nature portfolio

Corresponding author(s): Jacques Behmoaras

Last updated by author(s): Jul 16, 2026

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	All data used in this study were publicly available or from previously published studies, and therefore no software or code was used for this purpose.
Data analysis	R v4.4.3, Python v3.14.2 Packages: scanpy v1.11.4, anndata v0.8.0, Azimuth v0.5.0, Seurat v5.0.1, DEswan v0.0.0.9001, limma v3.62.1, limma v3.66.0, Mfuzz v.2.66.0, clusterProfiler v3.18.0, ggraph v2.2.1, MSigDB v7.5.1, speckle v1.10.0, minfi v1.44.0, minfi v1.52.0, ChAMP v2.28.0, Enmix v1.34.0, DMRcate v3.2.1, scikit-learn v1.3.0, scikit-learn v1,7.2, py-xgboost-gpu v2.1.4, PyTorch v2.5.1, shap v0.48.0. The codes are available via GitHub at https://github.com/HaileyHryPark/pbmc-aging-snakemake.

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

Data from OneK1K cohort was downloaded from CELLxGENE (https://cellxgene.cziscience.com/collections/dde06e0f-ab3b-46be-96a2-a8082383c4a1). Data from Asian Immune Diversity Atlas (AIDA) study was downloaded from CELLxGENE (https://cellxgene.cziscience.com/collections/ced320a1-29f3-47c1-a735-513c7084d508). Data from M et al. was downloaded from Synapse.org (https://www.synapse.org/Synapse:syn49637038/files/). Data of healthy and SLE patients from Perez et al. was downloaded from CELLxGENE (https://cellxgene.cziscience.com/collections/436154da-bcf1-4130-9c8b-120ff9a888f2). Data of healthy and COVID-19 patients from Ren et al. was downloaded from CELLxGENE (https://cellxgene.cziscience.com/collections/0a839c4b-10d0-4d64-9272-684c49a2c8ba). Data of healthy and COVID-19 patients from Wellcome cohort was downloaded from CELLxGENE (https://cellxgene.cziscience.com/collections/ddfad306-714d-4cc0-9985-d9072820c530). Data from COvid-19 Multi-omics Blood Atlas (COMBAT) Consortium was downloaded from CELLxGENE (https://cellxgene.cziscience.com/collections/8f126edf-5405-4731-8374-b5ce11f53e82). Glaucoma cohort data was downloaded from CELLxGENE (https://cellxgene.cziscience.com/collections/de2cde16-c8d3-4a6d-80be-1be9e879aaca). Rheumatoid arthritis cohort data was downloaded from CELLxGENE (https://cellxgene.cziscience.com/collections/de2cde16-c8d3-4a6d-80be-1be9e879aaca). Clonal hematopoiesis cohort data was downloaded from CELLxGENE (https://cellxgene.cziscience.com/collections/0aab20b3-c30c-4606-bd2e-d20dae739c45). Data from Immunobiology of Aging Cohort was downloaded from CELLxGENE (https://cellxgene.cziscience.com/collections/60a2676d-9f37-46cc-9b02-c7370a53be9c). Data from The Sound Life study cohort was downloaded from CELLxGENE (https://cellxgene.cziscience.com/collections/e9360edf-b0b7-4e01-bce8-e596814f13e7). DNA methylation data was downloaded from Gene Expression Omnibus (GEO) (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE87571). NHANES 2017-2018 data was downloaded from (https://wwwn.cdc.gov/nchs/nhanes/).

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	All analyses in this study were conducted with sex-stratification and/or with sex as a covariate.
Reporting on race, ethnicity, or other socially relevant groupings	Differences by race, ethnicity or other socially relevant groupings were not analyzed in this study. Self-reported ethnicity across the combined scRNA-seq datasets is shown in Extended Data Fig. 1a.
Population characteristics	Age and sex were included as variables in the analyses throughout the study.
Recruitment	N/A
Ethics oversight	N/A

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 size was determined by the number of individuals with exact age and sex information in the datasets included in this study. No a priori sample size calculation was performed. The sample size in this study is comparable to those of other human single-cell RNA sequencing studies.
Data exclusions	Individuals without exact age and sex information were excluded from the study.
Replication	No new data was generated in this study; therefore replication was not applicable to the study design.
Randomization	All data used in this study were publicly available or obtained from previously published studies; therefore randomization was not applicable.
Blinding	All data used in this study were publicly available or obtained from previously published studies; therefore blinding was not applicable. The internal and external datasets included in this study were independent of one another.

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.