nature research

Corresponding author(s): David Barford

Last updated by author(s): Aug 22, 2019

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☒	☐ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☒	☐ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐ 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 all covariates tested
☒	☐ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☒	☐ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Commercial software: EPU from Thermo Fisher Scientific was used for automated cryo-EM data collection.
Data analysis	Cryo-EM data were analyzed using the software MotionCor2 (version 2.1), GCTF (version 0.5), RELION2.1 (version 2.1), RELION3.0 (version 3.0), SIMPLEPRIME (version) and RESMAP (version 1.1.4). Model building and refinement were performed using COOT (version 0.8.9.2) and Phenix (version 1.15.2) and validated in COOT (version 0.8.9.2) and MolProbity (version 4.2). Visualization was performed with COOT (version 0.8.9.2), PyMOL (version 1.8.4.1, Chimera (version 1.8.1) and ChimeraX (version 0.8). Structure figures were generated use PyMOL (v1.8.4.1) and Chimera (version 1.8.1). Sequence alignments were performed and displayed with JALVIEW (version 1.0). Structure prediction was performed with the PHYRE2 web tool. Protein secondary structure and disordered regions were predicted with the PHYRE2 and PSIPred web tools. AUC data were analysed in SEDFIT v16.1 and Sednterp (version 20130813 beta). SEC-MALS data were analysed using ASTRA version 5.3.4.20 software (Wyatt Technologies) and data were plotted with the program PRISM (version 8.2.0) (GraphPad Software Inc.). XL-MS data were analysed using Proteome Discoverer 2.3 (version 2.3.0.522) and the xVis web tool.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

EM maps are deposited with EMDB with accession codes EMD-4580 (CCAN), EMD-4579 (CCAN-Cenp-ANuc), EMD-4581 (Mask1) and EMD-4971 (Mask2). Protein coordinates are deposited with RCSB with accession codes 6QLE (CCAN), 6QLD (CCAN-Cenp-ANuc) and 6QLF (Mask1). The cross-linking mass spectrometry raw files, the associated output and databases are deposited through the ProteomeXchange Consortium 48 via the PRIDE partner repository with the dataset identifier PXD013769.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	We collected 9002 cryo-EM images for the CCAN-Cenp-A dataset and 910 cryo-EM images for the Cenp-HIK dataset. The total number of particles for the CCAN-Cenp-A dataset was 1,796,061 and that for the Cenp-HIK dataset was 123,215. Sample sizes were estimated on the basis of previous studies using similar methods and analyses that are widely published.
Data exclusions	No data were excluded from the analysis.
Replication	All attempts at replication were successful and reproducible. At least three independent biological repeats per experiment where representative data are shown. Structure determination does not require replication.
Randomization	Samples were not allocated into groups. Randomization is not relevant to this study.
Blinding	Blinding was not relevant to this study because there was no group allocation.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	One primary antibody used: MOUSE ANTI STREP-TAG CLASSIC:HRP. Anti-Strep antibody (Source: Bio-Rad, Catalogue code:MCA2489P, Batch No: 147517). Dilution 1 to 1000.
Validation	Mouse anti Strep-Tag Classic antibody, clone Strep-tag II , also known as StrepMAB-Classic, recognizes Strep-tag II, a widely used tag in protein expression applications. This antibody recognizes both-C- and N-terminal Strep-tag II and is especially suited to Western blot applications. Validation: HCA182 specificity ELISA using various antigens (A: Human Serum, B: human IgG1/kappa from myeloma plasma, C: Rituximab, D: Ustekinumab, E: Infliximab, F: Adalimumab, G: Alemtuzumab and H: Bevacizumab) as coating components

followed by Human anti Avastin®(HCA182) and HRP conjugated Mouse anti Strep-tag (MCA2489P) as detection reagent.

Literature on Bio-Rad web-site:

1. Renzi F et al. (2015) Glycan-Foraging Systems Reveal the Adaptation of Capnocytophaga canimorsus to the Dog Mouth. MBio. 6 (2): .pii: e02507-14.
2. Gordon, C.A. et al. (2015) NUSAP1 expression is upregulated by loss of RB1 in prostate cancer cells. Prostate. 75 (5): 517-26.
3. Mavrakis, M. et al. (2016) Purification of recombinant human and Drosophila septin hexamers for TIRF assays of actin-septin filament assembly. Methods Cell Biol. 136: 199-220.
4. Oda, S. et al. (2015) Crystal Structure of Marburg Virus VP40 Reveals a Broad, Basic Patch for Matrix Assembly and a Requirement of the N-Terminal Domain for Immunosuppression. J Virol. 90 (4): 1839-48.
5. Renzi, F. et al. (2015) Glycan-foraging systems reveal the adaptation of Capnocytophaga canimorsus to the dog mouth. MBio. 6 (2): e02507.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	High-5 insect cell: Trichoplusia ni: expression system. S. cerevisiae strains: 1. AEY4992: with a chl4 deletion and cse4-R37A mutation (chl4Δ cse4-R37A), (MATα ade2-101 lys2 his3-11,15 trp1-1 leu2-3,112 ura3-1 can1-100 chl4Δ::kanMX Cse4R37A). 2. W303 (wild type strain: MATα ade2-101 his3-11,15 trp1-1 leu2-3,112 ura3-1).
Authentication	The High-5 insect cell line was not authenticated. The S. cerevisiae strains were authenticated refs 27, 41.
Mycoplasma contamination	The High-5 insect cell line was not tested for mycoplasma contamination. Yeast strains do not have mycoplasma and were not tested for mycoplasma contamination
Commonly misidentified lines (See ICLAC register)	None