nature portfolio

Corresponding author(s):	Martin Jinye Zhang
Last updated by author(s):	Jun 28, 2022

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	We did not collect data for this study.
Data analysis	Code is available at https://github.com/martinjzhang/scDRS and https://github.com/martinjzhang/scDRS_paper; customized software scdrs (v1.0.1) was used. We used public softwares MAGMA (v1.08; https://ctg.cncr.nl/software/magma) and LDSC (v1.0.1; https://github.com/bulik/ldsc). We used python and public Python packages numpy (v1.21.5), scipy (v1.5.1), pandas (v1.3.5), scanpy (v1.6.0), anndata (v0.7.8), scikit-misc (v0.1.3), statsmodels (v0.11.1), matplotlib (v3.3.0), seaborn (v0.11.0), gseapy (v0.10.2). We used R package VAM (v0.5.1).

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

We release our data at https://figshare.com/projects/Single-cell_Disease_Relevance_Score_scDRS_/118902 (instructions at https://github.com/martinjzhang/scDRS), including GWAS summary statistics of
the 74 diseases/traits, TMS FACS scRNA-seq data, reprocessed TMS FACS data (for T cells and hepatocytes), MAGMA and
gold-standard gene sets, and scDRS results for TMS FACS (disease scores and control scores for the 74 diseases/traits). The 16

scRNA-seq data sets were obtained as follows (15 out of 16 publicly available). The TMS FACS data and TMS droplet data was downloaded from the official release https://figshare.com/articles/dataset/Processed_files_to_use_with_scanpy_/8273102. The TS FACS data was downloaded from the official release https://figshare.com/articles/dataset/Tabula_Sapiens_release_1_0/14267219. The Cano-Gamez & Soskic et al. data was downloaded from https://www.opentargets.org/projects/effectorness. The Nathan et al. data was downloaded from https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE158769. The Zeisel & Muñoz-Manchado et al. data was downloaded from http://linnarssonlab.org/cortex/. The Zeisel et al. data was downloaded from http://mousebrain.org/adolescent/downloads.html. The Habib & Li et al. data and Habib, Avraham-Davidi, & Basu et al. data were downloaded from https://singlecell.broadinstitute.org/single_cell. The Ayhan et al. data was downloaded from https://cells.ucsc.edu/human-hippo-axis/. The Yao et al. data was downloaded from https://assets.nemoarchive.org/dat-jb2f34y. The Zhong et al. data was downloaded from https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE119212. The Aizarani et al. data was downloaded from https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE124395. The Halpern & Shenhav et al. data was downloaded from https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE84498. The Richter & Deligiannis et al. data (annotated count matrix) was obtained via communication with the authors (raw data publicly available via links in the paper). The Taychameekiatchai et al. data is not publicly available and was provided by co-authors A. Taychameekiatchai, P. Rao, and B. Wang. The MSigDB (v7.1) was downloaded from the official website http://www.gsea-msigdb.org/gsea/index.jsp. The Open Targets data was downloaded from the official website https://www.opentargets.org/.

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	We did not collect new data and the sample size corresponds to the actual sample size of the existing data sets.
Data exclusions	No data was excluded from the analyses.
Replication	We replicated our results by applying the same analysis procedures on additional publicly available data sets; all attempts at replication were successful.
Randomization	We did not perform randomization. Covariates such as age, sex, genotyping array, and ancestry PCs were controlled for the GWAS summary statistics in the original studies. We controlled for data-set-specific covariates, including age, sex, number of expressed genes, and mouse ID, for applying scDRS to the 16 single-cell data sets.
Blinding	We did not apply blinding because we only analyzed existing data sets.

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