nature research

Corresponding author(s): Rhiju Das

Last updated by author(s): May 26, 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	No code was used in data collection process.
Data analysis	Mapseeker v2.0 (https://github.com/eternagame/MAPseeker) was used to process RNA MAP-seq "Cloud Lab" datasets. HiTrace was used to process "DeepChemicalProfiling" datasets. Requirement: MATLAB 2011a. EternaBench code, used to run algorithms and perform statistical tests, is available at https://www.github.com/eternagame/EternaBench. EternaBench requirements: Python 3.7, numpy 1.19.5, seaborn 0.11.1, scipy 1.3.2, tqdm 4.60.0.

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 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 datasets used here for evaluation are available at https://www.github.com/eternagame/EternaBench. The original cloud lab datasets are available at the RNA Mapping Database28 under accession IDs ETERNA_R00_0000 (Round 00), ETERNA_R69_0000 (Round 01), ETERNA_R70_0000 (Round 02), ETERNA_R71_0000 (Round 03), ETERNA_R72_0000 (Round 04), ETERNA_R73_0000 (Round 05), ETERNA_R74_0000 (Round 06), ETERNA_R75_0000 (Round 07), ETERNA_R76_0000 (Round 08), ETERNA_R77_0002 (Round 09), ETERNA_R78_0001 (Round 10), ETERNA_R79_0001 (Round 11), ETERNA_R80_0001 (Round 12), ETERNA_R81_0001

(Round 13), ETERNA_R82_0001 (Round 14), ETERNA_R83_0003 (Round 15), ETERNA_R84_0000 (Round 16), ETERNA_R85_0000 (Round 17), ETERNA_R86_0000 (Round 18), ETERNA_R87_0001 (Round 19), ETERNA_R89_0000 (Round 20), ETERNA_R91_0000 (Round 21), ETERNA_R92_0000 (Round 22), ETERNA_R94_0000 (Round 23).
The riboswitch raw datasets are downloadable from the supporting information of Andreasson et al. PNAS (2022), at https://www.pnas.org/doi/10.1073/pnas.2112979119.
The "Deep Chemical Profiling" SHAPE and DMS test sets are available for download at https://github.com/DasLab/DeepChemicalProfiling.

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	No sample size calculations were performed. Sample sizes for high-throughput RNA structure mapping "cloud lab" experiments were determined by the maximum allowable size to achieve sufficient read depth on each library construct. The "Deep Chemical Profiling" constructs were selected to represent 13 RNAs of known structure from a range of biological contexts. The remaining datasets, the riboswitch datasets and chemical mapping datasets of natural RNAs, were taken from literature and therefore the sample size not determined.
Data exclusions	Chemical mapping data was processed to remove nucleotides with reactivity over the 98th percentile of reactivity values, also called "Winsorization". Constructs probed in the context of small molecules were removed from the Cloud Lab datasets in order to only compare chemical mapping data in the context of standard buffer conditions. Nucleotides in stretches of polyA > 6 were also removed due to demonstrated reduced signal from polyA stretches.
Replication	The algorithm rankings obtained in this work were evaluated over 24 independent structure mapping and 12 riboswitch datasets, as well as over more than 20 datasets collected from independent groups. Algorithm rankings replicated over all these contexts. Experimental replications were performed for the first cloud lab dataset, and the Deep Chemical Profiling data. This work did not include experimental replicates other than initial checks that the chemical mapping protocol was replicable.
Randomization	Randomization is not relevant because conditions were constructed and there was not subjective allocation of samples to experimental groups.
Blinding	Investigators were not blinded to the study, but all constructs considered were either designed by citizen scientists not involved in experimental characterization or analysis, or data taken from other independent peer-reviewed works.

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