nature research

Corresponding author(s): Peijun Zhang

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see Reporting Life Sciences Research. For further information on Nature Research policies, including our data availability policy, see Authors & Referees and the Editorial Policy Checklist.

Please do not complete any field with "not applicable" or n/a. Refer to the help text for what text to use if an item is not relevant to your study.
For final submission: please carefully check your responses for accuracy; you will not be able to make changes later.

Experimental design

1. Sample size

Describe how sample size was determined.

Our sample size was determined by those available in the two public data sets we chose to use for the evaluation of our methods.

2. Data exclusions

Describe any data exclusions.

No data were excluded from the analysis. The contribution of each sub-tomogram with a cross-correlation lower than the mean value of all cross-correlation scores is down weighted as previously described.

3. Replication

Describe the measures taken to verify the reproducibility of the experimental findings.

In addition to replication on various hardware configurations with variable image processing parameters by the first author, we also had a Postdoc in the group with knowledge of sub-tomogram averaging reproduce the results to a satisfactory degree.

4. Randomization

Describe how samples/organisms/participants were allocated into experimental groups.

The data were split into two independent (beyond 40 Angstrom resolution) half-sets randomly at the beginning of the sub-tomogram workflow.

5. Blinding

Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

While in principle the group identity of the individual sub-tomograms could be observed in the experimental meta data, the investigator, however, choose to be blinded to the group allocation during the process, and the software enforces separation at all stages prior to the final averaging.

Note: all in vivo studies must report how sample size was determined and whether blinding and randomization were used.

6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

n/a	Confirmed	
☒	☐	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
☒	☐	A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐	A statement indicating how many times each experiment was replicated
☒	☐	The statistical test(s) used and whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of any assumptions or corrections, such as an adjustment for multiple comparisons
☐	☒	Test values indicating whether an effect is present *Provide confidence intervals or give results of significance tests (e.g. P values) as exact values whenever appropriate and with effect sizes noted.*
☒	☐	A clear description of statistics including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
☒	☐	Clearly defined error bars in all relevant figure captions (with explicit mention of central tendency and variation)

See the web collection on statistics for biologists for further resources and guidance.

Software

Policy information about availability of computer code

7. Software

Describe the software used to analyze the data in this study.	The primary software used for data analysis is "emClarity version 1.0.0" which is the focus of the manuscript. Components of the "IMOD version 4.8.43" software from the University of Colorado at Boulder as well as "Chimera version 11.1.2" from the University of California San Fransisco were used for both data analysis and visualization.

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). *Nature Methods* guidance for providing algorithms and software for publication provides further information on this topic.

Materials and reagents

Policy information about availability of materials

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a third party.	We have used data that are freely available from the Electron Microscopy Pilot Image Archive (EMPIAR) https://www.ebi.ac.uk/pdbe/emdb/empiar/ with specific information detailed in the manuscript.

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).	No antibodies were used in this study.

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.	We did not make use of any eukaryotic cell lines in this study.
b. Describe the method of cell line authentication used.	No cell lines were used in this study.
c. Report whether the cell lines were tested for mycoplasma contamination.	No cell lines were used in this study.
d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by ICLAC, provide a scientific rationale for their use.	No cell lines were used in this study.

Animals and human research participants

Policy information about studies involving animals; when reporting animal research, follow the ARRIVE guidelines

11. Description of research animals

Provide all relevant details on animals and/or animal-derived materials used in the study.	No animals were used in this study.

Policy information about studies involving human research participants

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.	No human research patients were used in this study.