nature portfolio

Corresponding author(s): Chang Su, Jingfei Zhang

Last updated by author(s): Mar 3, 2025

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 *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 software was used.
Data analysis	The code that implements the proposed statistical method is available at https://github.com/ChangSuBiostats/scMultiMap. The code that reproduces all numerical results in the paper is provided at https://github.com/ChangSuBiostats/scMultiMap_analysis. We also used R packages Signac (v1.12.0), SCENT (v1.0.0), rtracklayer (v1.62.0), gprofiler2 (v0.2.2), gbm (v2.2.2), MACS2 (bioconda v2.2.9.1), JASPAR2020 (v0.99.10), Seurat (v5.0.1), dplyr (v1.1.3), ggplot2 (v3.4.4), pheatmap (v1.0.12). We used software LDSC (v1.0.1). R version 4.3.2 was used.

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 data used in this manuscript are publicly available. The PBMC data from 10x Genomics \cite{PBMC_10X,PBMC_10X_nextgem,PBMC_10X_nextgem_X} used in this study are available from the websites of 10x Genomics (URLs: https://www.10xgenomics.com/datasets/pbmc-from-a-healthy-donor-granulocytes-removed-

through-cell-sorting-3-k-1-standard-1-0-0, https://www.10xgenomics.com/datasets/pbmc-from-a-healthy-donor-granulocytes-removed-through-cell-sorting-10-k-1-standard-1-0-0, https://www.10xgenomics.com/datasets/10-k-human-pbm-cs-multiome-v-1-0-chromium-controller-1-standard-2-0-0, https://www.10xgenomics.com/datasets/10-k-human-pbm-cs-multiome-v-1-0-chromium-x-1-standard-2-0-0). The brain datasets \cite{anderson2023single,xiong2023epigenomic} used in this study are available in the GEO database under accession code GSE214979 [\url{https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE214979}] and available from URL \url{https://personal.broadinstitute.org/bjames/AD_snATAC/MFC/}. The ENCODE annotations of brain enhancers were based on subjects ENCDO980BZD and ENCDO871IWW and available on the ENCODE portal [\url{https://www.encodeproject.org}].

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 as no analysis involving sex and gender was performed. The lack of these analyses is due to the small sample size (number of subjects) of the datasets analyzed.
Reporting on race, ethnicity, or other socially relevant groupings	Not applicable as no analysis involving race, ethnicity, or other socially relevant groupings was performed.
Population characteristics	The 3 of the 4 PBMC datasets were collected from a healthy male donor aged 30-35. 1 of the 4 PBMC datasets were collected from a healthy female donor aged 25. For the two brain datasets we used, GSE214979 was collected from 7 AD patients and 8 unaffected donors; https://personal.broadinstitute.org/bjames/AD_snATAC/MFC/ contains 12 and 7 MFC samples from individuals with low and high epigenomic erosion in the prefrontal cortex. In the analysis of differential associations between AD and control groups, we estimate peak-gene associations in two groups separately (i.e. stratified by the covariate of disease status) and test for the difference. In all other analysis, we did not adjust for covariates due to the lack of other covariates' data and limited sample size.
Recruitment	We used publicly available data and no recruitment was conducted.
Ethics oversight	Not applicable as we used publicly available data.

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	Publicly available datasets were used, and no sample size calculation was performed by the authors. The selected sample sizes were the largest available at the time of the study. The results demonstrate that this sample size is sufficient to yield biologically meaningful findings.
Data exclusions	No data were excluded from analysis.
Replication	The experimental findings could be replicated with the description in the manuscript and the GitHub code.
Randomization	Publicly available datasets were used and no randomization was performed by the authors.
Blinding	Blinding is not relevant to our study because no intervention was given or studied.

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
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

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

Plants

Seed stocks

Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.

Novel plant genotypes

Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.

Authentication

Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosiacism, off-target gene editing) were examined.