3D super-resolution fluorescence microscopy maps the variable molecular architecture of the Nuclear Pore Complex

Vilma Jimenez Sabinina, M. Hossain, Jean-Karim Hériché, Philipp Hoess, Bianca Nijmeijer, Shyamal Mosalaganti, Moritz Kueblbeck, Andrea Callegari, Anna Szymborska, Martin Beck, Jonas Ries, and Jan Ellenberg

Corresponding author(s): Jan Ellenberg, European Molecular Biology Laboratory, Heidelberg, Germany

Review Timeline:

Submission Date:	2020-12-04
Editorial Decision:	2021-02-08
Revision Received:	2021-04-16
Accepted:	2021-05-05

Editor-in-Chief: Matthew Welch

Transaction Report:

(Note: With the exception of the correction of typographical or spelling errors that could be a source of ambiguity, letters and reports are not edited. The original formatting of letters and referee reports may not be reflected in this compilation.)

RE: Manuscript #E20-11-0728
TITLE: "3D super-resolution fluorescence microscopy maps the variable molecular architecture of the Nuclear Pore Complex"

Dear Jan,

Your paper has been reviewed by two experts in imaging and their comments are appended below. Basically, they both enthusiastically recommend publication after you provide more critical discussion of your data and give more description of the methodology. I look forward to seeing your revised paper soon.

Sincerely,

Jennifer Lippincott-Schwartz

Monitoring Editor
Molecular Biology of the Cell

Dear Dr. Ellenberg,

The review of your manuscript, referenced above, is now complete. The Monitoring Editor has decided that your manuscript requires minor revisions before it can be published in Molecular Biology of the Cell, as described in the Monitoring Editor's decision letter above and the reviewer comments (if any) below.

A reminder: Please do not contact the Monitoring Editor directly regarding your manuscript. If you have any questions regarding the review process or the decision, please contact the MBoC Editorial Office (mboc@ascb.org).

When submitting your revision include a rebuttal letter that details, point-by-point, how the Monitoring Editor's and reviewers' comments have been addressed. (The file type for this letter must be "rebuttal letter"; do not include your response to the Monitoring Editor and reviewers in a "cover letter.") Please bear in mind that your rebuttal letter will be published with your paper if it is accepted, unless you have opted out of publishing the review history.

Authors are allowed 180 days to submit a revision. If this time period is inadequate, please contact us immediately at mboc@ascb.org.

In preparing your revised manuscript, please follow the instruction in the Information for Authors (www.molbiolcell.org/info-for-authors). In particular, to prepare for the possible acceptance of your revised manuscript, submit final, publication-quality figures with your revision as described.

To submit the rebuttal letter, revised version, and figures, please use this link (please enable cookies, or cut and paste URL): Link Not Available

Authors of Articles and Brief Communications whose manuscripts have returned for minor revision ("revise only") are encouraged to create a short video abstract to accompany their article when it is published. These video abstracts, known as Science Sketches, are up to 2 minutes long and will be published on YouTube and then embedded in the article abstract. Science Sketch Editors on the MBoC Editorial Board will provide guidance as you prepare your video. Information about how to prepare and submit a video abstract is available at www.molbiolcell.org/science-sketches. Please contact mboc@ascb.org if you are interested in creating a Science Sketch.

Thank you for submitting your manuscript to Molecular Biology of the Cell. Please do not hesitate to contact this office if you have any questions.

Sincerely,

Eric Baker
Journal Production Manager
MBoC Editorial Office
mbc@ascb.org

Reviewer #1 (Remarks to the Author):

Recent advancements in cryo-electron microscopy have made it possible to elucidate the molecular structure of a large number of protein complexes including the nuclear pore. This manuscript presents an alternative approach relying on 3D super-resolution light microscopy to map the position of several nucleoporins/Nups within the nuclear pore complex of human cells at nanoscale resolution. To achieve this, the authors expand a computational pipeline previously used for sub-tomogram averaging to extract individual NPCs from 3D SMLM volumes, averaging and aligning the signal of fluorescently tagged Nups relative to the reference protein ELYS. From this, they build a three-dimensional localization-density map for several Nups, outlining their average axial and radial positions within the NPC.
In a next step, the authors employ an unbiased classification method, by clustering individual NPC sub-volumes based on the topological features of Nup localization point clouds to capture the structural heterogeny within the NPC-population. Interestingly, for several Nups different clusters were found, distinguished by a single morphological feature, pointing to different conformational states of the mature NPC or potentially, NPC assembly intermediates.

The method described here is an exciting technical advancement validated by the observation that Nup localization patterns match well with already known or predicted positions. Furthermore, it provides new information of less well characterized positions of the asymmetric Nups ELYS and TRP. In addition, this approach could have the potential to capture the structural heterogeneity within native complex-populations, something classical averaging-based methods fail to do. Unfortunately, the manuscript falls short in clearly explaining essential aspects of the presented method and generally does a rather poor job at discussing and addressing associated technical limitations. In addition, the authors often do not provide enough evidence to empower the reader to make judgments on the conclusions but instead make sometimes bold statements that are difficult to verify.

Major points:

1.) Conformational flexibility of ELYS. The authors chose the Nup ELYS as a reference. They do the averaging and alignment on the ELYS-signal and then propagate the transformations to the NUP channel that is examined. In such an experimental setup the resolution of the SNAP-tagged NUPs critically depend on the reference. One would assume that peripheral Nups are overall more prone to variability than the central ones and additionally the larger fluorophore displacement of the antibody labeling of ELYS will inevitably introduce more uncertainty. Do the authors know if the structured or the unstructured domain of ELYS is recognized by the ELYS-antibody? A lot of the observed variability might actually come from using the ELYS-antibody as reference rather than the respective NUPs? Why did the authors select ELYS and did not use one of the presumably more rigid Y-complex Nups as a reference?

2.) Biological vs. technical variability. Throughout the manuscript it is difficult to judge, which observations are biological relevant, and which are beyond the precision limit of the approach. The measurements in the manuscript are given in single digit nanometer precision, however the fact that the individual spokes of the SNAP-tagged Y-complex Nups and the Nup133 rings cannot be resolved suggest a resolution >1nm. The article would greatly benefit from a paragraph in the beginning of the manuscript establishing the technical resolution limit of the approach and discussing potential drawbacks of the method, so the reader can judge the soundness of the observed differences.
The fact that the eightfold rotational symmetry of the NPC cannot be resolved for any of the NUPs in the averaged localization-density maps is insufficiently addressed. The authors hypothesize that this is like due to the structural variability however there is no evidence provided in support of this claim. In fact, the observation that the two Nup133 rings cannot be resolved would argue against it. It would be important to know how the localization precision of ~ 7 nm was determined, and the authors should show evidence where this is tested. It is also stated that the heterogeneity explores a larger range of distances than can be explained by technical variability. What is the evidence for this?

3.) Classification method does not seem very robust. There seem to be distinct clusters for Nup107 and maybe TPR but the others are not clearly dominated by a single NUP. Again, the authors do not provide a statistical analysis, or any other evidence to support the statement that the observed differences are based on structural diversity instead of technical variability. Furthermore, it is not obvious why Nup107 and Seh1, which show very similar double ring structures can be clearly distinguished, whereas Nup133 that has a single ring cannot be distinguished well from the other Y-complex double-ring Nups. Also, why do some of the classes have way fewer particles than other classes?

4.) The weaker signal of the cytoplasmic ring. The fact that the cytoplasmic ring is weaker than the nucleoplasmic one needs to be addressed. One can see this both in the total averaged localization densities (Fig. 1) and especially pronounced in the sub-cluster densities (Fig. 4). Is this caused by the flexibility of using ELYS as a reference, causing the ring which is further away to blur out?

Minor points:

1. A schematic of the nuclear pore complex with an indication of the tagged components would help to make the manuscript more accessible for non-NPC specialists. Also, the workflow schematic (Fig 1A) should be made clearer.
2. Could the signal intensity be used to find out if the thicker Nup133 ring actually stems from double the number of molecules?
3. The authors should indicate the number of particles in each of the subclasses (Fig 4). This would enable the reader to judge what the dominant structural class is and what potential assembly intermediates or subclasses could be.
4. Please provide representative examples of single pores used for the analysis in the supplementary figures. One would assume that the average of all NPCs is perfectly round, yet the analysis shows that this is not the case. The authors should comment on this observation.

Reviewer #2 (Remarks to the Author):

In this study, Sabinina et al. have built used 3D SMLM to expand the our understanding of the NPC. They build up on previous ground breaking work from this group integrating many particle-averaging approaches from cryo-EM with specific tag localization to extract structural information well beneath the traditional power of superresolution microscopy. Using this, they expand the current model of the NPC to include several previously unresolved components, even confirming the proposed asymmetric distribution of ELYS and showing the highly dynamic nature of TPR, presumably to allow ELYS to bind to intended chromatin targets. They then take this one step farther using an unsupervised network classification to group individual NPCs based upon topological structure-remarkably identifying for the first time conformational dynamics of the entire NPC-truly a remarkable finding, in my opinion.

In general, I am extremely enthusiastic about this work-I am not an expert in the NPC so I am perhaps less qualified to comment on the novelty on that front, though to my uneducated perusing it seems to be a significant advance on the existing model. Regardless, the integration of unsupervised clustering to many particle averaging in SMLM in this way is a major advance for the SMLM field, and the work should be published based on that innovation alone. SMLM has largely been limited in utility within the field because of the inability to functionally average structures without defined symmetries or account for multiple states, this work marks the first major progress on that front in several years.

I have only one minor concern, which I hope the authors will address. While the paper is focused on the biological discoveries the authors have made, the implementation of unsupervised learning to cluster single NPCs is the most profound advance, in my opinion. It is a shame, then, that in this form the authors have not provided the reader with enough information to understand quantitatively how effective this was or to reproduce the technique in their own systems. I think the authors ought to address the following two things before publishing this work:

1. Figure 3 is not very useful to a reader, as it only gives a very qualitative sense for how the clustering worked. The authors should give us more data than this. Can we see quantification of how many Nups in each category were correctly grouped in the controls? How many were mis-targeted? This figure shows that it generally worked, but we cannot evaluate how well. (E.g it seems the TPR cluster has a lot of SEH1 in it? How much? And can we infer something from this?).
2. The authors need to expand on how "tree cutting" was performed. I imagine this was mostly done subjectively by eye, but if so that should be explicitly stated. It would also be nice to see how well the reconstructions work when the wrong number of groups are forced (i.e.-can you tell your number is representative of the data because of how well the fit converges or quantify the strain in the fit somehow?).

EMBL

EMBL Heidelberg · Meyerhofstr. 1 · 69117 Heidelberg · Germany

Jennifer Lippincott-Schwartz

Molecular Biology of the Cell

Jan Ellenberg
Head of Unit and Senior Scientist
Cell Biology & Biophysics Unit
T +49 6221 387 8328
F +49 6221 387 8512
Jan.ellenberg@embl.de
www.embl.org

EMBL Heidelberg
Meyerhofstraße 1
69117 Heidelberg
Germany

15. April 2021

Re-Submission of Jimenez Sabinina et al.

Dear Jennifer,

Thank you very much for your interest in publishing a revised version of our manuscript entitled "**3D super-resolution fluorescence microscopy maps the variable molecular architecture of the Nuclear Pore Complex**" by Vilma Jimenez Sabinina et al.

We very much appreciate the positive evaluation of our study by both reviewers and have implemented their very helpful suggestions for how to improve our study further in our revised manuscript. The revisions are described in detail in our point-by-point response attached to this letter.

In summary, we have included new figures and added more detail to the manuscript in order to assist the reader in navigating the data and make the clustering approach easier to understand. For example, we have added a new schematic figure to illustrate the position of all analyzed Nups and the particle averaging workflow (new Figure 1A) and have provided a confusion matrix to better support the performance of our clustering approach (new Figure 4B).

We very much hope that you will find our revised manuscript suitable for publication in Molecular Biology of the Cell and I am very much looking forward to your editorial decision.

All the best,

Jan

EMBL

Reviewer Point-by-Point response

Reviewer #1

Recent advancements in cryo-electron microscopy have made it possible to elucidate the molecular structure of a large number of protein complexes including the nuclear pore. This manuscript presents an alternative approach relying on 3D super-resolution light microscopy to map the position of several nucleoporins/Nups within the nuclear pore complex of human cells at nanoscale resolution. To achieve this, the authors expand a computational pipeline previously used for sub-tomogram averaging to extract individual NPCs from 3D SMLM volumes, averaging and aligning the signal of fluorescently tagged Nups relative to the reference protein ELYS. From this, they build a three-dimensional localization-density map for several Nups, outlining their average axial and radial positions within the NPC.
In a next step, the authors employ an unbiased classification method, by clustering individual NPC sub-volumes based on the topological features of Nup localization point clouds to capture the structural heterogeny within the NPC-population. Interestingly, for several Nups different clusters were found, distinguished by a single morphological feature, pointing to different conformational states of the mature NPC or potentially, NPC assembly intermediates.

The method described here is an exciting technical advancement validated by the observation that Nup localization patterns match well with already known or predicted positions. Furthermore, it provides new information of less well characterized positions of the asymmetric Nups ELYS and TRP. In addition, this approach could have the potential to capture the structural heterogeneity within native complex-populations, something classical averaging-based methods fail to do. Unfortunately, the manuscript falls short in clearly explaining essential aspects of the presented method and generally does a rather poor job at discussing and addressing associated technical limitations. In addition, the authors often do not provide enough evidence to empower the reader to make judgments on the conclusions but instead make sometimes bold statements that are difficult to verify.

We thank the reviewer for the careful assessment of our manuscript and for pointing out the technical advancements we made. We are also grateful that s/he alerts us to potential weaknesses which we have addressed as detailed below.

Major points:

1.) Conformational flexibility of ELYS. The authors chose the Nup ELYS as a reference. They do the averaging and alignment on the ELYS-signal and then propagate the transformations to the NUP channel that is examined. In such an experimental setup the resolution of the SNAP-tagged NUPs critically depend on the reference. One would assume that peripheral Nups are overall more prone to variability than the central ones and additionally the larger fluorophore displacement of the antibody labeling of ELYS will inevitably introduce more uncertainty. Do the authors know if the structured or the unstructured domain of ELYS is recognized by the ELYS-antibody? A lot of the observed variability might actually come from using the ELYS-antibody as reference rather than the respective NUPs? Why did the authors select ELYS and did not use one of the presumably more rigid Y-complex Nups as a reference?

As a reference we chose antibody labelling of endogenous Nups, to be universally applicable and comparable across our different homozygous knock-in SNAP tag cell lines. To select the most suitable antibody, we evaluated multiple anti-Nup antibodies for producing high labelling efficiency of clearly recognizable, ring-like structures with reproducible structural features such as ring diameter and regular subunit spacing. Based on these criteria, the anti ELYS antibody performed best compared to many other antibodies that we systematically evaluated previously (see e.g. Szymborska et al., 2013, Science 341:655-8). The ELYS antibody used was produced and validated by the Human Protein Atlas

EMBL

project and detailed information can be found here: https://www.proteinatlas.org/ENSG00000153207-AHCTF1/antibody#antigen_information. The antigen is located towards the C-terminus of the protein, between two domains, the ELYS domain and the C-terminal AT-hook domain. We have added this information to the methods section of the manuscript.

2.) Biological vs. technical variability. Throughout the manuscript it is difficult to judge, which observations are biological relevant, and which are beyond the precision limit of the approach. The measurements in the manuscript are given in single digit nanometer precision, however the fact that the individual spokes of the SNAP-tagged Y-complex Nups and the Nup133 rings cannot be resolved suggest a resolution >1nm. The article would greatly benefit from a paragraph in the beginning of the manuscript establishing the technical resolution limit of the approach and discussing potential drawbacks of the method, so the reader can judge the soundness of the observed differences.
The fact that the eightfold rotational symmetry of the NPC cannot be resolved for any of the NUPs in the averaged localization-density maps is insufficiently addressed. The authors hypothesize that this is like due to the structural variability however there is no evidence provided in support of this claim. In fact, the observation that the two Nup133 rings cannot be resolved would argue against it. It would be important to know how the localization precision of ~7 nm was determined, and the authors should show evidence where this is tested. It is also stated that the heterogeneity explores a larger range of distances than can be explained by technical variability. What is the evidence for this?

The localization precision of our single molecule localization microscopy was determined as described previously, using the open source SMAP software (Ries, 2020, Nat. Methods 17:870-872). In brief, the localization precision of our single molecule localization microscopy was determined as the Cramér-Rao Lower Bound (CRLB), calculated with our maximum likelihood fitting algorithm (Li, Y., Mund, M., Hoess, P., Deschamps, J., Matti, U., Nijmeijer, B., Sabinina, V. J., Ellenberg, J., Schoen, I. & Ries, J. Real-time 3D single-molecule localization using experimental point spread functions. Nature Methods 15, 367–369 (2018)). Consequently, images were rendered as localization histograms with a bin size of 7 nm. We have added the reference for the SMAP software to page 7 of the manuscript.
The eightfold rotational symmetry of all individual NUPs, with the exception of TPR is in fact recovered, when they are registered and averaged individually, independent of the reference. However, this is no longer the case when their registration is based on ELYS provided by the second fluorescence channel. Since preserving the relative spatial relationship of all Nups is necessary in order to build an integrative 3D map of the NPC, we used ELYS as common reference. The registration was therefore performed using ELYS and the same transformation was applied to the corresponding Nup of interest. We interpret the fact that with this method the structural features of individual Nups were lost, as evidence that their relative positions, at least towards ELYS, but potentially also towards each other, is variable. This is supported by our analysis of the heterogeneity of the structural features of each Nup, when averaged independently of a reference, which explore a range of several tens of nanometers, which is much larger than the precision of the super-resolution microscopy employed here. It is worth noting that previous studies have also reported variations in pore shape and size but that structural analyses have ignored these variations for the purpose of determining a single consensus structure based on averaging.

3.) Classification method does not seem very robust. There seem to be distinct clusters for Nup107 and maybe TPR but the others are not clearly dominated by a single NUP. Again, the authors do not provide a statistical analysis, or any other evidence to support the statement that the observed differences are based on structural diversity instead of technical variability. Furthermore, it is not obvious why Nup107 and Seh1, which show very similar double ring structures can be clearly distinguished, whereas Nup133 that has a single ring cannot be distinguished well from the other Y-complex double-ring Nups. Also, why do some of the classes have way fewer particles than other classes?

EMBL

Due to the demanding experimental workflow of single molecule localization microscopy, the complete data set is unfortunately not perfectly statistically balanced between the different Nups. In combination with the overall rather large size of the data set (over 10 000 particles), this indeed makes it challenging to evaluate the quality of the clustering. To support this better for the reader, we have now provided a confusion matrix (figure 4B) of our classification scheme based on persistence diagram clustering and have derived statistical measures to evaluate the ability of the clusters to predict labelled Nups. This analysis shows that the clustering performs well. The least well predicted Nup is SEH1, which is often misclassified as NUP133. The average cluster purity is 66% (ranging from 33-100%) and the overall accuracy of predicting Nups based on cluster membership is 53% which is significantly above the no-information rate (p-value < $2.2e^{-}16$). Average sensitivity and specificity are 42% (ranging from 13% for SEH1 to 71% for NUP107) and 88% (ranging from 76% for NUP133 to 97% for SEH1), respectively.

Because the topological signatures we used for clustering rely on all pairwise 3D distances, they capture structural aspects much beyond the presence of ring structures, which explains that two different double ring Nups can be separated. For example, it is expected that the size and shape of the rings would also be captured by the topological signatures and therefore influence the clustering. Such variability exists in the data as our measurements on individual particles for each Nup show for example that there is a range of circularities and some of the clusters reveal rings with partially filled-in central areas. While this structurally unbiased classification is very powerful, its limitation is that it does not allow an a priori interpretation of the clusters in terms of geometric features.

We have added a more detailed explanation of our approach to the manuscript on page 8.

4.) The weaker signal of the cytoplasmic ring. The fact that the cytoplasmic ring is weaker than the nucleoplasmic one needs to be addressed. One can see this both in the total averaged localization densities (Fig. 1) and especially pronounced in the sub-cluster densities (Fig. 4). Is this caused by the flexibility of using ELYS as a reference, causing the ring which is further away to blur out?

The weaker labelling of the cytoplasmic ring is independent of the ELYS reference, as it is also present in individual pores or average structures based only on the Nup of interest, independent of the reference. There appears to be systematically worse labelling efficiency of the cytoplasmic ring as compared to the nuclear ring in our data set. We have added this information to the methods section of the manuscript.

Minor points:

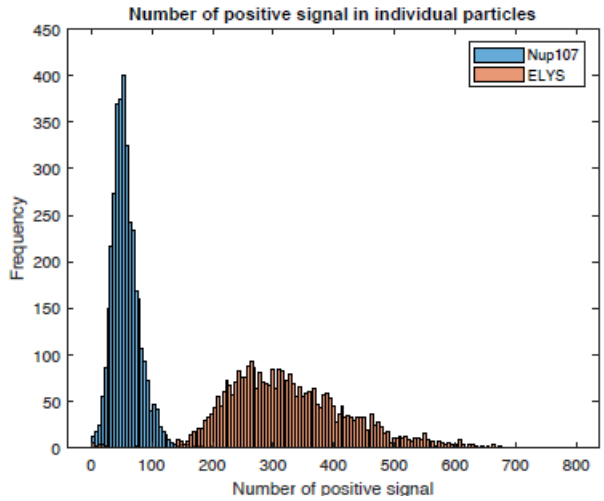
1. A schematic of the nuclear pore complex with an indication of the tagged components would help to make the manuscript more accessible for non-NPC specialists. Also, the workflow schematic (Fig 1A) should be made clearer.

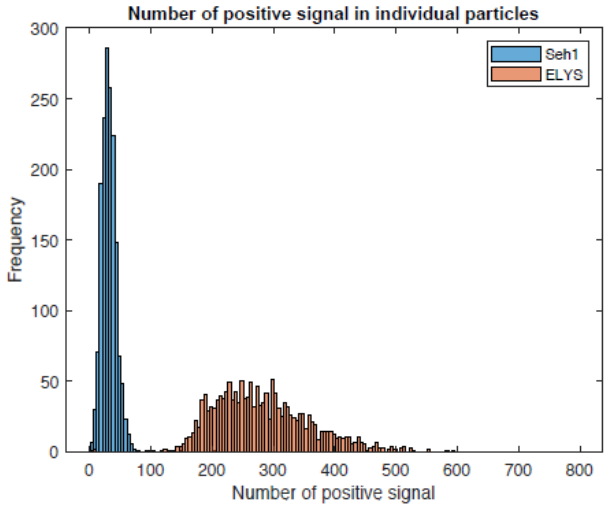
We've added a schematic representation of the pore indicating the localization of the studied proteins (Fig. 1A) and clarified the workflow representation (Fig. 1B)

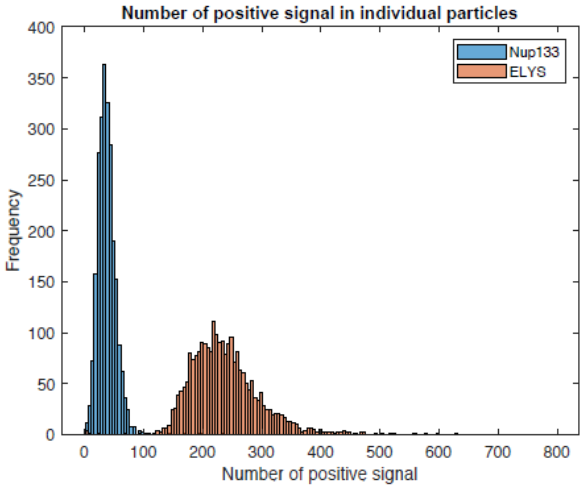
2. Could the signal intensity be used to find out if the thicker Nup133 ring actually stems from double the number of molecules?

We have carried out this analysis and in fact NUP133 has comparable signal intensity (number of localizations in each particle) as NUP107 and SEH1, as can be seen from the histograms below.

EMBL







3. The authors should indicate the number of particles in each of the subclasses (Fig 4). This would

EMBL

enable the reader to judge what the dominant structural class is and what potential assembly intermediates or subclasses could be.

We have added this information in a new table to the supplemental material as requested.

Table S1: Number of pores in each cluster. Numbers in parentheses indicate fractions of the total for the corresponding Nup. Totals are 862 pores for RANBP2, 3582 for NUP107 and 1585 for SEH1.

	cluster 1	cluster 2	cluster 3	cluster 4
RANBP2	747 (87%)	92 (11%)	n/a	n/a
NUP107	1282 (36%)	1603 (45%)	343 (10%)	252 (7%)
SEH1	737 (46%)	516 (33%)	228 (14%)	101 (6%)

4. Please provide representative examples of single pores used for the analysis in the supplementary figures. One would assume that the average of all NPCs is perfectly round, yet the analysis shows that this is not the case. The authors should comment on this observation.

We have added examples of individual pores in figure S1A.

The global averages in Fig. 2 (previously Fig. 1) show circular structures consistent with expectation. The plot in figure S1B shows that half of individual pores labelled on NUP107, SEH1 or NUP133 have a circularity below ~0.75 and RANBP2 and TPR have even lower median circularity. The averages for some of the clusters are not perfectly circular consistent with a partitioning based on structurally unbiased topological features that takes for example the shape into account.

EMBL

Reviewer #2

In this study, Sabinina et al. have built used 3D SMLM to expand the our understanding of the NPC. They build up on previous ground breaking work from this group integrating many particle-averaging approaches from cryo-EM with specific tag localization to extract structural information well beneath the traditional power of superresolution microscopy. Using this, they expand the current model of the NPC to include several previously unresolved components, even confirming the proposed asymmetric distribution of ELYS and showing the highly dynamic nature of TPR, presumably to allow ELYS to bind to intended chromatin targets. They then take this one step farther using an unsupervised network classification to group individual NPCs based upon topological structure-remarkably identifying for the first time conformational dynamics of the entire NPC-truly a remarkable finding, in my opinion.

In general, I am extremely enthusiastic about this work-I am not an expert in the NPC so I am perhaps less qualified to comment on the novelty on that front, though to my uneducated perusing it seems to be a significant advance on the existing model. Regardless, the integration of unsupervised clustering to many particle averaging in SMLM in this way is a major advance for the SMLM field, and the work should be published based on that innovation alone. SMLM has largely been limited in utility within the field because of the inability to functionally average structures without defined symmetries or account for multiple states, this work marks the first major progress on that front in several years.

We thank the reviewer for her/his enthusiastic words about the innovative nature of our work. We are also grateful that s/he alerts us to potential weaknesses which we have addressed as detailed below.

I have only one minor concern, which I hope the authors will address. While the paper is focused on the biological discoveries the authors have made, the implementation of unsupervised learning to cluster single NPCs is the most profound advance, in my opinion. It is a shame, then, that in this form the authors have not provided the reader with enough information to understand quantitatively how effective this was or to reproduce the technique in their own systems. I think the authors ought to address the following two things before publishing this work:

1. Figure 3 is not very useful to a reader, as it only gives a very qualitative sense for how the clustering worked. The authors should give us more data than this. Can we see quantification of how many Nups in each category were correctly grouped in the controls? How many were mis-targeted? This figure shows that it generally worked, but we cannot evaluate how well. (E.g it seems the TPR cluster has a lot of SEH1 in it? How much? And can we infer something from this?).

As the complete data set is unfortunately not perfectly statistically balanced between the different Nups and of overall rather large size (over 10 000 particles), this indeed makes it challenging to evaluate the quality of the clustering. To assist the reader in this, we are now providing the confusion matrix (figure 4B) of our classification scheme based on persistence diagram clustering and provide statistical measures to evaluate the ability of the clusters to predict labelled Nups. This analysis shows that the clustering performs well. The least well predicted Nup is SEH1, which is often misclassified as NUP133. The average cluster purity is 66% (ranging from 33-100%) and the overall accuracy of predicting Nups based on cluster membership is 53% which is significantly above the no-information rate (p-value < $2.2e^{-}16$). Average sensitivity and specificity are respectively 42% (ranging from 13% for SEH1 to 71% for NUP107) and 88% (ranging from 76% for NUP133 to 97% for SEH1).
The TPR region of the tree contains 13% of SEH1 with one subtree enriched up to 22%. The visual enrichment impression is likely due to the high contrast between the colours chosen to represent TPR and SEH1.

2. The authors need to expand on how "tree cutting" was performed. I imagine this was mostly done subjectively by eye, but if so that should be explicitly stated. It would also be nice to see how well the reconstructions work when the wrong number of groups are forced (i.e.-can you tell your number is

EMBL

representative of the data because of how well the fit converges or quantify the strain in the fit somehow?).

The global tree was cut at a fixed height, chosen to capture the smaller clusters produced due to the statistical imbalance of the data between individual Nups, which is caused by the experimental constraints of the demanding single molecule localization workflow. However, cluster purity was in fact stable around 65% over a wide range of tree cutting heights (from 50 to 500) and corresponding median cluster sizes (from 48 to a 434). To rule out that clustering is driven by a dominant Nup, we performed clustering of subsets of the data balanced between individual Nups, which yielded similar results. The extracted subtrees were arbitrarily cut to generate five clusters to ensure a sufficient number of particles in each cluster for subsequent particle averaging to explore their structural features.

Our clustering strategy cuts across the tree at a fixed height. From a structural similarity point of view this could probably be further improved by analysing clusters individually with orthogonal methods to topological clustering for similarity, or by detailed visual inspection. This could likely produce even better results but is currently difficult to automate or perform in an unbiased manner.

We have now added a more detailed description of the clustering methods to the manuscript on page 8.

RE: Manuscript #E20-11-0728R
TITLE: "3D super-resolution fluorescence microscopy maps the variable molecular architecture of the Nuclear Pore Complex"

Dear Jan,

Congratulations on excellent work. We are happy to accept your paper for publication.

Sincerely,
Jennifer Lippincott-Schwartz
Monitoring Editor
Molecular Biology of the Cell

Dear Dr. Ellenberg:

Congratulations on the acceptance of your manuscript.

A PDF of your manuscript will be published on MBoC in Press, an early release version of the journal, within 10 days. The date your manuscript appears at www.molbiolcell.org/toc/mboc/0/0 is the official publication date. Your manuscript will also be scheduled for publication in the next available issue of MBoC.

Within approximately four weeks you will receive a PDF page proof of your article.

Would you like to see an image related to your accepted manuscript on the cover of MBoC? Please contact the MBoC Editorial Office at mboc@ascb.org to learn how to submit an image.

Authors of Articles and Brief Communications are encouraged to create a short video abstract to accompany their article when it is published. These video abstracts, known as Science Sketches, are up to 2 minutes long and will be published on YouTube and then embedded in the article abstract. Science Sketch Editors on the MBoC Editorial Board will provide guidance as you prepare your video. Information about how to prepare and submit a video abstract is available at www.molbiolcell.org/science-sketches. Please contact mboc@ascb.org if you are interested in creating a Science Sketch.

We are pleased that you chose to publish your work in MBoC.

Sincerely,

Eric Baker
Journal Production Manager
MBoC Editorial Office
mbc@ascb.org
