Ripley, 1979). Critical values for these estimators can be found from Monte-Carlo simulations (Besag and Diggle, 1977) or calculated analytically (Dixon, 2001).

Ripley's *K* and *L* functions are being successfully used for analysis of SMLM data (Nicovich et al., 2017) including in the developments I have done in this work. Many experiments used the following procedure (Larghi et al., 2013; Owen et al., 2010; Rossy et al., 2013; Wang et al., 2014a; Williamson et al., 2011) (**Fig. 27–28**). First, the function *H(r)* for the experimental data is plotted and compared with the confidence envelope for *H(r)* for a random distribution of points. The confidence envelope is often calculated from Monte-Carlo simulations (see also section 3.2) by randomly distributing the same number of points as was in the experimental dataset over the experimental surface. This automatically takes into account edge effects and their correction is no more needed. The dataset is said to be clustered (dispersed) at some distances *r* when *H(r)* has values over (under) the confidence envelope at these *r* (Kiskowski et al., 2009). If the dataset is confirmed to exhibit clustering, to estimate the local density, the *L(r)* value is calculated for every point at some fixed distance (*e.g.*, *r* = 50 nm (Larghi et al., 2013; Rossy et al., 2013; Williamson et al., 2011)). These densities are interpolated to a regular grid in order to create a "density map". The density map is segmented (binarized) at some fixed threshold (*e.g.*, L(50 nm) = 78 (Rossy et al., 2013; Williamson et al., 2011) or L(50 nm) = 90 (Larghi et al., 2013)) in order to have clusters with defined shape. These clusters are then statistically evaluated for their size, shape, density of molecules, etc., and these values are compared for different biologically relevant situations.

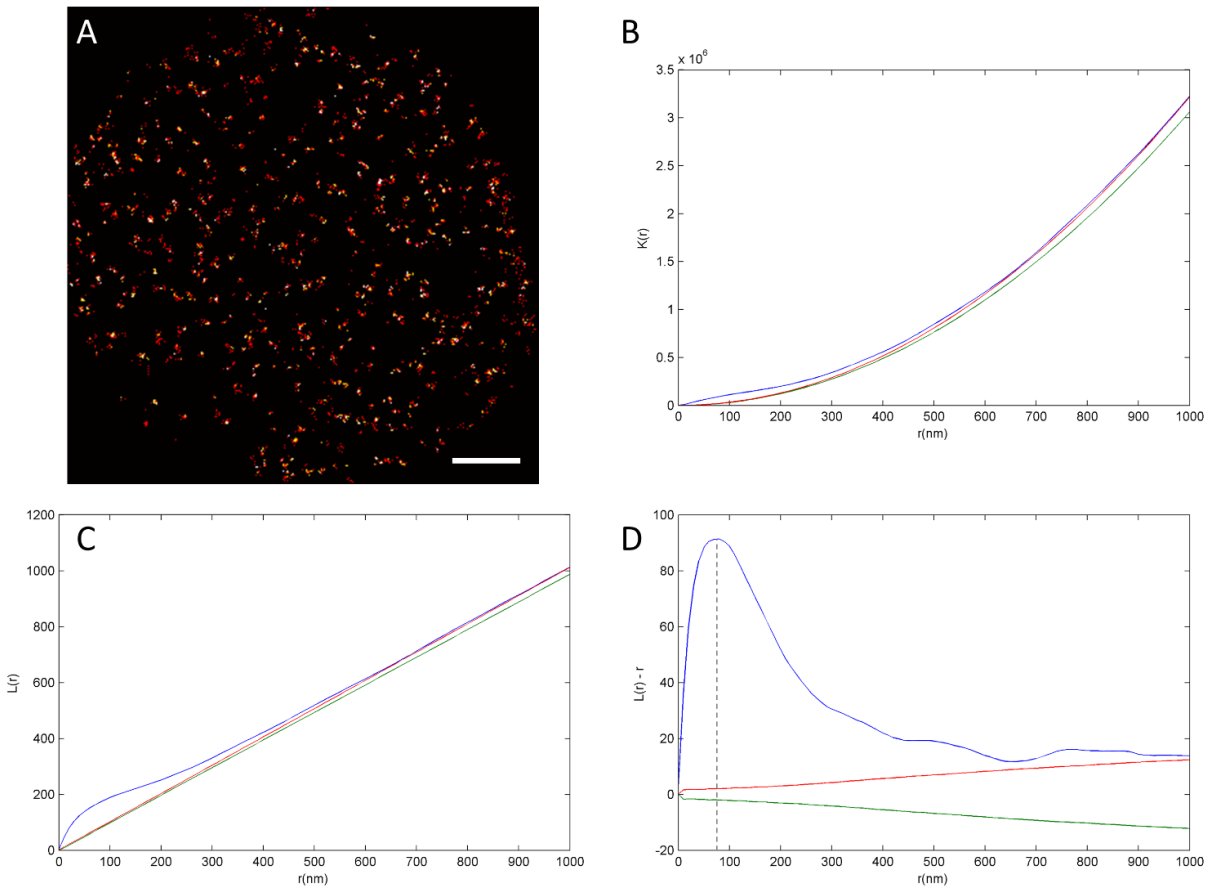


Figure 27. Example of Ripley's statistics for an SMLM dataset. (**A**) SR image of the TPR protein detected at the nuclear envelope of a U2OS cell (an about circular ROI is shown). (**B**) *K(r)* function for the experimental data (blue curve); the 99% confidence envelope obtained from 100 iterations of Monte-Carlo simulations for a complete spatially random distribution of the same number of points over the same experimental ROI (the region between the red and the green curves). (**C**) *L(r)* function for the same data. (**D**) *H(r)* function for the same data with a peak at $r \approx 75$ *nm*, which corresponds to the average size of clusters in the dataset. Scale bar, 1 µm. The data were acquired on the in-house Leica SR GSD system and were processed in ClusterViSu (Andronov et al., 2016b) for the Ripley's statistics. The curves are calculated with a 10 nm step in the search radius *r*.

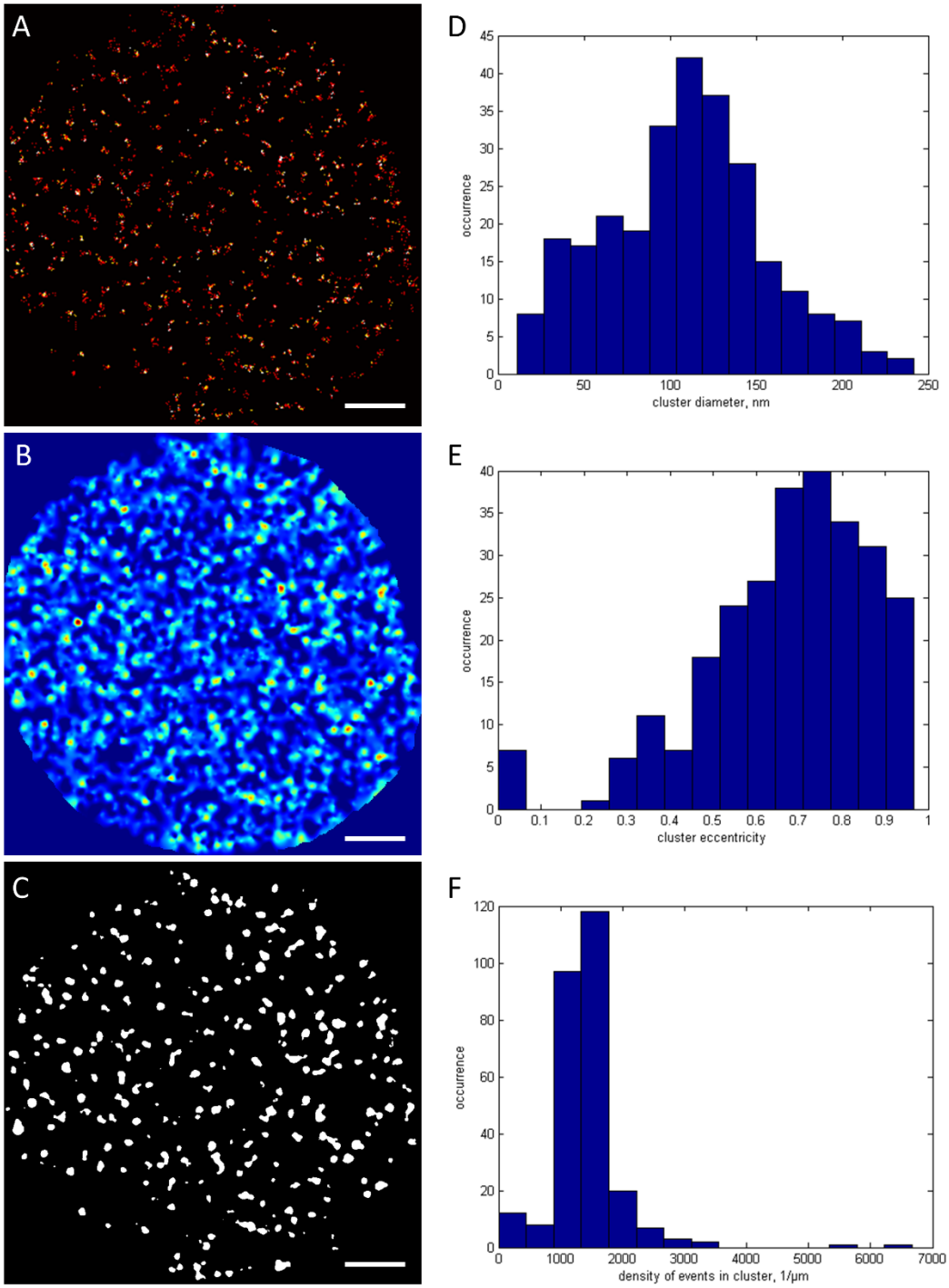


Figure 28. Detailed cluster analysis using the Ripley's statistics. (**A**) The ROI of the analyzed dataset, same as in **Fig. 27**. (**B**) A density map calculated as the *L(r = 50 nm)* function for every localization and interpolated over the FOV. (**C**) Binarized density map at the level of *L(50) = 78* allows to analyze individual clusters. (**D-F**) The equivalent diameter (**D**), the eccentricity (**E**) and the density of localizations (**F**) of the clusters found in the binarized image. Scale bars, 1 µm. The data were processed in ClusterViSu (Andronov et al., 2016b).

Even though Ripley's statistic is widely used for the analysis of SMLM data, it has some shortcomings: the Ripley's K- and L-functions reflect only global geometric behavior of the dataset without providing a direct access for detailed properties of individual clusters. The modifications of the method for local density estimation and for segmentation include some parameters (the radius for the local density estimation, the segmentation threshold) that cannot be set automatically and require a manual adjustment.

1.3.9.3. Pair correlation functions

As noted previously, one fluorophore can be activated several times before bleaching. As a consequence, in SMLM it will be detected every time as a new individual molecule. Because of localization imprecision, all localizations, originated from a single molecule, will be distributed with a normal distribution around the true molecule position. This will produce a cluster of a size, determined by formulas (8) and (10). Additionally, in the case of immunolabeling, every targeted protein will be depicted as a cluster of localizations with the size, determined by both the size of the antibodies and the localization precision. Because the Ripley's functions are integrative in distance, such an artificial clustering will propagate to a longer distance scale that may lead to a false positive clustering test and to false values for the size of the real clusters (Veatch et al., 2012).

A radial distribution function, or pair correlation function *g(r)* describes the probability of finding a particle at a distance r away from a given particle, relative to the probability for randomly distributed particles. Therefore, for a completely randomly distributed set of points $g(r) \approx 1$. The radial distribution function can also be represented as the derivative of the K-function (Stoyan and Stoyan, 1994):

$$g(r) = \frac{1}{2\pi r}\frac{d}{dr}K(r) \tag{16}$$

If the K-function can be seen as the normalized number of points inside a circle of radius *r*, the *g(r)* function can be seen as the normalized number of points on a ring of radius *r*, centered on an arbitrary point. Consequently, the pair correlation function allows to exclude the impact of typical for SMLM spurious short-range clustering and analyze data in a more reliable way (Sengupta and Lippincott-Schwartz, 2012; Sengupta et al., 2011, 2013; Veatch et al., 2012).

1.3.9.4. DBSCAN

The previously described methods provide information about global statistical behavior of a dataset averaging the properties of individual objects. This is useful when the dataset has clusters of a homogeneous size and density, but in the case of heterogeneous objects in the FOV those methods would not provide a meaningful result. Density-based clustering algorithms identify individual clusters in a dataset, based on a local density of points. The points, assigned to clusters, are therefore inside denser regions, separated by sparse regions that are assigned to background or outliers (Kriegel et al., 2011). Since this notion is in line with the common definition of clusters in microscopy, density-based algorithms have found their applications in SMLM. The most popular representative of such algorithms is Density-Based Spatial Clustering of Applications with Noise (DBSCAN) (Ester et al., 1996).

The algorithm classifies all the data points into core points, non-core points and outliers (noise) using two manually defined parameters, *minPts* and ε. A point p is a core point if at least *minPts* points, including p, are within the distance ε of it. Those neighbors are said to be directly reachable from p. Mutually reachable points form clusters. All points not reachable from any other point are outliers. Points, which are reachable from a core point, but do not have at least *minPts* neighbors within the distance ε, are situated at the borders of clusters and are called non-core points.

Even though DBSCAN is popular for SMLM cluster analysis (Endesfelder et al., 2013; Marjon et al., 2016; Pageon et al., 2016; Chen et al., 2016; Sinnen et al., 2017), its evident shortcoming is the requirement for two input parameters, which are not directly obtained from the data and need adjustment and thus can bias the results.

1.3.10. Colocalization analysis

Colocalization analysis has been an important method in fluorescence microscopy since its invention because of the possibility to label different proteins or organelles with spectrally different labels and then analyze their relative position in the cell. Due to the resolution limit, for conventional microscopy the colocalization analysis was commonly applied, *e.g.* for determination with which organelle a particular protein is associated or weather several proteins are colocalized with the same structure (Dunn et al., 2011). With super-resolution microscopy

such studies reach a new horizon – colocalization at the level of individual protein complexes (Szymborska et al., 2013).

1.3.10.1. Image-based analysis

Many studies tackled colocalization in super-resolution microscopy by using methods from the conventional fluorescence microscopy field. This includes a reconstruction of SR images for every color channel with following calculation of a colocalization value between the two images (Bielopolski et al., 2014; Chen et al., 2016; Zhao et al., 2013). The image-based colocalization value is commonly estimated by following criteria:

- Pearson's correlation coefficient (Manders et al., 1992; Pearson, 1896):

$$r_p = \frac{\sum_i (R_i - R_{aver}) \cdot (G_i - G_{aver})}{\sqrt{\sum_i (R_i - R_{aver})^2 \cdot \sum_i (G_i - G_{aver})^2}}, \tag{17}$$

where R_i and G_i are the gray values of the i-th pixel of the red and the green components of the double-color image, R_{aver} and G_{aver} are the mean values of the red and the green pixels.

- Manders overlap coefficient (Manders et al., 1993):

$$r_m = \frac{\sum_i R_i \cdot G_i}{\sqrt{\sum_i (R_i)^2 \cdot \sum_i (G_i)^2}} \tag{18}$$

- Manders colocalization coefficients (Manders et al., 1993):

$$M_1 = \frac{\sum_i R_{i,coloc}}{\sum_i R_i}, \qquad M_2 = \frac{\sum_i G_{i,coloc}}{\sum_i G_i}, \tag{19}$$

where $R_{i,coloc} = R_i$ if $G_i > 0$ and $R_{i,coloc} = 0$ if $G_i = 0$; $G_{i,coloc} = G_i$ if $R_i > 0$ and $G_{i,coloc} = 0$ if $G_i = 0$; M_1 and M_2 should be calculated after prior segmentation of both images (Dunn et al., 2011).

Since the image itself is not the primary data of SMLM and image rendering from point data is an ambiguous procedure, the image-based colocalization analysis is conceptually not the best way for getting optimal results.

1.3.10.2. Coordinate-based analysis

As noted previously, the conceptually optimal way for processing of SMLM data is working directly with SM coordinates. A coordinate-based method for colocalization analysis (CBC) of such data has been developed (Malkusch et al., 2012). The proposed concept estimates the colocalization value for each localization of each of the two color channels as a correlation between a univariate and a bivariate (cross-type) function, similar to Ripley's K(r) function. The univariate function is defined as $D_{A_i,A}(r) = \frac{N_{A_i,A}(r)}{N_{A_i,A}(R_{max})} \cdot \frac{R_{max}^2}{r^2}$ and the bivariate one as $D_{A_i,B}(r) = \frac{N_{A_i,B}(r)}{N_{A_i,B}(R_{max})} \cdot \frac{R_{max}^2}{r^2}$, where $N_{A_i,A}(r)$ is the number of localizations of species *A* within the distance *r* around the molecule A_i, $N_{A_i,B}(r)$ is the number of localizations of species *B* within the distance *r* around the molecule A_i and R_{max} is the maximal search distance for probing the correlation (Georgieva et al., 2016; Malkusch et al., 2012). In these functions the number of localizations is normalized in a way that a uniform distribution would give an expected value *D(r) = 1* for all *r*.

The colocalization value is then calculated as Spearman's rank correlation coefficient between $D_{A_i,A}$ and $D_{A_i,B}$, multiplied by a factor that reflects the distance to the nearest neighbor of the second species:

$C_{A_i} = S_{A_i} \cdot e^{(-\frac{E_{A_i,B}}{R_{max}})}$, where S_{A_i} is the Spearman's rank correlation coefficient between $D_{A_i,A}$ and $D_{A_i,B}$, and $E_{A_i,B}$ is the distance from A_i to the nearest neighbor from species *B*. C_{A_i} is therefore calculated for every localization *i* and can adopt values from -1 to 1. Similarly to the Pearson's correlation coefficient, positive values of C_{A_i} signify co-localization and negative ones stand for anti-localization or mutual exclusion. C_{A_i} also adopts negative values for the case when the localizations of the two species are in a close proximity but do not overlap (Georgieva et al., 2016).

A somewhat similar method with the bivariate *L(r)* function was also reported (Rossy et al., 2014). CBC analysis methods have already been used for studies of DNA repair pathways (Lemaître et al., 2014), nuclear factors (Georgieva et al., 2016), actin-binding proteins (Malkusch et al., 2012), arrestin proteins (Díez et al., 2014) and Fc receptors (Lopes et al., 2017).

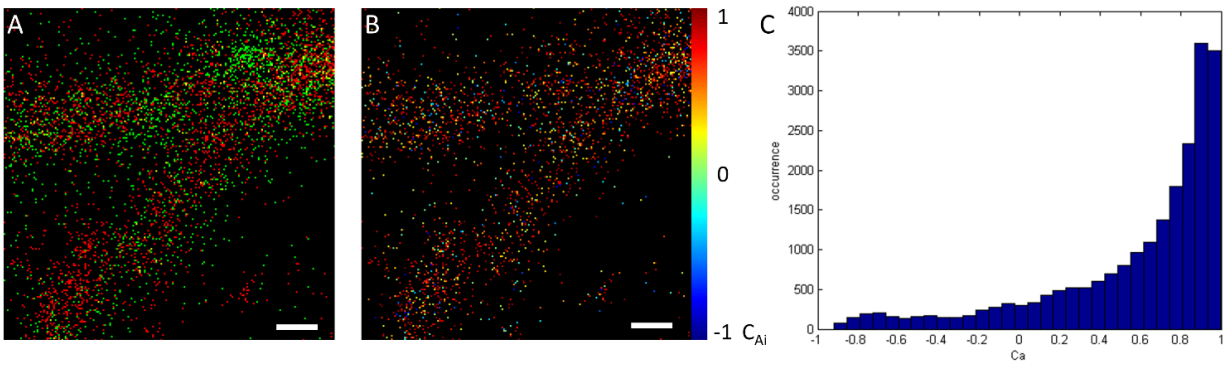


Figure 29. Coordinate-based colocalization analysis of SMLM data. (**A**) SMLM image of β-tubulin co-labeled with Alexa 647- (red) and Alexa 488- (green) conjugated secondary antibodies in a HeLa cell. (**B**) SR image displays the color-coded C_{Ai} values for the red channel localizations. (**C**) Histogram of C_{Ai} values indicates strong co-localization as expected for a co-labeled sample. Scale bars, 50 nm. The C_{Ai} values are calculated with the step of 3 nm up to $R_{max} = 300\ nm$. Bin size of the images, 2 nm. Data collected on the in-house Leica SR GSD system and processed using my implementation (Lemaître et al., 2014) of the CBC method (Malkusch et al., 2012).

1.4. Chromatin, histones and centromeres

In eukaryotic cells, DNA is organized in the form of chromatin. Besides DNA, chromatin primarily contains histones, proteins that regulate packing of DNA into higher-order structures. The detailed chromatin structure is still enigmatic, but in general actively transcribed DNA regions are loosely packed (euchromatin) and the DNA regions with inactive genes are packed more tightly (heterochromatin). Additionally, the chromatin structure undergoes significant transformations as the cell progresses in the cell cycle, which is especially evident in the course of cell division.

During their life, cells pass through a precise series of characteristic events, called the cell cycle (**Fig. 30**). In eukaryotes, after cell division, each of the newly formed daughter cells starts their life with the interphase. The cell cycle begins with the first period of the interphase, called Gap 1 (G_1), where the cell grows normally. If the cell is intended for a further division, it then enters S phase where it doubles its DNA content through DNA replication. RNA transcription and production of proteins, except histones, is low at this phase. After replication, each of the cell chromosomes has two sister chromatids. After S phase, the cell enters the G_2 phase, where it continues to grow with high rates of protein synthesis. Finally, after everything is set for the cell division, the cell enters mitosis (M phase), which itself is divided into several subphases (prophase, metaphase, anaphase, telophase), and then it proceeds with cytokinesis. During mitosis, chromatin condenses into chromosomes, the nuclear membrane breaks down and microtubules pull the sister chromatids at the level of their centromeres to opposite parts of the cell, which will then form two daughter cells. Each of the daughter cells will then enter interphase and the cycle can repeat again.

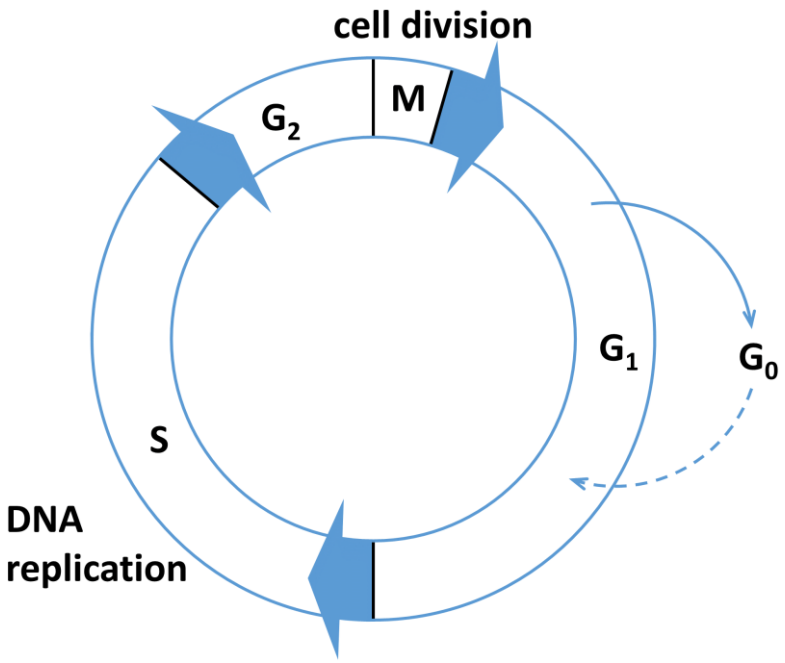


Figure 30. Schematic of the eukaryotic cell cycle. M, mitosis; G_1, Gap 1; G_0, Gap 0 or resting phase (cells in G_0 will not divide unless they re-enter G_1); S, synthesis; G_2, Gap 2.

The basic unit of chromatin is a nucleosome. It contains approximately 146 base pairs of DNA wrapped around a histone octamer consisting of four dimers of each of the core histones, H2A, H2B, H3 and H4 (Luger et al., 1997). This structure is supplemented with a linker histone H1 that binds and stabilizes the DNA ends (**Fig. 31**) (Bednar et al., 2017; Zhou et al., 2015).

DNA with histones form *in vitro* a "beads-on-a-string" fiber of 10 nm in diameter or more tightly packed 30-nm-diameter fibers under different conditions (Olins and Olins, 1974; Oudet et al., 1975; Woodcock et al., 1976). However, further studies of chromatin in cells did not confirm the presence of 30 nm fibers or other regularly ordered structures (Joti et al., 2012; Maeshima et al., 2010; Nishino et al., 2012). Currently, chromatin is imagined as an irregularly folded 10 nm fiber, which allows for easier and more dynamic DNA accessibility (Fussner et al., 2011; Joti et al., 2012). STORM microscopy demonstrated that nucleosomes are arranged along the "beads-on-a-string" fiber into domains of heterogeneous size and that less dense domains ("open" chromatin) are associated with RNA polymerase II and have smaller levels of H1 (Ricci et al., 2015).

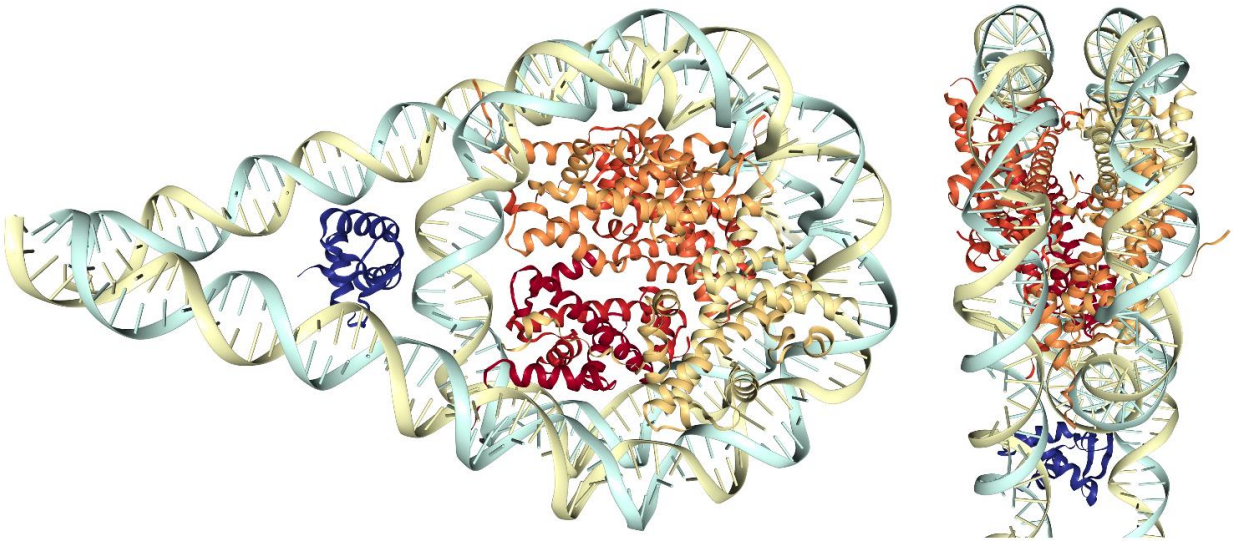


Figure 31. Nucleosome in complex with linker histone H1 (Bednar et al., 2017). DNA is wrapped around the histone octamer consisting of 2 copies of each of the core histones H2A, H2B, H3, H4. The linker histone H1 (shown in blue) stabilizes a compact nucleosome conformation. Image is created with NGL Viewer (Rose and Hildebrand, 2015).

The structure of chromatin obviously changes as cell progresses through the cell cycle. The most striking changes occur in mitosis, when chromatin condenses into distinct chromosomes, which are then split into two daughter cells. Mitotic chromosomes have a classic four arm structure and consist of two chromatids, which are joined at a region called centromere. On this region a large cellular structure builds up, kinetochore. During cell division, microtubules attach to the kinetochore from two sides and pull the sister chromatids apart, one for each of the two daughter cells (reviewed in Santaguida and Musacchio, 2009). Kinetochore is a complex structure consisting of around 100 different proteins in vertebrates (Samejima et al., 2015). Among them, 16 proteins form constitutive centromere-associated network (CCAN), the inner part of the kinetochore that is associated with the centromeres across the cell cycle and is the platform for the outer kinetochore recruitment in mitosis (Fukagawa and Earnshaw, 2014; McKinley and Cheeseman, 2016; Samejima et al., 2015).

In different organisms the centromeric regions have different extent on the chromosome. Some insects, plants and nematodes assemble a centromere along the entire chromosome. Such chromosomes are called holocentric (Guerra et al., 2010). The other eukaryotes form a centromere in a localized region (a monocentric chromosome). The position of the centromere on DNA is encoded epigenetically in most eukaryotes. The epigenetic marker for the centromere is an H3 histone variant, called Centromere protein A (CENP-A, also known as

CENH3) (Palmer et al., 1991). This protein replaces canonic histone H3 in some nucleosomes at the centromere. Monocentric chromosomes can have a point centromere, when only one CENP-A-containing nucleosome is present in the centromere (*e.g.*, in some budding yeasts (Pluta et al., 1995)), or regional centromeres with numerous CENP-A nucleosomes. In organisms with point centromeres the DNA sequence at the centromere is unique and it defines the centromeric position (McGrew et al., 1986). Within the regional centromeres, DNA usually contains repetitive sequences, known as satellite DNA, which are, however, not essential for the centromere function (McKinley and Cheeseman, 2016).

The amino acid sequence of CENP-A has substantial differences compared to H3, which allow centromere targeting (Black et al., 2007) and recruitment of core kinetochore proteins (Carroll et al., 2010) (**Fig. 32A**). It also explains structural differences of the CENP-A nucleosome as compared to the canonical H3 nucleosomes (the CENP-A nucleosome wraps ~20 bp less DNA than the canonical nucleosome (Miell et al., 2013)) (**Fig. 32B**) and, as a consequence, structural distinctions of the centromere chromatin, such as a stronger condensation (Geiss et al., 2014; Panchenko et al., 2011).

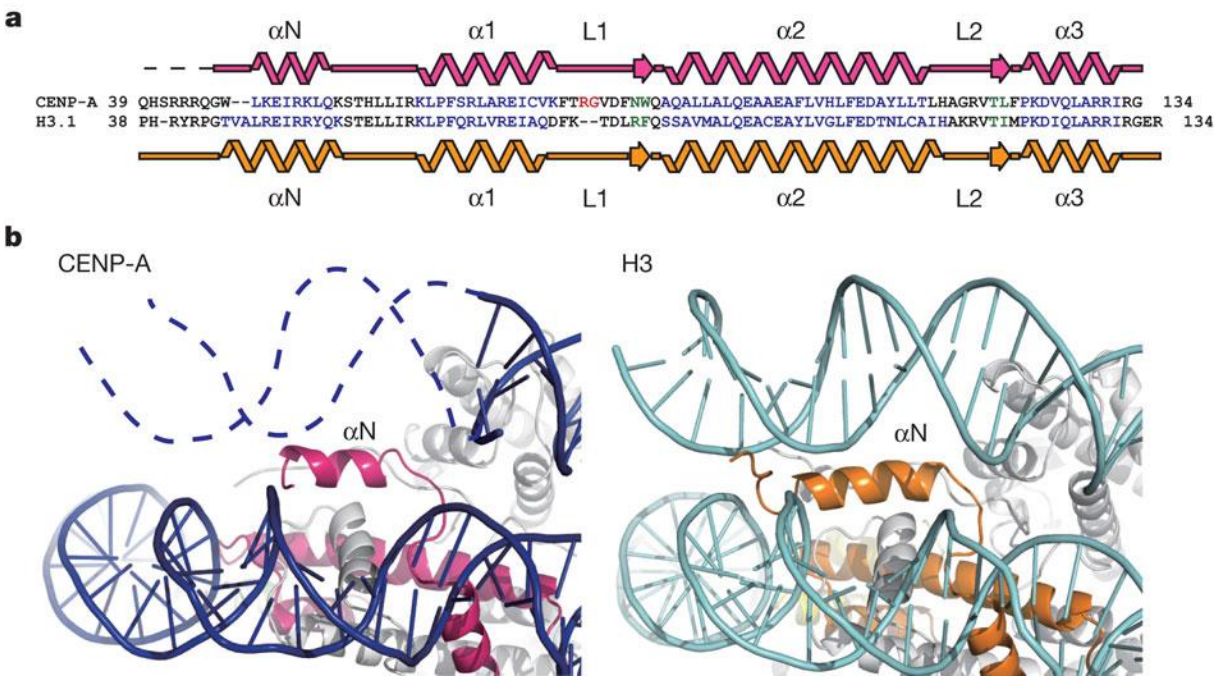


Figure 32. Structure of the DNA entrance and exit of the human CENP-A nucleosome. (**A**) Alignment of the CENP-A and H3 sequences. (**B**) Close-up views of the αN helices and the DNA edge regions of the CENP-A (left panel) and H3 (right panel) nucleosomes. Figure is adapted from (Tachiwana et al., 2011).

CENP-A is deposited with the help of a dedicated histone chaperone, HJURP (Holliday junction recognition protein) (Dunleavy et al., 2009; Foltz et al., 2009; Shuaib et al., 2010). Interestingly, the deposition of CENP-A into the centromeric chromatin does not occur during the DNA replication in S phase but it happens only after mitosis, in the early G1 phase (Bodor et al., 2013; Jansen et al., 2007). This fact raises the question how the structure of the centromere chromatin changes between the synthesis of new DNA and the deposition of new CENP-A (**Fig. 33**). In the current model, the CENP-A-containing nucleosomes are diluted between the daughter DNA strands during S phase and the missing places for the CENP-A nucleosomes are filled with H3.3-containing nucleosomes, which are replaced later in G1 phase with the CENP-A nucleosomes (Dunleavy et al., 2011). Previously, some data showed that the diluted CENP-A nucleosome may exist in a form of a tetramer (so-called hemisome), containing only one copy of each of H2A, H2B, CENP-A and H4 (Bui et al., 2012; Dalal et al., 2007; Dimitriadis et al., 2010) instead of the canonical form of a histone octamer with two copies of the core histones. More recent studies have shown, however, that the octameric form is the predominant form of the CENP-A nucleosomes throughout the cell cycle (Dunleavy et al., 2013; Miell et al., 2013; Padeganeh et al., 2013; Zhang et al., 2012).

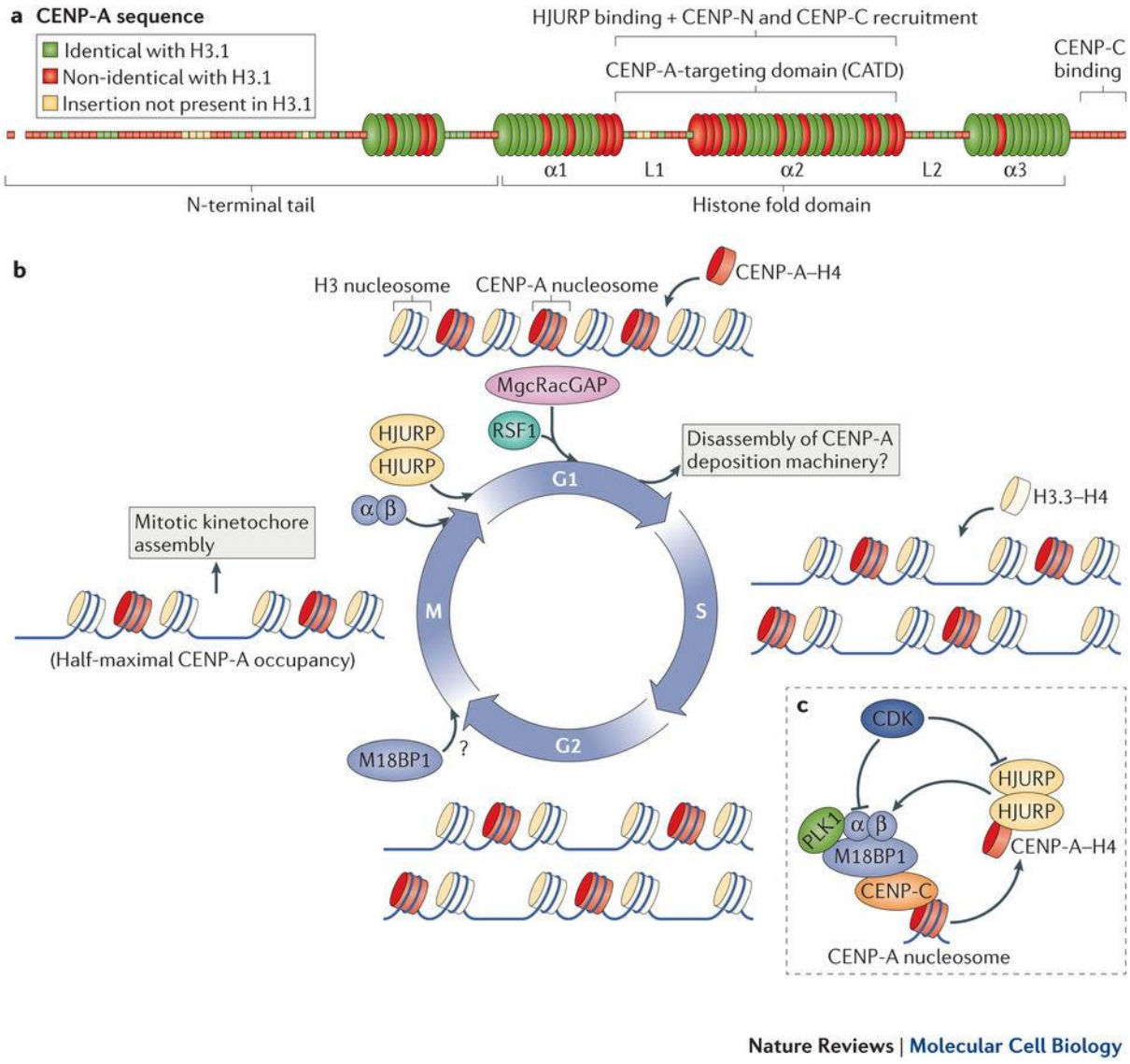


Figure 33. Function of CENP-A. (**A**) Primary and secondary structure of human CENP-A, indicating sequence conservation with H3. (**B**) Model of the transformation of the centromeric chromatin throughout the cell cycle. Figure is adapted from (McKinley and Cheeseman, 2016).

1.4.1. Super-resolution imaging of chromatin

SR microscopy gives a possibility to study the chromatin structure *in situ* or even *in vivo*. However, the high package density of chromatin prevents from resolving, *e.g.* individual DNA fibers or single histones even with most resolutive techniques; therefore the information about the chromatin structure has to be obtained in indirect ways.

Several studies have focused on topologically associated domains of chromatin. These are genomic regions within which physical interactions happen more frequently than between them. For example, by 3D STORM imaging of kilobase-to-megabase long genomic regions in *Drosophila* cells it was found that the volume of the transcriptionally active domains is larger than that of the inactive regions of the same length. The volume of the inactive regions is larger than that of Polycomb-repressed domains, which means that actively transcribing chromatin regions have a more open conformation than non-transcribing regions (Boettiger et al., 2016). Also, it was found that the actively transcribing, inactive and repressed regions have different size-scaling properties and 3D shape characteristics. The repressed domains showed a much stronger intermixing of chromatin within the region than the active and inactive domains. The spatial separation between neighboring domains of different epigenetic states also depends on their states (Boettiger et al., 2016). The 3D organization of transcriptionally inactive chromatin can be represented as the fractal globule state of a polymer (Lieberman-Aiden et al., 2009; Mirny, 2011), while Polycomb-repressed chromatin has much more compact packing (Boettiger et al., 2016). SMLM was also used for imaging of changes in the chromatin structure upon experimental ischemia of a cardiac muscle cell line and to show that chromatin is organized differently in different types of cancer cells (Szczurek et al., 2017).

When I started my PhD, little was known on SMLM of the centromeres. However, recently a few super-resolution studies were devoted to artificially unfolded centromeric chromatin in animals (Ribeiro et al., 2010) or to intact centromeres in plants (Ribeiro et al., 2017; Schubert et al., 2016). These studies allowed to propose an alternative model for the chromatin folding in kinetochores (**Fig. 34**) (Ribeiro et al., 2010; Schubert et al., 2016) and to observe ring structures in the centromeric regions of plant nuclei (**Fig. 35**) (Schubert et al., 2016). However, the understanding of the chromatin structure, including in the centromeres, is still very poor and further studies are needed in order to have a better model of structure-function features of chromatin.

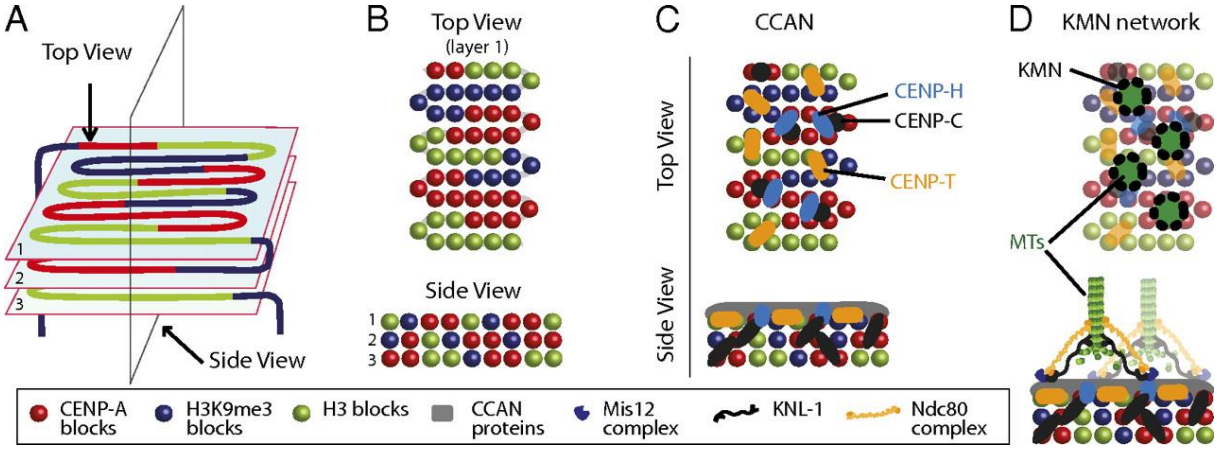


Figure 34. Model of centromeric chromatin organization based on SMLM imaging of unfolded centromeres (**A**) Layered structure of centromeric chromatin (**B**) Top and side view of A. (**C**) Distribution of CCAN proteins in the kinetochore. (**D**) Assembly of the KMN network. The proteins Mis12, KNL-1 and Ndc80 are the components of the outer kinetochore network KMN that links centromeric DNA with microtubules (MTs). Figure is adapted from (Ribeiro et al., 2010).

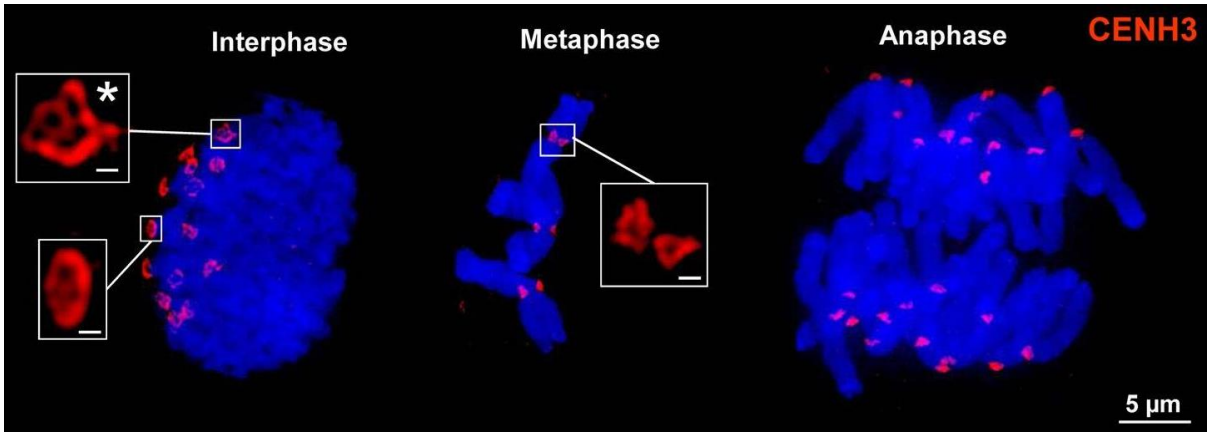


Figure 35. CENH3 chromatin is organized in ring structures throughout the cell cycle of rye cells. Scale bars in insets, 500 nm. Figure is adapted from (Schubert et al., 2016).

2. Objectives of the research project

The primary objective of my research project was development of methods for processing of SMLM data in order to make localization microscopy better suited for different applications in biology. At the beginning of my PhD, super-resolution microscopy was known for more than 10 years and localization microscopy for 7 years, so primary image processing methods, such as fitting of SM images, were already relatively well developed and dozens of different software for SM localization were available (reviewed in Sage et al., 2015; Small and Stahlheber, 2014). However, further processing of localization data, such as for corrections, evaluation, clustering, colocalization and other post-processing, was quite poorly developed. Only in a few localization tools authors started adding some features for basic corrections and post-processing. In such conditions, usage of SMLM for biological research was complicated because in order to obtain meaningful results, researchers had to use several different software tools for different steps of processing or even develop their own algorithms that would fit their experiments. This, of course, needs some programming experience and would keep many biologists away from using SR microscopy.

We therefore aimed at combining the post-processing steps that are most useful from our experience into a single tool. Namely, this tool would include: corrections of localization data for drift and chromatic aberrations, selection of localizations based on *e.g.* their localization precision, evaluation of the resolution by FRC, reconstruction of the corrected data into SR images, etc. The software should work with different input types of data and produce output in most useful formats.

One of essential applications of fluorescence microscopy is cluster analysis. However, the localization data of SMLM are not very usual for microscopy, this is why little or no algorithms were readily available for comprehensive cluster analysis of such data. The same is true for colocalization analysis of multicolor experiments. Localization data is extremely rich because it incorporates properties of individual molecules, and good algorithms should take full advantage of this information. An additional goal of my project was therefore development of such an algorithm, which would use relatively simple geometrical properties of SM localizations for extracting meaningful results on clustering and colocalization of fluorescent molecules, in a robust and as much as possible automated way.

While for many research projects 2D SMLM can be sufficient, going in three dimensions is essential for studies of the most complex objects, such as *e.g.* chromatin. Most of the data processing methods have been developed as 2D methods, some of them with the potential of an extension to 3D. Experimental methods for 3D SMLM data acquisition have been

successfully implemented; therefore, data processing methods should also be adapted for 3D. In fact, the extension of 2D methods for the 3D is not always straightforward because of the much higher complexity of 3D data as compared to 2D data. Along my PhD project it became clear that exploring 3D SMLM data processing was necessary. The next goal of my project was therefore extension of our 2D cluster analysis method on the third dimension and characterization of its performance.

Taking into account the current lack of understanding of the chromatin structure, we aimed at applying SMLM and our new data processing methods to imaging of chromatin in order to analyze its structure at the subdiffraction scale, in a time-dependent manner over the cell cycle of human cells. Since chromatin is a very densely packed object, SMLM imaging of it as a whole would be less successful than imaging of small fractions of it, namely because of the maximal allowable density of fluorophores in SMLM. We therefore focused on only a small and a well-defined fraction of chromatin, the centromeric region, which has a central role in cell function. The final objective of my thesis was therefore super-resolution imaging of centromeric chromatin in human cells over the cell cycle.

In summary, my PhD comprises four parts:

1) development of an integrative tool for SMLM data processing;
2) development of segmentation methods for SMLM data;
3) development of a 3D segmentation tool;
4) SMLM analysis of the centromeric chromatin region.

3. Results and discussion

3.1. SharpViSu, a software for SMLM data processing

The first part of my project was the development of a software for processing of SMLM data, which would integrate most useful processing steps in a single tool. I developed such a software, SharpViSu. It combines the following features that are necessary for getting best results from the pointillist data: support of one and double-color experiments, iterative cross-correlation based correction of drift, correction of chromatic aberrations, selection of localizations based on their brightness and/or frame number, reconstruction of 2D and 3D datasets in different visualization modes, resolution estimation of the data based on Fourier ring correlation, determination of the axial position of fluorophores in astigmatism-based 3D experiments. It also includes a plugin, ClusterViSu, which allows for comprehensive segmentation and cluster analysis of data using Ripley's functions and Voronoi diagrams. Details of the Voronoi diagram method are described in the next section, 3.2.

The results of the current section are presented below as an article, an application note in Bioinformatics (Andronov et al., 2016a). Methods and additional results and discussions are detailed later in the Supplementary material.

3.1.1. Publication 1 “SharpViSu: integrated analysis and segmentation of super-resolution microscopy data”

Bioinformatics, 32(14), 2016, 2239–2241
doi: 10.1093/bioinformatics/btw123
Advance Access Publication Date: 18 March 2016
Applications Note

OXFORD

Bioimage informatics

SharpViSu: integrated analysis and segmentation of super-resolution microscopy data

Leonid Andronov[1,2,3,4], Yves Lutz[1,2,3,4], Jean-Luc Vonesch[1,2,3,4] and Bruno P. Klaholz[1,2,3,4,*]

[1]Centre for Integrative Biology (CBI), Department of Integrated Structural Biology, IGBMC (Institute of Genetics and of Molecular and Cellular Biology), Illkirch, France, [2]Centre National de la Recherche Scientifique (CNRS) UMR 7104, Illkirch, France, [3]Institut National de la Santé et de la Recherche Médicale (INSERM) U964, Illkirch, France and [4]Université de Strasbourg, Strasbourg, France

*To whom correspondence should be addressed.
Associate Editor: Robert Murphy

Received and revised on January 29, 2016; accepted on February 28, 2016

Abstract

Summary: We introduce SharpViSu, an interactive open-source software with a graphical user interface, which allows performing processing steps for localization data in an integrated manner. This includes common features and new tools such as correction of chromatic aberrations, drift correction based on iterative cross-correlation calculations, selection of localization events, reconstruction of 2D and 3D datasets in different representations, estimation of resolution by Fourier ring correlation, clustering analysis based on Voronoi diagrams and Ripley's functions. SharpViSu is optimized to work with eventlist tables exported from most popular localization software. We show applications of these on single and double-labelled super-resolution data.
Availability and implementation: SharpViSu is available as open source code and as compiled stand-alone application under https://github.com/andronovl/SharpViSu.
Contact: klaholz@igbmc.fr
Supplementary information: Supplementary data are available at *Bioinformatics* online.

1 Introduction

The first step in processing of stochastic super-resolution microscopy data, the single-molecule localization, recently became a routine operation (Small and Stahlheber, 2014) and is often implemented by the manufacturer with the software operating the microscope. However, the further processing workflow of single-molecule localization microscopy (SMLM) data is much less standardized. Coordinates of individual molecules in space and time with their localization precisions are contained in a table of events obtained after fitting the Gaussian-like spots in the first step of the processing. Most available software for processing of single-molecule data such as rapidSTORM (Wolter *et al.*, 2012), QuickPALM (Henriques *et al.*, 2010), the Localization Microscopy plugin for μManager (Edelstein *et al.*, 2014), RainSTORM (Rees *et al.*, 2013) and ThunderSTORM (Ovesný *et al.*, 2014) are dedicated to fitting of camera images (Sage *et al.*, 2015), while few software, such as PALMsiever (Pengo *et al.*, 2015) and ViSP (El Beheiry and Dahan, 2013) are designed for processing of localization tables. The best way to analyze stochastic microscopy data is to work directly with eventlists (Deschout *et al.*, 2014) for which the development of new specialized and integrated tools is required.

2 Results and discussion

We have developed the SharpViSu software that combines the most important steps from our experience that are required for the treatment of localization data, namely: (i) multi-step correction of sample drift by cross-correlation with or without fiducial markers

© The Author 2016. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

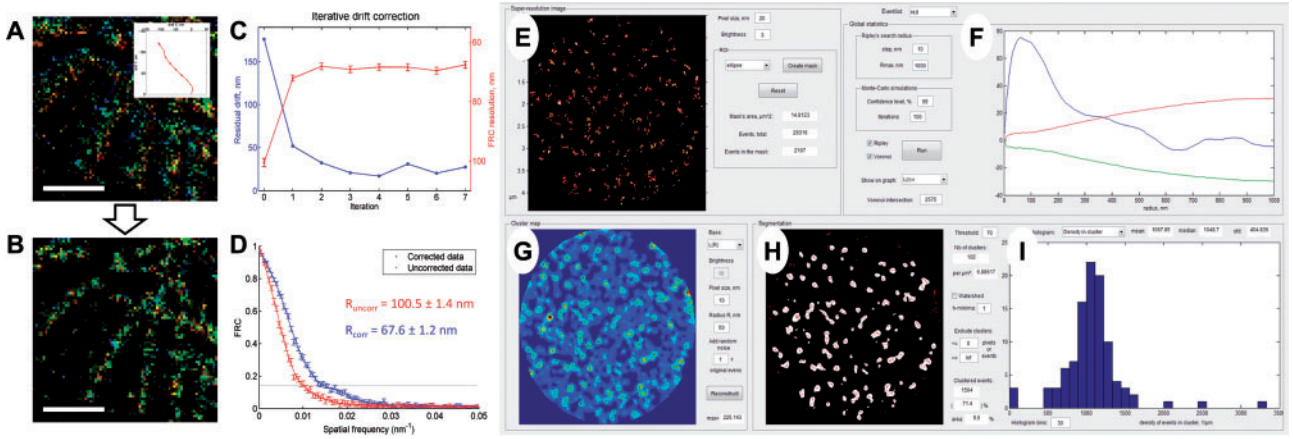


Fig. 1 Features of SharpViSu. (**A**, **B**) 1.5 μm × 1.5 μm fragment of a super-resolution image of β-tubulin in a HeLa cell reconstructed in the color-coded time mode before (**A**) and after 7 iterations of drift correction (**B**). The drift trace obtained by SharpViSu is shown in the inset. Scale bars: 500 nm. (**C**) Reduction of the estimated residual drift (blue) and corresponding improvement of FRC-resolution (red) by iterative drift correction. The curves converge after 2–4 iterations. (**D**) FRCs of the initial and the corrected datasets show statistically significant improvement in resolution. (**E–I**) Interface of ClusterViSu, a plugin for comprehensive segmentation of SMLM data. (**E**) Selected region of interest. (**F**) Statistics on localizations with Ripley's $L(r)$-r functions for the experimental data (blue) and 99% confidence interval for randomly distributed data (red and green) demonstrating statistically significant clustering. (**G**) Cluster density map calculated on the basis of Ripley's $L(R=50\,\text{nm})$ function. (**H**) Cluster map, binarized at the threshold $L=70$. (**I**) Histogram representing distribution of density of localizations in clusters. Data: nucleopore protein TPR, detected with Alexa-647-conjugated secondary antibodies (Lemaître *et al.*, 2014)

(Mlodzianoski *et al.*, 2011); (ii) sieving of event lists by merging consecutive events and removal of imprecise localizations; (iii) reconstruction of 2D super-resolution images in different modes (histogram, Gaussian (Huang *et al.*, 2008), quad-tree (Baddeley *et al.*, 2010), local density and hue-coded time) with selectable sampling; (iv) estimation of resolution by Fourier ring correlation (FRC; Banterle et al., 2013; Nieuwenhuizen *et al.*, 2013); (v) correction of chromatic aberrations for multi-color experiments (Erdelyi *et al.*, 2013); and (vi) reconstruction of 3D datasets with astigmatism (Huang *et al.*, 2008). The software also allows for calibration of localization data with chromatic aberrations and astigmatism. The output of SharpViSu can be saved in widespread formats such as .tiff (pictures), .png (graphs), .ascii or ViSP's .3dlp (El Beheiry and Dahan, 2013) (tables) allowing further analysis or preparation of publications.

SharpViSu provides a user-friendly integrated software package for corrections, analysis and visualization of super-resolution microscopy data (Supplementary Table S1). It uses localization tables as input which results in very small data sizes compared to raw time-lapse acquisitions, and a high precision of the contained information as compared to uncorrected, preliminarily reconstructed super-resolution images. For example, it handles iterative cross-correlation-based drift correction (without requiring fiducial markers) which shows progressive reduction of the estimated residual drift (Fig. 1A–D). The super-resolution image reconstructed from the corrected data looks much sharper, shows no detectable residual drift in the color-coded time representation (Fig. 1A, B) and demonstrates a significant improvement in resolution as quantified by FRC (Fig. 1D). SharpViSu allows correction of chromatic aberrations (Supplementary Fig. S2) and determination of Z-positions of fluorophores based on fitted data (Supplementary Fig. S4). Additionally, we introduced a novel local density visualization method based on Voronoi diagrams (Andronov *et al.*, 2016) that effectively improves the appearance of data and does not require any user-adjustable parameters that may be non-obvious to determine (Baddeley *et al.*, 2010). Finally, SharpViSu includes direct quantitative resolution evaluation with FRC.

The functionality of SharpViSu is extendable via plugins, such as ClusterViSu for comprehensive cluster analysis of SMLM data (Fig. 1E–I). It includes tools such as calculations of Voronoi and Ripley statistics (Owen *et al.*, 2010) with Monte-Carlo simulations, different modes of reconstruction (e.g. based on Gaussian blur or Ripley's functions) and segmentation of density maps, retrieval of geometrical properties of detected clusters, segmentation based on Voronoi tessellation (Andronov *et al.*, 2016; Levet *et al.*, 2015). SharpViSu is routinely used at the CBI/IGBMC for correction of super-resolution data and visualization of chromatin complexes (Lemaître *et al.*, 2014) and is largely applicable. SharpViSu is a timely contribution for the analysis of data from super-resolution microscopy, a research field in biology which is providing unprecedented insights into cellular and molecular function.

Funding

This work was supported by CNRS, the European Research Council (ERC Starting Grant 243296), Association pour la Recherche sur le Cancer (ARC), Université de Strasbourg and Investissement d'Avenir (IDEX), Institut National du Cancer (INCa), the French Infrastructure for Integrated Structural Biology (FRISBI) [ANR-10-INSB-05-01], Instruct as part of the European Strategy Forum on Research Infrastructures (ESFRI) and IGBMC facilities.

Conflict of Interest: none declared.

References

Andronov,L. *et al.* (2016) ClusterViSu, a method for clustering of protein complexes by Voronoi tessellation in super-resolution microscopy. *Sci. Rep.*, in press.

Baddeley,D. *et al.* (2010) Visualization of Localization Microscopy Data. *Microsc. Microanal.*, **16**, 64–72.

Banterle,N. *et al.* (2013) Fourier ring correlation as a resolution criterion for super-resolution microscopy. *J. Struct. Biol.*, **183**, 363–367.

Deschout,H. *et al.* (2014) Precisely and accurately localizing single emitters in fluorescence microscopy. *Nat. Methods*, **11**, 253–266.

Edelstein,A.D. *et al*. (2014) Advanced methods of microscope control using μManager software. *J. Biol. Methods*, **1**, e10.

El Beheiry,M. and Dahan,M. (2013) ViSP: representing single-particle localizations in three dimensions. *Nat. Methods*, **10**, 689–690.

Erdelyi,M. *et al*. (2013) Correcting chromatic offset in multicolor superresolution localization microscopy. *Opt. Express*, **21**, 10978–10988.

Henriques,R. *et al*. (2010) QuickPALM: 3D real-time photoactivation nanoscopy image processing in ImageJ. *Nat. Methods*, **7**, 339–340.

Huang,B. *et al*. (2008) Three-dimensional super-resolution imaging by stochastic optical reconstruction microscopy. *Science*, **319**, 810–813.

Lemaître,C. *et al*. (2014) Nuclear position dictates DNA repair pathway choice. *Genes Dev*., **28**, 2450–2463.

Levet,F. *et al*. (2015) SR-Tesseler: a method to segment and quantify localization-based super-resolution microscopy data. *Nat. Methods*, **12**, 1065–1071.

Mlodzianoski,M.J. *et al*. (2011) Sample drift correction in 3D fluorescence photoactivation localization microscopy. *Opt. Express*, **19**, 15009–15019.

Nieuwenhuizen,R.P.J. *et al*. (2013) Measuring image resolution in optical nanoscopy. *Nat. Methods*, **10**, 557–562.

Ovesný,M. *et al*. (2014) ThunderSTORM: a comprehensive ImageJ Plug-in for PALM and STORM data analysis and super-resolution imaging. *Bioinformatics*, **30**, 2389–2390.

Owen,D.M. *et al*. (2010) PALM imaging and cluster analysis of protein heterogeneity at the cell surface. *J. Biophoton*., **3**, 446–454.

Pengo,T. *et al*. (2015) PALMsiever: a tool to turn raw data into results for single-molecule localization microscopy. *Bioinformatics*, **31**, 797–798.

Rees,E.J. *et al*. (2013) Elements of image processing in localization microscopy. *J. Opt*., **15**, 094012.

Sage,D. *et al*. (2015) Quantitative evaluation of software packages for single-molecule localization microscopy. *Nat. Methods*, **12**, 717–724.

Small,A. and Stahlheber,S. (2014) Fluorophore localization algorithms for super-resolution microscopy. *Nat. Methods*, **11**, 267–279.

Wolter,S. *et al*. (2012) rapidSTORM: accurate, fast open-source software for localization microscopy. *Nat. Methods*, **9**, 1040–1041.

SharpViSu: integrated analysis and segmentation of super-resolution microscopy data

Supplementary Material

Leonid Andronov[1,2,3,4], Yves Lutz[1,2,3,4], Jean-Luc Vonesch[1,2,3,4] and Bruno P. Klaholz[1,2,3,4,*]

[1]Centre for Integrative Biology (CBI), Department of Integrated Structural Biology, IGBMC (Institute of Genetics and of Molecular and Cellular Biology), 1 rue Laurent Fries, Illkirch, France, [2]Centre National de la Recherche Scientifique (CNRS) UMR 7104, Illkirch, France, [3]Institut National de la Santé et de la Recherche Médicale (INSERM) U964, Illkirch, France and [4]Université de Strasbourg, Strasbourg, France.

*To whom correspondence should be addressed.

Supplementary Results and Discussion

To evaluate the precision of the drift correction in SharpViSu, we modulated data by adding a perfect circular-shape drift with a radius of 25 nm to a corrected dataset acquired with a β-tubulin-labelled sample. The direction and the amount of the drift obtained by cross-correlation is very similar to the introduced drift (Supplementary Fig. S1A). The algorithm was able to retrieve the original data with high accuracy: the resolution of the distorted dataset after 5 iterations of drift correction improved from 75.5 ± 1.1 nm to 50.2 ± 0.7 nm which equals to the initial resolution within the statistical error (Supplementary Fig. S1B).

Drift can be difficult to correct in the case of structure-less data. In the limit of randomly distributed localizations, cross-correlation peaks are located in random positions, leading to wrong values of the offset. Therefore, a redundancy in the experimental data is required for correction of drift by cross-correlations. Otherwise, fiducial markers should be introduced into the sample. The drift correction with fiducials by means of SharpViSu consists in choosing a small region of interest with the image of a bead, cross-correlation calculation of the drift from this small region, and application of the drift correction to the original dataset.

It should be noted that drift produces a shift between multi-color images in the case of sequential acquisitions. While the chromatic aberrations of the objective can be considered constant over time, drift is obviously variable. That is why simultaneous correction of both chromatic aberration and drift is indispensable for any multi-color experiment.

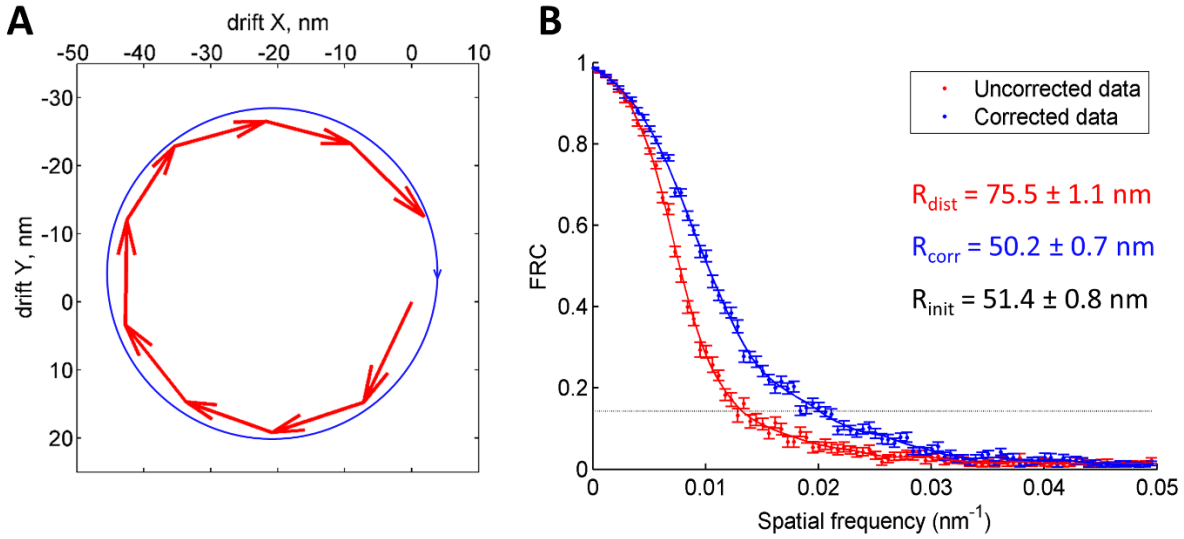
To demonstrate the ability of the software to correct data with chromatic aberrations we precalibrated our microscope using localization datasets with TetraSpeck multi-color beads (200 nm in diameter). After merging all the events within a radius of 100 nm and eliminating those with less than 10^5 photons, we obtained 26 events in each of the three channels, each event corresponding to the center of the point spread function (PSF) of a given bead (Supplementary Fig. S2A). The distances between the doublets of localizations were fitted with a 2-order polynomial function (Supplementary Fig. S2B). The second order of the polynomial was chosen, after testing, because it did not produce any edge effects. The quality of the fit is readily evaluable from our software.

To assess the precision of the fit and performance of double-color drift correction we simulated a typical super-resolution experiment by acquiring images of the beads first through the 642 nm channel and then through the 488 nm channel. The raw data from the microscope demonstrated significant chromatic aberrations and drift, in total up to 70 nm of distance between the images of each bead in the two colors (Supplementary Fig. S2C). This value is much worse than the localization precision, and would obviously lead to a significant loss in resolution and to misinterpretation of results. After correction for drift and chromaticity, the offset between channels dropped to less than 20 nm for the whole field of view, a value acceptable for super-resolution microscopy (Supplementary Fig. S2D).

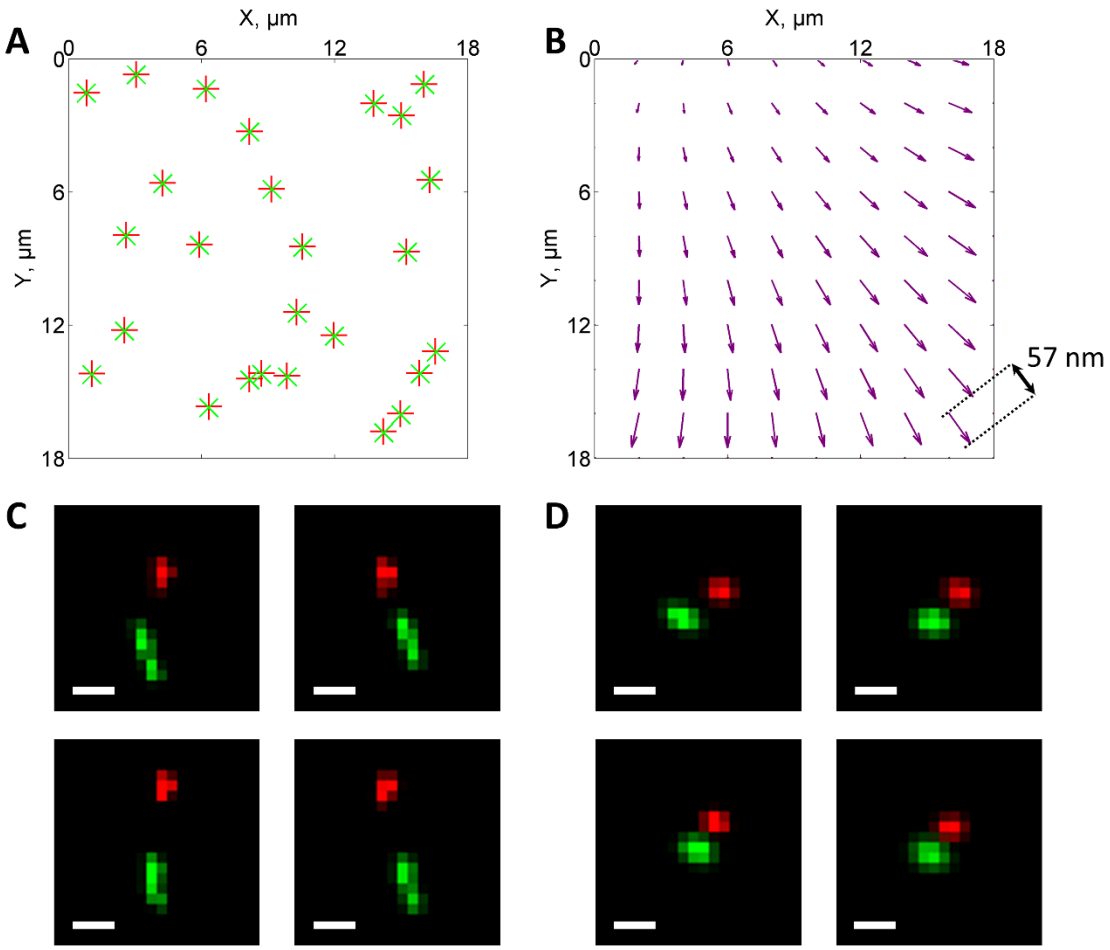
Multiple localizations of the same fluorophore can affect quantification of super-resolution experiments. This happens if the on-time of the fluorophore is greater than the exposure time of the acquisition frame, or if the boundary between two frames occurred during the shining of the dye, or if the fluorophore came to the “on-state” more than once during the time of acquisition. In the first two cases, the image of the dye appears on a few consecutive frames in the approximately same position. Signals from these localizations can be effectively merged into a single event, averaging the coordinates of the original events and summing the photon counts. The obtained localization has therefore higher precision.

We applied the described filtration to a drift-corrected β-tubulin dataset. As a result, the mean and the median values of the distribution of the number of photons collected in the event increased from 673 and 573 to 1428 and 889 respectively (Supplementary Fig. S3). This means that the average precision of localization improved 1.46 times. In other terms, before correction half of the localizations was determined with a precision better than 24 nm, after filtration this value improved to 19 nm. Precision here means the full width at half maximum of the corresponding Gaussian distribution with the standard deviation determined by formula (1, see the methods section). The less-precise localizations can be further rejected by means of the software imposing a minimum limit on the required photon counts. The regions with too high dye concentration that have not gone completely into the dark state appear on many consecutive frames and merge to a localization with a very high intensity. Such events can be rejected by introducing an upper limit on the photon count. In the above example, we only kept localizations with less than 10^4 photons (Supplementary Fig. S3B).

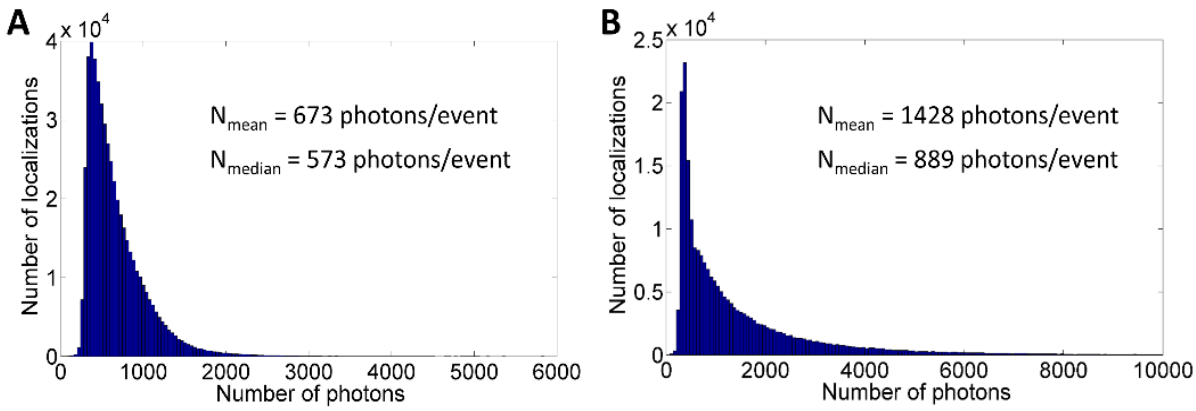
There are several methods to obtain the 3D coordinates of a fluorophore, including astigmatism (Huang *et al.*, 2008) bi-plane detection (Juette *et al.*, 2008) and double-helix PSF (Pavani *et al.*, 2009). SharpViSu provides a way to calibrate the astigmatic deformation of the PSF for the determination of the Z-position of dyes. The whole calibration process relies on localization tables with fitted positions and sizes (σ_x, σ_y or equivalent values) of the PSF for different Z-distances. The calibration should be done for every color channel that is intended for 3D imaging because of chromatic aberrations of the cylindrical lens and the objective. We provide two possibilities for multi-color calibration of the astigmatism. First, the imaging of multi-color beads should be done for all colors with the very same Z-positions. The obtained calibration curves $\Delta\sigma = \sigma_x - \sigma_y = f(z)$ can be readily applied to experimental data. Second, one can acquire calibration data for each color separately, keeping only the step in Z-positions of the objective constant. In order to consider the axial chromatic aberration, the offset in Z at the point where the PSF is closest to symmetrical should be measured separately and provided to the software. Alternatively, the offset in $\Delta\sigma$ near the $\Delta\sigma = 0$ in the same Z-position for all colors can be used (Supplementary Fig. S4).



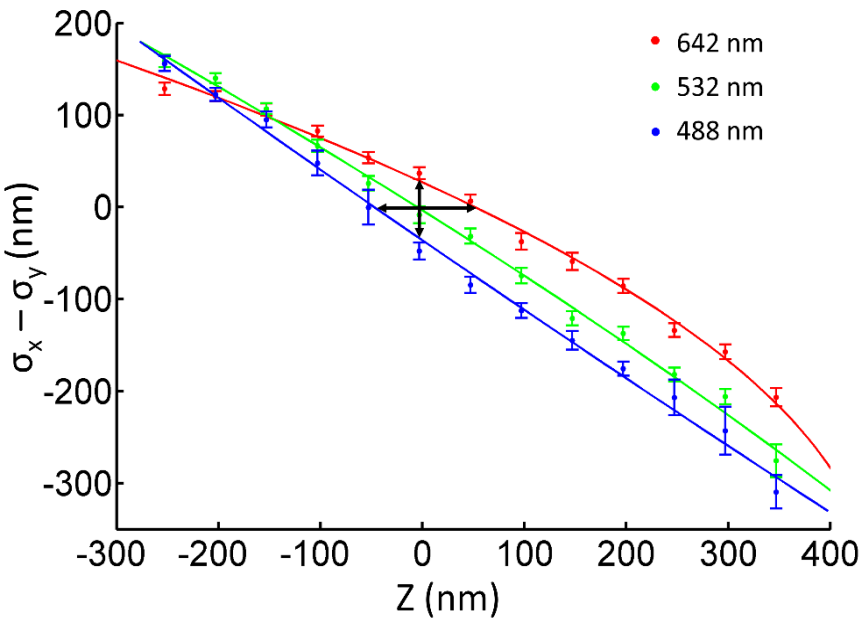
Supplementary Figure S1. Evaluation of drift correction with simulated circular-shape drift. (**A**) Red arrows show drift calculated by our algorithm from the data with artificial circular-shape drift. The blue circle with a radius of 25 nm shows the shape of the simulated ideal circular drift. (**B**) FRCs of the distorted data (red) and the corrected data (blue) show considerable improvement in resolution, with the resolution of the corrected data (R_{corr}) approaching the resolution of the initial non-distorted dataset (R_{init}); R_{dist}, the resolution of the artificially drifted dataset.



Supplementary Figure S2. Double-color correction of drift and chromatic aberrations. (**A**) Centers of images of 26 multi-color beads, detected with our algorithm. (**B**) Fit of the field of view of the objective with a second-degree polynomial. The arrows indicate the offset between the red and the green channels, multiplied 30 times for visualization. (**C-D**) Multi-color experiment simulated with Tetraspeck fluorescent beads. Each image shows a pair of images of a single bead, acquired through different color channels. The four images correspond to beads found in the corresponding corners of a single field of view. (**C**) Raw data demonstrate an offset due to chromatic aberrations and drift. (**D**) The same beads after correction of the chromatic aberrations using the fit (**B**) and the sequential drift correction. Red and green: imaging with excitation of 642-nm and 488-nm lasers accordingly. The images are in histogram representation with a pixel size of 5 nm. Scale bars: 20 nm.



Supplementary Figure S3. Sieving of datasets. (**A**) Histogram of the distribution of the number of photons per localization calculated from the raw data. (**B**) The histogram after merging consecutive events within a radius of 50 nm and subsequent removing of localizations with more than 10^4 photons demonstrates increased mean (N_{mean}) and median (N_{median}) photon counts, illustrating the improvement of the localization precision.



Supplementary Figure S4. Calibration of astigmatism by means of SharpViSu. The experimental points show mean values of $\Delta\sigma = (\sigma_x - \sigma_y)$ of a dataset, acquired at corresponding Z-positions of the objective with multi-color beads as sample. The error bars indicate standard deviations of the distribution of $\Delta\sigma$ for each position. The curves represent 2-order polynomial fits of the distributions. The horizontal arrow at $\Delta\sigma = 0$ nm indicates axial chromatic aberration of the system between the corresponding wavelengths. The vertical arrow at Z = 0 nm indicates the difference between $\Delta\sigma$ of the PSF of the corresponding color channels. These arrows represent the parameters that can be optionally used for calibration in SharpViSu.

	Sharp ViSu	PALM siever	ViSP	Quick PALM	Rapid STORM	LM µManager	Rain STORM	Thunder STORM
Platform	Matlab, stand-alone	Matlab	Stand-alone	ImageJ	Stand-alone	ImageJ	Matlab, stand-alone	ImageJ
Multi-channel	+	-	++	-	++	+	++	-
Drift correction	++	++	-	+	+	+	+	++
Chromatic correction	++	-	+	-	++	++	++	-
Fourier ring correlation	+	+	-	-	-	-	-	-
3D determination	+	-	-	+	+	+	-	+
3D viewer	+	+	++	-	-	-	-	+
Grouping of localizations	+	+	-	-	-	+	-	+
Local density visualization	+	+	+	-	+	-	-	+
Cluster analysis	++	+	+	-	+	-	-	-
Fitting of raw images	-	-	-	+	+	+	+	+

Supplementary Table S1. Comparison of SharpViSu with other open-source software for processing of localization microscopy data. Drift correction: "+", drift correction with fiducial markers; "++", also drift correction without fiducials. Chromatic correction: "+", only translation and scaling with manual parameters; "++", calibration with flexible transformation. 3D viewer: "+", slice viewer; "++", 3D volume viewer. Multi-channel: "+", 2 channels; "++", 3 or more channels. Grouping of localizations: averaging localizations within a given radius in consecutive frames. Cluster analysis: "++", comprehensive analysis including Ripley and Voronoi analysis with Monte-Carlo simulations and statistics on properties of detected cluster; "+", one of clustering methods available (either Ripley's K function, DBSCAN, or segmentation based on local density).

Process		Time
Loading dataset: 46 MB ASCII file exported from the Leica SR GSD software		1.5 s
Cross-correlation drift estimation and correction: histogram image (20 nm/pix), 11 subsets		6 s / iteration
Grouping consecutive events: R = 50 nm, ≤ 1 empty frame		15 s
FRC: histogram image (10 nm/pixel), 90 frequency points		5.5 s
Visualization	Histogram: 10 nm/pixel	0.5 s
	Gaussian: 10 nm/pixel	35 s
	Voronoi density: 10 nm/pixel	23 s
	Adaptive histogram: max bin capacity = 5	6 s

Supplementary Table S2. Computational time for some data processing steps in SharpViSu (stand-alone application) on example of a tubulin-labelled dataset with around $5 \cdot 10^5$ localizations. PC: i7-4770, 32GB RAM, Win7 x64.

Methods

Software

The software is written in Matlab 8.3 (The MathWorks Inc., 1984-2014). The GUI is built using GUIDE (The Matlab GUI Layout Editor). The software is designed to work with one or two sequentially or simultaneously imaged eventlists, called "red" (the firstly imaged channel) and "green" (the secondly imaged one, if present) for convenience. Besides that, for the "green" eventlist there is a possibility of choosing from two different imaging channels, e.g. for excitation laser lines 488 nm and 532 nm.

SharpViSu supports localization tables from Leica LAS AF, QuickPALM (Henriques *et al*., 2010), RapidSTORM (Wolter *et al*., 2012) or Localization Microscopy plugin for Micro-Manager (Edelstein *et al*., 2014).

For calibration of the objective for chromatic aberrations, one can use raw localization tables corresponding to super-resolution acquisitions of multi-color beads. To get the average positions of the centers of the beads throughout the frame, the events in consecutive frames are merged within a specified radius. Individual spread events are omitted by a threshold on the number of photons, controlled by a histogram. Finally, for the calibration only pairs of events located within a specified radius are taken into account. If there is more than one neighbor within the radius, this pair is excluded from the fit. The shift between the points is fitted with a 2D polynomial of a selectable degree (2-4, default 2) using Matlab's 'fitgeotrans' function. The obtained transformation objects are saved in the installation directory as files '488.mat' and '532.mat' describing the transformation of coordinates of the corresponding channels relative to the red one. The fit can be shown as a vector field on a regular grid. It is possible to load transformation objects from different fits to estimate the correction for the current experiment. During the correction, the transformation is applied to the 'green' eventlist with the Matlab function 'transformPointsInverse'. Optionally, it is possible to use predefined pairs of coordinates that are supposed to coincide on the corrected image.

Drift is calculated as follows: the entire eventlist is divided into n consecutive blocks with equal number of events in each of them (Szymborska, *et al*., 2013). From each of the blocks a histogram image is built for a 20 nm pixel size (by default, as the fastest and a robust enough method in our hands; the Gaussian and Voronoi representations are available with different sampling). Translation between each pair of consecutive images is calculated with subpixel precision by cross-

correlation using ‘dftregistration’ function (Guizar-Sicairos, *et al*., 2008). The (n-1) values of the offset are stored and used for graphical representation with vectors in the corresponding windows with ‘quiver’ function on an ‘axes’ graphics object. For representation, the values of the first and the last offset are multiplied by 1.5 to better account for the full acquisition range. For the correction of the calculated drift, the initial eventlist is divided into (n-1) sublists with borders on the centres of the n blocks used for the calculation of the drift. The first sublist contains the events from the beginning to the centre of the second block; the last sublist contains the events from the centre of the (n-1)th block until the end of the eventlist. The value of the total drift (including the drift in the preceding sublists and the drift from the beginning of the current sublist) is subtracted from the coordinates of each event. In the case of sequential two-color acquisition, the drift value for the events of the firstly imaged channel is calculated as the sum of drift in all the following frames and this value is added to the coordinates of the event in order to bring the drifted localizations towards the final frame of the channel. This procedure eliminates the additional shift between two sequentially imaged channels due to drift.

To merge consecutive events for a given localization the code checks whether there is an event in the following frame within a given Euclidean distance (default 50 nm). If so, it remembers the parameters of the consecutive event for comparison with the next frame. The procedure is iterated until there is no event in the next following frame. Then, the coordinates and standard deviations of all the retrieved events are averaged and the photon counts are added. The obtained average localization replaces the localization on the first frame in the series; the rest of the consecutive events are removed from the eventlist. If the number of “empty” frames is greater than zero, the procedure continues until there are no more than specified number of empty consecutive frames within the radius.

The histogram image representation is built with the ‘hist3’ function with edges generated for the specified pixel size. Quad-tree visualization image (Baddeley *et al*., 2010) is assembled from several histogram images with different pixel sizes, decreasing with a two-times step from half-size of the field of view until such a small size that no pixels have values greater than the required bin capacity. Then, all the images are resized to the maximal size, the excess overlapping regions are removed, and the final image is created by calculating the sum of the histogram images. The brightness of each cell in the final image is proportional to the number of events detected within its region, divided by the cell’s area.

The idea of Gaussian representation is to render the localization precision. It is achieved by displaying each event as a Gaussian kernel centered on (x, y) coordinates of the localization, with the standard deviation

$$\sigma = A/\sqrt{N} \tag{1}$$

where N – number of photons detected in the event; $A = \sigma_{psf} = 0.21 \cdot \lambda / NA$, σ_{psf} – standard deviation of the PSF of the microscope, λ – wavelength of detection, NA – numerical aperture of the objective (Abbe, 1873; Born & Wolf, 1999). The formula (1) approximates the localization precision as function of the number of detected photons (Thompson, *et al.*, 2002; Ober *et al.*, 2004). However, the real value of the localization precision depends on the accuracy of the localization software, calibration of the camera, type of noise, emitter properties etc. (Rieger & Stallinga, 2014; Deschout *et al.*, 2014; Endesfelder & Heilemann, 2014). We determined experimentally $A \approx 240$ nm on our system by measuring the spread of the distribution of localizations of a single object (fluorescent bead) for different N and λ. This gives sufficiently precise estimation of the localization precision for display purposes.

As opposed to Delaunay triangulation (Delauney, 1934), the Voronoi diagram has the advantage that each data point is situated inside an individual cell (Voronoi, 1908), so it is possible to use the value of the inverse area of the cell as the local density in the neighborhood of the data point. The Voronoi diagram is built using Matlab's function 'voronoin'. The values of the local densities are interpolated to a regular grid with selectable cell size using Natural Neighbor interpolation (Sibson, 1981; Matlab's function 'griddata' with option 'natural') which was the best option among those available in Matlab in terms of speed and precision. The interpolation avoids the usage of smoothening methods that are indispensable to get an interpretable image from a directly triangulated image plane (Baddeley *et al*., 2010). The grid is then transformed onto the image using the points of the grid as the centers of the corresponding pixels. To avoid edge effects during interpolation, the initial points lying near the borders of the field of view that have undefined Voronoi cells are considered to have zero density. If some points have exactly the same coordinates, only one point from this set is kept for the Voronoi triangulation, and the local density at this point is multiplied by the number of experimental points in the spot.

For the "time in color" representation, the image is firstly formed in hue-saturation-value (HSV) color space (Smith, 1978) with H values representing the time where H = 0° (red) corresponds to

the first frame and H = 240° (blue) corresponds to the last one, and V values representing the relative number of events in the pixel; the S value is always set to 1. The V values are multiplied by a specified multiplier to increase contrast.

To select a region of interest (ROI), a binary mask associated with a super-resolution image (histogram representation, pixelation chosen by user) is built interactively with the 'roipoly' function of Matlab. The software then checks in which pixel each event is situated; depending on the mask value in this pixel and selected option (Keep or Clear ROI), the event is either kept or removed from the eventlist.

The calculation of the Z-coordinate of events is based on the difference between standard deviations of Gaussian-shape spots of individual molecules in x and y directions that appears in the case of astigmatism. The fitted values of σ for each event should be present in the dataset. The astigmatism is calibrated as $f(z) = \sigma_x - \sigma_y$, fitted with a polynomial (2nd degree by default) (Henriques *et al*., 2010) using Matlab's 'polyfit'. The fit is used for determination of the Z-position of the fluorophore. The effect of the refractive index of the imaging buffer upon the magnification in the axial direction is taken into account by multiplying the one retrieved from the fit Z-coordinate by 0.79 (Egner and Hell, 2006), (Huang, et al., 2008) in case of mounting in an aqueous solution.

The FRC curve is calculated similarly to a previously described method (Nieuwenhuizen *et al*., 2013). First, the input eventlist is shuffled randomly using the Matlab function 'randperm'. The shuffled list is divided into two consecutive parts, approximately equal in length. From each, a histogram-mode image with sampling of 10 nm is built. A 2D Tukey window is applied to each image in the form of w2 = w * w`, where w = tukeywin (L, r), L is the size of the images in pixels (here L = 1800), r is a proportion of L that is equal to a part of a cosine function (here r = 0.25). The images are transformed into Fourier space by 'fftshift', and are divided into n rings with widths $w_{ri} = r_i - r_{i-1} = L/n$. For each ring, the FRC-value is calculated by the formula

$$FRC(q) = \frac{\sum_{\vec{q} \in circle} f_1(\vec{q})\, f_2(\vec{q})^*}{\sqrt{\sum_{\vec{q} \in circle} |\, f_1(\vec{q})|^2}\sqrt{\sum_{\vec{q} \in circle} |\, f_2(\vec{q})|^2}}$$

where $f_1(\vec{q})$ and $f_2(\vec{q})$ are the Fourier transforms of the two images. To get the resolution at 1/7 of FRC, the FRC-curve is smoothed by 'loess' (local regression using weighted linear least squares and a 2nd degree polynomial model) algorithm (Cleveland, 1979), and the first intersection of the smoothed curve with the value FRC = 1/7 is taken as the nominal resolution value.

Super-resolution imaging

HeLa cells were plated in a 4-compartment glass-bottom petri dish (CELLView, Greiner Bio-One) and fixed with 4% formaldehyde for 20 min in phosphate-buffered saline (PBS). After permeabilization with 0.1% Triton in PBS (PBS/Tx) twice for 10 min, the primary antibody (anti-β-tubulin monoclonal 1Tub-2A2, in house IGBMC) at 500x dilution in PBS/Tx was incubated overnight at 4 °C. The sample was then washed with PBS/Tx three times over 2 hours, and the secondary antibody (Goat anti-mouse Alexa Fluor-647 conjugated, Invitrogen) in dilution 4 µg/ml in PBS/Tx was incubated for 2 hours at room temperature. Subsequently, the cells were washed in PBS/Tx three times for 2 hours, then briefly three times in PBS. Prior to imaging, the sample was mounted in a PBS buffer that contained 10 mM of cysteamine and 25 mM of HEPES (pH 7.5).

The super-resolution experiments were performed on a Leica SR GSD system built on a base of DMI6000 B inverted wide-field microscope. We used the HCX PL APO 100x/1.47 Oil CORR TIRF PIFOC objective with 1.6x magnification lens that provides an equivalent pixel size of 100 nm on Andor iXon3 DU-897U-CS0-#BV EMCCD camera with a field of view of 18x18 µm in GSDIM mode. Continuous wave fiber lasers (MPBC Inc., 488 nm 300 mW, 532 nm 1000 mW, 642 nm 500 mW) and a diode laser (405 nm 30 mW) were utilized for excitation. The microscope is also equipped with the suppressed motion (SuMo) sample stage, which can reduce drift only to some extent, i.e. it cannot eliminate it (typical values 20-50 nm over 10 min).

The β-tubulin-labelled sample was first illuminated with the 100% laser power to quickly send the fluorophores into the dark state. The acquisition was started automatically after beginning of observation of single-fluorophore events ("blinking") that corresponded to the drop of the frame correlation value to 0.2 in the corresponding wizard in the LAS AF software. After a few minutes, as the number of blinking evens dropped, the sample started to be illuminated additionally by a 405 nm laser with gradual increase of its intensity in order to keep a nearly constant rate of single-molecular returns into the ground state. The acquisition was stopped after almost complete bleaching of the fluorophore.

For the multi-color imaging, the Tetraspeck multi-color beads were imaged sequentially: first by excitation with the 642-nm laser at 5% of its maximum power, then with the 488 nm one at 32% of power. The exposure time of a frame was 30 ms, the EM gain of the camera was 300 times.

The localization and fitting of single-molecule events were performed in real time during acquisitions in Leica LAS AF 3.2.0.9652 software with the "center of mass" fitting method. The

drift was detected using a pixel size of 20 nm, 10-11 consecutive sublists and histogram representation of sub-images for cross-correlation detection. For testing purposes, circular drift was added using the formula $x_i = x_{0i} + R \cdot \cos(2\pi n_i)$, $y_i = y_{0i} + R \cdot \sin(2\pi n_i)$, where x_{0i}, y_{0i} are the coordinates of an event i in the original dataset, R is the radius of the circle, $n_i = i/N$, N is the total number of events; R was set to 25 nm. In this case, the value of the angular shift is proportional to the number of the event; the drift circumscribes the full circle on the full duration of the acquisition.

The FRC curves were calculated in 90 concentric rings using histogram representation of half-datasets, unless stated differently. Prior to calculation the FRC, the eventlists were sieved by merging of repetitive localizations within a radius of 50 nm because these events may influence the FRC curve yielding too optimistic values (Banterle *et al*., 2013). For statistics, the FRC curves and the resolution values were calculated 50 times for each localization table. The FRC curves for histogram, Gaussian and Voronoi representations were also calculated by the FSC program (Image Science Software GmbH) (van Heel *et al*., 1996) confirming the results obtained with SharpViSu.

References

Andronov,L. *et al.* (2016) ClusterViSu, a method for clustering of protein complexes by Voronoi tessellation in super-resolution microscopy. *Sci. Rep.*, in press.

Abbe,E. (1873) Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung. *Arch. Mikrosk. Anat.*, **9**, 413–468.

Baddeley,D. *et al.* (2010) Visualization of Localization Microscopy Data. *Microsc. Microanal.*, **16**, 64–72.

Banterle,N. *et al*. (2013) Fourier ring correlation as a resolution criterion for super-resolution microscopy. *J. Struct. Biol.*, **183**, 363–367.

Born,M. and Wolf,E. (1999) Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light. Cambridge University Press, Cambridge, 461–475.

Cleveland,W.S. (1979) Robust Locally Weighted Regression and Smoothing Scatterplots. *J. Am. Stat. Assoc.*, **74**, 829–836.

Delaunay,B. (1934) Sur la sphère vide. A la mémoire de Georges Voronoï. *B. Acad. Sci. URSS*, **6**, 793–800.

Deschout,H. *et al.* (2014) Precisely and accurately localizing single emitters in fluorescence microscopy. *Nat. Methods* **11**, 253–266.

Edelstein,A.D. *et al*. (2014) Advanced methods of microscope control using μManager software. *J. Biol. Methods*, **1**, e10.

Egner,A. and Hell,S.W. (2006) Aberrations in confocal and multi-photon fluorescence microscopy induced by refractive index mismatch. In Pawley,J. (ed.) *Handbook of Biological Confocal Microscopy*, 3rd edn. University of Wisconsin, Madison, pp. 404–413.

Endesfelder,U. and Heilemann,M. (2014) Art and artifacts in single-molecule localization microscopy: beyond attractive images. *Nat. Methods* **11**, 235–238.

El Beheiry,M. and Dahan,M. (2013) ViSP: representing single-particle localizations in three dimensions. *Nat. Methods*, **10**, 689–690.

Guizar-Sicairos, M., Thurman, S. T. & Fienup, J. R., 2008. Efficient subpixel image registration algorithms. *Opt. Lett.* **33**, 156–158.

Henriques,R. *et al*. (2010) QuickPALM: 3D real-time photoactivation nanoscopy image processing in ImageJ. *Nat. Methods*, **7**, 339–340.

Huang,B. *et al.* (2008) Three-dimensional super-resolution imaging by stochastic optical reconstruction microscopy. *Science,* **319**, 810–813.

Juette,M.F. *et al.* (2008) Three-dimensional sub–100 nm resolution fluorescence microscopy of thick samples. Nat. Methods **5**, 527–529.

Nieuwenhuizen,R.P.J. *et al*. (2013) Measuring image resolution in optical nanoscopy. *Nat. Methods,* **10**, 557–562.

Ober,R.J. *et al.* (2004) Localization Accuracy in Single-Molecule Microscopy. *Biophys. J.* **86,** 1185–1200.

Ovesný,M. *et al.* (2014) ThunderSTORM: A Comprehensive ImageJ Plug-in for PALM and STORM Data Analysis and Super-Resolution Imaging. *Bioinformatics* **30**, 2389–90.

Pavani, S. R. P. *et al.* (2009) Three-dimensional, single-molecule fluorescence imaging beyond the diffraction limit by using a double-helix point spread function. *Proc. Natl. Acad. Sci. USA* **106,** 2995–2999.

Pengo,T. *et al.* (2015) PALMsiever: a tool to turn raw data into results for single-molecule localization microscopy. *Bioinformatics,* **31**, 797–798.

Rees,E.J. *et al.* (2013) Elements of image processing in localization microscopy. *J. Opt.*, **15**, 094012.

Rieger,B. and Stallinga,S. (2014). The Lateral and Axial Localization Uncertainty in Super-Resolution Light Microscopy. *ChemPhysChem* **15**, 664–670.

Sibson,R. (1981) A Brief Description of Natural Neighbour Interpolation. In Barnet V. (ed.) *Interpreting Multivariate Data*, Wiley, Chichester, pp. 21–36.

Smith,A.R. (1978) Color gamut transform pairs. *Comp. Graph.,* **12**, 12–19.

Szymborska,A. *et al*. (2013) Nuclear pore scaffold structure analyzed by super-resolution microscopy and particle averaging. *Science,* **341**, 655–658.

van Heel,M. *et al*. (1996) A new generation of the IMAGIC image processing system. *J. Struct. Biol.,* **116**, 17–24.

Voronoi,G. (1908) Nouvelles applications des paramètres continus à la théorie des formes quadratiques. *J. Reine Angew. Math.,* **133**, 97–178.

Wolter,S. *et al*. (2012) rapidSTORM: accurate, fast open-source software for localization microscopy. *Nat. Methods,* **9**, 1040–1041.

3.2. ClusterViSu, a method for 2D cluster analysis of SMLM data

The second part of my project was the development of a robust and automated method for segmentation and cluster analysis of SMLM data. I present these results below in a form of a publication in Scientific Reports (Andronov et al., 2016b).

In this segmentation method, we proposed to use Voronoi diagrams (Dirichlet, 1850; Voronoi, George, 1908) built on the coordinates of individual molecules. Voronoi diagrams represent a simple geometric concept and, like many other general mathematical concepts, are of use in a huge number of fields, in science, engineering, informatics, etc. (Okabe et al., 2000). Regarding SMLM, we noted that the inverse value of the surface area of a Voronoi polygon can be conveniently used as an approximation for the local density of molecules in the neighborhood of the molecule that belongs to the given polygon. Since a Voronoi diagram is virtually unique for a given set of points, this estimation is unambiguous unlike the commonly used estimation based on the number of points inside a circle of a manually defined radius. From these local density an image or a density map can be built. We showed that this density map by itself can be a useful way to represent SMLM data.

We show that the threshold for segmentation of SMLM data can be determined automatically by comparing the distribution of the Voronoi polygon sizes of the experimental dataset with that of the background noise model. Even though we used randomly distributed points as the default noise model for Monte-Carlo simulations, we demonstrate that the results of the segmentation can be more precise if this model is refined with *e.g.* a known number of re-localizations of fluorophores. This is because every fluorescent molecule may switch on and off several times during an experiment and this effect can distort clustering results if not accounted for (*e.g.*, this can produce a false positive clustering).

We have also shown that this method can be extended for colocalization analysis between distributions of two different species by analyzing the overlap between the two families of clusters determined by the Voronoi tessellations method.

For validation of the method we used simulated data as well as biological objects, namely microtubules, nuclear pores, histone H2B and RNA Polymerase II. Finally, we compared ClusterViSu with another recent method based on Voronoi tessellations, SR-Tesseler (Levet et al., 2015), we show advantages and shortcomings of both methods and we find that they can be complementary to each other.

3.2.1. Publication 2 “ClusterViSu, a method for clustering of protein complexes by Voronoi tessellation in super-resolution microscopy”

SCIENTIFIC REPORTS

OPEN

ClusterViSu, a method for clustering of protein complexes by Voronoi tessellation in super-resolution microscopy

Received: 2 September 2015
Accepted: 18 March 2016
Published: 12 April 2016

Leonid Andronov[1,2,3,4], Igor Orlov[1,2,3,4], Yves Lutz[1,2,3,4], Jean-Luc Vonesch[1,2,3,4] & Bruno P. Klaholz[1,2,3,4]

Super-resolution microscopy (PALM, STORM etc.) provides a plethora of fluorescent signals in dense cellular environments which can be difficult to interpret. Here we describe ClusterViSu, a method for image reconstruction, visualization and quantification of labelled protein clusters, based on Voronoi tessellation of the individual fluorescence events. The general applicability of this clustering approach for the segmentation of super-resolution microscopy data, including for co-localization, is illustrated on a series of important biological objects such as chromatin complexes, RNA polymerase, nuclear pore complexes and microtubules.

Single-molecule super-resolution microscopy recently brought fluorescence imaging into unprecedented levels of details by reaching the nanometer range of localization precision[1] and the ~30 nm range of actual resolution as estimated by Fourier ring correlation (FRC)[2,3]. As opposed to classical confocal microscopy, the data of stochastic super-resolution is essentially discontinuous because it comprises a set of points with molecular coordinates of the localization events, and hence it does not allow a direct segmentation. Special processing methods have been developed recently, including drift correction, visualization and estimation of co-localization[4–7]. However, considering the plethora of fluorescent signals in dense environments such as chromatin, the interpretation of localization data can become rather challenging with regards to image segmentation and cluster analysis. This is particularly true with respect to identifying and quantifying differently concentrated regions of labeled complexes. In some cases, super-resolution clustering on chromatin[8], neurons, lymphocytes and cell-surface receptors[9–11] has used Ripley's analysis for global overview of cluster properties in a given region; alternatively, pair correlation analysis[12] has been used, e.g. for studies of plasma membrane proteins[13]. For the estimation of local densities, for segmentation and cluster analysis methods such as Ripley's L function[9,10,14], median or Gaussian filtering of histogram images, k-means[8] and DBSCAN[15] clustering can provide some degree of visual clustering, but quantification is not straightforward. Recently, a Bayesian approach was developed for identification of clusters from a set of cluster proposals from Ripley's analysis[16]. Here we describe a method that we call ClusterViSu, which is based on Voronoi diagrams and tessellation of the individual fluorescence events for visual and quantitative clustering analysis of super-resolution microscopy data. When this manuscript was under review a similar method appeared using the same concept of segmentation based on Voronoi diagrams, called SR-Tesseler[17]. While ClusterViSu comprises additional features both studies complement each other; a more detailed comparison of the two methods including similarities and differences is made towards the end of this manuscript.

A Voronoi diagram, also known as Dirichlet decomposition, is a tessellation where a tile corresponding to a given data point is a locus of all points of space closest to this data point[18]. Applications of Voronoi tessellations are found in various fields from mathematics to natural sciences[19], e.g. for cluster detection in atom probe microscopy[20]. In the context of super-resolution microscopy as introduced here the Voronoi sites would correspond to the experimentally determined molecular coordinates of individual fluorophores. A Voronoi cell represents an area of influence of the data point it contains, and thus the local density in the proximity of a given point can be

[1]Centre for Integrative Biology (CBI), Department of Integrated Structural Biology, IGBMC (Institute of Genetics and of Molecular and Cellular Biology), 1 rue Laurent Fries, Illkirch, France. [2]Centre National de la Recherche Scientifique (CNRS) UMR 7104, Illkirch, France. [3]Institut National de la Santé et de la Recherche Médicale (INSERM) U964, Illkirch, France. [4]Université de Strasbourg, Strasbourg, France. Correspondence and requests for materials should be addressed to B.P.K. (email: klaholz@igbmc.fr)

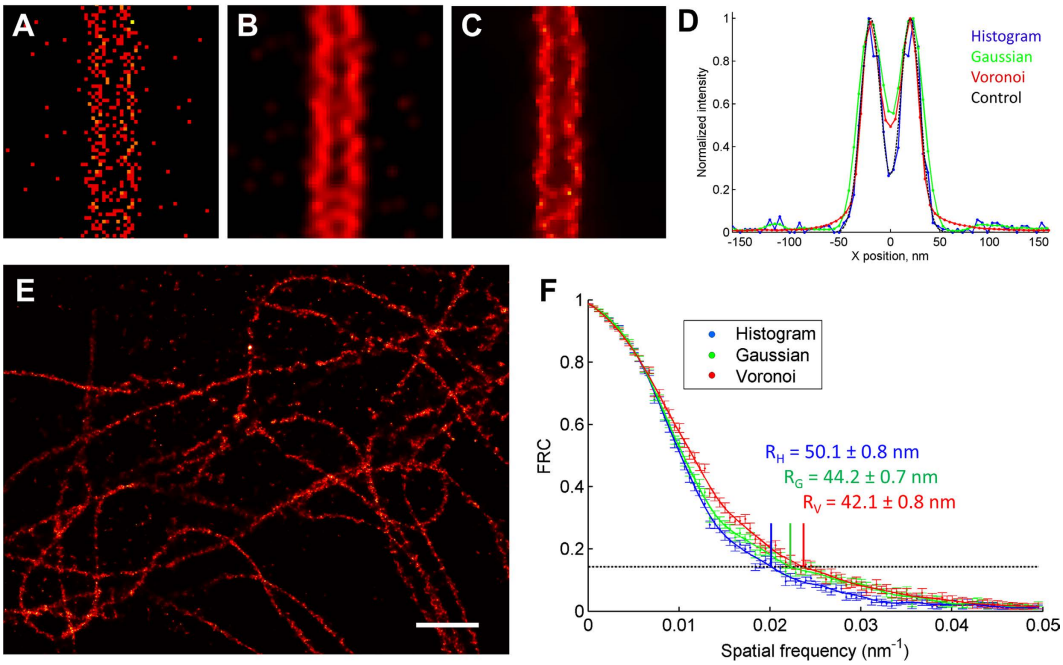


Figure 1. Comparison of different representation methods of localization events in super-resolution microscopy, showing that image representation by Voronoi diagrams fully preserves the resolution. (**A**–**C**) Simulated images of two lines separated by 40 nm and composed of localization events with a standard deviation of 10 nm in histogram (**A**), Gaussian (**B**) and Voronoi-based interpolated (**C**) local density representations. (**D**) Corresponding image projections generate back the underlying line structure – sum of two Gaussians (black dashed line); for line profiles see Fig. S1D. Each graph is normalized on its peak value; pixel size is 5 nm. (**E**) Voronoi density map of β-tubulin detected with Alexa-647 conjugated secondary antibodies. Pixel size is 10 nm, scale bar is 1 μm. (**F**) FRC curves calculated from a larger image containing (**E**) in histogram (blue), Gaussian (green) and Voronoi (red) representations with corresponding resolutions R_H, R_G and R_V obtained by the 1/7th FRC criterion, showing that the best resolution is obtained by Voronoi representation. The dataset used for calculation of FRC contained $2.3 \cdot 10^5$ localizations, the images in the three representations were reconstructed using a pixel size of 10 nm.

determined as the inverse of the cell area. This provides a direct precise measurement of the local density, unlike Ripley's analysis where the result depends on the chosen search radius. The values of local densities, interpolated to a regular grid, create a density map, which can be used for a direct image reconstruction and visualization of super-resolution data in the same manner as do standard histogram[21] or Gaussian[22] representation modes (Figs 1 and 2). The graphical properties of Voronoi diagrams and their mathematical propensity for potential quantification calculations prompted us to develop a method for clustering analysis based on Voronoi tessellations.

To validate the usage of the Voronoi diagram concept for super-resolution imaging, we first performed comparisons of image reconstructions using histogram[21] or Gaussian[22], and the Voronoi representation mode introduced here. Voronoi-based visualization both reduces visible noise and emphasizes features (Fig. 1A–D), and preserves the resolution upon image rendering equally well or better as compared to histogram and Gaussian mode image representations (Fig. 1F). The reason for this is that low-density and randomly distributed signals of the noise generate huge Voronoi cells, and the corresponding density values are much lower than those of highly dense regions. On the other hand, localizations at borders of dense areas have Voronoi cells significantly bigger than those of inner localization. This effect leads to muting of border localizations compared to internal ones, which is seen as effective increase of sharpness and contrast of the reconstructed image. On a test of visual resolution of a simulated structure with a linear density comparable to that of a labelled biological object ($0.5\,nm^{-1}$), the Voronoi visualization exhibits the best overall perception of the structure, discrimination between two closely located objects and complete suppression of noise (Figs 1A–E and S1). FRC curves calculated from Voronoi-density images lie at slightly higher frequencies than those of histogram- or Gaussian-mode images (Figs 1F and S1E–G). This resolution quantification indicates that the Voronoi-based visualization preserves the amount of useful information upon image rendering; this could be related with the fact that Voronoi-based reconstructions provide a better continuity of the signal (Fig. 1C), notably as compared to histogram-based reconstructions (Fig. 1A), which leads to a stronger correlation in FRC calculations.

Next, we extended the concept of Voronoi diagrams to perform cluster analysis, taking advantage of the intrinsic presence of Voronoi tessellations during the image reconstructions process (Fig. 2). A cluster is generally defined as a set of objects that are more similar to each other compared to objects in other sets. The properties of Voronoi cells (morphometric parameters such as shape, surface area, eccentricity etc.[19,23]) can be used as criteria for the association of several cells into clusters. We used intensity and spatial proximity parameters as

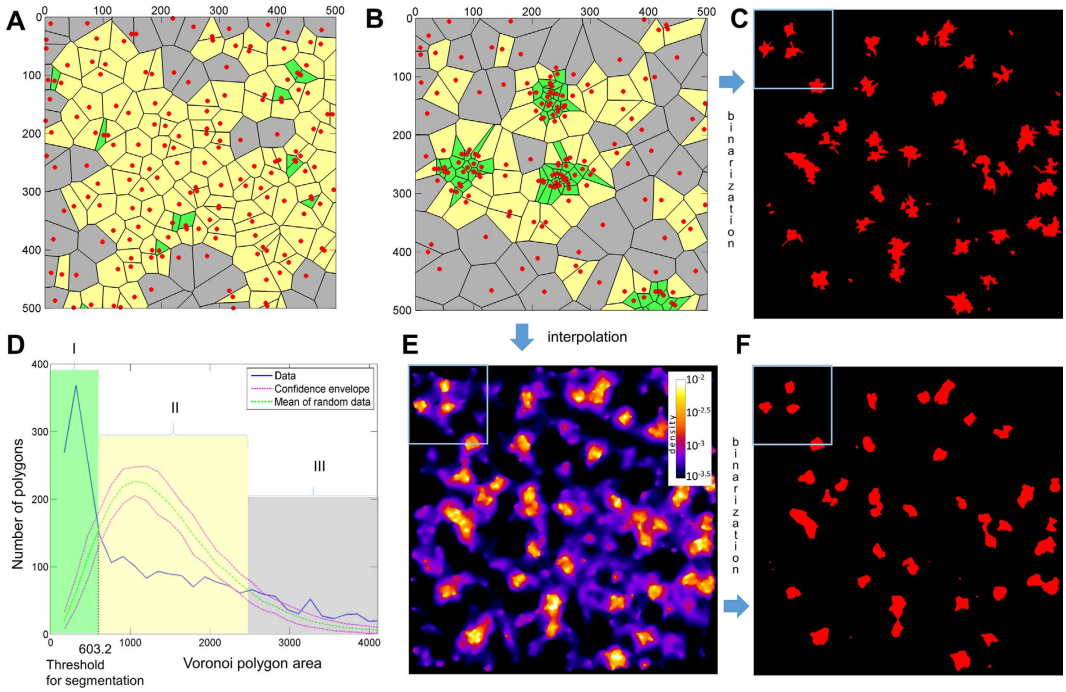


Figure 2. Principle of Voronoi-based image segmentation which allows visualization and quantification of clusters. (**A**) A region of simulated pointillist data with random distribution. (**B**) Simulated clustered distribution of the same number of points as in the random dataset. (**C**) Clustering obtained after direct segmentation of the Voronoi diagram allowing visualization of the clusters. (**D**) Distribution of Voronoi polygon areas of the clustered dataset (blue) and mean Voronoi polygon distribution of a random dataset (green) with confidence envelope (red) obtained from Monte-Carlo simulations which allows defining a threshold value for automated segmentation. The three characteristic regions: small clustered polygons (I, green); intermediate polygons corresponding to the random distribution (II, yellow); huge polygons corresponding to background in the clustered distribution (III, gray). (**E**) Interpolation of the local densities to pixel grid produces local density map. (**F**) Clusters, obtained by thresholding of the density map. The simulated dataset contained 48 clusters with a radius of 50 nm consisting of ~27 events each. After segmentation, the Voronoi tessellation allows quantification of the clusters (35 were detected, excluding small clusters that contained only one localization event, with 29 ± 16 events each and with equivalent radius of 49 ± 11 nm). The blue frames in (**C**,**E**,**F**) correspond to the region shown in panel (**B**).

determinants for clustering, allowing closely located fluorophore locations to be grouped according to the local fluorophore density. Thus, we define a cluster as a collection of neighboring Voronoi zones with areas smaller than a given threshold.

To test whether an experimental data set is significantly different from a spatially random distribution that has fluctuations of local density for a finite number of points, we performed Monte-Carlo simulations of randomly distributed points through the same experimental area using the same average density (Fig. 2; see methods). The probability distribution functions were calculated for experimental and for randomized data (including the mean value and associated 95% confidence interval; Fig. 2D). The comparison of distributions of Voronoi cell sizes of clustered and randomly distributed datasets reveals three characteristic regions (Fig. 2D): (i) small cell areas with an occurrence higher than in the random data set (cells describing clusters), (ii) intermediate regions with lower occurrence (containing Voronoi cell sizes similar to those of randomly distributed data), and (iii) regions of large Voronoi cells with higher occurrence corresponding to low-density event regions between clusters. The intermediate region (ii) can serve to define the boundaries of the clusters (i). The first intersection between the two density probability distribution functions (Fig. 2D) corresponds to the maximal cell size in the clusters, and this value can be used as a threshold value to automatically define regions of clustered polygons. Furthermore, calculation of the total area of the clustered polygons provides a direct measure of cluster occurrence which allows quantification of the fluorescence events within a given region. The inverse value of the obtained threshold can be used for binarization of the interpolated density map producing results similar to segmentation of Voronoi diagram (Fig. 2E,F). The difference is that clusters originated from the density map have smooth borders as the underlying Voronoi polygons are not displayed (Fig. 2F).

We applied the Voronoi tessellation strategy to visualize and quantify clustering observed in several important biological objects. A first example is the imaging of nuclear pore complexes (NPC; Fig. 3). Using Voronoi tessellation, individual NPCs (with a measured diameter of 114 ± 22 nm which corresponds to its known size[24]) can be distinguished even from sparse labeling (Fig. 3C,D), while this would not be obvious visually from a

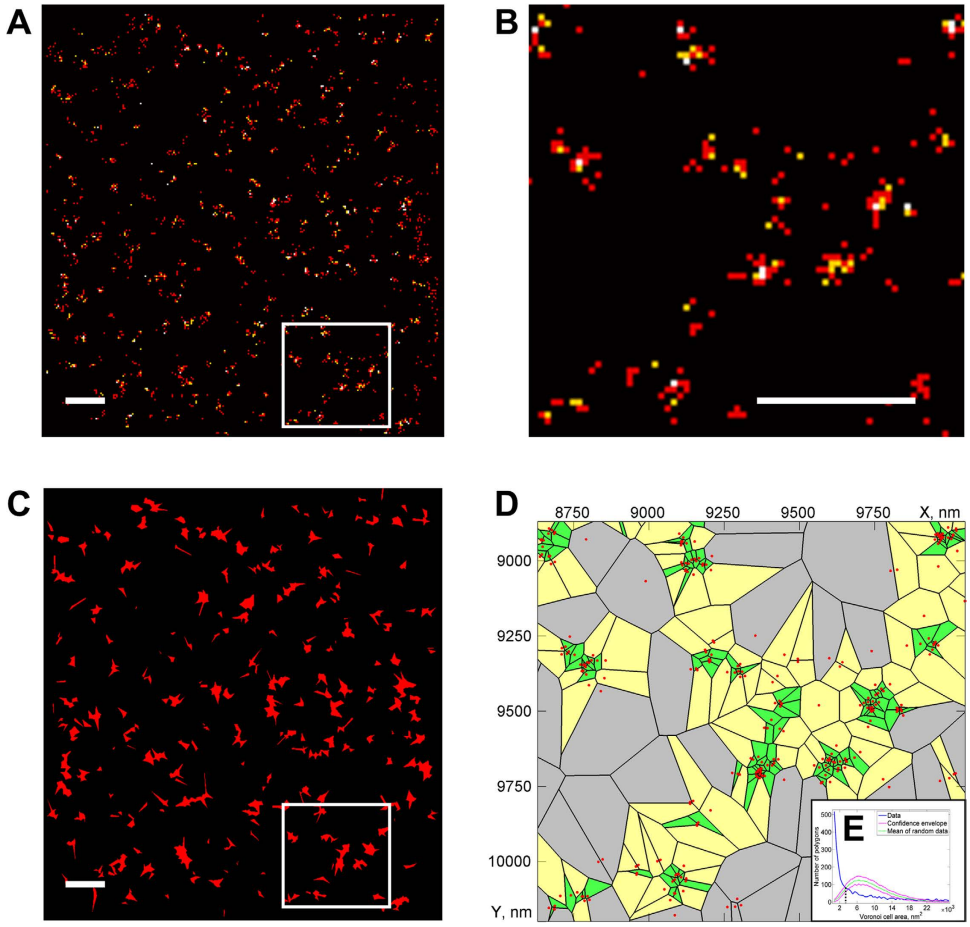


Figure 3. Segmentation and cluster analysis of nucleoporin protein TPR. (**A**) Histogram-based image reconstruction of TPR distribution at the nuclear envelope with a magnified region (white box) in (**B**). (**C**) Segmented Voronoi diagram calculated from the image in (**A**) with magnified region of the diagram shown in (**D**) which allows visualization and quantification of the TPR clusters present in the magnified region of the original image (panel **B**) in which quantification was not straightforward. (**E**) Voronoi polygon distribution of TPR (blue) demonstrating clustering as compared to random distributions (green and red). The clusters contain 9.7 ± 7.3 events with an equivalent radius of 57 ± 11 nm. Small clusters containing 2 or less events were excluded for this quantification. Scale bars correspond to 500 nm (**A**–**C**), pixel size is 20 nm (**A**,**B**).

histogram-based representation (Fig. 3B). The quantification shows that an average density of ~5.4 NPC complexes per μm^2 is found which is consistent with previous data[24,25].

We next extended the Voronoi tessellation method to multi-color super-resolution data to address the possibility of co-localization analysis (Fig. 4). For segmented multi-color images the co-localization value S/S_i for a given species i is defined by the ratio of the superposed area S between two colors relative to the total area S_i of clusters of a given species; this value is further compared with the confidence range for randomly distributed clusters S_{rand}/S_i. Thus, an S/S_i value in the interval from 0 to the lower boundary of the confidence range means anti-localization or mutual exclusion, and a value in the interval from the upper boundary of the confidence range to 1 corresponds to co-localization. To demonstrate the applicability of our co-localization method to a biological object we used tubulin simultaneously labeled with two dyes (Alexa Fluor-555 and Alexa Fluor-647). The Voronoi tessellation diagram calculated for each color mode separately or merged shows clusters along the microtubule fibers (Fig. 4A). For the tubulin sample, we observed relatively high co-localization values ($S/S_{red} = 0.416$, $S/S_{green} = 0.405$) compared to a random distribution ($S_{rand}/S_{red} = 0.123 \pm 0.021$, $S_{rand}/S_{green} = 0.120 \pm 0.020$) which confirms the visual impression of overlapping fiber structures. This, together with the simulations of different co-localization scenarios (Fig. S2), shows the validity of the Voronoi tessellation approach for co-localization analysis, as an alternative to coordinate-based methods[6,7].

To test the performance of Voronoi tessellation on dual-color labeling of different proteins we acquired super-resolution data on the cell nucleus (Fig. 4B) with RNA polymerase labeled with Alexa Fluor-488 and histone protein H2B labeled with Alexa Fluor-647. Interestingly, RNA polymerase is found more clustered in different regions of the cell nucleus (Fig. 4C), while histone H2B distribution (as compared to RNA polymerase) is more random within a given region (as illustrated by a shift of the density distribution compared to the random

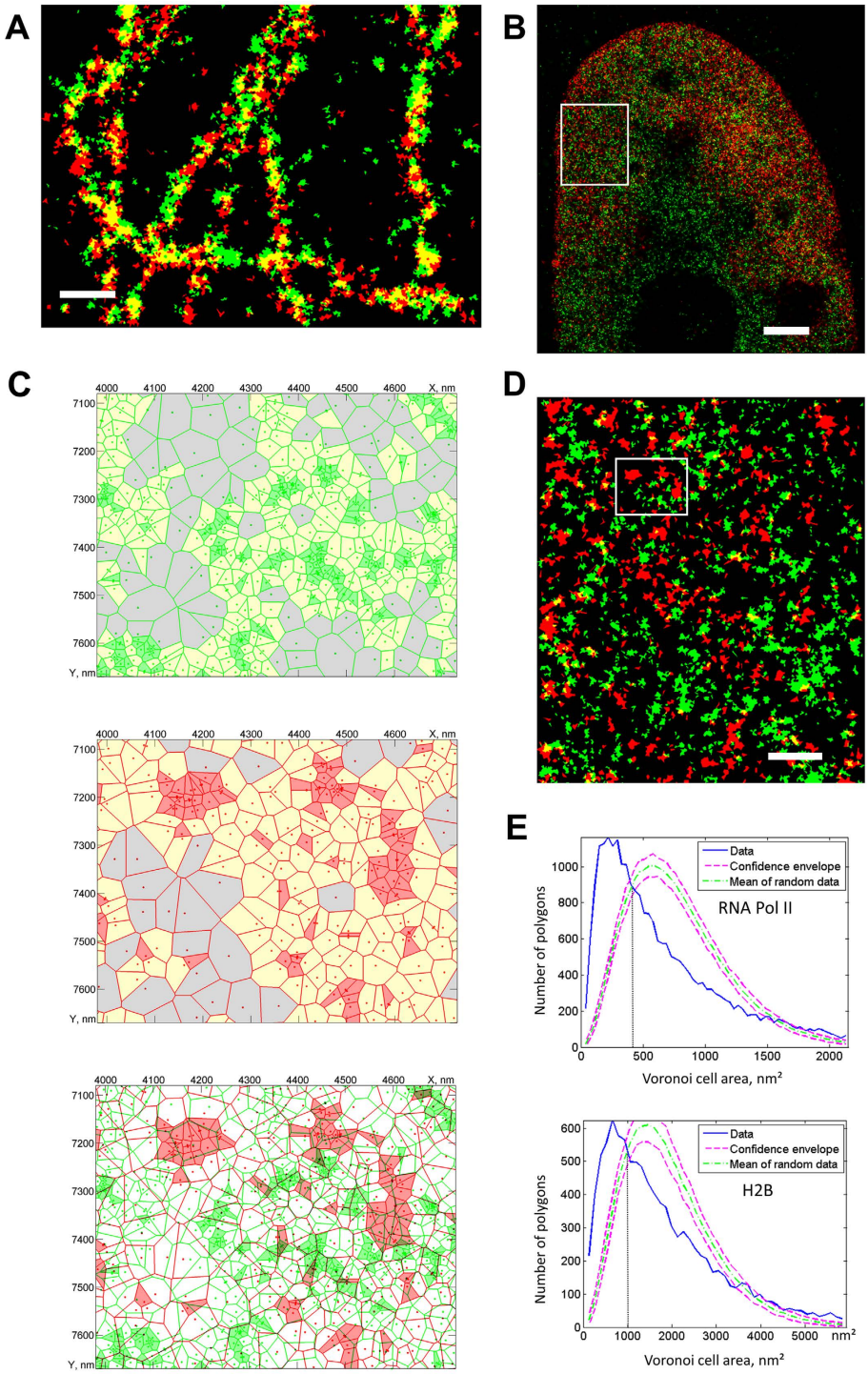


Figure 4. Co-localization analysis using Voronoi tessellations. (**A**) Segmented Voronoi diagram of β-tubulin, labelled with two different fluorophores (Alexa Fluor-555, green, and Alexa Fluor-647, red) demonstrating strong co-localization ($S/S_{red} = 0.416$, $S/S_{green} = 0.405$) compared to randomly distributed clusters ($S_{rand}/S_{red} = 0.123 \pm 0.021$, $S_{rand}/S_{green} = 0.120 \pm 0.020$). (**B**) A cell nucleus with labelled RNA Pol II (green) and histone H2B (red), represented as Voronoi local density map. (**C**) Example of Voronoi diagrams for RNA Pol II, H2B and their overlay, revealing no co-localization. (**D**) Segmented Voronoi diagram, used for determination of cluster properties (average cluster size: RNA Pol II 1818 nm^2, H2B 3058 nm^2) and for calculation of co-localization ($S/S_{Pol} = 0.090$, $S/S_{H2B} = 0.106$; random: $S_{rand}/S_{Pol} = 0.096 \pm 0.013$, $S_{rand}/S_{H2B} = 0.113 \pm 0.015$) revealing no correlation between the two distributions of clusters. (**E**) Voronoi polygon distributions of the clustered dataset (blue) and mean density distribution of the random datasets (green) with confidence envelopes (red) obtained from Monte-Carlo simulations for the RNA Pol II and H2B data. Scale bars 300 nm (**A**), 2 μm (**B**), 500 nm (**D**); panel (**D**) corresponds to the frame in panel (**B**) and (**C**) to that in panel (**D**).

distribution function, Fig. 4E). The clustering analysis and quantification using the Voronoi cell areas show that histone H2B labeling exhibits very weak co-localization with RNA polymerase II ($S/S_{Pol} = 0.090$; $S/S_{H2B} = 0.106$), equivalent to uncorrelated simulated data ($S_{rand}/S_{Pol} = 0.096 \pm 0.013$; $S_{rand}/S_{H2B} = 0.113 \pm 0.015$). Figure 4C shows that the clusters of the two proteins do not overlap. This observation appears consistent with the idea that DNA transcription by RNA polymerase occurs at H2A/H2B-depleted regions of less tightly compact chromatin[26–28], but this particular aspect would need to be analyzed in more detail in future studies, for example by H3 or H4 labeling. The Voronoi tessellation reveals an interesting non-random distribution of RNA polymerase molecules with a nearest neighbor distance between clusters ($d = 76 \pm 32$ nm). This suggests that they form some sort of superstructures which could correspond to actively transcribing regions in the cell nucleus.

An important property for super-resolution microscopy of Voronoi diagram is its geometric stability with respect to small changes of its sites[29], meaning that imprecise molecular localizations do not strongly influence results of segmentation. Indeed, we observed full retrieval of the number of clusters and their size in data with localization error $\sigma_{loc} \lesssim 3 \cdot \delta^{-1/2}$ and visually recognizable clustering with $\sigma_{loc} \lesssim 6 \cdot \delta^{-1/2}$ (estimated empirically) where σ_{loc} is the standard deviation of localization precision and δ is the density of clustered molecules (Fig. S3).

Multiple localizations of the same fluorescent molecules (through re-activation, blinking or multiple antibodies/fluorophores) may affect cluster analysis. By testing different simulated conditions we show that multiple localization (for example 2–6x more points distributed around a localization event) can improve the robustness of the segmentation under conditions of low background density (Fig. S4A) because more points are being used for defining clusters (*i.e.* the Voronoi cells become smaller and clusters show up better). In contrast, re-localizations of clusters with strong background lead to detection of spurious clusters in background regions (assuming multiple localization both for clustered points and for the background; Fig. S4B). Multiple localizations of randomly distributed points demonstrate false positive results in the Monte-Carlo simulations (Fig. S5), but if the number of re-localizations is known, it can be incorporated into the simulations leading to the correct results both in the evaluation of distributions for randomness and in segmentation (Fig. S5B,C).

Taken together, we present a robust method based on Voronoi tessellation for the convenient visualization and quantification of the localization and distribution of fluorescently labeled complexes, which allows segmentation, cluster analysis and estimation of the amount of co-localization. We show that image reconstruction using Voronoi diagrams preserves resolution equally well or better than Gaussian or histogram modes at the level of image rendering, as quantified by FRC. Voronoi tessellation allows performing a statistical analysis of the clusters, their occurrence and inter-cluster distance distribution, and works well including for the analysis of weak signals. Indeed, the general applicability of the method is illustrated here on a series of important biological objects including chromatin complexes, RNA polymerase, NPCs and microtubules. One of the advantages of the Voronoi tessellation method is that it does not require any *a priori* knowledge for the clustering (e.g. user-defined parameters such as the search radius for Ripley's analysis, the number of clusters in the case of k-means clustering or search distance and number of points for DBSCAN; these methods do not assign an area of influence to data points such as Voronoi tessellation does which also provides boundaries between clusters). Because the clustering uses an internal reference generated from randomized data to automatically determine the threshold value for forming clusters between neighboring Voronoi zones, it is fully automated for a given region of interest. We have implemented the Voronoi tessellation method as a ClusterViSu standalone software which can also be interfaced as a plugin with an interactive open-source software for processing of super-resolution fluorescence microscopy data (SharpViSu; ref. 30; https://github.com/andronovl/SharpViSu) which includes other standard tools such as corrections for drift and chromatic aberration and resolution estimation by FRC. In the future, the approach could be extended to 3D super-resolution data because Voronoi diagrams are also defined for multi-dimensional cases.

To our knowledge, this is the first application of Voronoi tessellation for the clustering analysis of super-resolution microscopy data, together with a related method published by Levet *et al.*[17] while the present manuscript was under review. To quantify differences between ClusterViSu and SR-Tesseler we run some simulations (see methods; Fig. S6). When compared with SR-Tesseler[17], our method shows a more complete detection of the cluster numbers and retrieval of their size and homogenous shape (Fig. S6A–E), especially at conditions of weak density of clustered points or weak background (Fig. S6C–E). In ClusterViSu, the detection of clusters is insensitive to background densities, while SR-Tesseler shows some artefacts for weak densities (Fig. S6E). The segmentation threshold determined automatically by Monte-Carlo simulations as proposed in our work is more stable over a large range of background densities (Fig. S6F) as opposed to thresholds determined by the average localization density[17]. In ClusterViSu, the threshold diminishes slightly with increasing background which favorably reduces spurious detection of clusters (Fig. S6F), while the increasing threshold values can lead to under/overestimation of the cluster number at low/high densities in SR-Tesseler. The usage of complete spatially random distributions of the background in our study as opposed to uniform distributions[17] provides more realistic simulations and explains the occurrence of spurious clusters that contain small numbers of events (in ClusterViSu these can be removed with a filter). ClusterViSu uses zero-rank density calculations, while SR-Tesseler uses first-rank as defined in ref. 17 which is more resistant to false detections at high cluster and high background densities (Fig. S6D) due to averaging of neighboring Voronoi polygons. ClusterViSu is also compatible with multiscale segmentation[17] (further segmentation of detected clusters), which can be useful for sub-classification of cellular structures. Additional differences are that our implementation already allows double-labeling and quantitative co-localization analysis using Voronoi-tessellation (Fig. 4) which is an important feature for super-resolution studies. In addition, we quantify the resolution of image rendering using Voronoi diagrams and show that it is at least as good as that of Gaussian and histogram mode reconstructions. Furthermore, the four biological examples that we give in this work extend the field of applications of Voronoi-based tessellation, segmentation and quantification of data.

Methods

Super-resolution imaging. HeLa cells were plated in a 4-compartment glass-bottom petri dish (CELLView, Greiner Bio-One) and fixed with 4% formaldehyde for 20 min in phosphate-buffered saline solution (PBS). After permeabilization with 0.1% Triton in PBS (PBS/Tx) twice for 10 min, the primary antibody (anti-β-tubulin monoclonal (1Tub-2A2, in house IGBMC) was used as mouse ascites fluid diluted 500× in PBS/Tx; RNA Pol II monoclonal antibody (1PB-7C2, in house IGBMC), directed against the CTD of the largest subunit of RNA Pol II (RPB I) was used as a purified IgG at 5 μg/ml in PBS/Tx; histone H2B monoclonal antibody (LG11-2) was used as a 500× dilution of mouse ascites fluid in PBS/Tx) was incubated overnight at 4 °C. The sample was then washed with PBS/Tx three times over 2 hours, and the secondary antibody (goat anti-mouse Alexa Fluor-647 or Alexa Fluor-555 conjugated, Invitrogen) in dilution 4 μg/ml in PBS/Tx was incubated for 2 hours at room temperature. Subsequently, the cells were washed in PBS/Tx three times for 2 hours, then briefly three times in PBS. For the TPR sample, cells were cultured on coverslips and fixed in 4% paraformaldehyde for 10 min, permeabilized in 0.5% Triton for 10 min, blocked in 1% BSA for 30 min, and incubated with primary antibodies (TPR rabbit polyclonal antibody, Abcam, ab84516) for 1h and with secondary antibodies (goat anti-rabbit Alexa Fluor-647) for 45 min[14]. The double-labelled β-tubulin sample was mounted in an imaging buffer[31] that contained 20% of Vectashield (Vector Laboratories), 70% of 2,2′-thiodiethanol (also known as thiodiglycol or TDE) and 10% of PBS 10× (the measured refractive index of this mounting medium is 1.491). The H2B/RNA Pol II, the TPR and the β-tubulin sample used for Fig. 1 were mounted in a PBS buffer with addition of 10 mM of cysteamine (also known as β-mercaptoethylamine or MEA) and 25 mM of HEPES (pH 7.5).

The super-resolution experiments were performed on a Leica SR GSD system built on a base of DMI6000 B inverted wide-field microscope. We used the HCX PL APO 100×/1.47 Oil CORR TIRF PIFOC objective with a 1.6× magnification lens that provides an equivalent pixel size of 100 nm on the Andor iXon3 DU-897U-CS0-#BV EMCCD camera with a field of view of 18 × 18 μm in super-resolution mode. Continuous wave fiber lasers (MPBC Inc., 488 nm 300 mW, 532 nm 1000 mW, 642 nm 500 mW) and a diode laser (405 nm 30 mW) were utilized for excitation. The microscope is also equipped with a suppressed motion (SuMo) sample stage, which reduces drift but does not eliminate it (typical values 20–50 nm over 10 min). The residual drift was corrected by data processing (see below).

The samples were first illuminated with the 100% power of the appropriate laser to quickly send the fluorophores into the dark state. The acquisition started after beginning of observation of single-fluorophore events ("blinking") that corresponded to the drop of the correlation value of consecutive frames to approximately 0.2 in the corresponding wizard in the LAS AF software. The time of exposition of a frame was 6.34 ms (H2B and RNA Pol II data) or 50 ms (β-tubulin and TPR data); the electron multiplying gain of the camera was 300 for H2B, RNA Pol II, TPR, 150 for β-tubulin-Alexa 647 and 63 for β-tubulin-Alexa 555; the laser power during the acquisition was 30% (H2B), 50% (TPR) or 100% (RNA Pol II and β-tubulin). After a few minutes, as the number of blinking evens dropped, the sample started to be illuminated additionally by a 405 nm laser with gradual increase of its intensity in order to keep a nearly constant rate of single-molecular returns into the ground state. The acquisition was stopped after almost complete bleaching of the fluorophore.

Data processing. The localization and fitting of single-molecule events were performed in real time during acquisitions in Leica LAS AF 3.2.0.9652 software with the "center of mass" fitting method. To reduce the number of localizations of the same fluorophore and improve localization precision the data were processed by averaging the coordinates of consecutive events within a radius of 50 nm around each localization. Subsequently, the data were processed in SharpViSu (ref. 30, in press; https://github.com/andronovl/SharpViSu). The drift was detected and corrected using cross-correlation-based approach. Briefly, the dataset was divided on several consecutive subsets, from each of them a histogram image with pixelation of 20–40 nm was build, the shift between these images was detected with subpixel precision and interpolated linearly throughout intermediate frames. The shift was then subtracted from every frame. The procedure was repeated iteratively 4 times assuring absence of detectable residual drift. For correction for the chromatic aberrations, the aberrations were detected beforehand with Tetraspeck multi-color beads, interpolated through the field of view with 2-order polynomial functions and subtracted from the molecular coordinates obtained from either 488 nm or 532 nm imaging channels. Regions of interest (ROI) for Voronoi analysis were selected manually allowing faster computations and more homogeneous distributions compared to the entire field of view.

The following analysis was performed in Matlab using customized code that is included as modules in the ClusterViSu software. Voronoi diagrams (vertices of polygons and connectivity order) were retrieved with either 'voronoi' or 'voronoin' functions. Areas of the cells were determined from the vertices with the function 'polyarea', the local density in each data point was defined as the inverse value of the area of the corresponding Voronoi polygon. To have a smooth appearance that can be used for visualization (Fig. 1C–E) or segmentation (Fig. 2E,F) the values of the local density were interpolated to a regular grid (pixels) using the 'griddata' function and the 'natural' interpolation method[32]. The spacing of the grid corresponds to the desired pixel size (in range of 2 to 20 nm in our case, it is indicated in the figure legends; at least 3–4 times smaller than the expected resolution to satisfy the Nyquist theorem).

For the comparison of visualization techniques, we used a Gaussian distribution of points arranged as two parallel lines, with a distance of 40 nm between the centers of the distributions and the standard deviation of each distribution $\sigma = 10$ nm in the X direction (Fig. 1A–D). The linear density of localization in each line was 0.5 nm^{-1}. To this dataset we added localizations with random (x, y) coordinates, with overall density of 400 μm^{-2}, comparable to background density at experimental data. The photon counts for all events were distributed with Gaussian distribution around the mean value of 1000 photons with standard deviation of 300. This is similar to a procedure described by Baddeley *et al.*[4] with following modifications: linear densities were chosen to be similar to those of typical experimental data, noise was added, and single line profiles (Fig. S1D) were used along with

projections (Fig. 1D; even though projections tend to not reflect discontinuities in the data, i.e. they generate back the underlying line structure). The Gaussian mode image was built representing each event as a Gaussian filter kernel with $\sigma = A/(N_{ph})^{1/2}$, where A = 240 nm (experimentally determined value for our system), and N_{ph} is the number of photons. The FRC curves were calculated in 90 concentric rings (resulting in 90 frequency values in the FRC graphs) using the corresponding type of image representation for half-datasets[3]. For statistics, the FRC curves and the resolution values were calculated 50 times for each localization table, using different random separations of the dataset on two parts. Standard deviations of the obtained values are shown in error bar for each frequency point. The FRC curves for histogram, Gaussian and Voronoi representations were also calculated by the FSC program (Image Science Software GmbH)[33] confirming the results obtained with our method.

The simulated cluster data on Fig. 2 were generated as randomly distributed points in circles with a radius of 50 nm. The density of points in the clusters was $3 \cdot 10^{-3}$. The positions of clusters and of low-density ($4 \cdot 10^{-4}$) background points were distributed randomly in the field of view (FOV). The distributions of the sizes of the Voronoi cells were built as histograms of the areas of the cells, for a range from 0 to 4 times the median value of the area, using $2N^{1/3}$ bins, where N is the number of the polygons, excluding infinity-sized polygons at the boundaries of the ROI. For Monte-Carlo simulations, random coordinates were generated throughout the ROI using the 'rand' function. To obtain the confidence envelope, the distributions were generated for 50 different random sets of points, the boundaries of the envelope were determined as $<n> \pm 1.96\,\sigma$ for each bin of the histogram, where $<n>$ is the average number of cells within the range of the bin and σ is the standard deviation of n, calculated from the 50 random datasets. The abscissas of the first and the second intersections between the curves of the experimental and the mean value of the randomized distributions were determined from the two points around the intersections in the linear approximation. The Voronoi cells smaller than the area corresponding to the first intersection were kept, and all touching cells (those that have at least one common vertex) were combined together yielding vertices of clusters. The clusters were rasterized with the 'poly2mask' function using a pixel size of 1 nm. For the nuclear pore data a threshold was used to set the minimum number of events to remove small, spurious clusters. In our current implementation calculations for medium-sized datasets, e.g. Monte-Carlo simulations of $5 \cdot 10^4$ localizations, take around a minute; calculations may be parallelized in future software versions.

The simulated data on Figs S3–S6 were generated as 10×10 clusters with radius of 50 nm through a FOV of $4 \times 4\,\mu m$. The positions of the centers of the clusters were distributed regularly in the FOV. The points inside the clusters and in background were placed randomly with the indicated average density. For Fig. S3 the positions of the points were picked from the normal distribution using the mean defined at the previous step and the standard deviation of σ_{loc}. For Figs S4 and S5, n (number of re-localizations) points were picked for every original point using the normal distribution with the standard deviation $\sigma_{loc} = 10$ nm and the mean value at the position of the original point. For Monte-Carlo simulations on Fig. S5 the randomized datasets were formed using N/n_{sim} seeds at spatially random positions and distributing n_{sim} points by Gaussian distribution with $\sigma_{loc} = 10$ nm around the seeds; N is the number of points in the original dataset, n_{sim} is the number of re-localizations used in the Monte-Carlo simulations.

For segmented multi-color images, the co-localization value for a given species is defined by the ratio of the superposed area S between two colors relative to the total area of clusters of a given species. To compare the obtained values with overlapping area S_{rand} of randomly distributed clusters we shuffled the experimentallydetermined clusters into randomized positions through the same field of view, independently for each color. Firstly, the list of cluster polygons was permutated randomly using Matlab's 'randperm' function. Then, we determined the centroids of the clusters by formulas: $C_x = \frac{1}{6A}\sum_{i=0}^{n-1}(x_i + x_{i+1})(x_i y_{i+1} - x_{i+1} y_i)$, $C_y = \frac{1}{6A}\sum_{i=0}^{n-1}(y_i + y_{i+1})(x_i y_{i+1} - x_{i+1} y_i)$; where C_x, C_y are the (x, y) coordinates of the centroid of a given polygon, x_i, y_i are the coordinates of the vertices, n is the number of the vertices, and the vertex i = 0 equals to the vertex i = n. A is the signed area of the polygon: $A = \sum_{i=0}^{n-1}(x_i y_{i+1} - x_{i+1} y_i)$. After subtraction of the centroid coordinates from coordinates of the vertices of the polygons, the new coordinates were determined as a random number situated in the FOV of the initial size. To avoid overlapping of newly placed clusters, we iteratively checked for each cluster being placed at a random spot if it was not overlapping with the previously placed clusters of the same color, in which case another random coordinate was chosen. The shuffling procedure was repeated 50 times with different random positions, and the 95% confidence range for the co-localization in the random case was obtained as mean $\pm 1.96\,\sigma$ of the corresponding ratios of the surfaces.

Pair distribution functions were calculated using Matlab's 'pdist2'. For the nearest neighbor distance, the smallest pairwise distance value was chosen for each data point. The equivalent radius of a cluster was calculated as the radius of a circle with the same surface area as the cluster. Quantified properties of clusters (number of events, equivalent radius, nearest distance between neighboring clusters) are represented as mean ± standard deviation of the corresponding values.

References

1. Szymborska, A. *et al.* Nuclear Pore Scaffold Structure Analyzed by Super-Resolution Microscopy and Particle Averaging. *Science* **341,** 655–658 (2013).
2. Banterle, N., Khanh, H. B., Lemke, E. A. & Beck, M. Fourier ring correlation as a resolution criterion for super-resolution microscopy. *J. Struct. Biol.* **183,** 363–367 (2013).
3. Nieuwenhuizen, R. P. J. *et al.* Measuring image resolution in optical nanoscopy. *Nat. Methods* **10,** 557–562 (2013).
4. Baddeley, D., Cannell, M. B. & Soeller, C. Visualization of Localization Microscopy Data. *Microsc. Microanal.* **16,** 64–72 (2010).
5. Wang, Y. *et al.* Localization events-based sample drift correction for localization microscopy with redundant cross-correlation algorithm. *Opt. Express* **22,** 15982–15991 (2014).
6. Malkusch, S. *et al.* Coordinate-based colocalization analysis of single-molecule localization microscopy data. *Histochem. Cell Biol.* **137,** 1–10 (2012).
7. Rossy, J., Cohen, E., Gaus, K. & Owen, D. M. Method for co-cluster analysis in multichannel single-molecule localisation data. *Histochem. Cell Biol.* **141,** 605–612 (2014).

8. Ricci, M. A., Manzo, C., Garcıa-Parajo, M. F., Lakadamyali, M. & Cosma, M. P. Chromatin Fibers Are Formed by Heterogeneous Groups of Nucleosomes *in Vivo*. *Cell* **160,** 1145–1158 (2015).
9. Ehmann, N. *et al.* Quantitative super-resolution imaging of Bruchpilot distinguishes active zone states. *Nat. Commun.* **5,** 4650 (2014).
10. Pageon, S. V. *et al.* Superresolution Microscopy Reveals Nanometer-Scale Reorganization of Inhibitory Natural Killer Cell Receptors upon Activation of NKG2D. *Sci. Signal.* **6,** ra62 (2013).
11. Gao, J. *et al.* Mechanistic insights into EGFR membrane clustering revealed by super-resolution imaging. *Nanoscale* **7,** 2511–2519 (2015).
12. Veatch, S. L. *et al.* Correlation functions quantify super-resolution images and estimate apparent clustering due to over-counting. *Plos One* **7,** e31457 (2012).
13. Sengupta, P. *et al.* Probing protein heterogeneity in the plasma membrane using PALM and pair correlation analysis. *Nat. Methods* **8,** 969–975 (2011).
14. Lemaître, C. *et al.* Nuclear position dictates DNA repair pathway choice. *Gene. Dev*. **28,** 2450–2463 (2014).
15. Ester, M., Kriegel, H.-P., Jörg, S. & Xu, X. A density-based algorithm for discovering clusters in large spatial databases with noise. Paper presented at *Second International Conference on Knowledge Discovery and Data Mining*, Portland, Oregon. AAAI (1996 August 2–4).
16. Rubin-Delanchy, P. *et al.* Bayesian cluster identification in single-molecule localization microscopy data. *Nat. Methods* **12,** 1072–1076 (2015).
17. Levet, F. *et al.* SR-Tesseler: a method to segment and quantify localization-based super-resolution microscopy data. *Nat. Methods* **12,** 1065–1071 (2015).
18. Voronoi, G. Nouvelles applications des paramètres continus à la théorie des formes quadratiques. *J. Reine Angew. Math.* **133,** 97–178 (1908).
19. Aurenhammer, F. Voronoi Diagrams – A Survey of a Fundamental Geometric Data Structure. *ACM Comput. Surv.* **23,** 345–405 (1991).
20. Felfer, P., Ceguerra, A., Ringer, S. & Cairney, J. Detecting and extracting clusters in atom probe data: A simple, automated method using Voronoi cells. *Ultramicroscopy* **150,** 30–36 (2015).
21. Egner, A. & Hell, S. W. Aberrations in confocal and multi-photon fluorescence microscopy induced by refractive index mismatch in *Handbook of Biological Confocal Microscopy* (ed. Pawley, J. B.) 404–412 (Springer, New York, 2006).
22. Betzig, E. *et al.* Imaging Intracellular Fluorescent Proteins at Nanometer Resolution. *Science* **313,** 1642–1645 (2006).
23. Ahuja, N. Dot pattern processing using Voronoi Neighborhoods. *IEEE T. Pattern Anal.* **PAMI-4,** 336–343 (1982).
24. Maul, G. G. Nuclear pore complexes. Elimination and reconstruction during mitosis. *J. Cell Biol.* **74,** 492–500 (1977).
25. Winey, M., Yarar, D., Giddings, J., H., T. & Mastronarde, D. N. Nuclear pore complex number and distribution throughout the Saccharomyces cerevisiae cell cycle by three-dimensional reconstruction from electron micrographs of nuclear envelopes. *Mol. Biol. Cell.* **8,** 2119–2132 (1997).
26. Baer, B. W. & Rhodes, D. Eukaryotic RNA polymerase II binds to nucleosome cores from transcribed genes. *Nature* **301,** 482–488 (1983).
27. Louters, L. & Chalkley, R. Exchange of histones H1, H2A, and H2B *in vivo*. *Biochemistry* **24,** 3080–3085 (1985).
28. Hamiche, A. *et al.* Interaction of the histone (H3-H4)2 tetramer of the nucleosome with positively supercoiled DNA minicircles: Potential flipping of the protein from a left- to a right-handed superhelical form. *Proc. Natl. Acad. Sci. USA* **93,** 7588–7593 (1996).
29. Reem, D. The geometric stability of Voronoi diagrams with respect to small changes of the sites. Paper presented at *27th Annual ACM Symposium on Computational Geometry (SoCG)*, Paris, France. Inria (2011 June 13–15).
30. Andronov, L., Lutz, Y., Vonesch, J.-L. & Klaholz, B. P. SharpViSu: integrated analysis and segmentation of super-resolution microscopy data. *Bioinformatics*, doi: 10.1093/bioinformatics/btw123 (2016).
31. Olivier, N., Keller, D., Rajan, V. S., Gönczy, P. & Manley, S. Simple buffers for 3D STORM microscopy. *Biomed. Opt. Express* **4,** 885–899 (2013).
32. Sibson, R. A brief description of natural neighbour interpolation in *Interpreting multivariate data* (ed. Barnet, V.) 21–36 (Wiley, Chichester, 1981).
33. van Heel, M., Harauz, G. & Orlova, E. V. A new generation of the IMAGIC image processing system. *J. Struct. Biol.* **116,** 17–24 (1996).

Acknowledgements

We thank Alexander Myasnikov for discussions and Remy Fritz and Jonathan Michalon for IT support. This work was supported by the European Research Council (ERC Starting Grant N_243296 TRANSLATIONMACHINERY), the Centre Nationale pour la Recherche Scientifique (CNRS), Association pour la Recherche sur le Cancer (ARC), Université de Strasbourg and Investissement d'Avenir (IDEX) and Institut National du Cancer (INCa). The super-resolution microscope setup is supported by the Alsace Region and by the French Infrastructure for Integrated Structural Biology (FRISBI) ANR-10-INSB-05-01, and Instruct as part of the European Strategy Forum on Research Infrastructures (ESFRI).

Author Contributions

L.A. wrote the software and performed imaging, L.A., I.O., J.-L.V. and B.P.K. worked on the concept, Y.L. performed fluorescence labelling, L.A., J.-L.V. and B.P.K. analysed the data and wrote the manuscript. All authors reviewed the manuscript.

Additional Information

Supplementary information accompanies this paper at http://www.nature.com/srep

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Andronov, L. *et al.* ClusterViSu, a method for clustering of protein complexes by Voronoi tessellation in super-resolution microscopy. *Sci. Rep.* **6**, 24084; doi: 10.1038/srep24084 (2016).

This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/

ClusterViSu, a method for clustering of protein complexes by Voronoi tessellation in super-resolution microscopy

Leonid Andronov[1,2,3,4], Igor Orlov[1,2,3,4], Yves Lutz[1,2,3,4], Jean-Luc Vonesch[1,2,3,4], Bruno P. Klaholz[1,2,3,4,*]

[1]Centre for Integrative Biology (CBI), Department of Integrated Structural Biology, IGBMC (Institute of Genetics and of Molecular and Cellular Biology), 1 rue Laurent Fries, Illkirch, France; [2]Centre National de la Recherche Scientifique (CNRS) UMR 7104, Illkirch, France; [3]Institut National de la Santé et de la Recherche Médicale (INSERM) U964, Illkirch, France; and [4]Université de Strasbourg, Strasbourg, France.

*To whom correspondence should be addressed. klaholz@igbmc.fr

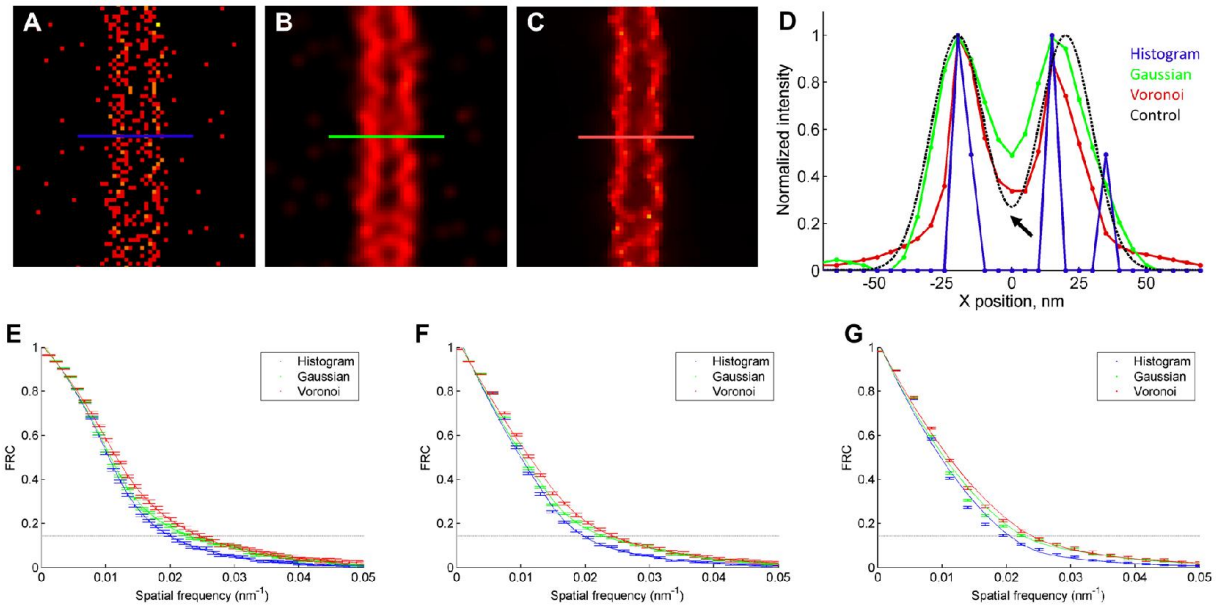


Figure S1 Evaluation of the quality of Voronoi image reconstruction

(**A-C**) Simulated images of two lines separated by 40 nm and composed of localization events with a standard deviation of 10 nm in histogram (**A**), Gaussian (**B**) and Voronoi-based interpolated (**C**) local density representations (same as **Figs. 1A-C**, but with line indicating the line profile). (**D**) One-pixel width profiles along the colored lines of images **A-C**. (**E-G**) FRC curves calculated from an image region in **Fig. 1E** in histogram (blue), Gaussian (green) and Voronoi local density (red) representations; pixel size of the image: 5 nm (**E**), 3 nm (**F**), 2nm (**G**). The graphs show only the frequency range from 0 to 0.05 nm^{-1} for consistency with **Fig. 1F**.

Suppl. Fig. S1

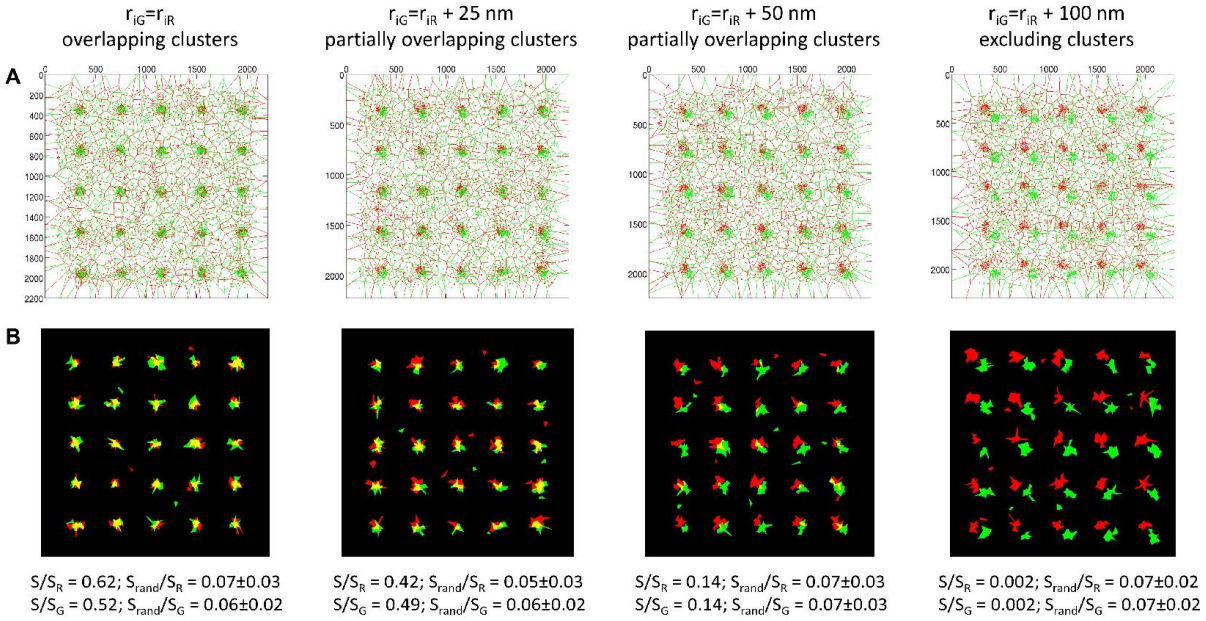


Figure S2 Simulations of different co-localization scenarios

(**A**) Voronoi diagrams of simulated double-color clusters. (**B**) Results of segmentation of the double-color clusters in ClusterViSu with the overlay area shown in yellow and the calculated co-localization values shown below. r_{iG} and r_{iR} are the Cartesian coordinates of the centers of the green and the red clusters, respectively. The simulated radius of clusters is 50 nm; density of events in clusters = $2.5 \cdot 10^{-3}$ nm^{-2}, in background = $1.25 \cdot 10^{-4}$ nm^{-2}. Localizations in the clusters and in the background are distributed randomly and independently for each color.

Suppl. Fig. S2

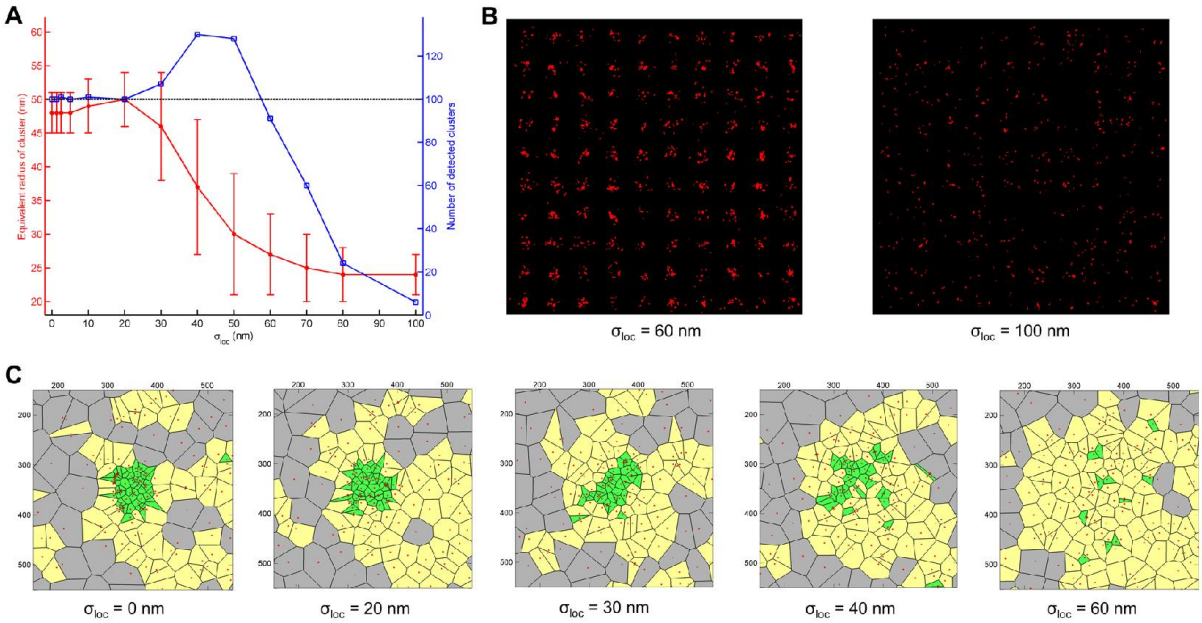


Figure S3 Effect of the localization error on Voronoi clustering

(**A**) Results of segmentation of a simulated cluster data with different amounts of localization error. σ_{loc} is the standard deviation of the normal distribution used for simulating the coordinates of the points. Clusters containing 5 or less molecules were removed. (**B**) Complete segmented FOVs of the simulated data with high localization errors σ_{loc} demonstrating remaining clusters at σ_{loc} = 60 nm and σ_{loc} = 100 nm. (**C**) The Voronoi diagrams of the upper-left corner of the simulated data with different localization errors. The density of clustered localizations is 0.01 nm^{-2}, the density of background is 0.002 nm^{-2}.

Suppl. Fig. S3

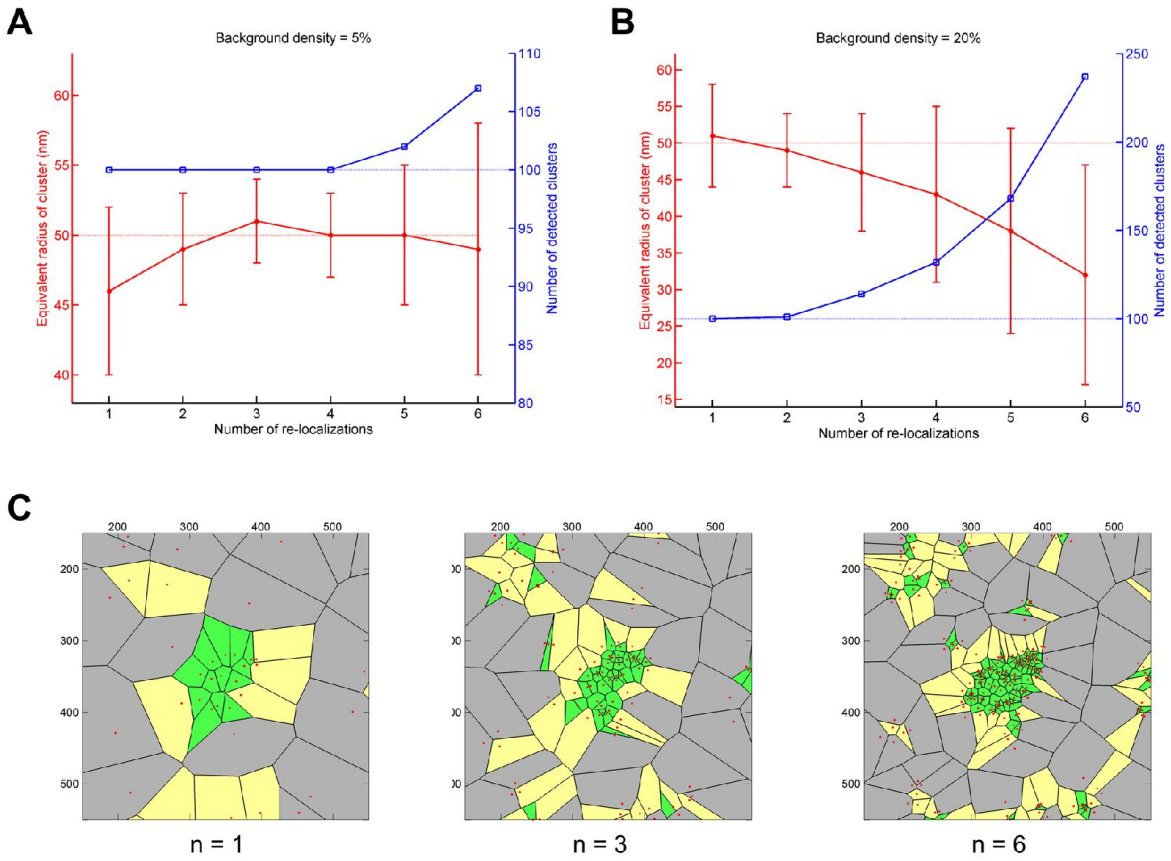


Figure S4 Effect of multiple localizations on Voronoi clustering

(**A&B**) Results of segmentation of simulated cluster data with different numbers of re-localizations. Clusters containing 5 or less molecules were removed. (**C**) Comparison of Voronoi diagrams with different numbers of re-localizations (n). The density of clustered localizations (at n = 1) is 0.0025 nm^{-2}), the density of background is 20% of the cluster density; the localization precision σ_{loc} = 10 nm.

Suppl. Fig. S4

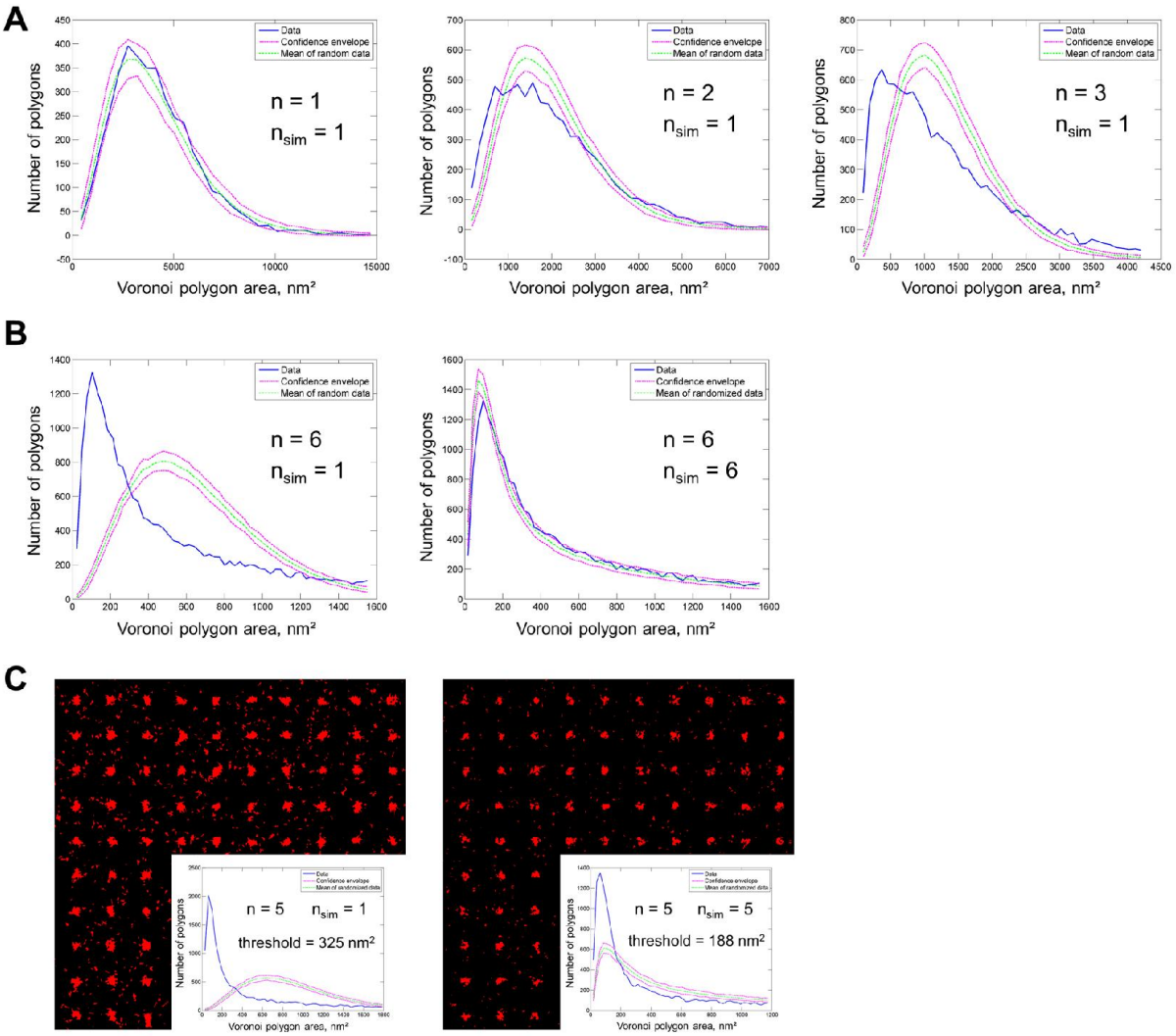


Figure S5 Effect of multiple localizations on Monte-Carlo simulations

(**A&B**) Voronoi polygon distributions of spatially random datasets with different numbers of re-localizations (n) and results of Monte-Carlo simulations taking into account multiple localizations (n_{sim}). (**C**) Simulated clustered datasets with re-localizations (n = 5) segmented by thresholds determined by simulations without (n_{sim} = 1) and with (n_{sim} = 5) multiple localizations. The density of localizations (at n = 1) is 0.001 nm^{-2} (**A&B**); the density of clustered localizations (at n = 1) is 0.0025 nm^{-2}, the density of background is 20% of the cluster density (**C**); the localization precision σ_{loc} = 10 nm (**A-C**).

Suppl. Fig. S5

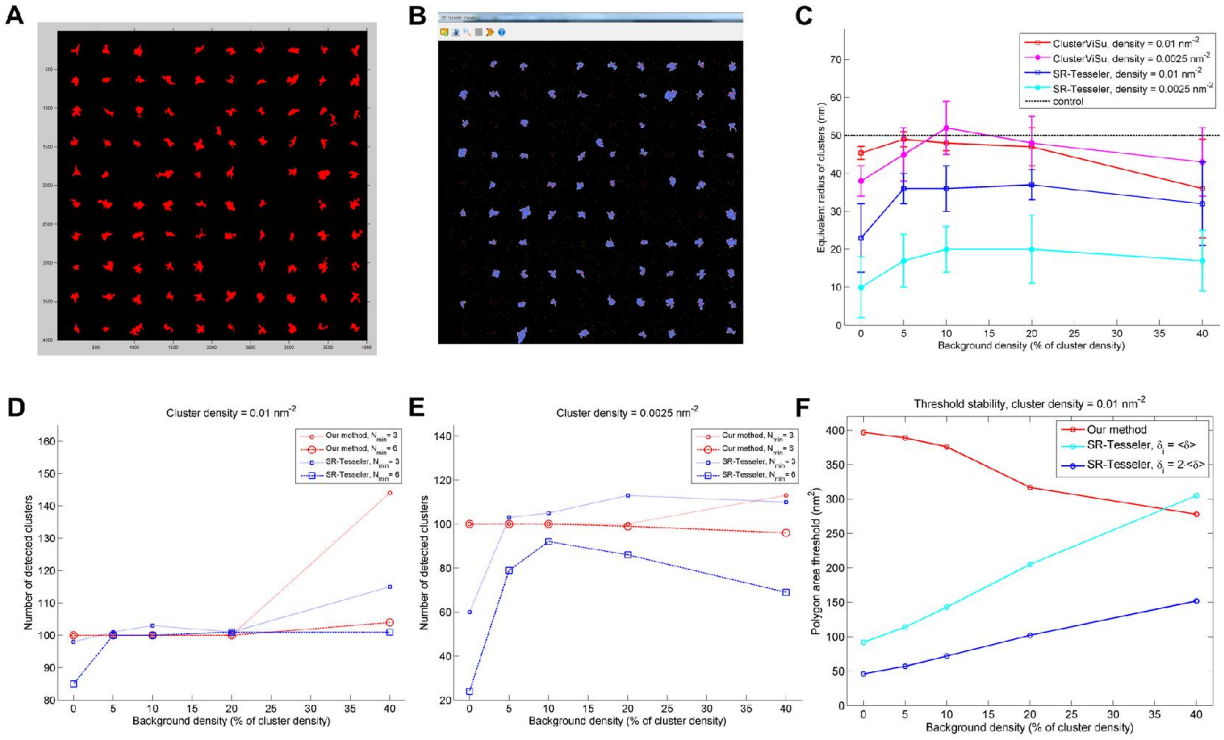


Figure S6 Comparison of ClusterViSu with SR-Tesseler.

(**A&B**) Clustered dataset segmented in ClusterViSu with threshold determined by Monte-Carlo simulations (**A**) and in SR-Tesseler with threshold determined by average localization density[17] $\delta^1_i = 2\cdot\langle\delta\rangle$ (**B**). Density of clustered localization is 0.0025 nm^{-2}, background density is 0.01 nm^{-2}. Clusters containing five or less molecules were removed. (**C**) Retrieval of the equivalent radius of clusters in ClusterViSu and SR-Tesseler. Clusters containing two or less molecules were removed. (**D&E**) Number of clusters detected by ClusterViSu (red curves) and SR-Tesseler (blue curves) as function of background density. N_{min}, minimal number of events in clusters kept for quantification. (**F**) Variation of segmentation threshold determined by Monte-Carlo simulations in ClusterViSu (red curve) or by average localization density in SR-Tesseler[17] (cyan and blue curves), depending on background density.

Suppl. Fig. S6

3.3. 3DClusterViSu, a method for 3D cluster analysis of SMLM data

In the third part of my research project we explored possibilities for segmentation of 3D SMLM data. I present the results below in the form of a manuscript submitted to Bioinformatics. The manuscript was also deposited on bioRxiv (Andronov et al., 2017).

In this work, first we emphasize the necessity of 3D experiments and 3D data processing by showing that processing of 3D information in only two dimensions can substantially distort the results or even make them meaningless. A similar situation may happen, if one tries to analyze, *e.g.* bulk chromatin in the nucleus with a 2D super-resolution method.

Then we describe the extension of the method reported in section 3.2, which is based on Voronoi tessellations of 3D data. Here we use 3D Voronoi diagrams that are more complex than 2D diagrams, *e.g.* the number of vertices of a polyhedron in a 3D Voronoi diagram is about 4.5 times bigger than the number of vertices of a polygon in a 2D Voronoi diagram (Okabe et al., 2000). We show that, despite the different geometry of 3D diagrams, the Monte-Carlo-based automatic threshold determination concept proposed by us for the 2D case, works as well for 3D data. We also note that visualization of data as density maps can be especially useful for 3D as it allows to analyze these data directly with software for 3D image processing.

We implemented the Voronoi diagram based method for segmentation of 3D SMLM data in a software, 3DClusterViSu. We also show how this method can be used for noise reduction. We validate the method with synthetic data and apply it for processing of super-resolution data of tubulin and histones H2B and CENP-A. The clustering of H2B determined by our method is consistent with that previously reported (Ricci et al., 2015). Finally, we show that CENP-A forms clusters in the centromeric regions of the cell nucleus and we determine the size of these clusters by using 3DClusterViSu.

3.3.1. Publication 3 “3DClusterViSu: 3D clustering analysis of super-resolution microscopy data by 3D Voronoi tessellations”

Bioinformatics, YYYY, 0–0
doi: 10.1093/bioinformatics/xxxxx
Advance Access Publication Date: DD Month YYYY
Manuscript Category

Bioimage informatics

3DClusterViSu: 3D clustering analysis of super-resolution microscopy data by 3D Voronoi tessellations

Leonid Andronov[1,2,3,4], Jonathan Michalon[1,2,3,4], Khalid Ouararhni[1,2,3,4], Igor Orlov[1,2,3,4], Ali Hamiche[1,2,3,4], Jean-Luc Vonesch[1,2,3,4] and Bruno P. Klaholz[1,2,3,4,*]

[1]Centre for Integrative Biology (CBI), Department of Integrated Structural Biology, IGBMC (Institute of Genetics and of Molecular and Cellular Biology), 1 rue Laurent Fries, Illkirch, France, [2]Centre National de la Recherche Scientifique (CNRS) UMR 7104, Illkirch, France, [3]Institut National de la Santé et de la Recherche Médicale (Inserm) U964, Illkirch, France and [4]Université de Strasbourg, Strasbourg, France.

*To whom correspondence should be addressed.

Associate Editor: XXXXXXX

Received on XXXXX; revised on XXXXX; accepted on XXXXX

Abstract

Motivation: Single-molecule localization microscopy (SMLM) can play an important role in integrated structural biology approaches to identify, localize and determine the 3D structure of cellular structures. While many tools exist for the 3D analysis and visualization of crystal or cryo-EM structures little exists for 3D SMLM data, which can provide unique insights but are particularly challenging to analyze in three dimensions especially in a dense cellular context.

Results: We developed 3DClusterViSu, a method based on 3D Voronoi tessellations that allows local density estimation, segmentation & quantification of 3D SMLM data and visualization of protein clusters within a 3D tool. We show its robust performance on microtubules and histone proteins H2B and CENP-A with distinct spatial distributions. 3DClusterViSu will favor multi-scale and multi-resolution synergies to allow integrating molecular and cellular levels in the analysis of macromolecular complexes.

Availability: 3DClusterViSu is available under http://cbi-dev.igbmc.fr/cbi/voronoi3D

Contact: andronov@igbmc.fr & klaholz@igbmc.fr

Supplementary information: Supplementary figures are available at *Bioinformatics* online.

1 Introduction

Structure-function studies of macromolecular complexes are increasingly moving towards cellular structural biology, which requires developing multi-resolution and multi-scale correlative approaches to address the molecular and cellular organization of living cells (Wolff et al., 2016; Orlov et al., 2017). The strength of such strategies is illustrated for example by correlative light and electron microscopy (CLEM) approaches (Koning et al., 2014; Schirra and Zhang, 2014; Karreman et al., 2016; Schorb et al., 2017; Arnold et al., 2016). Methods such as protein X-ray crystallography, NMR and cryo electron microscopy (cryo-EM) have had a strong impact over the past decades thanks to their ability to analyze and visualize macromolecules directly in 3D. These well-established 3D techniques encompass dedicated software tools to analyze and also visualize 3D structures, while no equivalent convenient tools exist up to now in the field of single-molecule localization microscopy (SMLM). SMLM includes techniques such as stochastic optical reconstruction microscopy (STORM (Huang et al., 2008)) or photo-activated localization microscopy (PALM (Betzig et al., 2006)) and gives access to the properties of each single fluorescent molecule. This allows precise determination of their lateral (X,Y) coordinates (Thompson et al., 2002; Fölling et al., 2008), spectral demixing (Lampe et al., 2012) and single-molecule FRET (Winckler et al., 2013), but also the possibility of determining the axial (Z) position of the fluorophores with sub-diffraction precision. Methods for 3D SMLM include bi-plane detection (Juette et al., 2008) or modifications of the point spread function (PSF) through either astigmatism (Huang et al., 2008) or double-helix PSF (Pavani et al., 2009), or the 4Pi optical setup (Aquino et al., 2011).

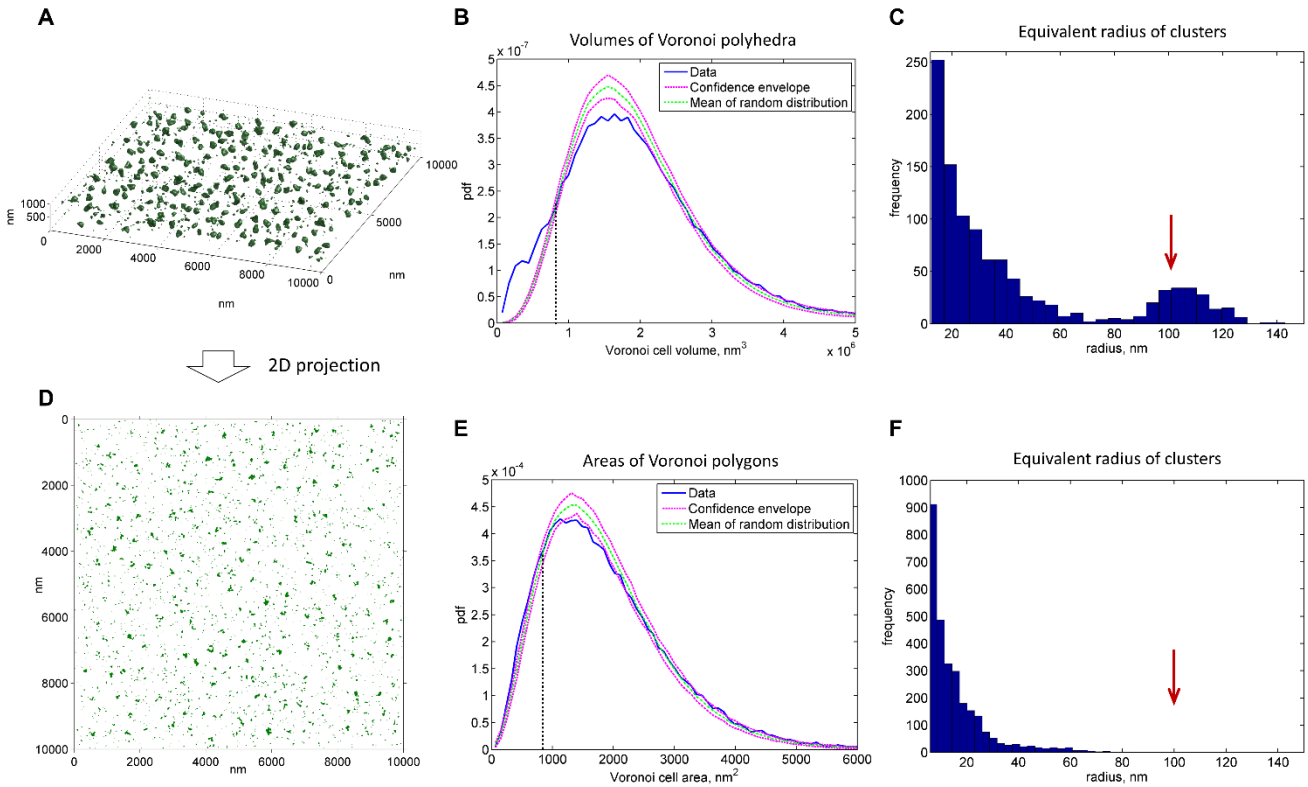


Fig. 1. Importance of 3D cluster analysis of localization data. Clusters in a 3D dataset (number of clusters = 200, density of localizations within clusters = $5{\cdot}10^{-6}$ nm^{-3}, in the background = $5{\cdot}10^{-7}$ nm^{-3}) can be easily segmented using the 3D Voronoi diagram approach described in the current work (A-C). The same dataset projected on the X-Y plane becomes close to a random distribution of points (E) and its processing with a 2D segmentation method (Andronov, Orlov, et al., 2016) does not allow to determine the correct shape, size and number of clusters (D, F, S1). In the left column are the clusters determined with the 3D Voronoi approach described in the current work (A) or with our previously reported 2D Voronoi approach (Andronov, Orlov, et al., 2016) (D). In the middle column are the probability density functions (pdf) of the volumes (B) or the surface areas (E) of Voronoi cells for the processed datasets (blue) in comparison with that of the background model, a spatially random distribution (magenta and green). In the right column (C, F) are the histograms of the equivalent radius of the detected clusters, to be compared with the simulated cluster radius of 100 nm (red arrows).

Recently, it also became possible to get sub-diffraction 3D information from 2D data using a photometry approach (Franke et al., 2017) or by fitting with experimental PSFs (Li et al., 2017). Despite experimental possibilities for acquiring 3D SMLM data, methods for processing these data in 3D are not well developed. For 2D SMLM data, means for visualization (Baddeley et al., 2010; Andronov, Orlov, et al., 2016), co-localization analysis (Malkusch et al., 2012; Andronov, Orlov, et al., 2016) and segmentation (Williamson et al., 2011; Rossy et al., 2014; Rubin-Delanchy et al., 2015; Pageon et al., 2016) have been reported. Data processing methods based on tessellations, namely on 2D Voronoi diagrams and Delaunay triangulations, have advantages over other approaches because they allow for efficient visualization (Andronov, Orlov, et al., 2016), unambiguous local density estimation (Andronov, Orlov, et al., 2016; Levet et al., 2015), noise reduction and multi-scale segmentation (Levet et al., 2015). Even though most of these approaches are described as potentially extendable to the 3D case, to date only DBSCAN (Ester et al., 1996; Barna et al., 2016) and Getis and Franklin's local point pattern analysis have been used for 3D segmentation of localization data (Owen et al., 2013; Beheiry and Dahan, 2013). For three-dimensional visualization, the ViSP software (Beheiry and Dahan, 2013) exists, which allows some basic segmentation with a rather simple estimation of the local density defined as the number of points inside a circle of a fixed radius. Other methods such as Bayesian clustering (Rubin-Delanchy et al., 2015) and Voronoi-based segmentations in fact remain 2D approaches as neither SR-Tesseler (Levet et al., 2015) nor ClusterViSu (Andronov, Orlov, et al., 2016) can process 3D data.

Processing data directly in 3D rather than in 2D turns out to be crucial to avoid artefacts due to the enhanced complexity in 3D as compared to the 2D case, especially in a dense cellular context. An analysis in 2D would lead to an overlap between neighboring structures which are otherwise separate in 3D space. Using test data with clusters distributed randomly in a 3D space with noise, we show that a simple 2D clustering analysis of the data can lead to unprecise results with erroneous size estimations (Figs. 1, S1). With increasing density of clusters and of background noise, the point distributions obtained for a 2D analysis become close to random due to the overlay of 3D structures in a 2D image (Figs. 1D-F, S1C), while 3D segmentation of the same data is robust and retrieves the clusters properly (Fig. 1A-C, 2C-H, S1A). This suggests that in order to avoid information loss and instead precisely determine properties of 3D objects which are labelled within a cell or a tissue, it is essential to acquire and also segment and represent the data in three dimensions, especially in crowded environments such as chromatin. In the following we describe the concept of three-dimensional Voronoi-based cluster analysis, which has never been applied to 3D SMLM data before and which carries interesting features useful for 3D analysis and 3D representation. We evaluate the performance of this method with synthetic and experimental data and provide examples of visualization and quantitative analysis of 3D SMLM data.

2 Methods

2.1 Cell culture and immunofluorescence

For the tubulin and H2B samples, HeLa cells were cultured in a 4-compartment glass-bottom petri dish (CELLView, Greiner Bio-One), washed in PBS and fixed with 4% formaldehyde for 20 min in phosphate-buffered saline solution (PBS). After permeabilization with 0.1% Triton X-100 in PBS (PBS/Tx) twice for 10 min, the primary antibody (anti-β-tubulin monoclonal (1Tub-2A2, in house IGBMC) used as mouse ascites fluid diluted 500x in PBS/Tx; histone H2B monoclonal antibody (LG11-2) used as a 500x dilution of mouse ascites fluid in PBS/Tx) was incubated overnight at 4 °C. The sample was then washed with PBS/Tx three times over 30 minutes, and the secondary antibody (goat anti-mouse Alexa Fluor-647) in dilution 4 µg/ml in PBS/Tx was incubated for 1 to 2 hours at room temperature (RT). Subsequently, the cells were washed in PBS/Tx three times for 30 minutes, then briefly three times in PBS. The samples were mounted in a PBS buffer with addition of 10 mM of cysteamine (also known as β-mercaptoethylamine or MEA) and 25 mM of HEPES (pH 7.5).

For the CENP-A samples, U2OS cells were synchronized using Thymidine-Nocodazole synchronization. Briefly, the cells were plated and 24 hours later the medium was replaced with prewarmed complete DMEM supplemented with 2.5 mM Thymidine for 18 hours. The cells were washed with DMEM and 100 ng/ml Nocodazole was added for 12-14 hours. Floating cells were collected and spin down washed twice with prewarmed DMEM and resuspended in prewarmed DMEM. Cell were collected after 8 hours, washed twice in PBS and fixed in 2% paraformaldehyde for 15 min at RT, washed twice in PBS, permeabilized with PBS/Tx for 15 min, washed twice in PBS, washed five times in PBS+0.5% BSA (PBB), blocked with 10% serum in PBB for 1 hour, washed five times in PBB, incubated with primary antibodies (EMD Millipore 07-574) in 400x dilution for 1h, washed five times in PBB, incubated with secondary antibodies in 400x dilution for 1h, washed five times with PBB and three times with PBS. The sample was mounted in an imaging buffer (Olivier et al., 2013) that contained 20% of Vectashield (Vector Laboratories), 70% of 2,2`-thiodiethanol (also known as thiodiglycol or TDE) and 10% of PBS 10x (the measured refractive index of this mounting medium is 1.49). The mounting was performed by incubating the sample in PBS solutions with gradually increased concentrations of TDE (10%, 25% and 50%), for 10 min each.

2.2 Super-resolution imaging

The super-resolution imaging of the β-tubulin and the H2B samples was performed on a Leica SR GSD 3D system in Wetzlar, Germany. The system was equipped with the HC PL APO 160x/1.43 Oil CORR GSD objective and the Andor iXon Ultra 897 EMCCD camera with a field of view of 18x18 µm in the high-power mode. Continuous wave fiber laser (MPBC Inc., 642 nm 500 mW) and a diode laser (405 nm 30 mW) were utilized for excitation. The objective was linked to the sample with help of a suppressed motion (SuMo) sample stage, which reduces lateral and axial drift. The astigmatism was induced with the cylindrical lens and calibrated using gold beads and the automatic built-in procedures.

The super-resolution imaging of the CENP-A samples was performed on our in-house Leica SR GSD system. We used the HCX PL APO 100x/1.47 Oil CORR TIRF PIFOC objective with a 1.6x magnification lens that provides an equivalent pixel size of 100 nm on the camera. The camera, the lasers and the stage were the same as on the Leica SR GSD 3D system. The astigmatism was induced with a MicAO 3DSR adaptive optics system installed between the microscope stage and the camera (the value of the astigmatism was set to 0.2 µm root mean square). The deformation of the PSF was calibrated by defocusing the objective with a step of 50 nm and imaging Tetraspeck fluorescent beads with diameter of 200 nm.

The samples were first illuminated with the 100% power of the 642 nm laser to quickly send the fluorophores into the dark state. The acquisition was started manually after observing first single-fluorophore events ("blinking"). The time of exposition of a frame was 6.9 ms (H2B), 10 ms (β-tubulin) or 25 ms (CENP-A); the electron multiplying gain of the camera was 300; the laser power during the acquisition was 28% (H2B) or 100% (β-tubulin and CENP-A). After a few minutes, as the number of blinking events dropped, the sample started to be illuminated additionally by a 405 nm laser with gradual increase of its intensity in order to keep a nearly constant rate of single-molecular returns into the ground state. The acquisition was stopped after almost complete bleaching of the fluorophores, lasting 9.15 min and producing 50419 frames (β-tubulin dataset); 3.88 min, 29913 frames (H2B dataset); ~12 min, ~28000 frames (CENP-A datasets).

2.3 Data processing

The localization and fitting of single-molecule events were performed in real time during acquisitions in Leica LAS AF software with the "direct fit" fitting method. The localization tables were then exported for further processing in the SharpViSu software workflow (Andronov, Lutz, et al., 2016) and with customized Matlab procedures. To reduce the number of localizations of the same fluorophore and improve localization precision the data were processed by averaging the coordinates of consecutive events within a radius of 50 nm around each localization. The drift was detected and corrected in three dimensions using cross-correlation-based approach. Briefly, the dataset was divided into several consecutive subsets, from each of them a histogram image with pixelation of 20 nm was build; next, the shift between these images was detected with subpixel precision and then interpolated linearly throughout intermediate frames. The shift value was then subtracted from the coordinates of every frame. The lateral drift was detected using the projection on the XY plane, and the axial drift was detected using the average shift, calculated from the XZ and YZ projections. The procedure was repeated iteratively to assure absence of detectable residual drift (Andronov, Lutz, et al., 2016). Regions of interest (ROI) for Voronoi analysis were selected manually allowing faster computations and more homogeneous distributions compared to the entire field of view (this avoids selecting regions with non-uniform background). Voronoi diagrams prevent edge effects because data points close to borders have big or infinite-sized Voronoi polyhedrons, thus local densities are low and do not contribute to clusters.

The following analysis was performed in Matlab using customized code (3DClusterViSu). Voronoi diagrams (vertices of polyhedrons and connectivity order) were retrieved with the 'voronoin' function. Volumes of the cells were determined from the vertices with the function 'convhulln' (Barber et al., 1996), the local density in each data point was defined as the inverse value of the volume of the corresponding Voronoi polyhedron. To account for non-isotropic distribution of experimental data points in Z-direction we introduced a correction factor for the calculated local density. Indeed, the intensity of the PSF diminishes while going out of focus and so does the number of localizations if one performs a 3D SMLM experiment using a PSF shape modifying method. Consequently, the local densities will be stronger in the center of the 3D volume compared to top and bottom regions in Z direction. To account for this distortion, we approximated the decreasing number of events in the axial direction as a Gaussian function with the fixed standard deviation $\sigma = 300$ nm, which corresponds to the axial resolution of the objective $R_z = 1.5 \cdot \lambda \cdot n/NA^2$, where λ is the wavelength of the detected light, n

is the refractive index of the sample and NA is the numerical aperture of the objective (Fig. S2). The densities associated with the experimental localizations were then multiplied by a factor proportional to the inverse value of the Gaussian function at the Z-position of the fluorophore. To have a smooth appearance for visualization or segmentation the values of the corrected local densities were interpolated to a regular grid (voxels) using the 'griddata' function and the 'natural' interpolation method (Sibson, 1981) that is also based on Voronoi diagrams. The spacing of the grid corresponds to the desired voxel size (20 nm in our case). Voronoi diagrams are used for quantifying densities and for Monte Carlo calculations, while the segmentation step is done on density maps obtained by interpolation to avoid clusters with star-like Voronoi polyhedron shapes (see Fig. 2F & 4F). The density maps can be readily exported to TIFF or MRC formats for further processing or visualization.

For processing of experimental data acquired with astigmatism, we used the following background model in the Monte-Carlo simulations: points with uniformly distributed random X, Y coordinates (the 'rand' function), and with normally distributed random Z coordinates picked from a Gaussian distribution with the standard deviation $\sigma = 300$ nm (the 'normrnd' function). The points were distributed over the manually defined field of view (FOV). The distributions were generated for 50 different random sets of points, the boundaries of the 95% confidence envelopes were determined as $\langle n \rangle \pm 1.96 \cdot \sigma$ for each bin of the histogram, where $\langle n \rangle$ is the average number of cells within the range of the bin and σ is the standard deviation of n, calculated from the 50 random datasets. The segmentation threshold, defined as the abscissa of the first intersection between the curves of the experimental and the mean value of the randomized distributions was determined from the two points around the intersection in the linear approximation.

For visualization of the β-tubulin data (Fig. 3), the 2D projections (Figs. 3E-F) were built as histogram images with the bin size of 20 nm. The signal-to-noise ratios were calculated as $\langle I \rangle/\sigma_I$, where $\langle I \rangle$ is the mean value of the non-zero pixels and σ_I is the standard deviation of all pixel values, yielding $\langle I \rangle/\sigma_I = 1.6$ (Fig. 3E) and $\langle I \rangle/\sigma_I = 2.2$ (Fig. 3F).

The simulated cluster data in Figs. 2C-H, S3 were generated as randomly distributed points inside spheres with the radius of 30 nm. The positions of clusters and of background points were distributed randomly in the FOV, with the following constraints: 1) distance from the centers of the clusters to the borders ≥ 75 nm; 2) distance between the centers of the clusters ≥ 150 nm. The borders of the clusters were smoothened such that the density of molecules at the distance r from the center of the cluster was modulated with the function $p(r) = ½(1+\mathrm{erf}(4\cdot(r_0-r)/r_0))$, where $r_0 = 30$ nm and erf(x) is the error function.

Even though the distribution of Voronoi cell areas of the 2D-projected dataset in Fig. S1C lies within the confidence envelope for the background model indicating no clustering, we segmented this dataset keeping only the regions with the local density stronger than the mean density calculated from the 2D Voronoi diagram (Levet et al., 2015). The resulting clusters, as expected, have random sizes much smaller than the original radius of 30 nm (Fig. S1C, left and right panels).

We have also implemented the averaging of the local densities using the first-rank (Levet et al., 2015) neighbors in 3D (scripts available under http://cbi-dev.igbmc.fr/cbi/first-rank-voronoi3D/. This option helps to remove spurious clusters in the background but it also leads to reduction of the volume of the detected clusters because the density of the near-border regions within clusters is averaged with the low-dense background (Fig. S4). By contrast, simple removal of small clusters as done by default (Fig. 2H) does not affect the shape of big clusters. Although not done here, it is possible to separate closely located clusters through methods such as watershed transformation, which are compatible with our clustering method and can be performed after binarization of the density map. Similarly, if experimentally evidenced, the background noise model could be adjusted to include re-localization events arising from molecules blinking several times.

For the CENP-A data, after the segmentation, centromeres touching each other and small clusters in the background were removed from the analysis. To determine the size of the CENP-A clusters, 5 different ROIs with 200 well separate centromeres in total were analyzed. The equivalent radius (diameter) of clusters was calculated as the radius (diameter) of the sphere with the same volume as the clusters. Quantified properties of clusters (number of events, equivalent radius or diameter) are represented as mean ± standard deviation of the corresponding values.

The standalone software for visualization of 3D Voronoi tessellations was written in Python using the Mayavi data visualization library (Ramachandran and Varoquaux, 2011) along with NumPy (Walt et al., 2011) and SciPy (Jones et al., 2001). SMLM datasets are loaded as a 3D CSV point list file. The user can then navigate through the 3D Voronoi cells determined by the input points. The cells are colored according to their size, from blue (big cells) to red (small cells). The user can set a cell size threshold to see only smaller cells where events are aggregated, use a clipping box to see inside the cells and toggle the drawing of the event points. The size of the cells is computed in parallel using QHull's "ConvexHull" function (Barber et al., 1996) from SciPy. Cells delimited by vertices outside of the event range are merely ignored. The Mayavi interface allows the user to change a lot of drawing parameters on-the-fly. Colors, opacity, orientations, lightning, clipping shape etc. are all tunable through the built-in parameter adjustment wizard.

As an example, it takes 6 hours to segment a large experimental dataset with a volume of ~200 μm^3 that contains 10^6 localizations using an i7 quad-core single processor with 32 Gb of memory. This is much faster in a single run than what could be done for example with the current 3D implementation of DBSCAN (Ester et al., 1996) in VividSTORM (Barna et al., 2016) where one single run takes 8 hours for 67 000 points (which is a rather small data set). In contrast, our Voronoi-based segmentation uses an automatically determined threshold value and takes in total 6 hours for 1 000 000 points (15 times as many) processed in one single run, i.e. it is more robust, non-dependent on user-set parameters and runs faster (in total range of 30-100 times faster depending on the number of trials needed and the complexity of the data). 3D DBSCAN was also implemented by us in Matlab to test DBSCAN at different background noise densities (Figs. S5B-D); scripts are available under http://cbi-dev.igbmc.fr/cbi/DBscan3D/. The volume of a cluster here was calculated as the volume of the convex hull of the points comprising the corresponding cluster. Since the DBSCAN parameters (ε and *MinPts*) need to be optimized for every dataset, we manually chose such values that produced best results at the intermediate density (background density = 20% of the cluster density): $\varepsilon=20$, *MinPts=5*.

3 Results

A Voronoi diagram for the 3D case is a space partitioning into polyhedron regions such that every polyhedron would be the locus of points closest to the corresponding seed (Aurenhammer, 1991). Similarly to the 2D case where each SMLM event forms a seed (Fig. 2A), in 3D the seeds are defined by the single-molecule (X,Y,Z) coordinates, while the Voronoi polyhedrons can be considered as the regions of influence of the corresponding fluorophore molecules (Fig. 2B). Because the geometrical properties of Voronoi regions reflect the environment characteristics of a given molecule, we can use their short-range collective properties determined from the diagram in addition to the experimentally determined coordinates and intensities of the fluorophores. One of the most

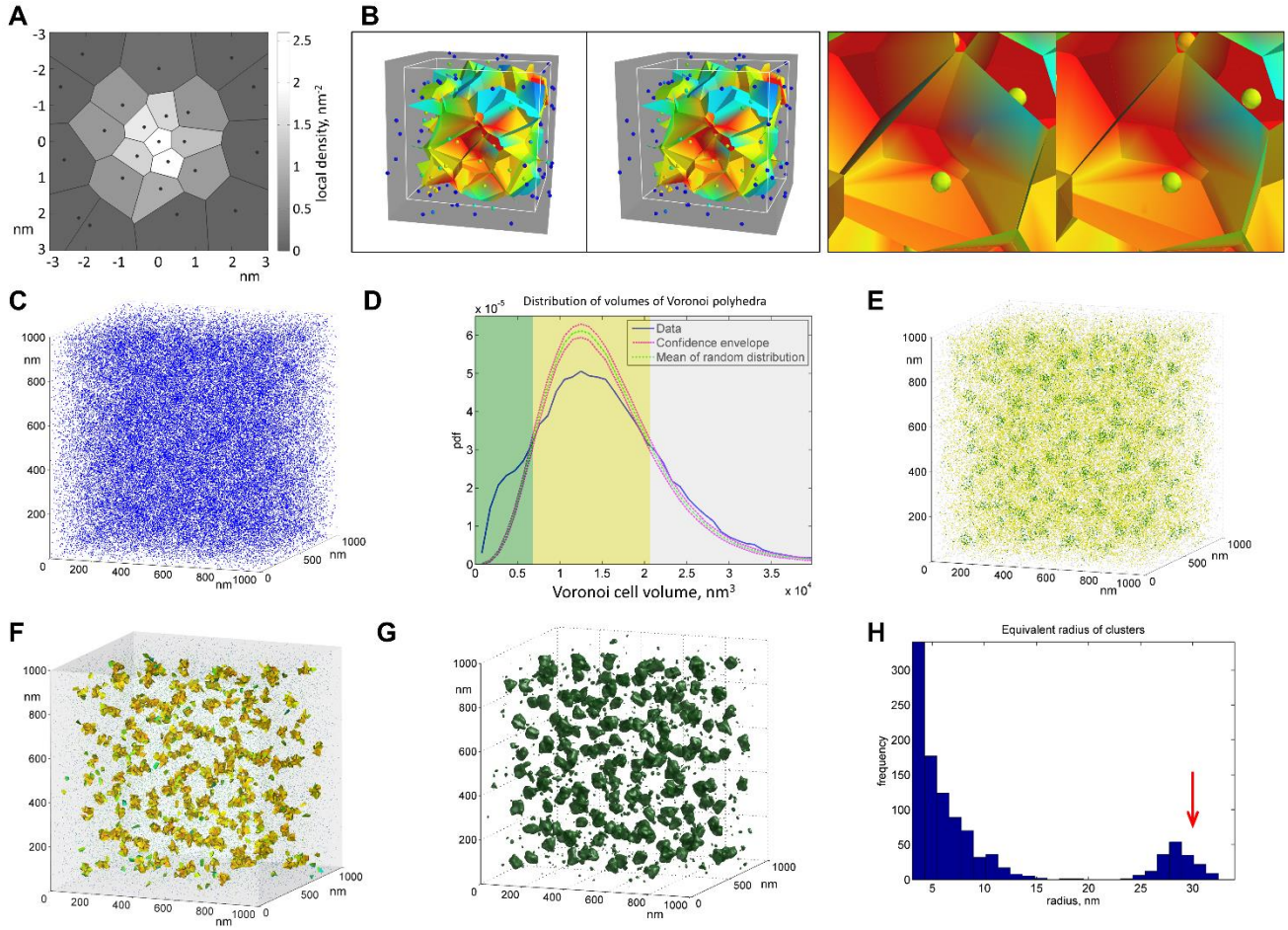


Fig. 2. Concept of 3D Voronoi-based segmentation of SMLM data. (A) Concept of two-dimensional Voronoi diagrams; the brightness of the polygons is proportional to the local density of points, determined as the inverse value of their surface area. (B) The concept of three-dimensional Voronoi tessellations highlights the strong increase in complexity upon transition from 2D to 3D data (stereo representations; zoomed region on the right; small spheres indicate the coordinates of the fluorophores; Voronoi diagram cells are colored according to their size, from blue (big cells) to red (small cells) and allow deriving the number of molecules present in a given 3D cluster with associated radius and volume values; 3D visualization tool available under http://cbi-dev.igbmc.fr/cbi/voronoi3D/tree/master/visualisation). (C) Simulated 3D volume of clustered points in a dense background. (D) Distribution of volumes of Voronoi polyhedrons of the clustered dataset (blue); mean values (green) and confidence envelope (red) of a set of distributions for datasets with randomly placed points obtained from Monte-Carlo simulations. The three characteristic regions: small clustered polyhedrons (green), intermediate and huge polygons corresponding to background (yellow and gray). (E) The original points displayed in colors accordingly to the three regions allow to visually delimit the clusters. (F) Small Voronoi polyhedrons correspond to the clusters. (G) Density map binarized at the level of the determined threshold allows cluster analysis. (H) Histogram of the equivalent radius of the clusters. The first peak corresponds to the small clusters originated from fluctuations in the dense background, the second peak with radius of (28.5±1.6) nm reveals the simulated clusters. The dataset on (C-H) consists of 175 clusters with a radius of 30 nm (red arrow). The density of molecules inside the clusters is $4{\cdot}10^{-4}$ nm^{-3}, the density of background localizations is $6{\cdot}10^{-5}$ nm^{-3}.

important properties, the density d in the neighborhood of the molecule i, can be determined simply as the inverse value of the volume V of the Voronoi cell: $d_i = 1/V_i$. Voronoi diagrams also prevent edge effects because data points close to borders have large or infinite-sized Voronoi polyhedrons, thus local densities are low and do not contribute to the clusters. Finally, Voronoi diagrams provide a direct estimate of the region of influence while Delaunay triangulations would not.

To segment a SMLM dataset, every localization should be assigned to a cluster or to the background. The minimal local density of molecules inside clusters, *i.e.* the segmentation threshold, can be determined automatically by comparing the distribution of Voronoi cell sizes of the experimental distribution with that of the background noise model, as has been recently shown by us for the 2D case (Andronov, Orlov, et al., 2016). Supposing that the background noise is distributed randomly and uniformly within the ROI, we can see how the distribution of the Voronoi cell volumes would look if all the experimental points were distributed randomly as well. For different random distributions of points the Voronoi polyhedrons would have slightly different sizes, therefore we perform many runs of randomization (Monte-Carlo simulations), in order to determine the confidence envelope, the region where 95% of all Voronoi cell sizes are situated. Then, if the experimental distribution is found outside the confidence envelope, the experimental dataset is said to be clustered. Otherwise, the experimental dataset cannot be distinguished from the random distribution and no clusters can be detected. If there is statistically significant clustering, we can classify the polyhedrons (together with the corresponding data points) into three regions with the boundaries at the intersections of the curves of the experimental and the randomized data (Fig. 2D). To keep only the clustered molecules, it is sufficient to segment the diagram at the level of the first intersection of the curves.

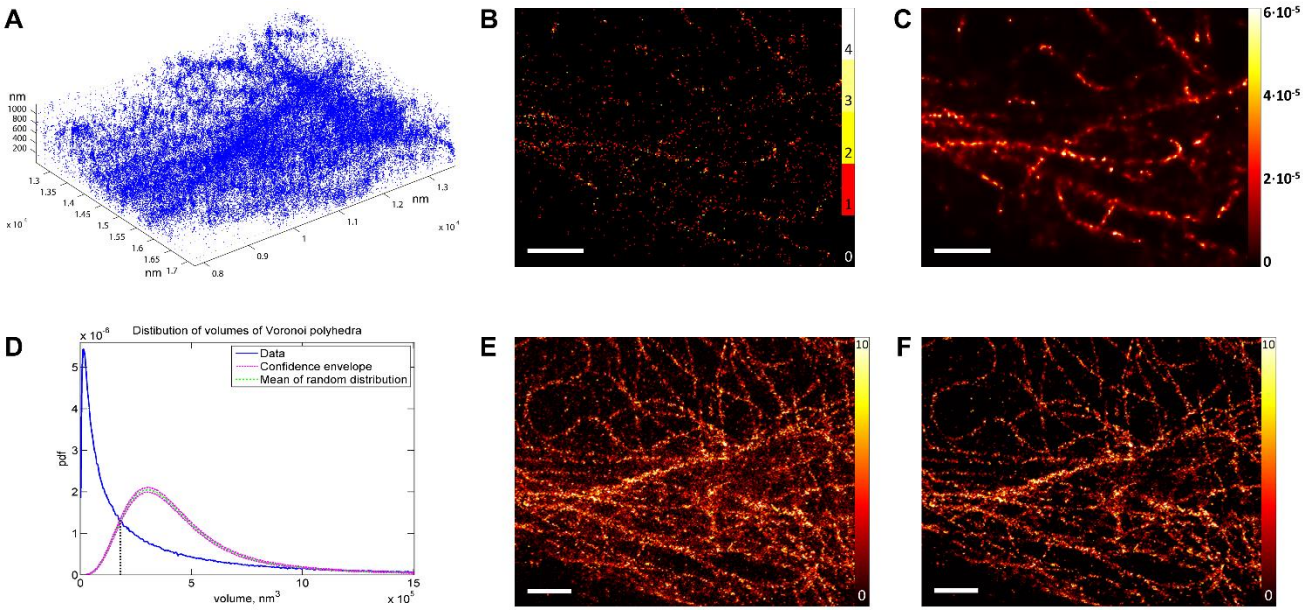


Fig. 3. Application of the 3D Voronoi diagram method for visualization and noise reduction on the example of β-tubulin. (A) Plot of fluorescent molecules in the form of points (scattered plot) does not allow for correct visualization of dense regions (β-tubulin detected with Alexa-647-labelled secondary antibodies). (B-C) A slice of the 3D volume shows more information content when calculated as a Voronoi density map (C) compared to a 3D histogram (B) (both panels show voxel-thick slices). (D-F) Distribution of volumes of Voronoi polyhedrons (D) allows noise reduction. (E-F) XY projections of the original dataset (E) and of the dataset containing only the points with the density over the threshold (F). Scale bars: 1µm. The color bars indicate number of localizations inside pixels for the histogram images (B, E-F) or the local density measured in nm^{-3} for the density map (C).

3.1 Evaluation on synthetic data

Using synthetic data to illustrate the 3D Voronoi tessellation concept (Fig. 2C; methods), we show that by randomly distributing the same number of points through the same-sized region we obtain a test dataset with Voronoi polyhedrons that should be compared with those of the clustered dataset. This allows determining the three characteristic regions which lie between the intersections of probability density functions from the distributions of Voronoi volumes of the two datasets (Fig. 2D). The first region contains small polyhedrons that are more frequent in the clustered dataset, therefore the corresponding molecules are found inside the clusters. The two other regions contain intermediate-sized polygons that are more frequent in the randomized datasets and big polygons that correspond to the low-dense background points. This can be easily seen with a color-coded display of the points (Fig. 2E) or with a Voronoi diagram (Fig. 2F). Next, interpolation of the local densities to a regularly spaced grid generates a 3D density map that can be segmented using the above determined threshold (Fig. 2G) to provide clusters with deterministic shape and properties such as their radius and number of molecules within a given cluster (Figs. 2H). The example shows many very small clusters that originate from the fluctuations of the local density in the background (Fig. 2H, first peak) and bigger clusters that correspond to the simulated clusters (Fig. 2H, second peak). The equivalent radius of these clusters (28.5±1.6 nm) matches their original radius (30 nm) and the number of events inside (36±6) compares well with the simulated number of localizations (45±3). At densities between $d_{bg} = 0.1 \cdot d_{cl}$ and $d_{bg} = 0.01 \cdot d_{cl}$ (ratio of clustered and noise events) the Voronoi volume distributions break up into two peaks which correspond to the cluster and the background regions, respectively (Fig. S3). The curve of the completely randomly distributed points crosses the other curves near the valley between the two peaks, which confirms the reliability of our automatic threshold calculation method. For a constant density of clustered molecules, the threshold decreases with increasing background density. This indicates that the cluster volume, which can be determined with confidence, decreases, while it also reduces the number of spurious clusters detected in dense background regions (Figs. S5 A-C).

3.2 Visualization of 3D SMLM data

Visualization of 3D data in the form of a 2D image can be difficult, especially in crowded environments with dense fluorophore occurrence due to the missing information in the Z-direction. The example of microtubules (Fig. 3) shows that the scatter plot is useful only for very low density data, otherwise the markers overlap and hide details (Fig. 3A). When displayed as slices of a 3D histogram, the density of molecules is usually not sufficient to ensure good continuity of the images (Fig. 3B). Moreover, when molecules are displayed as Gaussian clouds with a width dependent on the localization precision (Beheiry and Dahan, 2013) the resolution of the image can be impaired (Baddeley et al., 2010). To address these issues, it is useful to display 3D density maps (calculated from the local densities) as slices (Fig. 3C), as projections or as isosurfaces (Fig. 2G; see also later Fig. 4 panels B & E). Thanks to the properties of Voronoi polygons and after an interpolation step, the density value at any given voxel takes into account the distribution of molecules in a given 3D region around this voxel and thus ensures a good continuity of the interpolated 3D image. Additionally, the Voronoi 3D density maps can be conveniently viewed and analyzed with standard tools for processing confocal microscopy image stacks or electron microscopy maps. Another advantage of the method is noise reduction when keeping only the points with a density over the threshold determined from the Monte-Carlo simulations (Figs. 3D-F). The resulting β-tubulin dataset demonstrates a strong reduction of background noise without affecting the densely labelled microtubules, even when displayed as a 2D histogram image (Figs. 3E&F). As an example, the signal-to-noise ratio measured as $\langle I \rangle / \sigma_I$ of this image improved by 35%.

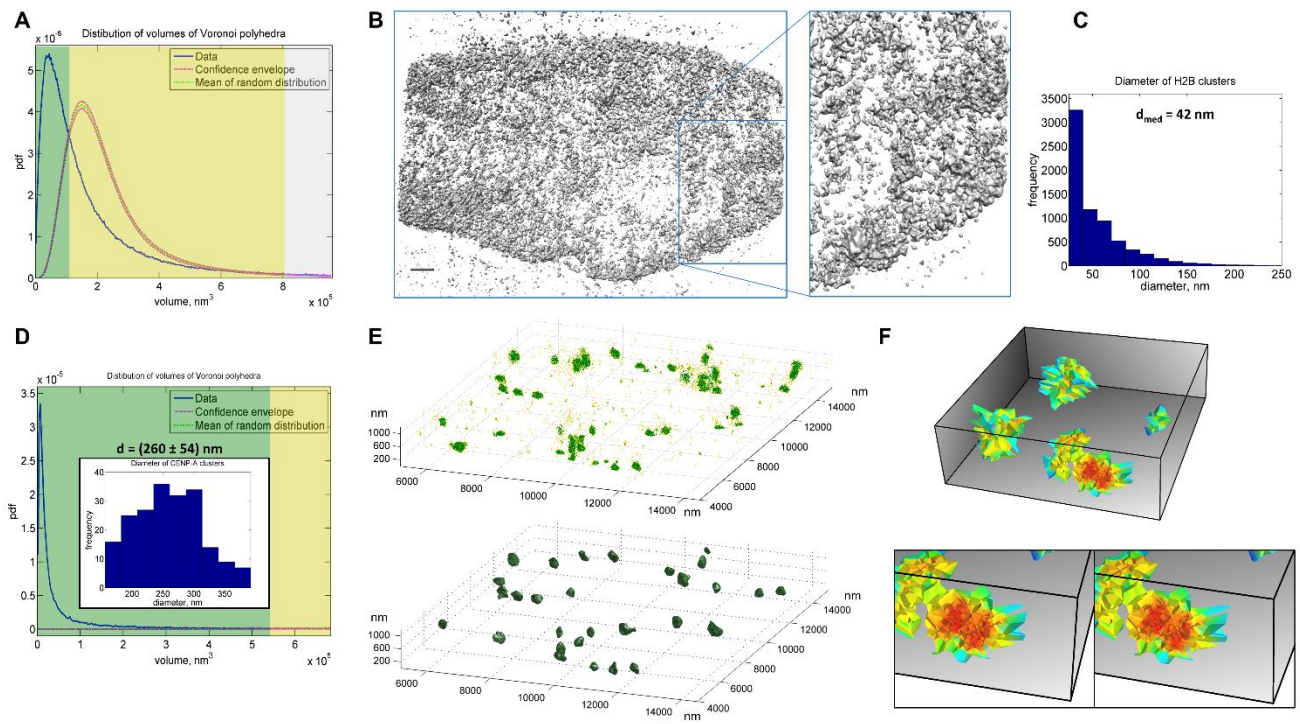


Fig. 4. Cluster analysis of histone proteins H2B and CENP-A by 3D Voronoi tesselation. (A-C) Histone H2B detected with Alexa-647-labelled secondary antibodies in a HeLa cell. (A) The distribution of volumes of Voronoi polyhedrons indicate clustering of histone H2B (color code as in Fig. 2D). (B) 3D density map displayed as an isosurface at the level of the threshold in the 3D software Chimera (Pettersen et al., 2004); scale bar, 1 µm.. (C) Histogram of the equivalent diameter of H2B clusters with the median value of 42.4 nm. (D-F) CENP-A detected with Alexa-647-labelled secondary antibodies in U2OS cells at late G1 phase. (D) The distribution of volumes of Voronoi polyhedrons indicate strong clustering of CENP-A which allows quantification of their properties. The histogram of the diameters of the centromeres is shown in the inset. (E) Scattered plot of molecules colored according to the local density (top; color code as in panels A & D); segmented density map displaying CENP-A clusters in the centromeric regions of the cell (bottom). (F) 3D representation of CENP-A clusters displayed as Voronoi tessellations in our 3D visualization tool (bottom: zoomed region in stereo representation; mean diameter of 260 nm containing an average of 418 localizations in a given 3D cluster).

3.3 Cluster analysis of histones

We applied the 3D Voronoi segmentation concept to the cluster analysis of two histone proteins, H2B and CENP-A with distinct spatial organizations (Fig. 4). H2B is known to be located inside the nucleus in the form of heterogeneous nanodomains with variable density at euchromatin and heterochromatin regions (Ricci et al., 2015). The distribution of Voronoi cells of an H2B-labelled dataset lies outside the confidence envelope for the noise model indicating statistically significant clustering (Fig. 4A). Segmented density maps of H2B show bigger and denser clusters at the periphery of the nucleus as expected for heterochromatin (Fig. 4B), the distribution of H2B cluster sizes being rather broad with a median diameter of 42.4 nm (Fig. 4C) in line with previously reported values (Ricci et al., 2015). In contrast, the histone variant CENP-A that replaces the canonical histone H3 (Palmer et al., 1987) localizes mostly at centromeric regions (Shuaib et al., 2010) where it plays a key role in determining chromosome location and segregation (Shuaib et al., 2010; Fukagawa and Earnshaw, 2014). Our analysis of SMLM CENP-A data shows that the distribution of Voronoi volumes is very different from that of the noise model and that the vast majority of the Voronoi cells is small compared to randomly distributed points indicating strong clustering (Fig. 4D). As can be seen from the comparison of the experimental distributions of the Voronoi cell volumes with those of the noise models, the CENP-A dataset has a very low background density, while the H2B dataset has high density of background (but still different from the random case) and the β-tubulin dataset has intermediate values of background density. Because of low background and low number of clusters in the CENP-A dataset, its projection in 2D still allows proper determination of the diameter of the clusters (Fig. S6), while for the H2B dataset clusters determined in 2D have substantially smaller sizes (13 nm as compared to 42 nm), which illustrates the risk of artefacts with 2D analysis as opposed to 3D analysis (Fig. S7). This is a general tendency when projecting volume data into a 2D image (see simulated data in Figs. 1, S1) because the distribution of points in 2D gets closer to random, which is seen on the graphs of the Voronoi cell sizes. Moreover, any volumetric features of clusters are lost when analyzed in 2D. Interestingly, we find that CENP-A forms clusters with relatively homogeneous sizes considering the large supra-structure of the centromeric region in the cell (diameter of 260 ± 54 nm, containing an average of 418 localizations within a cluster; Figs. 4D-F). The 3D Voronoi segmentation analysis provides for the first time evidence that histone variant CENP-A forms defined clusters in human cells and quantifies their size and occurrence.

4 Conclusions

Taken together, our novel cluster analysis method introduces a new concept for 3D (rather than 2D) segmentation of SMLM data based on the elegant mathematical properties of 3D Voronoi diagrams that allows deriving 3D cluster radius and volume values and associated molecule numbers. It provides automatic threshold determination and noise filtering. In addition, to avoid distortion effects due to non-isotropic distribution of experimental data points we introduced weighting in Z-direction (see methods; taken into account also for Monte-Carlo simulations). ViSP (Beheiry and Dahan, 2013) or a 3D implementation of DBSCAN (Barna et al., 2016) (see methods) both require multiple trials through individual runs to optimize parameters manually without additional a

priori knowledge. By contrast, our method proceeds in a single run with automatic threshold determination from Monte-Carlo simulations using a background model (into which re-localization events can be included if known) and it proves to be more robust with increasing background densities thanks to noise modelling (Figs. S5B-D; these include a statistical analysis of the results, the estimated cluster sizes can vary slightly depending on the background / cluster density). The method avoids artefactual clustering at very high densities, when the experimental distribution becomes undistinguishable from a random one. The Voronoi-based method was validated using simulated and experimental data, illustrating its applicability to different biological objects of interest including cases that are difficult to analyze such as fine chromatin structures in a dense cellular context and larger cellular structures such as tubulin. 3DClusterViSu will be of broad interest to the scientific community as it can be applied to any 3D SMLM data to analyze or re-analyze and quantify 3D information such as spatial organization, local density, volume and shape of labelled protein clusters in 3D and number of molecules within a given cluster. It allows SMLM analysis and also visualization of macromolecules directly in 3D like in the neighbor fields of cryo-EM and X-ray crystallography; therefore, it will be also a valuable tool for CLEM in integrated cellular structural biology approaches. Standalone software for 3D data processing with a GUI (3DClusterViSu) and a standalone Python-based 3D Voronoi visualization tool with a GUI are available under
http://cbi-dev.igbmc.fr/cbi/voronoi3D/repository/archive.zip

Acknowledgements

We thank Leica Microsystems for the access to the Leica SR GSD 3D system and Didier Hentsch, Pascal Kessler and Yves Lutz of the IGBMC Imaging Centre for support. This research was made possible by the adaptive optics plug-and-play accessory MicAO 3D-SR for PALM/STORM microscopes of Imagine Optic (www.imagine-optic.com).

Funding

This work was supported by CNRS, Association pour la Recherche sur le Cancer (ARC), Institut National du Cancer (INCa), Ligue nationale contre le cancer (Ligue) and Agence National pour la Recherche (ANR). The super-resolution microscope setup is supported by the Alsace Region and by the French Infrastructure for Integrated Structural Biology (FRISBI) ANR-10-INSB-05-01 and Instruct-ERIC as part of the European Strategy Forum on Research Infrastructures (ESFRI).

Conflict of Interest: none declared.

References

Andronov,L., Orlov,I., et al. (2016) ClusterViSu, a method for clustering of protein complexes by Voronoi tessellation in super-resolution microscopy. Sci. Rep., 6, 24084.

Andronov,L., Lutz,Y., et al. (2016) SharpViSu: integrated analysis and segmentation of super-resolution microscopy data. Bioinformatics, 32, 2239–2241.

Aquino,D. et al. (2011) Two-color nanoscopy of three-dimensional volumes by 4Pi detection of stochastically switched fluorophores. Nat. Methods, 8, 353–359.

Arnold,J. et al. (2016) Site-Specific Cryo-focused Ion Beam Sample Preparation Guided by 3D Correlative Microscopy. Biophys. J., 110, 860–869.

Aurenhammer,F. (1991) Voronoi Diagrams—a Survey of a Fundamental Geometric Data Structure. ACM Comput Surv, 23, 345–405.

Baddeley,D. et al. (2010) Visualization of Localization Microscopy Data. Microsc. Microanal., 16, 64–72.

Barber,C.B. et al. (1996) The Quickhull algorithm for convex hulls. Acm Trans. Math. Softw., 22, 469–483.

Barna,L. et al. (2016) Correlated confocal and super-resolution imaging by Vivid-STORM. Nat. Protoc., 11, 163–183.

Beheiry,M.E. and Dahan,M. (2013) ViSP: representing single-particle localizations in three dimensions. Nat. Methods, 10, 689–690.

Betzig,E. et al. (2006) Imaging Intracellular Fluorescent Proteins at Nanometer Resolution. Science, 313, 1642–1645.

Ester,M. et al. (1996) A density-based algorithm for discovering clusters in large spatial databases with noise. AAAI Press, pp. 226–231.

Fölling,J. et al. (2008) Fluorescence nanoscopy by ground-state depletion and single-molecule return. Nat. Methods, 5, 943–945.

Franke,C. et al. (2017) Photometry unlocks 3D information from 2D localization microscopy data. Nat. Methods, 14, 41–44.

Fukagawa,T. and Earnshaw,W.C. (2014) The Centromere: Chromatin Foundation for the Kinetochore Machinery. Dev. Cell, 30, 496–508.

Huang,B. et al. (2008) Three-Dimensional Super-Resolution Imaging by Stochastic Optical Reconstruction Microscopy. Science, 319, 810–813.

Jones,E. et al. (2001) SciPy: Open source scientific tools for Python.

Juette,M.F. et al. (2008) Three-dimensional sub–100 nm resolution fluorescence microscopy of thick samples. Nat. Methods, 5, 527–529.

Karreman,M.A. et al. (2016) Fast and precise targeting of single tumor cells in vivo by multimodal correlative microscopy. J Cell Sci, 129, 444–456.

Koning,R.I. et al. (2014) Correlative cryo-fluorescence light microscopy and cryo-electron tomography of Streptomyces. Methods Cell Biol., 124, 217–239.

Lampe,A. et al. (2012) Multi-colour direct STORM with red emitting carbocyanines. Biol. Cell, 104, 229–237.

Levet,F. et al. (2015) SR-Tesseler: a method to segment and quantify localization-based super-resolution microscopy data. Nat. Methods, 12, 1065–1071.

Li,Y. et al. (2017) Fast, robust and precise 3D localization for arbitrary point spread functions. bioRxiv, 172643.

Malkusch,S. et al. (2012) Coordinate-based colocalization analysis of single-molecule localization microscopy data. Histochem. Cell Biol., 137, 1–10.

Olivier,N. et al. (2013) Simple buffers for 3D STORM microscopy. Biomed. Opt. Express, 4, 885–899.

Orlov,I. et al. (2017) The integrative role of cryo electron microscopy in molecular and cellular structural biology. Biol. Cell, 109, 81–93.

Owen,D.M. et al. (2013) Quantitative Analysis of Three-Dimensional Fluorescence Localization Microscopy Data. Biophys. J., 105, L05–L07.

Pageon,S.V. et al. (2016) Clus-DoC: A combined cluster detection and colocalization analysis for single-molecule localization microscopy data. Mol. Biol. Cell, mbc.E16-07-0478.

Palmer,D.K. et al. (1987) A 17-kD centromere protein (CENP-A) copurifies with nucleosome core particles and with histones. J. Cell Biol., 104, 805–815.

Pavani,S.R.P. et al. (2009) Three-dimensional, single-molecule fluorescence imaging beyond the diffraction limit by using a double-helix point spread function. Proc. Natl. Acad. Sci., 106, 2995–2999.

Pettersen,E.F. et al. (2004) UCSF Chimera--a visualization system for exploratory research and analysis. J. Comput. Chem., 25, 1605–1612.

Ramachandran,P. and Varoquaux,G. (2011) Mayavi: 3D Visualization of Scientific Data. Comput. Sci. Eng., 13, 40–51.

Ricci,M.A. et al. (2015) Chromatin Fibers Are Formed by Heterogeneous Groups of Nucleosomes In Vivo. Cell, 160, 1145–1158.

Rossy,J. et al. (2014) Method for co-cluster analysis in multichannel single-molecule localisation data. Histochem. Cell Biol., 141, 605–612.

Rubin-Delanchy,P. et al. (2015) Bayesian cluster identification in single-molecule localization microscopy data. Nat. Methods, 12, 1072–1076.

Schirra,R.T. and Zhang,P. (2014) Correlative fluorescence and electron microscopy. Curr. Protoc. Cytom., 70, 12.36.1-10.

Schorb,M. et al. (2017) New hardware and workflows for semi-automated correlative cryo-fluorescence and cryo-electron microscopy/tomography. J. Struct. Biol., 197, 83–93.

Shuaib,M. et al. (2010) HJURP binds CENP-A via a highly conserved N-terminal domain and mediates its deposition at centromeres. Proc. Natl. Acad. Sci., 107, 1349–1354.

Sibson,R. (1981) A brief description of natural neighbour interpolation. In, Barnet,V. (ed), Interpreting multivariate data. New York, USA, pp. 21–35.

Thompson,R.E. et al. (2002) Precise nanometer localization analysis for individual fluorescent probes. Biophys. J., 82, 2775–2783.

Walt,S. van der et al. (2011) The NumPy Array: A Structure for Efficient Numerical Computation. Comput. Sci. Eng., 13, 22–30.

Williamson,D.J. et al. (2011) Pre-existing clusters of the adaptor Lat do not participate in early T cell signaling events. Nat. Immunol., 12, 655–662.

Winckler,P. et al. (2013) Identification and super-resolution imaging of ligand-activated receptor dimers in live cells. Sci. Rep., 3, 2387.

Wolff,G. et al. (2016) Towards correlative super-resolution fluorescence and electron cryo-microscopy. Biol. Cell, 108, 245–258.

Supplementary Figures

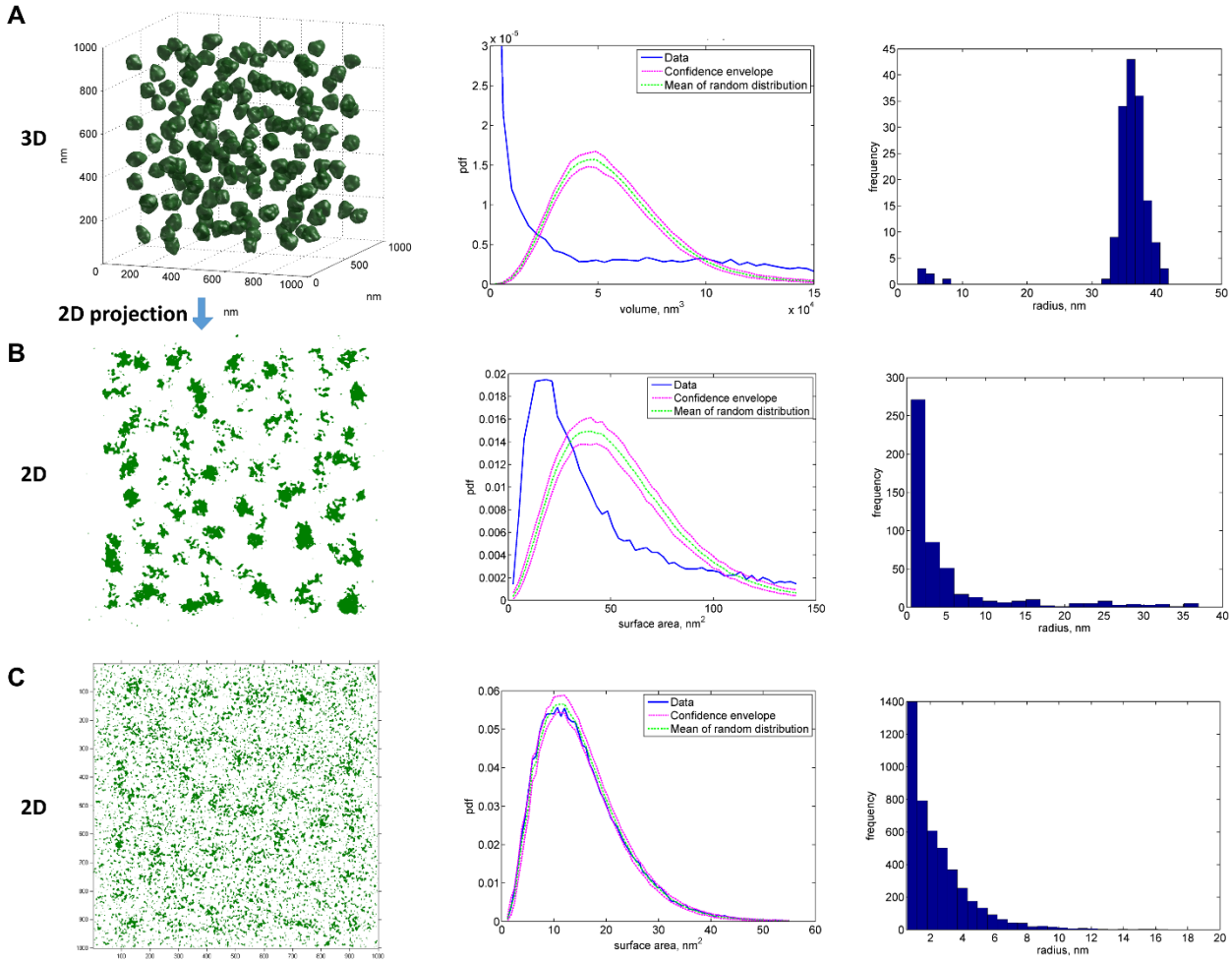


Figure S1. (A) Importance of 3D cluster analysis of localization data. Clusters in a 3D dataset (number of clusters = 150, density of localization within clusters = $6{\cdot}10^{-4}$ nm^{-3}, in the background = $8{\cdot}10^{-6}$ nm^{-3}) can be easily segmented using the 3D Voronoi diagram approach described in the current work. (B) The same dataset projected on the X-Y plane and processed with a 2D segmentation method (Andronov, Orlov, et al., 2016) does not allow to determine the correct shape, size and number of clusters. (C) Similar dataset with higher background density ($6{\cdot}10^{-5}$ nm^{-3}), lower cluster density ($4{\cdot}10^{-4}$ nm^{-3}) and higher number of clusters (175) becomes undistinguishable from a random distribution when projected in 2D, but still can be successfully processed in 3D (see later Fig. 2C-H). In the left column are the clusters determined with the 3D Voronoi approach described in the current work (A), with our previously reported 2D Voronoi approach (Andronov, Orlov, et al., 2016) (B) and using the mean local density as the threshold for binarization of 2D data (Levet et al., 2015) (C). In the middle column are probability density functions (pdf) of volumes (A) or surfaces (B, C) of Voronoi cells for the processed datasets (blue) in comparison with that of the background model, a spatially random distribution (magenta and green). The experimental curve inside the confidence envelope (C) indicates that the experimental dataset is undistinguishable from the background model and so no clusters can be detected with confidence. In the right column are histograms of the equivalent radius of the detected clusters, to be compared with the simulated cluster radius of 30 nm.

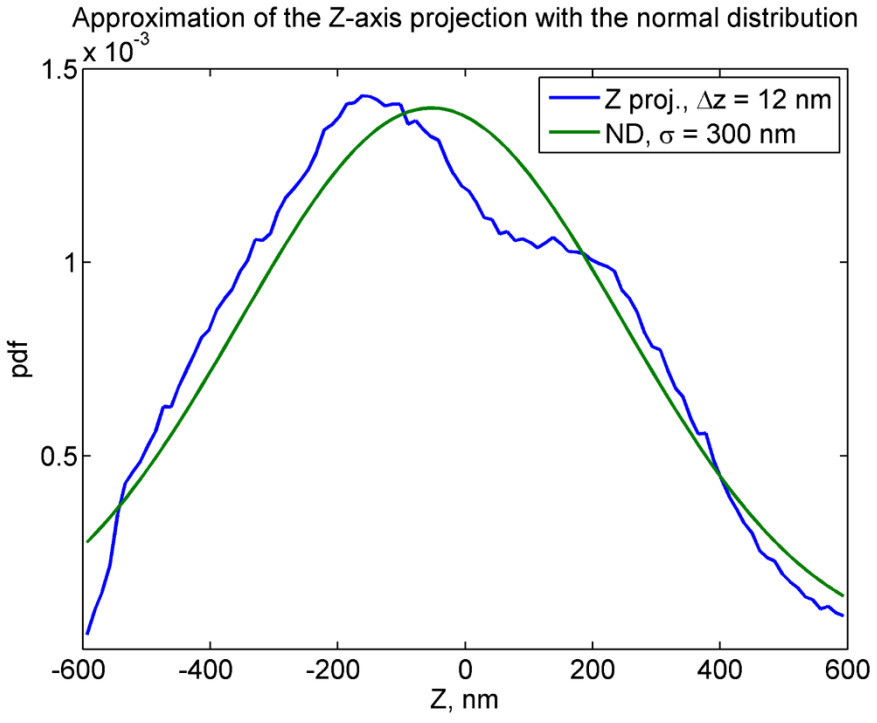


Figure S2. Approximation of the axial distribution of localizations in a 3D experiment acquired with astigmatism by a normal distribution (ND). σ, standard deviation; Δz, bin size of the histogram of the Z-positions of the dyes. Example of β-tubulin labelled with Alexa-647 conjugated antibodies (see methods).

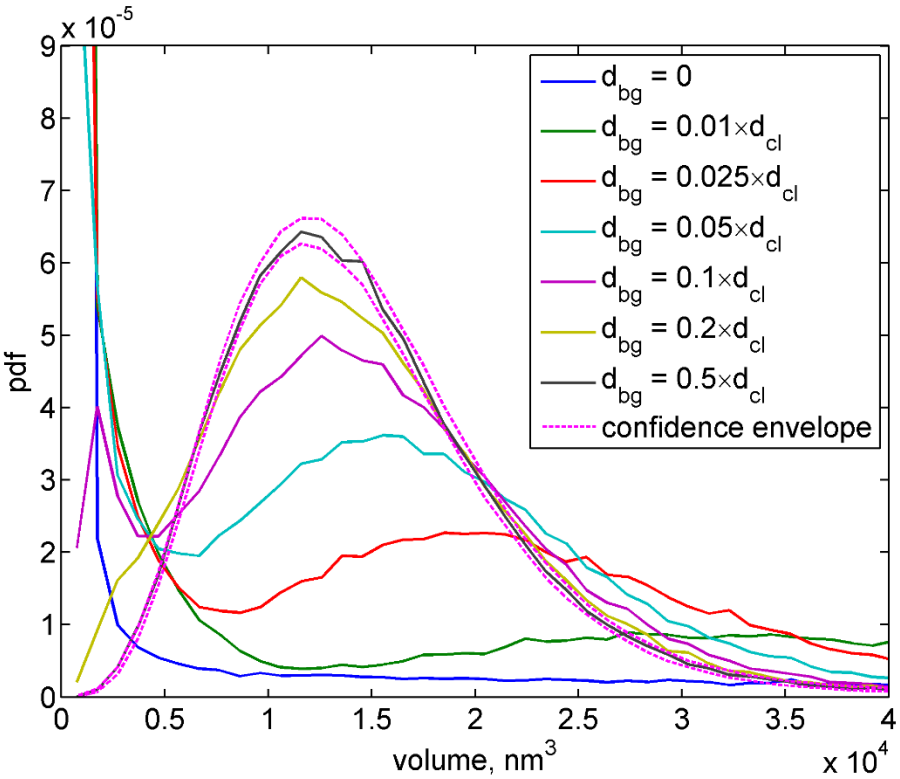


Figure S3. Behavior of volumes of Voronoi polyhedrons at different noise and cluster density ratios. The dataset contains a fixed total number of points ($7 \cdot 10^4$) and a fixed number of clusters (150) with fixed radius (30 nm).

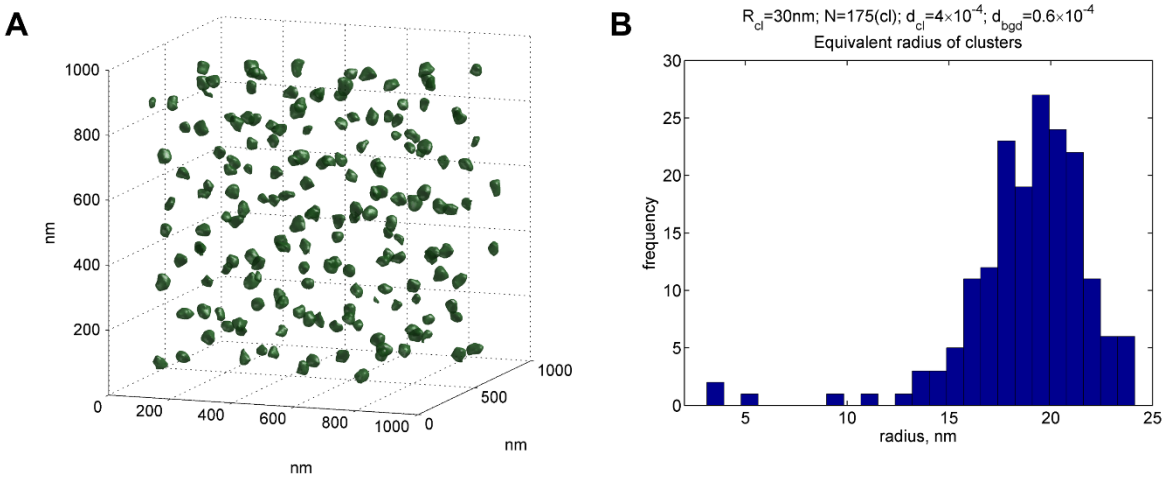


Figure S4. Segmentation of the dataset from Fig. 2C using first-rank neighbors for averaging the local densities using the 2D method described by Levet *et al.*, 2015, but implemented here in 3D for testing purposes. (A) General view of segmented clusters. (B) Histogram of the equivalent radius of the detected clusters: r=19+/-3nm, the simulated radius was 30 nm.

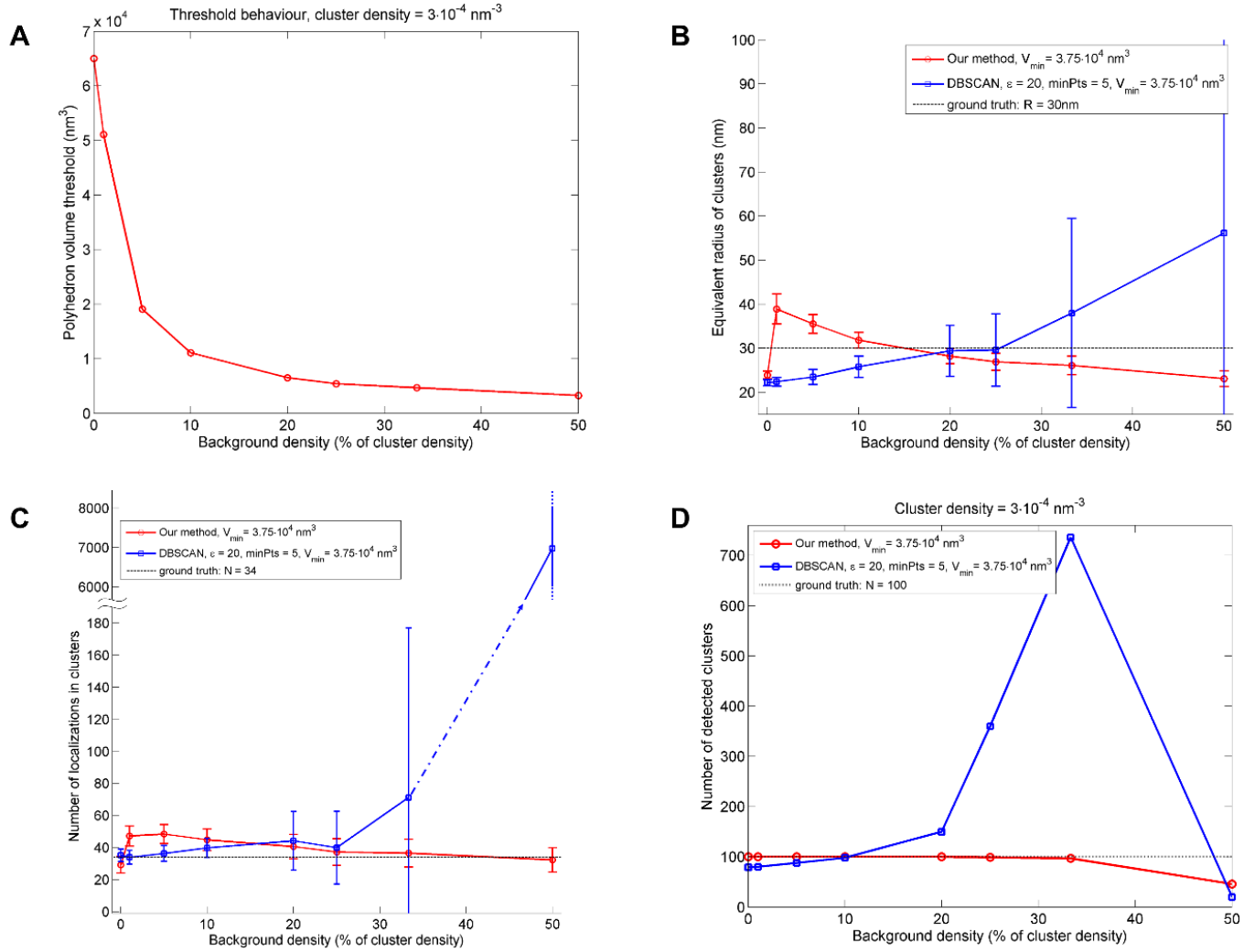


Figure S5. Evaluation of the performance of the 3D Voronoi tessellation method with synthetic data. A Behavior of the automatic segmentation threshold determination using Monte-Carlo simulations with varying background density. B-D Comparison of performance of our method with a 3D implementation of DBSCAN (with optimized parameters ε = 20, MinPts = 5) which behaves unstably with increasing background noise. Equivalent radius of clusters (B), number of localizations within clusters (C) and total number of clusters (D) detected from datasets with different amounts of noise. V_{min} is the minimal volume of an object to be considered as cluster (in order to remove spurious background clusters which are unavoidable with dense background noise). The data were generated with a constant density of clustered points ($3 \cdot 10^{-4}$ nm^{-3}) and a varying density of the background noise. Points in (B-C) indicate the mean values and the error bars indicate standard deviations. Maximum values rise up to 50% density and not beyond because a higher values the distributions become undistinguishable from random distributions (the experimental distribution of the Voronoi cell volumes is within a 95% confidence envelope for the randomly distributed points); this protects from erroneous segmentations at strong densities.

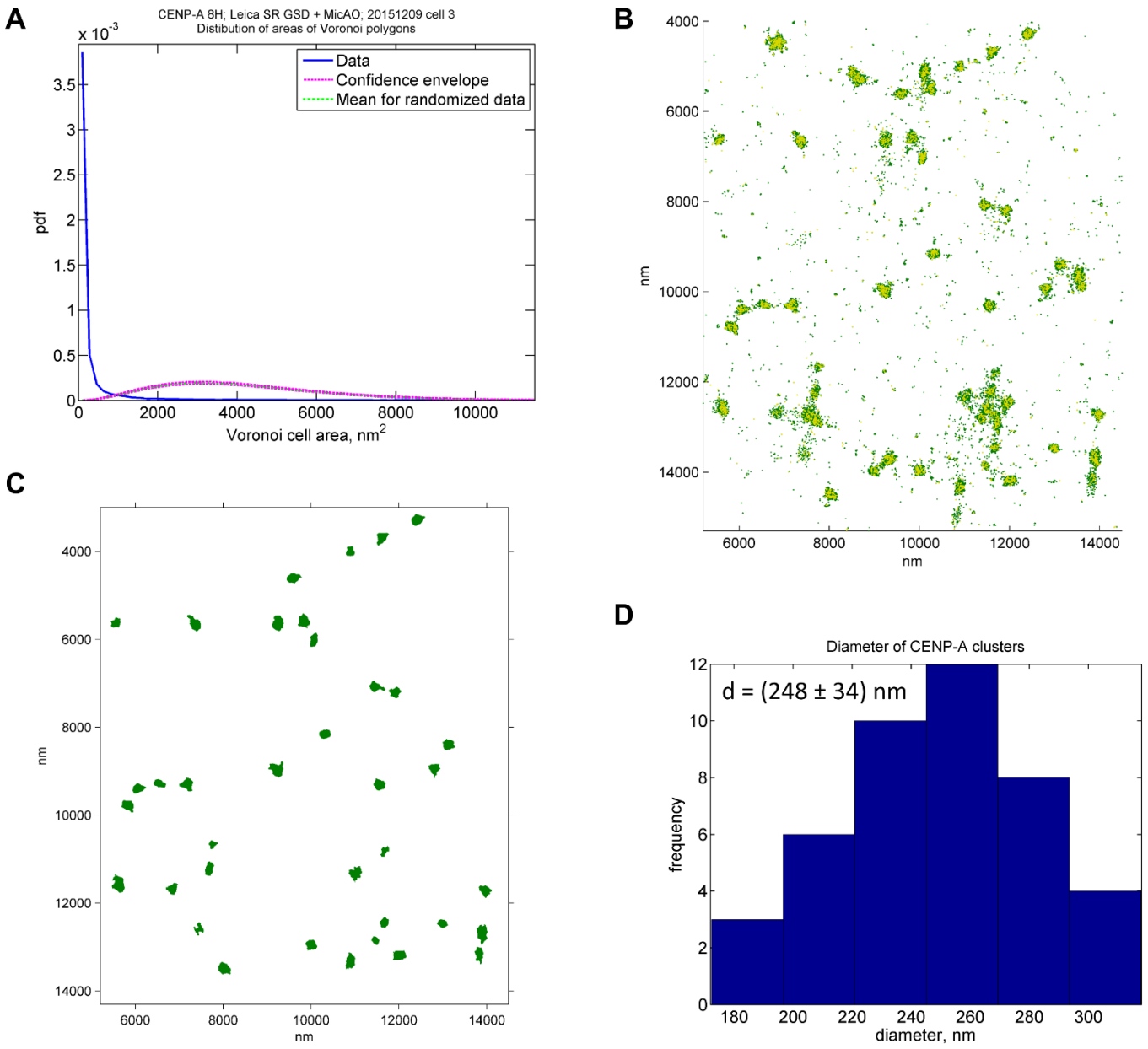


Figure S6. Segmentation of the CENP-A dataset (Figs. 4D-F) with the 2D method ClusterViSu. (A) Distribution of Voronoi cell areas indicates statistically significant clustering. (B) Scatter plot of experimental points, which correspond to clusters (green) or to background (yellow). (C) Segmented density map. (D) The distribution of the diameters of the clusters.

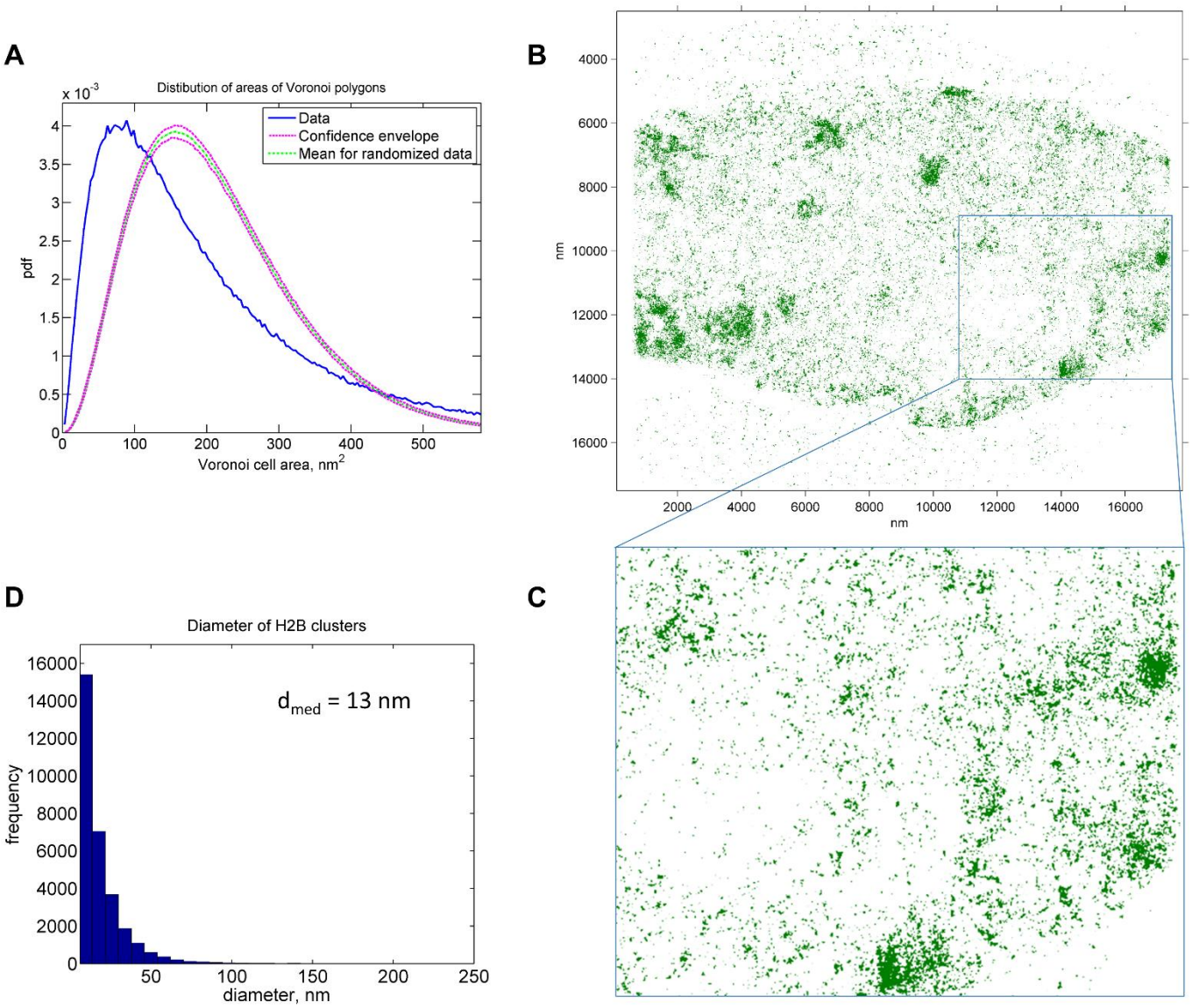


Figure S7. Segmentation of the H2B dataset (Figs. 4A-C) with the 2D method ClusterViSu. (A) Distribution of Voronoi cell areas indicates statistically significant clustering, but the distribution of molecules is closer to the random one than if analyzed in 3D (Fig. 4A). (B, C) Segmented density map (B) with a zoomed-in region (C). (D) The distribution of the diameters of the clusters is shifted toward small values as compared to the 3D analysis (Fig. 4C), in line with the general tendency seen on synthetic data (Figs. 1C, 1F, S1).

3.4. SR imaging of CENP-A throughout the cell cycle

For SR imaging of centromeric chromatin we chose to label, first of all, the epigenetic marker of centromere, the histone variant CENP-A. CENP-A is present only in some nucleosomes in the centromeric region (Blower et al., 2002). In human cells only around 400 CENP-A copies are found in each centromere, which represents ~ 4% of all centromeric nucleosomes (Bodor et al., 2014). For efficient single-molecule observation, the density of fluorophores should be under a certain limit, therefore the imaging of CENP-A instead of other histones would also facilitate SMLM because of a weaker density as compared to the histones present in every nucleosome (H2A, H2B, H4).

We initially imaged non-synchronized HeLa and U2OS cells and observed that CENP-A molecules form clusters in the centromere region of every chromosome (**Figs. 36–37**). The clusters have globally similar dimensions for most of the centromeres: globular shape with the diameter around 200–300 nm, as we reported previously (Andronov et al., 2017).

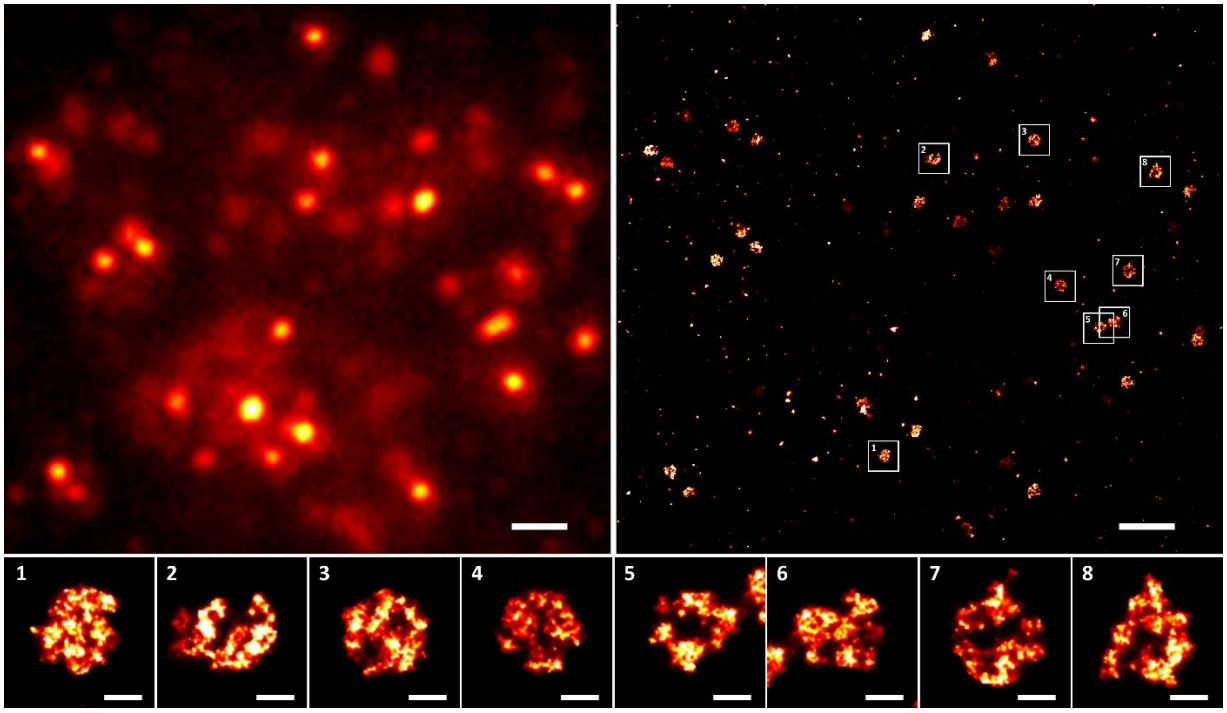


Figure 36. SMLM imaging of CENP-A in one single HeLa cell. (top left) Conventional wide-field epifluorescence image of CENP-A detected with Alexa 647 conjugated secondary antibodies. (top right) GSDIM image of the same region. (bottom) Zoomed view of several centromeres of this single HeLa cell. According to the shape of the clusters, this cell is in early G1 phase (see findings below). Scale bars, 1 µm (top), 100 nm (bottom).

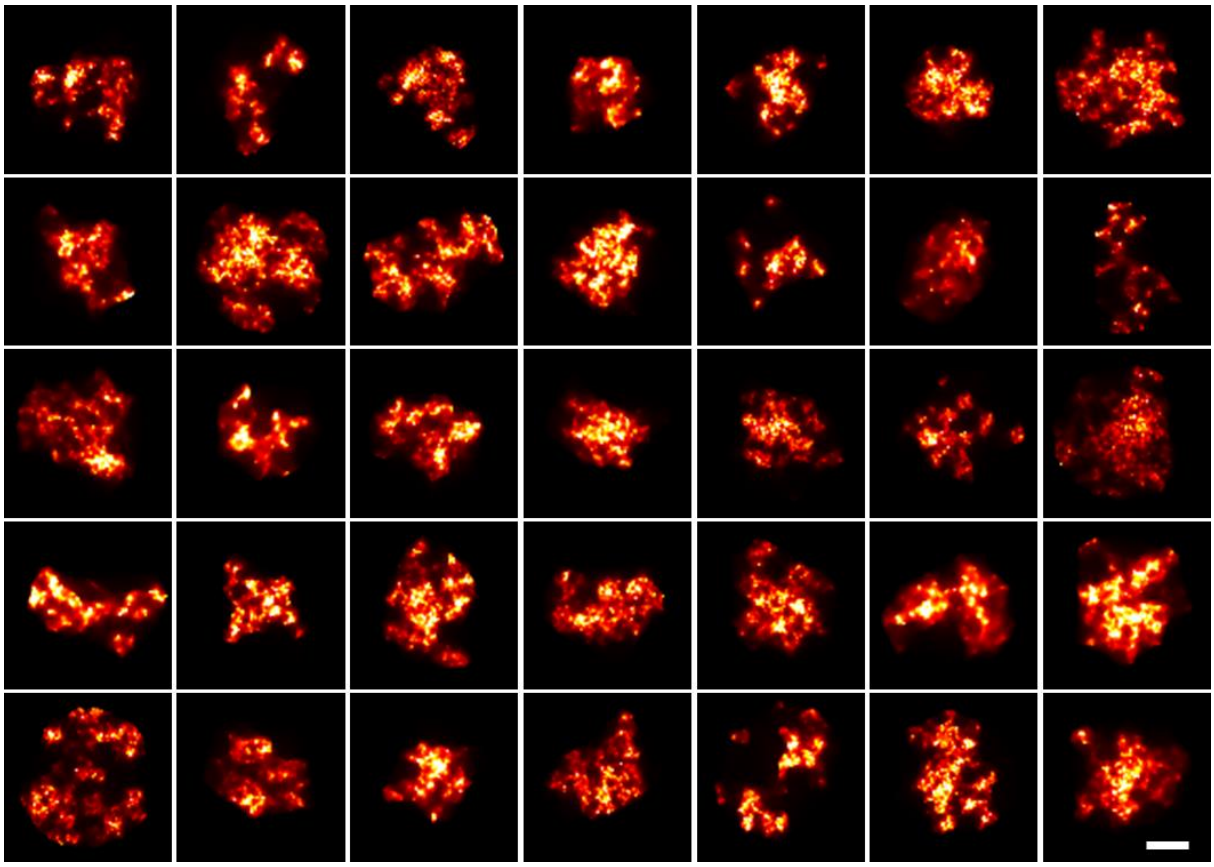


Figure 37. SMLM images of centromeres in different non-synchronized HeLa cells. The centromeres have been randomly selected from 218 imaged cells. Box size, 400x400 nm; scale bar, 100 nm.

Since centromeric chromatin experiences transformations in early G1 phase due to the deposition of new CENP-A in the nucleosomes, we decided to focus on this timepoint with super-resolution microscopy. We therefore synchronized U2OS cells, arrested and fixed them at different time points from 1.5 to 8 hours after mitosis. Already from reconstructed SR images we could notice a change in the shape of CENP-A clusters upon the progress of the cells in G1 phase. At the beginning of the cell cycle, until ~3 hours after mitosis, the centromeres are organized in more structured clusters (contrasted clusters with sharper edges, often composed of several small sub-clusters) with a hole in the middle. From ~5 hours after mitosis, the clusters adopt less structured cloud-like shape with more diffuse edges, and the central hole disappears (**Fig. 38**).

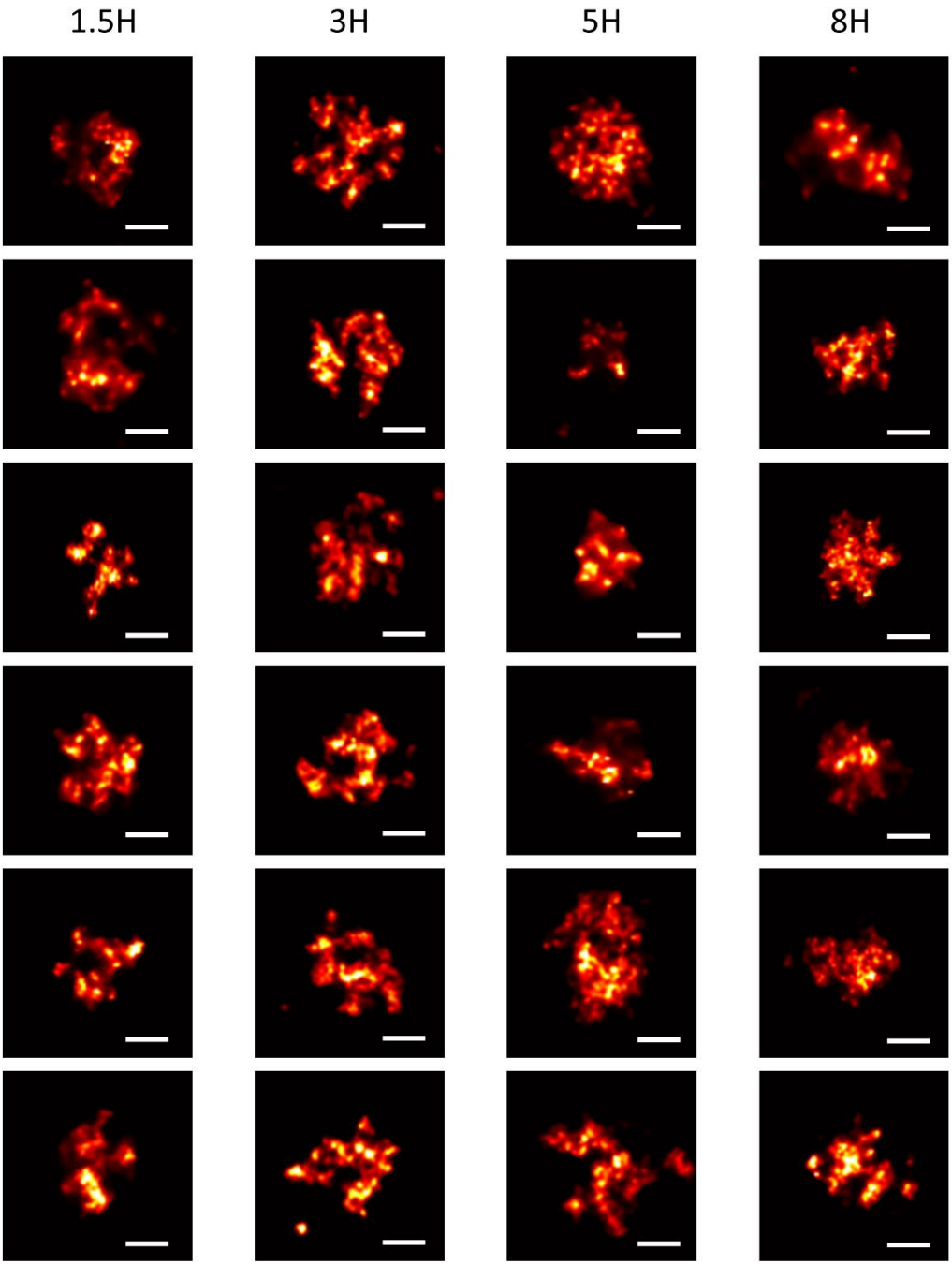


Figure 38. SMLM imaging of centromeres in synchronized U2OS cells. Cells were fixed at 1.5, 3, 5 and 8 hours after mitosis (the time point is shown at the top). For each time point, the six displayed centromeres were picked randomly. A transformation in the shape of the CENP-A clusters can be noticed when a cell progresses in the G1 phase. Scale bars, 100 nm.

To evaluate the clusters of CENP-A in 3D, we first tried to reconstruct 3D volumes from the 2D images supposing they are projections of the same object. I used a procedure from the cryo-EM field, based on the determination of the angles between projections with the common line

approach (Van Heel, 1987). Using two different software, Imagic (van Heel et al., 1996) and Simple (Elmlund and Elmlund, 2012), we could confirm that the shape of the clusters is hollow in 3D (**Fig. 39**). However, the reconstructions do not refine to high resolution, and each run of reconstruction produces a cluster with different higher-resolution details. This can be explained by the fact that centromeres in a cell probably do not have exactly the same structure, and, moreover, their images become more different when detected with indirect immunolabelling.

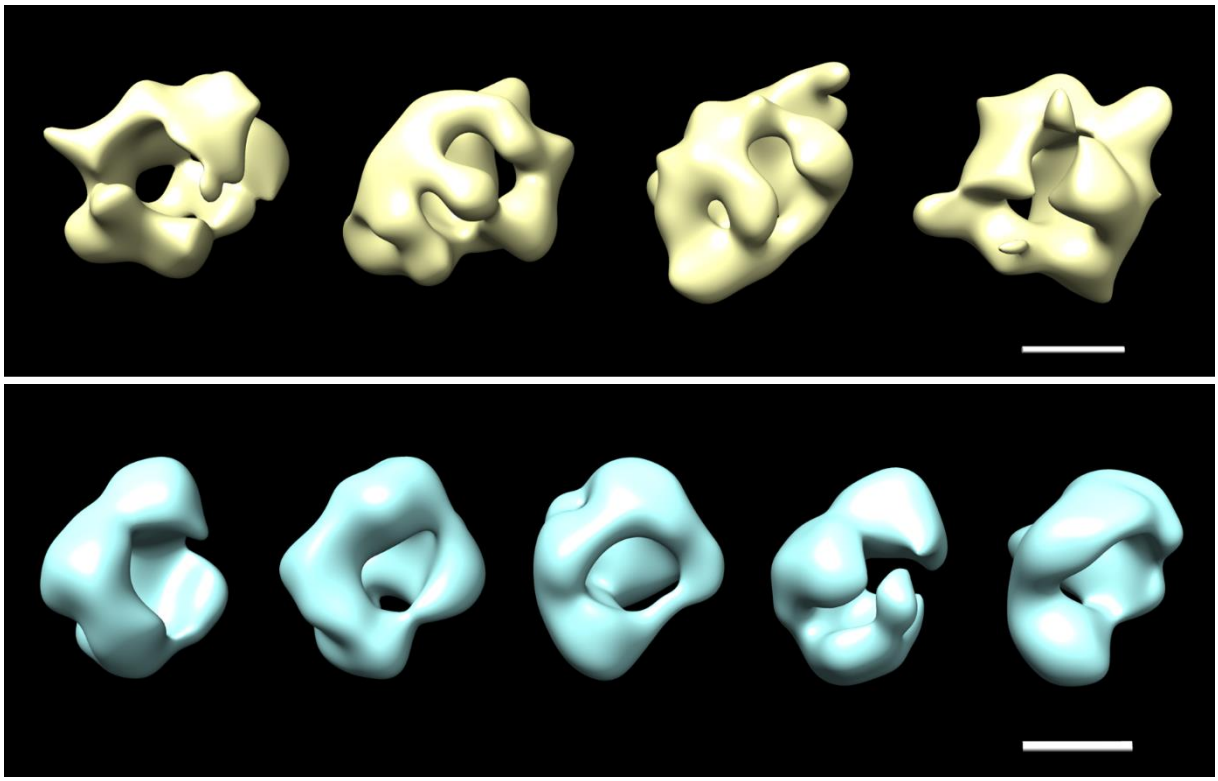


Figure 39. 3D reconstructions of the centromere at 1.5 hours after mitosis based on the common line approach. (top) 3D reconstructions in Imagic using different sets of particles for the C1-startup procedure. (bottom) 3D reconstructions from different runs in Simple. 2D images of centromeres from the same cell were used for both A and B. Scale bars, 300 nm.

To further investigate the structure of centromeric chromatin, we performed 3D SMLM with astigmatism (Huang et al., 2008a) on these samples. Using adaptive optics (Izeddin et al., 2012), which we implemented on the GSDIM device as part of my work, we could adjust the strength of astigmatism in order to obtain a more isotropic resolution. 3D SR imaging confirmed the findings about the transformation of the CENP-A clusters made with our previous 2D experiments (**Fig. 40**).

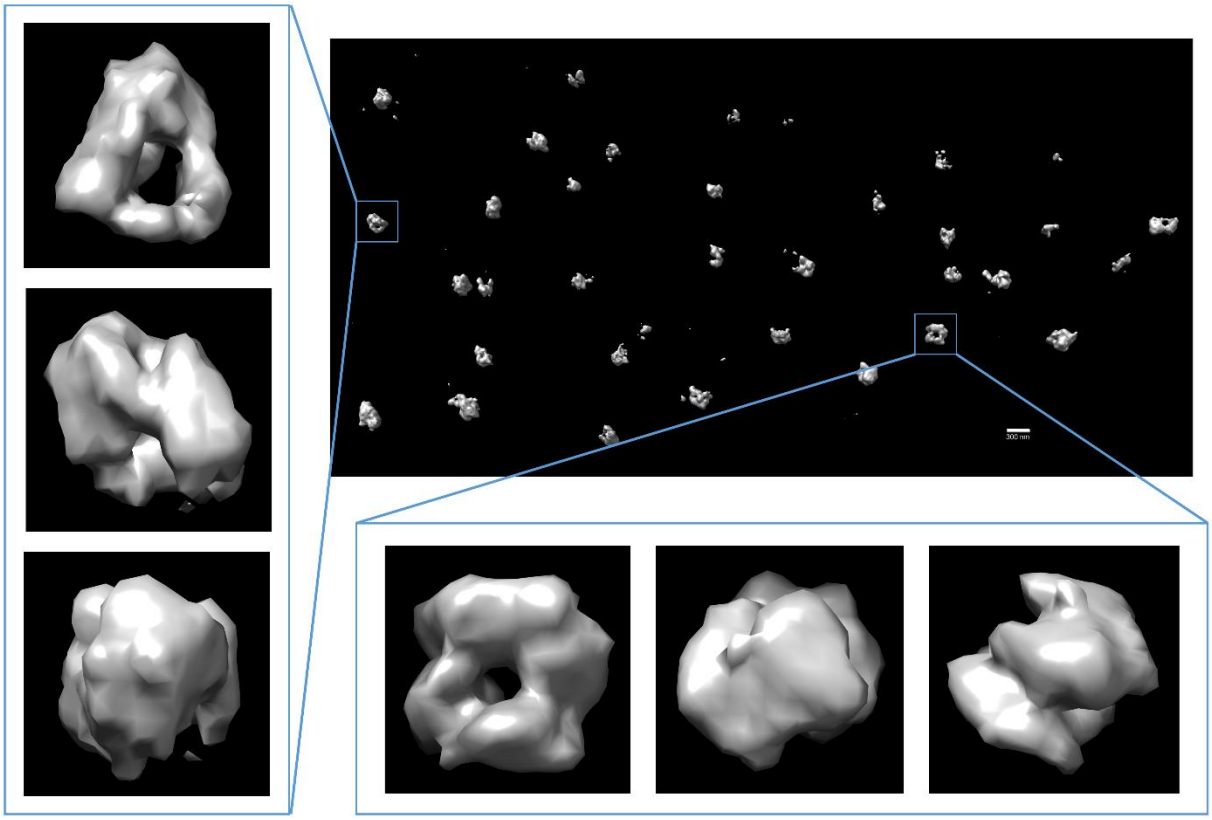


Figure 40. 3D SMLM imaging of CENP-A in a U2OS cell fixed at 1.5 hours after mitosis. Hollow clusters can be seen. Image is created as a 3D Voronoi-based density map (Andronov et al., 2017) and displayed with the Chimera software (Pettersen et al., 2004). Zoomed-in images display rotated views of a particle. Scale bar, 300 nm; box size, 600 nm.

To confirm these findings statistically, we picked the CENP-A clusters from 2D SMLM images of cells fixed at 1.5, 3, 5 and 8 hours after mitosis, centered and iteratively aligned them to their rotationally averaged sum in the Imagic software (van Heel et al., 1996), a procedure similar to that applied previously for structural studies of the nuclear pore complex (Szymborska et al., 2013). The rotationally averaged images of the sum of the aligned particles confirm that the structural transformation of the centromeric chromatin from the hollow to the filled shape at the beginning of G1 phase is statistically significant (**Fig. 41**).

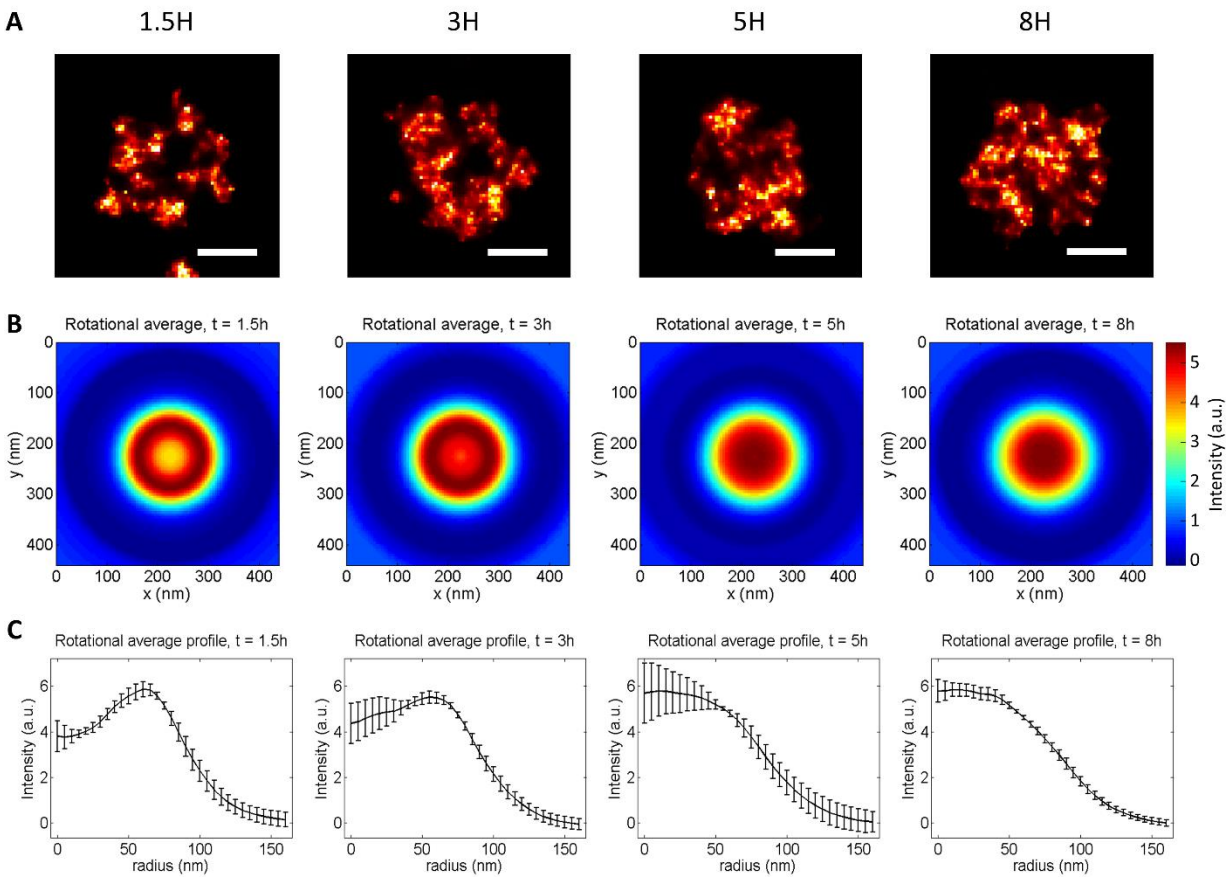


Figure 41. Statistical analysis of CENP-A clusters in fixed and synchronized U2OS cells. (**A**) Manually picked representative clusters for each of the time points (1.5, 3, 5 and 8 hours, shown on top). (**B**) Rotationally averaged images of the sum of many aligned clusters for every time point. (**C**) Profiles of the rotationally averaged images shown in B with the standard deviation displayed with the error bars. For this analysis, images of centromeres from 5 (1.5h – 5h) or 6 (8h) different cells were used: in total 211 clusters for the 1.5h time point, 196 clusters for 3h, 243 clusters for 5h and 214 clusters for 8h. The error bars indicate the SD between the rotationally averaged images of the sum of the CENP-A clusters of each analyzed cell for the given time point. Scale bars, 100 nm.

Since HJURP is a histone chaperone which allows for the deposition of CENP-A in chromatin and it was shown to co-localize with CENP-A in early G1 phase when imaged with the resolution of the confocal microscope ((Dunleavy et al., 2009), **Fig. 42A**), we studied its distribution with SMLM. Surprisingly, while at low resolution HJURP appears to co-localize with CENP-A, in reality it does not co-localize with CENP-A, which is evident when the two proteins are imaged with high resolution. Instead, HJURP is often situated in the empty center of the CENP-A clusters until 3 hours after mitosis (**Fig. 42B**).

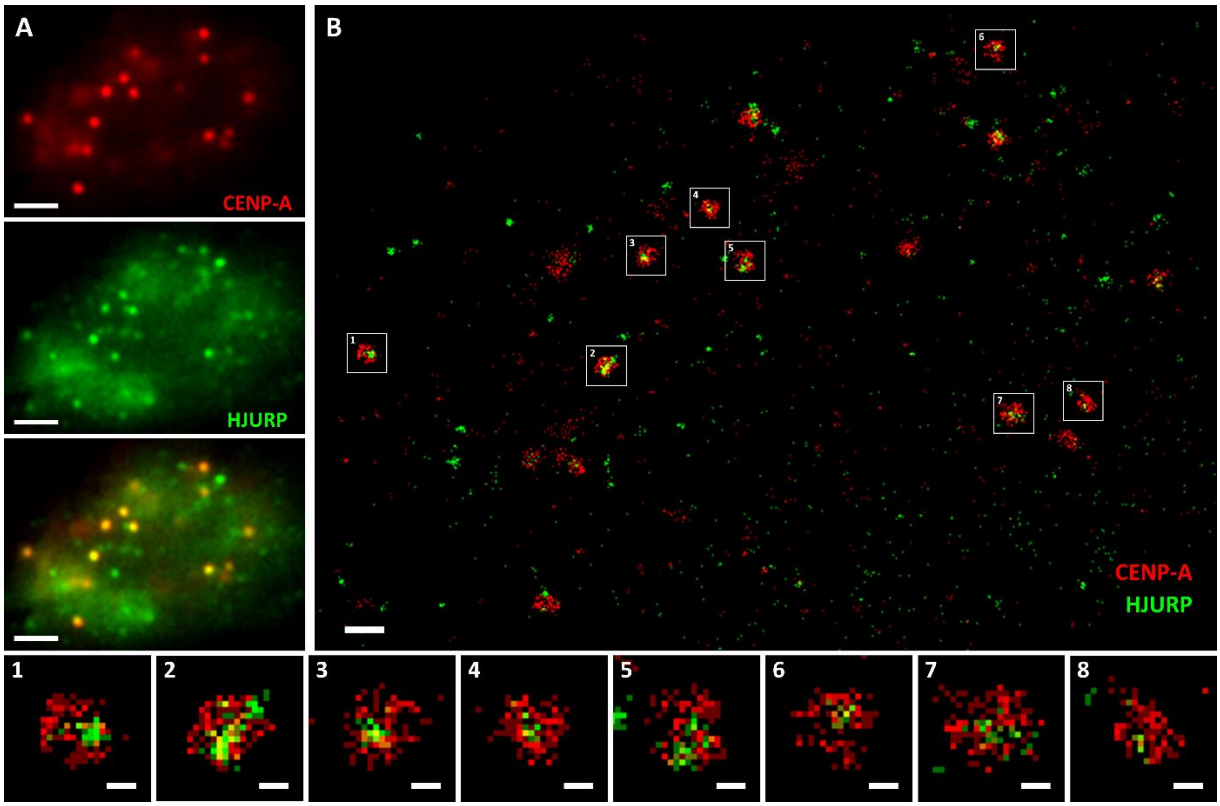


Figure 42. SMLM imaging of CENP-A with its chaperone HJURP. (**A**) Wide-field images of CENP-A and HJURP demonstrate their colocalization at centromeres. (**B**) SMLM demonstrates that CENP-A (red) and HJURP (green) form clusters of different shapes which do not completely colocalize. (bottom panels **1**–**8**) Zoomed in centromeres from image B demonstrate that an HJURP cluster is often situated in the empty center of a CENP-A cluster. CENP-A was labeled with Alexa 647- and HJURP with Alexa 555- conjugated secondary antibodies. SR images (B, 1–8) are reconstructed in the histogram mode with the bin size of 20 nm. Scale bars, 2 µm (A), 500 nm (B) and 100 nm (1–8).

Finally, we performed STED microscopy on U2OS cells synchronized in the same way. Despite lower resolution, STED imaging fully confirmed our previous findings about the centromeric chromatin, both about the shape of the clusters and the localization of HJURP in the middle of the CENP-A clusters at the beginning of G1 (**Fig. 43**).

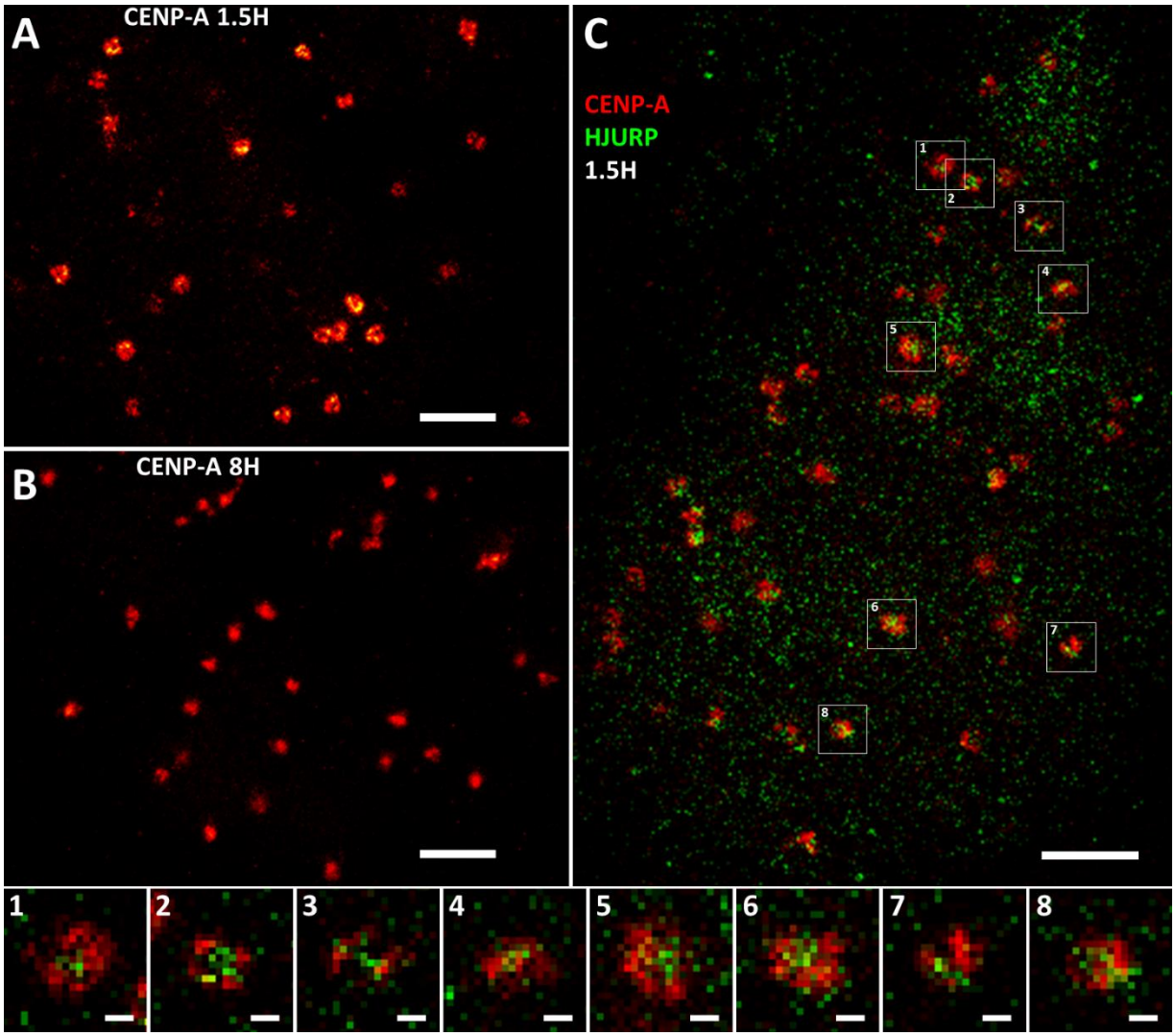


Figure 43. STED imaging of CENP-A and HJURP proteins. (**A**) CENP-A forms hollow clusters in a cell fixed at 1.5 hours after mitosis. (**B**) At 8 hours after mitosis clusters become less sharp without the central hole. (**C**) HJURP (green) forms smaller clusters that are often situated in the middle of the CENP-A clusters (red). (bottom panels **1**–**8**) Zoomed-in centromeres from image C. CENP-A was detected with Alexa 488- (A, B) or Alexa 555- (C) conjugated secondary antibodies. HJURP was labeled with Alexa 488-conjugated secondary antibodies. Scale bars, 1 µm (A-C) and 100 nm (1–8).

Taken together, using super-resolution fluorescence microscopy we show for the first time that CENP-A forms well-defined clusters in the centromeric regions of chromatin in interphase of human cells. Based on time points, we also describe structural changes of these clusters upon the deposition of newly synthetized CENP-A into the centromeric nucleosomes. We find that HJURP is situated in the center of these clusters where no CENP-A is located. Our results suggest that there is clearly a structural change in centromeric chromatin upon the CENP-A

deposition, not only at the level of individual nucleosomes, but also at the level of the high-order chromatin structure, at the range of ~50 nm.

The exact role of these structural changes in chromatin still remains to be explored. To further investigate this problem from the microscopic point of view it is necessary to image chromatin within the cell nucleus at a single nucleosome resolution, at least at ~10 nm in all three dimensions. Also, such imaging should be highly specific in order to distinguish *e.g.* different kinds of histones and other proteins within the overcrowded environment of chromatin. Despite huge progress of recent years, SR fluorescence microscopy is only slowly approaching this level. In this context of linking the high-resolution protein structures with their functions within a complex organelle, correlative SR fluorescence-electron microscopy may be useful; however it is an imaging technique, which is still on an early stage of its development and many technical issues are to be solved before it can be used in practice for such a complex task.

3.4.1. Methods

Cell culture and immunolabelling was performed as described previously (section 3.3), with the difference that the cells synchronized with the Thymidine-Nocodazole protocol were collected and fixed at 1.5, 3, 5 and 8 hours after mitosis. CENP-A was detected with a rabbit polyclonal antibody 07–574 (EMD Millipore). HJURP was detected with a mouse monoclonal antibody (produced in-house IGBMC). Unless otherwise specified, CENP-A was labeled with Alexa 647-conjugated secondary antibodies. SMLM imaging was performed in a buffer with the refractive index of 1.49 composed of Vectashield (Vector Laboratories), (20%), 2,2'-thiodiethanol (70%) and PBS 10x (10%) (Olivier et al., 2013). The samples for STED microscopy were mounted in ProLong Diamond Antifade (Thermo Fisher Scientific).

SMLM microscopy was performed as described previously (section 3.3). STED microscopy was done on a Leica TCS SP8 STED 3X microscope with the HC PL APO CS2 100x/1.40 OIL objective. The fluorophores were excited with a supercontinuum laser at 15% of its maximal power. Alexa 488 was depleted with a 592 nm continuous-wave laser set at 70%-90% of its maximal power. Alexa 555 was depleted with a 660 nm continuous-wave laser set at 75% of its maximal power. The fluorescence was detected with the HyD detectors working in the photon counting mode with time gating, detecting light from 0.5 to 6.0 ns after the excitation laser pulse. The detection of the fluorescence in a time window ("gating") is useful for STED microscopy with continuous-wave depletion lasers. Because in continuous STED the laser light

depletes fluorophores continuously, the periphery of the PSF is being suppressed continuously after the excitation pulse. That is, in the first nanoseconds after the excitation the PSF has still a size, close to the diffraction limit. Not detection of this first fraction of photons increases overall resolution of the image. This technique got the name “gated STED” (Auksorius et al., 2008; Vicidomini et al., 2011, 2013).

SM localization and drift correction were done as described in section 3.3. Unless otherwise stated, SMLM images were built as 2D or 3D density maps based on the Voronoi diagrams method (Andronov et al., 2016a, 2017; Levet et al., 2015). For the analysis of the CENP-A localization, the CENP-A clusters were picked from the images using the Boxer tool from the EMAN2 package (Tang et al., 2007). These images of the clusters were then normalized and iteratively aligned to their rotationally averaged sum in Imagic (van Heel et al., 1996). The resulting aligned particles were again rotationally averaged and their radial profile was calculated in Matlab (**Fig. 41**). For statistical evaluation of these data, all aligned particles originating from every individually analyzed cell were first rotationally averaged and summed separately for each cell. Radial profiles were calculated for each of these average images (one per cell). The standard deviation shown in **Fig. 41C** in error bars was calculated for every point of these radial profiles.

4. Conclusions and perspectives

4.1. General discussion

At the beginning of my thesis in 2013 methods for SM localization, *i.e.* for getting the coordinates of fluorophores from SMLM frames, were already relatively well established, but techniques for processing of these point data in the context of super-resolution microscopy were only on their early stage of development. Technical imperfections of SMLM were mostly understood and it was clear that some of these aspects, such as sample drift and chromatic aberrations can be improved not only on the hardware level but also by processing of localization data. These post-processing methods were often developed separately for every biological project, which obviously complicated research and more routine work.

I therefore tried to integrate important data processing steps in a single user-friendly tool, SharpViSu (Andronov et al., 2016a). It is conceived as a software that would conveniently lead a user through the corrections and processing of data, from the raw single-molecule coordinates up to the reconstructed SR image. These corrections include sample drift correction based on cross-correlation between images reconstructed from consecutive sublists of localizations. While most of drift correction tools available at that time needed fiducial markers in the sample for tracing their motion during acquisition, the cross-correlation method can also work without these markers using only the time-related information intrinsically present in any SMLM data. We also noted that this method performs best when applied iteratively 2–4 times and thus we implemented the iterative cross-correlation based drift correction. In SharpViSu user has the complete control over the drift estimation because it is possible to modify parameters, such as the number of consecutive sublists, the pixel size of images, ROI for cross-correlation calculation, number of iterations, etc. The estimated drift trace is readily shown on screen allowing to change parameters straight away if the estimation is not satisfactory.

Another problem for super-resolution microscopy is chromatic aberrations of the objective. While optics allow to correct them until a level unnoticeable in classical confocal microscopy, in SR microscopy the shift between the images taken with different wavelengths is usually way bigger than the localization precision. This, of course, compromises precise colocalization studies. To correct this aberration to a level acceptable for SMLM, we implemented in SharpViSu a fitting of the aberrations using positions of multicolor fluorescent beads, a fit than can be then applied for correction of experimental multicolor data. The corrected images acquired on our in-house Leica SR GSD microscope have a shift under 20 nm between the

images of different color channel, a value close to the nominal resolution of the system, while the uncorrected images had a lateral aberration of up to 70 nm.

Also, in SharpViSu the user can *e.g.* filter out imprecise localizations based on their brightness and merge localizations in consecutive frames originating from presumably same molecules that can be useful for following quantification of data. The resolution of data can be evaluated at any point with FRC that allows for tracing the resolution improvement with data corrections. After all the corrections and verifications are done, the data can be finally rendered as an image in one of numerous available modes, including histogram, Gaussian, quad-tree or Voronoi density map with flexible pixel size and brightness levels. Naturally, a color map for each of two supported channels can be flexibly chosen. The image can be rendered from the non-corrected and the corrected data at once allowing for visual evaluation of improvements. SharpViSu also supports 3D data as it provides means for calibration of experiments with astigmatism and allows viewing 3D data as slices in the histogram mode.

At the beginning of my PhD there was no ready-to-use tool with such reach, integrated, flexible and ready to use possibilities for basic analysis and corrections of SMLM data. Only some time later, in 2014–2015, software with similar capabilities, such as ThunderSTORM (Ovesný et al., 2014) or PALMsiever (Pengo et al., 2015), appeared. SharpViSu was therefore a timely contribution for the super-resolution community.

While tools for correction and visualization of data were being released as separate software and also some of these methods were integrated into commercial microscopy systems, the situation with advanced data processing was more difficult. Often, a separate method was developed for every new biological experiment that needed an advanced data processing. For cluster analysis Ripley's analysis was almost the only technique used in practice for SMLM data. This method has disadvantages in that it is not automated and includes several parameters that have to be set or adjusted manually by the user; furthermore, this method was not implemented in a ready-to-use form. I focused on developments of segmentation techniques for SMLM and first I implemented the Ripley's K and L functions method with Monte-Carlo simulations (Owen et al., 2010) in ClusterViSu, a tool that can be used as a separate software or as a plugin for SharpViSu.

Next I developed a novel method for cluster analysis based on Voronoi tessellations. This method has advantages over all previously reported methods in that it is automated for a given ROI thanks to noise modeling with Monte-Carlo simulations. While the previously used

modifications of the Ripley's function methods also include noise modeling with Monte-Carlo simulations, our Voronoi diagram-based method has advantages in that it allows for an unambiguous local density estimation as the inverse value of the Voronoi polygons. This method allows to determine such parameters as size and shape of clusters, density of molecules inside clusters, density of clusters in the dataset, etc. Also, we have shown an application of this method for the estimation of colocalization using the overlap between clusters of different species.

We have also calculated the FRC resolution of Voronoi density maps and shown that it can be better than that of the histogram or the Gaussian visualization mode. Therefore, the rendering of localization data in the form of Voronoi density maps has to become another useful way to build an image for SMLM experiments. This visualization mode emphasizes data features and allows their continuity in the image.

Another segmentation method based on Voronoi tessellation, SR-Tesseler (Levet et al., 2015), appeared when the development of our method was finished. In SR-Tesseler the authors defined the local density in the same way as we did but also they proposed an averaging of the local density using neighbors from different shells of a given Voronoi cells (so-called multi-rank neighbors). For the segmentation threshold, however, Levet et al. proposed to use a rather simple criterion of the average density. We showed on simulated data that our criterion based on noise modeling gives better results than that of SR-Tesseler, especially at conditions of weak density of clustered points or weak background. We have also shown that the segmentation with our method can be more precise if the noise model is refined with a known number of re-localizations of fluorescent molecules (Andronov et al., 2016b).

After validation of the Voronoi-based clustering concept on synthetic data, we applied this method for segmentation of biological data, namely for the nuclear pore complex, RNA polymerase II and histone H2B. The geometrical properties of the nuclear pore complexes determined with our method (their size and density on the nuclear membrane) correspond to the values known from the literature. Concerning RNA Pol II and H2B we have shown that both proteins form clusters in the cell nuclei; however the amount of the overlap between the clusters of RNA Pol II and H2B corresponds to that between two non-correlated distributions of clusters, therefore these two proteins neither co-localize nor anti-localize in chromatin when analyzed with a 2D method. Our Voronoi-based segmentation method is available in the ClusterViSu software.

At about the same time with our developments of ClusterViSu, several other methods for SMLM data segmentation appeared (Baumgart et al., 2016; Griffié et al., 2016; Levet et al., 2015; Rubin-Delanchy et al., 2015), but most of them were implemented as 2D methods often with a note that the 3D extension would be trivial. However, almost no techniques were implemented for 3D in practice in ready-to-use tools neither were discussed peculiarities of 3D data processing. We therefore decided to focus our attention on aspects of 3D cluster analysis of SMLM data. By analyzing the same dataset in 2D and in 3D I showed that 3D analysis is essential for getting correct results when imaging 3D objects (Andronov et al., 2017). Next, we extended our previously reported tessellation-based clustering method for 3D. We have shown that despite the increased complexity of 3D data the arguments known for 2D, *e.g.* the local density estimation and the determination of the segmentation threshold, are valid also for 3D. I implemented this method in a software, 3DClusterViSu, which became one of the first tools available for comprehensive cluster analysis of 3D localization data (Andronov et al., 2017).

At the beginning of my project, researchers only started applying SR microscopy for studies of chromatin. Due to the extreme complexity of the latter, simple SR imaging was not sufficient for drawing solid conclusions, therefore advanced data processing methods had to be developed before concrete results could be obtained. With development of such methods some important findings on the chromatin structure have been made with SMLM (Bohn et al., 2010; Ricci et al., 2015; Wang et al., 2014b), including studies of DNA repair pathways as function of the nuclear position, where I was involved for the SMLM part (Lemaître et al., 2014, see Appendix). Once I established SMLM data processing methods, we started to apply them for SR imaging of chromatin. Besides colocalization studies of H2B with RNA Pol II used as an illustration for ClusterViSu (Andronov et al., 2016b), we focused on the detailed structure of centromeric chromatin upon the CENP-A deposition.

Using first the data correction and visualization methods developed by me and then further applying advanced image processing methods, we demonstrated that (i) CENP-A forms clusters of a homogeneous size (with the diameter of around 260 nm) and shape in the centromeric regions of chromatin of human cells; (ii) these clusters have characteristic hollow shape in early G1 phase, which transforms into a filled shape with smooth borders upon the CENP-A deposition, according to the SMLM and cluster analysis of time points; (iii) CENP-A chaperone HJURP fills the cavity of the CENP-A clusters in early G1 phase. These are new findings that I report for the first time in the current thesis and which will be published in the near future.

4.2. Perspectives

Experimental techniques for further resolution improvement of fluorescence microscopy will continue to develop, therefore data processing methods have to keep pace with them. The SharpViSu software is designed with thought of further expansion when new processing algorithms will appear. On a short term, the support of 3D data can be improved in SharpViSu with 3D drift correction, more advanced 3D visualization and resolution estimation with Fourier shell correlation calculations. Processing speed of SharpViSu and 3DClusterViSu can be further optimized or parallelized. After polishing of 3DClusterViSu, it can be added as a plugin for SharpViSu as well.

The use of Voronoi tessellations can be developed further for *e.g.* 2D and 3D colocalization estimation directly from diagrams and not from segmented images as it is done currently (Andronov et al., 2016b).

We found that clusters of H2B do not correlate with that of RNA Pol II when imaged and analyzed with 2D SMLM methods. However, since RNA Pol II is associated with actively transcribed chromatin and H2B is denser in heterochromatin, the distributions of clusters of these proteins are likely to be mutually exclusive. As we have shown later, analyzing of 3D structures only in 2D can seriously distort results (Andronov et al., 2017), that may be the reason why no mutual exclusion could be found between the distributions of the two proteins. A similar experiment should be performed with a 3D SMLM and analyzed with *e.g.* 3DClusterViSu in order to refine the colocalization state of these proteins. We are going to do this in the near future.

Even though we found interesting patterns in the structure of centromeric chromatin at the resolution of 30~300 nm that surely correlate in time with the deposition of CENP-A, the underlying mechanism is still unknown. The structure of chromatin in general and of the centromeric chromatin in particular remains largely unclear and its study is complicated by the huge complexity of chromatin. In addition, centromeric chromatin has a turnover of CENP-A through the cell cycle, unusual for histones, which can obviously lead to a structure different from the rest of the chromatin that contains the classical histone H3. Many more studies have to be done before the structure of centromeric chromatin will be understood. This includes biochemical studies, purification of the components of centromeres, etc. We imaged synchronized cells only up to 8 hours after mitosis, but other shapes of CENP-A clusters than that present in the first 8 hours after mitosis were found in the non-synchronized cells. Hence,

it would make sense to continue imaging synchronized cells throughout the whole cell cycle. Some peculiarities of the centromeric chromatin structure can be found, *e.g.* during replication in S phase. Together, these studies could provide insights into the role of CENP-A and its histone-chaperone HJURP, the mechanism of insertion of CENP-A during replication and into the dynamic reorganization of chromatin structure during cell life.

5. References

Abbe, E. (1873). Ueber einen neuen Beleuchtungsapparat am Mikroskop. Arch. Für Mikrosk. Anat. *9*, 469–480.

Abraham, A.V., Ram, S., Chao, J., Ward, E.S., and Ober, R.J. (2009). Quantitative study of single molecule location estimation techniques. Opt. Express *17*, 23352–23373.

Aguet, F., Ville, D.V.D., and Unser, M. (2005). A maximum-likelihood formalism for sub-resolution axial localization of fluorescent nanoparticles. Opt. Express *13*, 10503–10522.

Airy, G.B. (1835). On the Diffraction of an Object-glass with Circular Aperture. Trans. Camb. Philos. Soc. *5*, 283.

Allen, J.R., Ross, S.T., and Davidson, M.W. (2013). Single molecule localization microscopy for superresolution. J. Opt. *15*, 094001.

Ambrose, E.J. (1956). A Surface Contact Microscope for the study of Cell Movements. Nature *178*, 1194–1194.

Ambrose, W.P., Basché, T., and Moerner, W.E. (1991). Detection and spectroscopy of single pentacene molecules in a p-terphenyl crystal by means of fluorescence excitation. J. Chem. Phys. *95*, 7150–7163.

Ambrose, W.P., Goodwin, P.M., and Nolan, J.P. (1999). Single-molecule detection with total internal reflection excitation: Comparing signal-to-background and total signals in different geometries. Cytometry *36*, 224–231.

Ando, R., Hama, H., Yamamoto-Hino, M., Mizuno, H., and Miyawaki, A. (2002). An optical marker based on the UV-induced green-to-red photoconversion of a fluorescent protein. Proc. Natl. Acad. Sci. *99*, 12651–12656.

Ando, R., Mizuno, H., and Miyawaki, A. (2004). Regulated Fast Nucleocytoplasmic Shuttling Observed by Reversible Protein Highlighting. Science *306*, 1370–1373.

Andronov, L., Lutz, Y., Vonesch, J.-L., and Klaholz, B.P. (2016a). SharpViSu: integrated analysis and segmentation of super-resolution microscopy data. Bioinformatics *32*, 2239–2241.

Andronov, L., Orlov, I., Lutz, Y., Vonesch, J.-L., and Klaholz, B.P. (2016b). ClusterViSu, a method for clustering of protein complexes by Voronoi tessellation in super-resolution microscopy. Sci. Rep. *6*, 24084.

Andronov, L., Michalon, J., Ouararhni, K., Orlov, I., Hamiche, A., Vonesch, J.-L., and Klaholz, B. (2017). 3D clustering analysis of super-resolution microscopy data by 3D Voronoi tessellations. bioRxiv 146456.

Aquino, D., Schönle, A., Geisler, C., Middendorff, C. v, Wurm, C.A., Okamura, Y., Lang, T., Hell, S.W., and Egner, A. (2011). Two-color nanoscopy of three-dimensional volumes by 4Pi detection of stochastically switched fluorophores. Nat. Methods *8*, 353–359.

Auksorius, E., Boruah, B.R., Dunsby, C., Lanigan, P.M.P., Kennedy, G., Neil, M.A.A., and French, P.M.W. (2008). Stimulated emission depletion microscopy with a supercontinuum source and fluorescence lifetime imaging. Opt. Lett. *33*, 113–115.

Axelrod, D. (1981). Cell-substrate contacts illuminated by total internal reflection fluorescence. J. Cell Biol. *89*, 141–145.

Backlund, M.P., Lew, M.D., Backer, A.S., Sahl, S.J., and Moerner, W.E. (2014). The Role of Molecular Dipole Orientation in Single-Molecule Fluorescence Microscopy and Implications for Super-Resolution Imaging. ChemPhysChem *15*, 587–599.

Baddeley, D., Cannell, M.B., and Soeller, C. (2010). Visualization of Localization Microscopy Data. Microsc. Microanal. *16*, 64–72.

Banterle, N., Bui, K.H., Lemke, E.A., and Beck, M. (2013). Fourier ring correlation as a resolution criterion for super-resolution microscopy. J. Struct. Biol. *183*, 363–367.

Bär, J., Kobler, O., Bommel, B. van, and Mikhaylova, M. (2016). Periodic F-actin structures shape the neck of dendritic spines. Sci. Rep. *6*, 37136.

Baranov, M.V., Revelo, N.H., Dingjan, I., Maraspini, R., Beest, M. ter, Honigmann, A., and Bogaart, G. van den (2016). SWAP70 Organizes the Actin Cytoskeleton and Is Essential for Phagocytosis. Cell Rep. *17*, 1518–1531.

Barlag, B., Beutel, O., Janning, D., Czarniak, F., Richter, C.P., Kommnick, C., Göser, V., Kurre, R., Fabiani, F., Erhardt, M., et al. (2016). Single molecule super-resolution imaging of proteins in living *Salmonella enterica* using self-labelling enzymes. Sci. Rep. *6*, srep31601.

Bates, M., Blosser, T.R., and Zhuang, X. (2005). Short-Range Spectroscopic Ruler Based on a Single-Molecule Optical Switch. Phys. Rev. Lett. *94*, 108101.

Bates, M., Huang, B., Dempsey, G.T., and Zhuang, X. (2007). Multicolor Super-Resolution Imaging with Photo-Switchable Fluorescent Probes. Science *317*, 1749–1753.

Baumgart, F., Arnold, A.M., Leskovar, K., Staszek, K., Fölser, M., Weghuber, J., Stockinger, H., and Schütz, G.J. (2016). Varying label density allows artifact-free analysis of membrane-protein nanoclusters. Nat. Methods *13*, 661–664.

Bednar, J., Garcia-Saez, I., Boopathi, R., Cutter, A.R., Papai, G., Reymer, A., Syed, S.H., Lone, I.N., Tonchev, O., Crucifix, C., et al. (2017). Structure and Dynamics of a 197 bp Nucleosome in Complex with Linker Histone H1. Mol. Cell *66*, 384–397.e8.

Bergermann, F., Alber, L., Sahl, S.J., Engelhardt, J., and Hell, S.W. (2015). 2000-fold parallelized dual-color STED fluorescence nanoscopy. Opt. Express *23*, 211–223.

Besag, J., and Diggle, P.J. (1977). Simple Monte Carlo Tests for Spatial Pattern. J. R. Stat. Soc. Ser. C Appl. Stat. *26*, 327–333.

Bethge, P., Chéreau, R., Avignone, E., Marsicano, G., and Nägerl, U.V. (2013). Two-Photon Excitation STED Microscopy in Two Colors in Acute Brain Slices. Biophys. J. *104*, 778–785.

Betzig, E. (1995). Proposed method for molecular optical imaging. Opt. Lett. *20*, 237–239.

Betzig, E., Patterson, G.H., Sougrat, R., Lindwasser, O.W., Olenych, S., Bonifacino, J.S., Davidson, M.W., Lippincott-Schwartz, J., and Hess, H.F. (2006). Imaging Intracellular Fluorescent Proteins at Nanometer Resolution. Science *313*, 1642–1645.

Bewersdorf, J., Schmidt, R., and Hell, S.W. (2006). Comparison of I5M and 4Pi-microscopy. J. Microsc. *222*, 105–117.

Bielopolski, N., Lam, A.D., Bar-On, D., Sauer, M., Stuenkel, E.L., and Ashery, U. (2014). Differential Interaction of Tomosyn with Syntaxin and SNAP25 Depends on Domains in the WD40 β-Propeller Core and Determines Its Inhibitory Activity. J. Biol. Chem. *289*, 17087–17099.

Black, B.E., Jansen, L.E.T., Maddox, P.S., Foltz, D.R., Desai, A.B., Shah, J.V., and Cleveland, D.W. (2007). Centromere Identity Maintained by Nucleosomes Assembled with Histone H3 Containing the CENP-A Targeting Domain. Mol. Cell *25*, 309–322.

Blower, M.D., Sullivan, B.A., and Karpen, G.H. (2002). Conserved Organization of Centromeric Chromatin in Flies and Humans. Dev. Cell *2*, 319–330.

Bock, H., Geisler, C., Wurm, C.A., Middendorff, C. von, Jakobs, S., Schönle, A., Egner, A., Hell, S.W., and Eggeling, C. (2007). Two-color far-field fluorescence nanoscopy based on photoswitchable emitters. Appl. Phys. B *88*, 161–165.

Bodor, D.L., Valente, L.P., Mata, J.F., Black, B.E., and Jansen, L.E.T. (2013). Assembly in G1 phase and long-term stability are unique intrinsic features of CENP-A nucleosomes. Mol. Biol. Cell *24*, 923–932.

Bodor, D.L., Mata, J.F., Sergeev, M., David, A.F., Salimian, K.J., Panchenko, T., Cleveland, D.W., Black, B.E., Shah, J.V., and Jansen, L.E. (2014). The quantitative architecture of centromeric chromatin. eLife *3*, e02137.

Boettcher, R., Haig, S.M., and Bridges, W.C. (1994). Behavioral Patterns and Nearest Neighbor Distances among Nonbreeding American Avocets. The Condor *96*, 973–986.

Boettiger, A.N., Bintu, B., Moffitt, J.R., Wang, S., Beliveau, B.J., Fudenberg, G., Imakaev, M., Mirny, L.A., Wu, C., and Zhuang, X. (2016). Super-resolution imaging reveals distinct chromatin folding for different epigenetic states. Nature *529*, 418–422.

Bohn, M., Diesinger, P., Kaufmann, R., Weiland, Y., Müller, P., Gunkel, M., von Ketteler, A., Lemmer, P., Hausmann, M., Heermann, D.W., et al. (2010). Localization Microscopy Reveals Expression-Dependent Parameters of Chromatin Nanostructure. Biophys. J. *99*, 1358–1367.

Bossi, M., Fölling, J., Belov, V.N., Boyarskiy, V.P., Medda, R., Egner, A., Eggeling, C., Schönle, A., and Hell, S.W. (2008). Multicolor Far-Field Fluorescence Nanoscopy through Isolated Detection of Distinct Molecular Species. Nano Lett. *8*, 2463–2468.

Bretschneider, S., Eggeling, C., and Hell, S.W. (2007). Breaking the Diffraction Barrier in Fluorescence Microscopy by Optical Shelving. Phys. Rev. Lett. *98*, 218103.

Brumberg, E.M., and Krylova, T.N. (1953). Primenenie interferentsionnykh delitelnykh zerkal v fluorestsentnoi mikroskopii. Zh Obshch Biol *14*, 461–464.

Bückers, J., Wildanger, D., Vicidomini, G., Kastrup, L., and Hell, S.W. (2011). Simultaneous multi-lifetime multi-color STED imaging for colocalization analyses. Opt. Express *19*, 3130–3143.

Bui, M., Dimitriadis, E.K., Hoischen, C., An, E., Quénet, D., Giebe, S., Nita-Lazar, A., Diekmann, S., and Dalal, Y. (2012). Cell-Cycle-Dependent Structural Transitions in the Human CENP-A Nucleosome In Vivo. Cell *150*, 317–326.

Burmeister, J.S., Olivier, L.A., Reichert, W.M., and Truskey, G.A. (1998). Application of total internal reflection fluorescence microscopy to study cell adhesion to biomaterials. Biomaterials *19*, 307–325.

Burns, D.H., Callis, J.B., Christian, G.D., and Davidson, E.R. (1985). Strategies for attaining superresolution using spectroscopic data as constraints. Appl. Opt. *24*, 154–161.

Carroll, C.W., Milks, K.J., and Straight, A.F. (2010). Dual recognition of CENP-A nucleosomes is required for centromere assembly. J. Cell Biol. *189*, 1143–1155.

Carter, A.R., King, G.M., Ulrich, T.A., Halsey, W., Alchenberger, D., and Perkins, T.T. (2007). Stabilization of an optical microscope to 0.1 nm in three dimensions. Appl. Opt. *46*, 421–427.

Castro, M.A.G. de, Höbartner, C., and Opazo, F. (2017). Aptamers provide superior stainings of cellular receptors studied under super-resolution microscopy. PLOS ONE *12*, e0173050.

Cheezum, M.K., Walker, W.F., and Guilford, W.H. (2001). Quantitative Comparison of Algorithms for Tracking Single Fluorescent Particles. Biophys. J. *81*, 2378–2388.

Chen, J., Gao, J., Zhang, M., Cai, M., Xu, H., Jiang, J., Tian, Z., and Wang, H. (2016). Systemic localization of seven major types of carbohydrates on cell membranes by dSTORM imaging. Sci. Rep. *6*, srep30247.

Chéreau, R., Saraceno, G.E., Angibaud, J., Cattaert, D., and Nägerl, U.V. (2017). Superresolution imaging reveals activity-dependent plasticity of axon morphology linked to changes in action potential conduction velocity. Proc. Natl. Acad. Sci. *114*, 1401–1406.

Choi, J., and Kim, D. (2013). Notch spatial filtering of image artifacts for structured illumination microscopy of cell-based assays. Opt. Commun. *308*, 142–146.

Choi, C.K., Vicente-Manzanares, M., Zareno, J., Whitmore, L.A., Mogilner, A., and Horwitz, A.R. (2008). Actin and alpha-actinin orchestrate the assembly and maturation of nascent adhesions in a myosin II motor-independent manner. Nat. Cell Biol. *10*, 1039-U36.

Chudakov, D.M., Belousov, V.V., Zaraisky, A.G., Novoselov, V.V., Staroverov, D.B., Zorov, D.B., Lukyanov, S., and Lukyanov, K.A. (2003). Kindling fluorescent proteins for precise in vivo photolabeling. Nat. Biotechnol. *21*, 191–194.

Churchman, L.S., Ökten, Z., Rock, R.S., Dawson, J.F., and Spudich, J.A. (2005). Single molecule high-resolution colocalization of Cy3 and Cy5 attached to macromolecules measures intramolecular distances through time. Proc. Natl. Acad. Sci. U. S. A. *102*, 1419–1423.

Clark, P.J., and Evans, F.C. (1954). Distance to Nearest Neighbor as a Measure of Spatial Relationships in Populations. Ecology *35*, 445–453.

Coltharp, C., Yang, X., and Xiao, J. (2014). Quantitative analysis of single-molecule superresolution images. Curr. Opin. Struct. Biol. *28*, 112–121.

Cordes, T., Strackharn, M., Stahl, S.W., Summerer, W., Steinhauer, C., Forthmann, C., Puchner, E.M., Vogelsang, J., Gaub, H.E., and Tinnefeld, P. (2010). Resolving Single-Molecule Assembled Patterns with Superresolution Blink-Microscopy. Nano Lett. *10*, 645–651.

Dalal, Y., Wang, H., Lindsay, S., and Henikoff, S. (2007). Tetrameric Structure of Centromeric Nucleosomes in Interphase Drosophila Cells. PLOS Biol. *5*, e218.

Danzl, J.G., Sidenstein, S.C., Gregor, C., Urban, N.T., Ilgen, P., Jakobs, S., and Hell, S.W. (2016). Coordinate-targeted fluorescence nanoscopy with multiple off states. Nat. Photonics *10*, 122–128.

Davidovits, P., and Egger, M.D. (1971). Scanning Laser Microscope for Biological Investigations. Appl. Opt. *10*, 1615–1619.

Dempsey, G.T., Vaughan, J.C., Chen, K.H., Bates, M., and Zhuang, X. (2011). Evaluation of fluorophores for optimal performance in localization-based super-resolution imaging. Nat. Methods *8*, 1027–1036.

Deschout, H., Zanacchi, F.C., Mlodzianoski, M., Diaspro, A., Bewersdorf, J., Hess, S.T., and Braeckmans, K. (2014). Precisely and accurately localizing single emitters in fluorescence microscopy. Nat. Methods *11*, 253–266.

D'Este, E., Kamin, D., Balzarotti, F., and Hell, S.W. (2017). Ultrastructural anatomy of nodes of Ranvier in the peripheral nervous system as revealed by STED microscopy. Proc. Natl. Acad. Sci. *114*, E191–E199.

Dickson, R.M., Cubitt, A.B., Tsien, R.Y., and Moerner, W.E. (1997). On/off blinking and switching behaviour of single molecules of green fluorescent protein. Nature *388*, 355–358.

Díez, L.T., Bönsch, C., Malkusch, S., Truan, Z., Munteanu, M., Heilemann, M., Hartley, O., Endesfelder, U., and Fürstenberg, A. (2014). Coordinate-based co-localization-mediated analysis of arrestin clustering upon stimulation of the C–C chemokine receptor 5 with RANTES/CCL5 analogues. Histochem. Cell Biol. *142*, 69–77.

Dimitriadis, E.K., Weber, C., Gill, R.K., Diekmann, S., and Dalal, Y. (2010). Tetrameric organization of vertebrate centromeric nucleosomes. Proc. Natl. Acad. Sci. *107*, 20317–20322.

Dirichlet, G.L. (1850). Über die Reduction der positiven quadratischen Formen mit drei unbestimmten ganzen Zahlen. J. Für Reine Angew. Math. *40*, 209–227.

Dixon, P. (2001). Ripley's K function. Encycl. Environmetrics 1796–1803.

Dunleavy, E.M., Roche, D., Tagami, H., Lacoste, N., Ray-Gallet, D., Nakamura, Y., Daigo, Y., Nakatani, Y., and Almouzni-Pettinotti, G. (2009). HJURP Is a Cell-Cycle-Dependent Maintenance and Deposition Factor of CENP-A at Centromeres. Cell *137*, 485–497.

Dunleavy, E.M., Almouzni, G., and Karpen, G.H. (2011). H3.3 is deposited at centromeres in S phase as a placeholder for newly assembled CENP-A in G1 phase. Nucleus *2*, 146–157.

Dunleavy, E.M., Zhang, W., and Karpen, G.H. (2013). Solo or doppio: how many CENP-As make a centromeric nucleosome? Nat. Struct. Mol. Biol. *20*, 648–650.

Dunn, K.W., Kamocka, M.M., and McDonald, J.H. (2011). A practical guide to evaluating colocalization in biological microscopy. Am. J. Physiol. - Cell Physiol. *300*, C723–C742.

Eggeling, C., Ringemann, C., Medda, R., Schwarzmann, G., Sandhoff, K., Polyakova, S., Belov, V.N., Hein, B., von Middendorff, C., Schönle, A., et al. (2009). Direct observation of the nanoscale dynamics of membrane lipids in a living cell. Nature *457*, 1159–1162.

Egner, A., Geisler, C., von Middendorff, C., Bock, H., Wenzel, D., Medda, R., Andresen, M., Stiel, A.C., Jakobs, S., Eggeling, C., et al. (2007). Fluorescence Nanoscopy in Whole Cells by Asynchronous Localization of Photoswitching Emitters. Biophys. J. *93*, 3285–3290.

Ellinger, P., and Hirt, A. (1929). Mikroskopische Beobachtungen an lebenden Organen mit Demonstrationen (Intravitalmikroskopie). Arch Exp Pathol Phar *147*, 63.

Elmlund, D., and Elmlund, H. (2012). SIMPLE: Software for ab initio reconstruction of heterogeneous single-particles. J. Struct. Biol. *180*, 420–427.

Enderlein, J., Toprak, E., and Selvin, P.R. (2006). Polarization effect on position accuracy of fluorophore localization. Opt. Express *14*, 8111–8120.

Endesfelder, U., Finan, K., Holden, S.J., Cook, P.R., Kapanidis, A.N., and Heilemann, M. (2013). Multiscale Spatial Organization of RNA Polymerase in Escherichia coli. Biophys. J. *105*, 172–181.

Engelhardt, J., Keller, J., Hoyer, P., Reuss, M., Staudt, T., and Hell, S.W. (2011). Molecular Orientation Affects Localization Accuracy in Superresolution Far-Field Fluorescence Microscopy. Nano Lett. *11*, 209–213.

Erdelyi, M., Rees, E., Metcalf, D., Schierle, G.S.K., Dudas, L., Sinko, J., Knight, A.E., and Kaminski, C.F. (2013). Correcting chromatic offset in multicolor super-resolution localization microscopy. Opt. Express *21*, 10978–10988.

Erdélyi, M., Sinkó, J., Kákonyi, R., Kelemen, A., Rees, E., Varga, D., and Szabó, G. (2015). Origin and compensation of imaging artefacts in localization-based super-resolution microscopy. Methods *88*, 122–132.

Ester, M., Kriegel, H.-P., Sander, J., and Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. (AAAI Press), pp. 226–231.

Fitzgerald, J.E., Lu, J., and Schnitzer, M.J. (2012). Estimation Theoretic Measure of Resolution for Stochastic Localization Microscopy. Phys. Rev. Lett. *109*, 048102.

Flors, C., Ravarani, C.N.J., and Dryden, D.T.F. (2009). Super-Resolution Imaging of DNA Labelled with Intercalating Dyes. ChemPhysChem *10*, 2201–2204.

Fölling, J., Bossi, M., Bock, H., Medda, R., Wurm, C.A., Hein, B., Jakobs, S., Eggeling, C., and Hell, S.W. (2008). Fluorescence nanoscopy by ground-state depletion and single-molecule return. Nat. Methods *5*, 943–945.

Foltz, D.R., Jansen, L.E.T., Bailey, A.O., Yates, J.R., Bassett, E.A., Wood, S., Black, B.E., and Cleveland, D.W. (2009). Centromere-Specific Assembly of CENP-A Nucleosomes Is Mediated by HJURP. Cell *137*, 472–484.

Foreman, M.R., Romero, C.M., and Török, P. (2008). Determination of the three-dimensional orientation of single molecules. Opt. Lett. *33*, 1020–1022.

Förster, R., Wicker, K., Müller, W., Jost, A., and Heintzmann, R. (2016). Motion artefact detection in structured illumination microscopy for live cell imaging. Opt. Express *24*, 22121–22134.

Fukagawa, T., and Earnshaw, W.C. (2014). The Centromere: Chromatin Foundation for the Kinetochore Machinery. Dev. Cell *30*, 496–508.

Fussner, E., Ching, R.W., and Bazett-Jones, D.P. (2011). Living without 30nm chromatin fibers. Trends Biochem. Sci. *36*, 1–6.

Geisler, C., Schönle, A., Middendorff, C. von, Bock, H., Eggeling, C., Egner, A., and Hell, S.W. (2007). Resolution of $\lambda/10$ in fluorescence microscopy using fast single molecule photo-switching. Appl. Phys. A *88*, 223–226.

Geisler, C., Hotz, T., Schönle, A., Hell, S.W., Munk, A., and Egner, A. (2012). Drift estimation for single marker switching based imaging schemes. Opt. Express *20*, 7274–7289.

Geiss, C.P., Keramisanou, D., Sekulic, N., Scheffer, M.P., Black, B.E., and Frangakis, A.S. (2014). CENP-A Arrays Are More Condensed than Canonical Arrays at Low Ionic Strength. Biophys. J. *106*, 875–882.

Gelles, J., Schnapp, B.J., and Sheetz, M.P. (1988). Tracking kinesin-driven movements with nanometre-scale precision. Nature *331*, 450–453.

Georgieva, M., Cattoni, D.I., Fiche, J.-B., Mutin, T., Chamousset, D., and Nollmann, M. (2016). Nanometer resolved single-molecule colocalization of nuclear factors by two-color super resolution microscopy imaging. Methods *105*, 44–55.

Giannone, G., Hosy, E., Levet, F., Constals, A., Schulze, K., Sobolevsky, A.I., Rosconi, M.P., Gouaux, E., Tampé, R., Choquet, D., et al. (2010). Dynamic Superresolution Imaging of Endogenous Proteins on Living Cells at Ultra-High Density. Biophys. J. *99*, 1303–1310.

Gordon, M.P., Ha, T., and Selvin, P.R. (2004). Single-molecule high-resolution imaging with photobleaching. Proc. Natl. Acad. Sci. U. S. A. *101*, 6462–6465.

Göttfert, F., Pleiner, T., Heine, J., Westphal, V., Görlich, D., Sahl, S.J., and Hell, S.W. (2017). Strong signal increase in STED fluorescence microscopy by imaging regions of subdiffraction extent. Proc. Natl. Acad. Sci. *114*, 2125–2130.

Griffié, J., Shannon, M., Bromley, C.L., Boelen, L., Burn, G.L., Williamson, D.J., Heard, N.A., Cope, A.P., Owen, D.M., and Rubin-Delanchy, P. (2016). A Bayesian cluster analysis method for single-molecule localization microscopy data. Nat. Protoc. *11*, 2499–2514.

Grotjohann, T., Testa, I., Leutenegger, M., Bock, H., Urban, N.T., Lavoie-Cardinal, F., Willig, K.I., Eggeling, C., Jakobs, S., and Hell, S.W. (2011). Diffraction-unlimited all-optical imaging and writing with a photochromic GFP. Nature *478*, 204–208.

Grove, J., Metcalf, D.J., Knight, A.E., Wavre-Shapton, S.T., Sun, T., Protonotarios, E.D., Griffin, L.D., Lippincott-Schwartz, J., and Marsh, M. (2014). Flat clathrin lattices: stable features of the plasma membrane. Mol. Biol. Cell *25*, 3581–3594.

Guerra, M., Cabral, G., Cuacos, M., González-García, M., González-Sánchez, M., Vega, J., and Puertas, M.J. (2010). Neocentrics and Holokinetics (Holocentrics): Chromosomes out of the Centromeric Rules. Cytogenet. Genome Res. *129*, 82–96.

Guizar-Sicairos, M., Thurman, S.T., and Fienup, J.R. (2008). Efficient subpixel image registration algorithms. Opt. Lett. *33*, 156–158.

Gurskaya, N.G., Verkhusha, V.V., Shcheglov, A.S., Staroverov, D.B., Chepurnykh, T.V., Fradkov, A.F., Lukyanov, S., and Lukyanov, K.A. (2006). Engineering of a monomeric green-to-red photoactivatable fluorescent protein induced by blue light. Nat. Biotechnol. *24*, 461–465.

Gustafsson, M.G.L. (2000). Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy. J. Microsc. *198*, 82–87.

Gustafsson, M.G.L. (2005). Nonlinear structured-illumination microscopy: Wide-field fluorescence imaging with theoretically unlimited resolution. Proc. Natl. Acad. Sci. U. S. A. *102*, 13081–13086.

Gustafsson, Agard, and Sedat (1999). I5M: 3D widefield light microscopy with better than 100 nm axial resolution. J. Microsc. *195*, 10–16.

Gustafsson, M.G.L., Agard, D.A., and Sedat, J.W. (1995). Sevenfold improvement of axial resolution in 3D wide-field microscopy using two objective lenses. pp. 147–156.

Gustafsson, M.G.L., Agard, D.A., and Sedat, J.W. (1996). 3D widefield microscopy with two objective lenses: experimental verification of improved axial resolution. pp. 62–66.

Gustafsson, M.G.L., Shao, L., Carlton, P.M., Wang, C.J.R., Golubovskaya, I.N., Cande, W.Z., Agard, D.A., and Sedat, J.W. (2008). Three-Dimensional Resolution Doubling in Wide-Field Fluorescence Microscopy by Structured Illumination. Biophys. J. *94*, 4957–4970.

Hamers-Casterman, C., Atarhouch, T., Muyldermans, S., Robinson, G., Hamers, C., Songa, E.B., Bendahman, N., and Hamers, R. (1993). Naturally occurring antibodies devoid of light chains. Nature *363*, 446–448.

Hanne, J., Göttfert, F., Schimer, J., Anders-Össwein, M., Konvalinka, J., Engelhardt, J., Müller, B., Hell, S.W., and Kräusslich, H.-G. (2016). Stimulated Emission Depletion

Nanoscopy Reveals Time-Course of Human Immunodeficiency Virus Proteolytic Maturation. ACS Nano *10*, 8215–8222.

Hänninen, P.E., Hell, S.W., Salo, J., Soini, E., and Cremer, C. (1995). Two-photon excitation 4Pi confocal microscope: Enhanced axial resolution microscope for biological research. Appl. Phys. Lett. *66*, 1698–1700.

Harke, B., Keller, J., Ullal, C.K., Westphal, V., Schönle, A., and Hell, S.W. (2008). Resolution scaling in STED microscopy. Opt. Express *16*, 4154–4162.

Harmsen, M.M., and Haard, H.J.D. (2007). Properties, production, and applications of camelid single-domain antibody fragments. Appl. Microbiol. Biotechnol. *77*, 13–22.

Harris, L.J., Skaletsky, E., and McPherson, A. (1998). Crystallographic structure of an intact IgG1 monoclonal antibody. J. Mol. Biol. *275*, 861–872.

van Heel, M., Keegstra, W., Schutter, W., and van Bruggen, E.F.J. (1982). Arthropod hemocyanin structures studied by image analysis. In Structure and Function of Invertebrate Respiratory Proteins, Life Chemistry Reports Suppl. 1, (Harwood, London: Wood, E.J.), pp. 69–73.

van Heel, M., Harauz, G., Orlova, E.V., Schmidt, R., and Schatz, M. (1996). A new generation of the IMAGIC image processing system. J. Struct. Biol. *116*, 17–24.

Heilemann, M., Margeat, E., Kasper, R., Sauer, M., and Tinnefeld, P. (2005). Carbocyanine Dyes as Efficient Reversible Single-Molecule Optical Switch. J. Am. Chem. Soc. *127*, 3801–3806.

Heilemann, M., van de Linde, S., Schüttpelz, M., Kasper, R., Seefeldt, B., Mukherjee, A., Tinnefeld, P., and Sauer, M. (2008). Subdiffraction-Resolution Fluorescence Imaging with Conventional Fluorescent Probes. Angew. Chem. Int. Ed. *47*, 6172–6176.

Heilemann, M., van de Linde, S., Mukherjee, A., and Sauer, M. (2009). Super-Resolution Imaging with Small Organic Fluorophores. Angew. Chem. Int. Ed. *48*, 6903–6908.

Heintzmann, R., Jovin, T.M., and Cremer, C. (2002). Saturated patterned excitation microscopy—a concept for optical resolution improvement. JOSA A *19*, 1599–1609.

Hell, S., and Stelzer, E.H.K. (1992). Properties of a 4Pi confocal fluorescence microscope. JOSA A *9*, 2159–2166.

Hell, S.W., and Kroug, M. (1995). Ground-state-depletion fluorscence microscopy: A concept for breaking the diffraction resolution limit. Appl. Phys. B *60*, 495–497.

Hell, S.W., and Wichmann, J. (1994). Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy. Opt. Lett. *19*, 780–782.

Hell, S.W., Lindek, S., Cremer, C., and Stelzer, E.H.K. (1994). Measurement of the 4Pi-confocal point spread function proves 75 nm axial resolution. Appl. Phys. Lett. *64*, 1335–1337.

Hell, S.W., Dyba, M., and Jakobs, S. (2004). Concepts for nanoscale resolution in fluorescence microscopy. Curr. Opin. Neurobiol. *14*, 599–609.

Hell, S.W., Schmidt, R., and Egner, A. (2009). Diffraction-unlimited three-dimensional optical nanoscopy with opposing lenses. Nat. Photonics *3*, 381–387.

Henriques, R., Lelek, M., Fornasiero, E.F., Valtorta, F., Zimmer, C., and Mhlanga, M.M. (2010). QuickPALM: 3D real-time photoactivation nanoscopy image processing in ImageJ. Nat. Methods *7*, 339–340.

Hess, S.T., Girirajan, T.P.K., and Mason, M.D. (2006). Ultra-High Resolution Imaging by Fluorescence Photoactivation Localization Microscopy. Biophys. J. *91*, 4258–4272.

Hirvonen, L., Mandula, O., Wicker, K., and Heintzmann, R. (2008). Structured illumination microscopy using photoswitchable fluorescent proteins. p. 68610L–68610L–8.

Holden, S.J., Uphoff, S., and Kapanidis, A.N. (2011). DAOSTORM: an algorithm for high-density super-resolution microscopy. Nat. Methods *8*, 279–280.

Holtzer, L., Meckel, T., and Schmidt, T. (2007). Nanometric three-dimensional tracking of individual quantum dots in cells. Appl. Phys. Lett. *90*, 053902.

Honigmann, A., Mueller, V., Ta, H., Schoenle, A., Sezgin, E., Hell, S.W., and Eggeling, C. (2014). Scanning STED-FCS reveals spatiotemporal heterogeneity of lipid interaction in the plasma membrane of living cells. Nat. Commun. *5*, 5412.

Hoogendoorn, E., Crosby, K.C., Leyton-Puig, D., Breedijk, R.M.P., Jalink, K., Gadella, T.W.J., and Postma, M. (2014). The fidelity of stochastic single-molecule super-resolution reconstructions critically depends upon robust background estimation. Sci. Rep. *4*, 3854.

Horn, A.E., Kugel, J.F., and Goodrich, J.A. (2016). Single molecule microscopy reveals mechanistic insight into RNA polymerase II preinitiation complex assembly and transcriptional activity. Nucleic Acids Res. *44*, 7132–7143.

Huang, B., Wang, W., Bates, M., and Zhuang, X. (2008a). Three-Dimensional Super-Resolution Imaging by Stochastic Optical Reconstruction Microscopy. Science *319*, 810–813.

Huang, B., Jones, S.A., Brandenburg, B., and Zhuang, X. (2008b). Whole-cell 3D STORM reveals interactions between cellular structures with nanometer-scale resolution. Nat. Methods *5*, 1047–1052.

Huang, F., Schwartz, S.L., Byars, J.M., and Lidke, K.A. (2011). Simultaneous multiple-emitter fitting for single molecule super-resolution imaging. Biomed. Opt. Express *2*, 1377–1393.

Huang, F., Hartwich, T.M.P., Rivera-Molina, F.E., Lin, Y., Duim, W.C., Long, J.J., Uchil, P.D., Myers, J.R., Baird, M.A., Mothes, W., et al. (2013). Video-rate nanoscopy using sCMOS camera-specific single-molecule localization algorithms. Nat. Methods *10*, 653–658.

Huang, F., Sirinakis, G., Allgeyer, E.S., Schroeder, L.K., Duim, W.C., Kromann, E.B., Phan, T., Rivera-Molina, F.E., Myers, J.R., Irnov, I., et al. (2016). Ultra-High Resolution 3D Imaging of Whole Cells. Cell *166*, 1028–1040.

Huff, J. (2015). The Airyscan detector from ZEISS: confocal imaging with improved signal-to-noise ratio and super-resolution. Nat. Methods *12*.

Irie, M., Fukaminato, T., Sasaki, T., Tamai, N., and Kawai, T. (2002). Organic chemistry: A digital fluorescent molecular photoswitch. Nature *420*, 759–760.

Izeddin, I., El Beheiry, M., Andilla, J., Ciepielewski, D., Darzacq, X., and Dahan, M. (2012). PSF shaping using adaptive optics for three-dimensional single-molecule super-resolution imaging and tracking. Opt. Express *20*, 4957.

Jabłoński, A. (1933). Efficiency of Anti-Stokes Fluorescence in Dyes. Nature *131*, 839–840.

Jansen, L.E.T., Black, B.E., Foltz, D.R., and Cleveland, D.W. (2007). Propagation of centromeric chromatin requires exit from mitosis. J. Cell Biol. *176*, 795–805.

Jones, S.A., Shim, S.-H., He, J., and Zhuang, X. (2011). Fast, three-dimensional super-resolution imaging of live cells. Nat. Methods *8*, 499–505.

Joti, Y., Hikima, T., Nishino, Y., Kamada, F., Hihara, S., Takata, H., Ishikawa, T., and Maeshima, K. (2012). Chromosomes without a 30-nm chromatin fiber. Nucleus *3*, 404–410.

Juette, M.F., Gould, T.J., Lessard, M.D., Mlodzianoski, M.J., Nagpure, B.S., Bennett, B.T., Hess, S.T., and Bewersdorf, J. (2008). Three-dimensional sub–100 nm resolution fluorescence microscopy of thick samples. Nat. Methods *5*, 527–529.

Juškaitis, and Wilson (1999). A method for characterizing longitudinal chromatic aberration of microscope objectives using a confocal optical system. J. Microsc. *195*, 17–22.

Kanchanawong, P., Shtengel, G., Pasapera, A.M., Ramko, E.B., Davidson, M.W., Hess, H.F., and Waterman, C.M. (2010). Nanoscale architecture of integrin-based cell adhesions. Nature *468*, 580–584.

Kao, H.P., and Verkman, A.S. (1994). Tracking of single fluorescent particles in three dimensions: use of cylindrical optics to encode particle position. Biophys. J. *67*, 1291–1300.

Kapuscinski, J. (1995). DAPI: a DNA-Specific Fluorescent Probe. Biotech. Histochem. *70*, 220–233.

Karadaglić, D., and Wilson, T. (2008). Image formation in structured illumination wide-field fluorescence microscopy. Micron *39*, 808–818.

Keppler, A., Gendreizig, S., Gronemeyer, T., Pick, H., Vogel, H., and Johnsson, K. (2003). A general method for the covalent labeling of fusion proteins with small molecules in vivo. Nat. Biotechnol. *21*, 86–89.

Kiskowski, M.A., Hancock, J.F., and Kenworthy, A.K. (2009). On the Use of Ripley's K-Function and Its Derivatives to Analyze Domain Size. Biophys. J. *97*, 1095–1103.

Klar, T.A., Jakobs, S., Dyba, M., Egner, A., and Hell, S.W. (2000). Fluorescence microscopy with diffraction resolution barrier broken by stimulated emission. Proc. Natl. Acad. Sci. *97*, 8206–8210.

Kozubek, M., and Matula, P. (2000). An efficient algorithm for measurement and correction of chromatic aberrations in fluorescence microscopy. J. Microsc. *200*, 206–217.

Kriegel, H.-P., Kröger, P., Sander, J., and Zimek, A. (2011). Density-based clustering. Wiley Interdiscip. Rev. Data Min. Knowl. Discov. *1*, 231–240.

Lacoste, T.D., Michalet, X., Pinaud, F., Chemla, D.S., Alivisatos, A.P., and Weiss, S. (2000). Ultrahigh-resolution multicolor colocalization of single fluorescent probes. Proc. Natl. Acad. Sci. *97*, 9461–9466.

Lakowicz, J.R. (2006). Fluorescence Anisotropy. In Principles of Fluorescence Spectroscopy, (Springer US), pp. 353–382.

Larghi, P., Williamson, D.J., Carpier, J.-M., Dogniaux, S., Chemin, K., Bohineust, A., Danglot, L., Gaus, K., Galli, T., and Hivroz, C. (2013). VAMP7 controls T cell activation by regulating the recruitment and phosphorylation of vesicular Lat at TCR-activation sites. Nat. Immunol. *14*, 723–731.

Leising, A.W., and Yen, J. (1997). Spacing mechanisms within light-induced copepod swarms. Mar. Ecol. Prog. Ser. *155*, 127–135.

Lemaître, C., Grabarz, A., Tsouroula, K., Andronov, L., Furst, A., Pankotai, T., Heyer, V., Rogier, M., Attwood, K.M., Kessler, P., et al. (2014). Nuclear position dictates DNA repair pathway choice. Genes Dev. *28*, 2450–2463.

Lesoine, M.D., Bose, S., Petrich, J.W., and Smith, E.A. (2012). Supercontinuum Stimulated Emission Depletion Fluorescence Lifetime Imaging. J. Phys. Chem. B *116*, 7821–7826.

Levet, F., Hosy, E., Kechkar, A., Butler, C., Beghin, A., Choquet, D., and Sibarita, J.-B. (2015). SR-Tesseler: a method to segment and quantify localization-based super-resolution microscopy data. Nat. Methods *12*, 1065–1071.

Lew, M.D., Thompson, M.A., Badieirostami, M., and Moerner, W.E. (2010). In vivo Three-Dimensional Superresolution Fluorescence Tracking using a Double-Helix Point Spread Function. Proc. SPIE-- Int. Soc. Opt. Eng. *7571*, 75710Z.

Lew, M.D., Lee, S.F., Ptacin, J.L., Lee, M.K., Twieg, R.J., Shapiro, L., and Moerner, W.E. (2011). Three-dimensional superresolution colocalization of intracellular protein superstructures and the cell surface in live Caulobacter crescentus. Proc. Natl. Acad. Sci. *108*, E1102–E1110.

Li, D., Shao, L., Chen, B.-C., Zhang, X., Zhang, M., Moses, B., Milkie, D.E., Beach, J.R., Hammer, J.A., Pasham, M., et al. (2015). Extended-resolution structured illumination imaging of endocytic and cytoskeletal dynamics. Science *349*, aab3500.

Liao, H.Y., and Frank, J. (2010). Definition and estimation of resolution in single-particle reconstructions. Struct. Lond. Engl. 1993 *18*, 768–775.

Lieberman-Aiden, E., Berkum, N.L. van, Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O., et al. (2009). Comprehensive Mapping of

Long-Range Interactions Reveals Folding Principles of the Human Genome. Science *326*, 289–293.

Linde, S. van de, Kasper, R., Heilemann, M., and Sauer, M. (2008). Photoswitching microscopy with standard fluorophores. Appl. Phys. B *93*, 725.

Linde, S. van de, Endesfelder, U., Mukherjee, A., Schüttpelz, M., Wiebusch, G., Wolter, S., Heilemann, M., and Sauer, M. (2009). Multicolor photoswitching microscopy for subdiffraction-resolution fluorescence imaging. Photochem. Photobiol. Sci. *8*, 465–469.

Lopes, F.B., Bálint, Š., Valvo, S., Felce, J.H., Hessel, E.M., Dustin, M.L., and Davis, D.M. (2017). Membrane nanoclusters of FcγRI segregate from inhibitory SIRPα upon activation of human macrophages. J Cell Biol jcb.201608094.

Lord Rayleigh (1879). XXXI. Investigations in optics, with special reference to the spectroscope. Philos. Mag. Ser. 5 *8*, 261–274.

Los, G.V., Encell, L.P., McDougall, M.G., Hartzell, D.D., Karassina, N., Zimprich, C., Wood, M.G., Learish, R., Ohana, R.F., Urh, M., et al. (2008). HaloTag: A Novel Protein Labeling Technology for Cell Imaging and Protein Analysis. ACS Chem. Biol. *3*, 373–382.

Luger, K., Mäder, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. Nature *389*, 251–260.

Lu-Walther, H.-W., Hou, W., Kielhorn, M., Arai, Y., Nagai, T., Kessels, M.M., Qualmann, B., and Heintzmann, R. (2016). Nonlinear Structured Illumination Using a Fluorescent Protein Activating at the Readout Wavelength. PLOS ONE *11*, e0165148.

Maeshima, K., Hihara, S., and Takata, H. (2010). New Insight into the Mitotic Chromosome Structure Irregular Folding of Nucleosome Fibers Without 30-nm Chromatin Structure. Cold Spring Harb. Symp. Quant. Biol. *75*, 439–444.

Malkusch, S., Endesfelder, U., Mondry, J., Gelléri, M., Verveer, P.J., and Heilemann, M. (2012). Coordinate-based colocalization analysis of single-molecule localization microscopy data. Histochem. Cell Biol. *137*, 1–10.

Manders, E.M., Stap, J., Brakenhoff, G.J., van Driel, R., and Aten, J.A. (1992). Dynamics of three-dimensional replication patterns during the S-phase, analysed by double labelling of DNA and confocal microscopy. J. Cell Sci. *103 (Pt 3)*, 857–862.

Manders, E.M.M., Verbeek, F.J., and Aten, J.A. (1993). Measurement of co-localization of objects in dual-colour confocal images. J. Microsc. *169*, 375–382.

Manley, S., Gillette, J.M., Patterson, G.H., Shroff, H., Hess, H.F., Betzig, E., and Lippincott-Schwartz, J. (2008). High-density mapping of single-molecule trajectories with photoactivated localization microscopy. Nat. Methods *5*, 155–157.

Manneville, J.B., Etienne-Manneville, S., Skehel, P., Carter, T., Ogden, D., and Ferenczi, M. (2003). Interaction of the actin cytoskeleton with microtubules regulates secretory organelle movement near the plasma membrane in human endothelial cells. J. Cell Sci. *116*, 3927–3938.

Marjon, K.D., Termini, C.M., Karlen, K.L., Saito-Reis, C., Soria, C.E., Lidke, K.A., and Gillette, J.M. (2016). Tetraspanin CD82 regulates bone marrow homing of acute myeloid leukemia by modulating the molecular organization of N-cadherin. Oncogene *35*, 4132–4140.

McGrew, J., Diehl, B., and Fitzgerald-Hayes, M. (1986). Single base-pair mutations in centromere element III cause aberrant chromosome segregation in Saccharomyces cerevisiae. Mol. Cell. Biol. *6*, 530–538.

McKinley, K.L., and Cheeseman, I.M. (2016). The molecular basis for centromere identity and function. Nat. Rev. Mol. Cell Biol. *17*, 16–29.

Miell, M.D.D., Fuller, C.J., Guse, A., Barysz, H.M., Downes, A., Owen-Hughes, T., Rappsilber, J., Straight, A.F., and Allshire, R.C. (2013). CENP-A confers a reduction in height on octameric nucleosomes. Nat. Struct. Mol. Biol. *20*, 763–765.

Mikhaylova, M., Cloin, B.M.C., Finan, K., Berg, R. van den, Teeuw, J., Kijanka, M.M., Sokolowski, M., Katrukha, E.A., Maidorn, M., Opazo, F., et al. (2015). Resolving bundled microtubules using anti-tubulin nanobodies. Nat. Commun. *6*, ncomms8933.

Minsky, M. (1961). Microscopy apparatus. US patent 3013467 A.

Mirny, L.A. (2011). The fractal globule as a model of chromatin architecture in the cell. Chromosome Res. *19*, 37–51.

Mlodzianoski, M.J., Schreiner, J.M., Callahan, S.P., Smolková, K., Dlasková, A., Šantorová, J., Ježek, P., and Bewersdorf, J. (2011). Sample drift correction in 3D fluorescence photoactivation localization microscopy. Opt. Express *19*, 15009.

Moneron, G., and Hell, S.W. (2009). Two-photon excitation STED microscopy. Opt. Express *17*, 14567–14573.

Nahidiazar, L., Agronskaia, A.V., Broertjes, J., Broek, B. van den, and Jalink, K. (2016). Optimizing Imaging Conditions for Demanding Multi-Color Super Resolution Localization Microscopy. PLOS ONE *11*, e0158884.

Neil, M. a. A., Juškaitis, R., and Wilson, T. (1997). Method of obtaining optical sectioning by using structured light in a conventional microscope. Opt. Lett. *22*, 1905–1907.

Nicovich, P.R., Owen, D.M., and Gaus, K. (2017). Turning single-molecule localization microscopy into a quantitative bioanalytical tool. Nat. Protoc. *12*, 453–460.

Nieuwenhuizen, R.P.J., Lidke, K.A., Bates, M., Puig, D.L., Grünwald, D., Stallinga, S., and Rieger, B. (2013). Measuring image resolution in optical nanoscopy. Nat. Methods *10*, 557–562.

Nishimune, H., Badawi, Y., Mori, S., and Shigemoto, K. (2016). Dual-color STED microscopy reveals a sandwich structure of Bassoon and Piccolo in active zones of adult and aged mice. Sci. Rep. *6*, 27935.

Nishino, Y., Eltsov, M., Joti, Y., Ito, K., Takata, H., Takahashi, Y., Hihara, S., Frangakis, A.S., Imamoto, N., Ishikawa, T., et al. (2012). Human mitotic chromosomes consist

predominantly of irregularly folded nucleosome fibres without a 30-nm chromatin structure. EMBO J. *31*, 1644–1653.

Oheim, M., Loerke, D., Stuhmer, W., and Chow, R.H. (1998). The last few milliseconds in the life of a secretory granule - Docking, dynamics and fusion visualized by total internal reflection fluorescence microscopy (TIRFM). Eur. Biophys. J. Biophys. Lett. *27*, 83–98.

Oijen, A.M. van, Köhler, J., Schmidt, J., Müller, M., and Brakenhoff, G.J. (1999). Far-field fluorescence microscopy beyond the diffraction limit. JOSA A *16*, 909–915.

Okabe, A., Boots, B., Sugihara, K., and Chiu, S.N. (2000). Spatial Tessellations: Concepts and Applications of Voronoi Diagrams (Chichester New York: Wiley).

Olichon, A., and Surrey, T. (2007). Selection of Genetically Encoded Fluorescent Single Domain Antibodies Engineered for Efficient Expression in Escherichia coli. J. Biol. Chem. *282*, 36314–36320.

Olins, A.L., and Olins, D.E. (1974). Spheroid Chromatin Units (ν Bodies). Science *183*, 330–332.

Olivier, N., Keller, D., Rajan, V.S., Gönczy, P., and Manley, S. (2013). Simple buffers for 3D STORM microscopy. Biomed. Opt. Express *4*, 885–899.

Opazo, F., Levy, M., Byrom, M., Schäfer, C., Geisler, C., Groemer, T.W., Ellington, A.D., and Rizzoli, S.O. (2012). Aptamers as potential tools for super-resolution microscopy. Nat. Methods *9*, 938–939.

Orlov, I., Myasnikov, A.G., Andronov, L., Natchiar, S.K., Khatter, H., Beinsteiner, B., Ménétret, J.-F., Hazemann, I., Mohideen, K., Tazibt, K., et al. (2017). The integrative role of cryo electron microscopy in molecular and cellular structural biology. Biol. Cell *109*, 81–93.

Oudet, P., Gross-Bellard, M., and Chambon, P. (1975). Electron microscopic and biochemical evidence that chromatin structure is a repeating unit. Cell *4*, 281–300.

Ovesný, M., Křížek, P., Borkovec, J., Švindrych, Z., and Hagen, G.M. (2014). ThunderSTORM: a comprehensive ImageJ plug-in for PALM and STORM data analysis and super-resolution imaging. Bioinformatics *30*, 2389–2390.

Owen, D.M., Rentero, C., Rossy, J., Magenau, A., Williamson, D., Rodriguez, M., and Gaus, K. (2010). PALM imaging and cluster analysis of protein heterogeneity at the cell surface. J. Biophotonics *3*, 446–454.

Oza, P., Jaspersen, S.L., Miele, A., Dekker, J., and Peterson, C.L. (2009). Mechanisms that regulate localization of a DNA double-strand break to the nuclear periphery. Genes Dev. *23*, 912–927.

Padeganeh, A., Ryan, J., Boisvert, J., Ladouceur, A.-M., Dorn, J.F., and Maddox, P.S. (2013). Octameric CENP-A Nucleosomes Are Present at Human Centromeres throughout the Cell Cycle. Curr. Biol. *23*, 764–769.

Pageon, S.V., Tabarin, T., Yamamoto, Y., Ma, Y., Nicovich, P.R., Bridgeman, J.S., Cohnen, A., Benzing, C., Gao, Y., Crowther, M.D., et al. (2016). Functional role of T-cell receptor nanoclusters in signal initiation and antigen discrimination. Proc. Natl. Acad. Sci. *113*, E5454–E5463.

Palmer, D.K., O'Day, K., Trong, H.L., Charbonneau, H., and Margolis, R.L. (1991). Purification of the centromere-specific protein CENP-A and demonstration that it is a distinctive histone. Proc. Natl. Acad. Sci. U. S. A. *88*, 3734–3738.

Panchenko, T., Sorensen, T.C., Woodcock, C.L., Kan, Z., Wood, S., Resch, M.G., Luger, K., Englander, S.W., Hansen, J.C., and Black, B.E. (2011). Replacement of histone H3 with CENP-A directs global nucleosome array condensation and loosening of nucleosome superhelical termini. Proc. Natl. Acad. Sci. *108*, 16588–16593.

Patterson, G.H., and Lippincott-Schwartz, J. (2002). A Photoactivatable GFP for Selective Photolabeling of Proteins and Cells. Science *297*, 1873–1877.

Pavani, S.R.P., and Piestun, R. (2008). High-efficiency rotating point spread functions. Opt. Express *16*, 3484–3489.

Pavani, S.R.P., Thompson, M.A., Biteen, J.S., Lord, S.J., Liu, N., Twieg, R.J., Piestun, R., and Moerner, W.E. (2009). Three-dimensional, single-molecule fluorescence imaging beyond the diffraction limit by using a double-helix point spread function. Proc. Natl. Acad. Sci. *106*, 2995–2999.

Pearson, K. (1896). Mathematical Contributions to the Theory of Evolution. III. Regression, Heredity, and Panmixia. Philos. Trans. R. Soc. Lond. Math. Phys. Eng. Sci. *187*, 253–318.

Pengo, T., Holden, S.J., and Manley, S. (2015). PALMsiever: a tool to turn raw data into results for single-molecule localization microscopy. Bioinforma. Oxf. Engl. *31*, 797–798.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. J. Comput. Chem. *25*, 1605–1612.

Pleiner, T., Bates, M., Trakhanov, S., Lee, C.-T., Schliep, J.E., Chug, H., Böhning, M., Stark, H., Urlaub, H., and Görlich, D. (2015). Nanobodies: site-specific labeling for super-resolution imaging, rapid epitope-mapping and native protein complex isolation. eLife *4*, e11349.

Ploem, J.S. (1967). The use of a vertical illuminator with interchangeable dichroic mirrors for fluorescence microscopy with incident light. Z Wiss Mikrosk *68*, 129–142.

Pluta, A.F., Mackay, A.M., Ainsztein, A.M., Goldberg, I.G., and Earnshaw, W.C. (1995). The Centromere: Hub of Chromosomal Activities. Science *270*, 1591–1594.

Prabhat, P., Ram, S., Ward, E.S., and Ober, R.J. (2004). Simultaneous imaging of different focal planes in fluorescence microscopy for the study of cellular dynamics in three dimensions. IEEE Trans. Nanobioscience *3*, 237–242.

Qu, X., Wu, D., Mets, L., and Scherer, N.F. (2004). Nanometer-localized multiple single-molecule fluorescence microscopy. Proc. Natl. Acad. Sci. U. S. A. *101*, 11298–11303.

Quan, T., Zhu, H., Liu, X., Liu, Y., Ding, J., Zeng, S., and Huang, Z.-L. (2011). High-density localization of active molecules using Structured Sparse Model and Bayesian Information Criterion. Opt. Express *19*, 16963–16974.

Rahbarizadeh, F., Rasaee, M.J., Forouzandeh-Moghadam, M., and Allameh, A.-A. (2005). High expression and purification of the recombinant camelid anti-MUC1 single domain antibodies in Escherichia coli. Protein Expr. Purif. *44*, 32–38.

Rego, E.H., Shao, L., Macklin, J.J., Winoto, L., Johansson, G.A., Kamps-Hughes, N., Davidson, M.W., and Gustafsson, M.G.L. (2012). Nonlinear structured-illumination microscopy with a photoswitchable protein reveals cellular structures at 50-nm resolution. Proc. Natl. Acad. Sci. *109*, E135–E143.

Reindl, J., Girst, S., Walsh, D.W.M., Greubel, C., Schwarz, B., Siebenwirth, C., Drexler, G.A., Friedl, A.A., and Dollinger, G. (2017). Chromatin organization revealed by nanostructure of irradiation induced γH2AX, 53BP1 and Rad51 foci. Sci. Rep. *7*, 40616.

Ribeiro, S.A., Vagnarelli, P., Dong, Y., Hori, T., McEwen, B.F., Fukagawa, T., Flors, C., and Earnshaw, W.C. (2010). A super-resolution map of the vertebrate kinetochore. Proc. Natl. Acad. Sci. *107*, 10484–10489.

Ribeiro, T., Marques, A., Novák, P., Schubert, V., Vanzela, A.L.L., Macas, J., Houben, A., and Pedrosa-Harand, A. (2017). Centromeric and non-centromeric satellite DNA organisation differs in holocentric Rhynchospora species. Chromosoma *126*, 325–335.

Ricci, M.A., Manzo, C., García-Parajo, M.F., Lakadamyali, M., and Cosma, M.P. (2015). Chromatin Fibers Are Formed by Heterogeneous Groups of Nucleosomes In Vivo. Cell *160*, 1145–1158.

Ries, J., Kaplan, C., Platonova, E., Eghlidi, H., and Ewers, H. (2012). A simple, versatile method for GFP-based super-resolution microscopy via nanobodies. Nat. Methods *9*, 582–584.

Ripley, B.D. (1976). The Second-Order Analysis of Stationary Point Processes. J. Appl. Probab. *13*, 255–266.

Ripley, B.D. (1979). Tests of `Randomness' for Spatial Point Patterns. J. R. Stat. Soc. Ser. B Methodol. *41*, 368–374.

Rittweger, E., Han, K.Y., Irvine, S.E., Eggeling, C., and Hell, S.W. (2009). STED microscopy reveals crystal colour centres with nanometric resolution. Nat. Photonics *3*, 144–147.

Rose, A.S., and Hildebrand, P.W. (2015). NGL Viewer: a web application for molecular visualization. Nucleic Acids Res. *43*, W576–W579.

Rossy, J., Owen, D.M., Williamson, D.J., Yang, Z., and Gaus, K. (2013). Conformational states of the kinase Lck regulate clustering in early T cell signaling. Nat. Immunol. *14*, 82–89.

Rossy, J., Cohen, E., Gaus, K., and Owen, D.M. (2014). Method for co-cluster analysis in multichannel single-molecule localisation data. Histochem. Cell Biol. *141*, 605–612.

Ruan, H., Yu, J., Yuan, J., Li, N., and Fang, X. (2016). Nanoscale Distribution of Transforming Growth Factor Receptor on Post-Golgi Vesicle Revealed by Super-resolution Microscopy. Chem. – Asian J. *11*, 3359–3364.

Rubin-Delanchy, P., Burn, G.L., Griffié, J., Williamson, D.J., Heard, N.A., Cope, A.P., and Owen, D.M. (2015). Bayesian cluster identification in single-molecule localization microscopy data. Nat. Methods *12*, 1072–1076.

Rust, M.J., Bates, M., and Zhuang, X. (2006). Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). Nat. Methods *3*, 793–796.

Sage, D., Kirshner, H., Pengo, T., Stuurman, N., Min, J., Manley, S., and Unser, M. (2015). Quantitative evaluation of software packages for single-molecule localization microscopy. Nat. Methods *12*, 717–724.

Samejima, I., Spanos, C., Alves, F. de L., Hori, T., Perpelescu, M., Zou, J., Rappsilber, J., Fukagawa, T., and Earnshaw, W.C. (2015). Whole-proteome genetic analysis of dependencies in assembly of a vertebrate kinetochore. J. Cell Biol. *211*, 1141–1156.

Santaguida, S., and Musacchio, A. (2009). The life and miracles of kinetochores. EMBO J. *28*, 2511–2531.

Saxton, W.O., and Baumeister, W. (1982). The correlation averaging of a regularly arranged bacterial cell envelope protein. J. Microsc. *127*, 127–138.

Schaefer, L.H., Schuster, D., and Schaffer, J. (2004). Structured illumination microscopy: artefact analysis and reduction utilizing a parameter optimization approach. J. Microsc. *216*, 165–174.

Schermelleh, L., Carlton, P.M., Haase, S., Shao, L., Winoto, L., Kner, P., Burke, B., Cardoso, M.C., Agard, D.A., Gustafsson, M.G.L., et al. (2008). Subdiffraction Multicolor Imaging of the Nuclear Periphery with 3D Structured Illumination Microscopy. Science *320*, 1332–1336.

Schermelleh, L., Heintzmann, R., and Leonhardt, H. (2010). A guide to super-resolution fluorescence microscopy. J. Cell Biol. *190*, 165–175.

Schmidt, T., Schütz, G.J., Baumgartner, W., Gruber, H.J., and Schindler, H. (1996). Imaging of single molecule diffusion. Proc. Natl. Acad. Sci. *93*, 2926–2929.

Schmied, J.J., Raab, M., Forthmann, C., Pibiri, E., Wünsch, B., Dammeyer, T., and Tinnefeld, P. (2014). DNA origami–based standards for quantitative fluorescence microscopy. Nat. Protoc. *9*, 1367–1391.

Schubert, V., Ruban, A., and Houben, A. (2016). Chromatin Ring Formation at Plant Centromeres. Plant Cell Biol. 28.

Selényi, P. (1939). Wide-Angle Interferences and the Nature of the Elementary Light Sources. Phys. Rev. *56*, 477–479.

Sengupta, P., and Lippincott-Schwartz, J. (2012). Quantitative analysis of photoactivated localization microscopy (PALM) datasets using pair-correlation analysis. BioEssays *34*, 396–405.

Sengupta, P., Jovanovic-Talisman, T., Skoko, D., Renz, M., Veatch, S.L., and Lippincott-Schwartz, J. (2011). Probing protein heterogeneity in the plasma membrane using PALM and pair correlation analysis. Nat. Methods *8*, 969–975.

Sengupta, P., Jovanovic-Talisman, T., and Lippincott-Schwartz, J. (2013). Quantifying spatial organization in point-localization superresolution images using pair correlation analysis. Nat. Protoc. *8*, 345–354.

Shaner, N.C., Campbell, R.E., Steinbach, P.A., Giepmans, B.N.G., Palmer, A.E., and Tsien, R.Y. (2004). Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein. Nat. Biotechnol. *22*, 1567–1572.

Sheppard, C.J.R., and Choudhury, A. (1977). Image Formation in the Scanning Microscope. Opt. Acta Int. J. Opt. *24*, 1051–1073.

Shimomura, O., Johnson, F.H., and Saiga, Y. (1962). Extraction, Purification and Properties of Aequorin, a Bioluminescent Protein from the Luminous Hydromedusan, Aequorea. J. Cell. Comp. Physiol. *59*, 223–239.

Shiraishi, I., Takamatsu, T., Minamikawa, T., and Fujita, S. (1992). 3-D observation of actin filaments during cardiac myofibrinogenesis in chick embryo using a confocal laser scanning microscope. Anat. Embryol. (Berl.) *185*, 401–408.

Shroff, H., Galbraith, C.G., Galbraith, J.A., and Betzig, E. (2008). Live-cell photoactivated localization microscopy of nanoscale adhesion dynamics. Nat. Methods *5*, 417–423.

Shtengel, G., Galbraith, J.A., Galbraith, C.G., Lippincott-Schwartz, J., Gillette, J.M., Manley, S., Sougrat, R., Waterman, C.M., Kanchanawong, P., Davidson, M.W., et al. (2009). Interferometric fluorescent super-resolution microscopy resolves 3D cellular ultrastructure. Proc. Natl. Acad. Sci. *106*, 3125–3130.

Shuaib, M., Ouararhni, K., Dimitrov, S., and Hamiche, A. (2010). HJURP binds CENP-A via a highly conserved N-terminal domain and mediates its deposition at centromeres. Proc. Natl. Acad. Sci. *107*, 1349–1354.

Sinnen, B.L., Bowen, A.B., Forte, J.S., Hiester, B.G., Crosby, K.C., Gibson, E.S., Dell'Acqua, M.L., and Kennedy, M.J. (2017). Optogenetic Control of Synaptic Composition and Function. Neuron *93*, 646–660.e5.

Small, A., and Stahlheber, S. (2014). Fluorophore localization algorithms for super-resolution microscopy. Nat. Methods *11*, 267–279.

Smith, C.S., Joseph, N., Rieger, B., and Lidke, K.A. (2010). Fast, single-molecule localization that achieves theoretically minimum uncertainty. Nat. Methods *7*, 373–375.

Spinelli, S., Frenken, L., Bourgeois, D., Ron, L. de, Bos, W., Verrips, T., Anguille, C., Cambillau, C., and Tegoni1, M. (1996). The crystal structure of a llama heavy chain variable domain. Nat. Struct. Mol. Biol. *3*, 752–757.

Stallinga, S., and Rieger, B. (2010). Accuracy of the Gaussian Point Spread Function model in 2D localization microscopy. Opt. Express *18*, 24461–24476.

Stallinga, S., and Rieger, B. (2012). Position and orientation estimation of fixed dipole emitters using an effective Hermite point spread function model. Opt. Express *20*, 5896–5921.

Stiel, A.C., Trowitzsch, S., Weber, G., Andresen, M., Eggeling, C., Hell, S.W., Jakobs, S., and Wahl, M.C. (2007). 1.8 Å bright-state structure of the reversibly switchable fluorescent protein Dronpa guides the generation of fast switching variants. Biochem. J. *402*, 35–42.

Stiel, A.C., Andresen, M., Bock, H., Hilbert, M., Schilde, J., Schönle, A., Eggeling, C., Egner, A., Hell, S.W., and Jakobs, S. (2008). Generation of Monomeric Reversibly Switchable Red Fluorescent Proteins for Far-Field Fluorescence Nanoscopy. Biophys. J. *95*, 2989–2997.

Stokes, G.G. (1852). Ueber die Veränderung der Brechbarkeit des Lichts. Ann. Phys. *163*, 480–490.

Stoyan, D., and Stoyan, H. (1994). Fractals, random shapes, and point fields: methods of geometrical statistics (Wiley).

Sun, Y., McKenna, J.D., Murray, J.M., Ostap, E.M., and Goldman, Y.E. (2009). Parallax: high accuracy three-dimensional single molecule tracking using split images. Nano Lett. *9*, 2676–2682.

Sundberg, E.J., and Mariuzza, R.A. (2002). Molecular recognition in antibody-antigen complexes. Adv. Protein Chem. *61*, 119–160.

Szczurek, A., Klewes, L., Xing, J., Gourram, A., Birk, U., Knecht, H., Dobrucki, J.W., Mai, S., and Cremer, C. (2017). Imaging chromatin nanostructure with binding-activated localization microscopy based on DNA structure fluctuations. Nucleic Acids Res. *45*, e56.

Szymborska, A., Marco, A. de, Daigle, N., Cordes, V.C., Briggs, J.A.G., and Ellenberg, J. (2013). Nuclear Pore Scaffold Structure Analyzed by Super-Resolution Microscopy and Particle Averaging. Science *341*, 655–658.

Tachiwana, H., Kagawa, W., Shiga, T., Osakabe, A., Miya, Y., Saito, K., Hayashi-Takanaka, Y., Oda, T., Sato, M., Park, S.-Y., et al. (2011). Crystal structure of the human centromeric nucleosome containing CENP-A. Nature *476*, 232–235.

Takasaki, K.T., Ding, J.B., and Sabatini, B.L. (2013). Live-Cell Superresolution Imaging by Pulsed STED Two-Photon Excitation Microscopy. Biophys. J. *104*, 770–777.

Talon, H., Fleury, L., Bernard, J., and Orrit, M. (1992). Fluorescence excitation of single molecules. JOSA B *9*, 825–828.

Tang, G., Peng, L., Baldwin, P.R., Mann, D.S., Jiang, W., Rees, I., and Ludtke, S.J. (2007). EMAN2: an extensible image processing suite for electron microscopy. J. Struct. Biol. *157*, 38–46.

Tang, J., Akerboom, J., Vaziri, A., Looger, L.L., and Shank, C.V. (2010). Near-isotropic 3D optical nanoscopy with photon-limited chromophores. Proc. Natl. Acad. Sci. *107*, 10068–10073.

Tang, Y., Wang, X., Zhang, X., Li, J., and Dai, L. (2014). Sub-nanometer drift correction for super-resolution imaging. Opt. Lett. *39*, 5685–5688.

Testa, I., Wurm, C.A., Medda, R., Rothermel, E., von Middendorf, C., Fölling, J., Jakobs, S., Schönle, A., Hell, S.W., and Eggeling, C. (2010). Multicolor Fluorescence Nanoscopy in Fixed and Living Cells by Exciting Conventional Fluorophores with a Single Wavelength. Biophys. J. *99*, 2686–2694.

Thompson, R.E., Larson, D.R., and Webb, W.W. (2002). Precise nanometer localization analysis for individual fluorescent probes. Biophys. J. *82*, 2775–2783.

Tilke, C. (1992). Statistics for spatial data : N.A.C. Cressie (1991) New York: John Wiley & Sons, 920 pp., ISBN 0-471-84336-9, 71 [pound sign] sterling. Comput. Stat. Data Anal. *14*, 547–544.

Tiwari, D.K., Arai, Y., Yamanaka, M., Matsuda, T., Agetsuma, M., Nakano, M., Fujita, K., and Nagai, T. (2015). A fast- and positively photoswitchable fluorescent protein for ultralow-laser-power RESOLFT nanoscopy. Nat. Methods *12*, 515–518.

Tønnesen, J., Nadrigny, F., Willig, K.I., Wedlich-Söldner, R., and Nägerl, U.V. (2011). Two-Color STED Microscopy of Living Synapses Using A Single Laser-Beam Pair. Biophys. J. *101*, 2545–2552.

Vale, R.D., Funatsu, T., Pierce, D.W., Romberg, L., Harada, Y., and Yanagida, T. (1996). Direct observation of single kinesin molecules moving along microtubules. Nature *380*, 451–453.

Van Heel, M. (1987). Angular reconstitution: A posteriori assignment of projection directions for 3D reconstruction. Ultramicroscopy *21*, 111–123.

Veatch, S.L., Machta, B.B., Shelby, S.A., Chiang, E.N., Holowka, D.A., and Baird, B.A. (2012). Correlation Functions Quantify Super-Resolution Images and Estimate Apparent Clustering Due to Over-Counting. PLOS ONE *7*, e31457.

Vicidomini, G., Moneron, G., Han, K.Y., Westphal, V., Ta, H., Reuss, M., Engelhardt, J., Eggeling, C., and Hell, S.W. (2011). Sharper low-power STED nanoscopy by time gating. Nat. Methods *8*, 571–573.

Vicidomini, G., Schönle, A., Ta, H., Han, K.Y., Moneron, G., Eggeling, C., and Hell, S.W. (2013). STED nanoscopy with time-gated detection: theoretical and experimental aspects. PloS One *8*, e54421.

Voronoi, George (1908). Nouvelles applications des paramètres continus à la théorie des formes quadratiques. J. Für Reine Angew. Math. *1908*, 97–102.

Wade, A.R., and Fitzke, F.W. (1998). A fast, robust pattern recognition system for low light level image registration and its application to retinal imaging. Opt. Express *3*, 190–197.

Wang, Y., Gao, J., Guo, X., Tong, T., Shi, X., Li, L., Qi, M., Wang, Y., Cai, M., Jiang, J., et al. (2014a). Regulation of EGFR nanocluster formation by ionic protein-lipid interaction. Cell Res. *24*, 959–976.

Wang, Y., Maharana, S., Wang, M.D., and Shivashankar, G.V. (2014b). Super-resolution microscopy reveals decondensed chromatin structure at transcription sites. Sci. Rep. *4*.

Ward, E.S., Güssow, D., Griffiths, A.D., Jones, P.T., and Winter, G. (1989). Binding activities of a repertoire of single immunoglobulin variable domains secreted from Escherichia coli. Nature *341*, 544–546.

Wegel, E., Göhler, A., Lagerholm, B.C., Wainman, A., Uphoff, S., Kaufmann, R., and Dobbie, I.M. (2016). Imaging cellular structures in super-resolution with SIM, STED and Localisation Microscopy: A practical comparison. Sci. Rep. *6*, srep27290.

Westphal, V., and Hell, S.W. (2005). Nanoscale Resolution in the Focal Plane of an Optical Microscope. Phys. Rev. Lett. *94*, 143903.

White, J.G., Amos, W.B., and Fordham, M. (1987). An evaluation of confocal versus conventional imaging of biological structures by fluorescence light microscopy. J. Cell Biol. *105*, 41–48.

Williamson, D.J., Owen, D.M., Rossy, J., Magenau, A., Wehrmann, M., Gooding, J.J., and Gaus, K. (2011). Pre-existing clusters of the adaptor Lat do not participate in early T cell signaling events. Nat. Immunol. *12*, 655–662.

Wilson, T. (2011). Resolution and optical sectioning in the confocal microscope. J. Microsc. *244*, 113–121.

Woodcock, C.L.F., Safer, J.P., and Stanchfield, J.E. (1976). Structural repeating units in chromatin. Exp. Cell Res. *97*, 101–110.

Xu, K., Babcock, H.P., and Zhuang, X. (2012). Dual-objective STORM reveals three-dimensional filament organization in the actin cytoskeleton. Nat. Methods *9*, 185–188.

Yang, F., Moss, L.G., and Phillips, G.N. (1996). The molecular structure of green fluorescent protein. Nat. Biotechnol. *14*, 1246–1251.

Yildiz, A., Forkey, J.N., McKinney, S.A., Ha, T., Goldman, Y.E., and Selvin, P.R. (2003). Myosin V Walks Hand-Over-Hand: Single Fluorophore Imaging with 1.5-nm Localization. Science *300*, 2061–2065.

Zhang, W., Colmenares, S.U., and Karpen, G.H. (2012). Assembly of Drosophila Centromeric Nucleosomes Requires CID Dimerization. Mol. Cell *45*, 263–269.

Zhang, X., Zhang, M., Li, D., He, W., Peng, J., Betzig, E., and Xu, P. (2016). Highly photostable, reversibly photoswitchable fluorescent protein with high contrast ratio for live-cell superresolution microscopy. Proc. Natl. Acad. Sci. *113*, 10364–10369.

Zhao, J., Bruck, S., Cemerski, S., Zhang, L., Butler, B., Dani, A., Cooper, J.A., and Shaw, A.S. (2013). CD2AP Links Cortactin and Capping Protein at the Cell Periphery To Facilitate Formation of Lamellipodia. Mol. Cell. Biol. *33*, 38–47.

Zhou, B.-R., Jiang, J., Feng, H., Ghirlando, R., Xiao, T.S., and Bai, Y. (2015). Structural Mechanisms of Nucleosome Recognition by Linker Histones. Mol. Cell *59*, 628–638.

6. Appendix

6.1. Publication 4 “Nuclear position dictates DNA repair pathway choice”

This paper answers the question, which DNA repair mechanisms are activated upon DNA double-strand brakes at different compartments of chromatin in human cells. My part in this work was SMLM imaging of lamina-associated domains of chromatin before and after DNA damage, with subsequent data processing (corrections for drift and chromatin aberrations in SharpViSu and estimation of colocalization using implemented by me in Matlab coordinate-based method (Malkusch et al., 2012)). The obtained by SMLM data suggest that unlike in yeast (Oza et al., 2009), in human cells lamina-associated domains do not seek to be repaired by the homologous recombination mechanism, which could be possible by their migration towards nuclear pores or nuclear interior, but they remain at the lamina and are repaired by alternative end-joining. The general conclusion of the paper is that the choice of DNA repair pathway is controlled by the spatial organization of DNA in the nucleus.

Downloaded from genesdev.cshlp.org on November 10, 2017 - Published by Cold Spring Harbor Laboratory Press

Nuclear position dictates DNA repair pathway choice

Charlène Lemaître,[1,2,3,4] Anastazja Grabarz,[1,2,3,4,7] Katerina Tsouroula,[1,2,3,4,7] Leonid Andronov,[1,2,3,4] Audrey Furst,[1,2,3,4] Tibor Pankotai,[1,2,3,4] Vincent Heyer,[1,2,3,4] Mélanie Rogier,[1,2,3,4] Kathleen M. Attwood,[5,6] Pascal Kessler,[1,2,3,4] Graham Dellaire,[5,6] Bruno Klaholz,[1,2,3,4] Bernardo Reina-San-Martin,[1,2,3,4] and Evi Soutoglou[1,2,3,4]

[1]Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), 67404 Illkirch CEDEX, France; [2]U964, Institut National de la Santé et de la Recherche Médicale (INSERM), 67404 Illkirch CEDEX, France; [3]UMR7104, Centre National de Recherche Scientifique (CNRS), 67404 Illkirch CEDEX, France; [4]Université de Strasbourg (UDS), 67404 Illkirch CEDEX, France; [5]Department of Pathology, [6]Department of Biochemistry and Molecular Biology, Dalhousie University, Halifax, Nova Scotia B3H 4R2, Canada

Faithful DNA repair is essential to avoid chromosomal rearrangements and promote genome integrity. Nuclear organization has emerged as a key parameter in the formation of chromosomal translocations, yet little is known as to whether DNA repair can efficiently occur throughout the nucleus and whether it is affected by the location of the lesion. Here, we induce DNA double-strand breaks (DSBs) at different nuclear compartments and follow their fate. We demonstrate that DSBs induced at the nuclear membrane (but not at nuclear pores or nuclear interior) fail to rapidly activate the DNA damage response (DDR) and repair by homologous recombination (HR). Real-time and superresolution imaging reveal that DNA DSBs within lamina-associated domains do not migrate to more permissive environments for HR, like the nuclear pores or the nuclear interior, but instead are repaired in situ by alternative end-joining. Our results are consistent with a model in which nuclear position dictates the choice of DNA repair pathway, thus revealing a new level of regulation in DSB repair controlled by spatial organization of DNA within the nucleus.

[*Keywords*: alternative end-joining; DNA repair; nuclear lamina; nuclear organization]

Supplemental material is available for this article.

Received July 2, 2014; revised version accepted October 14, 2014.

Cells continuously experience stress and damage from exogenous sources, such as UV light or irradiation, and endogenous sources, such as oxidative by-products of cellular metabolism (Jackson and Bartek 2009). To avoid subsequent genomic instability, several pathways evolved to detect DNA damage, signal its presence, and mediate its repair (Misteli and Soutoglou 2009). The two main pathways for double-strand break (DSB) repair are homologous recombination (HR) and nonhomologous end-joining (NHEJ) (Chapman et al. 2012).

DNA repair occurs in the highly compartmentalized nucleus, and emerging evidence suggests an important role of nuclear organization in the maintenance of genome integrity (Misteli and Soutoglou 2009). Observations in yeast suggest that distinct, dedicated DNA repair centers exist as preferential sites of repair (Lisby et al. 2003). Further evidence for spatially restricted repair in yeast comes from the observation that persistent DSBs migrate from their internal nuclear positions to the nuclear periphery, where they associate with nuclear pores (Therizols et al. 2006; Nagai et al. 2008; Oza et al. 2009). In mammalian cells, multiple DSBs on several chromosomes are repaired individually and do not meet on shared repair centers or move toward the nuclear periphery (Soutoglou et al. 2007). In line with these observations, spatial proximity of DSBs in the nucleus is a key parameter that affects the frequency of formation of chromosomal translocations in mammals (Roukos et al. 2013; Roukos and Misteli 2014). Therefore, in mammals, although nuclear organization has emerged as a key parameter in the formation of chromosomal translocations (for review, see Roukos and Misteli 2014), very little is known about how nuclear compartmentalization contributes to genome stability and whether DNA repair occurs throughout the nucleus with the same robustness and accuracy.

[7]**These two authors contributed equally to this work.**
Corresponding author: evisou@igbmc.fr
Article published online ahead of print. Article and publication date are online at http://www.genesdev.org/cgi/doi/10.1101/gad.248369.114. Freely available online through the *Genes & Development* Open Access option.

© 2014 Lemaître et al. This article, published in *Genes & Development*, is available under a Creative Commons License (Attribution 4.0 International), as described at http://creativecommons.org/licenses/by/4.0.

Downloaded from genesdev.cshlp.org on November 10, 2017 - Published by Cold Spring Harbor Laboratory Press

Here, we used an inducible system to create temporally and spatially defined DSBs in chromatin within different nuclear compartments and followed their fate. We show that the presence of heterochromatin at the nuclear lamina delays DNA damage response (DDR) and impairs HR. We further used live-cell imaging and superresolution microscopy to probe the spatial dynamics of these DSBs. We show that, contrary to what was observed in yeast, DNA DSBs within lamina-associated domains (LADs) do not migrate to more permissive environments for HR, like the nuclear pores or the nuclear interior. Instead, they are repaired in situ by NHEJ or alternative end-joining (A-EJ). Our data reveal a new level of regulation in DSB repair pathway choice controlled by spatial organization of DNA in the nucleus.

Results

To investigate the impact of nuclear compartmentalization on DNA repair, we induced DSBs in chromatin associated with the inner nuclear membrane and then tested the consequences of nuclear position in DDR kinetics and DNA repair efficiency. We generated I-U2OS19 cells that contain a stably integrated I-SceI restriction site flanked by 256 repeats of the lac operator DNA sequences (lacO) (Supplemental Fig. S1A). This cell line was also engineered to express the I-SceI endonuclease under the control of a doxycycline (Dox)-inducible promoter (pTRE-tight), allowing us to temporally control the induction of a DSB at the lacO/I-SceI locus (Supplemental Fig. S1A). Stable expression of the GFP lac repressor (lacI) enables the visualization of the lacO/I-SceI locus in the nucleus. We induced specific tethering of the lacO locus at the inner nuclear membrane by the expression of an Emerin C-terminal deletion (ΔEMD), which localizes at the nuclear lamina, fused to GFP-lacI (GFP-lacI-ΔEMD) (Supplemental Fig. S1A) as described in Reddy et al. (2008).

Consistent with previous results (Reddy et al. 2008), ΔEMD is sufficient to target the GFP-lacI-ΔEMD fusion protein to the nuclear membrane and relocate the lacO/I-SceI-containing chromosome at the nuclear lamina after one mitotic cycle (Supplemental Fig. S1B,C). Indeed, in cells expressing GFP-lacI-ΔEMD, we observed 70% of colocalization of the lacO array with laminB by immuno-FISH in the absence or presence of I-SceI, whereas in cells expressing GFP-lacI, this colocalization is as low as 10% (Supplemental Fig. S1B,C).

To determine whether tethering of the lacO/I-SceI locus to the nuclear lamina has an effect on the accessibility of the I-SceI endonuclease, we performed ligation-mediated PCR (LM-PCR) in cells expressing GFP-lacI or GFP-lacI-ΔEMD. We found that the cutting efficiency is equivalent in both environments (Supplemental Fig. S1D), demonstrating that the I-SceI endonuclease is able to recognize its target sequence and cleave its substrate regardless of its nuclear localization.

DSBs activate the DDR, which allows recognition of breaks and the activation of checkpoints. Consequently, cell cycle progression is paused, which allows time for the cell to repair the lesions before dividing (Misteli and Soutoglou 2009). DDR involves a megabase-wide spreading of a phosphorylated form of the histone variant H2AX (γ-H2AX) around them (Rogakou et al. 1998; Misteli and Soutoglou 2009).

To assess the impact of repositioning the lacO/I-SceI locus at the nuclear lamina compartment on DDR efficiency, we compared the kinetics of induction of γ-H2AX at the I-SceI break in cells expressing GFP-lacI or GFP-lacI-ΔEMD by immuno-FISH. Although repositioning of the lacO/I-SceI break at the nuclear lamina did not affect the maximal percentage of γ-H2AX, cells expressing GFP-lacI showed the highest percentage of γ-H2AX colocalization with the lacO/I-sceI locus 14 h after Dox addition, whereas GFP-lacI-ΔEMD cells only achieved the same level 24 h after Dox was added (Fig. 1A, B). This observation was further confirmed by chromatin immunoprecipitation (ChIP) experiments (Fig. 1C). We also investigated the recruitment of another DDR factor, 53BP1, which has been implicated in the choice of the DSB repair pathway (Bunting et al. 2010; Panier and Boulton 2014). Similarly to γ-H2AX, the recruitment of 53BP1 was also delayed and showed a maximal accumulation at 24 h after I-SceI expression in GFP-lacI-ΔEMD cells compared with 20 h in GFP-lacI cells (Fig. 1D,E). A similar difference was observed in a lacO/I-SceI system integrated in the I-Hela111 cell line (Supplemental Fig. S2A,B), suggesting that the effect is not tissue-specific but rather is a general mechanism. Taken together, these results reveal a general delay in DDR in lesions occurring in chromatin associated with the nuclear lamina and suggest that this compartment is a repressive microenvironment for DDR.

To rule out the possibility that this defect was due to the expression of the ΔEMD in the context of the GFP-lacI-ΔEMD fusion protein, we performed an immuno-FISH experiment in the presence of IPTG. Under these conditions, the GFP-lacI-ΔEMD fusion protein is expressed but does not bind to the lacO array, and the array is not relocalized at the nuclear lamina, which was confirmed by the markedly reduced colocalization of the array and laminB (Supplemental Fig. S3A–C). As shown in Supplemental Figure S3B and quantified in Supplemental Figure S3D, there was no difference in the degree of γ-H2AX at the I-SceI break in cells expressing either GFP-lacI or GFP-lacI-ΔEMD in the presence of IPTG and 14 h after Dox where there was the maximal difference in DDR between the two compartments (Fig. 1B), confirming that the decreased phosphorylation of H2AX is a consequence of a lesion induced at the nuclear lamina.

In light of the above observations, we investigated whether the delay in DDR at the I-SceI lesion at the nuclear membrane impacts on its repair. To evaluate the effect of the I-SceI break repositioning at the inner nuclear membrane on NHEJ, we compared the degree of colocalization of Ku80 (Britton et al. 2013) with the lacO/I-SceI array by immuno-FISH and the recruitment of XRCC4 by ChIP in cells expressing GFP-lacI and GFP-lacI-ΔEMD, two main proteins of the NHEJ pathway (Lieber 2010). We observed no difference in the recruitment of KU80 in I-U2OS19 (Fig. 2A; Supplemental Fig. S4A) and I-Hela111 (Supplemental

Downloaded from genesdev.cshlp.org on November 10, 2017 - Published by Cold Spring Harbor Laboratory Press

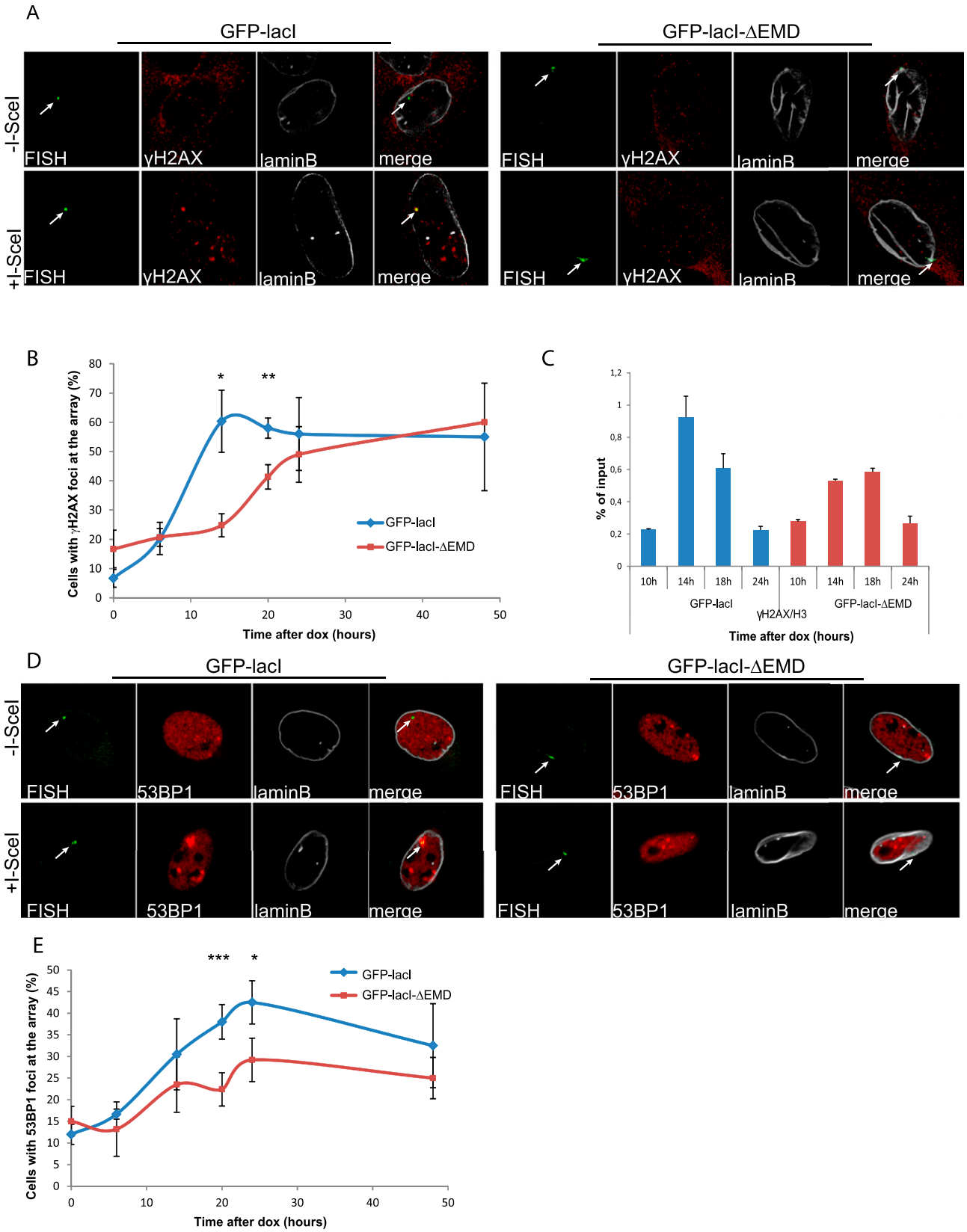


Figure 1. The DDR is delayed at the nuclear lamina. (*A*) Immuno-FISH single-Z confocal images of the lacO array (green), γ-H2AX (red), and laminB (gray) in I-U2OS19 cells expressing GFP-lacI or GFP-lacI-ΔEMD and treated or not with Dox for 14 h. (*B*) Time course of the percentage of colocalization of the lacO array with γ-H2AX. (*C*) γ-H2AX ChIP at the indicated time points after Dox addition in cells expressing GFP-lacI or GFP-lacI-ΔEMD. Values were normalized to input DNA and H3 ChIP and are representative of three independent experiments. (*D*) Immuno-FISH single-Z confocal images of the lacO array (green), 53BP1 (red), and laminB (gray) in I-U2OS19 cells expressing GFP-lacI or GFP-lacI-ΔEMD and treated or not with Dox for 20 h. (*E*) 53BP1 after Dox addition in I-U2OS19 cells expressing GFP-lacI or GFP-lacI-ΔEMD. Values represent mean ± SD of three independent experiments with $n > 50$ cells. For statistical analysis, a *t*-test was performed. (*) $P < 0.05$; (**) $P < 0.01$; (***) $P < 0.001$. In all figures, the arrow depicts the position of the lacO array.

Downloaded from genesdev.cshlp.org on November 10, 2017 - Published by Cold Spring Harbor Laboratory Press

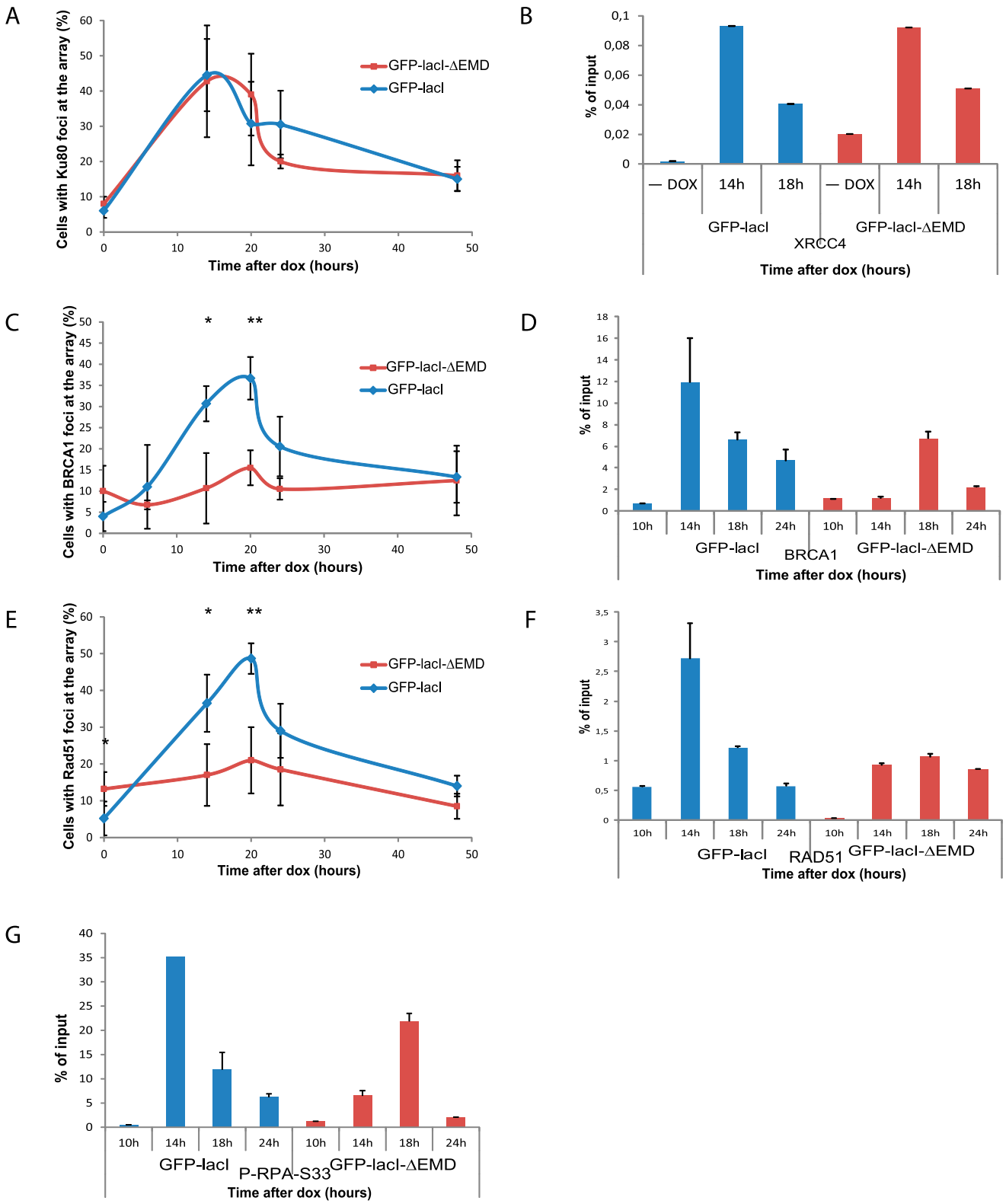


Figure 2. Recruitment of HR factors is impaired at the nuclear lamina. (*A*) Time course of the percentage of colocalization of the lacO array with Ku80 after Dox addition in I-U2OS19 cells expressing GFP-lacI or GFP-lacI-ΔEMD. Values represent mean ± SD of three independent experiments with $n > 50$ cells. ChIP for XRRC4 (*B*), BRCA1 (*D*), RAD51 (*F*), or P-RPAS33 (*G*) at the indicated times upon Dox addition in I-Hela111 cells (XRCC4) or I-U2OS19 cells expressing GFP-lacI or GFP-lacI-ΔEMD is shown. Values were normalized to input DNA and are representative of three independent experiments. The percentage of colocalization of the lacO array with BRCA1 (*C*) and Rad51 (*E*) at the indicated times after Dox addition in I-U2OS19 cells expressing GFP-lacI or GFP-lacI-ΔEMD is shown. Values represent mean ± SD of three independent experiments with $n > 50$ cells. For statistical analysis, a *t*-test was performed. (*) $P < 0.05$; (**) $P < 0.01$.

Downloaded from genesdev.cshlp.org on November 10, 2017 - Published by Cold Spring Harbor Laboratory Press

Fig. S5A–D) cells or XRCC4 at I-Hela111 (Fig. 2B) at the I-SceI break induced at the nuclear lamina compared with the nuclear interior, suggesting that NHEJ can occur efficiently in both compartments. Interestingly, the recruitment of NHEJ factors was not delayed, which is indicative of an uncoupling of DDR and repair by NHEJ.

HR is mainly active during the S phase of the cell cycle and uses the homologous sister chromatid as a template for error-free repair (San Filippo et al. 2008). Contrary to what was observed for NHEJ proteins, the recruitment of HR factors such as BRCA1, Rad51 (Fig. 2C–F; Supplemental Figs. S4B,C, S5B,C,E,F), and Rad54 (Supplemental Fig. S6A) at the broken lacO residing at the inner nuclear membrane was markedly decreased. Interestingly, the phosphorylation of RPA was delayed and less robust but not entirely abolished, suggesting a semifunctional resection pathway (Fig. 2G) and a more dramatic effect specific to late HR factors. To verify that this difference was not due to an impaired cell cycle progression in the cells expressing GFP-lacI-ΔEMD, we compared the cell cycle profiles of the two cell lines by flow cytometry and observed no difference (Supplemental Fig. S6B). Our results suggest that the nuclear lamina is a repressive environment for HR.

In the mammalian nucleus, chromatin is organized into structural domains by association with distinct nuclear compartments (Parada and Misteli 2002; Bickmore 2013). To gain insight into the cause of the DDR delay and HR repression promoted by the nuclear lamina environment, we considered the possibility that the repressive chromatin structure associated with the nuclear lamina (Padeken and Heun 2014) is involved in this phenomenon (Goodarzi and Jeggo 2012; Lemaître and Soutoglou 2014).

To test this hypothesis, we treated cells with an inhibitor of histone deacetylases, trichostatin A (TSA). This treatment resulted in an increase in histone acetylation (Supplemental Fig. S7A) and loss of heterochromatin in the nucleus, including perinuclear heterochromatin, leading to a homogenous chromatin state, as visualized by electron microscopy (Supplemental Fig. S7B–D). TSA treatment did not perturb the repositioning of the lacO/I-SceI locus at the inner nuclear membrane (Supplemental Fig. S7E,F). Interestingly, TSA treatment rescued the defect in γ-H2AX and recruitment of BRCA1 and RAD51 observed after the lacO locus relocalization at the inner nuclear membrane, pointing to an inhibitory role of chromatin compaction in DDR and HR (Fig. 3A–C; Supplemental Figs. S8, S9A,B). Our results are in line with previous studies that showed that reduced gene expression around the nuclear periphery after repositioning of the lacO array depends on the activity of histone deacetylases (Finlan et al. 2008).

To further confirm that the perinuclear heterochromatin in contact with the nuclear membrane is responsible for delayed DDR and repressed HR, we induced decondensation of the lacO/I-SceI chromatin by direct tethering of the chromatin remodeler BRG1. To this end, we expressed cherry-lacI-BRG1 in cells expressing GFP-lacI or GFP-lacI-ΔEMD (Supplemental Fig. S10A). As shown in Supplemental Figure S10B and quantified in Supplemental Figure S10C, tethering of BRG1 at the lacO array resulted in local chromatin decondensation, as visualized by an increased size of the array.

Similar to what we observed after global chromatin decondensation, local chromatin opening by BRG1 rescued the defect in γ-H2AX and the recruitment of BRCA1 and RAD51 upon lacO repositioning at the lamina (Fig. 3D–G; Supplemental Fig. S11A,B). Altogether, these results strongly suggest that the decreased recruitment of HR factors at the nuclear lamina is due to the highly compacted state of the surrounding chromatin.

To further examine whether the localization of a DSB within a nuclear compartment in relation to the state of the chromatin that surrounds the compartment can influence the DNA repair pathway choice, we assessed DSB repair at the nuclear pores, which are subcompartments of the nuclear periphery that represent a permissive environment for gene expression and other DNA-dependent nuclear transactions (Taddei et al. 2006; Ptak et al. 2014). To position the lacO/I-SceI locus at the nuclear pore compartment, we expressed GFP-lacI fused to the nucleoporin Pom121 (Supplemental Fig. S12A). We found that repositioning of the lacO array to the nuclear pores did not affect DDR, as visualized by H2AX phosphorylation and 53BP1 recruitment (Fig. 4A–C; Supplemental Fig. S12B). Furthermore, the recruitment of HR factors was similar in cells expressing GFP-lacI and GFP-lacI-Pom121 (Fig. 4D,E; Supplemental Fig. S12C,D). These observations suggest that in contrast to the nuclear lamina, nuclear pores represent a permissive microenvironment for DDR and DSB repair by HR. Therefore, although the nuclear lamina and nuclear pores are in very close proximity in the nuclear periphery, the difference in chromatin compaction associated with the two compartments regulates the choice of the repair pathway that will be prevalent in lesions occurring in each compartment.

It was previously shown that breaks inflicted at pericentric heterochromatin in *Drosophila* migrate at the periphery of the heterochromatin domain for HR repair in order to avoid recombination between repetitive sequences (Chiolo et al. 2011). Given that tethering of the lacO/I-SceI locus at the nuclear membrane using the GFP-lacI-ΔEMD might limit its potential mobility toward activating environments for DDR and repair, such as the nucleoplasm or the nuclear pores, we asked whether the lacO/I-SceI locus acquires mobility after break induction in the presence of IPTG when the lacI is not bound to the lacO array and cannot constrain its movement (Supplemental Fig. S13A). Surprisingly, we did not detect any migration of I-SceI breaks away from the compartment (Supplemental Fig. S13B).

To further investigate whether breaks occurring at the lamina migrate away from the lamina compartment toward the adjacent pores or the interior of the nucleus, we used an experimental system previously developed to visualize chromatin domains associated with laminB in single cells (Kind et al. 2013). This system uses DNA adenine methylation as a tag to visualize and track LADs

Downloaded from genesdev.cshlp.org on November 10, 2017 - Published by Cold Spring Harbor Laboratory Press

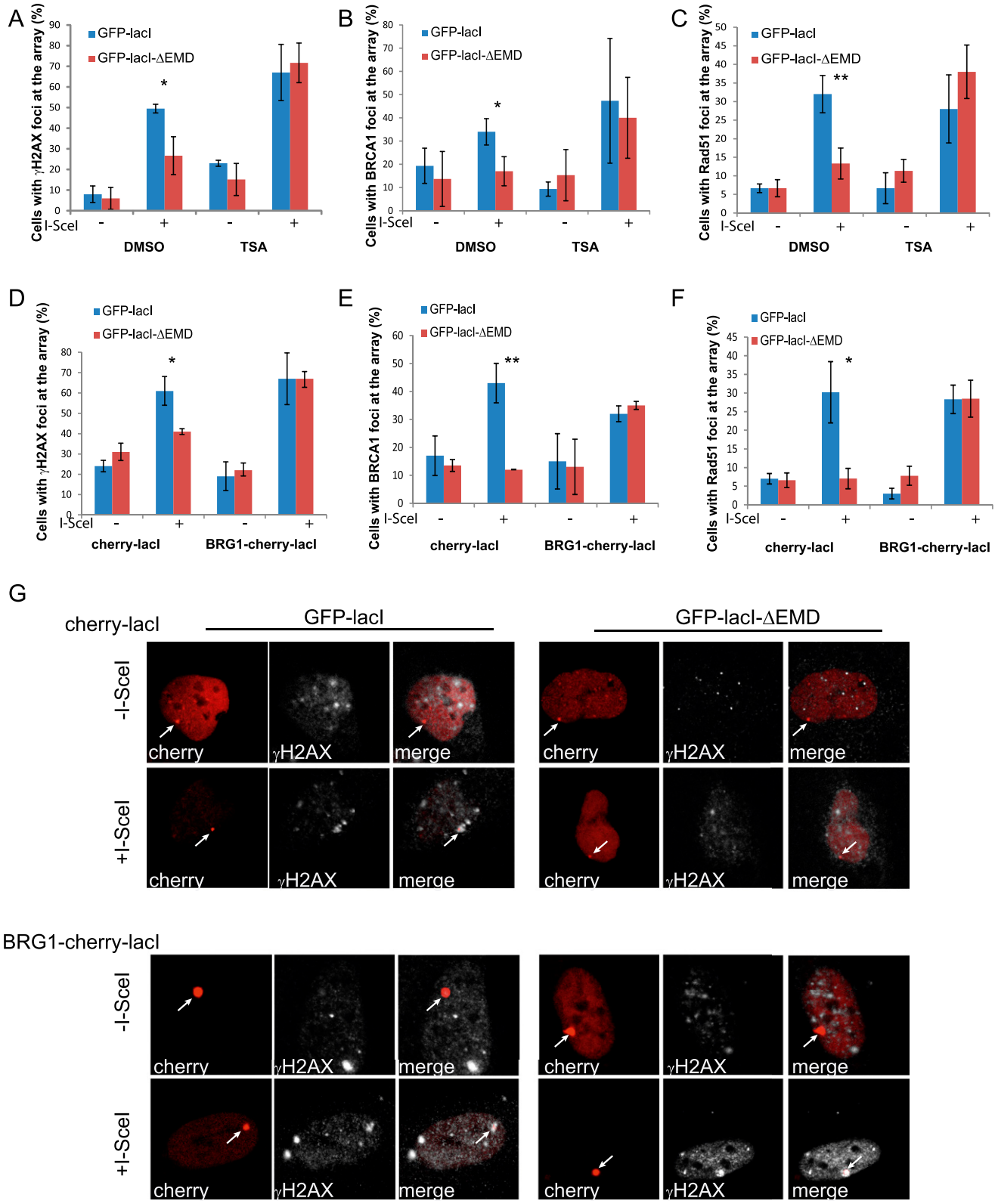


Figure 3. Chromatin decompaction restores DDR and the recruitment of HR factors at the nuclear lamina. Colocalization of the lacO array with γ-H2AX (*A*), BRCA1 (*B*), or RAD51 (*C*) in I-U2OS19 cells expressing GFP-lacI or GFP-lacI-ΔEMD and pretreated for 4 h with DMSO or TSA in the absence or presence of Dox for 14 h or 20 h is shown. The percentage of colocalization of the lacO array with γ-H2AX (*D*), BRCA1 (*E*), or RAD51 (*F*) in I-U2OS19 cells expressing GFP-lacI or GFP-lacI-ΔEMD and cherry-LacI or BRG1-cherry-lacI and treated or not with Dox for 14 h or 20 h is shown. (*G*) Immunofluorescence single-Z confocal images of γ-H2AX (gray) in I-U2OS19 cells expressing GFP-lacI or GFP-lacI-ΔEMD transfected with cherry-lacI or BRG1-cherry-lacI (red) and treated or not with Dox for 14 h. For statistical analysis, a *t*-test was performed. (*) $P < 0.05$; (**) $P < 0.01$.

Downloaded from genesdev.cshlp.org on November 10, 2017 - Published by Cold Spring Harbor Laboratory Press

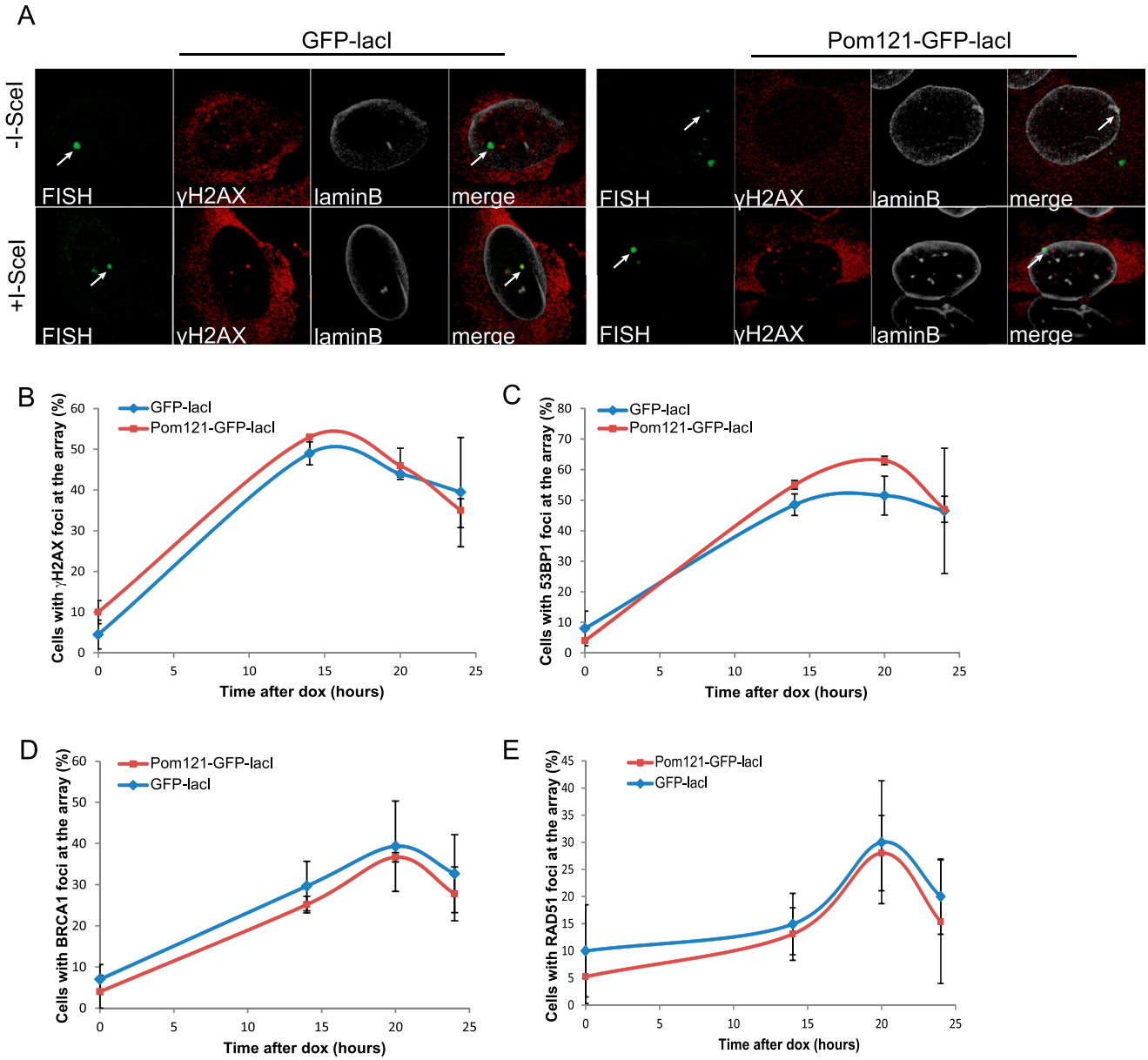


Figure 4. DDR and HR are not affected by tethering at the nuclear pores. (*A*) Immuno-FISH single-Z confocal images of the lacO array (green), γ-H2AX (red), and laminB (gray) in I-U2OS19 cells expressing GFP-lacI or Pom121-GFP-lacI and treated or not with Dox for 14 h. Time course of the percentage of colocalization of the lacO array with γ-H2AX (*B*), 53BP1 (*C*), BRCA1 (*D*), or RAD51 (*E*) in I-U2OS19 cells expressing GFP-lacI or Pom121-GFP-lacI cells after Dox addition is shown. Values represent mean ± SD of three independent experiments with $n > 50$ cells.

using a truncated version of the DpnI enzyme fused to GFP (m6a-Tracer), which recognizes methylated LADs in cells expressing LaminB-Dam (Kind et al. 2013). To probe the behavior of LADs in the presence of DNA damage, we followed the m6a-Tracer localization using live-cell imaging (Supplemental Fig. S13C) or confocal (Fig. 5A,B) or superresolution (Fig. 5C) microscopy. The infliction of DNA damage in the LADs was verified by γ-H2AX (Fig. 5A; Supplemental Fig. S13D). Interestingly, the partition of the LADs between the nuclear membrane and the nucleoplasm did not notably change before and after global DNA damage (Fig. 5A–C; Supplemental Fig. S13C), suggesting that DNA lesions do not lead to massive rearrangements of LADs within the nucleus.

In yeast, persistent DSBs migrate from their internal nuclear positions to the nuclear periphery, where they associate with nuclear pores (Therizols et al. 2006; Nagai et al. 2008; Oza et al. 2009). To more precisely assess the spatial proximity of LADs with laminB and nucleoporin of the nuclear basket TPR before and after DNA damage, we used two-color dSTORM superresolution microscopy (Folling et al. 2008). As expected, we observed juxtaposition and a certain degree of colocalization of LADs with LaminB but not with TPR

Downloaded from genesdev.cshlp.org on November 10, 2017 - Published by Cold Spring Harbor Laboratory Press

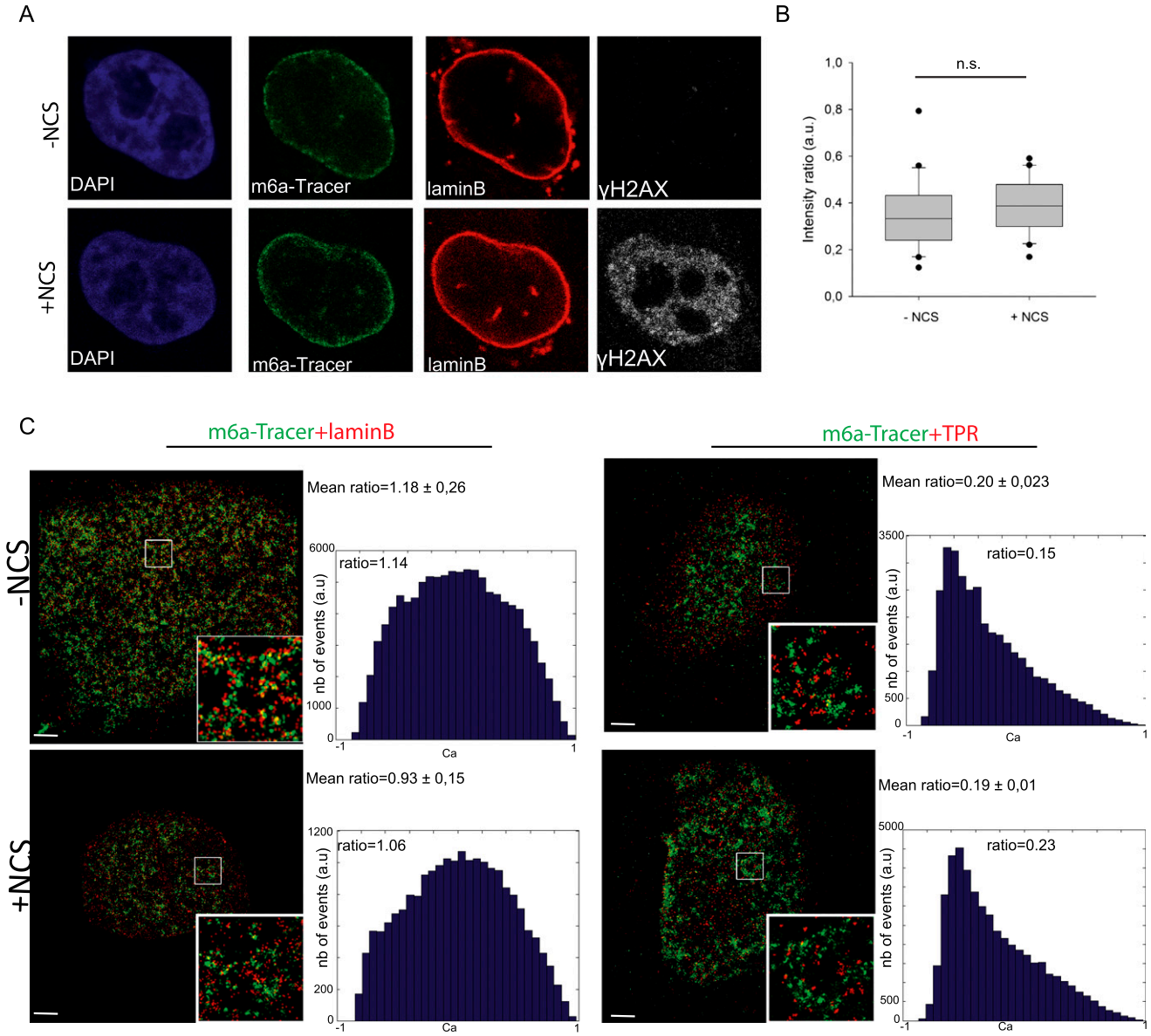


Figure 5. DSBs at the nuclear lamina are positionally stable. (*A*) Immunofluorescence of HT1080 cells expressing Dam-LaminB1 and m6A-Tracer 2 h after treatment (or not) with 50 ng/mL neocarzinostatin (NCS) for 15 min. (*B*) Box plot of GFP intensity ratios of the signal in the nucleoplasm versus the signal at the nuclear envelope in a HT1080-derived clonal cell line expressing a Dam-LaminB1 and the m6A-Tracer. The number of cells analyzed per condition was 20. For statistical analysis, χ^2 tests were performed. (n.s.) Nonsignificant. (*C*) dSTORM microscopy images of LADs (green) and laminB (*left* panel; red) or TPR (*right* panel; red) in the absence (*top* panel) or presence (*bottom* panel) of DNA damage (100 ng/mL NCS for 15 min and released for 2 h) in HT1080 cells expressing Dam-LaminB1 and m6A-Tracer. Images were taken from the bottom of the cells to allow better resolution of nuclear pores. Corresponding colocalization and the ratio of positive over negative colocalization events are displayed at the *right*. The mean ratios for all nuclei analyzed ($n \geq 8$) are displayed *above*.

(Fig. 5A). Interestingly, DNA damage did not induce changes in the proximity of LADs toward both compartments, which further pointed to the positional stability of LADs upon DNA damage (Fig. 5A). Taken together, these results suggest that contrary to what has been shown in yeast, breaks occurring on chromosomes that associate with the nuclear membrane do not travel and seek an environment permissive to HR repair, such as the nuclear pores.

To further investigate the contribution of NHEJ and HR in repairing the I-SceI breaks at the lamina or the nuclear interior, we assessed the degree of persistent breaks in GFP-lacI or GFP-lacI-ΔEMD cells depleted of XRCC4 and RAD51 (knockdown efficiencies verified in Supplemental Fig. S14A). Interestingly, in control cells, breaks were efficiently repaired in both nuclear compartments, which was exemplified by the decrease in γ-H2AX signal at the lacO array 24 h after break

Downloaded from genesdev.cshlp.org on November 10, 2017 - Published by Cold Spring Harbor Laboratory Press

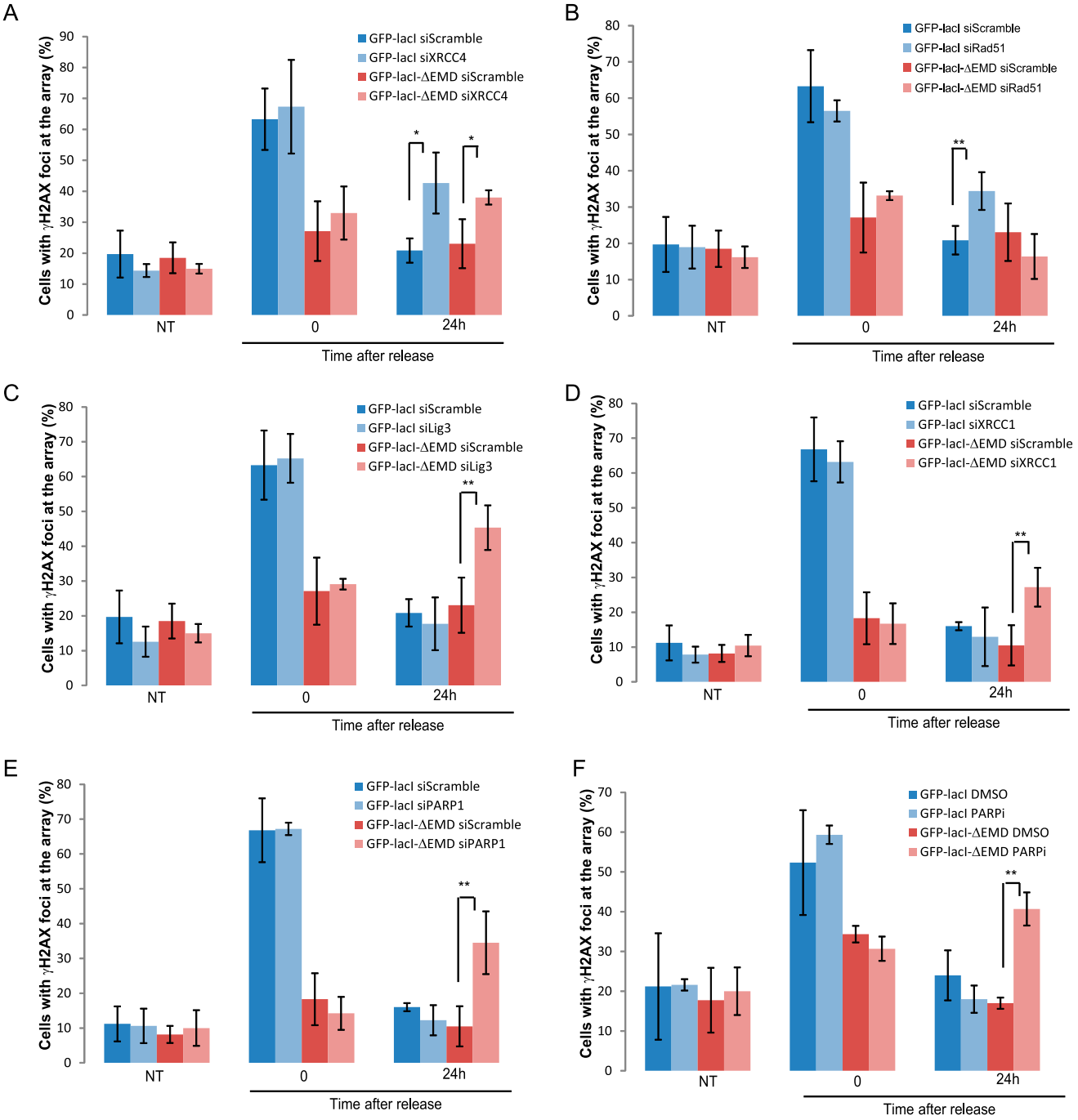


Figure 6. DSBs at the nuclear lamina are repaired by NHEJ or A-EJ. The percentage of colocalization of the lacO array with γ-H2AX in untreated cells (NT) or after 14 h of Dox (time point 0) and subsequent release for 24 h in I-U2OS19 cells expressing GFP-lacI or GFP-lacI-ΔEMD and transfected with XRCC4 (*A*), RAD51 (*B*), ligase 3 (*C*), XRCC1 (*D*), or PARP1-specific siRNAs (*E*) is shown. (*F*) The percentage of colocalization of the lacO array with γ-H2AX upon Dox treatment or release in I-U2OS19 cells expressing GFP-lacI or GFP-lacI-ΔEMD and treated with DMSO or a PARP inhibitor (PARPi, during the entire course of the experiment) is shown. Values represent mean ± SD of three independent experiments with $n > 50$ cells. For statistical analysis, a *t*-test was performed. (*) $P < 0.05$; (**) $P < 0.01$.

induction by a short pulse of Dox (Fig. 6A–E). Although depletion of XRCC4 led to persistent damage in both compartments (Fig. 6A), depletion of RAD51 did not affect the repair of breaks at the lamina (Fig. 6B). These results suggest that lesions at LADs do not depend on HR for their repair.

To test whether repositioning of the lacO/I-SceI break at the nuclear membrane affects the kinetics of repair, we performed LM-PCR in GFP-lacI and GFP-lacI-ΔEMD cells after a short pulse of Dox followed by release for 36 h. We found that breaks at both nuclear locations were efficiently repaired based on the marked decrease in PCR

Downloaded from genesdev.cshlp.org on November 10, 2017 - Published by Cold Spring Harbor Laboratory Press

signal (Supplemental Fig. S14B). These results strongly suggest that efficient DNA repair takes place at the lamina-associated I-SceI breaks even in the absence of functional HR.

Since resection is not abolished at lacO/I-SceI breaks when associated with the nuclear lamina, we sought to determine the fate of the lesions whereby resection has occurred but complete DNA repair by HR cannot occur. To answer this question, we assessed the contribution of the A-EJ pathway in the repair of breaks at the periphery. To this end, we quantified persistent γ-H2AX at the lacO/I-SceI locus 24 h after break induction in GFP-lacI and GFP-lacI-ΔEMD cells where ligase 3, XRCC1, or PARP1 had been depleted (knockdown efficiencies verified in Supplemental Fig. S14A,C) or PARP was inhibited. Interestingly, inhibition of the A-EJ pathway resulted in a repair delay for only breaks that were associated with the nuclear membrane (Fig. 6C–F; Supplemental Fig. S14D). These findings indicate that NHEJ and A-EJ, but not HR, are the most prevalent pathways of DNA repair for lesions occurring at nuclear membrane-associated chromatin and reveal for the first time that A-EJ takes place as a main pathway and not as a backup pathway activated solely in instances where there is a DNA repair factor deficiency (Frit et al. 2014).

Taken together, we showed that breaks occurring in chromatin that surrounds the nuclear membrane do not migrate to other regions of the nucleus, not even to other domains within the nuclear periphery, but rather are repaired within the lamina, where the break occurred by NHEJ and A-EJ.

Discussion

To preserve genomic integrity, different DNA repair pathways have evolved, and multiple layers of regulation like the cell cycle, specific proteins, or chromatin structure exist to ensure the tight balance between these pathways (Kass and Jasin 2010). Here, we propose another layer of regulation of DNA repair pathway choice imposed by nuclear compartmentalization. We show that the nuclear lamina restricts HR and allows NHEJ and A-EJ. These observations are in agreement with data in yeast showing that distinct nuclear compartments of the nuclear periphery like the nuclear pore or the inner nuclear membrane favor different repair outcomes (Nagai et al. 2008; Khadaroo et al. 2009; Oza et al. 2009; Horigome et al. 2014). Similar to what we observed, it was shown that binding of DSBs to Nup84 in yeast facilitates recombination through SUMO protease Ulp1 and the SUMO-dependent ubiquitin ligase Slx5/Slx8 (Nagai et al. 2008) using BIR and microhomology-mediated recombination. On the contrary, binding to the inner nuclear membrane protein Mps3 has two different outcomes: In the case of telomere tethering, it inhibits recombination by sequestering the DSBs from nonspecific interactions with chromatin (Oza et al. 2009; Schober et al. 2009), while in the case of persistent DSBs, it triggers repair by the classical HR pathway (Horigome et al. 2014).

We also found that the chromatin structure at the inner nuclear lamina is mainly responsible for inhibiting HR. This is in keeping with recent studies, which found that HR is activated at DSBs located within actively transcribed genes that reside in euchromatin (Aymard et al. 2014; Pfister et al. 2014). Given that the lacO locus is promoterless and not transcribed, our results indicate that HR is not regulated solely by the transcriptional status. Instead, the exact nature of the chromatin environment and chromatin accessibility appear to be major determinants of HR regulation (Jha and Strahl 2014; Pai et al. 2014). Indeed, other studies have shown that HR is a main pathway in repairing breaks within heterochromatin (Beucher et al. 2009; Geuting et al. 2013; Kakarougkas et al. 2013). However, our data point to the fact that not all heterochromatin domains within the nucleus behave in the same manner and that the specific type of heterochromatin at the nuclear lamina has distinct functions.

In most of the above studies, chromatin structure and histone modifications affect the very first step of the HR pathway that is DNA end resection. Aymard et al. (2014) show that H3K36me3 is essential for the recruitment of CtIP through LEDGF. On the other hand, H3K36me3 in yeast induces chromatin compaction and inhibits resection, as visualized by increased RPA foci when the methyltransferase responsible for this modification is absent (Pai et al. 2014). Here we observed that phosphorylation of RPA at S33 is delayed and not mounted properly at lesions occurring in chromatin associated with the inner nuclear membrane. We also show that BRCA1 recruitment is dramatically affected. Since BRCA1 is acting with CtIP to activate long-term resection (Chen et al. 2008), it is possible that DNA ends are not appropriately resected to create a proper template for recombination, and the short resection channels lesions to A-EJ as was proposed earlier (Zhang and Jasin 2011; Deng et al. 2014). The fact that resection at the lamina is not as dramatically affected as late steps of HR might also suggest that nuclear position dictates the DNA repair pathway choice by regulating only the recruitment of late HR proteins to DSBs.

The use of A-EJ, which is considered a highly mutagenic pathway, instead of the error-free HR pathway might seem dangerous for the maintenance of genomic stability. However, LADs are relatively gene-poor, have a repressive chromatin signature, and are demarcated by repetitive and AT-rich sequences (Meuleman et al. 2013). The inhibition of HR may represent a means to avoid genomic instability provoked by recombination between repetitive sequences, which is a mechanism that has been proposed for the repair of DSBs that form in heterochromatic regions in *Drosophila* (Chiolo et al. 2011). Moreover, activation of A-EJ that is an error-prone pathway might have less impact given that most of the genes that reside in LADs are not transcribed (Meuleman et al. 2013).

In *Drosophila*, breaks induced in the heterochromatic domain rapidly relocate outside of the domain, where HR is completed (Chiolo et al. 2011). A similar DSB relocation

Downloaded from genesdev.cshlp.org on November 10, 2017 - Published by Cold Spring Harbor Laboratory Press

was observed in mouse cells upon break induction by linear ion tracks in chromocenters (Jakob et al. 2011). On the contrary, we show that breaks occurring in chromatin associated with the inner nuclear lamina are positionally stable, suggesting that different heterochromatic compartments use different strategies to avoid recombination. One of the possible hypotheses to explain such a difference is a different chromatin composition or a difference in the regulation of chromatin mobility. Indeed, in yeast, DSBs were shown to have increased mobility (Dion and Gasser 2013). This mobility is facilitated by chromatin decompaction via chromatin remodelers (Neumann et al. 2012) and HR factors (Dion et al. 2012) and in turn allows the homology search step of HR (Mine-Hattab and Rothstein 2012). In mammalian cells, however, DSB mobility is limited and actively restricted by the NHEJ complex Ku70/Ku80 (Soutoglou et al. 2007; Roukos et al. 2013). In *Drosophila* cells, the relocation of DSBs outside of the heterochromatic domain is accompanied by decondensation of the domain (Chiolo et al. 2011), suggesting a mechanism similar to the one responsible for DSB mobility in yeast. At the nuclear lamina, however, this mechanism does not seem to be active, suggesting that an additional mechanism could repress DSB movement at the nuclear lamina. This hypothesis is in accordance with the observation that chromatin mobility is decreased for genomic loci associated with the nuclear lamina or the nucleoli (Chubb et al. 2002). Furthermore, laminA has recently been identified as a factor inhibiting DSB movement in mammalian cells (Mahen et al. 2013), further pointing to an active inhibition of DSB mobility at the nuclear lamina.

Another difference between our results and the results obtained in the heterochromatic compartment of *Drosophila* cells is the activation of DDR. In *Drosophila* cells, the activation of DDR was faster in heterochromatin compared with euchromatin (Chiolo et al. 2011). On the contrary, our results show a slower DDR activation at the nuclear lamina compared with the nuclear interior. Given the implication of the early steps of DDR in the initiation of resection by the ATM and MRN complexes, and the fact that resection facilitates DSB movement in yeast, one can hypothesize that the delayed DDR at the nuclear lamina inhibits DSB mobility.

Overall, our findings indicate that spatial positioning of a DSB is a new parameter to consider in the study of DSB repair, which has significant implications for our understanding of how the organization of repair in the highly compartmentalized nucleus contributes to maintaining genome stability and avoiding tumorigenesis.

Materials and methods

Cell lines, infections, transfections

I-U2OS19 GFP-lacI and GFP-lacI-ΔEMD cells were generated by infecting the U2OS19ptight13 cell line (Lemaître et al. 2012) with GFP-lacI (Soutoglou and Misteli 2008) and GFP-lacI-ΔEMD (Reddy et al. 2008) plasmids and after FACS sorting. Briefly, BOSC cells were transfected using FuGENE6 (Promega) according to the manufacturer's protocol with GFP-lacI or GFP-lacI-ΔEMD constructs and an amphotropic vector. Cell supernatants were harvested 48 h later and transferred to U2OS19ptight13 cells. Twenty-four hours after infection, cells were FACS-sorted for GFP-positive signal and cultured in the presence of 800 μg/mL G418 and 2 mM IPTG (inhibitor of the lacI/lacO interaction). Cells were plated in the absence of IPTG for 24 h prior to starting an experiment. To induce I-SceI expression, Dox was added to the cells at a concentration of 1 μg/mL. In Supplemental Figure S3, 2 mM IPTG was maintained during the whole experiment, and in Supplemental Figure S7, A and B, cells were plated in the absence of IPTG for 24 h and treated with Dox for 12 h. IPTG was then added for 2 h, while Dox was maintained until the end of the experiment.

Hela111 cells were obtained by transfection of lacO-I-SceI-hygro plasmid and subsequent clonal selection using 300 μg/mL hygromycin. I-HeLa111 cells were generated by transfection of Hela111 cells with pWHE320-HA-I-SceI and pWHE146-Tet activator plasmids and selection using 1 mg/mL G418. I-Hela111 GFP-lacI or GFP-lacI-ΔEMD cells were generated by infection of I-Hela111 cells with GFP-lacI and GFP-lacI-ΔEMD plasmids and FACS sorting for GFP-positive cells.

I-U2OS19 Pom121-GFP-lacI cells were obtained after infection of I-U2OS19 cells with Pom121-GFP-lacI and selection of GFP-positive cells using FACs sorting.

I-U2OS19 GFP-lacI and GFP-lacI-ΔEMD were transfected with cherry-lacI or BRG1-cherry-lacI by using FuGENE6 reagent according to the manufacturer's protocol. The cells were first plated in the absence of IPTG for 24 h and then transfected and treated with Dox 4 h after transfection.

I-U2OS19 GFP-lacI and GFP-lacI-ΔEMD cells were transfected with siRNA scramble (OnTarget Plus nontargeting pool siRNA; Dharmacon, D-001810-10-20), XRCC4 (Dharmacon, M-004494-02), Rad51 (Dharmacon, L-003530-00) or Lig3 (Dharmacon, L-009227-00) using oligofectamine reagent (Invitrogen) according to the manufacturer's protocol. Knockdown efficiency was analysed by Western blot or RT-qPCR. RNA was extracted using the RNeasy minikit (Qiagen) according to the manufacturer's protocol. RT-qPCRs were then processed as in (Pankotai et al. 2012). Proteins were extracted in RIPA buffer and analyzed by Western blot.

PARP inhibitor treatment

I-U2OS19 GFP-lacI and GFP-lacI-ΔEMD were plated in the absence of IPTG for 24 h and treated with PARPi (ABT-888, sc-202901A) at a 10 μM concentration or by DMSO.

TSA treatment

Cells were plated in the absence of IPTG for 24 h and subsequently treated with TSA at 0.5 μM or DMSO for control for 4 h. Dox was added after 4 h of treatment for the indicated time, while DMSO or TSA was maintained during the whole experiment.

Neocarzinostatin (NCS) treatment

Cells were plated in the presence of Shield for 20 h, treated for 15 min with 100 ng/mL NCS (N9162-100UG, Sigma), and fixed 2 h after treatment.

Cell cycle analysis

Cells were fixed in 70% EtOH overnight at −20°C and stained with 25 μg/mL propidium iodide. The acquisition was performed on a FACSCalibur. Results were analysed using FlowJo software.

Downloaded from genesdev.cshlp.org on November 10, 2017 - Published by Cold Spring Harbor Laboratory Press

LM-PCR

Cells were plated in the absence of IPTG for 24 h and subsequently treated with Dox for 14 h. DNA was then extracted with the DNeasy blood and tissue kit (Qiagen). Assymetric adaptator (S21, Phos-GCATCACTACGATGTAGGATG; and Lup, CATCCTACATCGTAGTGATGCTTAT) was annealed in TE for 5 min at 95°C and then allowed to reach room temperature slowly. One-hundred picomoles of assymetric adaptator was added to 1 μg of DNA extracted from cells. Ligation was performed using T4 DNA ligase overnight at 16°C. PCR was performed using Pfu enzyme (Agilent) with an annealing temperature of 58°C. The PCR primers used were LM-I-SceI (CAT CCTACATCGTAGTGATGC) and lacR (TTAATTAATCAAAC CTTCCTCT). The PCR product was then run on a 2% agarose gel.

Immunofluorescence, immuno-FISH, and microscopy

Cells were cultured on coverslips and fixed in 4% paraformaldehyde for 10 min, permeabilized in 0.5% Triton for 10 min, blocked in 1% BSA for 30 min, and incubated with primary antibody for 1 h (see the antibodies table in the Supplemental Material) and secondary antibodies for 45 min. Coverslips were incubated with DAPI and mounted on slides in Prolong Gold (Molecular Probes).

For Rad51 and Ku80 immunofluorescence or immuno-FISH, cells were pre-extracted in CSK buffer (10 mM Hepes at pH 7, 100 mM NaCl, 300 mM sucrose, 3 mM $MgCl_2$, 0.7% Triton X-100) containing 0.3 mg/mL RNase A prior to fixation (Britton et al. 2013).

For immuno-FISH, the same protocol was used, but after incubation with secondary antibodies, they were submitted to post-fixation in 4% formaldehyde for 20 min. Cells were washed for 5 min in 2× SSC and 45 min in 2× SSC with a increasing temperature from room temperature to 72°C. After one wash in 70% ethanol and two washes in absolute ethanol, coverslips were dried for 5 min at room temperature. They were subsequently incubated with 0.1 N NaOH for 10 min and washed in 2× SSC for 5 min. Coverslips were washed again in 70% ethanol and twice with absolute ethanol. After drying, cells were hybridized with DNA probe (see immuno-FISH probe preparation below) for 30 sec at 85°C and incubated overnight at 37°C.

The immuno-FISH probe was prepared by nick translation from the lacO-I-SceI plasmid that was used to create the I-Hela111 cell line. DNA probe (0.3 μg) was mixed with 9 μg of ssDNA and 3 μg of CotI human DNA (Roche) and precipitated with 2.5× vol of ethanol and 1/10 vol of 2.5 M sodium acetate for 30 min at −80°C. After 20 min of centrifugation, the supernatant was discarded, and the pellet was washed with 70% ethanol and centrifuged again for 5 min. The supernatant was discarded, and the pellet was dried. The pellet was resuspended in 20 μL of hybridization solution (50% formamide, 4× SSC, 10% dextran sulfate) per coverslip by vortexing for 1 h. The probe was denaturated for 5 min at 90°C and preannealed for at least 15 min at 37°C before hybridization with cells.

The day after hybridization, immuno-FISH was revealed. Coverslips were washed twice for 20 min at 42°C in 2× SSC and then incubated with secondary antibody and fluorescein anti-biotin (Vector Laboratories, SP-3040) at 1:100 dilution for 45 min. Coverslips were washed, incubated with DAPI, and mounted in Prolong Gold reagent (Molecular Probes).

Slides were observed, and colocalization counting was done in epifluorescence microscopy. Pictures were taken with confocal microscopy. For experiments with Pom121-GFP-lacI constructs, cells were always costained with laminB to evaluate relocalization of the lacO array at the nuclear pores. For experiments with BRG1-cherry-lacI or cherry-lacI transfections, colocalization was counted using confocal microscopy.

Time-lapse microscopy

Three-dimensional stacks were captured every 10 min for a total of 320 min upon NCS addition using the Leica DM6000 microscope with Leica CSU22 spinning disc and Andor Ixon 897 camera. Twenty different cells were imaged for each condition (±NCS).

Acknowledgements

We thank Harinder Singh (Genentech and University of California at San Francisco) for the ΔEMD-GFP-lacI plasmid, Naoko Imamoto for the POM121 plasmid (RIKEN Advanced Science Institute), Tom Misteli (National Cancer Institute, National Institutes of Health) for the cherry-lacI-BRG1 plasmid, Valérie Schreiber (École Supérieure de Biotechnologie de Strasbourg) for the XRCC1 and PARP1 antibodies, and Jop Kind (Hubrecht Institute) for the Dam-laminB/m6a-Tracer cell line. We thank Jiri and Claudia Lukas for critical reading of the manuscript, and the Soutoglou laboratory for helpful discussions. C.L. was supported by the Région Alsace and Fondation Association pour la Recherche sur le Cancer (ARC), K.T. received support from Marie Curie network address, and T.P. received support from Fondation pour la Recherche Médicale (FRM). Research in E.S.'s laboratory is supported by Agence Nationale de la Recherche (ANR), the Fondation Schlumberger pour l'Education et la Recherche (FSER), Fondation ARC (ARC fix), and La Ligue Contre le Cancer (Région). G.D.'s laboratory's contributions to this study were funded by a Discovery Grant (RGPIN 386049) from the Natural Sciences and Engineering Research Council (NSERC). K.M.A. is a recipient of a NSERC PGSD3 studentship and is a Killam Foundation Scholar. B.K.'s laboratory's contributions to this study were supported by the French Infrastructure for Integrated Structural Biology (FRISBI) (ANR-10-INSB-05-01) and Instruct as part of the European Strategy Forum on Research Infrastructures (ESFRI). B.R.-S.-M.'s laboratory's contributions to this study were supported by the FSER Fondation ARC and La Ligue Contre le Cancer.

References

Aymard F, Bugler B, Schmidt CK, Guillou E, Caron P, Briois S, Iacovoni JS, Daburon V, Miller KM, Jackson SP, et al. 2014. Transcriptionally active chromatin recruits homologous recombination at DNA double-strand breaks. *Nat Struct Mol Biol* **21:** 366–374.

Beucher A, Birraux J, Tchouandong L, Barton O, Shibata A, Conrad S, Goodarzi AA, Krempler A, Jeggo PA, Löbrich M. 2009. ATM and Artemis promote homologous recombination of radiation-induced DNA double-strand breaks in G2. *Embo J* **28:** 3413–3427.

Bickmore WA. 2013. The spatial organization of the human genome. *Annu Rev Genomics Hum Genet* **14:** 67–84.

Britton S, Coates J, Jackson SP. 2013. A new method for high-resolution imaging of Ku foci to decipher mechanisms of DNA double-strand break repair. *J Cell Biol* **202:** 579–595.

Bunting SF, Callén E, Wong N, Chen H-T, Polato F, Gunn A, Bothmer A, Feldhahn N, Fernandez-Capetillo O, Cao L, et al. 2010. 53BP1 inhibits homologous recombination in Brca1-deficient cells by blocking resection of DNA breaks. *Cell* **141:** 243–254.

Downloaded from genesdev.cshlp.org on November 10, 2017 - Published by Cold Spring Harbor Laboratory Press

Chapman JR, Taylor Martin RG, Boulton Simon J. 2012. Playing the end game: DNA double-strand break repair pathway choice. *Mol Cell* **47:** 497–510.

Chen L, Nievera CJ, Lee AY-L, Wu X. 2008. Cell cycle-dependent complex formation of BRCA1·CtIP·MRN is important for DNA double-strand break repair. *J Biol Chem* **283:** 7713–7720.

Chiolo I, Minoda A, Colmenares Serafin U, Polyzos A, Costes Sylvain V, Karpen Gary H. 2011. Double-strand breaks in heterochromatin move outside of a dynamic HP1a domain to complete recombinational repair. *Cell* **144:** 732–744.

Chubb JR, Boyle S, Perry P, Bickmore WA. 2002. Chromatin motion is constrained by association with nuclear compartments in human cells. *Curr Biol* **12:** 439–445.

Deng SK, Gibb B, de Almeida MJ, Greene EC, Symington LS. 2014. RPA antagonizes microhomology-mediated repair of DNA double-strand breaks. *Nat Struct Mol Biol* **21:** 405–412.

Dion V, Gasser SM. 2013. Chromatin movement in the maintenance of genome stability. *Cell* **152:** 1355–1364.

Dion V, Kalck V, Horigome C, Towbin BD, Gasser SM. 2012. Increased mobility of double-strand breaks requires Mec1, Rad9 and the homologous recombination machinery. *Nat Cell Biol* **14:** 502–509.

Finlan LE, Sproul D, Thomson I, Boyle S, Kerr E, Perry P, Ylstra B, Chubb JR, Bickmore WA. 2008. Recruitment to the nuclear periphery can alter expression of genes in human cells. *PLoS Genet* **4:** e1000039.

Folling J, Bossi M, Bock H, Medda R, Wurm CA, Hein B, Jakobs S, Eggeling C, Hell SW. 2008. Fluorescence nanoscopy by ground-state depletion and single-molecule return. *Nat Methods* **5:** 943–945.

Frit P, Barboule N, Yuan Y, Gomez D, Calsou P. 2014. Alternative end-joining pathway(s): Bricolage at DNA breaks. *DNA Repair* **17:** 81–97.

Geuting V, Reul C, Löbrich M. 2013. ATM release at resected double-strand breaks provides heterochromatin reconstitution to facilitate homologous recombination. *PLoS Genet* **9:** e1003667.

Goodarzi AA, Jeggo PA. 2012. The heterochromatic barrier to DNA double strand break repair: how to get the entry visa. *Int J Mol Sci* **13:** 11844–11860.

Horigome C, Oma Y, Konishi T, Schmid R, Marcomini I, Hauer MH, Dion V, Harata M, Gasser SM. 2014. SWR1 and INO80 chromatin remodelers contribute to DNA double-strand break perinuclear anchorage site choice. *Mol Cell* **55:** 626–639.

Jackson SP, Bartek J. 2009. The DNA-damage response in human biology and disease. *Nature* **461:** 1071–1078.

Jakob B, Splinter J, Conrad S, Voss KO, Zink D, Durante M, Löbrich M, Taucher-Scholz G. 2011. DNA double-strand breaks in heterochromatin elicit fast repair protein recruitment, histone H2AX phosphorylation and relocation to euchromatin. *Nucleic Acids Res* **39:** 6489–6499.

Jha DK, Strahl BD. 2014. An RNA polymerase II-coupled function for histone H3K36 methylation in checkpoint activation and DSB repair. *Nat Commun* doi: 1038/ncomms4965.

Kakarougkas A, Ismail A, Klement K, Goodarzi AA, Conrad S, Freire R, Shibata A, Lobrich M, Jeggo PA. 2013. Opposing roles for 53BP1 during homologous recombination. *Nucleic Acids Res* **41:** 9719–9731.

Kass EM, Jasin M. 2010. Collaboration and competition between DNA double-strand break repair pathways. *FEBS Lett* **584:** 3703–3708.

Khadaroo B, Teixeira MT, Luciano P, Eckert-Boulet N, Germann SM, Simon MN, Gallina I, Abdallah P, Gilson E, Geli V, et al. 2009. The DNA damage response at eroded telomeres and tethering to the nuclear pore complex. *Nat Cell Biol* **11:** 980–987.

Kind J, Pagie L, Ortabozkoyun H, Boyle S, de Vries SS, Janssen H, Amendola M, Nolen LD, Bickmore WA, van Steensel B. 2013. Single-cell dynamics of genome–nuclear lamina interactions. *Cell* **153:** 178–192.

Lemaître C, Soutoglou E. 2014. Double strand break (DSB) repair in heterochromatin and heterochromatin proteins in DSB repair. *DNA Repair* **19:** 163–168.

Lemaître C, Fischer B, Kalousi A, Hoffbeck AS, Guirouilh-Barbat J, Shahar OD, Genet D, Goldberg M, Betrand P, Lopez B, et al. 2012. The nucleoporin 153, a novel factor in double-strand break repair and DNA damage response. *Oncogene* **31:** 4803–4809.

Lieber MR. 2010. The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annu Rev Biochem* **79:** 181–211.

Lisby M, Mortensen UH, Rothstein R. 2003. Colocalization of multiple DNA double-strand breaks at a single Rad52 repair centre. *Nat Cell Biol* **5:** 572–577.

Mahen R, Hattori H, Lee M, Sharma P, Jeyasekharan AD, Venkitaraman AR. 2013. A-type lamins maintain the positional stability of DNA damage repair foci in mammalian nuclei. *PLoS ONE* **8:** e61893.

Meuleman W, Peric-Hupkes D, Kind J, Beaudry JB, Pagie L, Kellis M, Reinders M, Wessels L, van Steensel B. 2013. Constitutive nuclear lamina–genome interactions are highly conserved and associated with A/T-rich sequence. *Genome Res* **23:** 270–280.

Mine-Hattab J, Rothstein R. 2012. Increased chromosome mobility facilitates homology search during recombination. *Nat Cell Biol* **14:** 510–517.

Misteli T, Soutoglou E. 2009. The emerging role of nuclear architecture in DNA repair and genome maintenance. *Nat Rev Mol Cell Biol* **10:** 243–254.

Nagai S, Dubrana K, Tsai-Pflugfelder M, Davidson MB, Roberts TM, Brown GW, Varela E, Hediger F, Gasser SM, Krogan NJ. 2008. Functional targeting of DNA damage to a nuclear pore-associated SUMO-dependent ubiquitin ligase. *Science* **322:** 597–602.

Neumann FR, Dion V, Gehlen LR, Tsai-Pflugfelder M, Schmid R, Taddei A, Gasser SM. 2012. Targeted INO80 enhances subnuclear chromatin movement and ectopic homologous recombination. *Genes Dev* **26:** 369–383.

Oza P, Jaspersen SL, Miele A, Dekker J, Peterson CL. 2009. Mechanisms that regulate localization of a DNA double-strand break to the nuclear periphery. *Genes Dev* **23:** 912–927.

Padeken J, Heun P. 2014. Nucleolus and nuclear periphery: velcro for heterochromatin. *Curr Opin Cell Biol* **28:** 54–60.

Pai CC, Deegan RS, Subramanian L, Gal C, Sarkar S, Blaikley EJ, Walker C, Hulme L, Bernhard E, Codlin S, et al. 2014. A histone H3K36 chromatin switch coordinates DNA double-strand break repair pathway choice. *Nat Commun* doi: 10.1038.ncomms5091.

Panier S, Boulton SJ. 2014. Double-strand break repair: 53BP1 comes into focus. *Nat Rev Mol Cell Biol* **15:** 7–18.

Pankotai T, Bonhomme C, Chen D, Soutoglou E. 2012. DNAPKcs-dependent arrest of RNA polymerase II transcription in the presence of DNA breaks. *Nat Struct Mol Biol* **19:** 276–282.

Parada LA, Misteli T. 2002. Chromosome positioning in the interphase nucleus. *Trends Cell Biol* **12:** 425–432.

Pfister SX, Ahrabi S, Zalmas LP, Sarkar S, Aymard F, Bachrati CZ, Helleday T, Legube G, La Thangue NB, Porter AC, et al.

Downloaded from genesdev.cshlp.org on November 10, 2017 - Published by Cold Spring Harbor Laboratory Press

2014. SETD2-dependent histone H3K36 trimethylation is required for homologous recombination repair and genome stability. *Cell Reports* **7:** 2006–2018.

Ptak C, Aitchison JD, Wozniak RW. 2014. The multifunctional nuclear pore complex: a platform for controlling gene expression. *Curr Opin Cell Biol* **28:** 46–53.

Reddy KL, Zullo JM, Bertolino E, Singh H. 2008. Transcriptional repression mediated by repositioning of genes to the nuclear lamina. *Nature* **452:** 243–247.

Rogakou EP, Pilch DR, Orr AH, Ivanova VS, Bonner WM. 1998. DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139. *J Biol Chem* **273:** 5858–5868.

Roukos V, Misteli T. 2014. The biogenesis of chromosome translocations. *Nat Cell Biol* **16:** 293–300.

Roukos V, Voss TC, Schmidt CK, Lee S, Wangsa D, Misteli T. 2013. Spatial dynamics of chromosome translocations in living cells. *Science* **341:** 660–664.

San Filippo J, Sung P, Klein H. 2008. Mechanism of eukaryotic homologous recombination. *Annu Rev Biochem* **77:** 229–257.

Schober H, Ferreira H, Kalck V, Gehlen LR, Gasser SM. 2009. Yeast telomerase and the SUN domain protein Mps3 anchor telomeres and repress subtelomeric recombination. *Genes Dev* **23:** 928–938.

Soutoglou E, Misteli T. 2008. Activation of the cellular DNA damage response in the absence of DNA lesions. *Science* **320:** 1507–1510.

Soutoglou E, Dorn JF, Sengupta K, Jasin M, Nussenzweig A, Ried T, Danuser G, Misteli T. 2007. Positional stability of single double-strand breaks in mammalian cells. *Nat Cell Biol* **9:** 675–682.

Taddei A, Van Houwe G, Hediger F, Kalck V, Cubizolles F, Schober H, Gasser SM. 2006. Nuclear pore association confers optimal expression levels for an inducible yeast gene. *Nature* **441:** 774–778.

Therizols P, Fairhead C, Cabal GG, Genovesio A, Olivo-Marin JC, Dujon B, Fabre E. 2006. Telomere tethering at the nuclear periphery is essential for efficient DNA double strand break repair in subtelomeric region. *J Cell Biol* **172:** 189–199.

Zhang Y, Jasin M. 2011. An essential role for CtIP in chromosomal translocation formation through an alternative end-joining pathway. *Nat Struct Mol Biol* **18:** 80–84.

Materials and methods

Cloning and plasmids

Construct	Origin
GFP-lacI-ΔEMD	(Reddy et al. 2008)
GFP-lacI	(Reddy et al. 2008)
cherry-lacI	(Soutoglou and Misteli 2008)
BRG1- cherry-lacI	gift from Tom Misteli
Pom121 GFP lacI	(see below)
pEXPR-EF1α-Pom121A-Venus	gift from Naoko Imamoto
pWHE320-HA-I-SceI	(Lemaitre et al. 2012)
pWHE146-Tet activator	(Lemaitre et al. 2012)

pCXPA-POM121A-EGFP-LacI was assembled with the universal expression system (manuscript in preparation) in a single cloning reaction with 5 fragments and using type IIS restriction enzymes

Super-resolution imaging and analysis

Immunofluorescence was performed as described above. Postfixation in 4% formaldehyde for 20 min was performed prior to imaging. The super-resolution microscopy experiments were performed on a Leica SR GSD system that consists of: Leica DMI6000 B inverted microscope with HCX PL APO 100x/1.47 Oil CORR TIRF PIFOC objective and 1.6x magnification lens for resulting pixel size of 100 nm; Andor

iXon3 DU-897U-CS0-#BV EMCCD camera with field of view of 18x18 μm in GSDIM mode; continuous wave fibre lasers (MPBC Inc., 488 nm 300 mW, 532 nm 1000 mW, 642 nm 500 mW); a diode laser 405 nm 30 mW; supressed motion (SuMo) sample stage with reduced drift.

For the super-resolution imaging the samples were mounted in PBS buffer that contained 10 mM of cysteamine (Sigma) and that was adjusted to pH 7.5 with 25 mM of HEPES. MEA was dissolved at 1M in PBS and was stored at -20°C. The final dilution was done prior to imaging.

For imaging of Alexa-488 we used the 488 nm laser as excitation source, filter cube with excitation filter DBP 405/10 488/10, dichroic mirror LP 496 and emission filter BP 555/100. For Alexa-647 — 642 nm laser, DBP 405/10 642/10, LP 649 and BP 710/100, respectively. The two colour channels were imaged sequentially: first Alexa-647, then Alexa-488. The excitations were performed at 100% power of corresponding lasers; the acquisitions started after beginning of observation of single-fluorophore events ("blinking") that corresponded to 1-2 min of excitation for Alexa-488 and 1-5 s for Alexa-647. The time of exposition of a frame was 50 ms at 488 nm and 10 ms at 647 nm. After few minutes of acquisition, as number of blinking evens dropped, the sample started to be illuminated additionally by 405 nm laser with gradual increase of its intensity in order to keep a constant rate of single-molecular returns into the ground state. The acquisition stopped after complete bleaching of the fluorophore. Duration of acquisitions was typically 10-20 min for Alexa-488 and 7-10 min for Alexa-647.

The localization and fitting of single-molecular events were performed in Leica LAS AF 3.2.0.9652 software with "center of mass" fitting method. Close events on consecutive frames, most likely originating from the same fluorophore, were merged using a corresponding option in the software. Maximal number of events to merge was set to 10, radius – to 50 nm. The obtained event lists, containing for each event: frame ID, coordinates x y, fitted number of photons, standard deviations σ_x σ_y for fitted 2D-Gaussians, were exported in an .ascii file and analysed further using a custom software written in Matlab. Super-resolution images, were calclulated with grey value of a pixel as quantity of localizations detected in the pixel area.

In order to reduce chromatic aberrations, the microscope was calibrated with multi-colour fluorescent beads (Tetraspeck, d=200 nm). The same area of a coverslip with beads was excited by 488 nm and 642 nm laser light; obtained pair of images appeared shifted on 20-60 nm for each bead, depending on lateral position of the bead in the field of view. The values of the offset were fitted to the x and y position on the image by a 2-order polynom. The obtained fit was subtracted from coordinates of each event of the red channel, resulting in residual chromatic offset less than 25 nm through all the field of view.

In order to reduce a drift of the sample, each single-color acquisition was divided onto two successive parts with equal number of events. From each part, a super-resolution image was reconstructed. The shift between the two images was calculated with subpixel precision by cross-correlation using a Matlab function(Manuel Guizar-Sicairos 2008) . The obtained value was fitted linearly into full range of frames and was subtracted from each single-molecular localization. The red channel events were shifted towards the final frame of the red colour acquisition,

the green channel ones – towards the first frame of the green colour acquisition, in order to reduce an additional offset between two colours, produced by drift and sequential imaging.

We were not able to reliably calculate shifts between smaller datasets due to not enough quantity of localizations for reconstruction more than two resembling images. So with this approach only a constant component of drift may be reduced, that is yet the most significant on our system.

We performed coordinate-based colocalization analysis of single-molecule localization data of two species. For each single-molecular event A*i* we calculated a colocalization value C_{Ai} that adapts values from -1 (for anti-correlated distributions) through 0 (for non-correlated) to +1 (for perfectly correlated distributions)(Malkusch et al. 2012). For the calculation of C_{Ai} we took into account all the localizations of both colours around A*i* within radii from 2 nm to 500 nm with step of 2 nm.

A histogram of distribution of C_A showing overall colocalization level was calculated for each double-colour image. We also calculated a global colocalization value for each image by division the sum of all positive values C_{Ai} by the sum of all negative values.

Quantification of the distribution of m6A-Tracer intensity

The quantification of distribution of m6a-Tracer was done using a macro on ImageJ, available upon request. Ratio of intensity of m6a-Tracer in the nucleoplasm over the intensity at the nuclear enveloped was then calculated.

Antibodies

Antibodies	Company	Reference	Application
laminB	Santa Cruz	SC-6216	ImmunoFISH, IF
γH2AX	Abcam	Ab22551	ImmunoFISH, IF
53BP1	Novus	NB100-304	ImmunoFISH, IF
Brca1	Calbiochem	OP92+OP93	ImmunoFISH, IF
Rad51	Calbiochem	PC130	ImmunoFISH, ChIP
Rad54	Abcam	Ab11055	ImmunoFISH
Ku80	Santa Cruz	SC-56136	ImmunoFISH
γH2AX	Abcam	Ab2893	ChIP
RPA	Novus	NB600-565	ChIP
P-RPA	Bethyl	A-300 245A	ChIP
BRCA1	Santa Cruz	SC-642	ChIP
XRCC4	Abcam	Ab145	ChIP, WB
Tubulin	Sigma	DM1A	WB
GFP	IGBMC		IF prior to GSDIM
TPR	Abcam	Ab84516	IF prior to GSDIM

Electron spectroscopic imaging (ESI) of chromatin structure and variation. Human U2OS osteosarcoma cells were treated with either vehicle (0.1% DMSO) or with 500 nM trichostatin A (TSA) for 4 h before being fixed in 4% paraformaldehyde (EMS) for 10 min at room temperature (RT) prior to being permeabilized in PBS containing 0.5% Triton X-100 for 5 min. Cells were then "post fixed" in 1% glutaraldehyde

(EMS) for 5 min at RT to maintain chromatin structure during resin embedding. The cells were then dehydrated in an ethanol series and embedded in Quetol 651 (EMS) before being processed, sectioned and imaged by ESI as previously described (Dellaire G 2004)using a Tecnai 20 transmission electron microscope (FEI) equipped with an energy-filtering spectrometer (Gatan). Energy-filtered electron micrographs of nitrogen (N) and phosphorus (P) were collected, and non-chromosomal protein was segmented by subtracting the N from the P ESI micrograph, which was then false colored in cyan and combined in a composite image with the P ESI micrograph false colored in yellow in Photoshop CS6 (Adobe) to highlight chromatin. The composite elemental maps of N-P (cyan) and P (yellow) were then analyzed for thickness of nuclear-lamina-associated chromatin using Image J v1.48k software (NIH). Pixel measurements (50 measurements taken from 10 cells) were converted into microns (µm) and then averaged per cell, and the data was represented as mean chromatin thickness± SEM (where N=10). Statistical significance between cell lines was generated using the Student's *t* test in Excel (Microsoft). The mean coefficient of variation (CV) in chromatin density was calculated for chromatin within the nucleus of vehicle and TSA treated U2OS cells (N=5), using phosphorus-enriched 155 keV electron micrographs as previously described(Dellaire et al. 2009). Briefly, the mean and SD pixel intensities were first determined from 5 X 10 pixel-wide line scans per cell using Image J. Then for each cell the CV was determined by dividing the mean pixel intensity by the SD, after which the CVs were averaged for vehicle or TSA treated cells and represented as a percentage ± SEM.

Measurement of the size of the lacO array

The lacO array sizes at different conditions were measured on paraformaldehyde fixed samples. The images were taken by Leica DM6000 microscope with Leica CSU22 spinning disc and Andor Ixon 897 camera. For every condition at least 20 individual cells were imaged and analyzed. The Z planes were taken every 0.3 mm. For 3D reconstruction and quantification of volumes the Imaris software (Bitplane) was used.

ChromatinIP

The ChIP analysis was done following the Dynabeads ChIP protocol from Abcam(Pankotai et al. 2012) with a few modifications. Briefly, one 150-mm dish with cells that were 70% confluent was used for each time point. The cells were cross-linked for 30 min in 0.75% (v/v) paraformaldehyde and then sonicated in 1% (v/v) SDS-containing sonication buffer (50 mM HEPES, pH 8, 140 mM NaCl, 1 mM EDTA, 1% (v/v) TritonX-100, 1% (v/v) SDS and a protease inhibitor cocktail (Roche)). Thirty milligrams of chromatin were diluted in RIPA buffer (50 mM Tris-HCl, pH 8, 150 mM NaCl, 1 mM EDTA, 1% Triton X-100 (v/v), 0.1% sodium deoxycholate (w/v) and 0.1% SDS (v/v)) and were used in each immunoprecipitation by adding 4 μg of antibody and 50 μl Dynabeads M-280 (Invitrogen). The beads were washed for 5 min with low-salt buffer (20 mM Tris-HCl, pH 8, 150 mM NaCl, 2 mM EDTA, 1% (v/v) Triton X-100 and 0.1% (v/v) SDS), then 5 min with high-salt buffer (20 mM Tris-HCl, pH 8, 500 mM NaCl, 2 mM EDTA, 1% (v/v) Triton X-100 and 0.1% (v/v) SDS) and for 5 min with LiCl buffer (10 mM Tris-HCl, pH 8, 250 mM LiCl, 1 mM EDTA, 1% (v/v) NP-40 and 1% (w/v) sodium deoxycholate) and two times for 5 min with TE

buffer. The elution was done twice at 65 °C for 15 min. Cross-links were reversed by incubation at 65 °C for 6 h. The DNA was purified after proteinase K and RNaseA treatment by using phenol-chloroform extraction and was resuspended in 50 µl of TE buffer.

The signal in each experiment was calculated using the formula (immunoprecipitated sample–IgG control)/input, and each value represents a relative DNA concentration that is based on the standard curve of the input.

Dellaire G, Kepkay R, Bazett-Jones DP. 2009. High resolution imaging of changes in the structure and spatial organization of chromatin, γ-H2A.X and the MRN complex within etoposide-induced DNA repair foci. *Cell Cycle* **8**: 3750-3769.

Dellaire G NR, and Bazett-Jones DP 2004. Correlative light and electron spectroscopic imaging of chromatin in situ. *Methods Enzymol* **375**: 456-478.

Malkusch S, Endesfelder U, Mondry J, Gelléri M, Verveer P, Heilemann M. 2012. Coordinate-based colocalization analysis of single-molecule localization microscopy data. *Histochem Cell Biol* **137**: 1-10.

Manuel Guizar-Sicairos STT, and James R. Fienup 2008. Efficient subpixel image registration algorithms. *Opt Letter* **33**: 156-158.

Supplementary figure legends

Figure S1- Experimental system

(A) Schematic representation of the experimental system: the lac repressor (GFP-lacI) binds to the lac operator (lacO), which allows the relocalization of the lacO array to the nuclear lamina when fused to ΔEMD. The addition of Dox allows the expression of ISceI and induction of a DSB at the I-SceI restriction site, which is located next to the lacO repeats. **(B)** Quantification of the colocalization of the lacO array with lamin B in absence or presence of dox for 14h. Values represent means ± SD from three independent experiments (number of cells analyzed per experiment ≥ 50). **(C)** 3D Z-stacks confocal microscopy images of the lacO array (green) and lamin B (gray) in I-U2OS19 GFP-lacI cells (upper panel) or GFP-lacI-ΔEMD cells (lower panel) **(D)** LM-PCR in GFP-lacI and GFP-lacI- ΔEMD cells 14h after Dox addition. Products are shown (upper panel) and quantified (below) after 26 (left) or 28 (right) PCR cycles.

Figure S2- DDR is delayed at the nuclear lamina

(A) Quantification of the colocalization of the lacO array with γH2AX at the indicated times after dox addition in I-HeLa111 infected with GFP-lacI or GFP-lacI-ΔEMD. Values represent means ± SD of three independent experiments (number of cells analyzed per experiment ≥ 50). (B) Immuno-FISH single-Z confocal images of the lacO array (green) and γH2AX (red) in I-HeLa111 cells expressing GFP-lacI or GFP-lacI-ΔEMD and treated or not with Dox for 14h. For statistical analysis, t test was performed. P value are represented as follows : *<0.05, **<0.01

Figure S3- The expression of GFP-lacI-ΔEMD does not impair DDR activation

(A) Immuno-FISH single-Z confocal images of the lacO array (green) and lamin B (red) in I-U2OS19 GFP-lacI cells (left) or GFP-lacI-ΔEMD cells (right) in presence of 2mM IPTG and in the absence (upper panel) or presence (lower panel) of Dox for 14h. **(B)** Single-Z confocal microscopy images of the lacO array (green) and γH2AX (red) in I-U2OS19 GFP-lacI cells (left) or GFP-lacI-ΔEMD cells (right) in presence of 2mM IPTG and in the absence (upper panel) or presence (lower panel) of Dox for 14h, after immuno-FISH. **(C)** Percentage of the colocalization of the lacO array with lamin B in presence of 2mM IPTG, and in absence or presence of doxycycline for 14h. **(D)** Quantification of the colocalization of the lacO array with γH2AX in presence of 2mM IPTG, and in absence or presence of Dox for 14h. Values represent means ± SD of three independent experiments (number of cells analyzed per experiment ≥ 50).

Figure S4-Recruitment of DSB repair factors at the nuclear lamina in I-U2OS19

Immuno-FISH single-Z confocal images of the lacO array (green), laminB (gray) and (A) Ku80 (red), (B) BRCA1 (red), (C) Rad51 (red) in I-U2OS19 cells expressing GFP-lacI or GFP-lacI-ΔEMD and treated or not with Dox for 14 or 20h

Figure S5-Recruitment of DSB repair factors at the nuclear lamina in I-HeLa111

Immuno-FISH single-Z confocal images of the lacO array (green), laminB (gray) and (A) Ku80 (red), (B) BRCA1 (red), (C) Rad51 (red) in I-HeLa111 cells expressing GFP-lacI or GFP-lacI-ΔEMD and treated or not with Dox for 14 or 20h **(D-F)** Quantification of the colocalization of the lacO array with **(D)** Ku80, **(E)** BRCA1, **(F)** RAD51, at the indicated times after dox addition in I-HeLa111 infected with GFP-lacI or GFP-lacI-ΔEMD. Values represent means ± SD of three independent experiments (number of cells analyzed per experiment ≥ 50). For statistical analysis, t test was performed. P value are represented as follows : *<0.05, **<0.01

Figure S6-Recruitment of Rad54 at the nuclear lamina in I-U2OS19

(A) Quantification of the colocalization of the lacO array with Rad54 at the indicated times upon Dox addition in I-U2OS19 GFP-lacI or GFP-lacI-ΔEMD cells. Values represent means ± SD of three independent experiments (number of cells analyzed per experiment ≥ 50). For statistical analysis, t test was performed. P value are

represented as follows : *<0.05 **(B)** Cell cycle profiles of I-U2OS19 GFP-lacI and I-U2OS19 GFP-lacI-ΔEMD cells.

Figure S7- TSA treatment induces chromatin decompaction

(A) Immunofluorescence images depicting H4acetylation (red) in I-U2OS19 GFP-lacI cells treated with DMSO or TSA. Nuclear stain, DAPI (blue) **(B-C)** Human U2OS osteosarcoma cells were treated with **(B)** vehicle (0.1% DMSO) or **(C)** with 500 nM trichostatin A (TSA) for 4 h before fixation and processing for electron spectrocopic imaging (ESI). In each row a low magnification phosphorus-enriched (155 KeV) electron micrograph is shown at the left, a line-scan of phosphorus intensity across the cell nucleus (between the white arrows) is shown in the middle panel, and on the far right a high magnification ESI electron micrograph is shown of the region outlined by a white dashed box in the low magnification micrograph. The coefficient of variation (CV) is also shown for the phosphorus intensity across the nuclei of vehicle and TSA treated cells (n=5; ± SEM); which represents the degree of variability in chromatin density as a percentage, where a lower percentage indicates a more homogenous chromatin density. The ESI micrographs have been false coloured such that chromatin appears yellow and non-chromosomal protein (e.g. nucleopores, marked by white astericks) appears cyan. The thickness of the nuclear lamina associate chromatin is demarcated by white arrow heads, N = nucleoli, and the scale bars = 1 micron. **(D)** The mean thickness of condensed chromatin associated with the nuclear lamina for cells treated with vehicle or with 500 nM TSA and depicted as a bar graph. Error bars = SEM, N=10. *p < 0.001
(E) Immuno-FISH single-Z confocal images of the lacO array (green) and lamin B (red) in I-U2OS19 GFP-lacI cells (left) or GFP-lacI-ΔEMD cells (right) in presence of DMSO (uper panel) or TSA (lower panel). **(F)** Quantification of the colocalization of the lacO array with lamin B in absence or presence of dox for 14h in cells treated with DMSO or TSA for 4h. Values represent means ± SD of three independent experiments (number of cells analyzed per experiment ≥ 50).

Figure S8- H2AX phosphorylation at the nuclear lamina is rescued upon TSA treatment

(A) Immuno-FISH single-Z confocal images of the lacO array (green) and γH2AX (red) in I-U2OS19 GFP-lacI cells (left) or GFP-lacI-ΔEMD cells (right) in presence of DMSO (uper panel) or TSA (lower panel).

Figure S9- HR factors recruitment at the nuclear lamina is rescued upon TSA treatment

Immuno-FISH single-Z confocal images of the lacO array (green) and (A) BRCA1 (red) or (B) RAD51 in I-U2OS19 GFP-lacI cells (left) or GFP-lacI-ΔEMD cells (right) in presence of DMSO (uper panel) or TSA (lower panel).

Figure S10- BRG1 tethering induces chromatin decondensation

(A) Schematic representation of the experimental system. The lac repressor (GFP lacI/cherry-lacI) binding to the lac operator (lacO) allows the relocalization of the

lacO array at the nuclear lamina when fused to ΔEMD. The expression of BRG1-cherry-lacI allows local decondensation of the lacO/I-SceI locus. The addition of Dox allows the expression of I-SceI and induction of a DSB at the I-SceI restriction site, next to the lacO repeats. **(B)** Images of 3D reconstruction of nuclei (blue) and the lacO array (red). **(C)** Quantification of the volume of the lacO array, normalized to the volume of the nucleus in GFP-lacI and GFP-lacI-ΔEMD cells expressing cherry–lacI or BRG1-cherry-lacI.

Figure S11- BRG1 tethering rescues HR factors recruitment at the nuclear lamina

Immunofluorescence single-Z confocal images of (A) BRCA1 (gray) or (B) RAD51 (gray) in I-U2OS19 cells expressing GFP-lacI or GFP-lacI-ΔEMD, transfected with cherry-lacI or BRG1-cherry-lacI (red) and treated or not with Dox for 20h.

Figure S12- Recruitment of DDR and HR factors are not impaired by tethering at the nuclear pores

(A) Schematic representation of the experimental system (left panel) for relocalization of the lacO locus to nuclear pores. Expression of Pom121-GFP-lacI allows the repositioning of the lacO locus to the nuclear pores. Immuno-FISH single-Z confocal image (right upper panel) of the lacO array colocalizing with lamin B in Pom121-GFP-lacI expressing cells. D-Storm picture of Pom121-GFP-lacI (green) and nucleoporin TPR (red) showing colocalization of the lacO array with the nucleoporin TPR (right lower panel). Immuno-FISH single-Z confocal images of the lacO array (green), laminB (gray) and (B) 53BP1 (red), (C) BRCA1 (red), (D) RAD51 (red) in I-U2OS19 cells expressing GFP-lacI or GFP-lacI-ΔEMD and treated or not with Dox for 20h.

Figure S13- Positional stability of LADs upon DNA damage

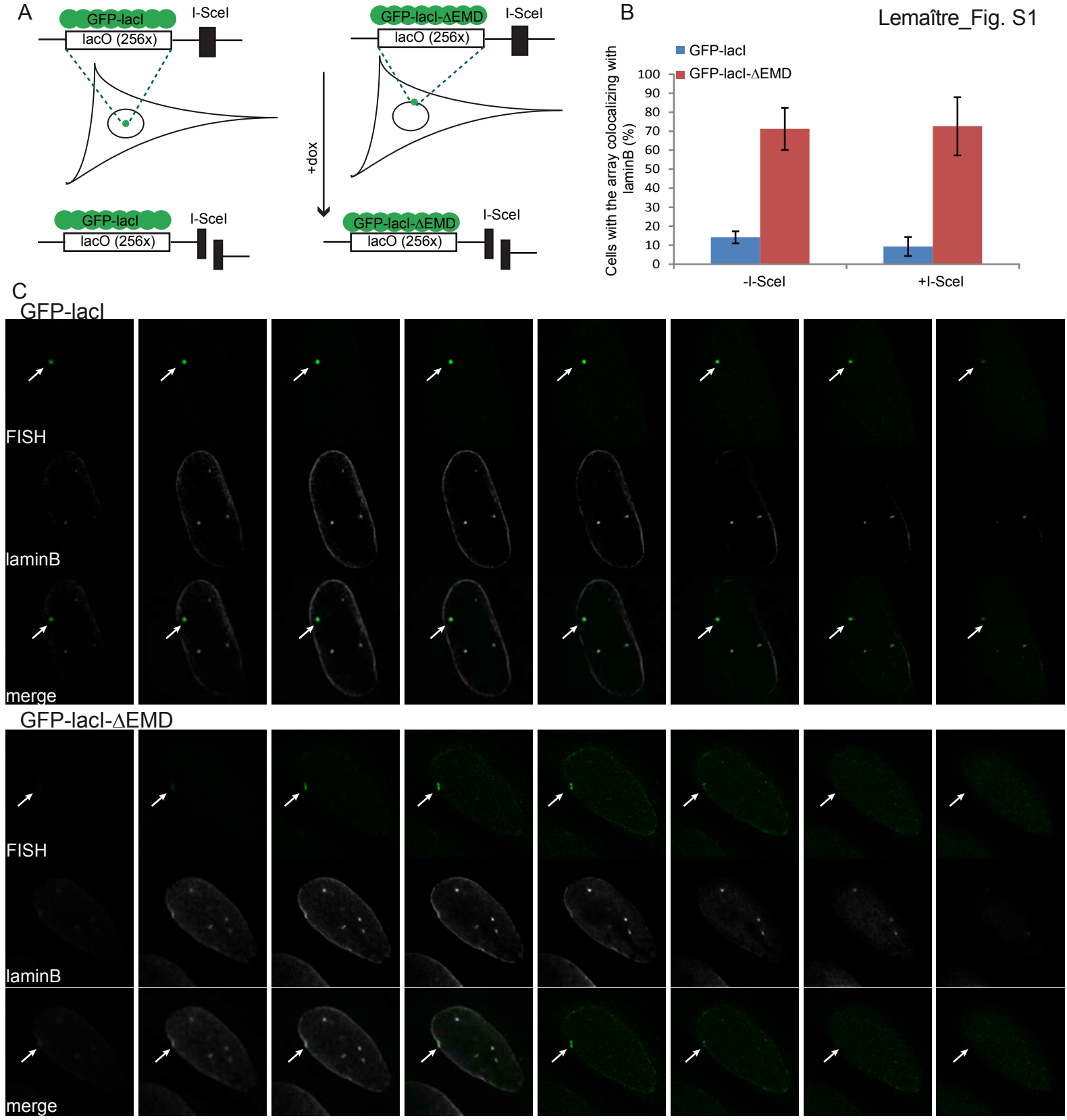
(A) Schematic representation of the experimental system. IPTG addition for 2h in GFP-lacI-ΔEMD after lacO repositioning to the periphery and DSB induction (with 14h dox treatment) allows the dissociation of the lacI from the lacO and a potential movement away from the nuclear lamina. **(B)** Percentage of colocalization of the lacO array with lamin B in absence or presence of dox (14h) in GFP-lacI- ΔEMD treated (for 2h) or not with IPTG. Values represent means ± SD of three independent experiments (number of cells analyzed per experiment ≥ 50). **(C)** Time lapse microscopy on HT1080 cells expressing Dam-laminB1 and m6a-Tracer (green) upon addition (or not) of 50ng/mL NCS for 15min. **(D)** D-STORM pictures of LADs colocalization with γH2AX (red) in HT1080 cells expressing Dam-laminB1 and m6a-Tracer (green) upon addition (or not) of 50ng/mL NCS for 15 min and release for 2 h.

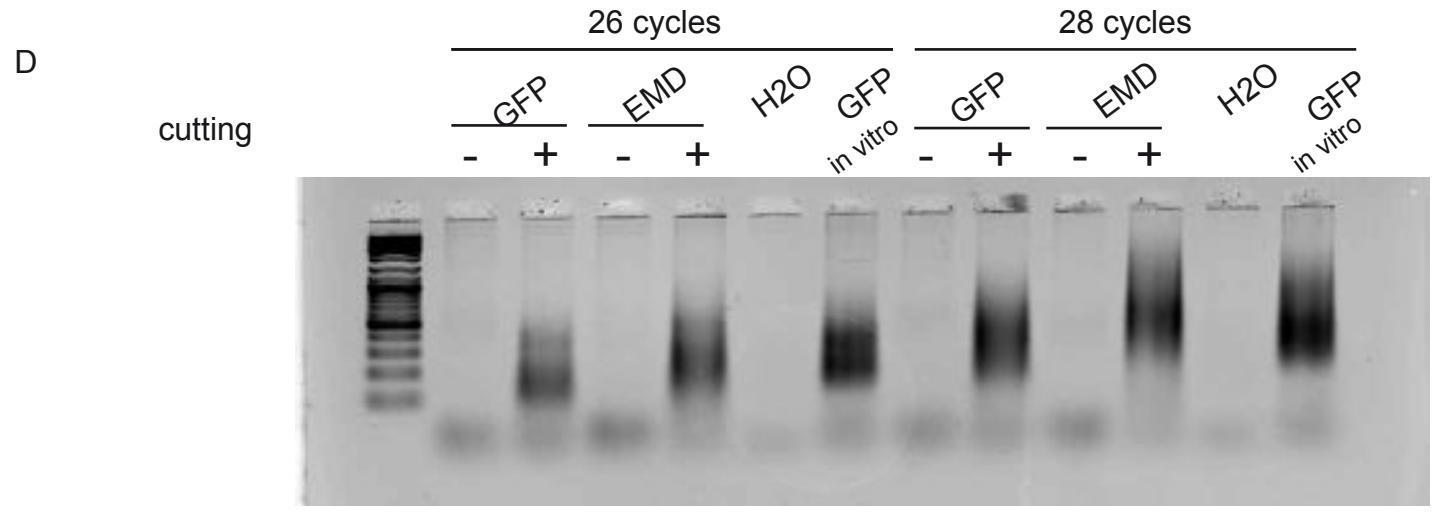
Figure S14- Validation of silencing of ligase 3, XRCC4, RAD51, PARP1, XRCC1 by siRNA

(A) Western blot for tubulin, XRCC4, Rad51, PARP1, XRCC1 in I-U2OS19 GFP-lacI or GFP-lacI-ΔEMD treated with corresponding siRNAs. **(B)** LM-PCR in GFP-lacI and GFP-lacI-ΔEMD cells non-treated, 14h after Dox addition or 36h after a 14h Dox

pulse. Products are shown (upper panel) and quantified (below) after 28 PCR cycles. The intensity of the products depicted is normalized to the products of the non-treated samples. (C) Quantitative RT-PCR analysis of ligase *3* expression levels in I-U2OS19 GFP-lacI or GFP-lacI-ΔEMD cells treated with siRNA that targets a scramble sequence (purple lines) and ligase 3 sequences (blue lines). (D) Percent colocalization of the lacO array with γH2AX in untreated cells (NT) or after 14h of Dox (time point 0) and subsequent release for 24h in I-U2OS19 cells expressing GFP-lacI or GFP-lacI-ΔEMD and transfected with different ligase3-specific siRNAs (siLig3-6 or siLig3-7).

Lemaître_Fig. S1

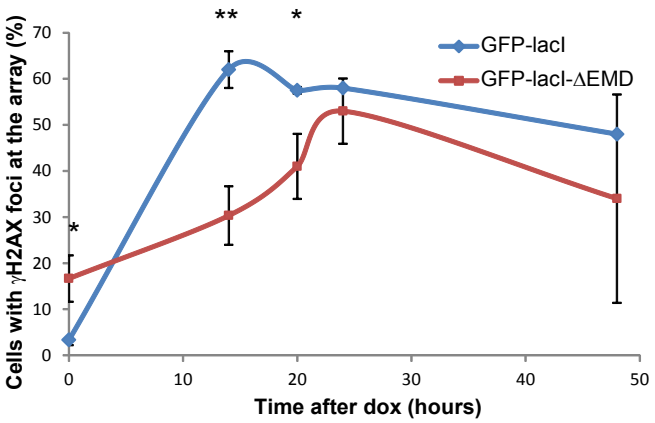




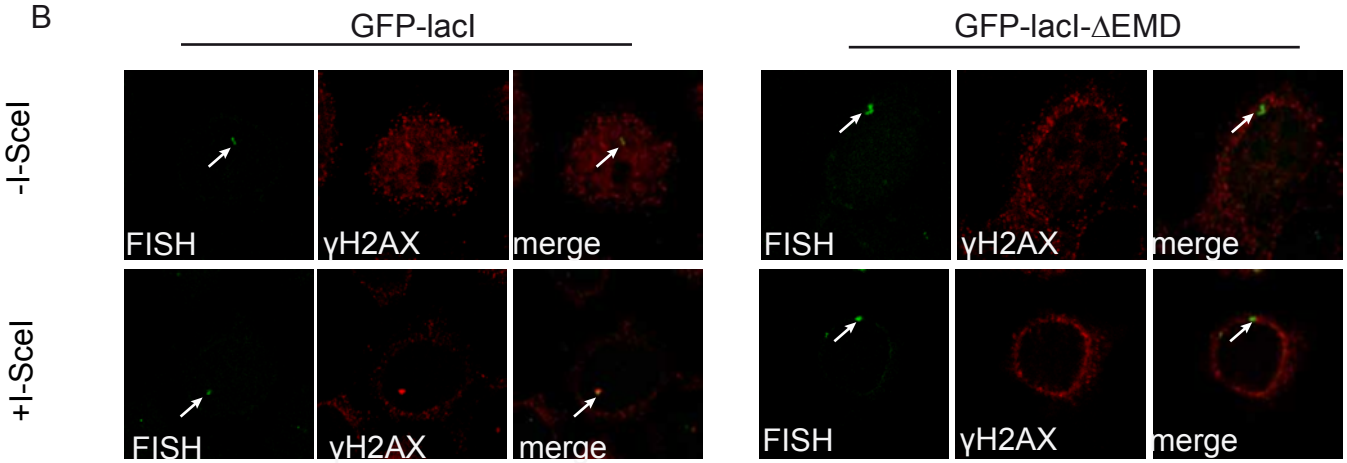
average intensity-background	1.7	**15.8**	-0.1	**22.3**	0	28.2	2.4	**28.6**	1.1	**32.5**	0	36.2

A

I-HeLa 111


Cells with γH2AX foci at the array (%)
Time after dox (hours)
GFP-lacI
GFP-lacI-ΔEMD
**
*
*
0
10
20
30
40
50
60
70

B


GFP-lacI
GFP-lacI-ΔEMD
-I-SceI
+I-SceI
FISH
γH2AX
merge

+IPTG

A

GFP-lacI

GFP-lacI-ΔEMD

-dox

+dox

FISH

laminB

merge

B

-dox

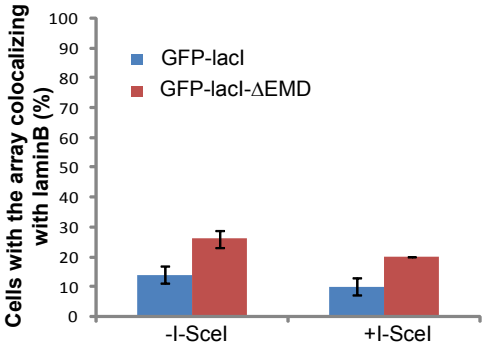
+dox

FISH

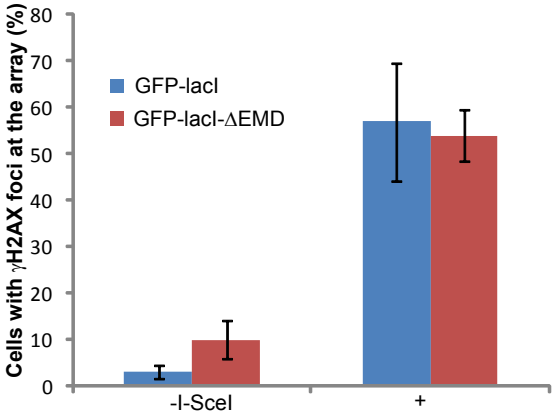
γH2AX

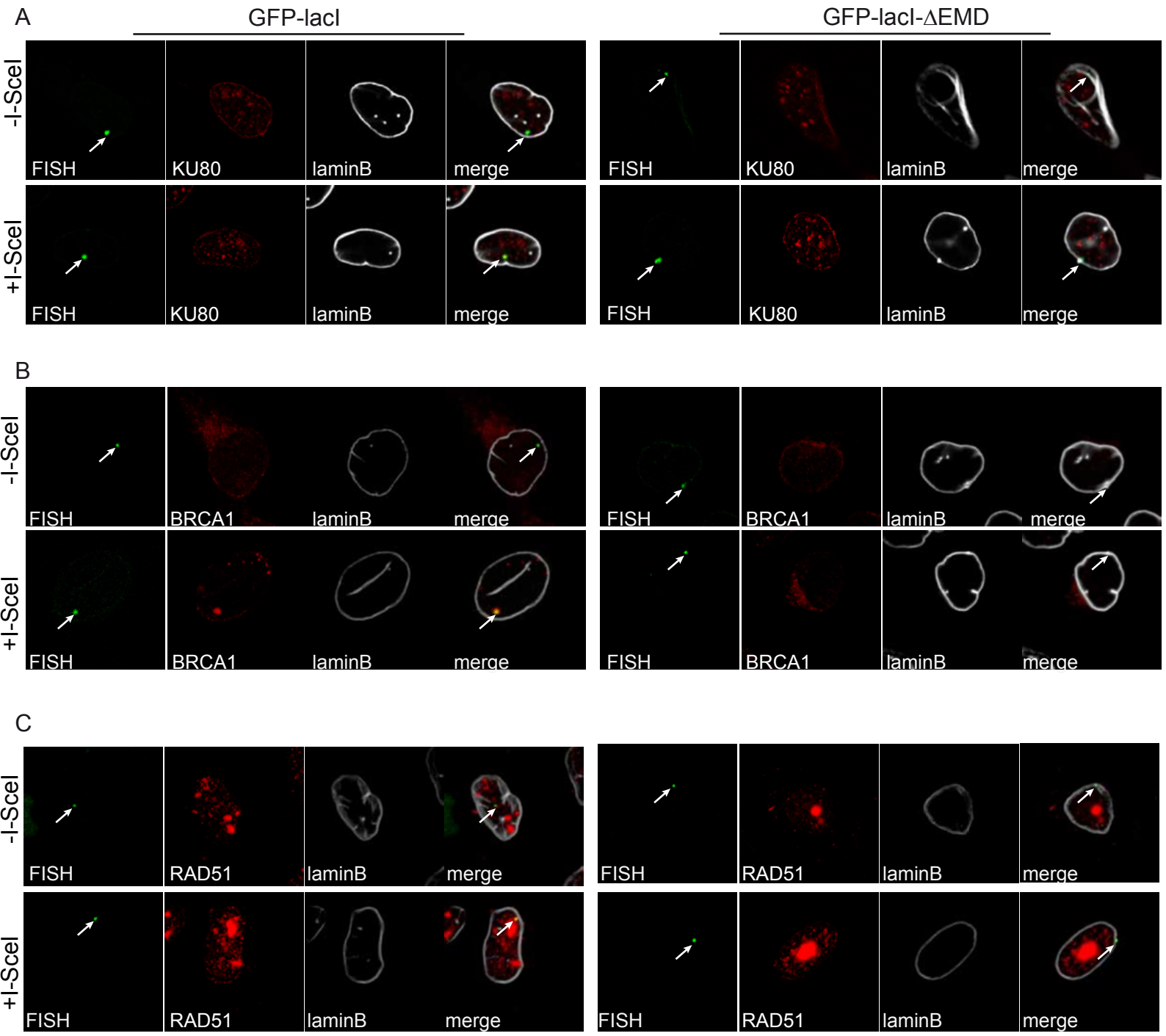
merge

C



D




A
GFP-lacI
GFP-lacI-ΔEMD
-I-SceI
+I-SceI
FISH
KU80
laminB
merge
B
BRCA1
C
RAD51

I-HeLa 111

Lemaître_Fig. S5

A

GFP-lacI

GFP-lacI-ΔEMD

-I-SceI

+I-SceI

FISH

Ku80

merge

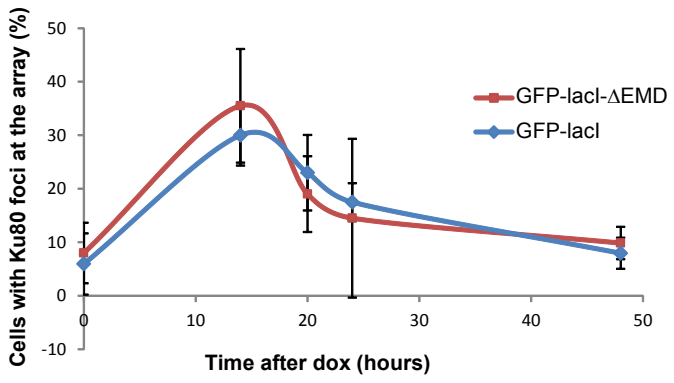
B

BRCA1

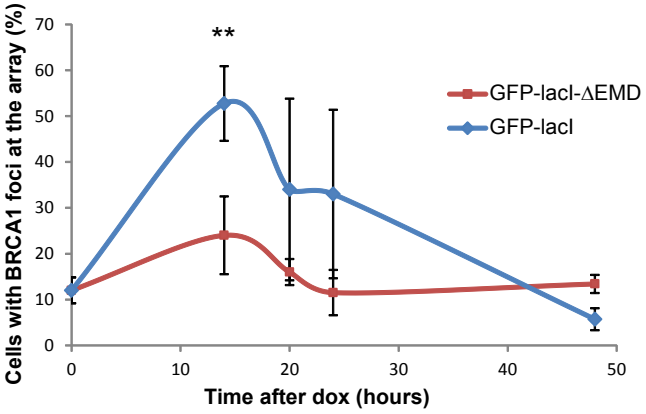
C

RAD51

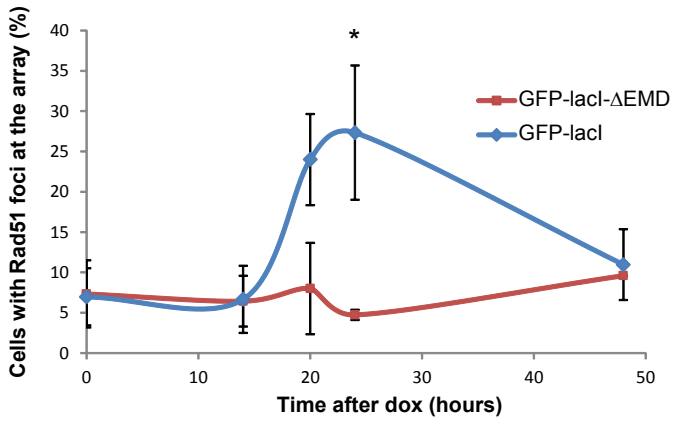
D



E



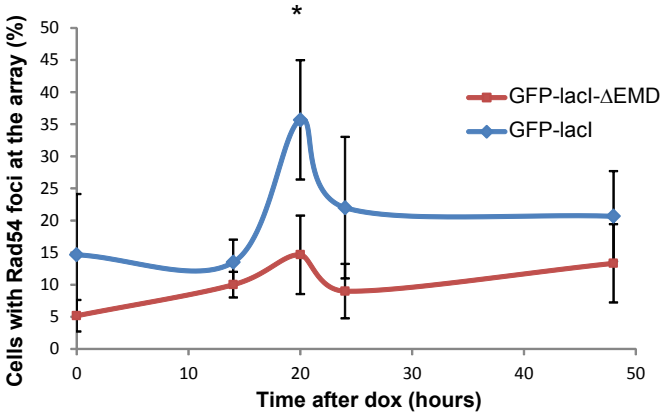
F



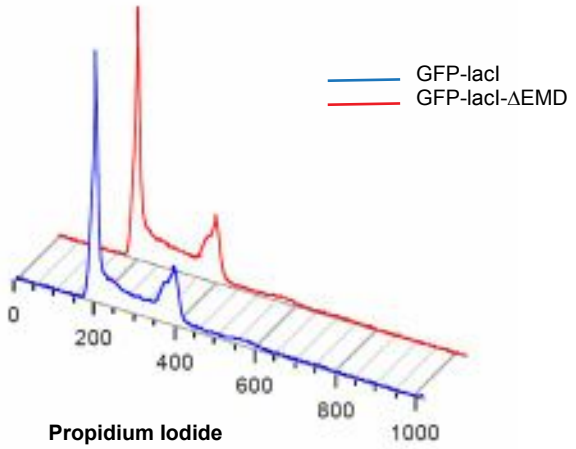
I-U2OS19

Lemaître_Fig. S6

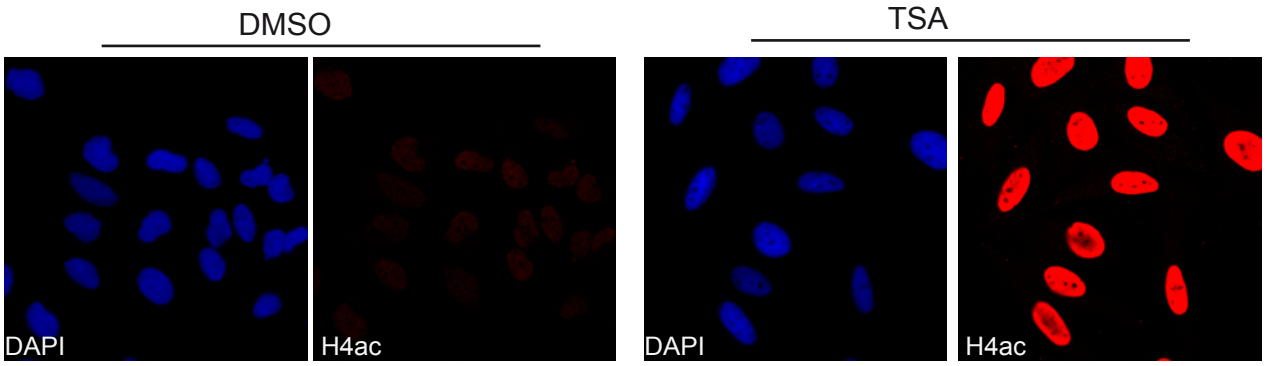
A



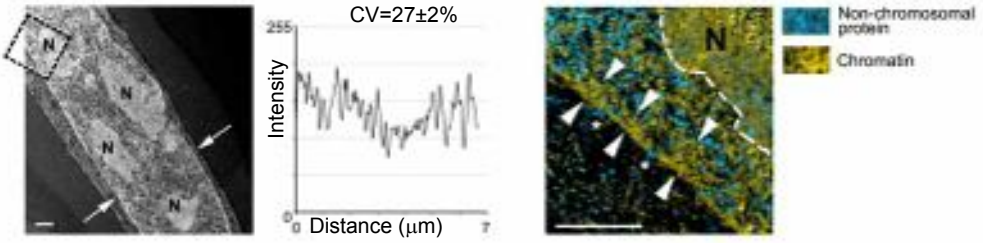
B



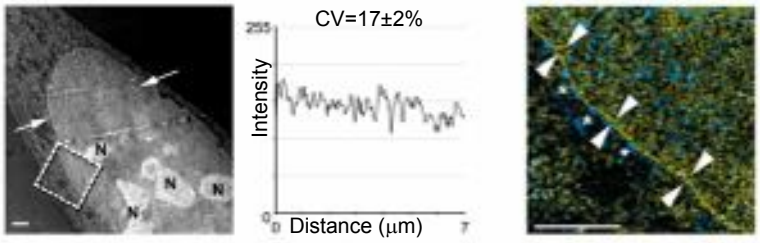
A


DMSO
TSA
DAPI
H4ac
DAPI
H4ac

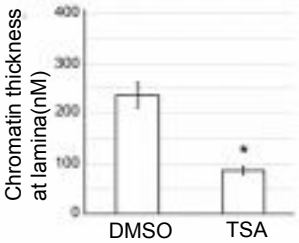
B


CV=27±2%
Intensity
Distance (µm)
Non-chromosomal protein
Chromatin

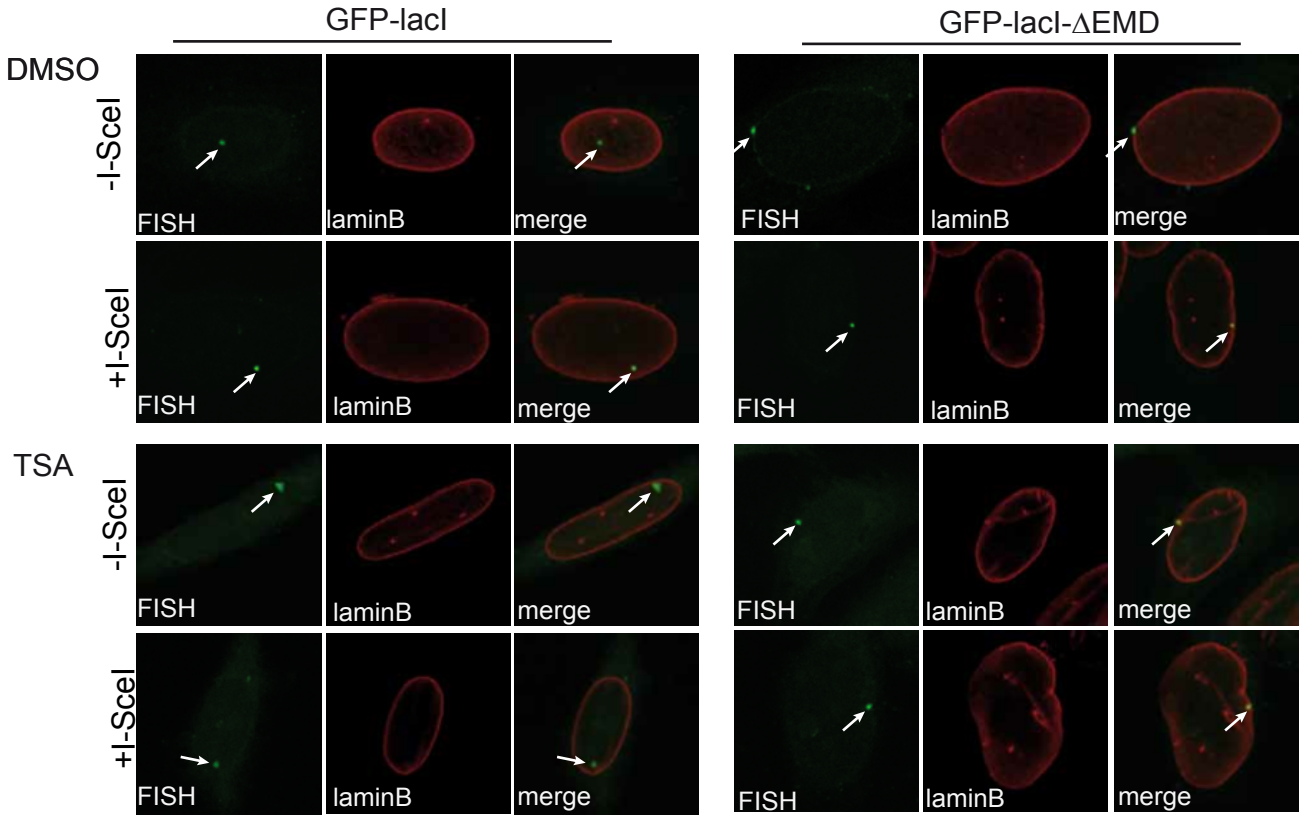
C


CV=17±2%
Intensity
Distance (µm)

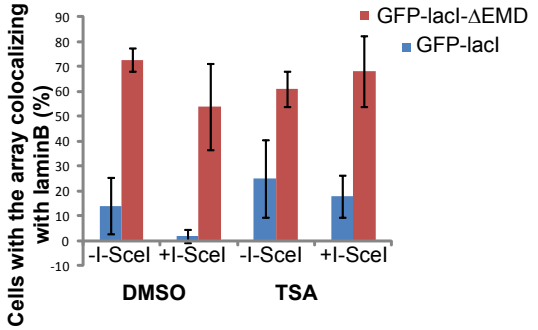
D

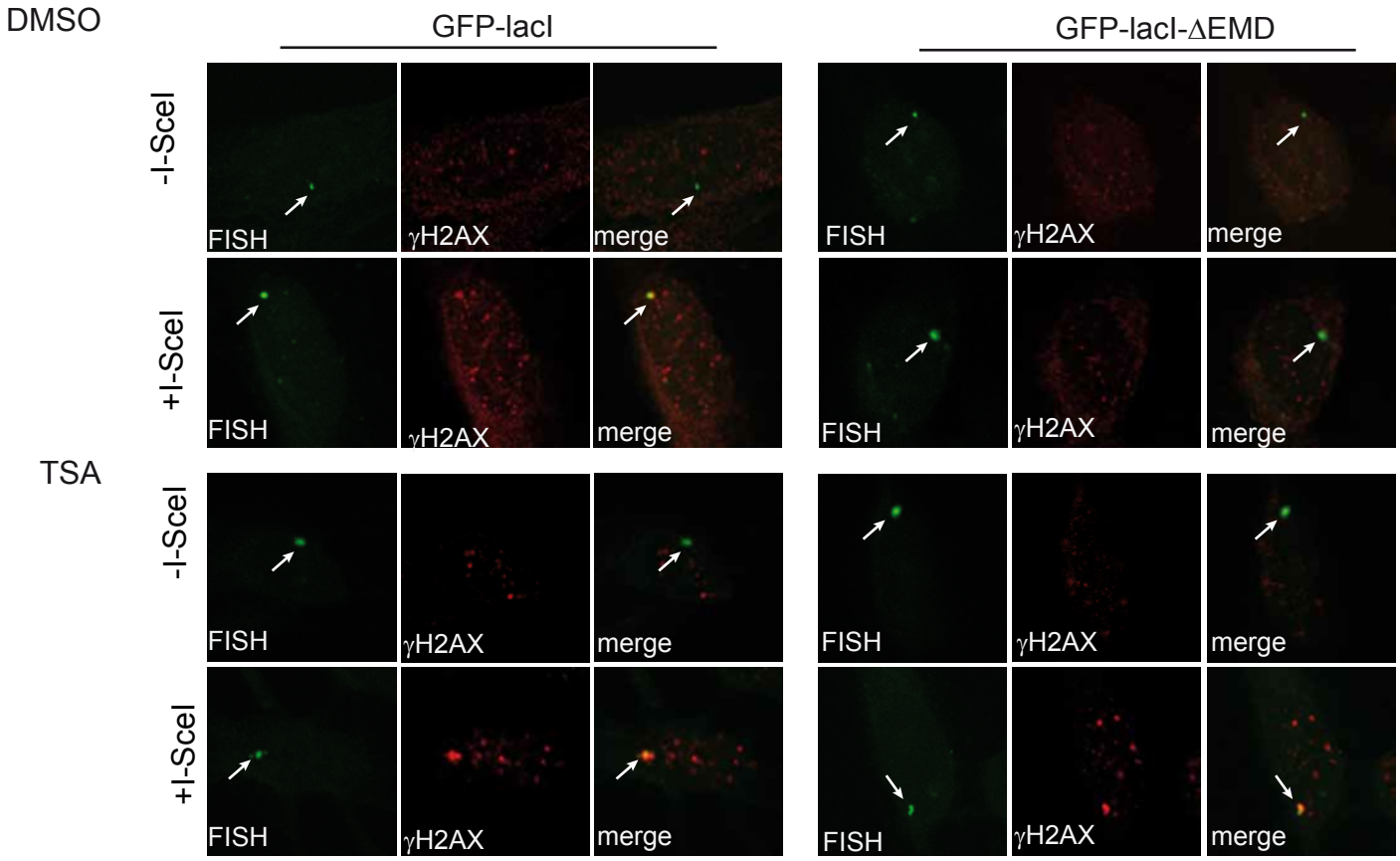

Chromatin thickness at lamina(nM)
DMSO
TSA

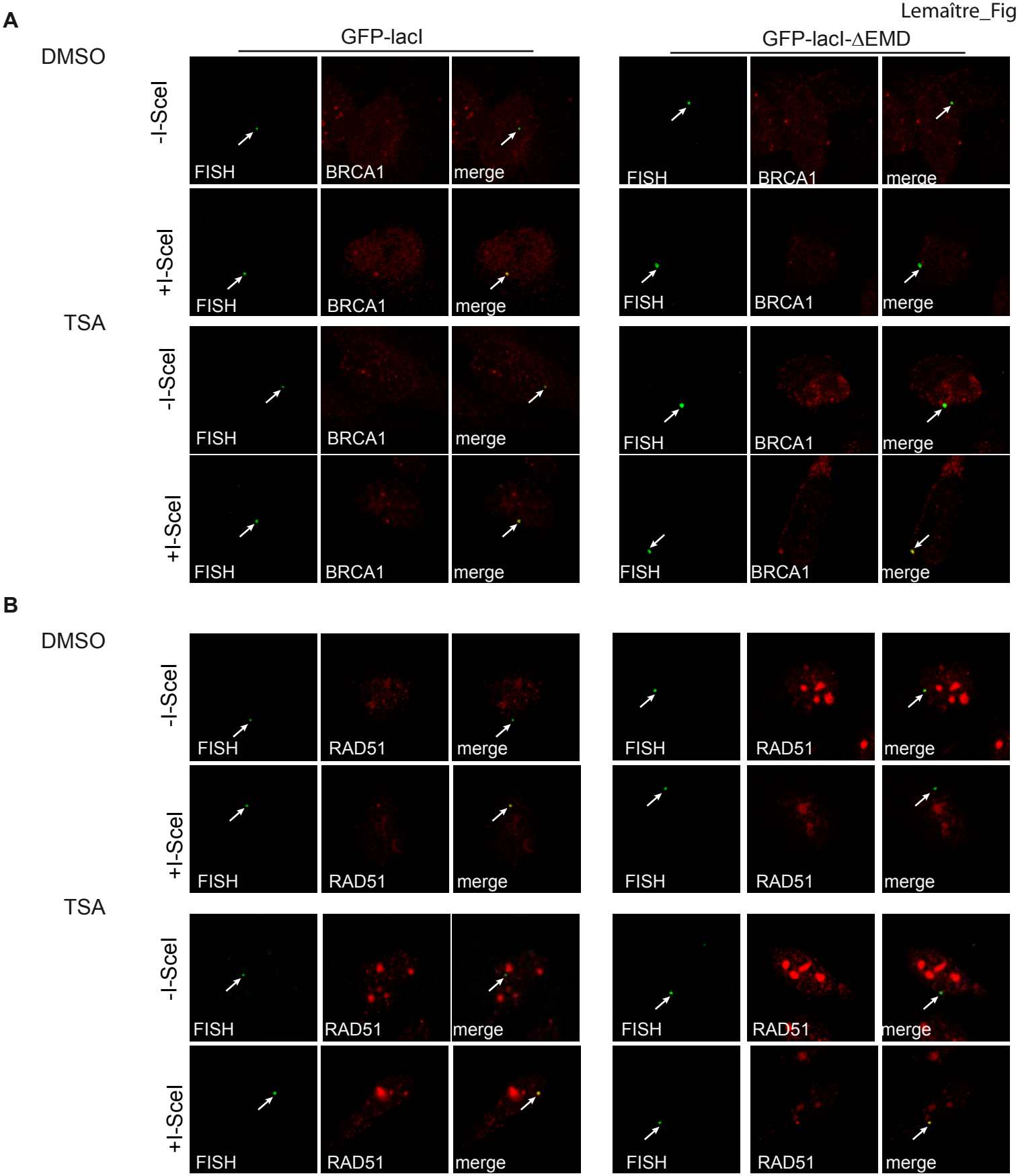
E


GFP-lacI
GFP-lacI-ΔEMD
DMSO
TSA
-I-SceI
+I-SceI
FISH
laminB
merge

F

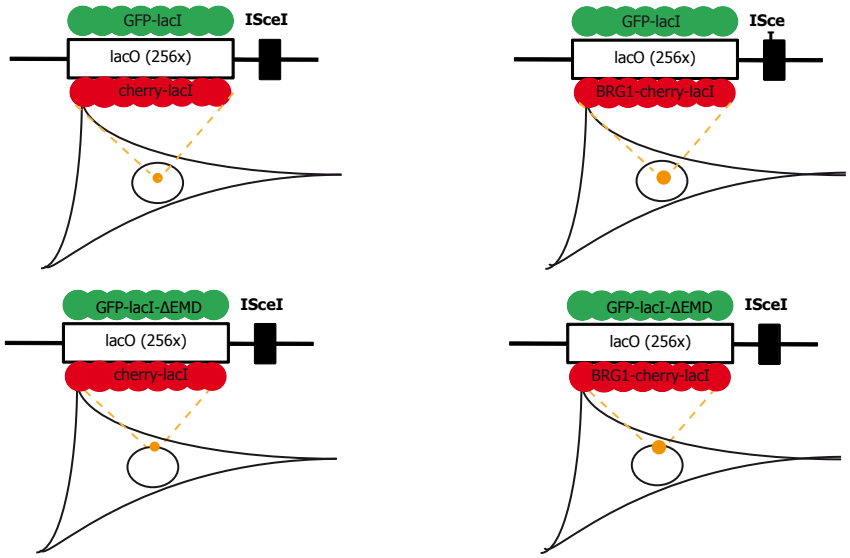

GFP-lacI-ΔEMD
GFP-lacI
Cells with the array colocalizing with laminB (%)
-I-SceI
+I-SceI
-I-SceI
+I-SceI
DMSO
TSA


DMSO
GFP-lacI
GFP-lacI-ΔEMD
-I-SceI
+I-SceI
TSA
-I-SceI
+I-SceI
FISH
γH2AX
merge

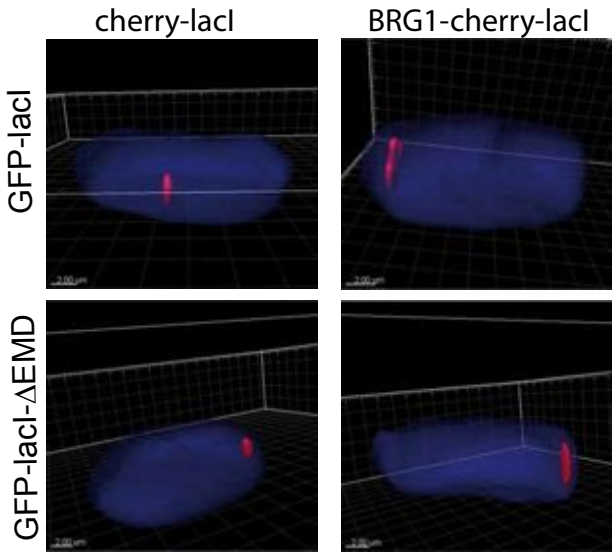

A
GFP-lacI
GFP-lacI-ΔEMD
DMSO
TSA
-I-SceI
+I-SceI
FISH
BRCA1
merge
B
RAD51

Lemaître_Fig. S10

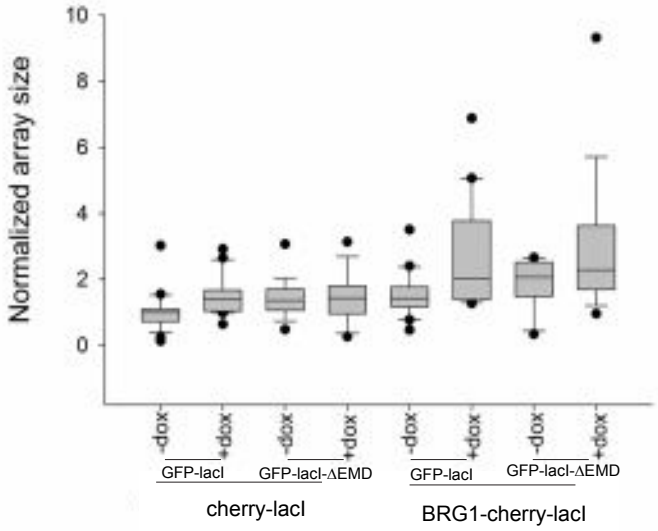
A

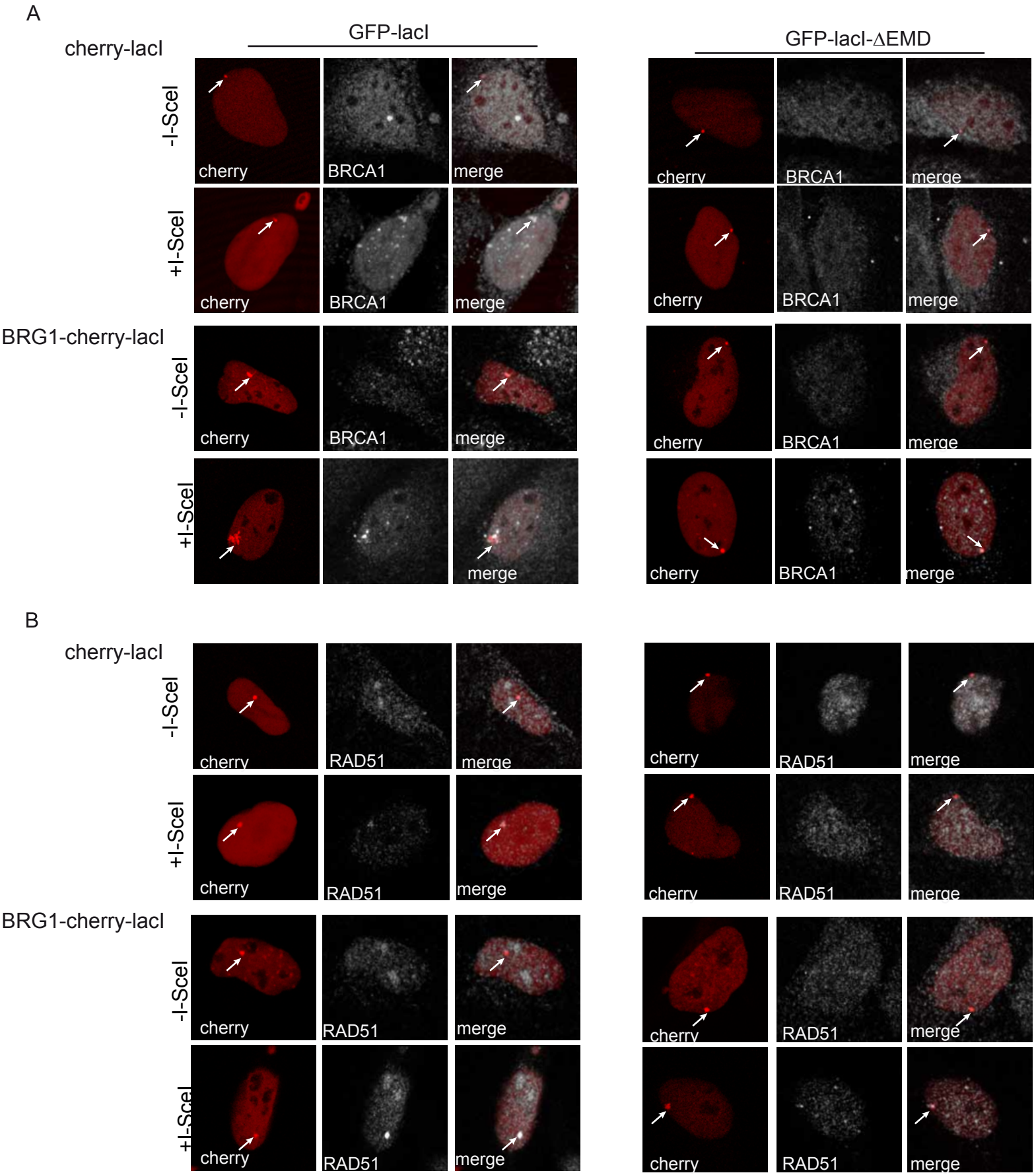


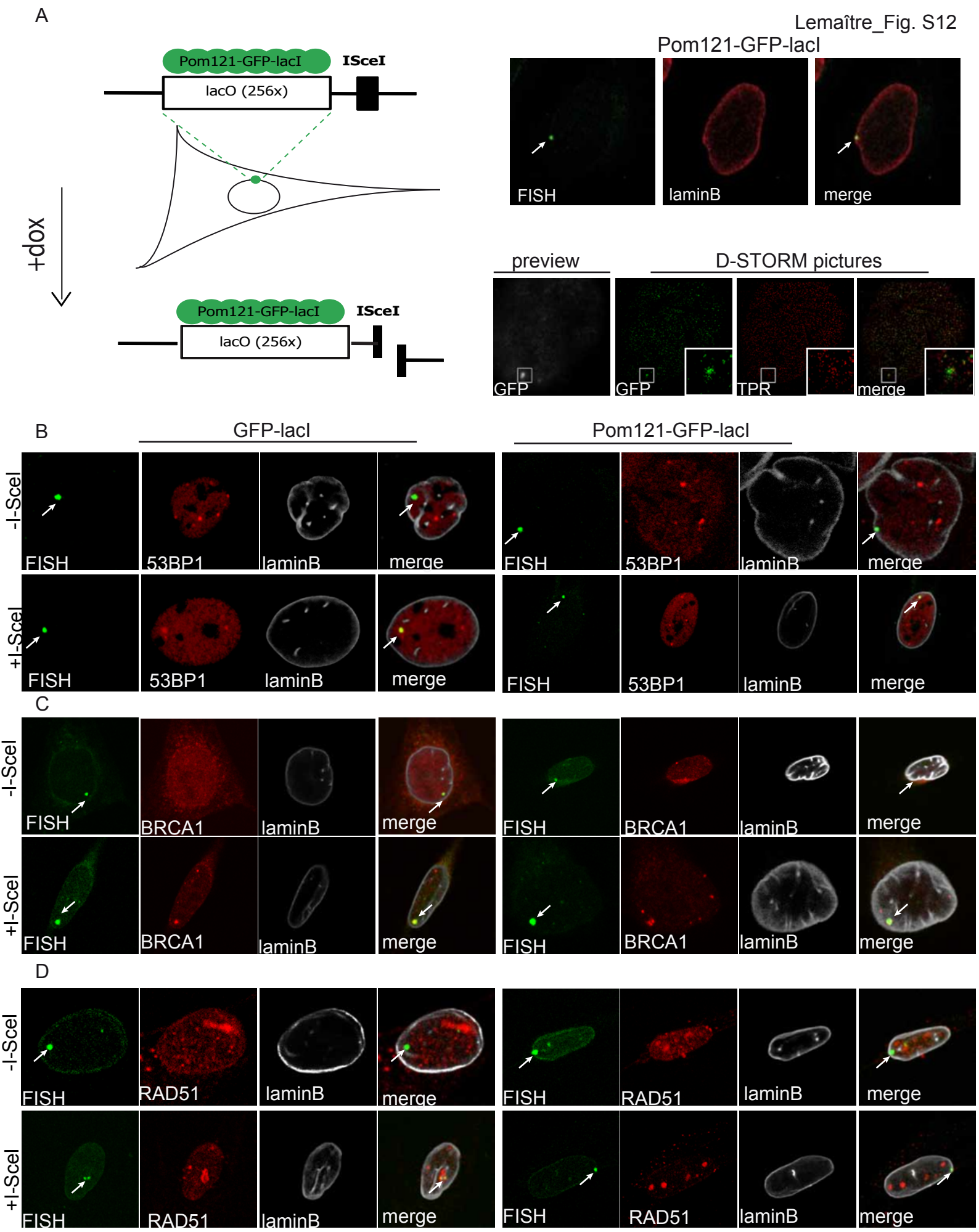
B

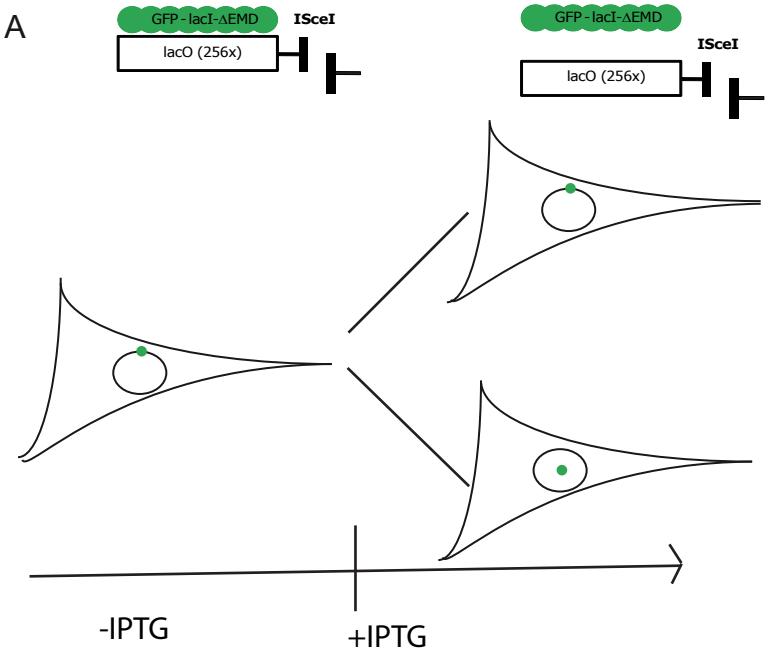


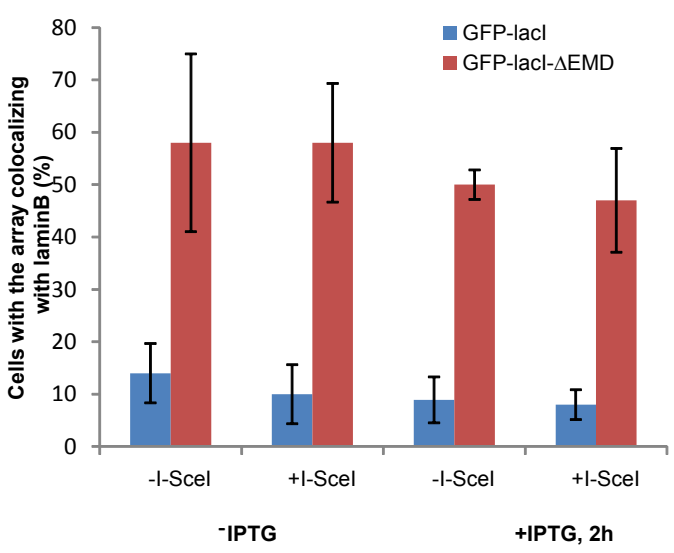
C

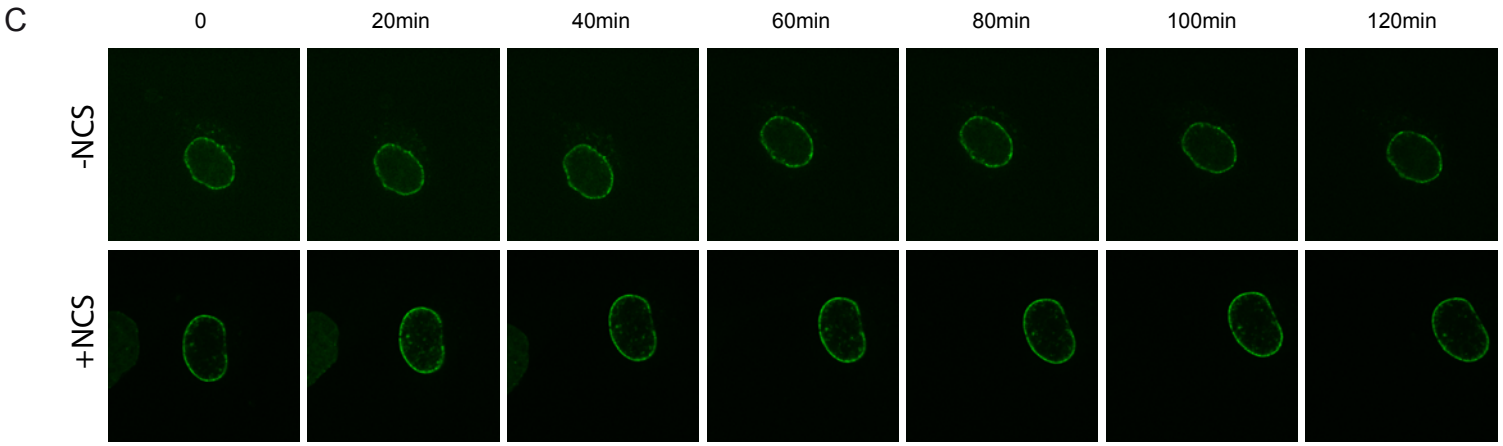


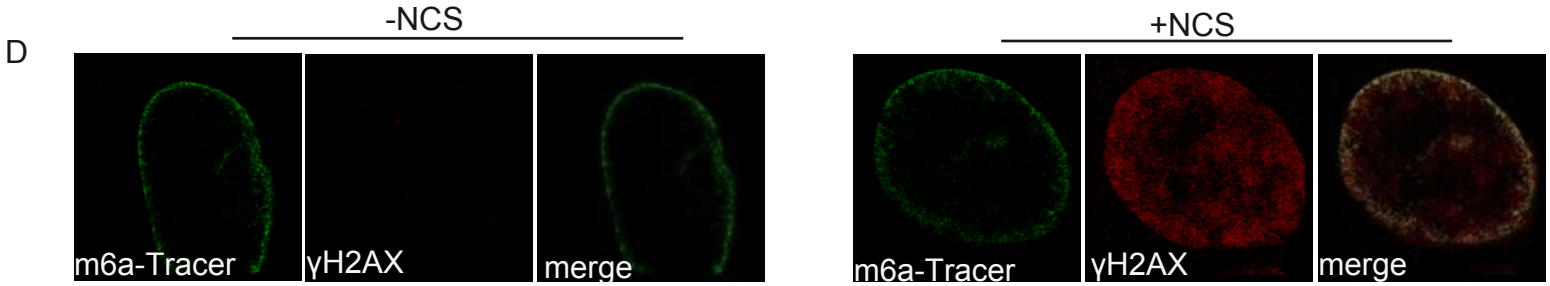

A
GFP-lacI
GFP-lacI-ΔEMD
cherry-lacI
-I-SceI
+I-SceI
BRG1-cherry-lacI
cherry
BRCA1
merge
B
RAD51


A
Pom121-GFP-lacI
ISceI
lacO (256x)
+dox
Pom121-GFP-lacI
ISceI
lacO (256x)
Pom121-GFP-lacI
FISH
laminB
merge
preview
D-STORM pictures
GFP
GFP
TPR
merge
B
GFP-lacI
Pom121-GFP-lacI
-I-SceI
+I-SceI
FISH
53BP1
laminB
merge
C
FISH
BRCA1
laminB
merge
D
FISH
RAD51
laminB
merge


A
GFP-lacI-ΔEMD
ISceI
lacO (256x)
GFP-lacI-ΔEMD
ISceI
lacO (256x)
-IPTG
+IPTG

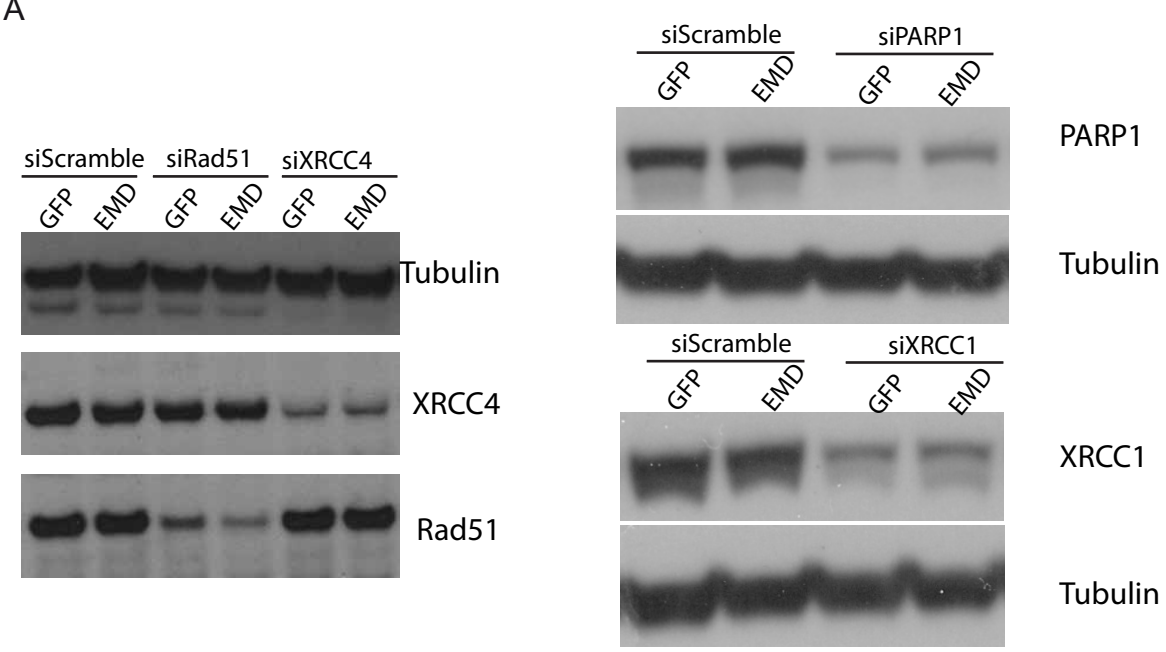

B
Cells with the array colocalizing with laminB (%)
80
70
60
50
40
30
20
10
0
GFP-lacI
GFP-lacI-ΔEMD
-I-SceI
+I-SceI
-I-SceI
+I-SceI
-IPTG
+IPTG, 2h


C
0
20min
40min
60min
80min
100min
120min
-NCS
+NCS

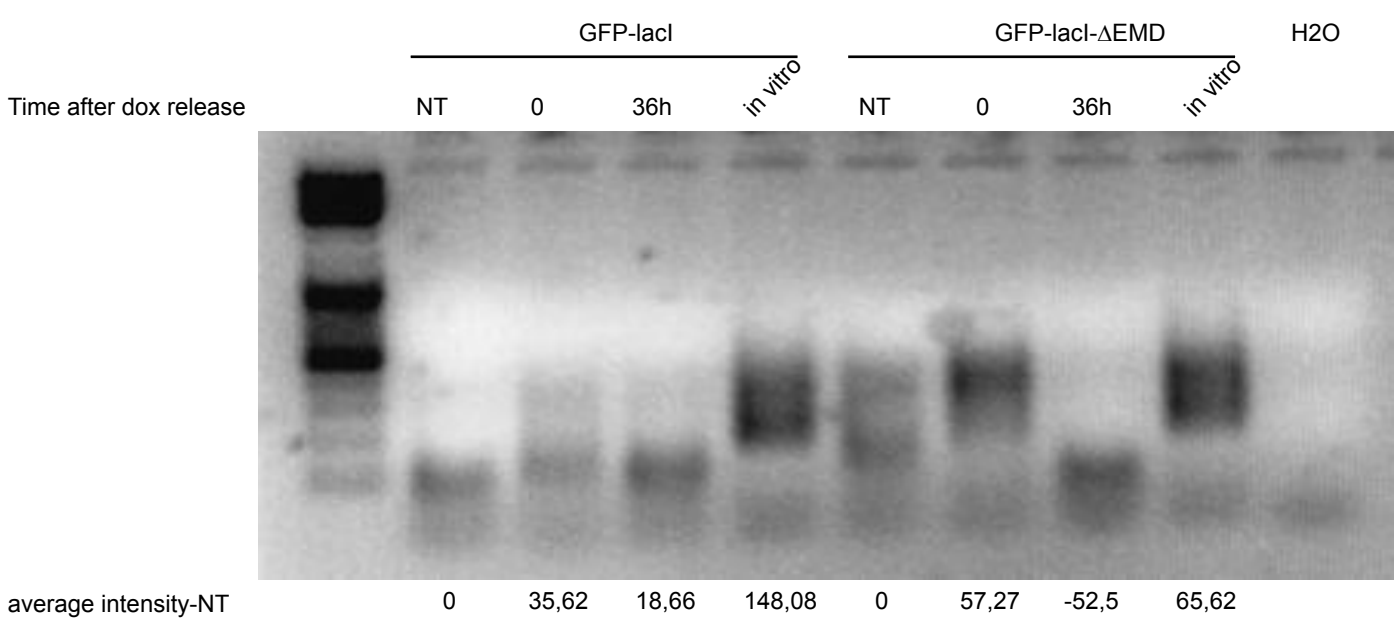

D
-NCS
+NCS
m6a-Tracer
γH2AX
merge
m6a-Tracer
γH2AX
merge

Lemaître_Fig. S14

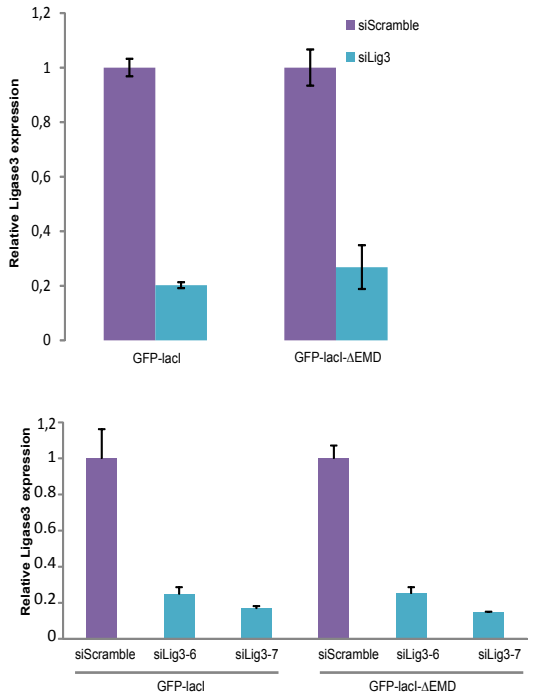
A



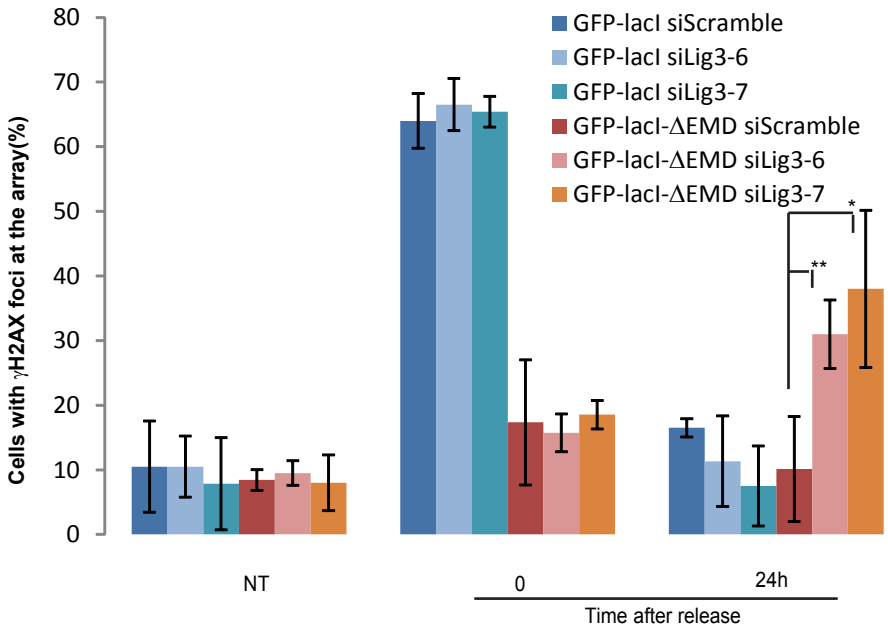
B



C



D



6.2. Review “The integrative role of cryo electron microscopy in molecular and cellular structural biology”

Biol. Cell (2017) **109,** 81–93 DOI: 10.1111/boc.201600042

Review

The integrative role of cryo electron microscopy in molecular and cellular structural biology

Igor Orlov*†‡§, Alexander G. Myasnikov*†‡§, Leonid Andronov*†‡§, S. Kundhavai Natchiar*†‡§, Heena Khatter*†‡§[2], Brice Beinsteiner*†‡§, Jean-François Ménétret*†‡§, Isabelle Hazemann*†‡§, Kareem Mohideen*†‡§, Karima Tazibt*†‡§, Rachel Tabaroni*†‡§, Hanna Kratzat*†‡§, Nadia Djabeur*†‡§, Tatiana Bruxelles*†‡§, Finaritra Raivoniaina*†‡§, Lorenza di Pompeo*†‡§, Morgan Torchy*†‡§, Isabelle Billas*†‡§, Alexandre Urzhumtsev*†‡§ and Bruno P. Klaholz*†‡§[1]

*Centre for Integrative Biology (CBI), Department of Integrated Structural Biology, IGBMC (Institute of Genetics and of Molecular and Cellular Biology), Illkirch, France, †Centre National de la Recherche Scientifique (CNRS) UMR 7104, Illkirch, France, ‡Institut National de la Santé et de la Recherche Médicale (INSERM) U964, Illkirch, France, and §Université de Strasbourg, Strasbourg, France

After gradually moving away from preparation methods prone to artefacts such as plastic embedding and negative staining for cell sections and single particles, the field of cryo electron microscopy (cryo-EM) is now heading off at unprecedented speed towards high-resolution analysis of biological objects of various sizes. This 'revolution in resolution' is happening largely thanks to new developments of new-generation cameras used for recording the images in the cryo electron microscope which have much increased sensitivity being based on complementary metal oxide semiconductor devices. Combined with advanced image processing and 3D reconstruction, the cryo-EM analysis of nucleoprotein complexes can provide unprecedented insights at molecular and atomic levels and address regulatory mechanisms in the cell. These advances reinforce the integrative role of cryo-EM in synergy with other methods such as X-ray crystallography, fluorescence imaging or focussed-ion beam milling as exemplified here by some recent studies from our laboratory on ribosomes, viruses, chromatin and nuclear receptors. Such multi-scale and multi-resolution approaches allow integrating molecular and cellular levels when applied to purified or *in situ* macromolecular complexes, thus illustrating the trend of the field towards cellular structural biology.

Introduction

The key event in cryo electron microscopy (cryo-EM) has been the introduction of cryo methods that allow preserving the biological sample in a hydrated and functional state (Dubochet et al., 1988). One of the reasons for the importance of cryo methods is that artefacts due to sample dehydration, fixation, adsorption and staining can be avoided, thus allowing the sample to be observed in a functionally relevant state. Sample preservation is thus an advantage of cryo methods that is worth considering both for single particle analysis of complexes extracted from the cellular context and the analysis of cell sections (*i.e.* in contrast to plastic embedding, fixation etc.). The second good reason to use cryo methods is that they allow performing high-resolution analysis, whereas conventional methods intrinsically limit the

[1]To whom correspondence should be addressed (email: klaholz@igbmc.fr)
[2]Present address: European Molecular Biology Laboratory (EMBL), Structural and Computational Biology Unit, Meyerhofstrasse 1, 69117 Heidelberg, Germany.

Key words: cryo electron microscopy, cryo electron tomography, Crystallography, Super-resolution microscopy, Structural biology.

Abbreviations: cryo-EM, cryo electron microscopy; cryo-ET, cryo electron tomography; CCD, charge-coupled device; CMOS, complementary metal oxide semiconductor; DDD, direct detection device; DQE, detective quantum efficiency; EcR, ecdysone receptor; FIB, focussed-ion beam; FRISBI, French Infrastructure for Integrated Structural Biology; Instruct, Integrated Structural Biology Infrastructure for Europe; kDa, kilo Dalton; MSA, multi-variate statistical analysis; mRNA, messenger RNA; rRNA, ribosomal RNA; SEM, scanning electron microscopy; SMLM, single-molecule localisation microscopy; tRNA, transfer RNA; 2D, two dimensional; 3D, three dimensional.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2016 The Authors. *Biology of the Cell* published by Wiley-VCH Verlag GmbH & Co. KGaA on behalf of Société Française des Microscopies and Société de Biologie Cellulaire de France.

attainable resolution. This second aspect was rather clear early on when cryo techniques were introduced in the field, as illustrated by the stunning amount of molecular structures visible by eye on cryo-EM images of various samples such as viruses, chromatin and ribosomes (Dubochet et al., 1988). However, it became obvious that the main limitation in seeing high-resolution details and being able to reconstruct them computationally in three dimensions (3D) was the ability to record appropriate images with good contrast and at the same time preserve the high-resolution information (*i.e.* avoid defocussing the microscope too much). For decades, the recording medium was photographic film (negatives, *e.g.* SO-163 from Kodak) which had the advantage of being able to record a large field of the specimen ($\sim 8 \times 10$ cm film support size) and at high resolution (grain size around 10 μm), and have a good detective quantum efficiency (DQE) as compared with charge-coupled devices (CCDs). Nevertheless, CCDs had the advantage of direct data acquisition, that is no need for digitising negatives on a scanning device, an aspect that facilitates automation of image processing. However, the strongest and latest breakthrough is the introduction of direct electron detectors (based on a complementary metal oxide semiconductor (CMOS) detector with a direct detection device (DDD) sensor) which have much higher sensitivity than film or CCD thanks to their direct measurement of electron events, requiring no amplification of the signal nor fibre or lens optics (*e.g.* CCDs operate through an electron to light conversion by a scintillator followed by coupled fibre optics). Latest-generation variants of these direct electron detectors comprise (i) back-thinning of the silicon chip to avoid multiple electron scattering events, (ii) counting-mode to measure individual electron events at high read-out speeds (tens to hundreds of frames per second), (iii) localisation of the electron impact position with sub-pixel precision (*i.e.* super-resolution mode). Specific characteristics of these cameras have been described (McMullan et al., 2009; Ruskin et al., 2013; McMullan et al., 2014; Kuijper et al., 2015; Spear et al., 2015) and include an overall high DQE across the entire frequency range, wherein the particularly increased amplitudes in the low frequency range provide a much improved image contrast. An additional feature is the high read-out speed that allows dose fractionation and movie processing (Brilot et al., 2012; Campbell et al., 2012; Li et al., 2013; Veesler et al., 2013; Allegretti et al., 2014; Scheres, 2014), that is beam-induced specimen drift can be motion-corrected by aligning a series of low-dose images taken on one given area of the specimen (Kunath et al., 1984); in addition, sub-frames with optimised dose can be selected for image processing (exposure filtering; Grant et al., 2015). It is these major technological developments of new-generation detectors that have recently introduced a 'revolution in resolution' in the cryo-EM field, analogous to the impact of the Pilatus & Eiger pixel detectors in the field of X-ray crystallography (Broennimann et al., 2006; Rajendran et al., 2011; Casanas et al., 2016). This has greatly contributed to a strong increase of the amount of structures determined by cryo-EM within the last few years (Figure 1).

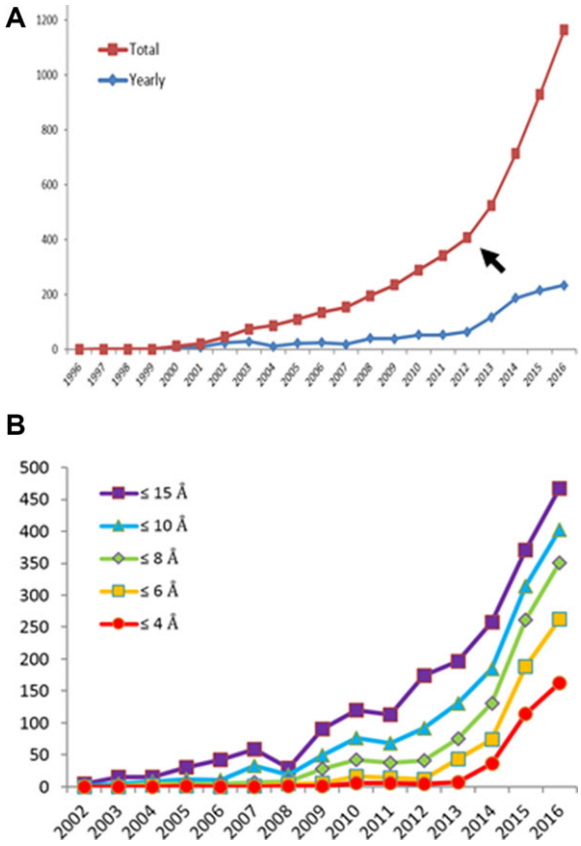
If the sample permits it, such cameras in principle allow – in synergy with advanced image processing informatics tools – to determine 3D structures in the 3 Å resolution range. Such resolution values allow building and refining atomic models against the cryo-EM maps. To illustrate this in more detail, we will now provide some examples of recent high-resolution cryo-EM analysis from projects of our laboratory (Figure 2) and we will discuss the integrative role of cryo-EM in synergy with other complementary methods. This reflects the general trend of the field towards cellular structural biology which relies more and more on multi-resolution and multi-scale approaches to address the molecular and cellular organisation of living cells (Figure 3).

Single particle cryo-EM analysis

The strength of current cryo-EM relies on the combination of cutting-edge instrumentation (high-resolution cryo electron microscopes equipped with direct electron detectors, such as those made available through the European and French infrastructures Instruct and FRISBI, see https://www.structuralbiology.eu/ and http://frisbi.eu/) and advanced image processing. Image processing and 3D reconstruction represent important components of modern cryo-EM because they allow getting unique and detailed 3D insights into the object of interest, much beyond a visual 2D description at low resolution. Images obtained on a transmission electron microscope are 2D projection images of the object and therefore contain all internal features of the 3D

Figure 1 | Recent increase of cryo-EM and cryo-ET studies as illustrated from the number of map depositions in the EMDB as a function of time and resolution

(**A, B**) Largely thanks to developments of high-sensitivity detectors, the year 2013 (arrow) marks a transition of the exponentially growing number of structures determined by cryo-EM or cryo-ET, that is the curve became bimodal with a steep increase in the last 4 years which is likely to continue considering the vast amount of biological objects that now become amendable to cryo-EM and cryo-ET analysis. This is also visible in the growing contribution of high-resolution cryo-EM structures over the four last years (B). The data are taken from the http://www.ebi.ac.uk/pdbe/emdb/ and http://www.rcsb.org/pdb/ websites (as of September 26, 2016).



object which is seen under different orientations. Thus, reversely, a 3D map of the object can be obtained from 2D projection images by back-projection, that is the views of the 3D object seen under different orientations (viewing angles) can be combined into a single 3D reconstruction; this concept applies to both single particle cryo-EM and cryo electron tomography (cryo-ET; for detailed basics on image processing and 3D reconstruction see for example reviews by van Heel et al., 2000; Ray et al., 2003; Briggs et al., 2013; Lučič et al., 2013; Milne et al., 2013; Scheres, 2014; Carazo et al., 2015; Carroni et al., 2016). An assumption in the 3D reconstruction process is that the input images describe the same object that it is functionally, structurally and conformationally homogenous. While this is the case for a single tomogram (recorded on one single object), it is rarely the case when averaging techniques are used, that is sub-tomogram averaging and single particle reconstruction using images describing (and tentatively merging) physically different objects. Sample heterogeneity can make the interpretation of 3D maps difficult or even impossible and might significantly limit the attainable resolution of the 3D reconstruction. Therefore, methods for 3D classifications to enable structure sorting are becoming an essential tool for the high-resolution analysis of single particle cryo-EM data, allowing the simultaneous analysis of several structures that are in equilibrium with each other. Examples of these are approaches based on cross correlation analysis using reference structures (template-based supervised classification, Gao et al., 2004) or based on multivariate statistical analysis (MSA) including local variance analysis in the particle images (White et al., 2004; Klaholz et al., 2004; Orlova & Saibil, 2010), resampling and bootstrapping methods to identify flexible regions in a macromolecular complex and perform 3D classifications (Penczek et al., 2006; Simonetti et al., 2008; Fischer et al., 2010; Klaholz, 2015; Liao et al., 2015), unsupervised classification (Fu et al., 2006) and maximum-likelihood (ML) based 3D classifications (Sigworth 1998; Scheres et al., 2010; Scheres et al., 2010; Lyumkis et al., 2013). The three categories of 2D/3D classification methods thus comprise (i) template-based methods which are intrinsically reference-biased, (ii) classification based on statistical analysis using MSA and bootstrapping methods and (iii) ML-based sorting; methods (ii) and (iii) are now commonly used as they turn out to be more robust during cryo-EM image processing of variable structures. These methods start from a low-resolution identification of larger conformational changes of the macromolecule of interest and then iteratively extend towards high-resolution sorting and

© 2016 The Authors. *Biology of the Cell* published by Wiley-VCH Verlag GmbH & Co. KGaA on behalf of Société Française des Microscopies and Société de Biologie Cellulaire de France.

Figure 2 | Examples of high-resolution cryo-EM and cryo-ET studies using direct electron detectors
(**A**) The introduction of CMOS-based direct electron detectors in the EM field has led to a 'revolution in resolution' thanks to their much increased sensitivity (as an example, here the Falcon camera from the FEI company; setup installed on the Polara and Titan Krios cryo electron microscopes at the Centre for Integrative Biology, IGBMC; available through infrastructure access at https://www.structuralbiology.eu/ and http://frisbi.eu/). (**B**) First high-resolution structure determination of the human 80S ribosome (Khatter et al., 2015). The map obtained from single particle cryo-EM image processing and 3D reconstruction (40S and 60S ribosomal subunits are labelled, the exit site tRNA is colored in red) allows visualising side-chains of amino acids and nucleic acids whose position needs to be known with precision to perform structure-based drug design (panels on the right; including a first human 80S complex with an antibiotic; Myasnikov et al., 2016). Two aspects can be highlighted there: (i) a resolution in the 3 Å range that is required to derive detailed atomic models can nowadays be obtained even for asymmetric objects, and (ii) cryo-EM has the potential to be used for structure-based drug design. (**C**) Comparison of cryo-ET reconstructions (after sub-tomogram averaging) obtained using either a CCD camera or a CMOS camera, illustrating the increased amount of structural features that can be visualised (poly-ribosome assembly, 60S and 40S ribosomal subunits are labelled in blue and yellow, respectively; Myasnikov et al., 2014). (**D**) Single particle cryo-EM reconstruction of the 1358 bacteriophage capsid (cross-section through the 3D reconstruction; Orlov et al., unpublished) in which secondary structure elements such as α-helices and residue details (side-chains; white arrow) can be visualised.

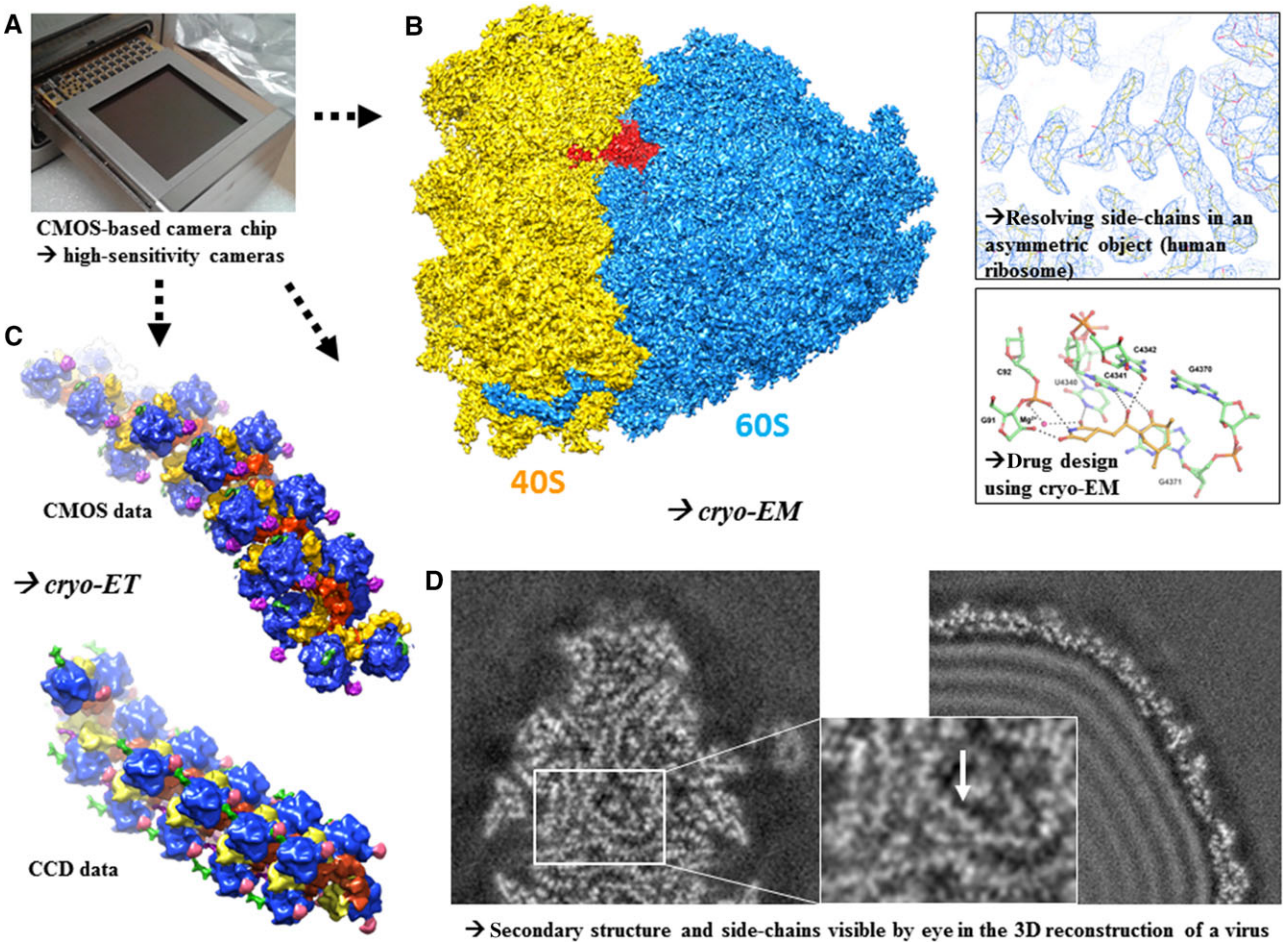
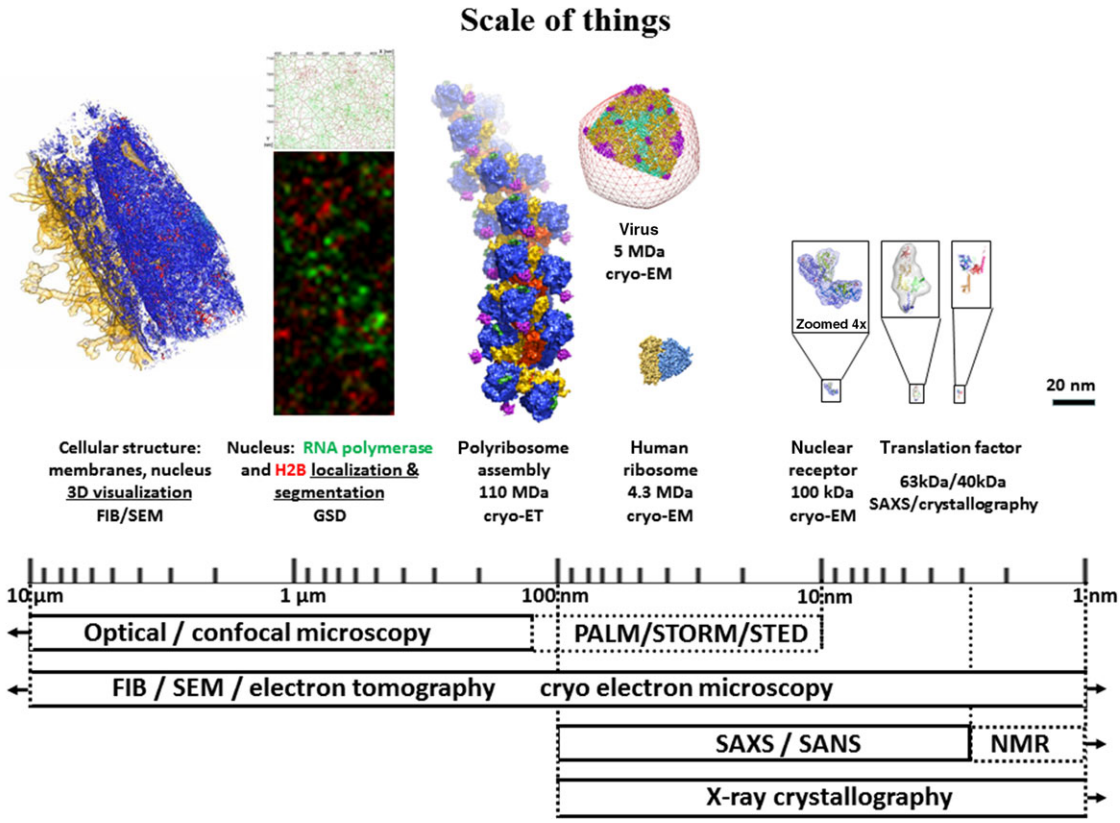


Figure 3 | Scale of things and the integrative role of cryo-EM in molecular and cellular structural biology Schematic representation of the multi-scale and multi-resolution approach in structural biology. Various examples of macromolecular complexes studied in our group are shown to illustrate the integrative role of cryo-EM/cryo-ET at the interface with complementary methods such as X-ray crystallography, NMR, SAXS (towards the molecular and atomic levels on the right) and FIB/SEM, super-resolution fluorescence microscopy (PALM, STORM/GSDIM, STED etc.) and optical/confocal microscopy (towards the cellular level on the left). Right to left: crystal structure of translation initiation factor IF2 and SAXS analysis of IF2 (Simonetti et al., 2013a), single particle cryo-EM structure of the 100 kDa nuclear receptor complex USP/EcR (Maletta et al., 2014); the structures are zoomed 4× for visualisation; middle: single particle cryo-EM analysis of the human 80S ribosome (Khatter et al., 2015) and the 1358 bacteriophage (Orlov et al., unpublished) and cryo-ET analysis of eukaryotic polyribosomes (Myasnikov et al., 2014); left: super-resolution imaging (GSDIM) of RNA polymerase (green) and histone H2B (red) and cluster analysis using new tools (Andronov et al., 2016a, 2016b); FIB/SEM analysis of HeLa cells (Orlov et al., 2015).



structure refinement (discussed in more detail in Klaholz, 2015).

Recent examples of high-resolution single particle cryo-EM analysis include work on various complexes such as viruses, ribosomes, proteasome, β-galactoside and so on (among many others: Chen et al., 2009; Zhang et al., 2013; Bai et al., 2013; Wong et al., 2014; Banerjee et al., 2016; Earl et al., 2016; Fischer et al., 2015; Khatter et al., 2015; Greber et al., 2015). Among the macromolecular complexes that we recently studied in our laboratory to high resolution (3 Å resolution range) using cryo-EM are the human ribosome and a virus (phage 1358); moreover, we also analysed a macromolecular assembly of poly-ribosomes by cryo-ET. These projects all benefited from the direct electron detector technology, high-resolution cryo electron microscopes and appropriate image processing (Figure 2).

Analysing the structure of the human ribosome can be considered to be the ultimate goal in structural studies of eukaryotic ribosomes, but it had to overcome the common *a priori* thinking that human

© 2016 The Authors. *Biology of the Cell* published by Wiley-VCH Verlag GmbH & Co. KGaA on behalf of Société Française des Microscopies and Société de Biologie Cellulaire de France.

ribosomes would be too inhomogenous and difficult to analyse. In order to address this issue, we developed biochemical protocols for the high-quality preparation of human ribosomes (Khatter et al., 2014) and their high-resolution structural analysis using advanced high-resolution cryo-EM and atomic model building (Figure 2; Khatter et al., 2015). The combination of data acquisition on a DDD camera, high-resolution image processing including 3D classifications and atomic model building and refinement using tools from the field of X-ray crystallography [PHENIX (Afonine et al., 2012), Chimera (Yang et al., 2012) and Coot (Emsley et al., 2010), among other available tools (*e.g.* Buster, REFMAC, CNS etc.; Smart et al., 2012; Brown et al., 2015; Adams et al., 2013; Natchiar et al., in prep.)] allowed determining the first atomic structure of the human ribosome. In this context, it is worth noting that the procedures for atomic model building and refinement into cryo-EM maps are evolving a lot recently, illustrating a strong synergy between crystallographic methods and cryo-EM (Brown et al., 2015; Barad et al., 2015; Sobolev et al., 2015; Natchiar et al., in prep.). The structural analysis of the human ribosome enabled full integration down to the atomic level, thus providing unprecedented insights into nucleotide and amino acid side-chains which can be analysed with respect to sequence conservation using a new tool for integrated sequence and 3D structure analysis of large macromolecular complexes (Beinsteiner et al., 2015). For example, it uncovered specific molecular interactions of the 28S rRNA with the universally conserved CCA end of the tRNA in the exit site. The structure opens the possibility of studying the molecular basis of interactions of drugs with the human ribosome, as well as functional complexes with mRNA, tRNA and protein factors. A follow-up study is the first structure of a ligand complex with the human ribosome that we recently determined, highlighting the potential of the human ribosome as a drug target for protein synthesis deregulations such as cancer (proliferating cells are dependent on elevated protein synthesis rates) and illustrating the capacity of high-resolution cryo-EM for structure-based drug design studies (Figure 2; Myasnikov et al., 2016). For comparison, our previous studies of bacterial translation initiation complexes were limited to a much lower resolution (8–10 Å; Simonetti et al., 2008; Simonetti et al., 2013a,2013b; Eiler et al., 2013; Simonetti et al., 2016) but nevertheless allowed addressing the localisation of initiation factors and tRNA, thus providing a solid basis for synergies with other methods including X-ray crystallography and solution SAXS.

Another example of a high-resolution single particle cryo-EM study is our ongoing structural analysis of the capsid of the lactococcal siphophage 1358 virion. Previously, negative staining EM and cryo-EM were used to determine the structure of the different parts of this tailed bacteriophage from the *Siphoviridae* family and to assemble an overall composite structure of the phage (Spinelli et al., 2014). The building units comprise the capsid, the connector, the tail and the base-plate. The structural analysis of the capsid using high-resolution cryo-EM now provides unprecedented insights in the 3–4 Å range (refinement ongoing; Orlov et al., unpublished). The quality of the structure is illustrated by the wealth of structural details that can be seen even by eye in the cross-section of the 3D reconstruction including secondary structure elements and side-chains. Indeed, zig-zag shapes in the densities reveal α-helices and protruding densities correspond to the amino acid side-chains (Figure 2, bottom right). This highlights the fact that a resolution has been reached that is sufficient for tracing the peptide backbone and positioning side-chains from the amino acid sequence, thus allowing an atomic model refinement against the cryo-EM map. Considering that phage 1358 infects *L. lactis* strains and has some similarity to *Listeria* phages, these structural investigations may be useful for medical and agricultural implications.

Cryo-ET analysis and sub-tomogram averaging

Cryo-ET is particularly useful to analyse unique cellular structures. Rather than obtaining a 3D reconstruction from different orientational views of physically different objects like in single particle cryo-EM, it is obtained from images acquired while rotating the sample with predefined angles. The images of a tilt series can then be aligned to reconstruct a tomogram (see for example Briggs, 2013; Lučič et al., 2013; Dubrovsky et al., 2015; Asano et al., 2016). If common sub-structures exist these can be extracted as sub-tomograms, aligned and averaged to form sub-tomogram averages with an improved signal-to-noise ratio (Wan et al., 2016; Galaz-Montoya et al., 2016).

Such approaches have been used in the past to analyse, for example, the cellular structure of the nuclear pore complex (NPC), of membrane-bound ribosomes, of actin filaments in combination with segmentation tools and template-based identification of the components (Frangakis et al., 2002; Medalia et al., 2002; Beck et al., 2004; Ortiz et al., 2006). Recently, stunning insights into cellular sub-structures have been obtained (Hagen et al., 2015; Nans et al., 2015; Chang et al., 2016; Irobalieva et al., 2016; Kosinski et al., 2016; Lin et al., 2016; Mahamid et al., 2016), achieving in a first case even side-chain resolution (Schur et al., 2016) using optimised cryo-ET, sub-tomogram averaging and a new dose-symmetric tilt acquisition scheme that preserves high-resolution data more isotropically (Hagen et al., 2016). We used cryo-ET in combination with sub-tomogram averaging and molecular modelling to address the supramolecular organisation of eukaryotic polyribosomes which can form large macromolecular assemblies (Brandt et al., 2010; Myasnikov et al., 2013; Afonina et al., 2014; Myasnikov et al., 2014; Afonina et al., 2015). This allowed deriving the 3D structure of one of the largest asymmetric complexes to date (~100 MDa, comprising over 20 ribosomes on the same mRNA molecule; Myasnikov et al., 2014). The structure allowed the visualisation of the three functional parts of the polysome assembly: the central core region that forms a rather compact left-handed supramolecular helix and the more open regions that harbour the initiation and termination sites at either ends. The helical region forms a continuous mRNA channel where the mRNA strand bridges neighbouring exit and entry sites of the ribosomes and prevents the mRNA from looping between ribosomes. This structure provides unprecedented insights into protein- and RNA-mediated inter-ribosome contacts that involve conserved sites through 40S ribosomal subunits and long protruding RNA expansion segments. These findings shed new light on the molecular machinery of the ribosome and its mode of action in the cellular context. The impact of direct electron detectors is illustrated by the improved resolution of the tomograms (Figure 2). In the future, contrast-increasing phase plates may be helpful to address more of the molecular details within polysomes, possibly in combination with double tilt cryo-ET which provides 3D reconstructions with less reconstruction artefacts (Myasnikov et al., 2013; one of the first examples of single particle cryo-ET together with some other examples: Dudkina et al., 2010; Murata et al., 2010; Wang et al., 2011).

Integrative role of cryo-EM and current trends in complementary approaches

Another aspect of cryo-EM is its integrative role in multi-scale and multi-resolution approaches that integrate molecular and cellular levels by combining various structural methods such as X-ray crystallography, fluorescence imaging or focussed-ion beam milling as exemplified here by some recent studies from our laboratory on ribosomes, viruses, chromatin and nuclear receptors (Figure 3). An important consideration is that challenging objects require the right choice of the method (Ménétret et al., 2013), each having intrinsic limitations, but when combined these methods can create synergies at the interfaces that overcome the individual limits. For example, in X-ray crystallography a typical bottleneck is crystallisation, but it is not much restricted to a particular object size as illustrated by the possibility of crystallising small organic compounds and large complexes such as ribosomes and viruses (including very large viruses with a diameter in the 1000 Å range). However, in cryo-EM, a typical limitation is the lower size of a complex because the lower the molecular weight of the macromolecule the smaller the image contrast is, which limits the possibility of accurately determining the structure. Nevertheless, there are clear trends in the cryo-EM field to move towards smaller complexes. An example of this is our recent study of a 100 kDa complex, the first structure of a full-length nuclear receptor bound to an inverted DNA repeat, the USP/EcR complex in insects where EcR is the ecdysone receptor (Maletta et al., 2014); it revealed an asymmetric organisation of the complex although the DNA is almost symmetric (Figure 3); note that the data were collected at a reduced acceleration voltage (100 kV) to increase the image contrast. This represents a technical advance in the field with respect of studying relatively small complexes by cryo-EM, thus opening the possibility to study many other biological complexes and drug targets previously believed to be too small for cryo-EM. Other examples in the field are the studies of a 50 kDa RNA (Baird et al., 2010) and most recently the 3.8 Å resolution structure of a 93 kDa protein (Merk et al., 2016),

© 2016 The Authors. *Biology of the Cell* published by Wiley-VCH Verlag GmbH & Co. KGaA on behalf of Société Française des Microscopies and Société de Biologie Cellulaire de France.

highlighting the fact that high-resolution structures can also be obtained on relatively small complexes, provided the biological samples behave well and can be imaged well (*i.e.* low aggregation, stability in relatively low-salt or low-detergent concentrations, good particle distribution etc.).

Among various structural biology methods that use averaging techniques, a common limitation is the sample homogeneity. While crystallisation (and re-crystallisation) can contribute to the high-level purification of a chemical compound or even a large macromolecular complex (*e.g.* PEG10000 precipitation of yeast and human ribosomes; Ben-Shem et al., 2010; Khatter et al., 2014; Khatter et al., 2015), sample homogeneity often limits crystallisation. However, the problem of sample heterogeneity can be turned into an advantage when using single-particle cryo-EM considering that with advanced image processing and 3D classification tools (see section Single particle cryo-EM analysis) different structures can be sorted and separated into distinct sub-populations, allowing to improve the refinement of a cryo-EM structure by local masking and refinement to high resolution and at the same time describe several conformational states of a complex that are in equilibrium with each other. Similarly, when using cryo-ET and sub-tomogram averaging of similar structures observed *in situ* in the cell (*e.g.* polyribosomes in Figure 2), it is possible to classify the aligned tomograms and address several structures at the same time (Heumann et al., 2011; Stölken et al., 2011; Frank et al., 2012; Kuybeda et al., 2013; Xu et al., 2013; Chen et al., 2014; Castaño-Díez et al., 2016; Obbineni et al., 2016). However, sorting is not possible on unique structures such as large regions of the cell, in which case methods such as focussed-ion beam/scanning electron microscopy (FIB/SEM) are more appropriate (*e.g.* chromatin in Figure 3; Orlov et al., 2015). FIB is a method originally emanating from material sciences, which, when applied to biological specimens, appears to be one of the most promising structural biology methods for future cell biology studies (Villa et al., 2013; Kizilyaprak et al., 2014; Rigort et al., 2015). In contrast to ultramicrotome sectioning of cells, which can exhibit knife cutting and compression artefacts, FIB milling allows obtaining high-quality cuts through the cell that can be observed by SEM (section by section to reconstitute a full tomogram, in the literal sense of τομός *tomos* (Greek: cut). Alternatively, thin lamella can be prepared by FIB milling and observed by cryo-ET (*e.g.* Mahamid et al., 2015; Schaffer et al., 2016; Zhang et al., 2016). Such analyses have the great advantage of being *in situ* (Lučič et al., 2013) and thus in the functional cellular context, but they require the identification of the complexes of interest by complementary methods such as fluorescence labelling. In an effort to allow the identification of chromatin structures, we have explored the possibilities of molecular imaging in cells through single-molecule localisation microscopy (SMLM; *e.g.* dSTORM, PALM; we used GSDIM, ground state depletion microscopy followed by individual molecule return). To make best use of super-resolution data, we developed an integrated software pipeline for image reconstruction, drift and chromatic aberration correction, co-localisation, resolution estimation, segmentation, clustering and quantification of labelled complexes (SharpViSu & ClusterViSu; Andronov et al., 2016a, 2016b). In addition to be very useful in the field of super-resolution microscopy, in the future such informatics tools may become particularly useful in the context of correlative light and electron microscopy (CLEM) approaches (Koning et al., 2014; Schirra et al., 2014; Arnold et al., 2016; Karreman et al., 2016; Schorb et al., 2016) to address the cellular fine structure, identify and localise protein complexes and visualise them at high resolution using cryo-EM and cryo-ET, that is perform an identification of the molecule-of-interest (MOI) rather than only the overall region-of-interest (ROI). Examples including SMLM (*e.g.* Kim et al., 2015) are studies of complexes such as RNA polymerase, nucleosomes, viruses or NPCs (Szymborska et al., 2013; Löschberger et al., 2014; Laine et al., 2015) some of which were reconstructed from super-resolution data using cryo-EM 3D reconstruction methods (Szymborska et al., 2013; Andronov et al., in prep.).

Current and near-future challenges in cryo-EM developments comprise key questions such as (i) how to push the resolution to the atomic level, (ii) how to analyse flexible complexes and (iii) how to integrate towards the cellular level. Two important areas to address these are (i) the instrumentation and (ii) software developments. Instrumentation (i) comprises for example high-resolution electron microscopes, direct electron detector cameras with higher DQE at high frequencies, energy filters to

remove inelastically scattered electrons and reduce the background in the images, correction of the spherical aberration (Cs) to improve the optical system in the column of the microscope, microelectron diffraction to determine the structure from small 3D crystals (Nederlof et al., 2013; Sawaya et al., 2016; Shi et al., 2016), spraying of small amounts on cryo-EM grids (Chen et al., 2015; Razinkov et al., 2016), cryo transfer between FIB and electron microscope for cryo-ET (Schaffer et al., 2016) and contrast-increasing phase plates placed in the back-focal plane of the microscope (Danev et al., 2014; Dai et al., 2014; Frindt et al., 2014; Walter et al., 2015; Chua et al., 2016; Danev et al., 2016; Glaeser, 2016; Rhinow 2016; Khoshouei et al., 2016a); phase plates could facilitate structure determination of relatively small complexes (Khoshouei et al., 2016b), which were difficult to address previously and usually limited to lower resolution (Baird et al., 2010; Orlov et al., 2012; Maletta et al., 2014) even though DDDs have helped a lot moving forward (Merk et al., 2016), suggesting that synergies will appear for example between phase plates, energy filters and high-sensitivity cameras to enable high-contrast high-resolution image acquisition (a feature that usually contradicts itself considering the requirement of defocussing during data collection to get some reasonable amount of image contrast required for image processing). Further software developments (ii) will be required for automatic data acquisition for massive data collection for single particle cryo-EM and cryo-ET (*e.g.* reviewed in Tan et al., 2016; automatic beam alignments including for Cs-corrected microscopes; remote control), and on-the-fly image processing during data collection, particle sorting and 3D classifications to address sample heterogeneity, automation in cryo-EM/cryo-ET structure determination, automation of backbone tracing like in X-ray crystallography, atomic model building and refinement into cryo-EM/cryo-ET maps of large macromolecular complexes to move towards larger cellular structures. Taken together, the scientific community is currently experiencing a very exciting era that is moving more and more towards the integration of various structural and imaging methods among which cryo-EM, cryo-ET and FIB will play a key role in multi-resolution integration. In close synergy with functional studies, this evolution towards cellular structural biology will lead in the coming years to unprecedented insights into cellular function and drug targets, including the analysis of the dynamic changes of macromolecular complexes that reflect their functional transitions in the cellular environment.

Funding

This work was supported by the European Research Council (ERC Starting Grant N_243296 TRANSLATIONMACHINERY), the Fondation pour la Recherche Médicale (FRM), the Association pour la Recherche sur le Cancer (ARC), the Alsace Region, the Agence National pour la Recherche, Université de Strasbourg and Investissement d'Avenir (IDEX), the Institut National du Cancer (INCa), the Centre National pour la Recherche Scientifique (CNRS), the Institut National pour la Recherche Médicale (Inserm), by the French Infrastructure for Integrated Structural Biology (FRISBI) ANR-10-INSB-05-01, and Instruct as part of the European Strategy Forum on Research Infrastructures (ESFRI).

Acknowledgements

We apologise to our colleagues for any publications not being cited which is due to space limitations and the cross-disciplinary nature of this review article. We would like to thank our colleagues with whom we collaborate on the projects mentioned in this review, notably Jean-Luc Vonesch, Alexander Spirin, Vladimir Shirokov, Dino Moras, Jean-François Peyron, Danièle Spehner and Christian Cambillau, and the members of our platforms and facilities for their support.

Conflict of interest statement

The authors have declared no conflict of interest.

References

Adams, P.D., Baker, D., Brunger, A.T., Das, R., DiMaio, F., Read, R.J., Richardson, D.C., Richardson, J.S. and Terwilliger, T.C. (2013) Advances, interactions, and future developments in the CNS, Phenix, and Rosetta structural biology software systems. Annu. Rev. Biophys. **42**, 265–287

Afonina, Z.A., Myasnikov, A.G., Shirokov, V.A., Klaholz, B.P. and Spirin, A.S. (2015) Conformation transitions of eukaryotic polyribosomes during multi-round translation. Nucleic Acids Res. **43**, 618–628

Afonina, Z.A., Myasnikov, A.G., Shirokov, V.A., Klaholz, B.P. and Spirin, A.S. (2014) Formation of circular polyribosomes on

© 2016 The Authors. *Biology of the Cell* published by Wiley-VCH Verlag GmbH & Co. KGaA on behalf of Société Française des Microscopies and Société de Biologie Cellulaire de France.

eukaryotic mRNA without cap-structure and poly(A)-tail: a cryo electron tomography study. Nucleic Acids Res. **42**, 9461–9469

Afonine, P.V., Grosse-Kunstleve, R.W., Echols, N., Headd, J.J., Moriarty, N.W., Mustyakimov, M., Terwilliger, T.C., Urzhumtsev, A., Zwart, P.H. and Adams, P.D. (2012) Toward automated crystallographic structure refinement with phenix.refine. Acta Cryst. **D68**, 352–367

Allegretti, M., Mills, D.J., McMullan, G., Kühlbrandt, W. and Vonck, J. (2014) Atomic model of the F420-reducing [NiFe] hydrogenase by electron cryo-microscopy using a direct electron detector. Elife **3**:e01963. doi: 10.7554/eLife.01963.

Andronov, L., Lutz, Y., Vonesch, J.-L. and Klaholz, B.P. (2016a) SharpViSu: integrated analysis and segmentation of super-resolution microscopy data. Bioinformatics **32**, 2239–2241

Andronov, L., Orlov, I., Lutz, Y., Vonesch, J.-L. and Klaholz, B.P. (2016b) ClusterViSu, a method for clustering of protein complexes by Voronoi tessellation in super-resolution microscopy. Sci. Rep. **6**, 24084

Arnold, J., Mahamid, J., Lucic, V., de Marco, A., Fernandez, J.J., Laugks, T., Mayer, T., Hyman, A.A., Baumeister, W. and Plitzko, J.M. (2016) Site-Specific Cryo-focused Ion Beam Sample Preparation Guided by 3D Correlative Microscopy. Biophys. J. **110**, 860–869

Asano, S., Engel, B.D. and Baumeister, W. (2016) In Situ Cryo-Electron Tomography: A Post-Reductionist Approach to Structural Biology. J. Mol. Biol. **428**, 332–343

Bai, X.C., Fernandez, I.S., McMullan, G. and Scheres, S.H. (2013) Ribosome structures to near-atomic resolution from thirty thousand cryo-EM particles. eLife **2**:e00461. doi: 10.7554/eLife.00461.

Baird, N.J., Ludtke, S.J., Khant, H., Chiu, W., Pan, T. and Sosnick, T.R. (2010) Discrete structure of an RNA folding intermediate revealed by cryo-electron microscopy. J. Am. Chem. Soc. **132**, 16352–163523

Banerjee, S., Bartesaghi, A., Merk, A., Rao, P., Bulfer, S.L., Yan, Y., Green, N., Mroczkowski, B., Neitz, R.J., Wipf, P., Falconieri, V., Deshaies, R.J., Milne, J.L., Huryn, D., Arkin, M. and Subramaniam, S. (2016) 2.3 Å resolution cryo-EM structure of human p97 and mechanism of allosteric inhibition. Science **351**, 871–875

Barad, B.A., Echols, N., Wang, R.Y., Cheng, Y., DiMaio, F., Adams, P.D. and Fraser, J.S. (2015) EMRinger: side chain-directed model and **map** validation for 3D **cryo**-electron microscopy. Nat. Methods **12**, 943–946.

Beck, M., Förster, F., Ecke, M., Plitzko, J.M., Melchior, F., Gerisch, G., Baumeister, W. and Medalia, O. (2004) Nuclear pore complex structure and dynamics revealed by cryoelectron tomography. Science **306**, 1387–1390

Beinsteiner, B., Michalon, J. and Klaholz B. P. (2015) IBiSS, a versatile and interactive tool for integrated sequence and 3D structure analysis of large macromolecular complexes. Bioinformatics **31**, 3339–3344

Ben-Shem, A., Jenner, L., Yusupova, G. and Yusupov, M. (2010) Crystal structure of the eukaryotic ribosome. Science **330**, 1203–1209

Brandt, F., Carlson, L.A., Hartl, F.U., Baumeister, W. and Grünewald, K. (2010) The three-dimensional organization of polyribosomes in intact human cells. Mol. Cell **39**, 560–569

Briggs, J.A. (2013) Structural biology in situ - the potential of subtomogram averaging. Curr. Opin. Struct. Biol. **23**, 261–267

Brilot, A.F., Chen, J.Z., Cheng, A., Pan, J., Harrison, S.C., Potter, C.S., Carragher, B., Henderson, R. and Grigorieff, N. (2012) Beam-induced motion of vitrified specimen on holey carbon film. J. Struct. Biol. **177**, 630–637

Broennimann, C., Eikenberry, E.F., Henrich, B., Horisberger, R., Huelsen, G., Pohl, E., Schmitt, B., Schulze-Briese, C., Suzuki, M., Tomizaki, T., Toyokawa, H. and Wagner, A. (2006) The PILATUS 1M detector. J. Synchrotron Radiat. **13**, 120–130

Brown, A., Long, F., Nicholls, R.A., Toots, J., Emsley, P. and Murshudov, G. (2015) Tools for macromolecular model building and refinement into electron cryo-microscopy reconstructions. Acta Crystallogr. D Biol. Crystallogr. **71**, 136–153

Campbell, M.G., Cheng, A., Brilot, A.F., Moeller, A., Lyumkis, D., Veesler, D., Pan, J., Harrison, S.C., Potter, C.S., Carragher, B. and Grigorieff, N. (2012) Movies of ice-embedded particles enhance resolution in electron cryo-microscopy. Structure **20**, 1823–1828

Carazo, J.M., Sorzano, C.O., Otón, J., Marabini, R. and Vargas, J. (2015) Three-dimensional reconstruction methods in Single Particle Analysis from transmission electron microscopy data. Arch. Biochem. Biophys. **581**, 39–48

Carroni, M. and Saibil, H.R. (2016) Cryo electron microscopy to determine the structure of macromolecular complexes. Methods **95**, 78–85

Casanas, A., Warshamanage, R., Finke, A.D., Panepucci, E., Olieric, V., Nöll, A., Tampé, R., Brandstetter, S, Förster, A., Mueller, M., Schulze-Briese, C., Bunk, O. and Wang, M. (2016) EIGER detector: application in macromolecular crystallography. Acta Crystallogr. D Struct. Biol. **72**, 1036–1048

Castaño-Díez, D., Kudryashev, M. and Stahlberg, H. (2016) Dynamo Catalogue: Geometrical tools and data management for particle picking in subtomogram averaging of cryo-electron tomograms. J. Struct. Biol. **16**, 30111–30113

Chang, Y.W., Rettberg, L.A., Treuner-Lange, A., Iwasa, J., Søgaard-Andersen, L. and Jensen, G.J. (2016) Architecture of the type IVa pilus machine. Science **351**, 6278:aad2001

Chen, B., Kaledhonkar, S., Sun, M., Shen, B., Lu, Z., Barnard, D., Lu, T.M., Gonzalez, R.L. Jr. and Frank, J. (2015) Structural dynamics of ribosome subunit association studied by mixing-spraying time-resolved cryogenic electron microscopy. Structure **23**, 1097–1105

Chen, J.Z., Settembre, E.C., Aoki, S.T., Zhang, X., Bellamy, A.R., Dormitzer, P.R., Harrison, S.C. and Grigorieff, N. (2009) Molecular interactions in rotavirus assembly and uncoating seen by high-resolution cryo-EM. Proc. Natl. Acad. Sci. U.S.A. **106**, 10644–10648

Chen, Y., Pfeffer, S., Fernández, J.J., Sorzano, C.O. and Förster, F. (2014) Autofocused 3D classification of cryoelectron subtomograms. Structure **22**, 1528–1537

Chua, E.Y., Vogirala, V.K., Inian, O., Wong, A.S., Nordenskiöld, L., Plitzko, J.M., Danev, R. and Sandin, S. (2016) 3.9 Å structure of the nucleosome core particle determined by phase-plate cryo-EM. Nucleic Acids Res. **44**, 8013–8019

Dai, W., Fu, C., Khant, H.A., Ludtke, S.J., Schmid, M.F. and Chiu, W. (2014) Zernike phase-contrast electron cryotomography applied to marine cyanobacteria infected with cyanophages. Nat. Protoc. **9**, 2630–2642

Danev, R. and Baumeister, W. (2016) Cryo-EM single particle analysis with the Volta phase plate. Elife **5**: pii: e13046

Danev, R., Buijsse, B., Khoshouei, M., Plitzko, J.M. and Baumeister, W. (2014) Volta potential phase plate for in-focus phase contrast transmission electron microscopy. Proc. Natl Acad. Sci. U.S.A. **44**, 15635–15640

Dubochet, J., Adrian, M., Chang, J.J., Homo, J.C., Lepault, J., McDowall, A.W. and Schultz, P. (1988) Cryo-electron microscopy of vitrified specimens. Q. Rev. Biophys. **21**, 129–228

Dubrovsky, A., Sorrentino, S., Harapin, J., Sapra, K,T. and Medalia, O. (2015) Developments in cryo-electron tomography for in situ structural analysis. Arch. Biochem. Biophys. **581**, 78–85

Dudkina, N.V., Kouril, R., Bultema, J.B. and Boekema, E.J. (2010) Imaging of organelles by electron microscopy reveals protein-protein interactions in mitochondria and chloroplasts. FEBS Lett. **584**, 2510–2515

Earl, L.A. and Subramaniam, S. (2016) Cryo-EM of viruses and vaccine design. Proc. Nat. Acad. Sci. U.S.A. **113**, 8903–8905

Eiler, D., Lin, J., Simonetti, A., Klaholz, B.P. and Steitz, T.A. (2013) IF2 Initiation factor 2 crystal structure reveals a different domain organization from eukaryotic initiation factor 5B and mechanism among translational GTPases. Proc. Nat. Acad. Sci. U.S.A. **110**, 15662–15667

Emsley, P., Lohkamp, B., Scott, W.G. and Cowtan, K. (2010) Features and development of Coot. Acta Crystallogr. D Biol. Crystallogr. **66**, 486–501

Fischer, N., Konevega, A.L., Wintermeyer, W., Rodnina, M.V. and Stark, H. (2010). Ribosome dynamics and tRNA movement by time-resolved electron cryomicroscopy. Nature **466**, 329–333

Fischer, N., Neumann, P., Konevega, A.L., Bock, L.V., Ficner, R., Rodnina, M.V. and Stark, H. (2015) Structure of the E. coli ribosome-EF-Tu complex at <3 Å resolution by Cs-corrected cryo-EM. Nature **520**, 567–570

Frangakis, A.S., Böhm, J., Förster, F., Nickell, S., Nicastro, D., Typke, D., Hegerl, R. and Baumeister, W. (2002) Identification of macromolecular complexes in cryoelectron tomograms of phantom cells. Proc. Natl. Acad. Sci. U.S.A. **99**, 14153–14158

Frank, G.A., Bartesaghi, A., Kuybeda, O., Borgnia, M.J., White, T.A., Sapiro, G. and Subramaniam, S. (2012) Computational separation of conformational heterogeneity using cryo-electron tomography and 3D sub-volume averaging. J. Struct. Biol. **178**, 165–176

Frindt, N., Oster, M., Hettler, S., Gamm, B., Dieterle, L., Kowalsky, W., Gerthsen, D. and Schröder, R.R. (2014) In-focus electrostatic Zach phase plate imaging for transmission electron microscopy with tunable phase contrast of frozen hydrated biological samples. Microsc. Microanal. **1**, 175–183

Fu, J., Gao, H. and Frank, J. (2006) Unsupervised classification of single particles by cluster tracking in multi-dimensional space. J. Struct. Biol. **157**, 226–239

Galaz-Montoya, J.G., Hecksel, C.W., Baldwin, P.R., Wang, E., Weaver, S.C., Schmid, M.F., Ludtke, S.J. and Chiu, W. (2016) Alignment algorithms and per-particle CTF correction for single particle cryo-electron tomography. J. Struct. Biol. **194**, 383–394

Gao, H., Valle, M., Ehrenberg, M. and Frank, J. (2004) Dynamics of EF-G interaction with the ribosome explored by classification of a heterogeneous cryo-EM dataset. J. Struct. Biol. **147**, 283–290

Glaeser, R.M. (2016) Protein complexes in focus. Elife **5**: pii: e16156

Grant, T. and Grigorieff, N. (2015) Measuring the optimal exposure for single particle cryo-EM using a 2.6 Å reconstruction of rotavirus VP6. Elife **4**:e06980

Greber, B.J., Bieri, P., Leibundgut, M., Leitner, A., Aebersold, R., Boehringer, D. and Ban, N. (2015) The complete structure of the 55S mammalian mitochondrial ribosome. Science **348**, 303–308

Hagen, C., Dent, K.C., Zeev-Ben-Mordehai, T., Grange, M., Bosse, J.B., Whittle, C., Klupp, B.G., Siebert, C.A., Vasishtan, D., Bäuerlein, F.J., Cheleski, J., Werner, S., Guttmann, P., Rehbein, S., Henzler, K., Demmerle, J., Adler, B., Koszinowski, U., Schermelleh, L., Schneider, G., Enquist, L.W., Plitzko, J.M., Mettenleiter, T.C. and Grünewald, K. (2015) Structural Basis of Vesicle Formation at the Inner Nuclear Membrane. Cell **163**, 1692–1701

Hagen, W.J., Wan, W. and Briggs, J.A. (2016) Implementation of a cryo-electron tomography tilt-scheme optimized for high resolution subtomogram averaging. J. Struct. Biol. **16**, 30113–30117

Heumann, J.M., Hoenger, A. and Mastronarde, D.N. (2011) Clustering and variance maps for cryo-electron tomography using wedge-masked differences. J. Struct. Biol. **175**, 288–299

Irobalieva, R.N., Martins, B. and Medalia, O. (2016) Cellular structural biology as revealed by cryo-electron tomography. J. Cell Sci. **129**, 469–476

Karreman, M.A., Mercier, L., Schieber, N.L., Solecki, G., Allio, G., Winkler, F., Ruthensteiner, B., Goetz, J.G. and Schwab, Y. (2016) Fast and precise targeting of single tumor cells in vivo by multimodal correlative microscopy. J. Cell Sci. **129**, 444–456

Khatter, H., Myasnikov, A.G., Mastio, L., Billas, I.M.L., Birck C., Stella, S. and Klaholz, B.P. (2014) Purification, characterization and crystallization of the human 80S ribosome. Nucleic Acids Res. **42**, 1–11

Khatter, H., Myasnikov, A.G., Natchiar, K. and Klaholz, B.P. (2015) Structure of the human 80S ribosome. Nature **520**, 640–645

Khoshouei, M., Pfeffer, S., Baumeister, W., Förster, F. and Danev, R. (2016a) Subtomogram analysis using the Volta phase plate. J. Struct. Biol. **16**, 30103–30104

Khoshouei, M., Radjainia, M., Phillips, A.J., Gerrard, J.A., Mitra, A.K., Plitzko, J.M., Baumeister, W. and Danev, R. (2016b) Volta phase plate cryo-EM of the small protein complex Prx3. Nat. Commun. **7**, 10534

Kim, D., Deerinck, T.J., Sigal, Y.M., Babcock, H.P., Ellisman, M.H. and Zhuang, X. (2015) Correlative stochastic optical reconstruction microscopy and electron microscopy. PLoS One **10**, e0124581.

Kizilyaprak, C., Daraspe, J. and Humbel, B.M. (2014) Focused ion beam scanning electron microscopy in biology. J. Microsc. **254**, 109–114

Klaholz, B.P. (2015) Structure sorting of multiple macromolecular states in heterogeneous cryo-EM samples by 3D multivariate statistical analysis. Open J. Stat. **5**, 820–836

Klaholz, B.P., Myasnikov, A.G. and van Heel, M. (2004) Visualization of release factor 3 on the ribosome during termination of protein synthesis. Nature **427**, 862–865

Koning, R.I., Celler, K., Willemse, J., Bos, E., van Wezel, G.P. and Koster, A.J. (2014) Correlative cryo-fluorescence light microscopy and cryo-electron tomography of Streptomyces. Methods Cell Biol. **124**, 217–239

Kosinski, J., Mosalaganti, S., von Appen, A., Teimer, R., DiGuilio, A.L., Wan, W., Bui, K.H., Hagen, W.J., Briggs, J.A., Glavy, J.S., Hurt, E. and Beck, M. (2016) Molecular architecture of the inner ring scaffold of the human nuclear pore complex. Science **352**, 363–365

Kuijper, M., van Hoften, G., Janssen, B., Geurink, R., De Carlo, S., Vos, M., van Duinen, G., van Haeringen, B. and Storms, M. (2015) FEI's direct electron detector developments: Embarking on a revolution in cryo-TEM. J. Struct. Biol. **192**, 179–187

Kunath, W., Weiss, K., Sack-Kongehl, H., Kessel, M., Zeitler, E. (1984). Time-resolved low-dose microscopy of glutamine synthetase molecules. Ultramicroscopy **13**, 241–252

Kuybeda, O., Frank, G.A., Bartesaghi, A., Borgnia, M., Subramaniam, S. and Sapiro, G.A. (2013) Collaborative framework for 3D alignment and classification of heterogeneous subvolumes in cryo-electron tomography. J. Struct. Biol. **181**, 116–127

Laine, R.F., Albecka, A., van de Linde, S., Rees, E.J., Crump, C.M. and Kaminski, C.F. (2015) Structural analysis of herpes simplex virus by optical super-resolution imaging. Nat. Commun. **6**, 5980

Li, X., Mooney, P., Zheng, S., Booth, C.R., Braunfeld, M.B., Gubbens, S., Agard, D.A. and Cheng, Y. (2013) Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. Nat. Methods **10**, 584–590

Liao, H.Y., Hashem, Y. and Frank, J. (2015) Efficient estimation of three-dimensional covariance and its application in the analysis of heterogeneous samples in cryo-electron microscopy. Structure **23**, 1129–1237

Lin, D.H., Stuwe, T., Schilbach, S., Rundlet, E.J., Perriches, T., Mobbs, G., Fan, Y., Thierbach, K., Huber, F.M., Collins, L.N., Davenport, A.M., Jeon, Y.E. and Hoelz, A. (2016) Architecture of the symmetric core of the nuclear pore. Science **352**, 6283

Löschberger, A., Franke, C., Krohne, G., van de Linde, S. and Sauer, M. (2014) Correlative super-resolution fluorescence and electron microscopy of the nuclear pore complex with molecular resolution. J. Cell Sci. **127**, 4351–4355

© 2016 The Authors. *Biology of the Cell* published by Wiley-VCH Verlag GmbH & Co. KGaA on behalf of Société Française des Microscopies and Société de Biologie Cellulaire de France.

Lučič, V., Rigort, A. and Baumeister, W. (2013) Cryo-electron tomography: the challenge of doing structural biology in situ. J. Cell Biol. **202**, 407–419
Lyumkis, D., Brilot, A.F., Theobald, D.L. and Grigorieff, N. (2013) Likelihood-based classification of cryo-EM images using FREALIGN. J. Struct. Biol. **183**, 377–388
Mahamid, J., Pfeffer, S., Schaffer, M., Villa, E., Danev, R., Cuellar, L.K., Förster, F., Hyman, A.A., Plitzko, J.M. and Baumeister, W. (2016) Visualizing the molecular sociology at the HeLa cell nuclear periphery. Science **351**, 969–972
Mahamid, J., Schampers, R., Persoon, H., Hyman, A.A., Baumeister, W. and Plitzko, J.M. (2015) A focused ion beam milling and lift-out approach for site-specific preparation of frozen-hydrated lamellas from multicellular organisms. J. Struct. Biol. **192**, 262–269
Maletta, M., Orlov, I.M., Roblin, P., Beck, Y., Moras, D., Billas, I.M.L. and Klaholz, B.P. (2014) The palindromic DNA-bound USP/EcR nuclear receptor adopts an asymmetric organization with allosteric domain positioning. Nat. Commun. **5**, 4139
McMullan, G., Chen, S., Henderson, R. and Faruqi, A.R. (2009) Detective quantum efficiency of electron area detectors in electron microscopy. Ultramicroscopy **109**, 1126–1143
McMullan, G., Faruqi, A.R., Clare, D. and Henderson, R. (2014) Comparison of optimal performance at 300keV of three direct electron detectors for use in low dose electron microscopy. Ultramicroscopy **147**, 156–163
Medalia, O., Weber, I., Frangakis, A.S., Nicastro, D., Gerisch, G. and Baumeister, W. (2002) Macromolecular architecture in eukaryotic cells visualized by cryoelectron tomography. Science **298**, 1209–1213
Ménétret, J-F., Khatter, H., Simonetti, A., Orlov, I., Myasnikov, A.G, Vidhya, K.V., Manicka, S., Torchy, M., Mohideen, K., Humm, A.-S., Hazemann, I., Urzhumtsev, A., Klaholz, B.P. (2013) Integrative structure-function analysis of large nucleoprotein complexes. In: D. Klostermeier and C. Hammann (Eds.). *RNA structure and folding*. de Gruyter.
Merk, A., Bartesaghi, A., Banerjee, S., Falconieri, V., Rao, P., Davis, M.I., Pragani, R., Boxer, M.B., Earl, L.A., Milne, J.L. and Subramaniam, S. (2016) Breaking Cryo-EM Resolution Barriers to Facilitate Drug Discovery. Cell **165**, 1698–1707
Milne, J.L., Borgnia, M.J., Bartesaghi, A., Tran, E.E., Earl, L.A., Schauder, D.M., Lengyel, J., Pierson, J., Patwardhan, A. and Subramaniam, S. (2013) Cryo-electron microscopy - a primer for the non-microscopist. FEBS J. **280**, 28–45
Murata, K., Liu, X., Danev, R., Jakana, J., Schmid, M.F., King, J., Nagayama, K. and Chiu, W. (2010) Zernike phase contrast cryo-electron microscopy and tomography for structure determination at nanometer and subnanometer resolutions. Structure **18**, 903–912
Myasnikov, A.G., Afonina, Z.A., Ménétret, J-F., Shirokov, V.A., Spirin, A.S. and Klaholz, B.P. (2014) The molecular structure of the left-handed supra-molecular helix of eukaryotic polyribosomes. Nat. Commun. **5**, 5294
Myasnikov, A.G., Afonina Z. and Klaholz, B.P. (2013) Single particle and molecular assembly analysis of polyribosomes by single- and double-tilt cryo electron tomography. Ultramicroscopy **126**, 33–39
Myasnikov, A.G., Natchiar, S.K., Nebout, M., Hazemann, I., Imbert, V., Khatter, H., Peyron, J-F. and Klaholz, B.P. (2016) Structure-function insights reveal the human ribosome as a cancer target for antibiotics. Nat. Commun. **7**, 12856
Nans, A., Kudryashev, M., Saibil, H.R. and Hayward, R.D. (2015) Structure of a bacterial type III secretion system in contact with a host membrane in situ. Nat. Commun. **6**, 10114
Nederlof, I., Li, Y.W., van Heel, M. and Abrahams, J.P. (2013) Imaging protein three-dimensional nanocrystals with cryo-EM. Acta Crystallogr. D Biol. Crystallogr. **69**, 852–859
Obbineni, J.M., Yamamoto, R. and Ishikawa, T. (2016) A simple and fast approach for missing-wedge invariant classification of subtomograms extracted from filamentous structures. J. Struct. Biol. **16**, 30172
Orlov, I., Rochel, N., Moras, D. and Klaholz, B.P. (2012) Structure of the full human RXR/VDR nuclear receptor heterodimer complex with its DR3 target DNA. EMBO J. **31**, 291–300
Orlov, I., Schertel, A., Zuber, G., Klaholz, B.P., Drillien, R., Weiss, E., Schultz, P. and Spehner, D. (2015) Live cell immunogold labelling of RNA polymerase II. Sci. Rep. **5**, 8324
Orlova, E.V. and Saibil, H.R. (2010) Methods for three-dimensional reconstruction of heterogeneous assemblies. Methods Enzymol. **482**, 321–41
Ortiz, J.O., Förster, F., Kürner, J., Linaroudis, A.A. and Baumeister, W. (2006) Mapping 70S ribosomes in intact cells by cryoelectron tomography and pattern recognition. J. Struct. Biol. **156**, 334–341
Penczek, P.A., Frank, J. and Spahn, C.M. (2006) A method of focused classification, based on the bootstrap 3D variance analysis, and its application to EF-G-dependent translocation. J. Struct. Biol. **154**, 184–194
Rajendran, C., Dworkowski, F.S., Wang, M. and Schulze-Briese, C. (2011) Radiation damage in room-temperature data acquisition with the PILATUS 6M pixel detector. J. Synchrotron Radiat. **18**, 318–328
Ray, P., Klaholz, B.P., Finn, R.D., Orlova, E.V., Burrows, P.C., Gowen, B., Buck, M. and van Heel, M. (2003) Determination of Escherichia coli RNA polymerase structure by single particle cryoelectron microscopy. Methods Enzymol. **370**, 24–42
Razinkov, I., Dandey, V.P., Wei, H., Zhang, Z., Melnekoff, D., Rice, W.J., Wigge, C., Potter, C.S. and Carragher, B. (2016) A new method for vitrifying samples for cryoEM. J. Struct. Biol. **195**, 190–198
Rhinow, D. (2016) Towards an optimum design for thin film phase plates. Ultramicroscopy **160**, 1–6
Rigort, A. and Plitzko, J.M. (2015) Cryo-focused-ion-beam applications in structural biology. Arch. Biochem. Biophys. **581**, 122–130
Ruskin, R.S., Yu, Z. and Grigorieff, N. (2013) Quantitative characterization of electron detectors for transmission electron microscopy. J. Struct. Biol. **184**, 385–393
Sawaya, M.R., Rodriguez, J., Cascio, D., Collazo, M.J., Shi, D., Reyes, F.E., Hattne, J., Gonen, T. and Eisenberg, D.S. (2016) Ab initio structure determination from prion nanocrystals at atomic resolution by MicroED. Proc. Natl. Acad. Sci. U.S.A. **2016**. pii: 201606287
Schaffer, M., Mahamid, J., Engel, B.D., Laugks, T., Baumeister, W. and Plitzko, J.M. (2016) Optimized cryo-focused ion beam sample preparation aimed at in situ structural studies of membrane proteins. J. Struct. Biol. **16**, 30151–30154
Scheres, S.H., Valle, M., Nuez, R., Sorzano, C.O., Marabini, R., Herman, G.T. and Carazo, J.M. (2005) Maximum likelihood multi-reference refinement for electron microscopy images. J. Mol. Biol. **22**, 139–149
Scheres, S.H. (2010) Classification of structural heterogeneity by maximum-likelihood methods. Methods Enzymol. **482**, 295–320
Scheres, S.H. (2014) Beam-induced motion correction for sub-megadalton cryo-EM particles. Elife **3**, e03665.
Schirra, R.T. Jr. and Zhang, P. (2014) Correlative fluorescence and electron microscopy. Curr. Protoc. Cytom. **70**, 1–10
Schorb, M., Gaechter, L., Avinoam, O., Sieckmann, F., Clarke, M., Bebeacua, C., Bykov, Y.S., Sonnen, A.F., Lihl, R. and Briggs, J.A. (2016) New hardware and workflows for semi-automated correlative cryo-fluoresence and cryo-electron microscopy/tomography. J. Struct. Biol. **16**, 30135–30136
Schur, F.K., Obr, M., Hagen, W.J., Wan, W., Jakobi, A.J., Kirkpatrick, J.M., Sachse, C., Kräusslich, H.G. and Briggs, J.A. (2016) An

atomic model of HIV-1 capsid-SP1 reveals structures regulating assembly and maturation. Science **353**, 506–508

Shi, D., Nannenga, B.L., de la Cruz, M.J., Liu, J., Sawtelle, S., Calero, G., Reyes, F.E., Hattne, J. and Gonen, T. (2016) The collection of MicroED data for macromolecular crystallography. Nat. Protoc. **11**, 895–904

Sigworth, F.J. (1998) A maximum-likelihood approach to single-particle image refinement. J. Struct. Biol. **122**, 328–339

Simonetti, A., Marzi, S., Fabbretti, A., Hazemann, I., Jenner, L., Urzhumtsev, A., Gualerzi, C.O. and Klaholz, B.P. (2013a) Structure of the protein core of translation initiation factor 2 in apo, GTP-bound and GDP-bound forms. Acta Cryst. **D69**, 925–933

Simonetti, A., Marzi, S., Billas, I.M.L., Tsai, A., Fabbretti, A., Myasnikov, A.G., Roblin, P., Vaiana, A.C., Hazemann, I., Eiler, D., Steitz, T.A., Puglisi, J.D., Gualerzi, C.O. and Klaholz, B.P. (2013b) Involvement of IF2 N domain in ribosomal subunit joining revealed from architecture and function of the full-length initiation factor. Proc. Nat. Acad. Sci. U.S.A. **110**, 15656–15661

Simonetti, A., Ménétret, J-F., Martin, F., Myasnikov, A.G., Vicens, Q., Prongidi-Fix, L., Natchiar, S.K., Klaholz, B.P. and Eriani, G. (2016) Ribosomal 18S rRNA base pairs with mRNA during eukaryotic translation initiation. Nat. Commun. **7**, 12622

Simonetti, A., Marzi, S., Myasnikov, A.G., Fabbretti, A., Yusupova, G., Yusupov, M., Gualerzi, C.O. and Klaholz, B.P. (2008) Structure of the 30S translation initiation complex. Nature **455**, 416–420

Smart, O.S., Womack, T.O., Flensburg, C., Keller, P., Paciorek, W., Sharff, A., Vonrhein, C. and Bricogne, G. (2012) Exploiting structure similarity in refinement: automated NCS and target-structure restraints in BUSTER. Acta Crystallogr. D Biol. Crystallogr. **68**, 368–380

Sobolev, O.V., Afonine, P.V., Adams, P.D. and Urzhumtsev, A. (2015) Programming new geometry restraints: parallelity of atomic groups. J. Appl. Crystallogr. **48**, 1130–1141

Spear, J.M., Noble, A.J., Xie, Q., Sousa, D.R., Chapman, M.S. and Stagg, S.M. (2015) The influence of frame alignment with dose compensation on the quality of single particle reconstructions. J. Struct. Biol. **192**, 196–203

Spinelli, S., Bebeacua C., Orlov, I., Tremblay, D., Klaholz, B.P., Moineau, S. and Cambillau, C. (2014) CryoEM structure of the lactococcal siphophage 1358 virion. J. Virol. **88**, 8900–8910

Stölken, M., Beck, F., Haller, T., Hegerl, R., Gutsche, I., Carazo, J.M., Baumeister, W., Scheres, S.H. and Nickell, S. (2011) Maximum likelihood based classification of electron tomographic data. J. Struct. Biol. **173**, 77–85

Szymborska, A., de Marco, A., Daigle, N., Cordes, V.C., Briggs, J.A. and Ellenberg, J. (2013) Nuclear pore scaffold structure analyzed by super-resolution microscopy and particle averaging. Science **341**, 655–658

Tan, Y.Z., Cheng, A., Potter, C.S. and Carragher, B. (2016) Automated data collection in single particle electron microscopy. Microscopy (Oxf.) **65**, 43–56

van Heel, M., Gowen, B., Matadeen, R., Orlova, E.V., Finn, R., Pape, T., Cohen, D., Stark, H., Schmidt, R., Schatz, M. and Patwardhan, A. (2000) Single-particle electron cryo-microscopy: towards atomic resolution. Q. Rev. Biophys. **33**, 307–369

Veesler, D., Campbell, M.G., Cheng, A., Fu, C.Y., Murez, Z., Johnson, J.E., Potter, C.S. and Carragher, B. (2013) Maximizing the potential of electron cryomicroscopy data collected using direct detectors. J. Struct. Biol. **184**, 193–202

Villa, E., Schaffer, M., Plitzko, J.M. and Baumeister, W. (2013) Opening windows into the cell: focused-ion-beam milling for cryo-electron tomography. Curr. Opin. Struct. Biol. **23**, 771–777

Walter, A., Steltenkamp, S., Schmitz, S., Holik, P., Pakanavicius, E., Sachser, R., Huth, M., Rhinow, D. and Kühlbrandt, W. (2015) Towards an optimum design for electrostatic phase plates. Ultramicroscopy **153**, 22–31

Wan, W. and Briggs, J.A. (2016) Cryo-electron tomography and subtomogram averaging. Methods Enzymol. **579**, 329–367

Wang, K., Fu, C.Y., Khayat, R., Doerschuk, P.C. and Johnson, J.E. (2011) In vivo virus structures: simultaneous classification, resolution enhancement, and noise reduction in whole-cell electron tomography. J. Struct. Biol. **174**, 425–433

White, H.E., Saibil, H.R., Ignatiou, A. and Orlova, E.V. (2004) Recognition and separation of single particles with size variation by statistical analysis of their images. J. Mol. Biol. **13**, 453–460

Wong, W., Bai, X.C., Brown, A., Fernandez, I.S., Hanssen, E., Condron, M., Tan, Y.H., Baum, J. and Scheres, S.H. (2014) Cryo-EM structure of the Plasmodium falciparum 80S ribosome bound to the anti-protozoan drug emetine. Elife **3**. doi: 10.7554/eLife.03080

Xu, M. and Alber, F. (2013) Automated target segmentation and real space fast alignment methods for high-throughput classification and averaging of crowded cryo-electron subtomograms. Bioinformatics **29**, 274–282

Yang, Z., Lasker, K., Schneidman-Duhovny, D., Webb, B., Huang, C.C., Pettersen, E.F., Goddard, T.D., Meng, E.C., Sali, A. and Ferrin, T.E. (2012) UCSF Chimera, MODELLER, and IMP: an integrated modeling system. J. Struct. Biol. **179**, 269–278

Zhang, J., Ji, G., Huang, X., Xu, W. and Sun, F. (2016) An improved cryo-FIB method for fabrication of frozen hydrated lamella. J. Struct. Biol. **194**, 218–223

Zhang, X., Ge, P., Yu, X., Brannan, J.M., Bi, G., Zhang, Q., Schein, S. and Zhou, Z.H. (2013) Cryo-EM structure of the mature dengue virus at 3.5-Å resolution. Nat. Struct. Mol. Biol., **20**, 105–110

Received: 12 July 2016; Revised: 4 October 2016; Accepted: 5 October 2016; Accepted article online: 12 October 2016

© 2016 The Authors. *Biology of the Cell* published by Wiley-VCH Verlag GmbH & Co. KGaA on behalf of Société Française des Microscopies and Société de Biologie Cellulaire de France.

6.3. Posters

1. Leonid Andronov, Yves Lutz, Jean-Luc Vonesch, Bruno Klaholz. SharpGSDIM, a software for processing of super-resolution microscopy data. Focus on Microscopy-2015. March 29 – April 1, 2015, Göttingen, Germany.

2. Leonid Andronov, Igor Orlov, Yves Lutz, Jean-Luc Vonesch, Bruno Klaholz. ClusterViSu, a method for clustering and colocalization analysis of protein complexes by Voronoi tessellation in single-molecule microscopy. ICON Europe 2016. June 07–10, 2016, Basel, Switzerland.

3. Leonid Andronov, Jonathan Michalon, Khalid Ouararhni, Igor Orlov, Yves Lutz, Jean-Luc Vonesch, Bruno Klaholz. Segmentation and cluster analysis of protein complexes by Voronoi tessellation in 3D localization microscopy. Focus on Microscopy-2017. April 09–12, 2017, Bordeaux, France.

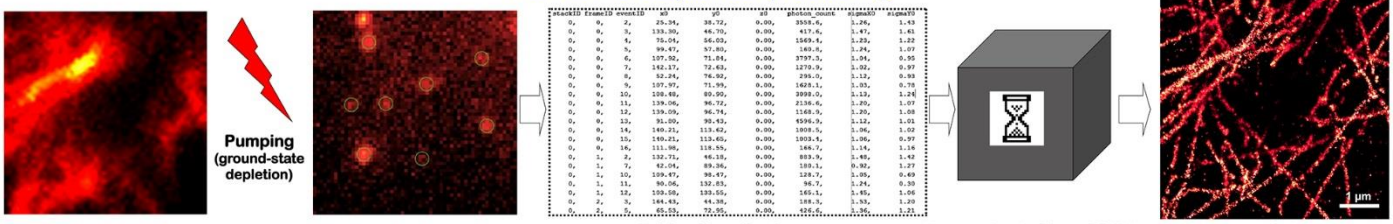
SharpGSDIM, a software for processing of super-resolution data

Leonid Andronov, Yves Lutz, Jean-Luc Vonesch, Bruno Klaholz
Centre de Biologie Intégrative, Institute de génétique et de biologie moléculaire et cellulaire (IGBMC), CNRS UMR 7104, Inserm U 964, 1 rue Laurent Fries / BP 10142 / 67404 Illkirch CEDEX, France
Université de Strasbourg, 67404 Illkirch, France
E-mail: andronov@igbmc.fr

The last decade in the light microscopy field was marked by the invention of super-resolution microscopy which overcomes the diffraction limit. Among other techniques, the most resolutive ones currently are stochastic-based methods reaching the 10-20 nm precision range. Resolution is limited mostly by localization accuracy, sample drift and labelling density. For multi-color imaging, the precision of colocalization is additionally impaired by chromatic aberrations of the microscope. Most of these aspects can be significantly improved through post-processing of localization data.

We introduce an interactive open-source software with a graphical user interface that allows performing processing steps for super-resolution data, such as correction of chromatic aberrations; cross-correlation-based iterative correction of drift; selection of localization events; reconstruction of super-resolution images in different modes, namely Voronoi, histogram, Gaussian and time-in-color; reconstruction of 3D-images; estimation of resolution by Fourier ring correlation. The software is optimized to work with eventlist tables exported from most popular localization tools. The output can be saved in widespread .tiff, .png, or .ascii formats allowing further analysis or preparation of publications.

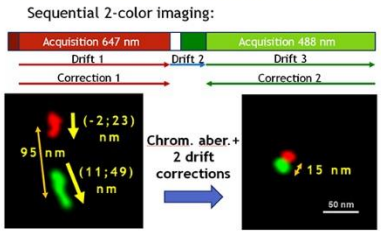
Workflow of acquisition-processing



Preview in epifluorescence mode | Detection of single-molecule events | Table of localizations | Processing in SharpGSDIM | Super-resolution image

Features of the SharpGSDIM software

Chromatic aberrations + drift



Time-in-color + drift correction

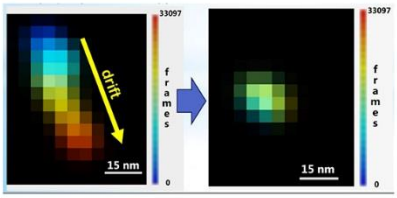
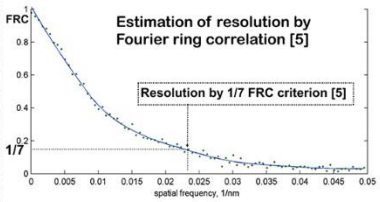
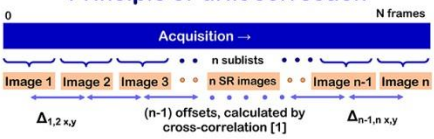


Image of a bead, drifted during long acquisition. The time dependence is coded in the colormap. The time dependence is mostly corrected. The bead appears green due to averaging.

Fourier ring correlation

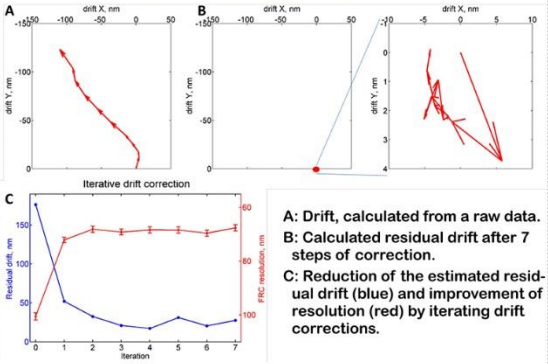


Principle of drift correction



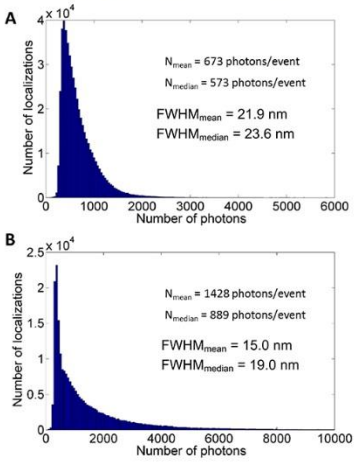
Regarding drift as constant on each sublist, move localizations towards the beginning (channel acquired secondly) or towards the end (channel acquired firstly) [2].

Iterative drift correction



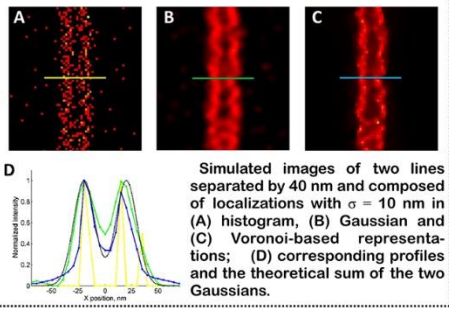
A: Drift, calculated from a raw data.
B: Calculated residual drift after 7 steps of correction.
C: Reduction of the estimated residual drift (blue) and improvement of resolution (red) by iterating drift corrections.

Averaging of consecutive events



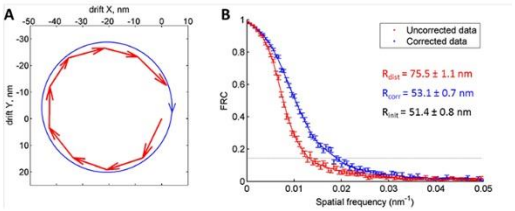
A: Histogram of the distribution of the number of photons per localization calculated from the raw data.
B: The histogram after merge consecutive events within a radius of 50 nm and subsequent removing of localizations with more than 10^4 photons demonstrating increased mean and median photon counts as well as increased average precision of localization.

Visualization techniques



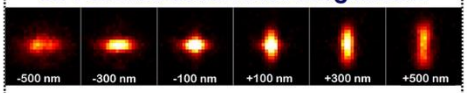
Simulated images of two lines separated by 40 nm and composed of localizations with σ = 10 nm in (A) histogram, (B) Gaussian and (C) Voronoi-based representations; (D) corresponding profiles and the theoretical sum of the two Gaussians.

Correction of simulated drift

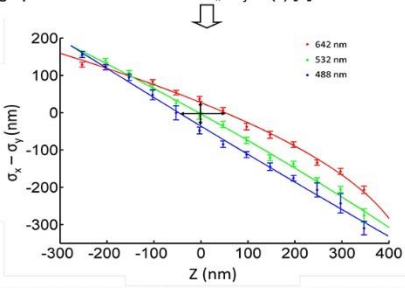


A: ○ shape of simulated ideal circular drift, d = 50 nm;
→ drift, calculated by SharpGSDIM from the distorted dataset.
B: FRCs of distorted and corrected data showing considerable improvement in resolution after the correction.

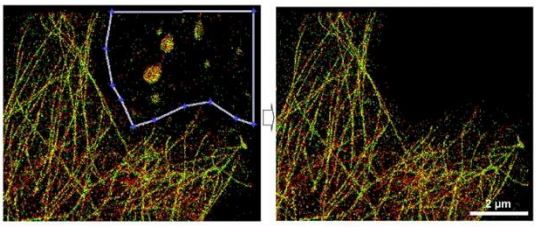
3D reconstruction with astigmatism



Deformation of the point spread function depending on Z-position of objective with a cylindrical lens inserted in the lightpath -> calibration curve $\sigma_x - \sigma_y = f(z)$ [4].



Selection of a region of interest



Software availability

The source code and a stand-alone application for Windows are freely available at https://github.com/andronovl/SharpGSDIM

References

1. M. Guizar-Sicairos, S. Thurman, J. Fienup, Opt. Lett. 33, 156-158 (2008).
2. A. Szymborska, A. de Marco *et al*, Science 341, 655-658 (2013).
3. E. Betzig, G. Patterson *et al*, Science 313, 1642–1645 (2006).
4. R. Henriques, M. Lelek *et al*, Nature Methods 7, 339 - 340 (2010).
5. R. Nieuwenhuizen, K. Lidke *et al*, Nature Methods 10, 557-562 (2013).

ClusterViSu, a method for clustering and colocalization analysis of protein complexes by Voronoi tessellation in single-molecule microscopy

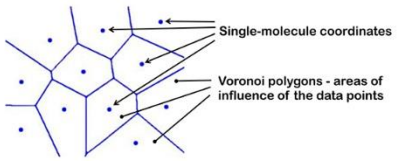
Leonid Andronov, Igor Orlov, Yves Lutz, Jean-Luc Vonesch, Bruno Klaholz

Centre de Biologie Intégrative (CBI), Institute de génétique et de biologie moléculaire et cellulaire (IGBMC), Université de Strasbourg, CNRS UMR 7104, Inserm U 964, 1 rue Laurent Fries / BP 10142 / 67404 Illkirch CEDEX, France. E-mail: andronov@igbmc.fr

As opposed to classical light microscopy, the data of single-molecule localization microscopy (SMLM) is essentially discontinuous because it comprises a set of points with molecular coordinates of the localization events, and hence it does not allow a direct segmentation. Considering the plethora of fluorescent signals in dense environments such as chromatin, the interpretation of localization data can become rather challenging with regards to image segmentation, clustering and colocalization analysis.

We present a robust method based on Voronoi tessellation [1] for the convenient visualization and quantification of the localization and distribution of fluorescently labeled complexes, which allows segmentation, cluster analysis and estimation of the amount of colocalization. We show that image reconstruction using Voronoi diagrams preserves resolution equally well or better than Gaussian or histogram modes at the level of image rendering, as quantified by Fourier ring correlation. Voronoi tessellation allows performing a statistical analysis of the clusters, their occurrence and inter-cluster distance distribution, and works well including for the analysis of weak signals. The general applicability of the method is illustrated on a series of important biological objects such as chromatin, nuclear pore complexes and microtubules. We demonstrate application of Voronoi diagrams to colocalization studies of double-color data on histone H2B and RNA polymerase: quantification shows that both protein complexes form separate clusters in the cell nucleus.

Voronoi diagram



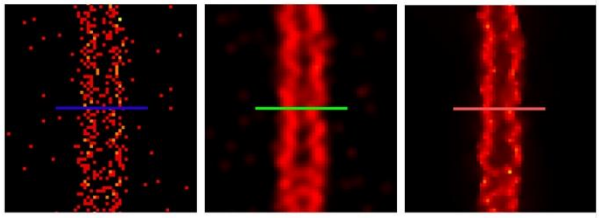
A Voronoi diagram allows for an unambiguous determination of the local density

2D case: $\delta_i = 1/S_i$;

δ_i is the local density in the neighborhood of the i-th point,

S is the area of the i-th Voronoi polygon

Image reconstruction using Voronoi diagrams



Histogram | Gaussian | Voronoi local density

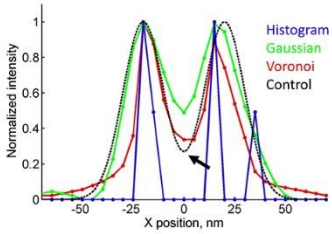
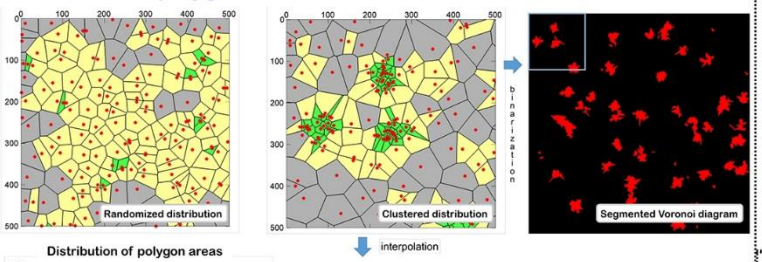
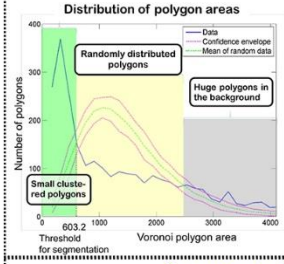


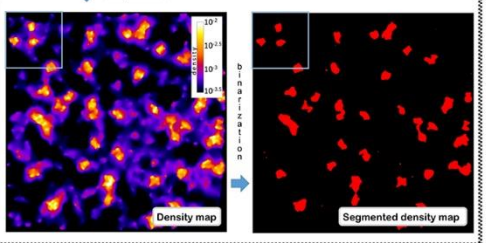
Image profiles

Voronoi polygons: clustered *vs* randomized distribution

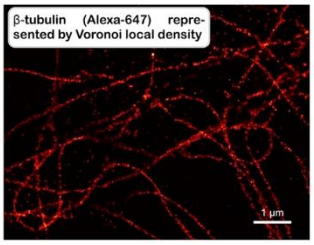


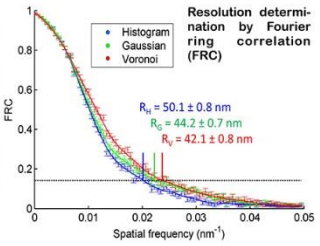
Distribution of polygon areas





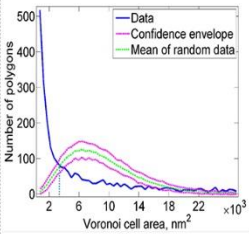
β-tubulin (Alexa-647) represented by Voronoi local density

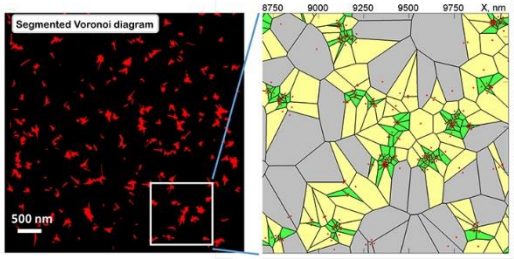




Resolution determination by Fourier ring correlation (FRC)

Segmentation of nuclear pore complexes

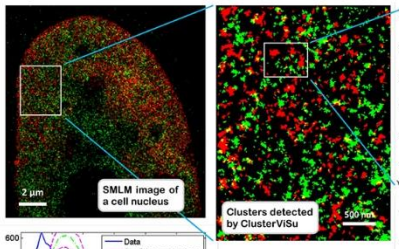


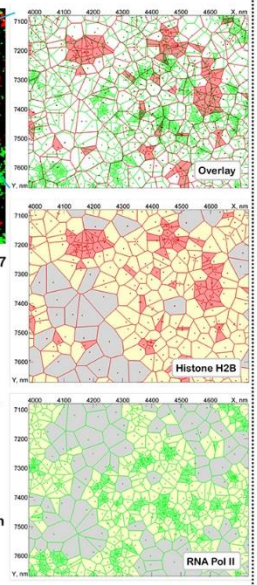


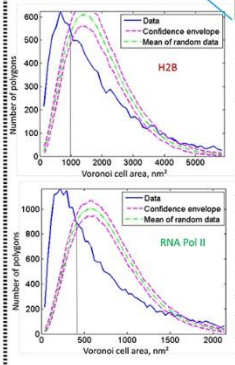
Labelling: TPR-Alexa-647

Results of segmentation: 9.7 ± 7.3 events/cluster; the equivalent radius of the clusters = 57 ± 11 nm.

Colocalization of H2B and RNA Pol II





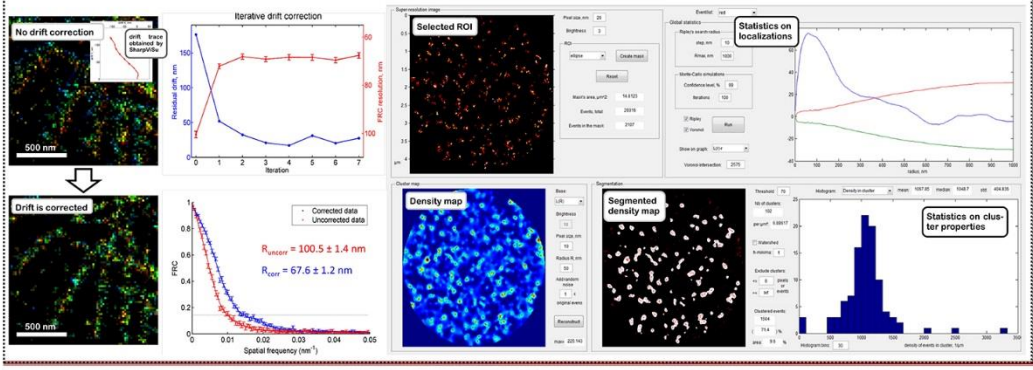


Labelling: H2B - Alexa-647
RNA Pol II - Alexa-488

Average cluster size:
RNA Pol II 1818 nm²,
H2B 3058 nm²

Colocalization quantification:
S/S_{Pol} = 0.090,
S_{rand}/S_{Pol} = 0.096 ± 0.013;
S/S_{H2B} = 0.106,
S_{rand}/S_{H2B} = 0.113 ± 0.015
=> no correlation between the H2B and RNA Pol II clusters

Features of SharpViSu, an integrated software for processing of SMLM data



Software availability

We have implemented the Voronoi tessellation method as a ClusterViSu standalone software [1] which can also be interfaced as a plugin with an interactive open-source software for processing of super-resolution fluorescence microscopy data, SharpViSu [2], which includes tools such as corrections for drift and chromatic aberrations and resolution estimation by FRC. Both software are available under https://github.com/andronovl/SharpViSu.

References

1. Andronov, L. *et al.* ClusterViSu, a method for clustering of protein complexes by Voronoi tessellation in super-resolution microscopy. *Sci. Rep.* **6**, 24084 (2016); doi: 10.1038/srep24084.

2. Andronov, L. *et al.* SharpViSu: integrated analysis and segmentation of super-resolution microscopy data. *Bioinformatics* (2016); doi:10.1093/bioinformatics/btw123.

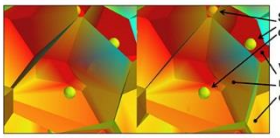
SEGMENTATION AND CLUSTER ANALYSIS OF PROTEIN COMPLEXES BY VORONOI TESSELLATION IN 3D LOCALIZATION MICROSCOPY

Leonid Andronov, Jonathan Michalon, Khalid Ouararhni, Igor Orlov, Yves Lutz, Ali Hamiche, Jean-Luc Vonesch, Bruno Klaholz
Centre de Biologie Intégrative (CBI), Institute de génétique et de biologie moléculaire et cellulaire (IGBMC), Université de Strasbourg, CNRS UMR 7104, Inserm U 964, 1 rue Laurent Fries / BP 10142 / 67404 Illkirch CEDEX, France. E-mail: andronov@igbmc.fr

As opposed to classical light microscopy, the data of single-molecule localization microscopy is essentially discontinuous because it comprises a set of points with molecular coordinates of the localization events and hence it does not allow a direct segmentation. Considering the plethora of fluorescent signals in dense environments such as chromatin, the interpretation of localization data can become rather challenging with regards to image segmentation, clustering and colocalization analysis. We have recently developed ClusterViSu, a robust pipeline for processing of 2D localization data based on Voronoi tessellations [1]-[2]. ClusterViSu is suitable for the convenient visualization and quantification of the localization and distribution of fluorescently labeled complexes, which allows segmentation, cluster analysis and estimation of the amount of co-localization. Voronoi diagrams allow performing a statistical analysis of the clusters, their occurrence and inter-cluster distance distribution, and work well including for the analysis of weak signals.

We now demonstrate an extension of this method to the 3D case, which is much more complex. We show that 3D Voronoi diagrams can be used for precise local density estimation, noise reduction, rendering and cluster analysis of 3D data. One of the advantages of the Voronoi tessellation method is that it does not require any *a priori* knowledge for the clustering. Because the clustering uses an internal reference generated with Monte-Carlo simulations of randomized data to automatically determine the threshold value for forming clusters between neighboring Voronoi zones, it is fully automated for a given region of interest. The general applicability of the method is illustrated on microtubules, histones H2B and CENP-A chromatin data [3].

3D Voronoi diagram

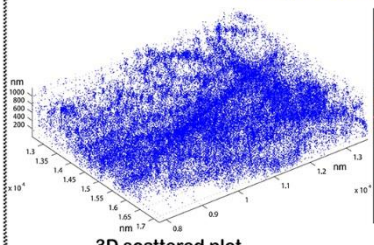


A Voronoi diagram allows for an unambiguous determination of the local density

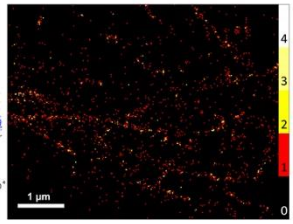
3D case: $\delta_i = 1/V_i$;

δ_i is the local density in the neighborhood of the i-th point,
V_i is the volume of the i-th Voronoi polygon

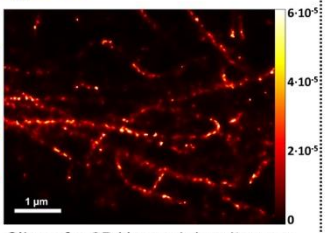
Image reconstruction using Voronoi diagrams



3D scattered plot

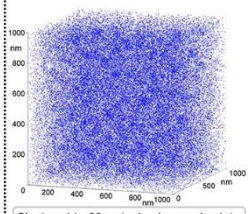


Slice of a 3D histogram

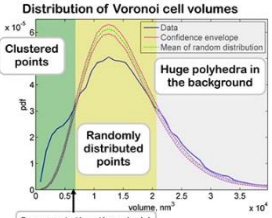


Slice of a 3D Voronoi density map

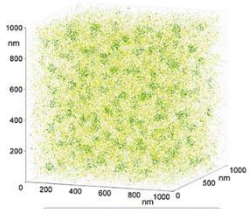
Voronoi polygons: clustered *vs* randomized distribution



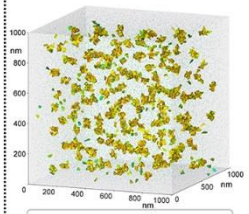
Clustered (r=30 nm) + background points



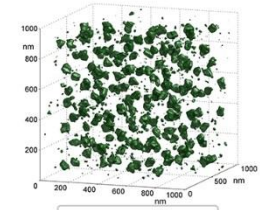
Segmentation threshold



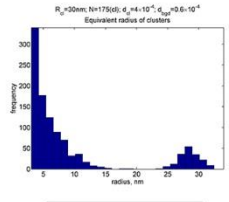
Clustered + background points



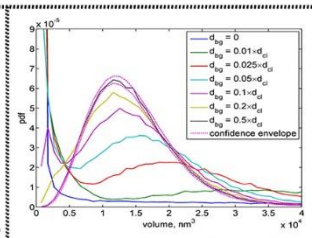
Voronoi polyhedra (V < threshold/2)

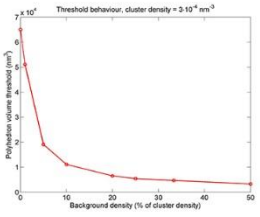


Segmented 3D density map

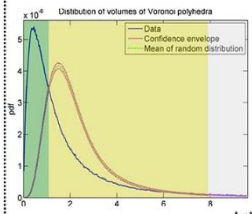


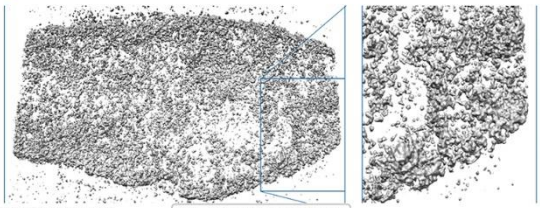
Radius of detected clusters





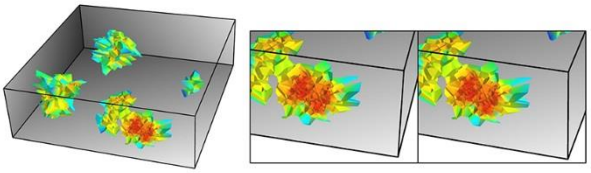
Segmentation of histone H2B

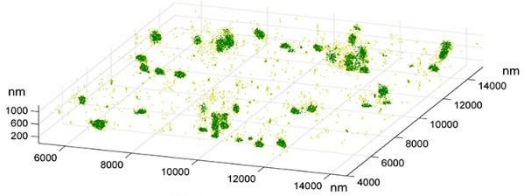


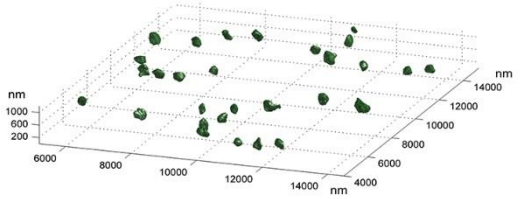


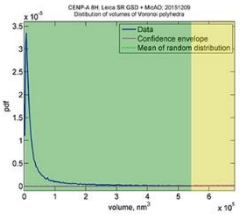
Labelling: H2B-Alexa-647 | Segmented 3D density map

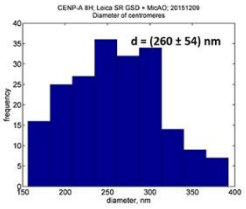
Segmentation of histone CENP-A



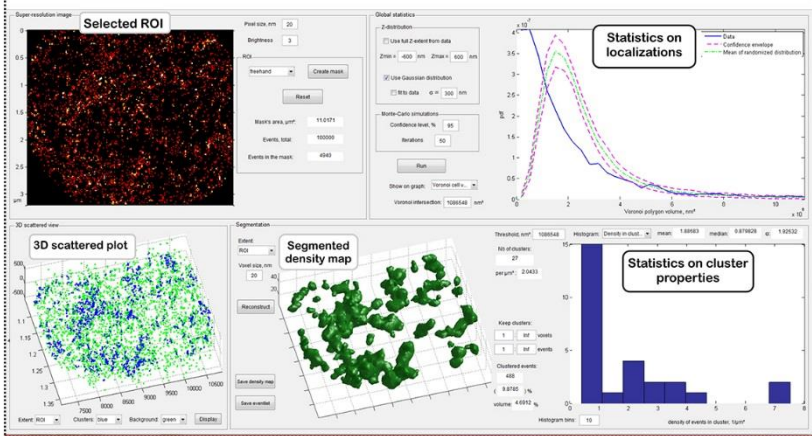








Software interface



References

1. Andronov L. *et al.* ClusterViSu, a method for clustering of protein complexes by Voronoi tessellation in super-resolution microscopy. *Sci. Rep.* **6**, 24084 (2016).

2. Andronov L. *et al.* SharpViSu: integrated analysis and segmentation of super-resolution microscopy data. *Bioinformatics* **32**(14), 2239-2241 (2016).

3. Andronov L. *et al.*, submitted.

http://cbi-dev.igbmc.fr/cbi/voronoi3D

UNIVERSITÉ DE STRASBOURG

Leonid ANDRONOV

École Doctorale des Sciences de la Vie et de la Santé STRASBOURG

Développement de méthodes avancées pour l'analyse et la segmentation de données de microscopie à super-résolution

Résumé

Parmi les méthodes de super-résolution, la microscopie par localisation de molécules uniques se distingue principalement par sa meilleure résolution réalisable en pratique mais aussi pour l'accès direct aux propriétés des molécules individuelles. Les données principales de la microscopie par localisation sont les coordonnées des fluorochromes, un type de données peu répandu en microscopie conventionnelle. Le développement de méthodes spéciales pour le traitement de ces données est donc nécessaire. J'ai développé les logiciels SharpViSu et ClusterViSu qui permettent d'effectuer les étapes de traitements les plus importantes, notamment une correction des dérives et des aberrations chromatiques, une sélection des événements de localisations, une reconstruction des données dans des images 2D ou dans des volumes 3D par le moyen de différentes techniques de visualisation, une estimation de la résolution à l'aide de la corrélation des anneaux de Fourier, et une segmentation à l'aide de fonctions K et L de Ripley. En plus, j'ai développé une méthode de segmentation de données de localisation en 2D et en 3D basée sur les diagrammes de Voronoï qui permet un clustering de manière automatique grâce à modélisation de bruit par les simulations Monte-Carlo. En utilisant les méthodes avancées de traitement de données, j'ai mis en évidence un clustering de la protéine CENP-A dans les régions centromériques des noyaux cellulaires et des transitions structurales de ces clusters au moment de la déposition de la CENP-A au début de la phase G1 du cycle cellulaire.

Mots-clés: microscopie à super-résolution, microscopie par localisation des molécules individuelles, dSTORM, diagrammes de Voronoï, traitement d'images, chromatine, histones, CENP-A, centromères

Résumé en anglais

Among the super-resolution methods single-molecule localization microscopy (SMLM) is remarkable not only for best practically achievable resolution but also for the direct access to properties of individual molecules. The primary data of SMLM are the coordinates of individual fluorophores, which is a relatively rare data type in fluorescence microscopy. Therefore, specially adapted methods for processing of these data have to be developed. I developed the software SharpViSu and ClusterViSu that allow for most important data processing steps, namely for correction of drift and chromatic aberrations, selection of localization events, reconstruction of data in 2D images or 3D volumes using different visualization techniques, estimation of resolution with Fourier ring correlation, and segmentation using K- and L-Ripley functions. Additionally, I developed a method for segmentation of 2D and 3D localization data based on Voronoi diagrams, which allows for automatic and unambiguous cluster analysis thanks to noise modeling with Monte-Carlo simulations. Using advanced data processing methods, I demonstrated clustering of CENP-A in the centromeric regions of the cell nucleus and structural transitions of these clusters upon the CENP-A deposition in early G1 phase of the cell cycle.

Keywords: super-resolution microscopy, single-molecule localization microscopy, dSTORM, Voronoi diagrams, image processing, chromatin, histones, CENP-A, centromeres