nature portfolio

# Peer Review File

# Universal Transcriptomic Hallmarks of Mammalian Aging and Mortality

Corresponding Author: Professor Vadim Gladyshev

Any redactions in this file are there to maintain patient confidentiality, the confidentiality of unpublished data, or to remove third-party material.

**This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.**

Version 0:

Reviewer comments:

Referee #1

(Remarks to the Author)
In this work, the authors perform a meta-analysis of publicly available gene expression data from rodents to develop transcriptomic clocks of aging, mortality, and maximum lifespan. They also generate additional RNA-seq data and test/validate their clocks in various contexts. This is a very timely topic as there is great interest in the field in developing clocks. The paper is generally well-written but I found it long and some aspects can be improved, as follows.

Generally speaking, developing clocks based on so much data - even if most of it has been published - is novel and impressive and I am sure the transcriptomic clocks developed by the authors will be useful for the scientific community. That said, I wonder how the clocks developed by the authors compare to other transcriptomic clocks? And how do they compare to epigenetic clocks, currently the benchmark in the field? There is a brief mention to this in line 299, but it could be clearer why would scientists use these transcriptomic clocks instead of epigenetic clocks.

The insights gained from the clocks are at times confirmatory, however, and I think the paper could be much more focused on the novel findings (see some specific comments on this below). Apart from the clocks, which are novel and interesting, what are the major breakthroughs from this work?

One major issue in the paper is that I think quite often the authors overinterpret their findings. "Remarkably" is mentioned 19 times in the manuscript. "Interestingly" is mentioned 27 times. Overall, the authors are finding correlations between molecular changes (gene expression) and phenotypes such as age and mortality, but correlations do not prove causality (see PMID 38242993 for a recent review in the context of aging), and thus I think some of the conclusions and statements are not supported by the results. Some specific examples:

Line 173: "This suggests that molecular mechanisms of aging and longevity..." - how do the authors know these signatures relate to molecular mechanisms of aging? Same for the abstract (line 53). The authors show correlations, but that does not by itself prove that molecular mechanisms of aging are underpinning the signatures.

Line 469: "Among top genes driving the pro-aging and pro-mortality effect of this model was Cdkn1a" - how do the authors know Cdkn1a is driving the phenotypes?

Line 552: "aging is accompanied by both global hallmarks of damage accumulation" - how does the authors' data show damage accumulation in normal mice? And related to this in Line 1036: "...reflect an aggregated level of molecular damage" - how do the authors know clocks reflect molecular damage? I suggest inducing different types of molecular damage in vitro, like DNA damage or protein aggregates, and assay its effect on the clocks. Or perhaps assay molecular damage in normal mice across the lifespan and determine how it correlates with the clocks? Either way, more data is necessary to support these conclusions.

Line 620: "Molecular mechanisms of biological age" - how do the authors know biological age has increased? A marker that correlates with age may or may not represent biological age. This is an issue in many other places and I suggest changing

terminology (ie, "biological age" to "transcriptomic age") throughout the manuscript.

Line 709: "Molecular mechanisms of rejuvenation induced by heterochronic parabiosis" - is parabiosis rejuvenation? I don't think the authors show this, and extraordinary claims require extraordinary evidence. Lifespan extension from parabiosis is modest, not what one would expect from rejuvenation.

Line 734: "The top gene driving the rejuvenating effect of HPB" - how do the authors know this gene is driving the rejuvenating effect?

Line 1085: "Several established models of organismal rejuvenation, including heterochronic parabiosis in old mice, early embryogenesis, and cellular reprogramming." - embryogenesis and cellular reprogramming are not organismal models. And as mentioned above I don't think heterochronic parabiosis is an established model of rejuvenation either.

Further specific comments:

Line 155: "this mortality model captures differences in survival dynamics across treatments and sexes, and can be used to estimate expected hazard rate and maximum lifespan" - did the authors validate their model in data not used for training the model? What about a different dataset?

Line 226: "9059 to 9167 signatures of chronological age" - you mean genes associated with age? Isn't that nearly half the genome associated with age?

Line 394: The co-expression analysis largely confirms previous results, such as inflammation and immune responses, mitochondria, ECM. I don't think it is particularly novel, unless I am missing something?

Line 513: The authors applied clocks to cell types using publicly available single cell data and show increases in transcriptomic age for most cell types. I think this is an interesting analysis and one of the strengths of the manuscript.

Line 667: again, this has been reported by many others (as acknowledged by the authors), can shorten this section in my view.

Line 874: "transcriptomic clocks of mortality predicted a statistically significant acceleration of biological age for patients diagnosed with all examined diseases compared to healthy people after adjustment for chronological age and sex" - if unhealthy individuals have an acceleration of clocks then I don't think we can say that biological age is accelerated but rather that the clock is not reflective of aging but of health. This goes back to my point above that I don't think the authors are showing that biological age is being affected and should tone down their conclusions.

Line 935: "At the same time, several interventions resulted in the decrease of biological age in murine tissues." - doesn't this include the ITP interventions from which the data was trained on?

Line 1008: what do you mean by "molecular hallmarks of aging"?

Line 1131: "Overall, our study resolves the challenge of predicting biological age and mortality by developing a comprehensive transcriptomic model encompassing both aging and lifespan- modulating interventions." - is there one model or multiple ones? I thought there were different clocks, for age, mortality and maximum lifespan, not one single clock. So in practice researchers will need to use different clocks for different purposes, not one single model or clock.

The website listed by the authors (http://app.gladyshevlab.org/TACO/) is not publicly available.

Overall, I think this is an interesting work and the transcriptomic clocks will make a valuable contribution to the field, though some aspects could be improved and I suggest not overselling the findings and shortening the manuscript.

It is possible that some of my comments reflect misunderstandings of mine. If so then I would suggest that the authors use my misunderstandings as an indication that such points might be made clearer in the manuscript.

It is my usual policy to reveal my identity to the authors: Joao Pedro de Magalhaes.

Referee #2

(Remarks to the Author)
This manuscript by Gladyshev and colleagues reports an extremely extensive, but also incredibly lengthy and meandering, study of transcriptional changes during rodent aging. The authors formulate a series of transcriptomic “aging clocks” which are quite interesting, especially the contrast between the three different types of clock they define (chronological, mortality, and maximum lifespan). Additionally, they reduce accurate clock predictions to sets of expression changes within gene modules (which have specific functional interpretations), addressing the widespread previous criticisms of aging clocks as predictive but without biological insight. While there have been previous studies of transcriptional aging clocks, their accuracy has been moderate-to-low and they have been applied to predict chronological age only, rather than the three aging types of the present work. Given this landscape, the results of the present manuscript are a significant contribution to

the aging clock literature.

Unfortunately, the communication of all of the results is overwhelming in volume. The paper comes in at 13,337 words of main text and a very extensive supplement with a grand total of 28 supplementary figures each with numerous subpanels. Despite this volume, the Methods sections needed to understand some of the key methodology are lacking in clarity and sufficient organization. Together, these factors made reading this manuscript frustrating and slow.

In summary, some of the science described in this paper is indeed quite exciting but it needs extensive revision to put it mildly.

Major concerns:

On balance, improving the organization of the methods section may be the lowest hanging fruit. For example, the calculation of transcriptomic signatures is spread across methods sections titled “Preprocessing and analysis of ITP cohort RNA-seq data” and “Transcriptomic signatures of aging, mortality and lifespan on the aggregated meta-dataset.” The WGCNA methods are also spread across multiple sections. It would be helpful if each analysis was discussed only in its own methods section. We frequently thought we had read all the methods related to a given analysis only to find another section further discussing it. We’d recommend splitting up the very long method subsections into descriptively titled shorter sections; creating a table that briefly defines each type of tAge, signature, or frequently used value and provides the name of the one methods section where the reader can find a further description of how this value was calculated; and a more intuitive naming of each value/variable.

There are an overwhelming number of different signatures and scores calculated from adjusted and unadjusted values related to lifespan. A table mapping each such value and signature to a brief description of how it was calculated is necessary. Relatedly, some of the names chosen for certain values are quite confusing (e.g., “Adj. lifespan-adjusted age”). Perhaps the table will be enough to clear this up but more intelligible names would be helpful.

The single sentence description (lines 1389-1390) of how the relative transcriptomic age clocks (and relative mortality clocks) were trained is insufficient. It seems from the main text that a randomly chosen reference group was used for each dataset, how does clock prediction accuracy vary depending on the random choice of reference group? If all datasets are normalized to the same reference group how do the results change? Is the performance of the relative clock robust to the random reference groups chosen (i.e., if the clock is re-trained N times with a different reference chosen for each group each time are age predictions of individual samples consistent?)

Before showing the new multi-species model (Figure 7) they should show how the mouse/rat clock(s) from the first 6 figures perform on human samples. If they do not perform well they should try to explain why. If the claim is that “transcriptomic hallmarks of aging and mortality appear to be generally conserved across species” (lines 877 - 878) ,why is it that they must include human samples in the training process?

Finally, the authors should consider consolidating or removing some of the results. For example Figures 5 and 6 might be abbreviated and combined. Additionally, the section on the discovery of novel mortality-regulating interventions was a weak point and could be removed or greatly abbreviated.

Minor concerns:

Figure 1 and Extended data Fig. 1:
4 associations are referenced in lines 162-164, but only 3 volcano plots are provided in Extended data Fig. 1D-F). The 4th volcano plot should also be shown.
Going from Extended data Fig. 1D-F to Extended data Fig. 1G/Figure 1E it seems that somehow the genes associated with each measure (chronological age, lifespan-adjusted age, expected hazard rate, and maximum lifespan) are integrated into a single “signature” score for that measure. The description of how this is done in the main text and methods is insufficient and should be clarified and expanded. Is each signature defined by the vector of coefficients for each gene coming from the linear mixed-effects model? Or maybe just the vector of coefficients for genes with a significant association? When transcriptomic signatures are compared (as in Extended data Fig. 1G) how is a correlation value defined between signatures, by measuring the correlation between coefficients (e.g., “chronological age slope”) of each gene between signatures? Answers to all of these questions should be made obvious.
The size of the circles in Figure 1M should be made proportional to the number of elements represented by the circles.
As the choice of normalization method scaling vs. yugene did not make much of a difference it does not add much to show both. Just one should be chosen to show in figure 1 and the other can be shown in the supplementary figures.

Extended data Fig. 2
The paragraph from line 176 to 194 should be broken into two paragraphs as the discussion of particular genes and then gene set enrichments are unrelated and it is confusing to have them grouped together.
It would be clarifying to have Extended data Fig. 2F come before Extended data Fig. 2E, as I believe Extended data Fig. 2E represents the pairwise correlations between the columns of Extended data Fig. 2F. Also the associated text is misleading; the associations between the signatures at the level of genes is shown in Extended data Fig. 1G while Extended data Fig. 2E shows the association between signatures at the level of pathways. This section should be rewritten to make this distinction between the heatmaps Extended data Fig. 1G, Extended data Fig. 2E, and Extended data Fig. 2F clear.

Extended data Fig. 3D is referenced before Extended data Fig. 3C.

Figure 2:
It would be illuminating to see how many features are selected for use in the final model (by penalized regression for example) by each of the clocks trained in Figure 3.
How did the tissue type used for transcriptomic profiling affect the accuracy of prediction of lifespan extension shown in Figure 2H-J and Extended data figure 6E-F? One way to show this would be to extend Figure 2H into a heatmap where each column is a tissue, each row is a clock, and each value is the association between that clock's prediction in that tissue and the effect on lifespan.

Extended data Fig. 5 panels are not referenced in order in the text. Nor in order with extended data figure 6.

Figure 3
The labels in Figure 3C are not legible due to their small size and insufficient image resolution. Some sort of legend mapping the thickness of the line to the correlation coefficient is needed.
The x-axis labels in Figure 3E are confusing, particularly for the right two barplots.
Extended Data Figures 9 and 10 are never referenced in the text.
Panels are not referenced in order in the text (I and H).
What do the last two columns refer to in Figure 3K? Were samples binned into high and low hazard groups and a GSEA was done comparing these groups?
The methods section describing the construction of the module specific clocks is insufficient. Were they relative clocks or not? Was the data yugene or scaled?

Figure 4
Panels F and G are referenced out of order. So are H and I.
How was tAge normalized in this figure?

Extended data figure 15
The assertion on lines 605-609, that the modules show commonalities between the klotho KO associated gene expression changes in proximal tubule cells and neurons, does not seem to be well supported by Extended data figure 15D-E. The most contributive module "Respiration/Mitochondrial translation" an increase in tAge in both cell types, but resulting from an opposite enrichment. Showing Extended data figure 15D-E as a scatterplot, as in Extended data figure 15C, would be more informative. To support your assertion it should be demonstrated that within positive and negative NES, separately, there is a significant positive correlation between the effects on tAge of each module.
The line of argument about inflammatory changes in the single cell klotho KO data (lines 608-618) is not persuasive because neither of the cited results in Extended Data Fig. 15F and Fig. 4K reached statistical significance. This section should be removed or supported with results that are significant.

Discussion
The claim "A 3-month treatment with ouabain was able to decrease the biological age of 24 month-old mice by ~10 months suggesting a first example of a radical in vivo rejuvenation effect induced by a single-molecule pharmacological intervention" (lines 1105-1106) is undermined by the fact that other single-molecule pharmacological interventions show similar if not greater "rejuvenations" in Extended Data Fig. 28A (e.g., metformin and other unlabeled points). This statement should either be walked back or expanded upon to explain why metformin and the other points shown in Extended Data Fig. 28A are less impressive than ouabain.

Data availability
A large amount of time and effort has clearly been devoted to identifying, downloading, cleaning, and combining numerous disparate transcriptomic datasets. To facilitate the replication of your results a script or set of scripts should be provided (along with the paper or on github) which download (from GEO) and combine the (public) analyzed datasets into the form analyzed in the paper. This will allow other researchers to benefit from, and build upon, the work you have done to aggregate these datasets. The TACO web-app is a good step on the way to this, but does not allow other researchers to directly analyze these dataset(s) for new insights.
The data availability section says the ITP RNA-seq dataset has been deposited on GEO but does not provide a GEO accession number.

Remarks on Code Availability: The code is well organized and sufficient.

Version 1:

Reviewer comments:

Referee #1

(Remarks to the Author)
The authors have addressed my suggestions, and the manuscript is much improved.

The only aspect I couldn't find more information about was that, in my previous comments, there were >9000 genes associated with aging, which is quite a lot, yet in the revised manuscript I couldn't find exactly how many genes are associated with aging or how many genes compose each of the transcriptomic clocks developed by the authors.

Referee #2

(Remarks to the Author)
The revised manuscript by Gladyshev and colleagues is significantly improved. All of the major issues identified in peer-review have been sufficiently addressed. The new analyses comparing the tAge model to epigenetic clocks are particularly illuminating and the significantly abridged manuscript length is a great improvement.

It is also appreciated that the authors now have published a processed version of their data for use by the greater scientific community.

However, there are two minor issues with this data which should be addressed before publication. Firstly, the metadata is provided as `.gsheet` files, which we are unfamiliar with and have been unable to open; these files would be better provided in a common file format like `.xlsx` or `.csv`. Similarly, the pre-processed expression data is provided as `.rds` files, which are cumbersome for non-R users to analyze; these would be more widely accessible if provided as `.xlsx` or `.csv`.

Trey Ideker and Zane Koch

Referee #3

(Remarks to the Author)
I co-reviewed this manuscript with one of the reviewers who provided the listed reports.

**Referee #1:**

In this work, the authors perform a meta-analysis of publicly available gene expression data from rodents to develop transcriptomic clocks of aging, mortality, and maximum lifespan. They also generate additional RNA-seq data and test/validate their clocks in various contexts. This is a very timely topic as there is great interest in the field in developing clocks. The paper is generally well-written but I found it long and some aspects can be improved, as follows.

We thank the reviewer for the positive assessment of our study and for recognizing its relevance to the field. Below we provide detailed responses to all comments.

Generally speaking, developing clocks based on so much data - even if most of it has been published - is novel and impressive and I am sure the transcriptomic clocks developed by the authors will be useful for the scientific community. That said, I wonder how the clocks developed by the authors compare to other transcriptomic clocks? And how do they compare to epigenetic clocks, currently the benchmark in the field? There is a brief mention to this in line 299, but it could be clearer why would scientists use these transcriptomic clocks instead of epigenetic clocks.

We would like to thank the reviewer for this advice. In the revised manuscript, we expanded the comparison between our clocks and existing transcriptomic and epigenetic clocks at several levels, highlighting their accuracy, applicability across species/tissues, and interpretability.

Most existing transcriptomic clocks predict chronological age and are limited to a single tissue and species (e.g., mouse brain, human blood or human fibroblasts) (PMIDs: 30567591, 26490707, 37118510). To our knowledge, our transcriptomic clocks are the first that (1) can predict not only chronological age but also expected mortality and maximum lifespan; (2) generalize across >25 tissues, individual cell types, and multiple mammalian species. During revision, we also enhanced our multi-species clock by including an additional species, crab-eating macaque (*Macaca fascicularis*; 2,623 samples, >30 tissues), so the final multi-species model in our study covers 4 species (mouse, rat, macaque, human) from 2 mammalian orders.

To benchmark our multi-species model against the pan-mammalian epigenetic clocks developed by Lu et al. (PMID: 37563227), we performed leave-one-fold-out (LOFO) validation following their protocol. Our multi-species multi-tissue transcriptomic clock of relative chronological age achieved r=0.952 across all samples and r≥0.94 within each species (Extended Data Fig. 3f, Review Fig. 1), comparable to the epigenetic clock reported by Lu et al. (r=0.953 across samples).

Using the Framingham Heart Study (FHS), we compared predictions of our multi-species transcriptomic clocks of expected mortality to human-specific transcriptomic (Peters, RNAAgeCalc) and epigenetic clocks (Horvath, Hannum, PhenoAge, DunedinPace, YingDamAge, etc.). Our composite universal mortality clock correlated with chronological age similarly to second-generation epigenetic clocks (e.g., PhenoAge, YingCausAge), and stronger than published transcriptomic clocks (Extended Data Fig. 8j). Transcriptomic age estimated by several module-specific mortality clocks (e.g., chromatin, inflammation, oxidative phosphorylation) showed strong positive associations with time-to-death, comparable to state-of-the-art epigenetic clocks including DunedinPace and PhenoAge (Fig. 5f, Extended Data Fig. 8i-j, Review Fig. 2a). Moreover, predictions from transcriptomic and epigenetic clocks in human blood were significantly correlated after adjustment for age and sex (Fig. 5g, Review Fig. 2b). The highest concordance was seen for the chromatin-related transcriptomic clock, indicating linkage between mortality-related changes at gene expression and DNA methylation levels.

Finally, unlike epigenetic clocks, transcriptomic clocks allow mechanistic interpretation of observed effect by identifying genes, pathways, and cellular modules that contribute to age acceleration or deceleration.

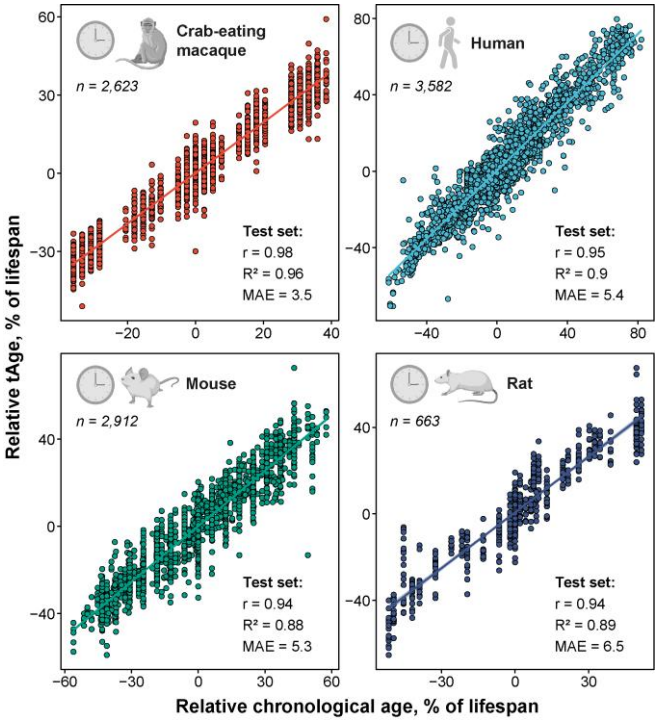
During revision, we included additional case studies (Fig. 3-5, Extended Data Fig. 5-7) demonstrating how module-specific clocks can characterize molecular pathways associated with the effects of detrimental or geroprotective interventions *in vitro* and *in vivo*.

We have also revised the main text to:

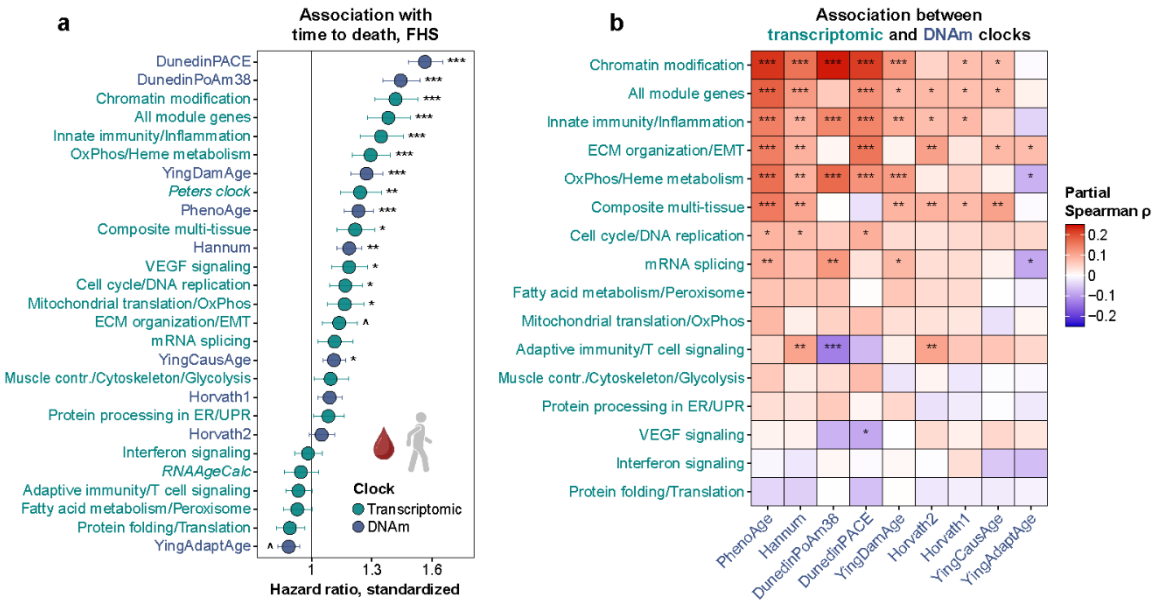
- Clarify benchmarking our transcriptomic clocks against other existing molecular clocks (lines 292-301; 595-608);

- Highlight novelty (multi-species, multi-tissue, mortality-focused, interpretable clocks) (lines 729-740);

- Make the rationale for using transcriptomic clocks both clearer and more compelling (lines 741-749; 819-823).




**Review Figure 1.** Leave-one-fold-out (LOFO) performance of EN multi-species multi-tissue clocks of relative chronological age (YuGene normalization) across mouse (n=2,912), rat (n=663), crab-eating macaque (n=2,623), and human (n=3,582) samples. Only test predictions are shown. Pearson's r, $R^2$, and MAE are indicated.




**Review Figure 2. a,** Association of transcriptomic (green) and human DNAm (blue) clock predictions in human blood with time to death in Framingham Heart Study (n=1,698–3,698). Standardized hazard ratios were estimated via Cox proportional hazard models, with standardized molecular age adjusted for age and sex as an independent variable. Data are exp(logHR ± s.e.), and corresponding p-values are denoted with asterisks. Tested transcriptomic clocks included EN universal mortality clocks (composite, multi-module, and module-specific) and previously published human chronological clocks (Peters clock and RNAAgeCalc, italicized). **b,** Partial Spearman correlations between DNAm (columns) and EN universal mortality transcriptomic (rows) clocks, adjusted for chronological age and sex. ^ p.adj < 0.1; * p.adj < 0.05; ** p.adj < 0.01; *** p.adj < 0.001.

The insights gained from the clocks are at times confirmatory, however, and I think the paper could be much more focused on the novel findings (see some specific comments on this below). Apart from the clocks, which are novel and interesting, what are the major breakthroughs from this work?

Thank you for bringing up this question. In the revised manuscript, we updated the abstract and Results section to emphasize the most central findings beyond clock development.

Below we summarize the key advances of this work:

- Accurate and interpretable multi-species, multi-tissue clocks of chronological age and expected mortality trained on >11,000 tissue samples from >25 tissues across 4 species. These clocks robustly predict lifespan-extending interventions, chronic diseases, and time-to-death, providing a practical and interpretable tool for discovery and characterization of health-modulating interventions complement to other established molecular clocks.
- Conserved transcriptomic biomarkers of aging across mammalian species and individual cell types.
- A modular architecture of aging and mortality-associated transcriptomic changes, reflecting the state of key cellular subsystems. The strongest health-related modules include innate immunity/inflammation, mitochondrial translation and oxidative phosphorylation, chromatin organization, and ECM/differentiation.
- Cell type- and tissue-specific mortality-associated responses to interventions. Health-modulating interventions (e.g., *Klotho* knockout, caloric restriction) can induce changes in molecular age that are not uniform across tissues, cell types, or functional subsystems. Module-specific clocks and single-cell analyses reveal pathways and cell types that are particularly responsive or resistant to harmful or lifespan-extending interventions.

- Consistency of transcriptomic signatures of aging and mortality across *in vivo* and *in vitro* biological models. The same molecular biomarkers associated with organismal aging and health are reproduced in models of replicative senescence, γ-irradiation, metabolic disruption, inflammation, and other stress models, indicating that they reflect systemic processed of damage accumulation at multiple biological levels of organization.
- Convergence between transcriptomic and epigenetic clocks in human blood. Deviations in transcriptomic and DNA methylation age were correlated, particularly for the module clock associated with chromatin modification, linking two major modalities of aging biomarkers.
- Transcriptomic age deceleration induced by heterochronic parabiosis and early embryogenesis, supporting the "ground zero" hypothesis. Embryonic reduction of tAge is specific to intraembryonic lineages, whereas extraembryonic endoderm does not demonstrate rejuvenation.
- Identification of conserved molecular markers of systemic damage and mortality. Several individual genes and proteins, including *Cdkn1a*, *Gpnmb*, and *Lgals3*, are consistently associated with transcriptomic age deviation across numerous *in vivo* and *in vitro* models, including aging, diseases, stress, heterochronic parabiosis, caloric restriction, and embryogenesis, suggesting a conserved molecular signature of damage accumulation.

One major issue in the paper is that I think quite often the authors overinterpret their findings. "Remarkably" is mentioned 19 times in the manuscript. "Interestingly" is mentioned 27 times. Overall, the authors are finding correlations between molecular changes (gene expression) and phenotypes such as age and mortality, but correlations do not prove causality (see PMID 38242993 for a recent review in the context of aging), and thus I think some of the conclusions and statements are not supported by the results. Some specific examples:

We thank the reviewer for pointing out this issue. We have carefully revised the manuscript to reduce the use of words such as "remarkably" and "interestingly" and have rephrased several statements to ensure they are fully supported by the data and do not imply causality beyond what our results demonstrate.

Line 173: "This suggests that molecular mechanisms of aging and longevity..." - how do the authors know these signatures relate to molecular mechanisms of aging? Same for the abstract (line 53). The authors show correlations, but that does not by itself prove that molecular mechanisms of aging are underpinning the signatures.

Thank you for this comment. We agree that different fundamental mechanisms could eventually result in convergent molecular signatures. We have replaced "mechanisms" with "signatures" or "biomarkers" in line 138 and other related instances to clarify that our study identifies correlated molecular features rather than definitive mechanistic drivers.

Line 469: "Among top genes driving the pro-aging and pro-mortality effect of this model was Cdkn1a" - how do the authors know Cdkn1a is driving the phenotypes?

We appreciate the reviewer's point. To avoid implying causality, we have replaced "driving" with "contributing" (line 423), highlighting that *Cdkn1a* has one of the highest contributions in the model of age and mortality prediction without claiming it has a causal effect on the phenotypes.

Line 552: "aging is accompanied by both global hallmarks of damage accumulation" - how does the authors' data show damage accumulation in normal mice? And related to this in Line 1036: "...reflect an aggregated level of molecular damage" - how do the authors know clocks reflect molecular damage? I suggest inducing different types of molecular damage in vitro, like DNA damage or protein aggregates, and assay its effect

on the clocks. Or perhaps assay molecular damage in normal mice across the lifespan and determine how it correlates with the clocks? Either way, more data is necessary to support these conclusions.

We would like to thank the reviewer for this insightful suggestion. Following it, we applied our transcriptomic clocks to multiple *in vitro* and *in vivo* models of induced molecular damage, including:

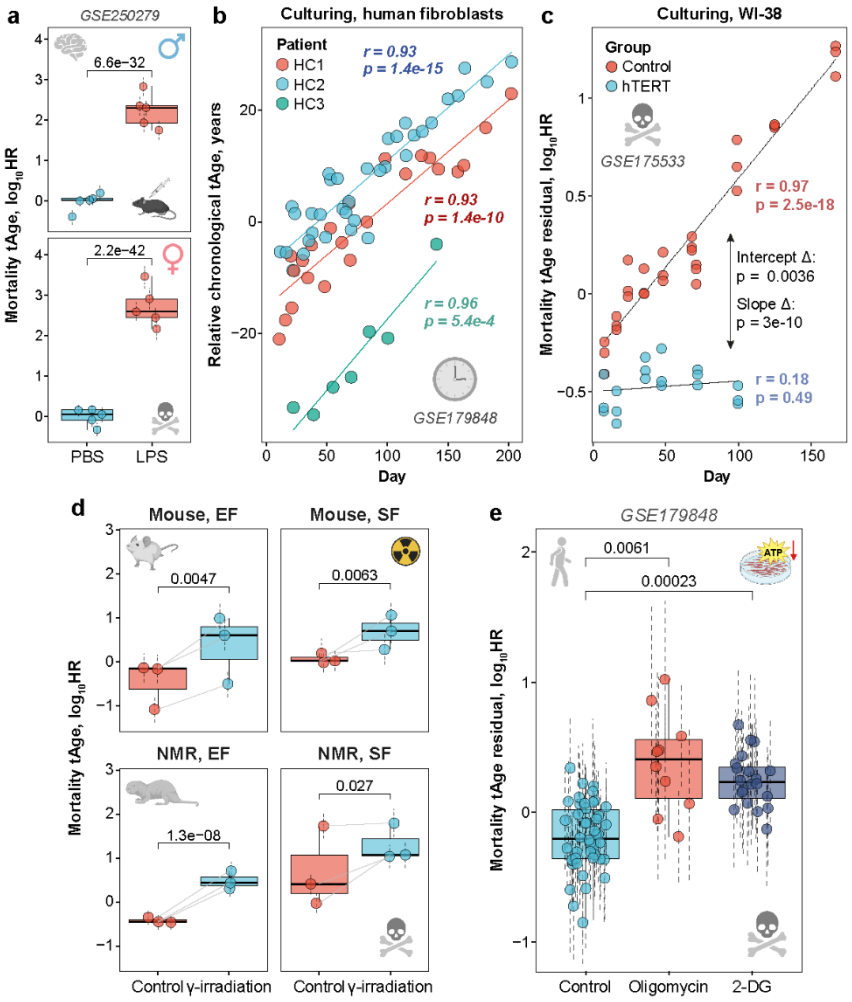
- Inflammatory stress induced by LPS injection in mouse brain (Fig. 3d-e, Extended Data Fig. 5h-j, Review Fig. 3a);
- Replicative senescence in human primary fibroblasts and WI-38 cells (Fig. 4a-b, Extended Data Fig. 7a-e, Review Fig. 3b-c);
- DNA damage induced by γ-irradiation in embryonic and skin fibroblasts from mice and naked mole rats (Fig. 4f-g, Extended Data Fig. 7h, Review Fig. 3d);
- Metabolic stress induced by oligomycin or 2-deoxyglucose (2-DG) treatments (Fig. 4e, 4j, Extended Data Fig. 7g, Review Fig. 3e).

In all cases, we observed statistically significant acceleration of transcriptomic age. This was frequently accompanied by upregulation of genes associated with systemic damage, such as *Cdkn1a* and *Eda2r* (Fig. 4k), supporting the interpretation that transcriptomic clocks capture not only changes related to aging dynamics and probability of death, but rather overall damage status of the biological system, reproduced even at the level of individual cells *in vitro*. Importantly, interventions that mitigated these stresses also reduced transcriptomic age acceleration. For example, hTERT induction in WI-38 cells prolonged proliferation and stabilized transcriptomic age, linking biomarkers of molecular damage and functional phenotypes.

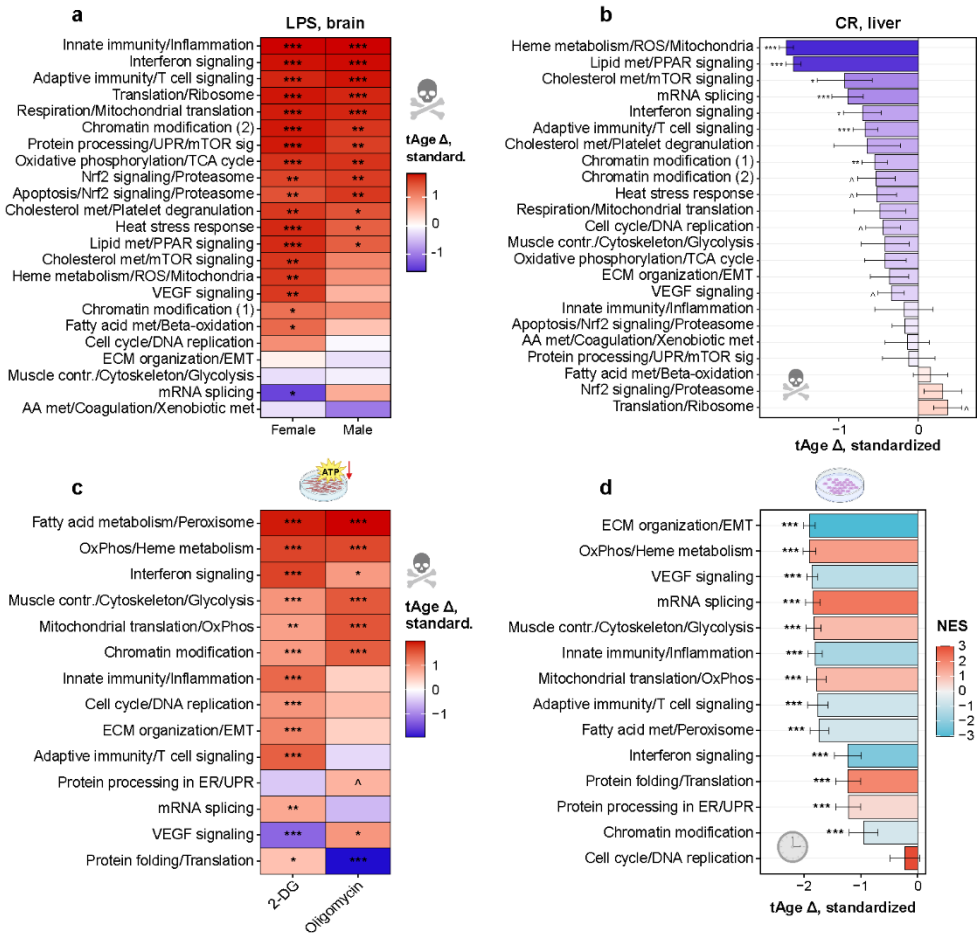
To further characterize these effects, we developed module-specific multi-species clocks of chronological age and expected mortality, trained on individual gene expression modules shared across rodents and primates (Fig. 4h, Extended Data Fig. 7j-l). Consistent with expected biology:

- LPS accelerated molecular age of immune-related modules (inflammation, interferon signaling, etc.) (Fig. 3e, Extended Data Fig. 5i-j, Review Fig. 4a);
- Caloric restriction in mouse liver produced geroprotective effect on metabolic modules, including lipid metabolism, heme metabolism, cholesterol metabolism (Fig. 3e, Extended Data Fig. 6a-b, Review Fig. 4b);
- Oligomycin and 2-DG induced the strongest pro-aging effect on metabolic modules, including fatty acid metabolism, oxidative phosphorylation, heme metabolism (Fig. 4j, Review Fig. 4c);
- Cellular reprogramming most strongly reduced transcriptomic age associated with differentiation/ECM-related modules (Fig. 4i, Extended Data Fig. 7m, Review Fig. 4d).

These results show that while various stressors may converge in accelerating composite transcriptomic age, they may have different primary targets and differentially impact specific functional subsystems. Our tools can be used to identify most responsive modules, providing improved interpretability and mechanistic insight compared to traditional molecular clocks.




**Review Figure 3. a,** Mortality tAge of brains from 5-6-month-old male (left) and female (right) mice injected with PBS or lipopolysaccharide (LPS), estimated with the BR rodent multi-tissue mortality clock. BH-adjusted p-values, estimated using mixed-effect models, are shown in text. Data are tAges ± s.d. **b,** Chronological tAge dynamics during culturing of human primary fibroblasts (n=65), estimated with the EN universal chronological clock. Pearson's r between tAge and culturing time for each patient and corresponding p-values are shown in colored text. **c,** Mortality tAge dynamics during culturing of control (n=30) and hTERT-induced (n=18) WI-38 cells, estimated with the EN universal mortality clock. Batch-adjusted tAge is plotted on the y-axis. Pearson's r between tAge residual and culturing time for each condition and corresponding p-values are shown in colored text. Differences in intercept and slope were assessed using linear models; corresponding p-values are labeled in black. **d,** Mortality tAge of mouse and naked mole rat (NMR) skin (SF) and embryonic (EF) fibroblasts under control and γ-irradiation conditions (n=3 per group), estimated with the BR rodent multi-tissue mortality clock. Comparisons were performed with a mixed-effect model, adjusted for cell line ID; BH-adjusted p-values are shown. Data are tAges ± s.d. **e,** Mortality tAge residuals of control (n=46), oligomycin-treated (n=10), and 2-deoxyglucose-treated (2-DG; n=20) human primary fibroblasts, estimated with the BR universal mortality clock. Only samples cultured ≥40 days were included. Mixed-effect models were used to assess pairwise differences, adjusted for donor and culturing time; BH-adjusted p-values are shown. Data are tAge residuals ± s.d.




**Review Figure 4. a,** Standardized mortality tAge differences induced by LPS in brain, assessed with module-specific rodent multi-tissue mortality clocks. Data are mean standardized tAge differences ± s.e. **b,** Aggregated standardized mortality tAge differences induced by caloric restriction (CR) in liver, estimated using module-specific rodent multi-tissue mortality clocks and aggregated across datasets, strains, and sexes with mixed-effect models. Data are standardized tAge differences ± s.e. **c,** Standardized mortality tAge changes after oligomycin or 2-DG treatment, assessed with module-specific universal mortality clocks. Significance was assessed with ANOVA, adjusted for patient ID and culturing time. **d,** Standardized chronological tAge changes induced by reprogramming of human fibroblasts, assessed with module-specific universal chronological clocks. Differences were tested using ANOVA, adjusted for donor ID. Color indicates general shift of module expression, measured by NES through GSEA. Data are estimates of standardized tAge difference ± s.e. ^ p.adj < 0.1; * p.adj < 0.05; ** p.adj < 0.01; *** p.adj < 0.001.

Line 620: "Molecular mechanisms of biological age" - how do the authors know biological age has increased? A marker that correlates with age may or may not represent biological age. This is an issue in many other places and I suggest changing terminology (ie, "biological age" to "transcriptomic age") throughout the manuscript.

Thank you for this valuable suggestion. To improve clarity and avoid potential ambiguity, we have replaced the term “biological age” with “transcriptomic age” throughout the manuscript. In addition, we have introduced a glossary (Extended Data Table 1) defining all key terms used in this study and describing how each was calculated, to ensure consistent terminology and interpretation.

Line 709: "Molecular mechanisms of rejuvenation induced by heterochronic parabiosis" - is parabiosis rejuvenation? I don't think the authors show this, and extraordinary claims require extraordinary evidence. Lifespan extension from parabiosis is modest, not what one would expect from rejuvenation.

We would like to thank the reviewer for this important point. Multiple studies have demonstrated that heterochronic parabiosis can reverse aging-related phenotypes in old mice, resulting in enhanced stem cell function and remyelination, increased neurogenesis and angiogenesis, improved bone repair, reduced cardiac hypertrophy, and decreased epigenetic age, often restoring these features to levels typical of animals younger than at treatment onset (PMID: 24797482, 25988592, 22226359, 24797481, 23663781, 37500973). However, we agree that the term "rejuvenation" may overstate these effects. To ensure scientific rigor and avoid overinterpretation, we have replaced "rejuvenation" with more conservative terms such as "age deceleration" or "anti-aging effect" when referring to the effects of heterochronic parabiosis (lines 630, 632, etc.).

Line 734: "The top gene driving the rejuvenating effect of HPB" - how do the authors know this gene is driving the rejuvenating effect?

Thank you for pointing it out. Consistent with our response to similar comments above, we have replaced "driving" with "contributing" (line 648), indicating that this gene had one of the strongest associations with the observed anti-aging effect according to the clock model, without implying a causal role in affecting age deceleration.

Line 1085: "Several established models of organismal rejuvenation, including heterochronic parabiosis in old mice, early embryogenesis, and cellular reprogramming." - embryogenesis and cellular reprogramming are not organismal models. And as mentioned above I don't think heterochronic parabiosis is an established model of rejuvenation either.

We thank the reviewer for pointing out this omission. To ensure scientific accuracy and avoid misinterpretation, we removed the term "organismal" and replaced "rejuvenation" with "age deceleration" (line 787). The revised text now refers more broadly to models associated with age deceleration at cellular or tissue levels, without implying full organismal rejuvenation.

Further specific comments:

Line 155: "this mortality model captures differences in survival dynamics across treatments and sexes, and can be used to estimate expected hazard rate and maximum lifespan" - did the authors validate their model in data not used for training the model? What about a different dataset?

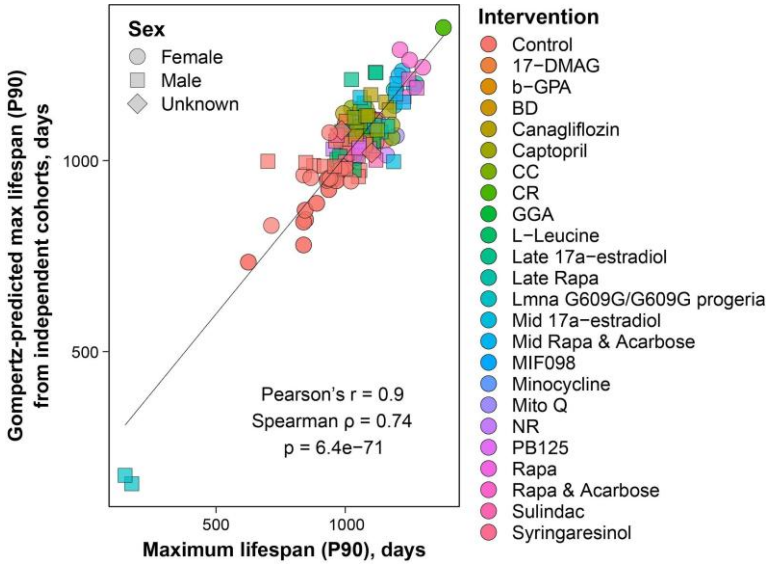
We thank the reviewer for this important comment. We would like to clarify that the Gompertz mortality model was not used as a predictive model per se, but rather as a way to derive expected mortality rate (log-hazard) for each gene expression sample. This expected mortality served as a target variable for training transcriptomic clocks. Thus, Gompertz modeling was used to quantify mortality rather than predict it in independent datasets.

Nevertheless, to ensure robustness and reproducibility of estimated maximum lifespan values (P90) across independent survival cohorts for the same strain, sex, and intervention, we performed a leave-one-cohort-out (LOCO) validation. Specifically, for each experimental group represented by at least 2 cohorts, we fitted Gompertz models to survival data from all but one cohort, averaged Gompertz parameters across these training cohorts, and then used them to estimate maximum lifespan for the remaining cohort. Predicted and observed values were highly correlated (Pearson's $r = 0.9$, Spearman's $\rho = 0.74$; Review Fig. 5), indicating overall consistency of mortality estimates across independent cohorts. Residual inter-cohort variability reflects biological and experimental variation in survival studies rather than limitations of the Gompertz model, as similar variability across cohorts was observed when maximum lifespan was calculated directly from the survival data. To minimize this variability when assigning expected mortality to

transcriptomic samples, we used the average of cohort-specific and aggregated Gompertz estimates for each animal.

Importantly, the transcriptomic clocks trained using these Gompertz-derived mortality values were extensively validated with independent datasets. Using a leave-one-dataset-out approach, the clocks showed high accuracy when applied to unseen datasets (Pearson's $r \geq 0.81$; Extended Data Fig. 2g, 2j–k). To further test generalizability across species, we performed leave-one-species-out validation, where clocks were trained on 3 species and tested on the remaining one. Even under this stringent setting, clocks reliably captured aging trajectories (Pearson's $r \geq 0.66$; Extended Data Fig. 3g).

Together, these analyses indicate that Gompertz-based mortality metrics are reproducible across independent survival cohorts, and transcriptomic clocks trained using these metrics generalize across datasets, tissues, and species.




**Review Figure 5.** Correlation between observed maximum lifespan (P90) and values predicted by Gompertz models fitted to survival data from independent cohorts across rodent strains, sexes, and interventions.

Line 226: "9059 to 9167 signatures of chronological age" - you mean genes associated with age? Isn't that nearly half the genome associated with age?

This is correct. Because we analyzed over 3,500 rodent samples, the large statistical power allowed us to detect even subtle age-related changes in gene expression, resulting in many statistically significant signatures of aging (with BH-adjusted $p < 0.05$; Fig. 1g). To ensure transparency and enable detailed evaluation of these biomarkers in terms of their statistical significance and effect size, we provide the full results of signature meta-analysis in Supplementary Table 2.

Line 394: The co-expression analysis largely confirms previous results, such as inflammation and immune responses, mitochondria, ECM. I don't think it is particularly novel, unless I am missing something?

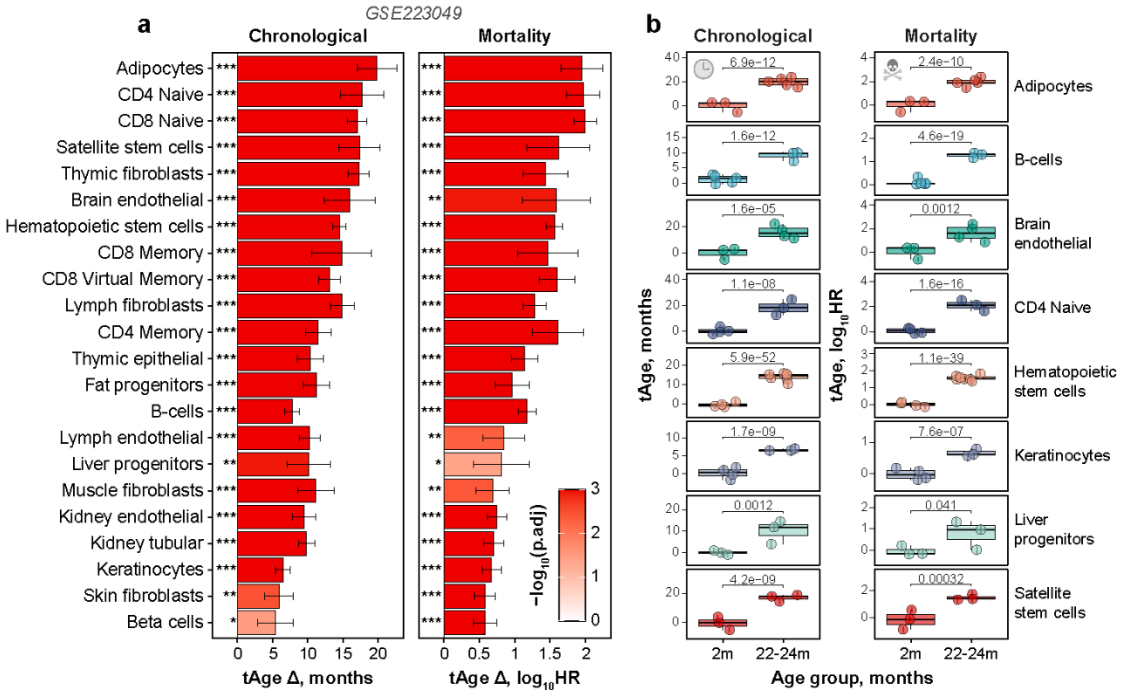
We thank the reviewer for this comment. We agree that the associations of inflammatory, interferon-related, and mitochondrial modules with aging and mortality are consistent with our signature-based results (Fig. 1i) and with previous studies in the field (PMID: 19189975, 33611312, 37269831, 33454277). The main purpose of this section is not to highlight these known associations, but rather to define unbiased, co-regulated gene modules that can serve as the foundation for a compact panel of module-specific clocks of chronological age and expected mortality, each reflecting the behavior of a distinct cellular subsystem. As

demonstrated in the manuscript, this panel enables systematic and interpretable assessment of biological pathways most responsive to given health-modulating interventions.

In this context, the association analysis was performed primarily to validate the functional coherence of the derived modules by confirming that they exhibit expected relationships with aging and longevity. To better reflect this intent, we have revised the opening paragraph of this section accordingly.

Line 513: The authors applied clocks to cell types using publicly available single cell data and show increases in transcriptomic age for most cell types. I think this is an interesting analysis and one of the strengths of the manuscript.

We thank the reviewer for appreciating the importance of this result. We agree that this is one of the most compelling findings of our study. To further support this conclusion, we additionally applied our multi-tissue chronological and mortality clocks to bulk transcriptomes of FACS-sorted cell populations from young (2-month-old) and old (22–24-month-old) C57Bl/6 females. Both clocks significantly distinguished young from old samples across all tested cell types (BH-adjusted $p < 0.05$; Extended Data Fig. 4c-d, Review Fig. 6), including cell types not present in Tabula Muris, such as adipocytes, hematopoietic stem cells, and pancreatic β-cells. This further supports the presence of a consistent transcriptomic aging signature shared across diverse cell types that can be captured by our multi-tissue models.




**Review Figure 6. a,** tAge difference of individual FACS-sorted cell types from young (2-months) and old (22-24-month) C57Bl/6 female mice estimated with BR multi-tissue chronological (left) and mortality (right) clocks (n=3-5 per group). Differences between age groups were tested with mixed-effect models. **b,** tAge of individual representative FACS-sorted cell types from young and old mice estimated with the BR multi-tissue chronological (left) and mortality (right) clocks. BH-adjusted p-values from mixed-effect models are shown in text. Data are tAges ± s.d. * p.adj < 0.05; ** p.adj < 0.01; *** p.adj < 0.001.

Line 667: again, this has been reported by many others (as acknowledged by the authors), can shorten this section in my view.

Thank you for the suggestion. We agree and have shortened this paragraph accordingly.

Line 874: "transcriptomic clocks of mortality predicted a statistically significant acceleration of biological age for patients diagnosed with all examined diseases compared to healthy people after adjustment for chronological age and sex" - if unhealthy individuals have an acceleration of clocks then I don't think we can say that biological age is accelerated but rather that the clock is not reflective of aging but of health. This goes back to my point above that I don't think the authors are showing that biological age is being affected and should tone down their conclusions.

We thank the reviewer for raising this important point. We agree that the term “biological age” can be interpreted in different ways and may imply specificity to aging rather than overall health status. In our study, the clocks capture transcriptomic signatures associated with age-related functional decline and disease burden, and, therefore, reflect an individual’s health status and the level of damage accumulation. To avoid ambiguity, we have replaced “biological age” with “transcriptomic age” in line 592 and throughout the manuscript and clarified this interpretation in both the main text and the glossary (Extended Data Table 1).

Line 935: "At the same time, several interventions resulted in the decrease of biological age in murine tissues." - doesn't this include the ITP interventions from which the data was trained on?

Some interventions with lifespan-extending effects, such as rapamycin, were indeed included in the ITP datasets used during model training. However, our screening also identified multiple interventions reducing transcriptomic age that were not part of the training data (e.g., estrogen, progesterone, ouabain). Nonetheless, in response to another reviewer’s suggestion to streamline the manuscript and focus on core findings, we have removed the section on intervention screening using the Clockbase platform.

Line 1008: what do you mean by "molecular hallmarks of aging"?

Thank you for bringing up this question. In our manuscript, the term “transcriptomic hallmarks of aging” refers to the collection of gene expression signatures reflecting aging of different cellular pathways and functional subsystems.

Line 1131: "Overall, our study resolves the challenge of predicting biological age and mortality by developing a comprehensive transcriptomic model encompassing both aging and lifespan- modulating interventions." - is there one model or multiple ones? I thought there were different clocks, for age, mortality and maximum lifespan, not one single clock. So in practice researchers will need to use different clocks for different purposes, not one single model or clock.

We thank the reviewer for this clarification. In our study, we developed several transcriptomic models predicting different target traits: chronological age, normalized age, expected mortality, and expected maximum lifespan. Different models are intended for different applications, and therefore we have made all of them publicly available. At the same time, the mortality clock integrates both age-related decline and the effects of lifespan-modulating interventions, and therefore is proposed as the primary model for assessing overall organismal health status, as supported by its high performance in predicting both chronological age and the effect on lifespan (Fig. 1o-p; Fig. 3c; Fig. 4h). We have revised the Discussion to make this distinction and rationale clearer (lines 729-740).

The website listed by the authors (http://app.gladyshevlab.org/TACO/) is not publicly available.

We apologize for the earlier limitation. The TACO web application is now publicly accessible at the original link without requiring a password: http://app.gladyshevlab.org/TACO/. Additionally, to facilitate local analysis, we have developed the R package tAge, which enables data preprocessing and application of all transcriptomic clocks presented in this study: https://github.com/Gladyshev-Lab/tAge.

Overall, I think this is an interesting work and the transcriptomic clocks will make a valuable contribution to the field, though some aspects could be improved and I suggest not overselling the findings and shortening the manuscript.

We are thankful to the reviewer for his positive assessment and constructive suggestions. In response, we have carefully revised the manuscript to remove overly strong claims and improve clarity. We have also substantially shortened the text, reducing the number of main figures from 7 to 6, supplementary figures from 28 to 10, and decreasing the main text from 13,337 to 8,575 words.

It is possible that some of my comments reflect misunderstandings of mine. If so then I would suggest that the authors use my misunderstandings as an indication that such points might be made clearer in the manuscript.

We found many of the comments extremely useful, and they helped us to improve the paper.

**Referee #2:**

This manuscript by Gladyshev and colleagues reports an extremely extensive, but also incredibly lengthy and meandering, study of transcriptional changes during rodent aging. The authors formulate a series of transcriptomic "aging clocks" which are quite interesting, especially the contrast between the three different types of clock they define (chronological, mortality, and maximum lifespan). Additionally, they reduce accurate clock predictions to sets of expression changes within gene modules (which have specific functional interpretations), addressing the widespread previous criticisms of aging clocks as predictive but without biological insight. While there have been previous studies of transcriptional aging clocks, their accuracy has been moderate-to-low and they have been applied to predict chronological age only, rather than the three aging types of the present work. Given this landscape, the results of the present manuscript are a significant contribution to the aging clock literature.

We sincerely thank the reviewer for recognizing the scientific significance of our work and its contribution to the field of aging biology. We are glad that the reviewer found the development of transcriptomic clocks for different dimensions of aging novel and valuable, and appreciated our efforts to link these clocks to interpretable biological modules.

To further strengthen this aspect, we have extended our analysis and developed module-specific multi-species clocks trained on conserved co-regulated gene clusters shared between rodents and primates (Fig. 4h; Extended Data Fig. 7j-l). We also validated the module-specific clocks across several established *in vivo* and *in vitro* models, demonstrating their ability to reveal molecular mechanisms mainly targeted by the interventions:

- LPS injection in mouse brain, accelerating transcriptomic age of immune-related modules (inflammation, interferon signaling, etc.) (Fig. 3e, Extended Data Fig. 5i-j, Review Fig. 4a);
- Caloric restriction in mouse liver, producing a geroprotective effect on metabolic modules (lipid metabolism, heme metabolism, cholesterol metabolism) (Fig. 3e, Extended Data Fig. 6a-b, Review Fig. 4b);
- Oligomycin and 2-DG treatment, inducing strong pro-aging effect on metabolic modules, including fatty acid metabolism, oxidative phosphorylation, and heme metabolism (Fig. 4j, Review Fig. 4c);
- Cellular reprogramming, markedly reducing transcriptomic age through differentiation/ECM-related modules (Fig. 4i, Extended Data Fig. 7m, Review Fig. 4d).

Together, these analyses provide use-case examples of module-specific clocks and demonstrate their power to uncover molecular pathways that mediate changes in transcriptomic age induced by a given intervention.

Below, we provide detailed responses to all other comments and suggestions raised by the reviewer.

Unfortunately, the communication of all of the results is overwhelming in volume. The paper comes in at 13,337 words of main text and a very extensive supplement with a grand total of 28 supplementary figures each with numerous subpanels. Despite this volume, the Methods sections needed to understand some of the key methodology are lacking in clarity and sufficient organization. Together, these factors made reading this manuscript frustrating and slow.

We agree with the reviewer that the original version of the manuscript was overly long and dense, complicating readability. In the revised version, we have substantially condensed and reorganized the text, reducing the main text from 13,337 to 8,575 words, the number of main figures from 7 to 6, and the number of supplementary figures from 28 to 10. Following the reviewer's advice, we have also restructured and updated the Methods section to improve clarity and accessibility. Together, these changes have made the manuscript more concise, better organized, and easier to follow while preserving all key results and interpretations.

In summary, some of the science described in this paper is indeed quite exciting but it needs extensive revision to put it mildly.

Major concerns:

On balance, improving the organization of the methods section may be the lowest hanging fruit. For example, the calculation of transcriptomic signatures is spread across methods sections titled "Preprocessing and analysis of ITP cohort RNA-seq data" and "Transcriptomic signatures of aging, mortality and lifespan on the aggregated meta-dataset." The WGCNA methods are also spread across multiple sections. It would be helpful if each analysis was discussed only in its own methods section. We frequently thought we had read all the methods related to a given analysis only to find another section further discussing it. We'd recommend splitting up the very long method subsections into descriptively titled shorter sections; creating a table that briefly defines each type of tAge, signature, or frequently used value and provides the name of the one methods section where the reader can find a further description of how this value was calculated; and a more intuitive naming of each value/variable.

We thank the reviewer for this valuable suggestion. Following their advice, we have substantially reorganized the Methods section to improve its structure and clarity. Specifically, we (1) consolidated descriptions of related analyses into unified subsections (e.g., WGCNA-related methods are now grouped together); (2) divided lengthy subsections into shorter, descriptively titled parts (e.g., "*Development of transcriptomic clocks*" now has multiple subsections, including "*Training data and feature preparation*", "*Supervised machine learning models*", "*Training procedure and cross-validation strategy*", "*Nested cross-validation performance assessment*", etc.); and (3) added a comprehensive glossary table (Extended Data Table 1) that defines key terms and metrics used throughout the paper, briefly describes how they are calculated, and references the corresponding Methods subsection for more details. Additionally, we introduced Supplementary Table 4, which lists all composite and module-specific transcriptomic clocks developed in this study, along with their main characteristics (predicted traits, utilized tissues, applied machine-learning algorithms, test-set performance, number of non-zero features, etc.). Together, these changes improve the organization and readability of the Methods section, making it easier for readers to navigate and follow the analyses.

There are an overwhelming number of different signatures and scores calculated from adjusted and unadjusted values related to lifespan. A table mapping each such value and signature to a brief description of how it was calculated is necessary. Relatedly, some of the names chosen for certain values are quite confusing (e.g., "Adj. lifespan-adjusted age"). Perhaps the table will be enough to clear this up but more intelligible names would be helpful.

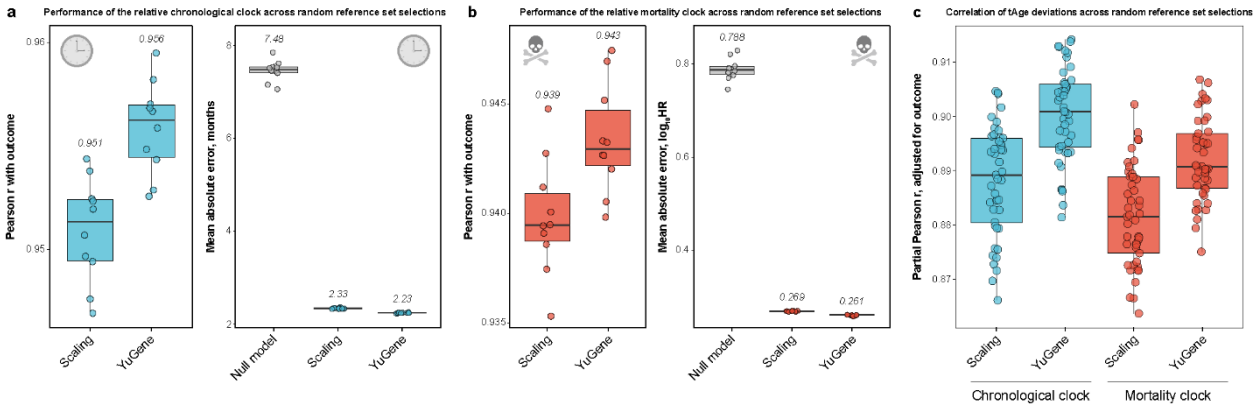
Thank you for this helpful comment. To enhance clarity, we have added a detailed glossary table (Extended Data Table 1) that defines key traits and signatures used in the study and provides concise descriptions of how each variable is calculated (with cross-references to the relevant Methods subsections). This addition facilitates navigation through the terminology and understanding the relationships between derived values. In line with the reviewer's recommendation, we also replaced the term "Lifespan-adjusted age" with the term "Normalized age" and ensured consistent terminology throughout the manuscript and figures.

The single sentence description (lines 1389-1390) of how the relative transcriptomic age clocks (and relative mortality clocks) were trained is insufficient. It seems from the main text that a randomly chosen reference group was used for each dataset, how does clock prediction accuracy vary depending on the random choice of reference group? If all datasets are normalized to the same reference group how do the results change? Is the performance of the relative clock robust to the random reference groups chosen (i.e., if the clock is re-trained N times with a different reference chosen for each group each time are age predictions of individual samples consistent?)

We thank the reviewer for raising this important point. A detailed description of how relative expression datasets were derived has been added to the Methods "*Relative expression adjustment for batch correction*" subsection (lines 1303-1321). Briefly, within each dataset and tissue, we randomly selected an age-matched reference group of control samples and subtracted their median log-expression from all samples in the same dataset and tissue. These relative expression values were then used to derive transcriptomic signatures of chronological age, normalized age, expected mortality, and expected maximum lifespan, as well as to train the corresponding transcriptomic clocks.

We agree that the robustness of this approach to the choice of reference group is an important consideration. To address this, we systematically tested the effect of random reference group selection on clock performance, following the reviewer's suggestion. Specifically, we re-trained the EN rodent multi-tissue clocks of chronological age and expected mortality 10 times, each time randomly reselecting reference groups within each dataset and tissue. For each iteration, we performed leave-one-fold-out (LOFO) cross-validation, iteratively training the clock on 9 folds and evaluating it on the remaining one, thereby generating test predictions across the full meta-dataset.

Clock performance remained highly consistent across all iterations, with median Pearson's correlations between predicted and true outcomes of 0.951-0.956 for chronological age and 0.939-0.943 for mortality (Extended Data Fig. 2c,i; Review Fig. 7a-b), and standard deviations of Pearson's $r \leq 0.0027$ for all clocks. Furthermore, partial correlations of clock predictions across iterations after adjustment for the outcome value were strongly positive (median $r \geq 0.88$; Review Fig. 7c), indicating that deviations in transcriptomic age are also stable with respect to the reference group selection. Thus, relative-expression-based training strategy appears to be both robust and effective. This approach substantially improves clock performance, particularly on independent test datasets, by reducing batch effects and inter-tissue variability (Extended Data Fig. 2f-g). By calibrating each dataset and tissue to its own internal reference group, it minimizes technical variability and preserves generalizability across diverse biological contexts.




**Review Figure 7. a-b,** Performance of relative EN rodent multi-tissue clocks of chronological age **(a)** and expected mortality **(b)** across 10 random selections of reference groups. Pearson's r and MAE are shown for each iteration; medians across the 10 runs are indicated in text; null model MAE is in grey. **c,** Partial Pearson correlations of tAge predictions after adjustment for the outcome values between all pairs of reference group iterations for the EN rodent multi-tissue clocks of chronological age (left) and expected mortality (right).

Before showing the new multi-species model (Figure 7) they should show how the mouse/rat clock(s) from the first 6 figures perform on human samples. If they do not perform well they should try to explain why. If the claim is that "transcriptomic hallmarks of aging and mortality appear to be generally conserved across species" (lines 877 - 878), why is it that they must include human samples in the training process?

We thank the reviewer for this insightful question. To address whether transcriptomic biomarkers of aging are conserved across mammalian species, we performed 2 complementary analyses.

(1) Correlation of transcriptomic signatures across species.

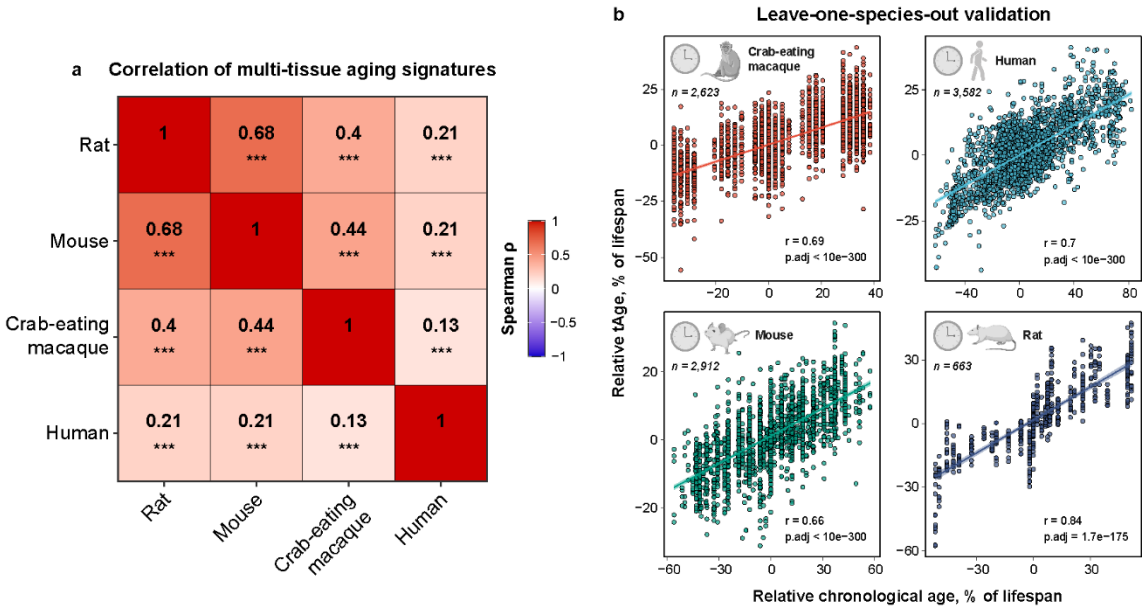
To increase the comprehensiveness of our multi-species meta-dataset, we further expanded it by incorporating 2,623 RNA-seq samples from >30 tissues of male and female crab-eating macaques (*Macaca fascicularis*) of different ages (from 4 to 19 years old; data from PMID: 39270656). For each species (mouse, rat, human, and macaque), we independently identified transcriptomic signatures of aging both across all tissues and within representative tissues (brain and skeletal muscle). We then compared age-associated slopes of log-expression across species and observed significant positive correlations for every pair of species, both across tissues (Fig. 2c; Review Fig. 8a) and within individual tissues (Extended Data Fig. 3h-i). Moreover, many previously identified rodent biomarkers of aging and mortality, including *Gpnmb*, *Cdkn1a*, *Vsig4*, and *Nrep*, showed consistent age-related changes across all 4 species (Fig. 2b), indicating conservation of transcriptomic hallmarks even at the level of individual genes.

(2) Leave-one-species-out (LOSO) validation of transcriptomic clocks.

To directly test whether transcriptomic clocks trained in one set of species can predict aging in an unseen species, we implemented LOSO cross-validation by iteratively training the EN multi-tissue chronological clock on 3 species and testing it on the remaining one (Extended Data Fig. 3g; Review Fig. 8b). Remarkably, the clock maintained robust predictive power even when applied to new species completely excluded from training, with Pearson's $r \geq 0.66$ between predicted and true ages in each species.

Together, these results demonstrate that transcriptomic biomarkers of aging are overall shared among mammals and that clocks trained on several species can successfully generalize to others. Including additional species in the training process further constrains the model to rely on the most conserved hallmarks rather than species-specific effects, thereby enhancing cross-species generalizability while also improving precision for each included species. We therefore trained the final multi-species chronological and mortality clocks on all 4 species, achieving consistently high accuracy in each (Fig. 2a; Extended Data Fig. 3f; Review Fig. 1). These integrative universal clocks provide a robust framework for quantifying

conserved transcriptomic hallmarks of aging and mortality. In this study, we applied them to characterize disease, stress, and rejuvenation processes in both *in vivo* and *in vitro* human systems (Fig. 4a-e,i-j; Fig. 5f-g; Extended Data Fig. 7a-g,m-o; Extended Data Fig. 8h-k).




**Review Figure 8. a,** Pairwise Spearman correlations between gene expression aging slopes across species. Each correlation was calculated using the union of top 500 most age-associated genes (with the lowest p-value) in the pair. Correlation coefficients are shown with text. **b,** Leave-one-species-out performance of EN multi-species multi-tissue clocks of relative chronological age (YuGene normalization). Pearson's r and BH-adjusted p-values are shown in text.

Finally, the authors should consider consolidating or removing some of the results. For example Figures 5 and 6 might be abbreviated and combined. Additionally, the section on the discovery of novel mortality-regulating interventions was a weak point and could be removed or greatly abbreviated.

We thank the reviewer for this valuable suggestion. Following their advice, we substantially streamlined the manuscript by reducing the number of main figures from 7 to 6 and the number of supplementary figures from 28 to 10. For example, Figures 1 and 2 from the original version were merged into a single figure in the revised manuscript (Figure 1). We also agree that the section describing the discovery of novel mortality-regulating interventions via ClockBase was relatively weaker than other parts of the paper. To maintain focus on the most robust and conceptually central findings, we removed this section in accordance with the reviewer's recommendation.

At the same time, we further strengthened several key result sections with additional analyses. In particular, using Framingham Heart Study (FHS) blood gene expression data, we expanded Figure 5 to evaluate how universal transcriptomic clocks are associated with time-to-death in humans and whether their predictions align with those from established DNA methylation clocks. We found that tAge estimated by multiple module-specific mortality clocks (e.g., chromatin modification, inflammation, oxidative phosphorylation) showed strong positive associations with time-to-death, comparable to state-of-the-art DNAm clocks such as DunedinPace and PhenoAge (Fig. 5f; Extended Data Fig. 8i–j; Review Fig. 2a). Moreover, transcriptomic and epigenetic clock predictions in human blood were significantly correlated after adjusting for age and sex (Fig. 5g; Review Fig. 2b), with the strongest concordance observed for the chromatin-related transcriptomic clock, highlighting the link between mortality-associated changes in gene expression and DNA methylation.

Overall, during revision we focused the manuscript on the most novel and scientifically grounded results while shortening or removing less essential sections.

Minor concerns:

Figure 1 and Extended data Fig. 1:

4 associations are referenced in lines 162-164, but only 3 volcano plots are provided in Extended data Fig. 1D-F). The 4th volcano plot should also be shown.

We thank the reviewer for noticing this point. In the revised manuscript, to improve clarity and focus on the most central findings, we removed all volcano plots of signatures derived from the ITP cohort. However, to ensure transparency and accessibility of these results, we now provide a comprehensive Supplementary Table 2 that includes the complete statistical output for all signatures evaluated in the study, including those calculated for the ITP liver data (previously shown on volcano plots), rodent multi-tissue signatures, and multi-tissue aging signatures for all species in our study.

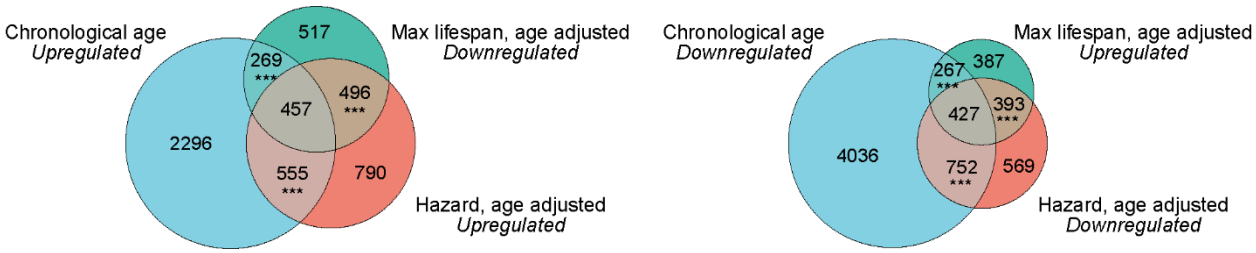
Going from Extended data Fig. 1D-F to Extended data Fig. 1G/Figure 1E it seems that somehow the genes associated with each measure (chronological age, lifespan-adjusted age, expected hazard rate, and maximum lifespan) are integrated into a single "signature" score for that measure. The description of how this is done in the main text and methods is insufficient and should be clarified and expanded. Is each signature defined by the vector of coefficients for each gene coming from the linear mixed-effects model? Or maybe just the vector of coefficients for genes with a significant association? When transcriptomic signatures are compared (as in Extended data Fig. 1G) how is a correlation value defined between signatures, by measuring the correlation between coefficients (e.g., "chronological age slope") of each gene between signatures? Answers to all of these questions should be made obvious.

Thank you for these important questions. We agree that the methodology required further clarification and have expanded this section in the revised manuscript.

To compare transcriptomic signatures (e.g., in Extended Data Fig. 1d-e), each signature was defined as the vector of gene-specific slope coefficients (β estimates) from the linear mixed-effects model for a given trait (chronological age, normalized age, expected mortality, or maximum lifespan). These coefficients reflect both the direction and magnitude of association for each gene. To minimize noise from genes with weak or no association, correlations were calculated using only the union of the top 1,000 genes most significantly associated with each trait in the pair (with the lowest p-values). Thus, for each pairwise comparison, Spearman correlations were computed between two coefficient vectors of 1,000-2,000 genes, depending on the overlap between signatures. This denoised signature comparison strategy has been tested and applied previously (PMID: 37269831, 37500973, 37086720). To compare signatures at the pathway level (e.g., Extended Data Fig. 1i), we calculated correlations between vectors of normalized enrichment scores (NES) derived from GSEA across all gene sets in the KEGG, REACTOME, and Hallmark databases. We have now updated the Methods section to clearly describe these procedures in detail (lines 1430-1436; 1450-1459; 1543-1545).

The size of the circles in Figure 1M should be made proportional to the number of elements represented by the circles.

We thank the reviewer for this helpful suggestion. In the revised version, we updated the Venn diagram, making sizes of the circles proportional to the number of statistically significant genes, as recommended (Extended Data Fig. 1p, Review Fig. 9).




**Review Figure 9.** Venn diagrams of genes significantly associated (p.adjusted < 0.05) with chronological age (blue), expected mortality (red), and maximum lifespan (green) adjusted for age. Circle size is proportional to the number of significant genes. Pairwise overlaps were tested with Fisher's exact test.

As the choice of normalization method scaling vs. yugene did not make much of a difference it does not add much to show both. Just one should be chosen to show in figure 1 and the other can be shown in the supplementary figures.

Thank you for this suggestion. To streamline the figure, we followed the reviewer's advice and retained only one normalization method (YuGene) for displaying mortality clock coefficients in Fig. 1n, since (1) its performance was demonstrated in Fig. 1m, and (2) the coefficients were highly consistent across normalization methods (Extended Data Fig. 2m). As for the leave-one-tissue-out validation of the multi-tissue clock (Fig. 1l), we kept both normalization methods on the figure as we observed subtle but meaningful differences across them: while the scaling-based clock showed slightly lower median accuracy across tissues overall, it outperformed the YuGene-based clock in blood, skin, and stomach (tissues where the YuGene-based clock showed the lowest accuracy). Therefore, we believe it is informative to present both normalization methods for this specific analysis.

Extended data Fig. 2

The paragraph from line 176 to 194 should be broken into two paragraphs as the discussion of particular genes and then gene set enrichments are unrelated and it is confusing to have them grouped together.

We thank the reviewer for this suggestion. In the revised manuscript, we have separated the discussion of top individual genes from enriched pathways associated with aging, mortality, and lifespan into two distinct paragraphs (lines 142-155).

It would be clarifying to have Extended data Fig. 2F come before Extended data Fig. 2E, as I believe Extended data Fig. 2E represents the pairwise correlations between the columns of Extended data Fig. 2F.

Thank you for this helpful suggestion. We have reordered the figures so that the functional enrichment results (Extended Data Fig. 1h) now precede the pathway-level correlation matrix (Extended Data Fig. 1i), making the logic of the analysis clearer.

Also the associated text is misleading; the associations between the signatures at the level of genes is shown in Extended data Fig. 1G while Extended data Fig. 2E shows the association between signatures at the level of pathways. This section should be rewritten to make this distinction between the heatmaps Extended data Fig. 1G, Extended data Fig. 2E, and Extended data Fig. 2F clear.

We thank the reviewer for pointing out this omission. We have revised the corresponding text in the manuscript (lines 134-155) to clearly distinguish between the results shown in these figures: gene-level correlations between signatures (Extended Data Fig. 1d), enriched pathways (Extended Data Fig. 1h), and pathway-level correlations between signatures (Extended Data Fig. 1i). In addition, we added explicit panel titles to the figures to make these distinctions clear even without referring to the legends.

Extended data Fig. 3D is referenced before Extended data Fig. 3C.

Thank you for pointing this out. We have updated the figure and reordered the panels (Extended Data Fig. 1m-n) so that they are now referenced in the correct order.

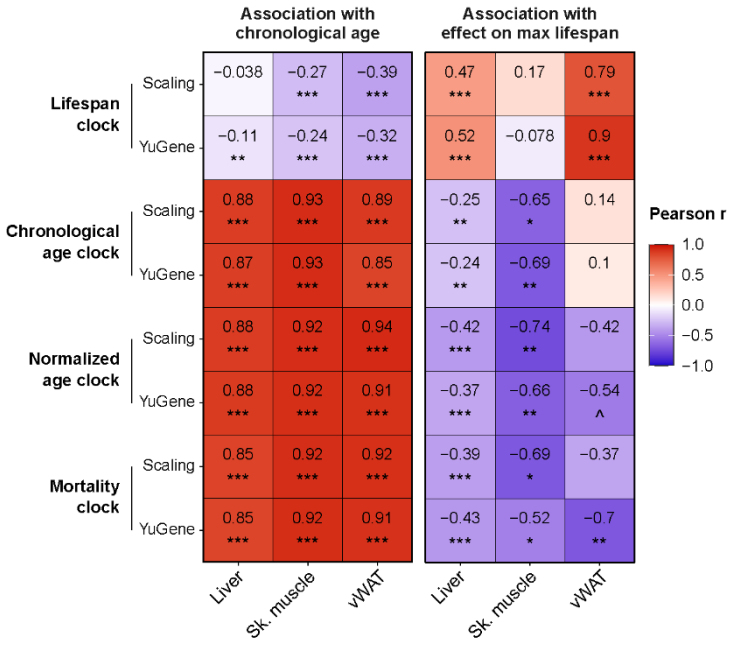
Figure 2:

It would be illuminating to see how many features are selected for use in the final model (by penalized regression for example) by each of the clocks trained in Figure 3.

We would like to thank the reviewer for this great suggestion. We have added Supplementary Table 4, which provides detailed information for each composite and module-specific transcriptomic clock developed in this study, including the number of non-zero features selected by the Elastic Net models as well as the performance of each clock on the test set (based on nested cross-validation).

How did the tissue type used for transcriptomic profiling affect the accuracy of prediction of lifespan extension shown in Figure 2H-J and Extended data figure 6E-F? One way to show this would be to extend Figure 2H into a heatmap where each column is a tissue, each row is a clock, and each value is the association between that clock's prediction in that tissue and the effect on lifespan.

We thank the reviewer for this question. Following their advice, we evaluated the performance of the multi-tissue transcriptomic clocks in predicting chronological age and intervention-induced effect on maximum lifespan across individual tissues. Specifically, we analyzed 3 tissues most represented in our meta-dataset in terms of available lifespan-modulating interventions, including liver, skeletal muscle, and visceral WAT (Extended Data Fig. 3d, Review Fig. 10). Consistent with the overall performance patterns shown in Fig. 1o, the clocks of normalized age and expected mortality (particularly those based on YuGene normalization) demonstrated the strongest and most consistent performance, accurately predicting both chronological age and lifespan effects across tissues.




Review Figure 10. Performance of EN rodent multi-tissue transcriptomic clocks in predicting chronological age (left) and intervention's effects on maximum lifespan (log-ratio, right) in individual tissues (liver, skeletal muscle, visceral WAT). ^ $p < 0.1$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

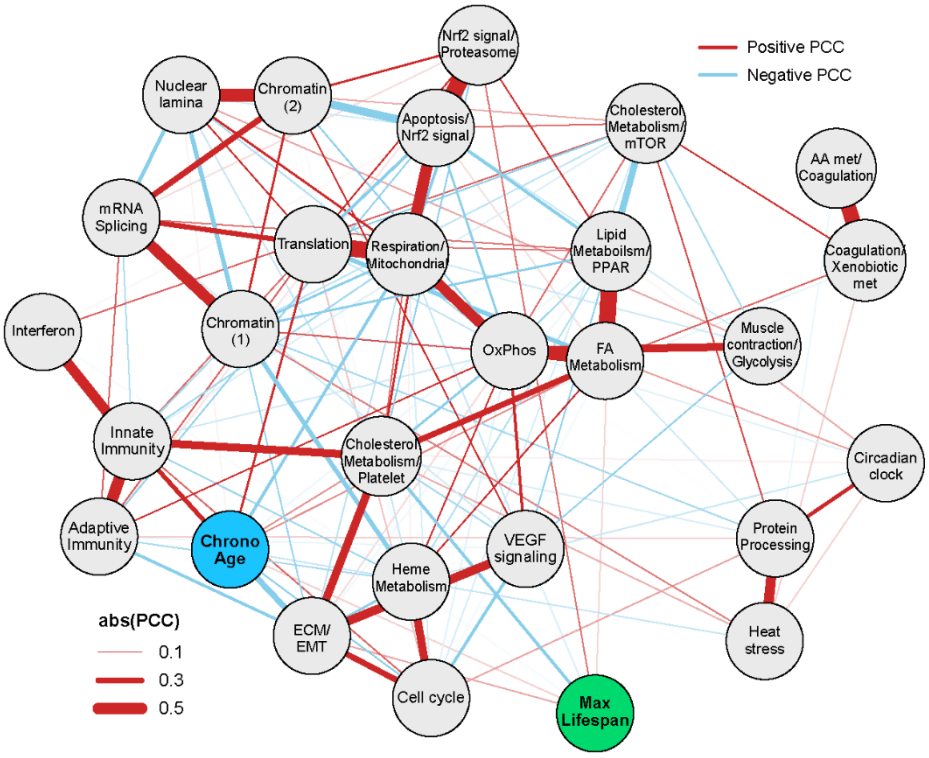
Extended data Fig. 5 panels are not referenced in order in the text. Nor in order with extended data figure 6.

We thank the reviewer for pointing out this issue. In the revised manuscript, we have carefully reordered the Extended Data Figures and ensured that all panels are referenced in the text according to their order of appearance.

Figure 3

The labels in Figure 3C are not legible due to their small size and insufficient image resolution. Some sort of legend mapping the thickness of the line to the correlation coefficient is needed.

We thank the reviewer for this helpful comment. We have revised all network figures (Extended Data Fig. 5d, f-g; Review Fig. 11) to improve readability by increasing label font size and adding a legend showing the correspondence between line thickness and the absolute value of partial correlation coefficient.




Review Figure 11. Partial correlation network of 1st principal components (PCs) of co-regulated modules (grey), chronological age (blue), and expected maximum lifespan (green). Edge color and width indicate the sign and magnitude of partial correlation coefficients (PCC). Modules are labeled by representative enriched functions.

The x-axis labels in Figure 3E are confusing, particularly for the right two barplots.

Thank you for highlighting this issue. We have updated the x-axis labels in this figure (Fig. 3c) to make them clearer and more self-explanatory. They are now labeled “Pearson r with predicted outcome” and “Pearson r with trait.” We also revised the figure legend to provide a detailed explanation: “Left: test predictive accuracy (Pearson r) of clocks of relative chronological age (blue) and expected mortality (red). Right: Pearson correlations between tAge predicted by mortality clocks and chronological age (blue) or age-adjusted expected maximum lifespan (green)”.

Extended Data Figures 9 and 10 are never referenced in the text.

To decrease the number of supplementary materials, we have removed these figures from the revised manuscript. The pathway enrichment information they contained is presented more compactly in Extended Data Fig. 5c, and the complete enrichment results for all modules are available in Supplementary Table 6.

Panels are not referenced in order in the text (I and H).

We thank the reviewer for pointing this out. We have revised the figures to ensure that all panels appear in the same order as they are referenced in the text.

What do the last two columns refer to in Figure 3K? Were samples binned into high and low hazard groups and a GSEA was done comparing these groups?

We thank the reviewer for this clarifying question. The last two columns in the original figure represented the overall transcriptomic signatures of expected mortality (unadjusted and adjusted for chronological age) derived from the rodent meta-dataset (as in Fig. 1i). Since the pathway enrichment of these signatures was already presented in Fig. 1i, we removed them from this panel in the revised version. In addition, to make the enrichment analysis more complementary to the module-specific clocks, we reran it using the same module gene sets employed for clock construction. Consequently, the revised heatmap (Extended Data Fig. 6i) now reflects the overall up- or downregulation of genes associated with each module, while the module-specific clocks (Fig. 3h; Extended Data Fig. 6g) indicate the association between module expression changes and expected mortality.

The methods section describing the construction of the module specific clocks is insufficient. Were they relative clocks or not? Was the data yugene or scaled?

We thank the reviewer for noting this omission. We have expanded the Methods section to clarify the procedure for developing the module-specific clocks in detail. All module-specific clocks of chronological age and expected mortality were trained as relative models using the same data as the corresponding composite clocks. Because the clusters of co-regulated genes were identified from scaled data, we trained all module-specific clocks on data normalized with the scaling approach. These and other methodological details have been added to the revised Methods "*Development of module-specific clocks*" subsection (lines 1731-1743).

Figure 4

Panels F and G are referenced out of order. So are H and I.

Thank you for pointing this out. We have updated the text and the order of the panels, ensuring that they are now referenced in the same order as they appear in the figure.

How was tAge normalized in this figure?

We thank the reviewer for raising this question. In the scatterplot panel (Fig. 2h), tAge estimates were standardized across all samples within each cell type: each tAge value was divided by the standard deviation of all tAges calculated for the corresponding cell type and clock. We have updated the y-axis label to clarify this normalization and added a description of this procedure to the Methods "*Aging dynamics across individual cell types*" subsection (lines 1838-1840). To improve data interpretability, we also revised the accompanying barplot (Fig. 2g), which now shows Pearson correlation between tAge and chronological

age on the x-axis. In the gene contribution barplot (Fig. 2i), the x-axis now represents $\log_{10}$(HR) per year (contribution to mortality tAge per unit of time).

Because the identification of conserved, aging-associated changes across multiple cell types is one of the central findings of our study, we further validated these results using an independent data source. In this dataset, cell types from young (2-month-old) and old (22-24-month-old) C57BL/6 female mice were isolated via FACS prior to bulk RNA sequencing. We observed a statistically significant increase in tAge with age across all tested cell types according to both chronological and mortality clocks (BH-adjusted $p < 0.05$; Extended Data Fig. 4c-d, Review Fig. 6), including cell types not represented in the Tabula Muris dataset, such as adipocytes, hematopoietic stem cells, and pancreatic β-cells.

To make the analysis of genes contributing to mortality tAge more robust, we aggregated results from both Tabula Muris and this additional dataset (Fig. 2i). This analysis confirmed that many genes identified as conserved biomarkers of aging and mortality, such as *Lgals3* and *Cdkn1a*, are consistently associated with increased tAge across multiple cell types, supported by two independent data sources encompassing >50 individual cell types.

Extended data figure 15

The assertion on lines 605-609, that the modules show commonalities between the klotho KO associated gene expression changes in proximal tubule cells and neurons, does not seem to be well supported by Extended data figure 15D-E. The most contributive module "Respiration/Mitochondrial translation" an increase in tAge in both cell types, but resulting from an opposite enrichment. Showing Extended data figure 15D-E as a scatterplot, as in Extended data figure 15C, would be more informative. To support your assertion it should be demonstrated that within positive and negative NES, separately, there is a significant positive correlation between the effects on tAge of each module.

We would like to thank the reviewer for this helpful comment and fully agree that the statement would require additional statistical analyses to be adequately supported. Since this section is not central to the main focus of the study, and given the reviewer's earlier recommendation to streamline the manuscript, we decided to remove this paragraph entirely. The revised version now retains only the overall effects of the *Klotho* KO model on individual cell types in the brain and kidney (Fig. 3i-j; Extended Data Fig. 6j-n), which are directly relevant to the core findings.

The line of argument about inflammatory changes in the single cell klotho KO data (lines 608-618) is not persuasive because neither of the cited results in Extended Data Fig. 15F and Fig. 4K reached statistical significance. This section should be removed or supported with results that are significant.

Thank you for this observation. Although genes from the inflammatory module showed a statistically significant negative contribution to mortality tAge when assessed with the composite mortality clock (Extended Data Fig. 15e in the original version of the manuscript), we agree that this effect is relatively modest compared to our main findings. To keep the manuscript focused and concise, we therefore removed this section as suggested, retaining only the summary of the overall effects of the *Klotho* KO model on tAge across cell types in the brain and kidney (Fig. 3i-j; Extended Data Fig. 6j-n).

Discussion

The claim "A 3-month treatment with ouabain was able to decrease the biological age of 24 month-old mice by ~10 months suggesting a first example of a radical in vivo rejuvenation effect induced by a single-molecule pharmacological intervention" (lines 1105-1106) is undermined by the fact that other single-molecule pharmacological interventions show similar if not greater "rejuvenations" in Extended Data Fig. 28A (e.g., metformin and other unlabeled points). This statement should either be walked back or expanded

upon to explain why metformin and the other points shown in Extended Data Fig. 28A are less impressive than ouabain.

We thank the reviewer for this comment. We agree that the original wording may have overstated the uniqueness of ouabain, especially considering that other single-molecule interventions demonstrated comparable effects in Extended Data Fig. 28A. In line with another reviewer's suggestion, we have removed this screening section to shorten the paper and to maintain the manuscript's focus on the most robust and interesting findings.

Data availability

A large amount of time and effort has clearly been devoted to identifying, downloading, cleaning, and combining numerous disparate transcriptomic datasets. To facilitate the replication of your results a script or set of scripts should be provided (along with the paper or on github) which download (from GEO) and combine the (public) analyzed datasets into the form analyzed in the paper. This will allow other researchers to benefit from, and build upon, the work you have done to aggregate these datasets. The TACO web-app is a good step on the way to this, but does not allow other researchers to directly analyze these dataset(s) for new insights.

Thank you for this suggestion. We fully agree that enabling reproducibility and user-friendly application of our aggregated datasets is essential for advancing the field. In addition to the TACO web application, we have implemented the following measures to facilitate further use of our data and models:

- We developed the *tAge* R package, which allows users to locally reproduce our preprocessing pipeline and apply all transcriptomic clocks described in the manuscript to any gene expression data. The package includes scripts for dataset preprocessing (including scaling and YuGene normalization) and model application: https://github.com/Gladyshev-Lab/tAge.

- To make the data aggregated in this study directly accessible for the field, we have uploaded the preprocessed rodent gene expression meta-dataset (Fig. 1e), along with comprehensive sample metadata: https://bit.ly/4a3pzer.

These resources allow other researchers to (i) reproduce our data integration workflow, (ii) apply our transcriptomic clocks locally or in custom analyses beyond the TACO interface, and (iii) explore the aggregated dataset to generate new hypotheses on molecular signatures of aging and lifespan regulation.

The data availability section says the ITP RNA-seq dataset has been deposited on GEO but does not provide a GEO accession number.

We would like to thank the reviewer for pointing out this omission. We have updated the manuscript, providing a GEO accession number related to the bulk and snRNA-seq data generated throughout the study: SuperSeries ID GSE292885 (accession numbers for individual datasets: GSE292843, GSE292884, GSE293503). We will make them available as soon as the manuscript is accepted for publication. Currently they can be accessed with the following secure token: **stynowyufxqnfer**.

Remarks on Code Availability: The code is well organized and sufficient.

Thank you for these comments, which greatly helped us to improve the paper.

**Referee #1:**

The authors have addressed my suggestions, and the manuscript is much improved.

We thank the reviewer for helpful suggestions and for the positive assessment of the revised manuscript.

The only aspect I couldn't find more information about was that, in my previous comments, there were >9000 genes associated with aging, which is quite a lot, yet in the revised manuscript I couldn't find exactly how many genes are associated with aging or how many genes compose each of the transcriptomic clocks developed by the authors.

We further revised the manuscript to report the exact range of genes significantly associated with aging and other traits (BH-adjusted p-value < 0.05). We also added the complete lists of genes associated with each individual trait in Supplementary Table 2. The number of genes comprising each transcriptomic clock (including the composite and module-specific clocks) is provided in Supplementary Table 4 and highlighted in Fig. 3c and Fig. 4h.

**Referee #2:**

The revised manuscript by Gladyshev and colleagues is significantly improved. All of the major issues identified in peer-review have been sufficiently addressed. The new analyses comparing the tAge model to epigenetic clocks are particularly illuminating and the significantly abridged manuscript length is a great improvement.

It is also appreciated that the authors now have published a processed version of their data for use by the greater scientific community.

We thank the reviewer for the thoughtful and positive evaluation of the revised manuscript. We appreciate the reviewer's recognition of the improvements made during revision. Below, we address the remaining minor issues regarding data file formats.

However, there are two minor issues with this data which should be addressed before publication. Firstly, the metadata is provided as `.gsheet` files, which we are unfamiliar with and have been unable to open; these files would be better provided in a common file format like `.xlsx` or `.csv`. Similarly, the pre-processed expression data is provided as `.rds` files, which are cumbersome for non-R users to analyze; these would be more widely accessible if provided as `.xlsx` or `.csv`.

We thank the reviewer for identifying these accessibility issues. We have updated the deposited metadata file and now provide it in .csv format. We have also uploaded the pre-processed expression matrices in .csv format to improve accessibility for users who do not work in R. All processed data files are now stored on Zenodo (https://doi.org/10.5281/zenodo.18763485).

**Referee #3:**

I co-reviewed this manuscript with one of the reviewers who provided the listed reports.

We thank the reviewer for co-reviewing the manuscript and providing valuable suggestions.