

Reporting Summary

Nature Portfolio 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 Portfolio 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 (<i>n</i>) 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 <i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i>
☒	☐ 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. <i>F</i> , <i>t</i> , <i>r</i>) with confidence intervals, effect sizes, degrees of freedom and <i>P</i> value noted <i>Give P values as exact values whenever suitable.</i>
☒	☐ 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 <i>d</i> , Pearson's <i>r</i>), 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	The code to compile, access and process the Human Protein Atlas data can be found at https://github.com/uhlerlab/PUPS/ .
Data analysis	Our code is deposited in Github along with a list of software dependencies: https://github.com/uhlerlab/PUPS/. The following python packages are used: <pre>_libgcc_mutex=0.1=main _openmp_mutex=5.1=1_gnu _tflow_select=2.3.0=mkl abseil-cpp=20211102.0=hd4dd3e8_0 absl-py=1.4.0=py39h06a4308_0 aiohttp=3.8.3=py39h5eee18b_0 aiosignal=1.2.0=pyhd3eb1b0_0 anyio=3.7.0=pypi_0 appdirs=1.4.4=pyhd3eb1b0_0 arrow=1.2.3=pypi_0 asciitree=0.3.3=pypi_0 asttokens=2.2.1=pyhd8ed1ab_0 astunparse=1.6.3=py_0 async-timeout=4.0.2=py39h06a4308_0 attrs=22.1.0=py39h06a4308_0 backcall=0.2.0=pyh9f0ad1d_0 backports=1.0=pyhd8ed1ab_3</pre>

```

backports.functools_lru_cache=1.6.4=pyhd8ed1ab_0
beautifulsoup4=4.12.2=pypi_0
bio=1.5.3=pypi_0
biopython=1.81=pypi_0
biothings-client=0.3.0=pypi_0
biotite=0.34.1=pypi_0
blas=1.0=mkl
blessed=1.20.0=pypi_0
blinker=1.4=py39h06a4308_0
brotlipy=0.7.0=py39h27cfd23_1003
bzip2=1.0.8=h7b6447c_0
c-ares=1.19.0=h5eee18b_0
ca-certificates=2023.5.7=hbcca054_0
cachetools=4.2.2=pyhd3eb1b0_0
certifi=2023.5.7=pyhd8ed1ab_0
cffi=1.15.1=py39h74dc2b5_0
charset-normalizer=2.0.4=pyhd3eb1b0_0
click=8.0.4=py39h06a4308_0
cloudpickle=2.2.1=pypi_0
contourpy=1.1.0=pypi_0
croniter=1.3.15=pypi_0
cryptography=39.0.1=py39h9ce1e76_0
csbdeep=0.7.4=pypi_0
cudatoolkit=11.1.1=ha002fc5_10
cycler=0.11.0=pypi_0
dateutils=0.6.12=pypi_0
debugpy=1.5.1=py39h295c915_0
decorator=5.1.1=pyhd8ed1ab_0
deepdiff=6.3.0=pypi_0
dnspython=2.3.0=pypi_0
docopt=0.6.2=pypi_0
ensembl-rest=0.3.3=pypi_0
entrypoints=0.4=pyhd8ed1ab_0
et-xmlfile=1.1.0=pypi_0
exceptiongroup=1.1.1=pypi_0
executing=1.2.0=pyhd8ed1ab_0
fair-esm=2.0.0=pypi_0
fastapi=0.88.0=pypi_0
fasteners=0.19=pypi_0
fastjsonschema=2.17.1=pypi_0
ffmpeg=4.3=hf484d3e_0
filelock=3.9.0=py39h06a4308_0
flatbuffers=2.0.0=h2531618_0
fonttools=4.40.0=pypi_0
freetype=2.12.1=h4a9f257_0
frozenset=1.3.3=py39h5eee18b_0
fsspec=2023.6.0=pypi_0
gast=0.4.0=pyhd3eb1b0_0
giflib=5.2.1=h5eee18b_3
gmp=6.2.1=h295c915_3
gmpy2=2.1.2=py39heeb90bb_0
gnutls=3.6.15=he1e5248_0
goatools=1.3.11=pypi_0
google-auth=2.6.0=pyhd3eb1b0_0
google-auth-oauthlib=0.4.4=pyhd3eb1b0_0
google-pasta=0.2.0=pyhd3eb1b0_0
grpc-cpp=1.46.1=h33aed49_1
grpcio=1.42.0=py39hce63b2e_0
h11=0.14.0=pypi_0
h5py=3.7.0=py39h737f45e_0
hdf5=1.10.6=h3ffc7dd_1
hiddenlayer=0.3=pypi_0
icu=58.2=he6710b0_3
idna=3.4=py39h06a4308_0
imageio=2.31.1=pypi_0
imaris-ims-file-reader=0.1.8=pypi_0
importlib-metadata=6.0.0=py39h06a4308_0
importlib-resources=5.12.0=pypi_0
iniconfig=2.0.0=pypi_0
inquirer=3.1.3=pypi_0
intel-openmp=2021.4.0=h06a4308_3561
intervaltree=3.1.0=pypi_0
ipykernel=6.15.0=pyh210e3f2_0
ipython=8.14.0=pyh41d4057_0
ipython-genutils=0.2.0=pypi_0
ipywidgets=8.0.6=pypi_0

```

itsdangerous=2.1.2=pypi_0
 jedi=0.18.2=pyhd8ed1ab_0
 jinja2=3.1.2=py39h06a4308_0
 joblib=1.2.0=pypi_0
 jpeg=9e=h5eee18b_1
 jsonschema=4.17.3=pypi_0
 jupyter_client=7.0.6=pyhd8ed1ab_0
 jupyter_core=4.12.0=py39hf3d152e_0
 jupyterlab-widgets=3.0.7=pypi_0
 keras=2.11.0=py39h06a4308_0
 keras-preprocessing=1.1.2=pyhd3eb1b0_0
 kiwisolver=1.4.4=pypi_0
 krb5=1.20.1=h568e23c_1
 lame=3.100=h7b6447c_0
 lcms2=2.12=h3be6417_0
 ld_impl_linux-64=2.38=h1181459_1
 lerc=3.0=h295c915_0
 libblas=3.9.0=12_linux64_mkl
 libcbblas=3.9.0=12_linux64_mkl
 libclang=16.0.0=pypi_0
 libcurl=8.1.1=h91b91d3_1
 libdeflate=1.17=h5eee18b_0
 libedit=3.1.20221030=h5eee18b_0
 libev=4.33=h7f8727e_1
 libffi=3.3=he6710b0_2
 libgcc-ng=11.2.0=h1234567_1
 libgfortran-ng=11.2.0=h00389a5_1
 libgfortran5=11.2.0=h1234567_1
 libgomp=11.2.0=h1234567_1
 libiconv=1.16=h7f8727e_2
 libidn2=2.3.4=h5eee18b_0
 liblapack=3.9.0=12_linux64_mkl
 libnghttp2=1.52.0=ha637b67_1
 libpng=1.6.39=h5eee18b_0
 libprotobuf=3.20.3=he621ea3_0
 libsodium=1.0.18=h36c2ea0_1
 libssh=1.10.0=h37d81fd_2
 libstdcxx-ng=11.2.0=h1234567_1
 libtasn1=4.19.0=h5eee18b_0
 libtiff=4.5.0=h6a678d5_2
 libunistring=0.9.10=h27cfd23_0
 libwebp=1.2.4=h11a3e52_1
 libwebp-base=1.2.4=h5eee18b_1
 lightning=2.0.1=pypi_0
 lightning-cloud=0.5.37=pypi_0
 lightning-utilities=0.8.0=pypi_0
 llvmlite=0.40.1=pypi_0
 lz4-c=1.9.4=h6a678d5_0
 markdown=3.4.1=py39h06a4308_0
 markdown-it-py=3.0.0=pypi_0
 markupsafe=2.1.1=py39h7f8727e_0
 matplotlib=3.7.1=pypi_0
 matplotlib-inline=0.1.6=pyhd8ed1ab_0
 mdurl=0.1.2=pypi_0
 mkl=2021.4.0=h06a4308_640
 mkl-service=2.4.0=py39h7f8727e_0
 mkl_fft=1.3.1=py39hd3c417c_0
 mkl_random=1.2.2=py39h51133e4_0
 mpc=1.1.0=h10f8cd9_1
 mpfr=4.0.2=hb69a4c5_1
 mpmath=1.2.1=py39h06a4308_0
 msgpack=1.0.5=pypi_0
 multidict=6.0.2=py39h5eee18b_0
 mygene=3.2.2=pypi_0
 nbformat=5.9.0=pypi_0
 ncurses=6.4=h6a678d5_0
 nest-asyncio=1.5.6=pyhd8ed1ab_0
 nettle=3.7.3=hbbd107a_1
 networkx=3.1=pypi_0
 numba=0.57.1=pypi_0
 numcodecs=0.12.1=pypi_0
 numpy=1.23.3=pypi_0
 nvidia-cublas-cu11=11.10.3.66=pypi_0
 nvidia-cuda-nvrtc-cu11=11.7.99=pypi_0
 nvidia-cuda-runtime-cu11=11.7.99=pypi_0
 nvidia-cudnn-cu11=8.5.0.96=pypi_0

oauthlib=3.2.2=py39h06a4308_0
 opencv-python=4.8.0.74=pypi_0
 openh264=2.1.1=h4ff587b_0
 openpyxl=3.1.2=pypi_0
 openssl=1.1.1u=h7f8727e_0
 opt_einsum=3.3.0=pyhd3eb1b0_1
 ordered-set=4.1.0=pypi_0
 packaging=23.1=pyhd8ed1ab_0
 pandas=1.4.4=pypi_0
 parso=0.8.3=pyhd8ed1ab_0
 patsy=0.5.6=pypi_0
 pexpect=4.8.0=pyh1a96a4e_2
 pickleshare=0.7.5=py_1003
 pillow=9.5.0=pypi_0
 pip=23.1.2=py39h06a4308_0
 plotly=5.15.0=pypi_0
 pluggy=1.5.0=pypi_0
 pooch=1.4.0=pyhd3eb1b0_0
 prompt-toolkit=3.0.38=pyha770c72_0
 prompt_toolkit=3.0.38=hd8ed1ab_0
 propy3=1.1.1=pypi_0
 protobuf=3.19.6=pypi_0
 psutil=5.9.0=py39h5eee18b_0
 ptyprocess=0.7.0=pyhd3deb0d_0
 pure_eval=0.2.2=pyhd8ed1ab_0
 pyasn1=0.4.8=pyhd3eb1b0_0
 pyasn1-modules=0.2.8=py_0
 pycparser=2.21=pyhd3eb1b0_0
 pydantic=1.10.9=pypi_0
 pydot=2.0.0=pypi_0
 pygments=2.15.1=pyhd8ed1ab_0
 pyjwt=2.4.0=py39h06a4308_0
 pymongo=4.3.3=pypi_0
 pynndescent=0.5.10=pypi_0
 pyopenssl=23.0.0=py39h06a4308_0
 pyparsing=3.1.0=pypi_0
 pyrsistent=0.19.3=pypi_0
 pysocks=1.7.1=py39h06a4308_0
 pytest=8.2.2=pypi_0
 python=3.9.13=haa1d7c7_2
 python-dateutil=2.8.2=pyhd8ed1ab_0
 python-editor=1.0.4=pypi_0
 python-flatbuffers=2.0=pyhd3eb1b0_0
 python-graphviz=0.20.3=pypi_0
 python-multipart=0.0.6=pypi_0
 python_abi=3.9=2_cp39
 pytorch=2.0.1=py3.9_cpu_0
 pytorch-lightning=2.0.4=pypi_0
 pytorch-mutex=1.0=cpu
 pytz=2023.3=pypi_0
 pywavelets=1.4.1=pypi_0
 pyyaml=6.0=pypi_0
 pyzmq=19.0.2=py39hb69f2a1_2
 re2=2022.04.01=h295c915_0
 readchar=4.0.5=pypi_0
 readline=8.2=h5eee18b_0
 requests=2.28.1=pypi_0
 requests-oauthlib=1.3.0=py_0
 rich=13.4.2=pypi_0
 rsa=4.7.2=pyhd3eb1b0_1
 scienceplots=2.1.1=pypi_0
 scikit-image=0.19.3=pypi_0
 scikit-learn=1.2.2=pypi_0
 scipy=1.9.1=pypi_0
 seaborn=0.12.1=pypi_0
 setuptools=67.8.0=py39h06a4308_0
 shap=0.42.1=pypi_0
 simplejson=3.19.1=pypi_0
 six=1.16.0=pyhd3eb1b0_1
 sklearn=0.0.post5=pypi_0
 slicer=0.0.7=pypi_0
 snappy=1.1.9=h295c915_0
 sniffio=1.3.0=pypi_0
 sortedcontainers=2.4.0=pypi_0
 soupsieve=2.4.1=pypi_0
 sqlite=3.41.2=h5eee18b_0

```

stack_data=0.6.2=pyhd8ed1ab_0
stardist=0.8.5=pypi_0
starlette=0.22.0=pypi_0
starsessions=1.3.0=pypi_0
statsmodels=0.14.1=pypi_0
sympy=1.11.1=py39h06a4308_0
tenacity=8.2.2=pypi_0
tensorboard=2.11.0=py39h06a4308_0
tensorboard-data-server=0.6.1=py39h52d8a92_0
tensorboard-plugin-wit=1.8.1=py39h06a4308_0
tensorflow=2.11.0=mkl_py39hea27e05_0
tensorflow-base=2.11.0=mkl_py39he5f8e37_0
tensorflow-estimator=2.11.0=py39h06a4308_0
tensorflow-io-gcs-filesystem=0.32.0=pypi_0
termcolor=2.1.0=py39h06a4308_0
threadpoolctl=3.1.0=pypi_0
tifffile=2023.4.12=pypi_0
tk=8.6.12=h1ccaba5_0
tomli=2.0.1=pypi_0
torch=1.13.0=pypi_0
torchaudio=2.0.2=py39_cpu
torchlens=0.1.15=pypi_0
torchmetrics=0.11.4=pypi_0
torchrecorder=1.0.3=pypi_0
torchsummary=1.5.1=pypi_0
torchvision=0.14.0=pypi_0
torchviz=0.0.2=pypi_0
tornado=6.1=py39hb9d737c_3
tqdm=4.64.1=pypi_0
traitlets=5.9.0=pyhd8ed1ab_0
typing_extensions=4.6.3=py39h06a4308_0
tzdata=2023c=h04d1e81_0
umap-learn=0.5.2=pypi_0
uniprot=1.3=pypi_0
urllib3=1.26.16=py39h06a4308_0
uvicorn=0.22.0=pypi_0
wcwidth=0.2.6=pyhd8ed1ab_0
websocket-client=1.6.0=pypi_0
websockets=11.0.3=pypi_0
werkzeug=2.2.3=py39h06a4308_0
wheel=0.38.4=py39h06a4308_0
widgetsnextension=4.0.7=pypi_0
wrapt=1.14.1=py39h5eee18b_0
xlsxwriter=3.1.9=pypi_0
xz=5.4.2=h5eee18b_0
yarl=1.8.1=py39h5eee18b_0
zarr=2.17.0=pypi_0
zeromq=4.3.4=h9c3ff4c_1
zipp=3.11.0=py39h06a4308_0
zlib=1.2.13=h5eee18b_0
zstd=1.5.5=hc292b87_0

```

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 Portfolio [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 description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

All images from the Human Protein Atlas (HPA)3 version 16.0 to 22.0 can be downloaded from the HPA website (<https://www.proteinatlas.org/humanproteome/subcellular>).

The images generated from the experimental validation experiments are available in Zenodo at <https://doi.org/10.5281/zenodo.15059321>.

Human research participants

Policy information about [studies involving human research participants and Sex and Gender in Research](#).

Reporting on sex and gender

N/A

Population characteristics

N/A

Recruitment

N/A

Ethics oversight

N/A

Note that full information on the approval of the study protocol must also be provided in the manuscript.

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

For each pair of cell line and protein in our analysis, we used all available data in the Human Protein Atlas. The number of images used for each cell line and protein pair are listed in Supplementary Tables 1-3. No sample size calculation was performed for the experimental validation and, instead, we obtained a sample size that is comparable to the HPA for each cell line and protein pair. For the experimental validation, six field of views of 205.8×205.8 µm were acquired for each cell line and HPA antibody combination, which is comparable to the sample size in HPA.

Data exclusions

No data from the experimental validation was excluded.

Replication

One validation experiment was performed for 9 proteins on 5 cell lines. The results of the validation experiments agreed with the predictions of our computational model. No further replication was performed.

Randomization

We did not separate samples into experimental groups.

Blinding

We did not separate samples into experimental groups.

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
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

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

Antibodies

Antibodies used

EIF4G1 (Atlas Antibodies, HPA028487, poly)
DDIT3 (Atlas Antibodies, HPA058416, poly)
N4BP2 (Atlas Antibodies, HPA042607, poly)

Validation

CHID1 (Atlas Antibodies, HPA039374, poly)
 COPA (Atlas Antibodies, HPA028024, poly)
 ATP13A5 (Atlas Antibodies, HPA031772, poly)
 MESD (MESDC2) (Atlas Antibodies, HPA039414, poly)
 RBM23 (Atlas Antibodies, HPA004144, poly)
 PSME3IP1 (FAM192A) (Atlas Antibodies, HPA054382, poly)
 a-Tubulin (Abcam, ab7291, DM1A)
 calreticulin (Abcam, ab2908, poly)
 Anti-rabbit Alexa488 (Thermo, A11034, poly)
 Anti-mouse Alexa555 (Thermo, A21424, poly)
 Anti-chicken Alexa647 (Thermo, A21449, poly)

Primary validations were performed by the respective companies that produces these antibodies, which can be found on the manufacturers' website. In additionally, the 9 primary antibody we used were extensively validated by Human Protein Atlas (HPA) and available online:
 EIF4G1 (Atlas Antibodies, HPA028487, poly), validated on human by ICC, IHC, WB and PA: <https://www.proteinatlas.org/ENSG00000114867-EIF4G1/summary/antibody>
 DDIT3 (Atlas Antibodies, HPA058416, poly), validated on human by ICC, and PA: <https://www.proteinatlas.org/ENSG00000175197-DDIT3/summary/antibody>
 N4BP2 (Atlas Antibodies, HPA042607, poly), validated on human by ICC, and PA: <https://www.proteinatlas.org/ENSG00000078177-N4BP2/summary/antibody>
 CHID1 (Atlas Antibodies, HPA039374, poly), validated on human by IHC, and WB: <https://www.proteinatlas.org/ENSG00000177830-CHID1/summary/antibody>
 COPA (Atlas Antibodies, HPA028024, poly), validated on human by ICC, IHC, WB and PA: <https://www.proteinatlas.org/ENSG00000122218-COPA/summary/antibody>
 ATP13A5 (Atlas Antibodies, HPA031772, poly), validated on human by ICC, and PA: <https://www.proteinatlas.org/ENSG00000187527-ATP13A5/summary/antibody>
 MESD (MESDC2) (Atlas Antibodies, HPA039414, poly), validated on human by IHC, WB and PA: <https://www.proteinatlas.org/ENSG00000117899-MESD/summary/antibody>
 RBM23 (Atlas Antibodies, HPA004144, poly), validated on human by ICC, IHC and PA: <https://www.proteinatlas.org/ENSG00000100461-RBM23/summary/antibody>
 PSME3IP1 (FAM192A) (Atlas Antibodies, HPA054382, poly), validated on human by ICC, IHC and PA: <https://www.proteinatlas.org/ENSG00000172775-PSME3IP1/summary/antibody>

The other primary antibodies and 2nd antibodies validation can be found on the respective companies' website:

a-Tubulin (Abcam, ab7291, DM1A), validated on human, mouse, and rat by Flow Cyt (Intra), ICC, IHC and WB
 calreticulin (Abcam, ab2908, poly), validated on human by WB
 Anti-rabbit Alexa488 (Thermo, A11034, poly), validated on mouse IgG by Flow Cyt and ICC
 Anti-mouse Alexa555 (Thermo, A21424, poly), validated on mouse IgG by Flow Cyt, ICC and IHC
 Anti-chicken Alexa647 (Thermo, A21449, poly), validated on chicken IgY by ICC, IHC and WB

We then performed the cell culture, fixation, permeabilization and antibody staining with the same protocol as the HPA Cell Atlas Standard Immunostaining Protocol (dx.doi.org/10.17504/protocols.io.x2dfqa6) with the same antibodies reported as HPA, and all parallel for 5 cell line and 9 antibodies in the same run. Staining results were then visually validated with the HPA subcellular human lines staining as reference to confirm the staining, where imaging data could be found in Data Availability.

Eukaryotic cell lines

Policy information about [cell lines and Sex and Gender in Research](#)

Cell line source(s)

Cell lines used in the research are: MCF7, GAMG, HEK293FT, A375, and HeLa.
 HeLa was acquired from ATCC (CRM-CCL-2). A375 cell was acquired from ATCC (CRL-1619).
 HEK293FT was obtained from Thermo Fisher (R70007).
 GAMG and MCF7 were obtained from the Broad-Novartis Cancer Cell Line Encyclopedia (CCLE) project (Nature 483.7391 (2012): 603-607.).

Authentication

None of the cell line used are specifically authenticated, though their morphology and growth dynamic matched with expected cell lines.

Mycoplasma contamination

The cell lines were not tested for mycoplasma contamination, but the IF images took from all DAPI stained cell lines did not reveal existence of mycoplasma.

Commonly misidentified lines
(See [ICLAC](#) register)

None of the cell lines used in this study is listed on the commonly misidentified list.