nature research

Corresponding author(s): Janne Lehtiö, Lukas M. Orre

Last updated by author(s): Dec 15, 2021

# 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 our Editorial Policies 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<br>*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<br>*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 | Proteomics RAW data were collected by LC-MS analysis (Thermo Fisher Scientific, Orbitrap model, QE and QE-HF), the protein/peptide data were generated by proteowizard from msConvert (v.3.0.19127), MSGF+ (v10072), Percolator (v2.08) and IsobaricAanlyzer (v2.0) |
| Data analysis | Software for SubCellBarCode R-package (SubCellBarCode_1.4.0 ) analysis: any operating system Windows, Linux, OSX (Version 10.12.6 or higher); R (version 3.6 or later), Bioconductor (3.9 or later).<br>The package and package versions used in the method are given in Box 1 in the main manuscript. |

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 and 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

The mass spectrometry proteomics data for the analysis of the HeLa cell line have been deposited to the ProteomeXchange Consortium via the jPOST partner repository with the dataset identifier PXD022533.
The SubCellBarCode R-package and manual is freely available through the Bioconductor repository (doi: 10.18129/B9.bioc.SubCellBarCode).
The mass spectrometry proteomics data for the previous analysis of five different cell lines have been deposited to the ProteomeXchange Consortium via the PRIDE

partner repository with the dataset identifier PXD006895.
All described processed datasets can be easily accessed, explored and visualized at the SubCellBarCode portal (https://www.subcellbarcode.org).

# 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 | this study only focused on HeLa cell line. |
| Data exclusions | no data exclusions in this study |
| Replication | this study has duplicate subcellular fractionation for HeLa cell line. |
| Randomization | not relevant to this study |
| Blinding | not relevant to this study |

# 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 and archaeology |
| ☒ | ☐ Animals and other organisms |
| ☒ | ☐ Human research participants |
| ☒ | ☐ Clinical data |
| ☒ | ☐ Dual use research of concern |

## Methods

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

## Antibodies

| | |
|---|---|
| Antibodies used | 1, anti-GAPDH, cat. no. sc-69778 (Santa Cruz Biotechnology), Mouse mAb, RRID: AB_1124759 (https://scicrunch.org/resolver/RRID:AB_1124759)<br>2, anti-ATP1B1, cat. no. D8W8J (Cell Signaling Technology), Rabbit mAb, RRID: AB_2799633 (https://scicrunch.org/resolver/RRID:AB_2799633)<br>3, anti-Histone H4, cat. no. sc-8658-R (Santa Cruz Biotechnology), Rabbit pAb, RRID:AB_2011538 (https://scicrunch.org/resolver/RRID:AB_2011538)<br>4, anti-TOM20, cat. no. sc-11415 (Santa Cruz Biotechnology), Rabbit pAb, RRID:AB_2207533 (https://scicrunch.org/resolver/RRID:AB_2207533) |
| Validation | Antibodies were validated according to the manufacturers' instructions |

## Eukaryotic cell lines

Policy information about cell lines

| | |
|---|---|
| Cell line source(s) | HeLa (CCL-2, ATCC), https://scicrunch.org/resolver/RRID:CVCL_0030 |
| Authentication | HeLa cell line was authenticated by Eurofins Genomics |
| Mycoplasma contamination | The cell line was mycoplasma-free, which is tested by MycoAlert mycoplasma detection kit (Lonza, Walkersville, MD, USA) |

| | |
|---|---|
| Commonly misidentified lines<br>(See ICLAC register) | No commonly misidentified cell lines were used in this study. |