nature research

Corresponding author(s): Nima Aghaeepour

Last updated by author(s): Mar 4, 2023

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	N/A
Data analysis	Data processing and analysis was performed in Python (version 3.9.1), R (4.0.3), and Julia (1.5.3) as described in the methods section. Current packages were installed from conda-forge and Bioconductor.

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

We provide PREE dataset via Zenodo (https://doi.org/10.5281/zenodo.7713898). The data contains 32 samples from 23 women of at least 18 years of age in their first trimester of a singleton pregnancy were recruited to the study after providing their informed consent and under Institutional Review Board (IRB) approved protocols.

The multiomics Pregnancy (PREG) dataset used in Section "Results" and Section "Large-scale multiomics correlation analysis across pregnancy" is available from a public repository (https://doi.org/10.5281/zenodo.7713898). The dataset contains 68 samples from 17 women of at least 18 years of age in their first trimester of a singleton pregnancy were recruited to the study after providing their informed consent and under IRB approved protocols. Also, intermediate data to produce

The cancer dataset used in Section "Results" is derived from a multiomics study available from LinkedOmics (http://linkedomics.org/data_download/TCGA-STAD/). In particular, we integrate the omic datasets Methylation (CpG-site level, HM450K), Methylation (Gene level, HM450K), Mutation (Gene level), RNAseq (HiSeq, Gene level), RPPA (Analyte Level), and SCNV (Gene level, log-ratio). After aligning omics and dropping features with missing or only homogeneous values, the final dataset consisted of samples from 258 patients.

The single cell dataset used to derive the benchmark dataset "Single Cell (SING)" as well as to support the findings in Section "Correlated functional changes across immune cells" is available from FlowRepository (http://flowrepository.org/id/FR-FCM-ZY3Q). Preprocessed benchmarking data as well as intermediate data to produce Figure 3 is provided through a public repository (https://doi.org/10.5281/zenodo.7713898). The dataset contains 68 mass cytometry samples from 17 women of at least 18 years of age in their first trimester of a singleton pregnancy were recruited to the study after providing their informed consent and under IRB-approved protocols.

Download instructions and preprocessing scripts are available at: https://doi.org/10.5281/zenodo.7714039 and https://nalab.stanford.edu/corals/

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	The study uses all available samples available in the datasets (Preeclampsia: 32, Pregnancy 68, Cancer: 258, Single cell: 72). In our prior studies we have demonstrated that these sample sizes are adequate for capturing the required effect size (Aghaeepour, N. et al. 2017, Ghaemi, M. S. et al. 2019, Marić, I. et al. 2022).
Data exclusions	For the Preeclampsia dataset, after aligning omics and dropping features with missing or only homogeneous values, 32 samples with 16,897 features where obtained. For the pregnancy dataset, after aligning omics and dropping features with missing or only homogeneous values, 32,211 features where obtained. For the cancer dataset, after aligning omics and dropping features with missing or only homogeneous values, the dataset consisted of samples from 258 patients. For the single cell dataset only unstimulated samples were used in order to keep cells functionally comparable.
Replication	None of the datasets included replication data. This is not relevant for the current manuscript as it focuses on computational aspects.
Randomization	See publications about individual datasets. This is not relevant for the current manuscript as it focuses on computational aspects.
Blinding	See publications about individual datasets. This is not relevant for the current manuscript as it focuses on computational aspects.

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