

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 (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
<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. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P 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 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

We manually collected and processed approximately 20,000 raw epigenomic data samples of multiple types from numerous data sources. (1) Transcription regulation databases like TcoFbase (<https://bio.liclab.net/TcoFbase/>), CRdb (<https://bio.liclab.net/crdb/>), and TFTG (<http://tf.liclab.net/TFTG/>) have amassed a vast amount of transcription regulator data. (2) Epigenomic databases such as SEdb (<https://bio.liclab.net/sedb/>) and ATACdb (<https://bio.liclab.net/ATACdb/>) have compiled the most comprehensive collections of enhancer and chromatin accessibility data. (3) We also utilized 570 TR knockdown/knockout datasets from the KnockTF database (<https://bio.liclab.net/KnockTFv2/>).

Data analysis

The TRAPT algorithm is implemented in Python 3.11. The source code of TRAPT is available at <https://github.com/TOSTRING-Z/TRAPT>.

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 the datasets analyzed in this study are publicly available. TR knockdown/knockout datasets from KnockTF. The protein-protein interaction (PPI) networks from the STRING database (<https://string-db.org/>). The breast cancer scRNA-seq expression profiles from TCGA (<https://portal.gdc.cancer.gov/>). The ESR1 knockdown RNA-seq datasets are available in the Gene Expression Omnibus (GEO) repository under accession number GSE37820 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE37820>). The GWAS dataset from causaldB (<http://www.mulinlab.org/causaldB>) and the Alzheimer's disease-related H3K27ac data are available under GSE65159 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE65159>). The human hematopoietic stem cell dataset is available on GitHub (<https://gitlab.com/cvejic-group/integrative-scrna-scatac-human-foetal#data>), the human embryonic stem cells dataset is available under GSE75748 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE75748>), and the normal human tissues expression profiles from GTEx (<https://www.gtexportal.org/home/>).

Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

Reporting on sex and gender	Not applicable.
Reporting on race, ethnicity, or other socially relevant groupings	Not applicable.
Population characteristics	Not applicable.
Recruitment	Not applicable.
Ethics oversight	Not applicable.

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	Sample sizes were based on data availability from multiple databases for comprehensive coverage. We collected 1,739 H3K27ac ChIP-seq, 2,723 chromatin accessibility, 11,056 transcription factors, 4,246 transcription co-factors, and 2,591 chromatin regulators datasets. Filtering redundant samples, we maintained 17,227 unique peak files, ensuring robust analysis.
Data exclusions	We excluded redundant data with a peak correlation value of 1 to ensure uniqueness, resulting in 17,227 unique peak files. These criteria were pre-established to eliminate duplicates and enhance data integrity for robust analysis. No other data were excluded from the analyses.
Replication	This is not relevant to our study.
Randomization	This is not relevant to our study.
Blinding	This is not relevant to our study.

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	Involvement in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

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

Plants

Seed stocks

Not applicable.

Novel plant genotypes

Not applicable.

Authentication

Not applicable.