Corresponding author(s): Alan Saghatelian (asaghatelian@salk.edu)
Thomas F Martinez (tmartinez@salk.edu)

Last updated by author(s): Oct 15, 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	Raw sequencing data was demultiplexed and converted into FASTQ files using CASAVA (v1.8.2). Reads were trimmed of adaptor sequences using fastx_clipper (0.0.13) and aligned with STAR (2.5.2b). Read alignments were processed using samtools (1.3).
Data analysis	Ribo-Seq data was analyzed for smORF translation using RibORF (v0.1). Translated smORFs were filtered against annotated CDS regions using bedtools (2.25). Differential expression analysis was performed using DESeq2 (1.14), and differential translation was analyzed using Xtail (1.1.5). Mass spectrometry data was searched using ProLuCID with DTASelect2 using the web interface from Integrated Proteomics Pipeline (IP2) and pFind (3.1.3). De novo transcriptome assembly was carried out using Cufflinks. Generation of ORF databases from the de novo assembled transcriptome of each cell line was done using a custom java script, GTFtoFASTA. Bar graphs and curves were illustrated by GraphPad Prism 7 and R (3.3).

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

All sequencing datasets generated in this study are available through GEO (GSE125218).

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	Human smORF annotation by Ribo-Seq is novel and therefore the field does not have accepted guidelines for how many replicates to use. Because we found hundreds to thousands of new smORFs with each of the first 3 replicates, we collected and analyzed a minimum of 3 replicates for each cell line. Beyond that we found fewer novel smORFs to be detected. For transcriptome assembly using RNA-Seq, two biological replicates each prepared using two different RNA-Seq library methods (4 total libraries) were utilized for each cell line in order to ensure comprehensive coverage and accuracy.
Data exclusions	No data points were excluded from our analyses.
Replication	For smORF annotation experiments, the overlap was measured and reported for all experiments. For differential mRNA expression and translation experiments, known ER stress genes were checked for similar changes in expression or translation. Differential expression and translation software calculate adjusted p-values for changes. For differential expression $p\_adj < 0.05$ was used, while $p\_adj < 0.1$ was used for differential translation.
Randomization	Randomization was not relevant to this study. Only human cell lines were studied.
Blinding	Blinding was not relevant to this study as the measurements used are all collected and analyzed by computational software.

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

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	All cell lines used were purchased directly from ATCC (HeLa-S3), GE Life Sciences (HEK293T), and Sigma Aldrich (K562).
Authentication	The cell lines are authenticated by the supplier.
Mycoplasma contamination	Cell lines were not tested for mycoplasma after receiving from the supplier.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used.