nature research

Corresponding author(s): Ge Gao, Xiaochen Bo, Hebing Chen

Last updated by author(s): Jan 29, 2022

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 *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	Codes for reproduction of the results reported in this article are available from the GitHub repository, link: https://github.com/gao-lab/HERE.
Data analysis	Codes for reproduction of the results reported in this article are available from the GitHub repository, link: https://github.com/gao-lab/HERE.

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 datasets supporting the conclusions of this article are available in the Zenodo repository, DOI: 10.5281/zenodo.5804921; link: https://zenodo.org/record/5804921. Codes for reproduction of the results reported in this article are available from the GitHub repository, link: https://github.com/gao-lab/HERE. All raw sequencing datasets are available from original publication (Supplementary Table 1), and the compiled editome (and some intermediate results) is available from Zenodo, DOI: 10.5281/zenodo.5804921, link: https://zenodo.org/record/5804921.

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	In addition to including human embryonic RNA-Seq datasets whose A-to-I editomes have been studied previously, we used GEOmetadb to search GEO for all RNA-Seq samples submitted before October 1st, 2020, using the keyword "embryo" and the species restriction of Homo sapiens. We filtered the datasets identified by this search to identify paired-end RNA-Seq data with read length >= 75 × 2 bp, to increase the accuracy of A-to-I RNA editome identification. For single-cell RNA-Seq datasets, we required that the sequencing technology not be based on cell barcoding. This process yielded a total of 2,071 samples (1,852 normal and 274 abnormal) from 18 datasets, which were sent to the A-to-I RNA editome identification pipeline. We believe that results anallyzed from these samples represent the majority, if not all, of the current human knowledge to the transcriptomic dynamics of human early embryonic development.
Data exclusions	No data were excluded from the analyses.
Replication	All results from this study are purely in silico and can be reproduced by the code available here: https://github.com/gao-lab/HERE.
Randomization	This study focuses on identifying and analyzing recurrent (>=50% chances of occurring in a given stage) A-to-I RNA editing in human early embryogenesis based on thousands of human RNA-Seq datasets from 18 published studies. In this way the allocation for the normal samples per early stage is (at least approximately) random.
Blinding	Because this study is a re-analysis of most previously published human early embryo RNA-Seq datasets, blinding is not relevant here.

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