nature portfolio

Corresponding author(s): Alexander Tyshkovskiy, Vadim Gladyshev

Last updated by author(s): Mar 3, 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	Publicly available datasets were downloaded from the repositories described in the Data Availability section, including GEO, ArrayExpress, SRA, GSA, and dbGaP. No dedicated software was used for data collection beyond standard download utilities and repository-provided tools.
Data analysis	RNAseq data were mapped and counted using STAR (v2.7.11b) and featureCounts (v2.0.6). R (v4.4.0) and Python (v3.8.18) were used to analyze the data. Packages and libraries utilized for data analysis include: numpy (v1.24.4), pandas (v2.0.3), scikit-learn (v1.3.2), lightgbm (v4.1.0), Cell Ranger (v7.0.0), Seurat (v5.0.1), Scanpy (v1.9.6), wot (v1.0.8.post2), flexsurv (v2.3), metafor (v4.6-0), edgeR (v4.2.0), limma (v4.2.0), lme4 (v1.1.35.3), lmerTest (v3.1.3), bootnet (v1.6), qgraph (1.9.8), fgsea (v1.30.0), RNAAgeCalc (v1.16.0), survival (v3.5-8), clusterProfiler (v4.12.0). Resources developed in the study include (1) the TACO interactive web application (http://app.gladyshevlab.org/TACO/); (2) trained transcriptomic clock models (Zenodo: https://doi.org/10.5281/zenodo.18763485); (3) the tAge R package (version 1.0.0) and custom scripts for data preprocessing and clock application (GitHub: https://github.com/Gladyshev-Lab/tAge; Zenodo: https://doi.org/10.5281/zenodo.19039290).

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

RNA-seq and snRNA-seq datasets generated in this study have been deposited in the NCBI Gene Expression Omnibus under accession numbers GSE292843, GSE292884, and GSE293503, and are linked under the SuperSeries GSE292885. Publicly available gene expression data were obtained from published studies (PMID: 37289866, 34306034, 34666007, 29432174, 30787436) and public repositories (GEO, ArrayExpress, SRA, and GSA) using the following identifiers: GSE3150, GSE6591, GSE11291, GSE34378, GSE27625, GSE53960, GSE66715, GSE132040, GSE146977, GSE145480, GSE3129, GSE55272, GSE131754, GSE81959, GSE36838, GSE46895, GSE51108, GSE26267, GSE92486, GSE124772, GSE101657, GSE117188, GSE32609, GSE1093, GSE40977, GSE48333, GSE48331, GSE49000, GSE97074, GSE108379, GSE122116, GSE121395, GSE122080, GSE122085, GSE122243, GSE122367, GSE126814, GSE127758, GSE128724, GSE129083, GSE137504, GSE139204, GSE140286, GSE143304, GSE145972, GSE146796, GSE147666, GSE148647, GSE149569, GSE155407, GSE156762, GSE158980, GSE165409, GSE154832, GSE166615, GSE166778, GSE168211, GSE168610, GSE175571, GSE178770, GSE69952, GSE78130, GSE83931, GSE93833, GSE124483, GSE127475, GSE190939, GSE201207, GSE141252, GSE134781, GSE134780, GSE219203, GSE149029, GSE39699, GSE75574, GSE89272, GSE54853, GSE155064, GSE86882, GSE234563, GSE51202, GSE84390, GSE123293, GSE234667, GSE52794, GSE241904, GSE75192, GSE103232, GSE113957, GSE123658, GSE134080, GSE17612, GSE183701, GSE190125, GSE21935, GSE221178, GSE226189, GSE36192, GSE40645, GSE5086, GSE53890, GSE674, GSE75337, GSE149590, GSE223049, GSE250279, GSE179848, GSE175533, GSE224378, GSE39897, GSE142633, GSE137482, GSE148350, GSE129586, GSE102558, GSE189377, GSE197699, GSE120977, GSE26538, GSE135055, GSE104704, GSE166925, GSE157712, E-MTAB-3374, E-MEXP-2320, E-MEXP-153, E-MTAB-3037, PRJNA281127, PRJNA516151, and CRA012195. The preprocessed rodent gene expression meta-dataset is deposited on Zenodo (https://doi.org/10.5281/zenodo.18763485).
In addition, individual-level human data from the MESA (phs001416.v3.p1), Framingham Heart Study (FHS; phs000974.v6.p5), and the UK Biobank were used in the study. These data are subject to controlled access and can be requested via dbGaP (for MESA and FHS) or the UK Biobank access management system, in accordance with their data use policies.

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	No new human data were generated in this study. Publicly available human gene expression data collected from females and males were utilized in the study to ensure that developed models and findings can be generalized across sexes.
Reporting on race, ethnicity, or other socially relevant groupings	No new human data were generated in this study. Race, ethnicity, and other social groupings were not considered in the study.
Population characteristics	No new human data were generated in this study. For development of the multi-species multi-tissue clocks, we used publicly available human gene expression data comprising tissue samples from 4,003 individuals (2,209 males and 1,794 females) aged 1-106 years (median 56 years; IQR 37-69 years), including healthy donors without diagnosed chronic disease and individuals with Hutchinson-Gilford progeria syndrome (HGPS; n=10), Down syndrome (n=346) or type 1 diabetes (n=65). Pooled tissues include blood (n=2,571), brain (n=1,030), skin (n=278) and skeletal muscle (n=124). Detailed information on sex, tissue and source composition of the human datasets is provided in Supplementary Table 1C. When assessing association between transcriptomic age and disease state or time to death, and between plasma protein concentrations and health outcomes, we adjusted for chronological age and sex of participants.
Recruitment	Participants were not recruited in the study. Only publicly available human gene expression data were used.
Ethics oversight	This study did not involve work with human participants or biological material. Only publicly available human data were used.

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	No a priori statistical power calculations were performed. For mouse bluk RNA-seq data generated in this study (ITP and Klotho KO cohorts), sample sizes were determined by the number of biologically independent animals available from the respective cohorts and by sample sizes commonly used in analogous studies of transcriptomic longevity signatures and molecular clock applications (e.g., PMID 31353263, 37500973,

28380383). For the ITP analyses, we included interventions from the 2015-2017 cohorts and used ≥3 animals per sex and intervention where available, aggregated across program sites, to support replication across data sources and improve generalizability across treatments. For analyses based on publicly available datasets, sample sizes were determined by the source studies, and all eligible samples passing QC were included as described.

Data exclusions

For the ITP cohort data analysis, samples were considered outliers based on Principal Component Analysis (PCA) performed separately for each sex, if their 1st or 2nd principal component (PC) values were more than 1.5 interquartile ranges (IQR) above the third quartile or below the first quartile. Following this rule, 21 samples were identified as outliers and filtered out from the dataset to remove potential tehnical noise associated with the low data quality.
To remove tissue samples with the low data quality prior to the meta-analysis and training of the transcriptomic clocks, for every sample in the meta-dataset we calculated correlation between its expression profile and median expression profiles across all samples corresponding to the same tissues. Samples with Spearman $\rho < 0.5$ were removed from the data as outliers. In addition, for meta-analysis we further excluded samples with values of the 1st or 2nd PCs falling outside the boundaries of the quartiles ±1.5 IQR according to PCA. These quality control procedures were utilized to ensure that the samples utilized for statistical analysis and machine learning do not bias the results due to technical noise introduced during sample preparation or sequencing.

Replication

Most analyses in this study were based on publicly available datasets; we did not perform additional independent experimental repeats ourselves beyond generating mouse RNA-seq data for the ITP and Klotho KO cohorts. Replication of our results was assessed computationally by evaluating reproducibility across independent data sources and biological contexts: (i) ITP-derived signatures were derived separately in males and females and concordance was assessed across sexes; (ii) clock performance and signature robustness were evaluated using nested cross-validation and holdout analyses across >100 independent datasets and by testing in external tissues, species and datasets not used for training; (iii) intervention-, disease-, and damage-associated signatures were assessed for consistency across independent datasets and across tissues, cell lines, sexes, strains, and doses where applicable; (iv) aging-associated effects of Klotho KO and embryogenesis were further replicated across data modalities (bulk and single-cell/single-nucleus); (v) robustness of top contributing genes was assessed across models and molecular modalities (gene expression and proteomics). The number of biologically independent samples per analysis is reported in the figure legends and Methods section.

Randomization

For experiments generated in this study, allocation followed the original study designs. In the ITP cohorts, cages were randomly allocated to treatment groups by the ITP program. Within each group, animals were selected for sacrifice according to the study schedule, and a subset of liver samples was randomly selected for RNA-sequencing. In the Klotho KO experiment, age-matched wild-type and Klotho KO male mice were randomly selected for tissue collection. For analyses based on publicly available datasets, group allocation was determined by the source studies and was not controlled by the authors. Where random allocation was not applicable (e.g., meta-analysis and cross-study comparisons), potential confounding was addressed by including relevant covariates (dataset, tissue, cell line, age, sex, batch or donor where available) in the statistical models, as described in Methods.

Blinding

Blinding was not relevant to this study, since the analysis was performed on molecular data sequenced and processed automatically with the predetermined protocols.

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

Animals and other research organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research, and Sex and Gender in Research

Laboratory animals

For Interventions Testing Program (ITP) cohort analysis, liver samples from young (4-6 month old) and old (22 month old) female and male mice (Mus musculus) from genetically heterogeneous UM-HET3 strain were utilized. Samples were acquired from collections of University of Michigan Medical School (UM), University of Texas (UT) and The Jackson Laboratory (TJL) obtained from animals of 2015, 2016 and 2017 ITP cohorts. The detailed information about age, sex and treatments applied to the animals is provided in Methods section and Supplementary Table 1A. All mice were fed ad libitum with the same diet (Purina 5LG6) made in the same commercial diet kitchen (TestDiet, Richmond, IN, USA). Mice represented genetically heterogenous UM-HET3 strain, produced by crossing female CByB6F1/J and male C3D2F1/J animals. Mice were housed in plastic ventilated cages under similar environmental conditions in all sites (21-23°C temperature, 40-60% humidity, 12-hour light/dark cycle).
For Klotho knockout experiment, tissue samples from 8-week-old male C57BL6/J mice (Mus musculus) and α-klotho knockout age-

matched males purchased from CLEA Japan were utilized. Mice were housed under specific pathogen-free conditions, at a temperature of 22 ± 2 °C, with a relative humidity of 50 ± 10%, and on a 12-hour light/dark cycle. Mice had free access to tap water and standard chow.

Wild animals

The study did not involve wild animals.

Reporting on sex

To ensure that findings of the study are generalized across sexes, both female and male samples were examined in this work.
For analysis of ITP cohort data, liver samples from 94 males and 88 females were sequenced. The detailed information about age, sex and treatments applied to the animals is provided in Methods section and Supplementary Table 1A. Signatures of aging, maximum lifespan, and mortality were identified separately for each sex and for pooled data (with sex included as a covariate) to ensure that discovered biomarkers reflect universal associations with examined traits, shared across sexes.
For meta-analysis and development of transcriptomic clocks, female and male gene expression data were aggregated from publicly available datasets. In total, our rodent meta-dataset included 4,539 animals, including 3,876 mice (1,145 females, 2,593 males, and 138 animals with unreported sex) and 663 rats (160 females, 491 males, and 12 animals with unreported sex). To identify signatures of chronological age, lifespan, and mortality shared across sexes, we performed statistical analysis for pooled data, adjusting for sex by including it as a factorial covariate into the model. To ensure that transcriptomic clocks are applicable to both sexes and are not biased by sex-specific expression patterns, we trained them on the pooled data. The details about species, strain, age, sex, genotype, treatment, and source information for each animal sample included in the meta-analysis are provided in Supplementary Table 1B.

Field-collected samples

The study did not involve samples collected from the field.

Ethics oversight

For the original ITP studies, the animal protocols were approved by the Institutional Animal Care and Use Committee of each site (The Jackson Laboratory in Bar Harbor (TJL), the University of Michigan at Ann Arbor (UM), and the University of Texas Health Science Center at San Antonio (UT)).
For the Klotho KO study, the animal protocols were approved by the Tohoku University Institutional Animal Care and Use Committee. The animal procedures performed conform to the NIH guidelines (Guide for the Care and Use of Laboratory Animals).

Note that full information on the approval of the study protocol must also be provided in the manuscript.

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.