

Transcriptomic reprogramming screen identifies SRSF1 as a regulator of aging

Corresponding Author: Professor George Church

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

In this work, Plesa and colleagues developed a model to infer age from transcriptomic data and implemented a screening to identify age reversal factors through doxycycline-induced gene overexpression. A main motivation was to identify new factors that induce cell rejuvenation without the safety concerns of the classic Yamanaka or OSKM factors (i.e., loss of cell identity and neoplasia). The authors identify SRSF1 as the most promising candidate to induce rejuvenation effects in vitro and in vivo.

This is a very interesting study describing useful tools to assess cell rejuvenation, and the results obtained with SRSF1 overexpression are promising. However, in my opinion, the study is not fully comprehensive and some key pieces of information necessary to fully evaluate the potential of SRSF1 and their approach are missing. Such concerns and suggestions are summarized below.

1. In Figure 1 the authors should confirm the aging trajectories achieved by their model with other methods. For example, well-established DNA methylation clocks of other available transcriptomic clocks (<https://www.biorxiv.org/content/10.1101/2024.07.04.601982v1>). The only direct comparison provided in their manuscript (Figure 1f) was performed with an insufficient n number (n=1/time-point).
2. In Supplementary Figure 1-f, the age of iPSC and ESC controls is negative (~-100 to -150 years). However, it has been well-established by others that reprogramming to pluripotency reverses biological age to ground zero. Can the authors explain this result? A direct comparison to a more established method (e.g., DNA methylation clocks) could help explain whether this is a problem of the proposed model.
3. Sample sizes are missing in many figure legends and should be included.
4. Figure 2c: exposure to doxycycline can have significant effects in gene expression. Have the authors controlled for this? In addition of BFP controls that would also be treated with dox, comparing the gene expression of WT cell with and without the drug would be informative.
5. In Figure 2d the transcriptomic age data is missing for some gene candidates in the M65 sample. Similarly, flow cytometry data for SA- β -gal expression, proteasome activity, and membrane potential were not performed in this cell line due to insufficient cell numbers, according to the authors. This should be completed, especially given that the overall response was rather different in different donor cell lines.
6. Figure 3a provides the age effect (predicted age – age) after overexpression of the top 5 candidate genes in 6 different donor cell lines. However, the experiment was performed with n=2 sample size, which does not allow for assessing the variability of the intervention within a fixed cell line. While increasing the sample size for all five interventions and six cell lines would be very costly, I recommend that the authors perform a larger-sample study for SRSF1 and at least one of the cell lines. How many days after gene induction was the data collected? This information should be stated in the figure legend.
7. Similarly, confirming the reduction of biological age shown in Figure 3a via a DNA methylation assay would add to the robustness of this study.
8. In Figure 3d, can the authors explain why the overexpression of candidate genes induces MSC signatures? Are those signatures also enriched in the donor cells pre-overexpression, and if not, would that not be considered loss of cell identity in a way?
9. The data in Figure 3e does not add much information. The cells treated with OKSM look already different on day 1. Day 0 images should be included as well as morphometric quantification of the cell body area. Scale bars are missing.
10. Figure 3f shows that the induction of SRSF1 for 2-3 days does not achieve long-lasting reversal of transcriptomic age. In fact, most of the effect is lost within 12 days. The authors propose that further dosage and timing optimization needs to be tested. However, the authors do not test further dosages and therefore there is no proof that they would not induce loss of cell identity. This could mean that there is no advantage of SRSF1 against OKSM (many studies have shown that 2-day expression of OKSM does not induce de-differentiation) and therefore needs to be tested.
11. How long do the functional changes (e.g., decrease in s- β -gal activity and others) last after induction? This assessment

is important and should be included, as ultimately, it is cellular function that matters the most.

12. The in vivo experiments (Figure 4), in general, are not comprehensive enough. For example, in Figure 4d, SRSF1 is administered intradermally but there is no information on the cell types that express SRSF1 in vivo upon viral transduction. It is also not clear for how long SRSF1 is expressed for. Characterization of the cells post-SRSF1 expression (e.g., transcriptomic age and age-related phenotype) is also missing. The absence of these data makes it very difficult to interpret the results provided on wound healing in a mechanistic manner.

13. Similarly, for the studies with systemic administration of SRSF1, it is not described how long the expression of the gene lasts after viral transduction. Some kinetics at the gene or protein expression levels are necessary. In fact, in vivo levels of SRSF1 are not assessed at all. Instead, the authors tested the expression of rTA as a proxy, can they explain why SRSF1 levels are not investigated?

14. There is also no mechanistic explanation behind the increased lifespan and reduced frailty of the mice. Since the highest expression (of rTA) is localized in the liver, have the authors tested for liver function biomarkers? Are there paracrine effects expected from SRSF1 in other organs?

(Remarks on code availability)

Reviewer #2

(Remarks to the Author)

In their study, Pleša et al. present a transcriptomic clock based on human fibroblasts and use it to identify genes whose overexpression may reverse ageing-associated gene expression patterns. Their findings highlight SRSF1 as a potential rejuvenation factor that may reverse transcriptomic ageing signatures, reduce SA- β -Gal staining and proteasome activity in human cells, and enhance wound healing, frailty metrics, and lifespan in mice.

While the biological findings are interesting and point to a promising new regulator of ageing, the machine learning pipeline suffers from serious methodological flaws, such as data leakage, which likely lead to overestimated model performance. In addition, many figures lack appropriate indicators of statistical significance (e.g., p-values), hindering interpretation. The in vivo experiments are also limited by small cohort sizes (fewer than 11 treated animals per group) and lack mechanistic validation to support the observed phenotypes. Detailed comments to improve the rigor, clarity, and reproducibility of the study are provided below. Additionally, minor concerns, such as clarifying methods and figures, sharing generated data, and adopting a more cautious interpretation of results, require resolution to enhance the study's credibility.

Major points:

1. During the clock development, PCA, “weak” predictor training, and pathway/gene selection were all performed on the entire dataset, including the folds later used for cross-validation, so information from the validation folds leaked into model construction. Given the multi-layer pipeline and extensive feature selection, this design almost certainly inflates performance. Besides, for the dataset used to train the clock, the authors haven't reported the quality on a test set. The original dataset must be partitioned into separate training and test sets (or assessed via nested leave-one-out approach), with every dimensional-reduction and selection step fitted exclusively on the training split. Model accuracy should then be reported on unseen data, not the training set (Extended Data Fig. 1C).

2. The application of the trained clock to external datasets is also affected by data leakage. According to the Methods, the model is retrained separately for each new dataset, which limits its generalizability and undermines its utility. Ideally, a single, fixed model should be developed and shown to perform robustly across external datasets without retraining. Moreover, before retraining, the authors pool training and external data for TMM normalization and batch correction. Since these procedures use information from the entire dataset, including the test samples, they introduce data leakage by allowing the model to indirectly access the structure of the test data. This likely inflates the reported performance of the clock on external datasets ($r = 0.93$). To rigorously assess predictive performance, normalization and batch correction parameters must be derived solely from the training data and then applied unchanged to external samples. Performance should also be evaluated using a single, pre-trained model.

3. In Fig. 2b the list of age-associated pathways surprisingly omits several established ageing hallmarks, including inflammation and mitochondrial function (e.g., see 10.1093/bioinformatics/btp073). The authors should explain this omission or demonstrate, using a leakage-free analysis, that these signatures do in fact emerge. It would also be informative to compare the pathways identified in fibroblasts with ageing signatures reported at the tissue level (e.g., blood, brain, skeletal muscle) to evaluate how well the fibroblast clock captures organism-wide ageing trends.

4. Many figures lack statistical significance estimates (p-values): Fig. 1d-e; Fig. 2d-g; Fig. 3a; Extended Data Fig. 1e-f; Extended Data Fig. 2a-c, 2e-k; Extended Data Fig. 4; Extended Data Fig. 6b, f; and Extended Data Fig. 8f-h, k. The authors should provide appropriate p-values and apply multiple testing correction where relevant (e.g., in Fig. 2d-g, Fig. 3a). In several instances, effects are described in the text but are not clearly supported by the figures. These claims should be more precisely stated and backed by statistical analysis. For example, the authors assert that predicted age increases in fibroblasts exposed to sun-induced UV damage, but Extended Data Fig. 2b shows no consistent pattern: there is no difference in young donors, an increase in middle-aged, and a decrease in older individuals. Such discrepancies should be clearly addressed and statistically validated.

5. Fig. 2c: The authors use their transcriptomic clock to evaluate the impact of gene overexpression across a perturbation screen. However, some of the perturbed genes may have been included in the clock itself. In such cases, the resulting predicted age changes are likely to be trivial – for instance, overexpressing a gene with a negative coefficient in the model will obviously reduce predicted age. To avoid circular reasoning, the authors should exclude each tested gene from the clock when evaluating its perturbation effect.

6. Fig. 3d: The authors report cell identity enrichment scores for several perturbations and control lines compared to a reference fibroblast group. However, it is strange that WT (NTg) fibroblasts display significant pro-fibroblast enrichment, as they should be nearly identical to the reference fibroblasts and serve as negative controls. This finding suggests that the external reference fibroblast dataset may not be an appropriate comparator, potentially due to batch effects. To reduce this confounding, the authors should consider using NTg or BFP cells as the internal reference group for calculating enrichment scores.

7. Line 241: The authors claim that SRSF1 overexpression does not lead to loss of cellular identity, in contrast to OSKM reprogramming. To support this claim, they should present expression levels of canonical pluripotency markers, such as Nanog, Oct4, and Sox2, across WT, BFP, SRSF1, and OSKM1 cell lines. This would provide a clearer comparison of reprogramming versus rejuvenation.

8. Fig. 4a: The authors report that SRSF1 activation reduces proteasome activity by ~40%. However, this appears inconsistent with Extended Data Fig. 6b, where no clear difference is observed and no p-value is provided. The authors should clarify whether these figures represent different experimental assays or if this discrepancy is due to an error. If distinct assays were used, details of their respective methodologies and sensitivities should be included.

9. Line 257: The authors report that “a transient 3-day SRSF1 mRNA treatment induced ... a 30% decrease in ROS levels upon H₂O₂ exposure.” However, according to Fig. 4b, this effect does not reach statistical significance. The text should be revised to reflect this.

10. Fig. 4d: The authors conclude that SRSF1 induction significantly improves wound healing in old mice, based on the observed wound closure dynamics between SRSF1- and GFP-treated groups. However, Extended Data Fig. 8f suggests that the Tet-GFP plasmid alone impairs wound healing in both young and old mice when compared to untreated controls. This raises the concern that the apparent benefit of SRSF1 may be due to the deleterious effects of the GFP plasmid rather than a true therapeutic effect.

To address this, the authors should compare wound healing in SRSF1-treated old mice to that in untreated old WT mice. Furthermore, while the authors report no significant benefit from constitutive (CAG) SRSF1 expression, it seems that CAG-GFP animals show better healing than Tet-GFP animals, whereas SRSF1 treatment shows similar effects in both contexts. This suggests that the lack of observed benefit from CAG-SRSF1 may result from improved baseline healing in CAG-GFP controls, rather than from reduced SRSF1 efficacy.

The authors should carefully reanalyze their wound healing data using appropriate control groups and clearly distinguish between beneficial effects of SRSF1 and detrimental effects of specific plasmid constructs. Additionally, if AAV delivery itself impairs physiological outcomes, as appears likely from the GFP control wound healing data, its potential impact on other *in vivo* results, such as the reduced lifespan in GFP-treated mice compared to SRSF1 group, should also be critically evaluated.

11. Extended Data Fig. 8l: The authors report an improvement in the frailty index of SRSF1-treated mice but present the data as the age at which each animal reaches a frailty index threshold of 0.18. This threshold appears arbitrary, and the rationale for its selection is not explained. Additionally, the frequency of frailty index measurements for each mouse is not specified and should be clearly detailed in the Methods section.

Typically, frailty index data are presented either as pairwise comparisons between treatment groups at specific timepoints (e.g., before and 3 months after treatment) or as longitudinal trajectories over time. If multiple measurements per animal were collected, the authors should plot the frailty index dynamics for GFP and SRSF1 groups on the same graph and provide statistical analyses (p-values) for group comparisons at each timepoint. This approach would better illustrate the rate of frailty progression in treated versus control mice. Furthermore, it would be valuable to analyze and report the individual components of the frailty index, identifying which specific features most strongly contribute to differences between groups.

12. Extended Data Fig. 9: The authors performed RNA sequencing on tissues from mice treated with SRSF1 vectors. However, the analysis presented is limited. Key details such as the number of differentially expressed genes per tissue are missing. Moreover, it would be important to assess how transcriptomic changes induced by SRSF1 in these tissues compare to those observed in fibroblasts to evaluate the consistency of the observed effects.

13. Extended Data Fig. 9d: The authors propose a model in which SRSF1 exerts its anti-aging effects by modulating translation and histone modifications, thereby reprogramming the transcriptome of aged cells toward a younger gene expression profile. However, these mechanistic claims are not supported by experimental evidence. For example, the authors could directly assess global histone methylation levels using western blotting or examine other epigenetic modifications, such as DNA methylation changes, induced by SRSF1 treatment. Additionally, validating the anti-aging effects of SRSF1 *in vivo* using established biomarkers, such as DNA methylation clocks, would strengthen their conclusions.

14. The study relies extensively on external datasets, but the specific sources or database accession numbers are not consistently referenced in the figure legends. Instead, only a full list is provided in Supplementary Table 7. For improved

clarity and reproducibility, the authors should cite the exact source or accession ID for each dataset used directly in the corresponding figure legends.

15. The authors have not provided the GEO accession number for the generated transcriptomic data.

Minor points:

1. Extended Data Fig. 1c: The panel reports the correlation coefficient as r , while the text refers to the same value as R^2 . This appears to be a typo, and one of these terms should be corrected for consistency.

2. Extended Data Fig. 2c: To assess predicted age deviations in fibroblasts from HGPS donors, these samples should be compared against fibroblasts from healthy donors within the same dataset, with adjustment for chronological age. Without appropriate controls from the same dataset, it is unclear whether the observed higher predicted age in HGPS cells reflects true biological differences or batch effects causing overestimation of predicted age.

3. Extended Data Fig. 3d: The authors report a negative correlation between baseline gene expression and its logFC in plasmid-treated cells. It is unclear which cells were used to calculate baseline expression — were these the same control cells used to calculate logFC? If so, this negative correlation may be artificially inflated, as the same dataset contributes to both log(TPM) and logFC but with opposite signs. For example, correlating a variable A with $(B - A)$ will by chance produce a negative correlation due to shared sampling in A . To test this hypothesis more objectively, the authors should plot log(TPM) in control cells versus log(TPM) in treated cells. Observation of a plateau in treated cells would support the claim of a transcriptional output limit.

4. Extended Data Fig. 3e-f are not referenced anywhere in the manuscript and should be either discussed or removed.

5. Fig. 2c: It is unclear which cells are used as the reference group for comparing predicted age in cells overexpressing various genes. Are WT cells used as the baseline? The authors should explicitly specify this. Additionally, while the authors apply a p-value threshold of 0.05 to identify significant changes in predicted age across 95 tested plasmids, they must adjust for multiple comparisons (e.g., using FDR correction) to control for false positive artifacts.

6. Fig. 2d: Also unclear which cells serve as the reference for this comparison. Are predicted ages compared to WT cells or BFP cells? The authors should clarify this. Additionally, statistical significance for each boxplot (testing difference from the 0% level) should be indicated with asterisks reflecting p-values adjusted for multiple comparisons, accounting for the number of plasmids used in the screening.

7. Fig. 2e: The abbreviation NTg appears here and in several other figures but is never defined in the manuscript. The authors should provide an explanation of this term.

8. Fig. 3a: The authors plot the difference between predicted and chronological age for each cell line. However, if there is a common batch effect causing systematic under- or overestimation of predicted ages, this difference may be biased. A less biased approach would be to compare predicted ages to those of WT cells from the same patient sequenced in the same batch.

9. Fig. 3c: Since WT samples cluster separately in Fig. 3b, it would be informative to calculate the number of differentially expressed genes in the WT group compared to BFP and include this information in the scatterplot. If many genes are differentially expressed between WT and BFP, the authors should discuss this finding and specify which transcriptomic changes are induced by the screening platform, including any enriched pathways.

10. Fig. 3f: It is surprising that 2 days of SRSF1 induction results in a rejuvenation effect more than twice as strong as that observed after 3 days, while 1 day and 3 days of induction produce similar effects. If the groups are correctly labeled, the authors should discuss possible reasons for this unexpected pattern. Additionally, the method used for normalization of predicted age in this figure should be clearly explained.

11. Extended Data Fig. 5b-f: The chosen heatmap visualization is difficult to interpret. For example, the suggested differences in the effects of SIRT1 and KDM6A overexpression are not clearly visible. Displaying boxplots summarizing enrichment for each pathway across all cell lines would provide a clearer visual representation. This would also enable calculation and display of statistical significance for overall pathway changes in response to each perturbation. P-values should be adjusted for multiple comparisons (the number of pathways) and indicated with asterisks. Furthermore, it is unclear which cells were used as controls for assessing each perturbation's effect (WT, BFP, or both); this information should be clearly stated in the figure legend.

12. Extended Data Fig. 8i: The authors present muscle regeneration data for CAG-driven SRSF1 overexpression, but comparable data for Tet-SRSF1 plasmid are missing. Including this data would be informative, especially since other results in the manuscript suggest Tet-SRSF1 yields better outcomes.

13. Extended Data Fig. 9b: Same recommendations as for Extended Data Fig. 5b-f.

14. Extended Data Fig. 9e: It is unclear from the legend whether this GSEA plot shows pathways enriched for differentially spliced genes or for differentially expressed genes in response to SRSF1 treatment. If the plot represents splicing changes,

it would be valuable to perform the same analysis for differentially expressed genes as well. Furthermore, indicating the directionality of pathway changes would enhance interpretation: rather than showing only significance on the y-axis, the authors could multiply $-\log_{10}(\text{p-adjusted})$ by the sign of the NES, placing upregulated pathways above and downregulated pathways below the axis.

15. Line 485: For the single-sample GSEA, the authors calculated TPM values from counts for each sample and transformed them into z-scores for enrichment analysis. Did the authors perform a log-transformation of TPM values prior to calculating z-scores? Since gene counts and TPM values are typically highly right-skewed, log-transformation is important to approximate a normal distribution before z-scoring. The authors should also specify the source of the reference iPSC dataset used for the “gained identities (2)” analysis.

16. Line 521: The authors mention that predictions were “validated using the Human Ageing Genomic Resources databases,” but do not specify how this validation was performed. Details on the validation approach should be provided.

(Remarks on code availability)

Reviewer #3

(Remarks to the Author)

In this manuscript, the author develops a transcriptome-based screening assay to identify genetic factors that can rejuvenate human dermal fibroblasts. SRSF1 is identified as a key regulator of cellular rejuvenation and aging, with its overexpression shown to extend lifespan in mice. While the study presents some interesting findings, several concerns need to be addressed:

1. A major contribution of this work is the development of an advanced method to assess cellular aging based on transcriptomic profiles. However, the author should compare and discuss how this method performs relative to recently published approaches, particularly the SenCID algorithm (Cell Metabolism, 2024). It would also be informative to evaluate whether SRSF1 would have been identified using SenCID.

2. In addition to transcriptomic data, the study uses SA- β Gal staining, proteasome activity, and mitochondrial membrane potential to assess cellular aging. However, established markers such as cell cycle arrest, SASP factors, and DNA damage responses are notably absent and should be considered to strengthen the phenotypic characterization.

3. The authors should show SA- β Gal staining images. WB or immunostaining images for aging markers is also highly recommended to provide a more comprehensive characterization of cell aging in the main figures. Data quality in Extended Data Fig. 7 is marginally at best.

3. In the abstract, the author claims that SRSF1 can improve wound healing; however, the data presented only include an in vitro wound healing assay measuring cell migration. This is insufficient to support the broader claim of enhanced wound healing. The authors should either tone down this statement in the abstract or provide additional in vivo evidence to substantiate the claim.

4. The author claims to have developed a high-throughput aging assay; however, only 89 genes were screened in this study, which does not qualify as high-throughput by current standards. The authors should clarify this discrepancy and, at a minimum, discuss the potential for applying their transcriptomic aging evaluation method in truly high-throughput screening platforms. For example, it would be valuable to explore whether this method could be integrated with transcriptomic screening technologies such as DRUG-seq or PHD-seq.

Minor comment:

1. It is unusual that SRSF1, the central finding of the study, is not mentioned in the title. The current title, “Transcriptomic reprogramming screen for cell aging,” suggests a focus on screening methodology, while the manuscript primarily centers on the role of SRSF1.

(Remarks on code availability)

Reviewer #4

(Remarks to the Author)

Plesa et al. present a transcriptional reprogramming strategy aimed at cellular rejuvenation, leading to the identification of SRSF1, a splicing factor whose overexpression rejuvenates aspects of the transcriptome and yields multiple beneficial functional outcomes.

The paper identified the role in aging of SRSF1 by the following steps:

1. Development of a Transcriptional Clock: The authors subset the transcriptome into cellular processes whose transcriptional changes correlate with age in normal human dermal fibroblasts (NHDFs) aged 1–96 years, identifying 146 age-associated genes. The validity of this gene set was benchmarked against chemical perturbations and DNA methylation

clocks.

2. Identification of Age-Reversal Candidates: Using the connectivity of their transcriptional clock, the authors derived a second set of genes whose overexpression could theoretically reverse the transcriptional aging signature in fibroblasts.

3. Screening for Effective Rejuvenation Factors: The 89 candidate genes were overexpressed in three fibroblast lines (M55, M65, and M79), leading to the identification of SRSF1 as the most consistent factor in reducing the transcriptional age metric.

4. Validation of Transcriptional Rejuvenation: Additional overexpression lines were generated to validate the rejuvenating effects of SRSF1 and other candidates (e.g., SIRT1, RNASEL).

5. Functional Validation in Cell Culture: The rejuvenating effects of SRSF1 were confirmed using assays for β -galactosidase activity, proteasome function, collagen production, and in vitro wound healing.

6. In Vivo Functional Validation: The effects of SRSF1 were further validated in vivo through wound healing assays in mice and lifespan studies in both mice and *C. elegans*.

Overall, this study represents an interesting computational-experimental approach to identify genes that, when overexpressed, revert transcriptional aging signatures with functional consequences for aging in vivo.

The study is rigorously conducted and is well-written, with no major concerns. The different aspects of the paper are validated by using multiple independent lines of evidence. However, several minor issues should be addressed:

Minor Revisions (Necessary to Address)

- Sex of Mice and Cell Lines: The methods indicate that only female mice were used. If male mice were also tested but did not show rejuvenation effects, those results should be included. Otherwise, explicitly state that no experiments were conducted on males.
- X Chromosome Considerations: Several transcriptional clock and age-reversal genes (e.g., HDAC6) are located on the X chromosome. Given the potential erosion of X-inactivation with age, expression changes may differ between sexes. To clarify this, please:
 - Provide a table that specifies the sex distribution of the fibroblast cell lines or color-code the dots in Figure 8a by sex.
 - Indicate the sex and age of fibroblast lines M55, M65, and M79 in the text.

I apologize if this sounds obsessive-compulsive with regards to sex differences, but if we ever aim to elucidate the age differences between sexes, it would be very valuable if broad studies such as this one provide sexes to be mined.

Points Needing Clarification

- Line 215: The phrase “between our 270 cell lines” is unclear. It appears to refer to 30 cell lines tested under 3×3 conditions, totaling 270 samples. Please revise for clarity.
- Figure 4b: Why was the method of SRSF1 overexpression changed from tetracycline induction to mRNA treatment?
- Figure 4c: Are all genes in this panel overexpressed using the Tet-On system, or only SRSF1? Please clarify.
- Figure 4d & Extended Figures 8b, 8g, 8h: Figure 4d and extended figure 8b,g h: From extended figure 8g, it seems that keeping DOX during wound healing might contribute to beneficial effects unless there is also DOX added to the GFP control. It would be beneficial to indicate precisely when DOX was present in each figure. This information is essential to judge how much the improved wound healing is due to DOX treatment or to exclude that possibility. Figure 4e: Please indicate that the mice used were female only

(Remarks on code availability)

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

In this revised version by Plesa et al, the authors have addressed most of my suggestions and concerns. I thank the authors for this and recognize that the revisions have elevated the quality of the work. However, in my judgement, there are still a few aspects of the work that need to be revisited before I can recommend publication in Nature Aging, which I summarize below. Major concerns:

- In Figure 4, the authors claim a “partially sustained rejuvenation effect” of SRSF1. However, the wash-off period after SRSF1 induction is only two days (after five days induction). I would not consider such a time frame appropriate to claim a “partially sustained” effect. I recommend that either the authors tone down their claim, or conduct a study to directly investigate for how long the reduction in transcriptomic age is maintained.

Minor concerns:

- Figure 1e. Not enough details are provided on how OKSM reprogramming was induced (e.g. method of exogenous expression, for how where the factors expressed), neither of the relationship between the transcriptomic and epigenetic age that we see in the graph and the different phases of reprogramming. When did iPSC-colonies appeared in the culture? Where these cells confirmed iPSCs? This information is important to contextualize the experiment. Eg, is the transcriptomic

tool predicting age zero even before cells turn into bona fide iPSCs?

- Figure 3a. It would also be important to provide details on how OKSM and SRSF1 were overexpressed in this experiment, as it may provide some insights on the greater variability in age reduction induced by OKSM. Was a similar system for induction used? Was the expression of these factors confirmed in all cell lines? Were similar levels of expression reached? These data should be included at least in the extended data collection.
- Extended data Figure 6. In panel “a”, I recommend including close-up/higher magnification images (at least for the Day 5 time point) as with the current versions it is difficult to confirm the morphological differences at all.
- Figure 6a. When does doxycycline administration start in the wound healing model? This label seems to be missing.

(Remarks on code availability)

Reviewer #2

(Remarks to the Author)

In their review, Pleša et al. added important mouse frailty dynamics data, as well as several additional experiments supporting the improvement of selected cellular functional and molecular hallmarks and the lack of cell identity changes induced by SRSF1 overexpression in human fibroblasts. These additions substantially strengthen the manuscript. The proposed gene target appears to be an interesting regulator associated with cellular ageing, and the data are likely to be important for the geroscience field. However, while the study has substantially improved compared to the previous version, several methodological and textual issues should still be addressed prior to publication.

Major points:

1. The authors state in their response that the revised clock development pipeline is now free of data leakage: the dataset was split a priori into training and held-out test sets, and all dimensionality reduction, weak predictor training, feature selection, preprocessing, and model optimization were performed exclusively within the training set. If implemented exactly as described, this largely addresses the original concern about data leakage. However, the revised manuscript itself should make this equally clear. As currently written, the Fig. 1 legend still states that 464 samples were used to train weak age predictors and select features for the final model. In their response, the authors state that this set of 464 samples was divided into a training set (306 samples) and a test set (158 samples). Thus, the Fig. 1 legend appears misleading, as it implies that the entire dataset was used for feature selection, whereas the authors' response suggests the opposite. The Methods section also does not explicitly specify that dimensionality reduction, preprocessing, and feature selection were performed exclusively on the training set. The Methods and figure legend should be updated to state this explicitly.
2. The presentation of clock performance in Fig. 1c and in the main text remains misleading. The main text states that the model “accurately predicted age across more than 300 independent fibroblast samples (Pearson $r = 0.95$; MAE = 5.22)”. However, the authors state in their response that only 158 samples were used as an independent held-out test set, while the Fig. 1c legend indicates that the reported performance was calculated using both training and test samples. Training samples cannot be described as independent validation samples, and combined training + test performance is not an unbiased estimate of model accuracy, since models usually overfit to training data, as also illustrated by the authors in Extended Data Fig. 1c ($r = 0.99$). The main figure and main text should instead report performance on the held-out test set as the primary estimate of model accuracy. The simplest solution would be to replace Fig. 1c with Extended Data Fig. 1c.
3. From the manuscript, I had the impression that clock performance was assessed on completely independent external datasets (for example, that D10 was used only for clock training, whereas D1 was used only for testing) because the manuscript and figure legends refer to “independent” and “external” datasets in the relevant sections. However, based on the code, it appears that samples from the same data sources may have been split between the training and test sets, and that the authors define a “dataset” as a training/test set of samples rather than as a separate data source (D1, D2, etc.). This sample-level splitting strategy is acceptable for held-out validation, but I suggest referring to the groups as the “training set” and “test set” rather than as “independent datasets” to avoid misinterpretation. Besides, the authors should clarify in the Methods section which data sources were used for training, which were used for held-out testing, and whether any source contributed samples to both sets.
4. The revised analysis provides statistical evidence that Old-Srsf1 mice show improved wound repair compared with Old-GFP controls at selected time points. However, because the AAV-GFP control itself shows a trend toward impaired wound healing, even though this reaches significance only in young mice, the authors should clarify whether Old-Srsf1 improves wound repair relative to untreated Old-WT mice, not only relative to Old-GFP controls. This distinction is important for interpreting the result as rescue of ageing-related functional impairment in otherwise untreated aged animals rather than mitigation of AAV-associated burden.
5. The authors claim systemic rejuvenation of mice in response to AAV8-Tet-Srsf1. While they show improvement in selected functional parameters, frailty index dynamics, and lifespan extension, the term “systemic rejuvenation” remains too strong. Rejuvenation would imply that ageing-associated deterioration is reversed (i.e. improved relative to a baseline aged state), rather than merely slowed down. From the frailty-index trajectories, functional impairment appears to accumulate more slowly rather than being reversed. True rejuvenation would be expected to produce a negative slope, rather than a smaller positive slope. Therefore, more precise wording would be “healthspan extension”, “attenuated late-life decline”, or “improved late-life function”, rather than “rejuvenation”.
6. The added mouse RNA-seq analyses improve the manuscript, but the interpretation would be stronger if Srsf1-induced molecular changes were directly compared with normal ageing trajectories in the same mouse tissues. The authors report changes in pathways related to bioenergetics, immune signaling, cellular homeostasis, sterol metabolism, and other processes, and state that these pathways are often dysregulated during ageing. However, without a tissue-matched ageing comparison, it is difficult to determine whether these changes represent anti-ageing, pro-ageing, or ageing-orthogonal

effects. Therefore, it would be informative to perform DEG and pathway enrichment analyses using external mouse ageing gene expression datasets for the same organs, and to show the effects of Srsf1 alongside the corresponding pathway changes observed during normal ageing. This would improve the interpretation of Srsf1 effects in the in vivo mouse model from an ageing perspective.

7. The authors report significant extension of both median and maximum lifespan in mice following Srsf1 overexpression, but they support these claims using only the log-rank test, which examines overall differences between survival curves rather than specifically testing maximum lifespan. If the authors wish to claim a statistically significant effect on maximum lifespan, they should perform a prespecified upper-tail survival test, such as the Wang-Allison test or another appropriate late-life survival test. Otherwise, the claim should be softened.

Minor points:

1. The authors claim age reversal in response to FOXM1 overexpression in NHDFs (Fig. 1f), but this result is based on a single sample ($n = 1$) and is not statistically supported. It should therefore be described as a non-significant trend rather than as confirmed age reversal.

2. Extended Data Fig. 2b: the statistical comparisons appear to compare different fibroblast lines under the same treatment rather than comparing treatment versus matched control conditions within the same cell line. If the goal is to show that metabolic or mitochondrial perturbations increase transcriptomic age, the more direct test would compare each treatment against the corresponding control condition within the same fibroblast line. As currently presented, the significant comparisons are difficult to interpret as evidence of treatment-induced transcriptomic ageing.

3. Extended Data Fig. 2c: the authors assess the effect of mitoNUTs and DEX on epigenetic age in human fibroblasts using the DNAm DunedinPACE clock, while using the Skin & Blood clock in other analyses (e.g., Extended Data Fig. 2d). The rationale for using DunedinPACE in fibroblasts is not clear, given that this model was trained on blood data. The Skin & Blood clock, which the authors use elsewhere, would be a more natural choice for this analysis. The authors should also indicate in the manuscript that the DNAm clock used in Extended Data Fig. 6e is the Skin & Blood clock. This is mentioned in their response to reviewers but appears to be missing from the manuscript itself.

4. Extended Data Fig. 5i: the changes in alternative splicing events induced by SRSF1 overexpression are intriguing, but they are compared with the NTg group rather than with BFP-expressing cells. The authors report that the screening platform itself induces substantial transcriptomic differences between NTg and BFP controls. Therefore, comparing SRSF1-overexpressing cells with NTg cells may confound SRSF1-specific splicing effects with effects of the transgene system burden. A comparison of SRSF1 with BFP (and with OSKM as an alternative rejuvenative treatment) would better isolate SRSF1-specific splicing changes.

5. The manuscript should clearly state the exact age of mice and the time after treatment at which mouse DNAm and RNA-seq profiling were performed. This information seems to be currently lacking.

6. Extended Data Fig. 12c: the mouse liver DNAm analysis is based on a mouse Liver Interventions Clock. It would be informative to also report results for more established chronological liver and pan-tissue mouse DNAm clocks to determine whether the findings are consistent across different clocks. This would improve the reliability and interpretability of the result.

(Remarks on code availability)

Reviewer #3

(Remarks to the Author)

The authors have substantially improved the manuscript by addressing the concerns raised during the previous round of review. I have no concern any more.

(Remarks on code availability)

Reviewer #4

(Remarks to the Author)

A: In the provided manuscript Plesa et al screen for partial reprogramming factors that improve cell function and rejuvenate transcriptional patterns, but don't do so by eroding cell identity markers as classical Yamanaka factors would. They develop an age predictor based on RNA expression and then use the predictor to conduct a cDNA overexpression screen to identify potential regulators that rejuvenate cells. They identify the splicing factor SRSF1 as a central regulator that, when overexpressed, rejuvenates multiple aspects of cellular physiology. The data show that in vivo the induction of SRSF1 in mice accelerates wound healing and extends the lifespan of C57BL6 mice.

B. This is a very interesting and very original study that expands upon the idea of the Yamanaka reprogramming factors but aims to identify factors that, rather than erasing the differentiation pattern of a cell, restore the differentiation pattern to its level seen when the cells are young. While the screen is done in cell culture, they then establish in vivo evidence by expressing SRSF1 using AAV vectors and show that indeed driving transcriptional rejuvenation leads to improved wound healing and an increase in lifespan.

C. The evidence for these claims are as follows. Figure 1 shows the development of a transcriptional aging model to distinguish chronological from biological age based on gene expression data and validate it using human fibroblasts aging 0 to 96 years. Figure 2 then validates their transcriptional age predictor by testing its ability to detect the rejuvenation effect of rapamycin treatment on four different human fibroblasts of different ages.

They then screen a series of transcription and splicing factors for their ability to reduce cellular aging and identify factors that

increase cellular age, such as Notch, and factors that decrease cellular age, such as SRSF1. In Figures 3 to 5 they then evaluate SRSF1 for its ability to reprogram the transcription of aged fibroblasts into that of young fibroblasts and validate that SRSF1 preserves fibroblast identity and does not lead to dedifferentiation or pluripotency. Figure 6 then expands the finding to mouse and *C. elegans* in vivo settings.

D. During the past review, there were several concerns regarding the number of samples, effect sizes, and controls, specifically controls that include doxycycline treatment. The authors have made an extensive effort to address these concerns and have considerably updated the previous manuscript

E-F. The conclusions are well supported by a considerable body of data. The only remaining concern from my part is that the lifespan of the mice is relatively short compared to other studies. I think requesting a repeat of these studies at the current stage would be unfair, but I believe it would be in the interest of both the authors and the readers to point out that the lifespans of the mice are somewhat shorter than what is published in the general literature.

(Remarks on code availability)

Version 2:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

The authors have thoroughly replied to my comments and requests for clarifications. I have no further feedback other than congratulating the authors for their work and recommending the article for publication as is.

(Remarks on code availability)

Reviewer #2

(Remarks to the Author)

In their review, Pleşa et al. addressed most of my concerns, which substantially strengthened the manuscript. Their study and the collected data are likely to be valuable for the field of ageing biology, and I support publication after one minor textual issue is addressed:

1. In their response, the authors agreed that the in vivo effects of SRSF1 overexpression in mice may reflect attenuation of age-associated changes rather than rejuvenation. This distinction is important because the treated mice were compared with age-matched controls, but post-treatment molecular profiles were not compared with pre-treatment baseline measurements. Therefore, the current data do not directly demonstrate that ageing-associated markers were reversed relative to the baseline aged state. However, the term “rejuvenation” is still used when discussing the effects of SRSF1 on the mouse tissue transcriptome (e.g., in lines 474-475, 678, 1392 and 1406). I suggest replacing this term with more cautious wording, such as “ageing-attenuating” or “anti-ageing” effects.

(Remarks on code availability)

Open Access This Peer Review File is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.

The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

Response to reviewers

We thank the editor and all reviewers for their valuable comments and suggestions. In response, we have revised the manuscript by incorporating new analyses and data, clarifying statistical methods and experimental design, and refining the interpretation and presentation of the results. We believe these changes significantly strengthen the manuscript and fully address the reviewers' concerns. Please find below our detailed response to all the reviewer comments.

Reviewer #1

We thank Reviewer #1 for the thoughtful mechanistic comments and constructive feedback. We address each point as follows:

Major

1. In Figure 1 the authors should confirm the aging trajectories achieved by their model with other methods. For example, well-established DNA methylation clocks of other available transcriptomic clocks (<https://www.biorxiv.org/content/10.1101/2024.07.04.601982v1>). The only direct comparison provided in their manuscript (Figure 1f) was performed with an insufficient n number (n=1/time-point).

To address this, we generated new DNA methylation and RNA-seq data in BFP- and *SRSF1*-expressing lines, comparing our transcriptomic age predictions with the well-validated Skin & Blood Horvath clock¹. Aggregate trends across three biological replicates supported age reduction: two lines (M55, M79) showed significant epigenetic age reversal (−8.1 and −7.9 years, respectively), while the third showed a modest non-significant increase (+3 years) within expected biological and technical variance (Extended Data Fig. 6e).

Furthermore, we evaluated intervention effects using the DunedinPACE² epigenetic measure on DNAm datasets corresponding to Extended Data Fig. 2b. We selected DunedinPACE because our primary goal is to quantify *intervention-induced changes in aging dynamics* rather than absolute chronological age. DunedinPACE captures the pace of biological aging and is highly sensitive to short-term physiological perturbations, making it particularly well-suited for proliferative *in vitro* systems where absolute DNAm age is confounded by replication history. Interpreting DunedinPACE comparatively rather than in absolute units, we observed widespread concordance with our transcriptomic age predictor (Extended Data Fig. 2c), including the recapitulation of donor-specific pro-aging and rejuvenation responses (e.g., following dexamethasone and mitoNUTs treatment).

Together, these analyses demonstrate that the transcriptomic aging model captures biologically meaningful aging dynamics that are supported by orthogonal epigenetic measures.

2. In Supplementary Figure 1-f, the age of iPSC and ESC controls is negative (~-100 to -150 years). However, it has been well-established by others that reprogramming to pluripotency reverses biological age to ground zero. Can the authors explain this result? A direct comparison to a more established method (e.g., DNA methylation clocks) could help explain whether this is a problem of the proposed model.

We agree that negative predicted ages were counterintuitive. As the reviewer insightfully suggested, comparing our approach to established DNA methylation clocks highlighted the root of this issue: our initial model fit chronological age using a simple Gaussian distribution. Because chronological age is bounded at zero, but the linear model is not, natural variance in the transcriptomic data of embryonic and reprogrammed cells extrapolated into deeply negative values.

To address this, we revised our model to incorporate the exact mathematical age transformation introduced by the Horvath DNA methylation clock model³. This transformation uses a shifted log-linear model to properly bound early-life and ground-state ages near zero. Applying this established epigenetic clock methodology to our transcriptomic model successfully constrained the output, substantially reducing the magnitude and frequency of negative (most ages >-1 year, Fig. 1d), thereby improving interpretability without affecting the reported conclusions.

3. Sample sizes are missing in many figure legends and should be included.

We have updated all figure legends to include sample sizes and statistical tests used.

4. Figure 2c: exposure to doxycycline can have significant effects in gene expression. Have the authors controlled for this? In addition of BFP controls that would also be treated with dox, comparing the gene expression of WT cell with and without the drug would be informative.

We agree with the reviewer that Dox can significantly affect gene expression. We have now updated Fig. 2c to include labels for untreated samples (NTg), which were also treated with doxycycline (Dox) to ensure direct comparability across conditions and to isolate the effects of transgene induction rather than drug exposure. In these Dox-treated NTg controls, predicted ages deviated from donor chronological age with no consistent direction (-4.5 , $+15.1$, and $+9.6$ years across M55, M65, and M79, respectively). However, when evaluated relative to the empirical null distribution derived from the full perturbation screen, these values corresponded to Z-scores of -1.1 , 0.36 , and 0.63 , respectively, and therefore did not meet criteria for a significant pro-aging or rejuvenating effect.

In addition, we analyzed the M55, M65, and M79 lines cultured without Dox from an independent experiment. In this context, predicted ages were within ± 3 years of donor age. This indicates that while Dox exposure introduces measurable transcriptomic variance and shifts in predicted age, it does not, by itself, generate effects comparable in magnitude to the bona fide age-modulating perturbations identified in our screen. We now clarify this experimental design and interpretation in the figure legend and Methods and explicitly note that all screening comparisons are made within Dox-treated conditions to control for drug-related transcriptional effects (Lines 167-168, Lines 672-675).

5. In Figure 2d the transcriptomic age data is missing for some gene candidates in the M65 sample. Similarly, flow cytometry data for SA-bgal expression, proteasome activity, and membrane potential were not performed in this cell line due to insufficient cell numbers, according to the authors. This should be completed, especially given that the overall response was rather different in different donor cell lines.

We thank the reviewer for noting the missing data. These measurements were not feasible in the initial M65 line due to limited live cell yield. While we understand the desire for complete matched datasets, we elected not to repeat the measurements in newly derived M65 cultures, as inherent batch, passage, and baseline differences would strictly confound comparability to the original screen.

Instead, our downstream validation explicitly accounted for this donor-specific variability. As detailed in the manuscript (Fig. 3, 5), we functionally validated our main hits across an expanded cohort of six independent NHDF lines, which already included a newly derived M65 line. This robust existing validation across diverse donors comprehensively addresses the concern regarding donor variability, providing stronger support for *SRSF1*-mediated rejuvenation than retroactively forcing batch-confounded M65 replicates into the initial screening dataset (Fig. 2).

6. Figure 3a provides the age effect (predicted age – age) after overexpression of the top 5 candidate genes in 6 different donor cell lines. However, the experiment was performed with n=2 sample size, which does not allow for assessing the variability of the intervention within a fixed cell line. While increasing the sample size for all five interventions and six cell lines would be very costly, I recommend that the authors perform a larger-sample study for *SRSF1* and at least one of the cell lines. How many days after gene induction was the data collected? This information should be stated in the figure legend.

Following the reviewer's suggestion, we assessed the technical variability within a single donor line by measuring the transcriptomic age across four independent replicates (n = 4) of the M79 line expressing BFP or *SRSF1*. The results (Extended Data Fig. 5b) demonstrate highly reproducible *SRSF1* age reversal (mean effect of –17.2 years) with remarkably low variability across replicates (SEM = 1.7 years). This confirms that the rejuvenation effect is robust and not an artifact of small sample size within the line. Regarding the timeline, data were collected following a three-day induction, which has now been added to the relevant figure legends (Lines 162, 224, 289, 294, 357, 360, 1186).

7. Similarly, confirming the reduction of biological age shown in Figure 3a via a DNA methylation assay would add to the robustness of this study.

We agree with the reviewer that the DNAm assay would provide an orthogonal validation of the biological age changes. However, because DNAm dynamics are slower than transcriptional changes (Fig 1e), we reasoned that our standard three-day induction might be insufficient to detect robust DNAm changes.

To appropriately test this, we extended the *SRSF1* induction to five days, followed by a two-day washout period, in three distinct donor lines (M55, M68, M79). As detailed in our response to Reviewer 1 (Major Point 1), the resulting DNAm data (Extended Data Fig. 6e) measured via the Skin & Blood Horvath clock¹ showed that *SRSF1* induction drove robust epigenetic age reversal in two of the lines (–8.1 years in M55 and –7.9 years in M79), while the third line (M68) showed a modest shift (+3 years) falling well within expected biological and technical variance. Together, this orthogonal assay confirms that *SRSF1* induces a sustained epigenetic effect that is consistent with our observed transcriptomic rejuvenation.

8. In Figure 3d, can the authors explain why the overexpression of candidate genes induces MSC signatures?. Are those signatures also enriched in the donor cells pre-overexpression, and if not, would that not be considered loss of cell identity in a way?

We agree that the original presentation of Fig. 3d (now Fig. 4a) was insufficiently clear. The analysis shown in this panel was performed using iPSCs as the reference background, rather than fibroblasts as mistakenly stated in the original legend. As a consequence of this baseline, even untreated somatic samples inherently exhibit strong mesenchymal/fibroblast signatures relative to iPSCs, which created the illusion of an induced MSC signature.

To resolve this and improve interpretability, we revised the analysis to focus on well-established marker genes for each cell type. Specifically, we curated marker sets for fibroblasts (*COL1A1*, *COL1A2*, *COL3A1*, *COL5A1*, *DCN*, *LUM*, *FN1*, *PDGFRA*, *SPARC*, *THY1*) and for pluripotency (*SOX2*, *OCT4*, *NANOG*, *LIN28A*, *PRDM14*, *ZFP42*, *SALL4*) and visualized their average normalized expression across conditions (BFP, *SRSF1*, OSKM). Using this corrected framework, OSKM samples show the expected induction of pluripotency markers (e.g., *SOX2*, *NANOG*, *ZFP42*) alongside reduced expression of fibroblast markers (e.g., *PDGFRA*, *THY1*, *DCN*), indicating partial reprogramming and loss of somatic identity. In contrast, *SRSF1*-expressing cells do not exhibit these changes and closely resemble BFP controls. These results

definitively show that unlike OSKM, *SRSF1* upregulation achieves rejuvenation without eroding key cell identity markers or inducing a loss of somatic identity.

9. The data in Figure 3e does not add much information. The cells treated with OKSM look already different on day 1. Day 0 images should be included as well as morphometric quantification of the cell body area. Scale bars are missing.

Following the reviewer's comment, we repeated the experiment and initiated the measurements at day -1. The new representative images (Extended Data Fig. 6a), which now include the required scale bars, confirm that OSKM induction drives a progressive reduction in cell body area, whereas *SRSF1* does not induce sustained morphological changes.

To quantitatively support this, we performed the requested morphometric analysis of cell body area over a five-day induction (Extended Data Fig. 6b, c). This analysis revealed a significant, sustained reduction in cell body area for OSKM versus BFP starting at Day 3 ($P < 0.001$, Dunnett's test). In contrast, *SRSF1*-expressing cells, across both M55 and M79 lines, remained statistically indistinguishable from BFP by the end of the time course ($P = 0.25$ at Day 5).

10. Figure 3f shows that the induction of *SRSF1* for 2-3 days does not achieve long-lasting reversal of transcriptomic age. In fact, most of the effect is lost within 12 days. The authors propose that further dosage and timing optimization needs to be tested. However, the authors do not test further dosages and therefore there is no proof that they would not induce loss of cell identity. This could mean that there is no advantage of *SRSF1* against OKSM (many studies have shown that 2-day expression of OKSM does not induce de-differentiation) and therefore needs to be tested.

We agree that short-term *SRSF1* induction (2–3 days) does not result in durable transcriptomic age reversal, with much of the effect dissipating after withdrawal (Extended Data Fig. 6d). Accordingly, we do not claim that brief induction is sufficient for long-lasting rejuvenation.

To address whether longer induction compromises cell identity, we evaluated a five-day on/two-day off *SRSF1* induction paradigm in the same fibroblast lines. Under these conditions, *SRSF1*-expressing cells remained viable (Extended Data Fig. 6a-c), preserved fibroblast identity (Fig. 4d-f), and exhibited sustained reductions in transcriptomic and epigenetic aging measures (Fig. 4c, Extended Data Fig. 6e). In contrast, matched five-day OSKM induction resulted in rapid loss of fibroblast identity (Fig. 4d), acquisition of pluripotency markers (Fig. 4e), and poor viability in standard culture conditions (Extended Data Fig. 6a-b), consistent with initiation of dedifferentiation, precluding durable post-induction measurements. Thus, within an induction window where OSKM already destabilizes cell identity, *SRSF1* can be induced for longer durations without lineage disruption.

While further dose and timing optimization remains an important future direction, these data demonstrate that *SRSF1*-mediated rejuvenation is achievable under conditions that preserve cellular identity and viability. Moreover, modulation of a single factor offers practical translational advantages (i.e. feasibility of activation via AAV or oligonucleotides) over multi-factor reprogramming approaches. Together, these findings support *SRSF1* as a mechanistically and practically distinct rejuvenation strategy from OSKM.

11. How long do the functional changes (e.g., decrease in s-b-gal activity and others) last after induction? This assessment is important and should be included, as ultimately, it is cellular function that matters the most.

We agree that assessing the persistence of functional rejuvenation markers (e.g., SA- β -gal activity, proteasome activity, mitochondrial membrane potential) would provide additional insight into the durability of *SRSF1* effects. However, robust post-induction assessment of

these endpoints *in vitro* is technically challenging and difficult to interpret in our system. Following transient induction, cell proliferation rates are altered, making delayed measurements highly sensitive to passage- and density-dependent confounders.

Importantly, we do not claim that transient *SRSF1* induction produces durable functional rejuvenation under the *in vitro* conditions tested. Rather, our conclusion is that short-term *SRSF1* induction is sufficient to immediately elicit transcriptomic and functional rejuvenation processes. We agree that defining conditions that yield sustained functional benefits is an important next step. Our current efforts therefore prioritize *in vivo* models, where long-term regulation and physiological relevance can be assessed without limitations of *in vitro* cultures. We have clarified these points and moderated claims accordingly in the revised manuscript (Lines 306-315, 330-331, 520-525).

12. The *in vivo* experiments (Figure 4), in general, are not comprehensive enough. For example, in Figure 4d, *SRSF1* is administered intradermally but there is no information on the cell types that express *SRSF1* *in vivo* upon viral transduction. It is also not clear for how long *SRSF1* is expressed for. Characterization of the cells post-*SRSF1* expression (e.g., transcriptomic age and age-related phenotype) is also missing. The absence of these data makes it very difficult to interpret the results provided on wound healing in a mechanistic manner.

We agree with the reviewer that a more comprehensive mechanistic characterization of *SRSF1* expression and the specific responding cell types *in vivo* is an exciting future direction. However, these *in vivo* experiments were purposefully designed to establish translational proof-of-concept and physiological relevance, rather than to provide a deep molecular dissection of the highly complex wound-healing environment.

In our intradermal delivery paradigm, the duration and cellular distribution of transgene expression are constrained by vector design, tissue turnover, and the lack of longitudinal sampling, which together limit the feasibility of transcriptomic or single-cell profiling post-delivery. Nonetheless, the observed macroscopic improvement in wound healing is entirely consistent with the functional rejuvenation induced by *SRSF1* in fibroblasts *in vitro*.

To address the reviewer's specific question regarding the duration of expression, we have now explicitly clarified the induction timeline in Fig. 6a (formerly Fig. 4) and its corresponding legend. We have also revised the manuscript to make it unambiguously clear that the *in vivo* experiments are intended as a physiological proof-of-relevance rather than a mechanistic analysis, and we have carefully tempered our mechanistic interpretations accordingly. We believe this accurately reflects the scope and intent of the current study while highlighting the need for detailed *in vivo* tracking in future follow-up work (Lines 384-385, 441-442, 518-519).

13. Similarly, for the studies with systemic administration of *SRSF1*, it is not described how long the expression of the gene lasts after viral transduction. Some kinetics at the gene or protein expression levels are necessary. In fact, *in vivo* levels of *SRSF1* are not assessed at all. Instead, the authors tested the expression of rtTA as a proxy, can they explain why *SRSF1* levels are not investigated?

We agree that direct measurement of *SRSF1* expression kinetics *in vivo* would strengthen mechanistic interpretation. However, longitudinal quantification of *SRSF1* mRNA or protein was not feasible without repeating the systemic delivery experiments, which require more than one year to obtain survival outcomes. To address this point with available material, we analyzed bulk RNA-seq data from terminal tissues and detected robust rtTA expression predominantly in the liver (Extended Data Fig. 12a), consistent with the hepatic tropism of AAV8 at the administered dose (1×10^{11} vg).

Exogenous *SRSF1* transcripts were detected in only two terminal samples (liver and heart from a single animal), indicating that endpoint RNA-seq is not suitable for inferring peak or early transgene expression in this system. Because *SRSF1* is an essential endogenous gene with high baseline expression, low levels of virally derived *SRSF1* transcripts at late time points are likely masked by this high natural background. In contrast, the bacterial rtTA transcript has no mammalian background, making it a significantly more sensitive marker for confirming the long-term presence of the AAV vector. Accordingly, rtTA expression was used solely as an indicator of vector biodistribution rather than a proxy for sustained *SRSF1* expression. We have now clarified this point in the revised manuscript.

Prior studies using AAV8 vectors at comparable doses report relatively high early transgene expression with attenuation over weeks to months before reaching a lower steady state^{4,5}. Such kinetics likely explain why *SRSF1* transcripts are not consistently detectable at terminal time points despite potentially meaningful earlier expression following vector delivery. We have now added these references to provide appropriate context. Finally, we note that the conclusions of our study do not rely on sustained systemic overexpression of *SRSF1*, and that the observed physiological benefits are compatible with a model in which transient or tissue-restricted expression is sufficient to initiate downstream effects. We now explicitly acknowledge this limitation in the Discussion. (Lines 406-410, 518-522, 1288-1921).

14. There is also no mechanistic explanation behind the increased lifespan and reduced frailty of the mice. Since the highest expression (of rtTA) is localized in the liver, have the authors tested for liver function biomarkers? Are there paracrine effects expected from SRSF1 in other organs?

We thank the reviewer for this insightful comment. As noted above, the AAV8–Tet system used in this study resulted in the highest vector expression in the liver, raising the possibility that hepatic effects could contribute to the observed improvements in lifespan and frailty.

The primary objective of the *in vivo* experiments was to determine whether systemic induction of *Srsf1* can modulate late-life physiological decline and survival, rather than to dissect organ-specific mechanisms. Accordingly, our analyses focused on organism-level phenotypes, including longitudinal frailty measurements and survival (Fig. 6c, d), together with molecular readouts such as transcriptomic and DNA methylation profiling (Extended Data Fig. 12c–e). These analyses revealed enrichment of metabolic pathways and reduced inflammatory signaling, consistent with systemic physiological improvements.

Endpoint RNA-seq analysis indicated that transgene expression was largely restricted to the liver (Extended Data Fig. 12a), suggesting that extrahepatic effects are likely to arise indirectly. However, future work is needed to dissect whether the systemic benefits arise from improved liver homeostasis, altered secretion of liver-derived factors, or downstream modulation of immune and metabolic pathways. We therefore avoid mechanistic claims in the revised manuscript and note that further studies can bridge the gap between cellular state changes and systemic physiology (Lines 427-436, 522-525).

Reviewer #2

We thank Reviewer #2 for identifying critical methodological flaws. We fully agree with the comments and have comprehensively corrected them:

Major

1. During the clock development, PCA, “weak” predictor training, and pathway/gene selection were all performed on the entire dataset, including the folds later used for cross-validation, so information from the validation folds leaked into model construction. Given the multi-layer pipeline and extensive feature selection, this design almost certainly inflates performance. Besides, for the dataset used to train the clock, the authors haven’t reported the quality on a test set. The original dataset must be partitioned into separate training and test sets (or assessed via nested leave-one-out approach), with every dimensional-reduction and selection step fitted exclusively on the training split. Model accuracy should then be reported on unseen data, not the training set (Extended Data Fig. 1C).

We thank the reviewer for identifying this important issue. In response, we have fully redesigned the training and evaluation strategy for the transcriptomic aging predictor to eliminate information leakage and provide a rigorous assessment of generalization. Specifically, we assembled an expanded dataset of 464 fibroblast transcriptomes spanning ages 1–94 years and partitioned it *a priori* into a training set ($n = 306$; $\sim 2/3$) and a held-out validation set ($n = 158$; $\sim 1/3$). All dimensionality reduction, pathway selection, weak predictor training, and feature selection steps were performed exclusively within the training set, with model hyperparameters optimized by 10-fold cross-validation restricted to these samples. The validation set was not used at any stage of model construction and served solely for performance evaluation. Owing to the increased training cohort, the final model now includes 292 informative genes across the most predictive biological processes. Performance metrics are reported separately for the training and validation sets in Extended Data Fig. 1c, demonstrating robust accuracy on unseen data. These changes fully address the concern regarding data leakage and inflated performance estimates.

2. The application of the trained clock to external datasets is also affected by data leakage. According to the Methods, the model is retrained separately for each new dataset, which limits its generalizability and undermines its utility. Ideally, a single, fixed model should be developed and shown to perform robustly across external datasets without retraining. Moreover, before retraining, the authors pool training and external data for TMM normalization and batch correction. Since these procedures use information from the entire dataset, including the test samples, they introduce data leakage by allowing the model to indirectly access the structure of the test data. This likely inflates the reported performance of the clock on external datasets ($r = 0.93$). To rigorously assess predictive performance, normalization and batch correction parameters must be derived solely from the training data and then applied unchanged to external samples. Performance should also be evaluated using a single, pre-trained model.

We agree that improper normalization and batch effects can inflate apparent performance and have been underappreciated sources of variability among transcriptomic aging clocks. While our original approach sought to mitigate batch effects by jointly normalizing training and test datasets, we acknowledge that this strategy could introduce information leakage.

To address this issue, we have implemented a revised workflow that enables application of a single, fixed model to new datasets without retraining or joint normalization. In the revised pipeline, all samples are first independently TPM-normalized and filtered to the gene set

defined by the training data. Batch effects are then adjusted using frozen surrogate variable analysis (fSVA), with surrogate variables estimated exclusively from the training set and applied unchanged to external samples. Importantly, external datasets are no longer pooled with training data at any stage, and no model parameters are re-estimated. The pre-trained predictor is applied directly to transformed external samples, thereby preventing information leakage and enabling a rigorous assessment of generalization. All reported external performance metrics in the revised manuscript reflect this leakage-free evaluation strategy.

3. In Fig. 2b the list of age-associated pathways surprisingly omits several established ageing hallmarks, including inflammation and mitochondrial function (e.g., see 10.1093/bioinformatics/btp073). The authors should explain this omission or demonstrate, using a leakage-free analysis, that these signatures do in fact emerge. It would also be informative to compare the pathways identified in fibroblasts with ageing signatures reported at the tissue level (e.g., blood, brain, skeletal muscle) to evaluate how well the fibroblast clock captures organism-wide ageing trends.

The primary goal of this study was to provide a proof-of-concept for using transcriptomic reprogramming screens to identify age-reversal interventions in an accessible, highly controlled *in vitro* system. We agree with the reviewer that cultured fibroblasts do not capture the full complexity of organismal aging. For instance, the absence of canonical inflammatory signatures in our clock is largely expected, as isolated fibroblast monocultures lack the immune cell populations that drive "inflammaging" signatures in whole-tissue bulk RNA-seq.

Furthermore, while our updated, leakage-free pipeline (described in Reviewer 2 Major Point 1) slightly refined our predictive feature set, it remains fundamentally tailored to cell-intrinsic fibroblast biology rather than organism-wide trends. Motivated by the ubiquitous expression of SRSF1 across cell types, we transitioned to *in vivo* models to investigate its broader systemic effects. We have revised the manuscript to explicitly discuss the biological differences between *in vitro* fibroblast clocks and tissue-level aging signatures, ensuring readers do not interpret our cellular model as a direct surrogate for organismal aging (Lines 522-525).

4. Many figures lack statistical significance estimates (p-values): Fig. 1d-e; Fig. 2d-g; Fig. 3a; Extended Data Fig. 1e-f; Extended Data Fig. 2a-c, 2e-k; Extended Data Fig. 4; Extended Data Fig. 6b, f; and Extended Data Fig. 8f-h, k. The authors should provide appropriate p-values and apply multiple testing correction where relevant (e.g., in Fig. 2d-g, Fig. 3a). In several instances, effects are described in the text but are not clearly supported by the figures. These claims should be more precisely stated and backed by statistical analysis. For example, the authors assert that predicted age increases in fibroblasts exposed to sun-induced UV damage, but Extended Data Fig. 2b shows no consistent pattern: there is no difference in young donors, an increase in middle-aged, and a decrease in older individuals. Such discrepancies should be clearly addressed and statistically validated.

We have added appropriate statistical tests and corresponding *P* values to all figures involving quantitative comparisons. Where multiple hypotheses are tested (e.g., perturbation screens and multi-condition comparisons), we now apply appropriate multiple-testing corrections, as specified in the figure legends. In addition, we have revised the text to ensure that all reported effects are directly supported by the corresponding statistical analysis shown in the figures.

Regarding the reviewer's specific observation about the UV damage cohort (previously Extended Data Fig. 2b), we agree with your assessment. Upon applying our updated, more rigorous statistical framework, the UV damage data did not demonstrate a consistent, statistically significant pattern across all age groups. Consequently, to maintain evidence rigor, we have removed this specific analysis and its associated text from the revised manuscript.

5. Fig. 2c: The authors use their transcriptomic clock to evaluate the impact of gene overexpression across a perturbation screen. However, some of the perturbed genes may have been included in the clock itself. In such cases, the resulting predicted age changes are likely to be trivial – for instance, overexpressing a gene with a negative coefficient in the model will obviously reduce predicted age. To avoid circular reasoning, the authors should exclude each tested gene from the clock when evaluating its perturbation effect.

We thank the reviewer for this insightful comment regarding potential circularity. We agree that overexpression of genes included in the aging predictor could, in principle, artifactually influence the predicted age.

To directly assess this possibility, we performed a censoring analysis in which each tested gene that appears in the predictor (*CTNNB1*, *GSK3B*, *KAT7*, *KDM6A*, *VCP*) was individually removed from the input prior to age prediction. Because our updated transcriptomic model relies on a distributed network of 292 informative genes, the removal of a single feature generally exerts minimal influence on the aggregate prediction. Indeed, for four of the five genes, censoring had negligible impact on the predicted ages (*CTNNB1*: +6.7 years; *GSK3B*: -1.0 years; *KDM6A*: -0.6 years; *VCP*: -7.1 years).

Only *KAT7* showed a substantial dependence, with uncensored samples predicted to be ~24 years older than censored counterparts, indicating partial circularity in that specific case. We now explicitly note this limitation in the Results (Lines 147-148). These analyses confirm that the primary conclusions of Fig. 2c are not driven by self-referential contributions of predictor genes.

6. Fig. 3d: The authors report cell identity enrichment scores for several perturbations and control lines compared to a reference fibroblast group. However, it is strange that WT (NTg) fibroblasts display significant pro-fibroblast enrichment, as they should be nearly identical to the reference fibroblasts and serve as negative controls. This finding suggests that the external reference fibroblast dataset may not be an appropriate comparator, potentially due to batch effects. To reduce this confounding, the authors should consider using NTg or BFP cells as the internal reference group for calculating enrichment scores.

We agree with the reviewer that the original presentation of Fig. 3d (now Fig. 4a) could be misleading. In the initial analysis, the enrichment scores shown were calculated using iPSCs as background rather than fibroblasts, which explains why NTg fibroblasts display strong fibroblast enrichment.

To address this issue, we have revised both the analysis and figure to rely on direct expression of established cell-type marker genes rather than enrichment against an external reference dataset. Specifically, we curated marker sets for fibroblasts (*COL1A1*, *COL1A2*, *COL3A1*, *COL5A1*, *DCN*, *LUM*, *FNI*, *PDGFRA*, *SPARC*, *THY1*) and for pluripotency (*SOX2*, *OCT4*, *NANOG*, *LIN28A*, *PRDM14*, *ZFP42*, *SALL4*) and visualized their average normalized expression across conditions (BFP, *SRSF1*, OSKM).

Using this corrected framework, OSKM samples show the expected induction of pluripotency markers (e.g., *SOX2*, *NANOG*, *ZFP42*) alongside reduced expression of fibroblast markers (e.g., *PDGFRA*, *THY1*, *DCN*), indicating partial reprogramming and loss of somatic identity. In contrast, *SRSF1*-expressing cells do not exhibit these changes and closely resemble BFP controls. These revisions remove batch-related confounding and directly support the conclusion that *SRSF1* induction preserves fibroblast identity.

7. Line 241: The authors claim that *SRSF1* overexpression does not lead to loss of cellular identity, in contrast to OSKM reprogramming. To support this claim, they should present expression levels of canonical pluripotency markers, such as Nanog, Oct4, and Sox2. across

WT, BFP, SRSF1, and OSKM1 cell lines. This would provide a clearer comparison of reprogramming versus rejuvenation.

We have updated the cell identity analysis (now Fig. 4b) to explicitly show the expression of canonical pluripotency markers, including *SOX2*, *POU5F1* (*OCT4*), and *NANOG*. As expected, OSKM induction leads to clear upregulation of pluripotency genes and loss of fibroblast marker expression. In contrast, *SRSF1* overexpression does not induce pluripotency markers and closely mirrors BFP controls. Together with the revised marker-based analysis described above, these data support the claim that *SRSF1*-based rejuvenation does not involve dedifferentiation.

8. Fig. 4a: The authors report that SRSF1 activation reduces proteasome activity by ~40%. However, this appears inconsistent with Extended Data Fig. 6b, where no clear difference is observed and no p-value is provided. The authors should clarify whether these figures represent different experimental assays or if this discrepancy is due to an error. If distinct assays were used, details of their respective methodologies and sensitivities should be included.

We thank the reviewer for noting this apparent discrepancy. Fig. 4a (now Fig 5b) and Extended Data Fig. 8b reflect results obtained using distinct experimental systems. In Fig. 4a/5a, proteasome activity was measured in a stable inducible system with tightly controlled *SRSF1* expression. In contrast, Extended Data Fig. 8b shows data from transient mRNA-mediated *SRSF1* delivery, which is inherently more variable due to differences in transfection efficiency, mRNA stability, and donor-specific sensitivity to lipid-based transfection reagents. Primary fibroblasts are particularly sensitive to lipofection-induced stress, which can transiently perturb proteasome activity and increase variance. We have now clarified these methodological differences in the results and figure legend and added statistical analyses and *P* values to Extended Data Fig. 8b, 9 (Lines 346-351, 1233-1237, 1244-1253). These clarifications explain the observed variability and resolve the apparent inconsistency between the two panels.

9. Line 257: The authors report that “a transient 3-day SRSF1 mRNA treatment induced ... a 30% decrease in ROS levels upon H₂O₂ exposure.” However, according to Fig. 4b, this effect does not reach statistical significance. The text should be revised to reflect this.

We thank the reviewer for pointing out this discrepancy. Upon re-examining the source data, we identified an error in the originally reported effect size and *P* value in the text. The data shown in Fig. 4b (now Extended Data Fig. 9b) demonstrate a statistically significant mean reduction in ROS levels of approximately 20% (*P* = 0.028), rather than the previously stated nonsignificant 30% decrease. We have corrected the text and figure legend accordingly to accurately report the observed effect size and significance (Lines 370-371, 1246-1248).

10. Fig. 4d: The authors conclude that SRSF1 induction significantly improves wound healing in old mice, based on the observed wound closure dynamics between SRSF1- and GFP-treated groups. However, Extended Data Fig. 8f suggests that the Tet-GFP plasmid alone impairs wound healing in both young and old mice when compared to untreated controls. This raises the concern that the apparent benefit of SRSF1 may be due to the deleterious effects of the GFP plasmid rather than a true therapeutic effect. To address this, the authors should compare wound healing in SRSF1-treated old mice to that in untreated old WT mice. Furthermore, while the authors report no significant benefit from constitutive (CAG) SRSF1 expression, it seems that CAG-GFP animals show better healing than Tet-GFP animals, whereas SRSF1 treatment shows similar effects in both contexts. This suggests that the lack of observed benefit from CAG-SRSF1 may result from improved baseline healing in CAG-GFP controls, rather than from reduced SRSF1 efficacy. The authors should carefully reanalyze their wound healing data using appropriate control groups and clearly distinguish between beneficial effects of SRSF1

and detrimental effects of specific plasmid constructs. Additionally, if AAV delivery itself impairs physiological outcomes, as appears likely from the GFP control wound healing data, its potential impact on other *in vivo* results, such as the reduced lifespan in GFP-treated mice compared to SRSF1 group, should also be critically evaluated.

We thank the reviewer for this critical observation and agree that vector-associated toxicity must be carefully accounted for. The reviewer correctly notes that AAV-GFP administration delayed wound closure relative to untreated WT animals. While this delay trended visually across both age groups, post-hoc statistical analysis confirmed it only reached significance in the young cohort (e.g., Day 5, $P < 0.0001$; Extended Data Fig. 11c). This observation is consistent with prior reports that local AAV delivery can transiently impair tissue repair, likely through inflammatory or stress-related mechanisms⁶, and highlights the critical necessity of using matched viral controls to accurately evaluate baseline healing.

Because the old mice are already at a physiological baseline of impaired healing, the AAV-GFP vector did not significantly exacerbate their deficit compared to Old-WT controls ($P > 0.05$ at all time points). Importantly, all AAV-treated groups in our study were rigorously matched for viral dose, promoter usage, and injection procedure, ensuring that any vector-related stress was equivalent across conditions. Within this controlled framework, *Srsf1* expression significantly accelerated healing compared to Old-GFP controls (Days 5 and 11, $P < 0.01$, Fig. 6b) and restored wound closure kinetics to levels statistically indistinguishable from Young-GFP controls throughout the entire trajectory. Because the GFP plasmid did not artificially depress the baseline healing of the aged cohort, we conclude that the reproducible improvements observed with *Srsf1* treatment represent a true functional rejuvenation, rather than reflecting a nonspecific effect of the GFP control.

Regarding the comparison between CAG- and Tet-driven constructs, we completely agree with the reviewer that differences in baseline wound healing kinetics preclude direct cross-promoter comparisons. These experiments are not directly comparable, as the Tet-driven studies utilized a ten-fold higher viral dose. Furthermore, we observed that high levels of exogenous *Srsf1* trigger a negative feedback loop *in vivo* that normalizes *Srsf1* levels, consistent with known autoregulatory feedback mechanisms that constrain *Srsf1* levels *in vivo*⁷.

Accordingly, we have revised the manuscript to temper mechanistic claims, framing the wound-healing experiments as evidence of biological activity and functional rescue rather than a super-physiological enhancement of wound healing. We believe this interpretation more accurately reflects the scope and limitations of the data and aligns with the reviewer's concern (Lines 384-386, 394-399).

11. Extended Data Fig. 8l: The authors report an improvement in the frailty index of SRSF1-treated mice but present the data as the age at which each animal reaches a frailty index threshold of 0.18. This threshold appears arbitrary, and the rationale for its selection is not explained. Additionally, the frequency of frailty index measurements for each mouse is not specified and should be clearly detailed in the Methods section. Typically, frailty index data are presented either as pairwise comparisons between treatment groups at specific timepoints (e.g., before and 3 months after treatment) or as longitudinal trajectories over time. If multiple measurements per animal were collected, the authors should plot the frailty index dynamics for GFP and SRSF1 groups on the same graph and provide statistical analyses (p-values) for group comparisons at each timepoint. This approach would better illustrate the rate of frailty progression in treated versus control mice. Furthermore, it would be valuable to analyze and report the individual components of the frailty index, identifying which specific features most strongly contribute to differences between groups.

We thank the reviewer for this constructive critique of the frailty index analysis. The FI threshold of 0.18 was selected based on prior validation in mice, where an FI < 0.18 denotes non-frail and FI $\geq$ 0.18 denotes frail⁸. In our cohort, this value also corresponds to the upper 95% confidence interval of the mean FI score. We now explicitly state this rationale, alongside the biweekly frequency of FI measurements, in the Methods (Lines 846-850).

In response to the reviewer's suggestion, we now present longitudinal frailty trajectories for GFP and *Srsf1*-treated groups on the same plot, with accompanying statistical analyses (Fig. 6d). These data show that *Srsf1* induction significantly slows the rate of frailty accumulation compared to controls. To quantify this effect while accounting for mortality-related dropout, we applied a shared-parameter Bayesian joint model, which revealed a 44% reduction in the rate of frailty progression in *Srsf1*-treated mice ($P = 0.046$). We also analyzed all 29 individual frailty index components longitudinally (Supplementary Table 6). While nominal differences were observed in select parameters (reduced accumulation of deficits in breathing rate and dermatitis), no single component remained significant after multiple-testing correction. Importantly, this indicates that the observed benefit reflects a coordinated, systemic attenuation of frailty rather than an isolated improvement in a single physiological domain.

12. Extended Data Fig. 9: The authors performed RNA sequencing on tissues from mice treated with SRSF1 vectors. However, the analysis presented is limited. Key details such as the number of differentially expressed genes per tissue are missing. Moreover, it would be important to assess how transcriptomic changes induced by SRSF1 in these tissues compare to those observed in fibroblasts to evaluate the consistency of the observed effects.

We thank the reviewer for pointing out this limitation. In the revised manuscript, we now explicitly report the number of differentially expressed genes (DEGs) per tissue and discuss both the associated biological processes and how these changes compare to those observed in human fibroblasts (Lines 428-436). Consistent with our observation that the AAV8 vector primarily transduces the liver, this tissue exhibited the highest number of DEGs (879). In contrast, the brain, spleen and heart displayed markedly fewer changes (~200 DEGs each). Focusing on the liver, we found that *Srsf1* induction broadly modulated biological processes that are dysregulated during aging (Extended Data Fig. 12d). Comparison with the transcriptional changes observed in human fibroblasts revealed a shared downregulation of stress response, inflammation, and cell cycle pathways, alongside cell type- and tissue-specific effects (Extended Data Fig. 12e).

13. Extended Data Fig. 9d: The authors propose a model in which SRSF1 exerts its anti-aging effects by modulating translation and histone modifications, thereby reprogramming the transcriptome of aged cells toward a younger gene expression profile. However, these mechanistic claims are not supported by experimental evidence. For example, the authors could directly assess global histone methylation levels using western blotting or examine other epigenetic modifications, such as DNA methylation changes, induced by SRSF1 treatment. Additionally, validating the anti-aging effects of SRSF1 *in vivo* using established biomarkers, such as DNA methylation clocks, would strengthen their conclusions.

Following the reviewer's suggestions, we have added DNA methylation analyses in both human fibroblasts and mouse liver, which provide orthogonal validation of *SRSF1*'s rejuvenation effect. Specifically, *SRSF1* induction significantly reduces DNA methylation age in fibroblasts and decreases epigenetic age acceleration in mouse liver, supporting the conclusion that *SRSF1* promotes epigenetic rejuvenation *in vitro* and *in vivo* (Extended Data Fig. 6e and Extended Data Fig. 12c).

Regarding histone modifications, our proposed model is based on the observation that the histone methylation transcriptomic predictor module is changing towards a younger state. Moreover, several key chromatin regulators are both differentially expressed and alternatively spliced following *SRSF1* induction, suggests a potential link between *SRSF1* RNA-processing activity and pathways governing histone methylation. While we completely agree that direct biochemical measurements of global histone methylation would be highly informative, performing a comprehensive spatial and temporal characterization of these specific epigenetic marks *in vitro* and *in vivo* extends substantially beyond the scope of the present study. To avoid overinterpretation, we have clarified in the revised manuscript that we provide a mechanistic hypothesis based on transcriptomic and splicing evidence (Lines 254-261).

14. The study relies extensively on external datasets, but the specific sources or database accession numbers are not consistently referenced in the figure legends. Instead, only a full list is provided in Supplementary Table 7. For improved clarity and reproducibility, the authors should cite the exact source or accession ID for each dataset used directly in the corresponding figure legends.

We agree with the reviewer that a direct connection of each figure panel with the corresponding external datasets improves clarity and reproducibility throughout the manuscript. Since several panels have been created using multiple datasets, we kept Supplementary Table 7 (now Supplementary Table 2) and introduced a dataset identifier (D1 – D11). These identifiers have been included for all figure panel legends that are based on external datasets (Lines 96, 100, 102, 105, 1128, 1130, 1138, 1146, 1151).

15. The authors have not provided the GEO accession number for the generated transcriptomic data.

We agree with the reviewer and are fully committed to data transparency. Due to the large scale of the transcriptomic data generated in this study (>500 individual samples), we are currently finalizing the extensive metadata curation. We are actively processing this submission and will forward the resulting GEO accession number and a secure reviewer access token directly to the editorial office the moment they are generated. The formal accession numbers will also be included in the final Data Availability Statement prior to publication.

Minor

1. Extended Data Fig. 1c: The panel reports the correlation coefficient as r , while the text refers to the same value as R^2 . This appears to be a typo, and one of these terms should be corrected for consistency.

We appreciate the reviewer for pointing out this typo. The value initially reported was R^2 , but we changed it to r for both the figure and the legend for consistency (Extended Data Fig. 1c).

2. Extended Data Fig. 2c: To assess predicted age deviations in fibroblasts from HGPS donors, these samples should be compared against fibroblasts from healthy donors within the same dataset, with adjustment for chronological age. Without appropriate controls from the same dataset, it is unclear whether the observed higher predicted age in HGPS cells reflects true biological differences or batch effects causing overestimation of predicted age.

After revising the manuscript and implementing the updated analysis pipeline, we repeated the analysis using the improved framework. In this setting, the difference was no longer statistically significant. To avoid overinterpretation, we have removed this analysis from the revised manuscript.

3. Extended Data Fig. 3d: The authors report a negative correlation between baseline gene expression and its logFC in plasmid-treated cells. It is unclear which cells were used to calculate baseline expression — were these the same control cells used to calculate logFC? If so, this negative correlation may be artificially inflated, as the same dataset contributes to both log(TPM) and logFC but with opposite signs. For example, correlating a variable A with (B – A) will by chance produce a negative correlation due to shared sampling in A. To test this hypothesis more objectively, the authors should plot log(TPM) in control cells versus log(TPM) in treated cells. Observation of a plateau in treated cells would support the claim of a transcriptional output limit.

We thank the reviewer for their rigorous assessment of the mathematical coupling. The log₂ fold changes were calculated with respect to the BFP control, and we initially used all the cells in our dataset to calculate the baseline endogenous expression. However, we acknowledge that outliers could still skew this analysis, and, to objectively test for a transcriptional output limit, we decoupled our analysis in two ways:

- a. Independent baseline: We recalculated the relationship using No-Transgene (NTg) samples as the baseline (Extended Data Fig. 3d). Since these NTg samples are biologically and statistically independent from the BFP controls used for fold change calculations, the persistent negative correlation ($r = -0.66$, $P = 5.78 \times 10^{-13}$) confirms the effect is biological rather than a mathematical artifact.
- b. Direct visualization and housekeeping benchmarks: As suggested, we plotted Treated (Plasmid) vs Control (NTg) TPMs (Extended Data Fig. 3e). This visualization reveals a transcriptional plateau. Crucially, the induced genes converge at approximately 12.5–12.8 log₂ TPM, levels that mirror the endogenous expression of essential housekeeping genes like *GAPDH* and *ACTB*. This supports a model where the cell's transcriptional machinery or resource pool imposes a physical cap on total mRNA abundance for any single locus.

4. Extended Data Fig. 3e-f are not referenced anywhere in the manuscript and should be either discussed or removed.

We removed the Extended Data Fig. 3e-f. They served as gating examples for flow cytometry.

5. Fig. 2c: It is unclear which cells are used as the reference group for comparing predicted age in cells overexpressing various genes. Are WT cells used as the baseline? The authors should explicitly specify this. Additionally, while the authors apply a p-value threshold of 0.05 to identify significant changes in predicted age across 95 tested plasmids, they must adjust for multiple comparisons (e.g., using FDR correction) to control for false positive artifacts.

We clarify that the Z-scores in Fig. 2c were calculated relative to the distribution of all tested genes within each donor line (Line 165). This approach was used as a primary discovery screen to identify candidate genes that deviated significantly from the library mean. To address the concern regarding multiple comparisons, we have now applied the FDR correction to the discovery screen results (Fig. 2d, Lines 169-171). Furthermore, the primary hits identified (e.g., *SRSF1*, *WDTCI*) were subsequently subjected to high-stringency validation with multiple independent donor lines, where age-reversal effects were quantified and verified using FDR-adjusted *P*-values (Fig. 3a, Lines 224-226).

6. Fig. 2d: Also unclear which cells serve as the reference for this comparison. Are predicted ages compared to WT cells or BFP cells? The authors should clarify this. Additionally, statistical significance for each boxplot (testing difference from the 0% level) should be indicated with asterisks reflecting p-values adjusted for multiple comparisons, accounting for the number of plasmids used in the screening.

In accordance with the reviewer's suggestion, we have updated Fig. 2d to explicitly indicate statistical significance using asterisks directly on the plot. We have also updated the figure legend to clearly specify the reference comparison group (BFP cells) and to confirm that the reported *P* values have been adjusted for multiple hypotheses (FDR correction, Lines 169-171).

7. Fig. 2e: The abbreviation NTg appears here and in several other figures but is never defined in the manuscript. The authors should provide an explanation of this term.

We have now explicitly defined No Transgene (NTg) in the figure legend for Fig. 2 (Line 167).

8. Fig. 3a: The authors plot the difference between predicted and chronological age for each cell line. However, if there is a common batch effect causing systematic under- or overestimation of predicted ages, this difference may be biased. A less biased approach would be to compare predicted ages to those of WT cells from the same patient sequenced in the same batch.

We completely agree with the reviewer that shared batch effects could systematically bias these differences. To mitigate this, we have updated our calculation to compare predicted ages directly against a batch-specific, line-specific NTg baseline (Fig 3a, Line 224).

9. Fig. 3c: Since WT samples cluster separately in Fig. 3b, it would be informative to calculate the number of differentially expressed genes in the WT group compared to BFP and include this information in the scatterplot. If many genes are differentially expressed between WT and BFP, the authors should discuss this finding and specify which transcriptomic changes are induced by the screening platform, including any enriched pathways.

We thank the reviewer for this suggestion. We have now explicitly highlighted the No-Transgene (NTg) controls in the scatterplot (Extended Data Fig. 5d). The screening platform itself induced substantial transcriptomic changes (>8,000 DEGs between NTg and BFP controls), which underscores the critical necessity of using BFP-transfected cells as our baseline for all comparisons. Over-representation analysis of these platform-induced DEGs (Extended Data Fig. 5e) revealed an upregulation in fibroblast proliferation and hypoxia responses, alongside a downregulation in GTPase signaling and immune responses. We have added a brief discussion of these baseline findings to the Results section (Lines 239-243).

10. Fig. 3f: It is surprising that 2 days of *SRSF1* induction results in a rejuvenation effect more than twice as strong as that observed after 3 days, while 1 day and 3 days of induction produce similar effects. If the groups are correctly labeled, the authors should discuss possible reasons for this unexpected pattern. Additionally, the method used for normalization of predicted age in this figure should be clearly explained.

We confirmed that the groups are correctly labeled. Upon further quality control analysis of the transcriptomic data, we observed moderate variability in RNA-seq library quality across specific induction time points and biological replicates, which likely accounts for the non-linear rejuvenation pattern observed between 1, 2, and 3 days of induction. In the interest of transparency, we retained this exploratory time course (now Extended Data Fig 6d) but have de-emphasized changes in rejuvenation based on different induction times, concluding instead that short-term *SRSF1* induction generally initiates a temporary rejuvenation program of similar magnitude (Lines 306-308). To further strengthen this finding with higher-quality samples and a more robust protocol, we repeated the experiment using a five-day on/two-day off regimen, which demonstrates consistent and significant age reversal across multiple models (Fig. 4c). Finally, predicted ages in this figure are normalized to BFP and Day 0 baseline, which we have now explicitly detailed in the corresponding figure legend (Lines 1211-1212).

11. Extended Data Fig. 5b-f: The chosen heatmap visualization is difficult to interpret. For example, the suggested differences in the effects of SIRT1 and KDM6A overexpression are not clearly visible. Displaying boxplots summarizing enrichment for each pathway across all cell lines would provide a clearer visual representation. This would also enable calculation and display of statistical significance for overall pathway changes in response to each perturbation. P-values should be adjusted for multiple comparisons (the number of pathways) and indicated with asterisks. Furthermore, it is unclear which cells were used as controls for assessing each perturbation's effect (WT, BFP, or both); this information should be clearly stated in the figure legend.

We agree with the reviewer that the heatmap visualization was difficult to interpret. As suggested, we have replaced heatmaps with boxplots to clearly illustrate process activity variability across cell lines (Fig 3d, Extended Data Fig. 5g, h). Furthermore, we have updated the figure panels and legends to explicitly state the control condition (BFP) and to display statistical significance, utilizing FDR-adjusted *P* values to correct for multiple comparisons across the tested pathways (Lines 231-233, 237-238, 1196-1199).

12. Extended Data Fig. 8i: The authors present muscle regeneration data for CAG-driven SRSF1 overexpression, but comparable data for Tet-SRSF1 plasmid are missing. Including this data would be informative, especially since other results in the manuscript suggest Tet-SRSF1 yields better outcomes.

We agree with the reviewer that a direct comparison of muscle regeneration outcomes between CAG-driven and Tet-inducible SRSF1 expression would be informative. In the current study, the CAG-SRSF1 experiment was performed in parallel with the initial Tet-SRSF1 studies, prior to establishing that Tet-inducible delivery consistently yielded more favorable and tunable outcomes across *in vivo* assays. Unfortunately, establishing a new, dedicated longitudinal animal cohort to perform these specialized muscle injury and regeneration assays falls outside the feasible scope and timeframe of the current revision process.

However, Extended Data Fig. 8i (now Extended Data Fig 11g) is used to show that strong, ubiquitous *SRSF1* overexpression is not overtly detrimental to muscle regenerative capacity in this assay. We have therefore limited our interpretation of these data to a safety and tolerability context, as well as contrasting it with our inducible system to highlight the importance of expression dynamics. We have revised the text to clarify this point and to avoid any implication of efficacy for the constitutive system (Lines 494-496).

13. Extended Data Fig. 9b: Same recommendations as for Extended Data Fig. 5b-f.

As requested, we have similarly updated these panels to display boxplots with FDR-adjusted statistical significance and explicit control baselines, precisely mirroring the improvements described in our response to Minor Point 11.

14. Extended Data Fig. 9e: It is unclear from the legend whether this GSEA plot shows pathways enriched for differentially spliced genes or for differentially expressed genes in response to SRSF1 treatment. If the plot represents splicing changes, it would be valuable to perform the same analysis for differentially expressed genes as well. Furthermore, indicating the directionality of pathway changes would enhance interpretation: rather than showing only significance on the y-axis, the authors could multiply $-\log_{10}(p\text{-adjusted})$ by the sign of the NES, placing upregulated pathways above and downregulated pathways below the axis.

We thank the reviewer for these concrete suggestions to improve the interpretability of this figure. We have replaced the previous visualization with a directional dot-plot (Extended Data

Fig. 12d) that displays selected enriched biological processes and clearly indicates whether they are up- or down-regulated based on normalized enrichment scores (NES). In addition, dot size now represents $-\log_{10}(P_{adj})$. Finally, we clarified in the figure legend that this analysis was based on $\log_2$ fold changes from differential expression analysis (Line 1295-1296).

15. Line 485: For the single-sample GSEA, the authors calculated TPM values from counts for each sample and transformed them into z-scores for enrichment analysis. Did the authors perform a log-transformation of TPM values prior to calculating z-scores? Since gene counts and TPM values are typically highly right-skewed, log-transformation is important to approximate a normal distribution before z-scoring. The authors should also specify the source of the reference iPSC dataset used for the “gained identities (2)” analysis.

We appreciate the reviewer pointing out these important methodological considerations regarding the ssGSEA transformations. To provide a more direct, biologically intuitive, and robust assessment of cell state that avoids these normalization caveats, we have replaced the ssGSEA-based identity analysis with a direct comparison of established, raw marker gene expression for pluripotency and fibroblast identity (Fig. 4a, b, d, e).

16. Line 521: The authors mention that predictions were “validated using the Human Ageing Genomic Resources databases,” but do not specify how this validation was performed. Details on the validation approach should be provided.

We have expanded the Methods section to provide precise details on this validation approach, alongside the reference to the original study that generated the candidate genes (Lines 636-637). Specifically, we employed a hypergeometric test to determine whether the overlap between our predicted candidates and these established aging-related genes was significantly greater than would be expected by chance, benchmarking against the top-ranked DEGs.

Reviewer #3

We thank Reviewer #3 for the thoughtful and constructive comments, which helped improve the rigor, clarity, and positioning of the study. We have addressed each point below and revised the manuscript accordingly.

Major

1. A major contribution of this work is the development of an advanced method to assess cellular aging based on transcriptomic profiles. However, the author should compare and discuss how this method performs relative to recently published approaches, particularly the SenCID algorithm (Cell Metabolism, 2024). It would also be informative to evaluate whether *SRSF1* would have been identified using SenCID.

We thank the reviewer for suggesting the SenCID algorithm. To evaluate whether this methodology could identify *SRSF1*, we ran the pipeline on all samples from our initial screen and quantified the difference in SID-scores compared to our BFP controls (Fig. R1).

The results differ significantly from those obtained by our transcriptomic predictor. SenCID correctly identified *CDKN2A* as pro-aging in M79 cells and TP53(DN) (dominant negative p53) as anti-aging in M65 cells. However, OSKM had scores opposite to its experimentally validated biological effects, and the algorithm did not identify *SRSF1* as a rejuvenating hit.

SenCID appears highly optimized for identifying specific senescence signatures, whereas our clock is designed to capture holistic rejuvenation patterns. Because a formal performance evaluation of SenCID in our system is limited by the small number of known rejuvenating/pro-aging genes, we have provided this analysis exclusively here for the reviewer's reference.

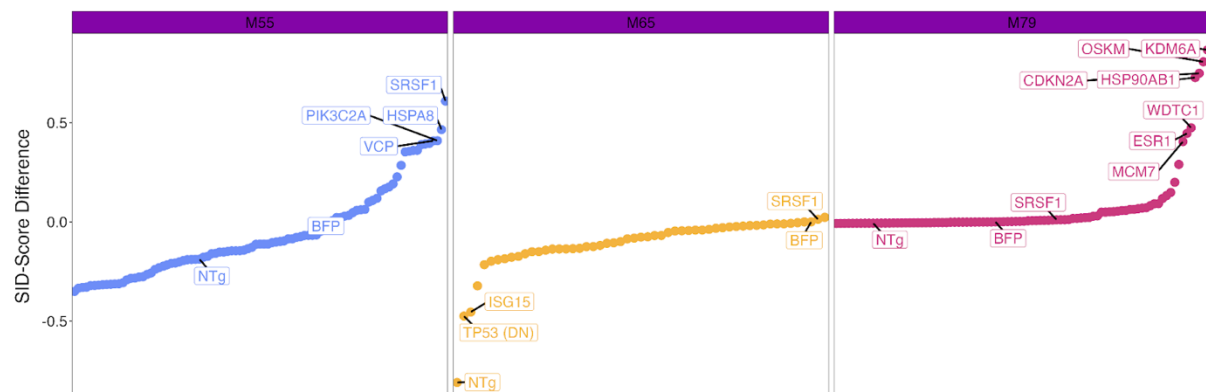


Figure R1. SID-score differences between candidate gene overexpression and BFP controls. Positive scores indicate a higher SID-score obtained for the candidate gene while negative scores correspond to higher SID-scores in BFP samples. Genes with an absolute SID-score difference of at least 0.4 are labeled, alongside with no transgene (NTg), BFP controls, and *SRSF1* across all samples.

2. In addition to transcriptomic data, the study uses SA- β Gal staining, proteasome activity, and mitochondrial membrane potential to assess cellular aging. However, established markers such as cell cycle arrest, SASP factors, and DNA damage responses are notably absent and should be considered to strengthen the phenotypic characterization.

We thank the reviewer for this suggestion and agree that the inclusion of established, orthogonal aging markers significantly strengthens our phenotypic characterization. In the revised manuscript, we have now explicitly included readouts for DNA damage response, SASP-associated signaling, and cell cycle arrest. Specifically, we quantified NF- κ B nuclear localization to assess SASP-associated inflammatory activation (Fig. 5c) and performed γ -

H2AX immunostaining to assess DNA damage (Fig. 5d). We observed a consistent reduction in these age-associated phenotypes following *SRSF1* overexpression, including a statistically significant decrease in NF- κ B activation.

To robustly contextualize these imaging-based results, we also performed gene set enrichment analyses of pathways related to SASP signaling and DNA repair (Fig. 5g, h). These analyses revealed significant downregulation of SASP-associated gene sets and upregulation of DNA repair pathways in the *SRSF1* group, directly corroborating the phenotypic trends. Regarding cell cycle arrest, we assessed p16 immunostaining and observed a very low fraction of p16⁺ cells (<4%) across all conditions, with no significant difference between *SRSF1* and control groups (Extended Data Fig. 8c). This demonstrates that the cells analyzed were largely non-arrested and actively cycling, consistent with their proliferative capacity observed *in vitro*. We have updated the Results section to comprehensively incorporate these new assays (Lines 337-340).

3. The authors should show SA- β Gal staining images. WB or immunostaining images for aging markers is also highly recommended to provide a more comprehensive characterization of cell aging in the main figures. Data quality in Extended Data Fig. 7 is marginally at best.

We appreciate the reviewer's emphasis on visual validation and agree that providing immunostaining and SA- β Gal images offers a more comprehensive characterization of the aging phenotypes. To directly address this, we have substantially expanded and upgraded our imaging data. First, to strengthen our initial flow-based readouts, we performed a histochemical SA- β Gal assay, which showed a robust (>50%) reduction in senescence-associated staining in *SRSF1*-overexpressing cells. Representative images are now included in Extended Data Fig. 7a, alongside the quantitative analysis in Extended Data Fig. 8c.

Second, we added representative immunostaining for canonical aging markers (γ -H2AX, NF- κ B, and p16) to provide a more comprehensive visual characterization supporting the main figures (Extended Data Fig. 7a, b). Finally, we agree with the reviewer that the original images for ROS stress and collagen deposition were visually suboptimal. We have replaced these with much clearer, high-fidelity examples derived from the same experiments (Extended Data Fig. 7c, d). Together, these additions provide direct visual confirmation of our quantitative readouts and substantially improve the transparency of our phenotypic characterization.

4. In the abstract, the author claims that *SRSF1* can improve wound healing; however, the data presented only include an *in vitro* wound healing assay measuring cell migration. This is insufficient to support the broader claim of enhanced wound healing. The authors should either tone down this statement in the abstract or provide additional *in vivo* evidence to substantiate the claim.

We agree with the reviewer that claims regarding wound healing must be rigorously supported and explicitly defined. We respectfully clarify that *in vivo* wound-closure data were indeed included in the original submission (previously Fig. 4c, now Fig. 6b), alongside extensive methodological details, controls, and dosing information in Extended Data Fig. 11a–e.

Nevertheless, we completely agree with the reviewer that our abstract should not overgeneralize these findings. To ensure our claims are strictly proportional to the data, we have revised the abstract and main text to explicitly distinguish between the enhanced cell migratory capacity observed *in vitro* and the functional rescue of wound-closure dynamics *in vivo*. The *in vivo* experiments are now strictly framed as a proof of physiological relevance rather than a broad claim of super-physiological enhancement (Lines 384-385, 397-399). These revisions ensure consistency across the manuscript and align the stated conclusions with the scope and limitations of the data presented.

5. The author claims to have developed a high-throughput aging assay; however, only 89 genes were screened in this study, which does not qualify as high-throughput by current standards. The authors should clarify this discrepancy and, at a minimum, discuss the potential for applying their transcriptomic aging evaluation method in truly high-throughput screening platforms. For example, it would be valuable to explore whether this method could be integrated with transcriptomic screening technologies such as DRUG-seq or PHD-seq.

We appreciate this clarification and agree that the terminology requires refinement. In the revised manuscript, we have softened our claims regarding throughput. While our targeted screen successfully evaluated 89 candidate genes and 6 controls across over 480 samples, establishing a robust functional framework, we agree this represents an initial implementation rather than a fully scaled high-throughput screen.

The core contribution of this work is the development of a transcriptomic aging evaluation framework that is inherently compatible with highly scalable perturbation platforms. As recommended by the reviewer, we have updated our Discussion to explicitly highlight how this framework can be integrated with true high-throughput transcriptomic screening technologies (Lines 509-514). Specifically, we now note that screening breadth can be exponentially expanded using highly multiplexed transcriptomic approaches (such as DRUG-seq or PHDs-seq) alongside single-cell resolution technologies like Perturb-seq. We thank the reviewer for highlighting these specific platforms, which will clearly define the next generation of large-scale reprogramming screens.

Minor

1. It is unusual that SRSF1, the central finding of the study, is not mentioned in the title. The current title, “Transcriptomic reprogramming screen for cell aging,” suggests a focus on screening methodology, while the manuscript primarily centers on the role of SRSF1.

We agree with the reviewer that highlighting our central biological finding significantly improves the clarity and impact of the manuscript. Accordingly, we have revised the title to: “Transcriptomic reprogramming screen identifies *SRSF1* as a central regulator of aging and lifespan”.

Reviewer #4

We thank Reviewer #4 for these careful and constructive clarifications. We have addressed each point below and revised the manuscript accordingly to improve transparency and precision.

Major

1. Sex of Mice and Cell Lines: The methods indicate that only female mice were used. If male mice were also tested but did not show rejuvenation effects, those results should be included. Otherwise, explicitly state that no experiments were conducted on males.

We confirm that all *in vivo* experiments in this study were performed exclusively in female mice; no experiments were conducted in male mice. Conversely, all human fibroblast lines used *in vitro* were male-derived, as indicated by their nomenclature (e.g., M55 = male, 55 years). No female fibroblast lines were included in this study. These details are now explicitly stated in the Methods and indicated in Fig 6a for all *in vivo* mouse experiments (Lines 439, 772).

2. X Chromosome Considerations: Several transcriptional clock and age-reversal genes (e.g., HDAC6) are located on the X chromosome. Given the potential erosion of X-inactivation with age, expression changes may differ between sexes. To clarify this, please:

- *Provide a table that specifies the sex distribution of the fibroblaste cell lines or color-code the dots in Figure 8a by sex.

- *Indicate the sex and age of fibroblast lines M55, M65, and M79 in the text.

Because all primary fibroblast lines used were male-derived, sex-specific effects related to X-chromosome inactivation or erosion are not a confounding variable in this study. We now explicitly indicate the sex and donor age of lines M55, M65, and M79 in the main text (Line 131).

3. Line 215: The phrase “between our 270 cell lines” is unclear. It appears to refer to 30 cell lines tested under 3×3 conditions, totaling 270 samples. Please revise for clarity.

We agree that the original phrasing was imprecise and appreciate the opportunity to clarify. The original statement referred to the distinct transgenic lines generated by introducing our library of genetic perturbations into three primary fibroblast donor backgrounds (M55, M65, and M79). In total, this effort yielded approximately 240 distinct transgenic and control cell lines. These lines were subsequently evaluated in duplicate, generating a final screening cohort of 480 transcriptomic samples. We have revised the text and figure legend to explicitly describe that the UMAP analysis was applied to this complete 480-sample screening cohort (Lines 216, 227-228, 1187).

4. Figure 4b: Why was the method of SRSF1 overexpression changed from tetracycline induction to mRNA treatment?

We deliberately utilized both Tet-inducible transgenic lines and mRNA transfection to demonstrate that *SRSF1*-mediated rejuvenation is robust across distinct delivery modalities. Furthermore, the mRNA approach represents a transient, translationally relevant method that aligns with our ongoing transition to LNP-based delivery *in vivo*. However, to streamline the narrative and minimize confusion regarding delivery methods in the main text, we have relocated the RNA-based validation data to Extended Data Figs. 8 and 9. We have also clarified this rationale in the Results section (Lines 346-347).

5. Figure 4c: Are all genes in this panel overexpressed using the Tet-On system, or only *SRSF1*? Please clarify.

Yes, all genes shown in Fig. 4c (now Fig. 6b) were overexpressed exclusively using the Tet-On inducible system. Relocating the RNA-based perturbation data to the Extended Data Figs. 8 and 9 (as noted above) resolves this ambiguity in the main figures.

6. Figure 4d & Extended Figures 8b, 8g, 8h: Figure 4d and extended figure 8b,g h: From extended figure 8g, it seems that keeping DOX during wound healing might contribute to beneficial effects unless there is also DOX added to the GFP control. It would be beneficial to indicate precisely when DOX was present in each figure. This information is essential to judge how much the improved wound healing is due to DOX treatment or to exclude that possibility. Figure 4e: Please indicate that the mice used were female only

We appreciate the opportunity to clarify this critical control. We confirm that Dox was administered to all induced groups, including the GFP controls, in all Tet-On in vivo studies. Therefore, the improved wound healing is specifically driven by *SRSF1* induction, and we can definitively exclude Dox exposure as a confounding variable.

Consistent with this properly controlled design, Extended Data Fig. 8g (now Extended Data Fig. 11e) demonstrates that sustained *SRSF1* induction (continuous Dox) throughout the wound-healing period confers a significantly greater benefit than transient pre-induction (Dox withdrawal at the time of injury). We have revised the figure legends and schematics to explicitly detail the exact timing and presence of Dox across all experimental and control cohorts (Lines 442, 1276-1279). Finally, as requested, we have added a statement confirming the exclusive use of female mice in the legend for Figure 4e (now Fig. 6a, Line 439).

References

1. Horvath, S. *et al.* Epigenetic clock for skin and blood cells applied to Hutchinson Gilford Progeria Syndrome and ex vivo studies. *Aging* <https://doi.org/10.18632/aging.101508> (2018) doi:10.18632/aging.101508.
2. Belsky, D. W. *et al.* DunedinPACE, a DNA methylation biomarker of the pace of aging. *Elife* **11**, (2022).
3. Horvath, S. DNA methylation age of human tissues and cell types. *Genome Biol.* <https://doi.org/10.1186/gb-2013-14-10-r115> (2013) doi:10.1186/gb-2013-14-10-r115.
4. Muhuri, M., Levy, D. I., Schulz, M., McCarty, D. & Gao, G. Durability of transgene expression after rAAV gene therapy. *Mol. Ther.* **30**, 1364–1380 (2022).
5. Greig, J. A. *et al.* Integrated vector genomes may contribute to long-term expression in primate liver after AAV administration. *Nat. Biotechnol.* **42**, 1232–1242 (2024).
6. Gernoux, G. *et al.* AAV8 locoregional delivery induces long-term expression of an immunogenic transgene in macaques despite persisting local inflammation. *Mol. Ther. Methods Clin. Dev.* **20**, 660–674 (2021).
7. Gonçalves, V. & Jordan, P. Posttranscriptional regulation of splicing factor SRSF1 and its role in cancer cell biology. *BioMed Research International* Preprint at <https://doi.org/10.1155/2015/287048> (2015).
8. Kane, A. E. *et al.* A Comparison of Two Mouse Frailty Assessment Tools. *J. Gerontol. A Biol. Sci. Med. Sci.* **72**, 904–909 (2017).

Response to reviewers

We thank the editor and all reviewers for their continued engagement and constructive feedback. We are glad the previous round of revisions was viewed as substantially strengthening the work, and the remaining concerns focus largely on methodological clarification, presentation, and refinement of specific claims. We have addressed each by incorporating additional analyses, new figure panels, statistical tests, and refined wording. Please find our point-by-point response below.

Reviewer #1

We thank Reviewer #1 for thorough and constructive comments. We address each point as follows:

Major

1. In Figure 4, the authors claim a “partially sustained rejuvenation effect” of SRSF1. However, the wash-off period after SRSF1 induction is only two days (after five days induction). I would not consider such a time frame appropriate to claim a “partially sustained” effect. I recommend that either the authors tone down their claim, or conduct a study to directly investigate for how long the reduction in transcriptomic age is maintained.

We agree with the reviewer that "partially sustained" overreaches given the two-day post-withdrawal observation window. We have removed this characterization throughout the manuscript and replaced it with precise statements of what was measured. The main text (Lines 332, 340-341 and 357-358), the Fig. 4c legend (Lines 319-321), and the Extended Data Fig 6 title and legend (Lines 1361, 1371-1372) now describe the effect as a transcriptomic age reduction observed two days after Dox withdrawal, without implying durability beyond that endpoint.

Minor

1. Figure 1e. Not enough details are provided on how OKSM reprogramming was induced (e.g. method of exogenous expression, for how long where the factors expressed), neither of the relationship between the transcriptomic and epigenetic age that we see in the graph and the different phases of reprogramming. When did iPSC-colonies appeared in the culture? Where these cells confirmed iPSCs? This information is important to contextualize the experiment. Eg, is the transcriptomic tool predicting age zero even before cells turn into bona fide iPSCs?

We thank the reviewer for these points. The Fig. 1e trajectories were generated from publicly deposited RNA-seq and DNAm data (Ohnuki *et al.*, ref. 66; SRA PRJNA243926; Dataset D9). The Methods (Lines 665-679) and Fig. 1e legend (Lines 105-110) now document the source protocol: retroviral delivery of OSKM via pMXs vectors into HDFs expressing mouse *Slc7a1*, with day 0 = day of transduction. Because pMXs is an integrating retroviral system, OSKM factors were expressed constitutively from day 0 throughout the time course, with progressive endogenous silencing during maturation/stabilization rather than a controlled withdrawal.

Sample identity at each timepoint is explicitly stated: d3 is an EGFP+ post-transduction population, d7–d28 are TRA-1-60(+) intermediate reprogrammed cells, and d35–d42 are established iPSC clones at feeder-plate passages 1–2. iPSC colony morphology was first reported in the source protocol at day 25; only the d35–d42 clones were subjected to functional pluripotency validation. The d3–d28 samples are designated as "intermediate reprogrammed cells" in the source publication. We have also added shaded background regions to Fig. 1e

delineating the three sample categories (parental HDFs, intermediate reprogrammed cells, confirmed iPSCs), allowing direct visual comparison with the validated iPSC reference.

Regarding whether the predictor reaches age zero before cells become bona fide iPSCs, our trajectory shows that this is the case for the transcriptomic age, which approaches iPSC-comparable values (d35–d42) by d7, within the intermediate phase and well before functional pluripotency was confirmed. The Skin & Blood DNAm clock follows a slower trajectory and converges on iPSC values by d20. This early transcriptomic/epigenetic divergence is consistent with prior reports that transcriptional rejuvenation precedes stable acquisition of pluripotency (Olova *et al.*, *Aging Cell* 2019, ref. 16; Roux *et al.*, *Cell Syst.* 2022, ref. 17) and motivates the screening strategy developed below, designed to identify factors that induce transcriptomic rejuvenation without committing cells to dedifferentiation.

2. Figure 3a. It would also be important to provide details on how OKSM and SRSF1 were overexpressed in this experiment, as it may provide some insights on the greater variability in age reduction induced by OKSM. Was a similar system for induction used? Was the expression of these factors confirmed in all cell lines? Were similar levels of expression reached? These data should be included at least in the extended data collection.

We thank the reviewer for these points. All four constructs in Fig. 3a (SRSF1, WDTC1, OSKM cassette, BFP control) were delivered using the same piggyBac Dox-inducible system used for the primary screen, in the six donor fibroblast lines (M55, M65, M67, M68, M69, M79), under identical induction conditions. We have added an explicit statement to the Results (Lines 226-230) confirming uniform delivery across constructs, and the construct architecture and screening platform are described in the Methods (Lines 733-744).

To directly address whether OSKM's greater between-line variability in age reduction reflects uneven transgene expression, we quantified per-line transgene expression for all four constructs from construct-specific 20-bp barcode reads in the screen RNA-seq data (Methods, Lines 604-608). The results are shown as a new panel in Extended Data Fig. 5b. Transgene expression was uniform across donor lines for all four constructs, with cross-line coefficients of variation of 0.23 (OSKM), 0.28 (BFP), 0.36 (SRSF1), and 0.38 (WDTC1). OSKM exhibited expression levels similar to SRSF1 and the lowest between-line variability of the four constructs, indicating that the greater between-line variability of OSKM-induced age reduction likely reflects differential biological response across donor backgrounds

3. Extended data Figure 6. In panel “a”, I recommend including close-up/higher magnification images (at least for the Day 5 time point) as with the current versions it is difficult to confirm the morphological differences at all.

We agree with the reviewer that the original wide-field images did not clearly resolve individual cell morphology. We have updated Extended Data Fig. 6a to include 10× magnified close-ups alongside the existing 4× wide-field views for days 1, 3, and 5 across all three conditions (BFP, SRSF1, OSKM), with boxed regions in the wide-field images marking the magnified areas. The figure legend has been updated accordingly (Lines 1364-1368).

4. Figure 6a. When does doxycycline administration start in the wound healing model? This label seems to be missing.

We thank the reviewer for catching this omission. We have added a "Dox treatment start" label to the wound healing timeline in Fig. 6a, indicating that dox administration begins concurrent with skin punch biopsy/wounding. The legend (Lines 468-470) has been updated accordingly. Alternative dox-timing conditions for the wound healing model are described in Extended Data Fig. 11e.

Reviewer #2

We thank Reviewer #2 for the continued rigorous evaluation of the manuscript. We have addressed each remaining concern through additional analyses, refined statistical reporting, and clearer presentation.

Major

1. The authors state in their response that the revised clock development pipeline is now free of data leakage: the dataset was split a priori into training and held-out test sets, and all dimensionality reduction, weak predictor training, feature selection, preprocessing, and model optimization were performed exclusively within the training set. If implemented exactly as described, this largely addresses the original concern about data leakage. However, the revised manuscript itself should make this equally clear. As currently written, the Fig. 1 legend still states that 464 samples were used to train weak age predictors and select features for the final model. In their response, the authors state that this set of 464 samples was divided into a training set (306 samples) and a test set (158 samples). Thus, the Fig. 1 legend appears misleading, as it implies that the entire dataset was used for feature selection, whereas the authors' response suggests the opposite. The Methods section also does not explicitly specify that dimensionality reduction, preprocessing, and feature selection were performed exclusively on the training set. The Methods and figure legend should be updated to state this explicitly.

We thank the reviewer for this careful reading of the manuscript. We have revised both the figure legend and the Methods section to make the two-stage pipeline and its data boundaries fully explicit, as detailed below.

The revised Methods section now describes the pipeline in two dedicated sub-sections. The first covers the process selection (Lines 609-639), which was carried out following a three-step procedure. First, a dataset comprising wild-type (WT) fibroblasts (Dataset D10), HGPS fibroblasts (Dataset D10), in-house iPSCs, and rapamycin-treated fibroblasts (Dataset D12) was compiled, normalized, and batch-corrected. Samples were normalized jointly using batch-aware variance stabilizing transformation, with batch effects regressed out while preserving biological variation between WT, HGPS, rapamycin-treated fibroblasts, and iPSCs. Second, WT fibroblasts from dataset D10 were extracted from the corrected data and subset to the genes composing each candidate cellular process gene set. Principal component analysis was applied to each gene subset, and a weak age predictor (generalized linear model, GLM) was trained per process. Finally, the resulting predictions were passed as features to a GLM with lasso regularization to identify the processes that are most informative of fibroblast aging. This step is entirely separate from the final model training.

The second sub-section covers final model training (Lines 640-659). A dataset of 464 fibroblast samples was compiled in which each sample was TPM normalized and log1p transformed. Basic quality control was performed by removing genes with an average expression of less than 0.5 TPM. Subsequently, the data was split into a training set (n=306) and a held-out test set. Surrogate variable estimation and final model training were performed exclusively on the training set (n=306). Importantly, after splitting the dataset, the held-out test set contained samples from Dataset D10 and D12, which were used for process selection. Although this step estimates no parameters that feed into the final model, we conservatively excluded these samples from the held-out test set in the revised version of this manuscript to eliminate any potential ambiguity. The held-out test set therefore comprises 105 samples in the revised manuscript. Notably, model performance was essentially unchanged ($r = 0.85$ without D10 versus $r = 0.82$ with D10), confirming that the original conclusions are unaffected by this adjustment.

We have revised the figure legend (Lines 91-102) accordingly to explicitly state how the weak age predictors were trained and to which parts of the data the normalization, feature selection and dimensionality reduction steps for the final model estimation apply.

2. The presentation of clock performance in Fig. 1c and in the main text remains misleading. The main text states that the model “accurately predicted age across more than 300 independent fibroblast samples (Pearson $r = 0.95$; MAE = 5.22).” However, the authors state in their response that only 158 samples were used as an independent held-out test set, while the Fig. 1c legend indicates that the reported performance was calculated using both training and test samples. Training samples cannot be described as independent validation samples, and combined training + test performance is not an unbiased estimate of model accuracy, since models usually overfit to training data, as also illustrated by the authors in Extended Data Fig. 1c ($r = 0.99$). The main figure and main text should instead report performance on the held-out test set as the primary estimate of model accuracy. The simplest solution would be to replace Fig. 1c with Extended Data Fig. 1c.

We agree with the reviewer that it could be misleading to show the combined performance in the main figure. We therefore followed the suggestion and replaced Fig. 1c with Extended Data Fig. 1c.

3. From the manuscript, I had the impression that clock performance was assessed on completely independent external datasets (for example, that D10 was used only for clock training, whereas D1 was used only for testing) because the manuscript and figure legends refer to “independent” and “external” datasets in the relevant sections. However, based on the code, it appears that samples from the same data sources may have been split between the training and test sets, and that the authors define a “dataset” as a training/test set of samples rather than as a separate data source (D1, D2, etc.). This sample-level splitting strategy is acceptable for held-out validation, but I suggest referring to the groups as the “training set” and “test set” rather than as “independent datasets” to avoid misinterpretation. Besides, the authors should clarify in the Methods section which data sources were used for training, which were used for held-out testing, and whether any source contributed samples to both sets.

We thank the reviewer for this careful observation. We agree that the nomenclature used in the original manuscript was potentially misleading and have revised it accordingly.

First, we have homogenized the terminology throughout the manuscript: Samples used for model training are now consistently referred to as the ‘training set’ and the samples used for validation are referred to as the ‘held-out test set’. All references to “independent datasets” in this context have been removed.

Second, we have clarified in the Methods section that datasets D1-D4, D6-D8, D10-D12 were combined into a single pool of samples, which was then split into a training set ($n=306$) and a held-out test set ($n=105$) (Lines 641-649). This clarification makes explicit that individual data sources may contribute samples to both sets, and that the split was performed at the sample level rather than at the data source level.

Third, we have retained references to external datasets (e.g. D1, D2, etc.) elsewhere in the manuscript, as they allow readers to trace specific results back to their underlying data sources. In addition, we added the number of samples from each dataset contributing to the training and held-out test set to Supplementary Table 2.

4. The revised analysis provides statistical evidence that Old-Srsf1 mice show improved wound repair compared with Old-GFP controls at selected time points. However, because the AAV-GFP control itself shows a trend toward impaired wound healing, even though this reaches

significance only in young mice, the authors should clarify whether Old-Srsf1 improves wound repair relative to untreated Old-WT mice, not only relative to Old-GFP controls. This distinction is important for interpreting the result as rescue of ageing-related functional impairment in otherwise untreated aged animals rather than mitigation of AAV-associated burden.

We thank the reviewer for raising this important distinction. The original analysis already included untreated WT controls at both ages (Old-WT, Young-WT) in the linear mixed model alongside the AAV-treated cohorts (Old-GFP, Old-SRSF1, Young-GFP). To directly address the reviewer's concern, we have now explicitly reported the relevant pairwise contrasts in the Results (Lines 426-434) and Extended Data Fig. 11c legend (Lines 1432-1436).

The reviewer is correct that AAV-GFP imposes some healing burden, most clearly in young animals (Young-WT vs. Young-GFP at day 5, $P < 0.0001$). A smaller, non-significant trend is also present in aged animals (Old-WT vs. Old-GFP at day 9, $P = 0.056$, Tukey-adjusted). Crucially, the magnitude of SRSF1-mediated improvement (Old-SRSF1 vs. Old-GFP, day 5 mean difference = 0.17) exceeds the AAV-GFP burden observed in matched WT comparisons (Old-WT vs. Old-GFP, day 5 mean difference = 0.10), and Old-SRSF1 also trends toward faster healing than untreated Old-WT controls (day 5 mean difference = 0.07, $P = 0.22$; day 11 mean difference = 0.08, $P = 0.094$). These comparisons are consistent with interpretation of the SRSF1 effect as rescue of age-related functional impairment in otherwise untreated aged animals, rather than purely mitigation of AAV-associated burden, although we acknowledge that the Old-SRSF1 vs. Old-WT contrast does not reach statistical significance and the inference is therefore based on effect-size comparison rather than a direct statistical test of that contrast. Finally, Old-SRSF1 wound closure kinetics are statistically indistinguishable from Young-GFP across the entire time course (all timepoints n.s.), indicating that SRSF1 induction restores aged wound repair to the level of young AAV-matched controls.

5. The authors claim systemic rejuvenation of mice in response to AAV8-Tet-Srsf1. While they show improvement in selected functional parameters, frailty index dynamics, and lifespan extension, the term “systemic rejuvenation” remains too strong. Rejuvenation would imply that ageing-associated deterioration is reversed (i.e. improved relative to a baseline aged state), rather than merely slowed down. From the frailty-index trajectories, functional impairment appears to accumulate more slowly rather than being reversed. True rejuvenation would be expected to produce a negative slope, rather than a smaller positive slope. Therefore, more precise wording would be “healthspan extension”, “attenuated late-life decline”, or “improved late-life function”, rather than “rejuvenation”.

We thank the reviewer for this important clarification. We agree that “systemic rejuvenation” is too strong for the organism-level claims given that frailty-index trajectories show attenuated accumulation rather than reversal. We have revised the manuscript throughout to use more precise language for organismal phenotypes: “healthspan extension,” “lifespan extension,” “attenuation of decline,” and “delayed frailty onset.” Changes can be found in the section title (Line 415), Fig. 6 title (Line 467), Results (Lines 490, 492), Discussion (Lines 578-579), and Extended Data Fig. 12 title (Line 1450).

6. The added mouse RNA-seq analyses improve the manuscript, but the interpretation would be stronger if Srsf1-induced molecular changes were directly compared with normal ageing trajectories in the same mouse tissues. The authors report changes in pathways related to bioenergetics, immune signaling, cellular homeostasis, sterol metabolism, and other processes, and state that these pathways are often dysregulated during ageing. However, without a tissue-matched ageing comparison, it is difficult to determine whether these changes represent anti-

ageing, pro-ageing, or ageing-orthogonal effects. Therefore, it would be informative to perform DEG and pathway enrichment analyses using external mouse ageing gene expression datasets for the same organs, and to show the effects of *Srsf1* alongside the corresponding pathway changes observed during normal ageing. This would improve the interpretation of *Srsf1* effects in the in vivo mouse model from an ageing perspective.

We thank the reviewer for this suggestion, which substantially strengthens the manuscript. We compared SRSF1-induced pathway changes against normal-aging transcriptomic trajectories in tissue-matched samples using the Tabula Muris Senis bulk RNA-seq atlas (GSE132040; 21- vs 3-month-old C57BL/6J). For each SRSF1-significant pathway in our mouse data, we classified the effect as Rejuvenating (R; opposes aging direction), Pro-Aging (A; mimics aging direction), or Orthogonal (O; SRSF1-specific).

SRSF1 induction produced predominantly rejuvenating signatures in heart (80% R; 15% O; 5% A) and liver (64% R; 28% O; 8% A), mixed rejuvenating-orthogonal effects in the spleen (38% R; 54% O; 8% A), and directionally mixed effects in brain (35% R; 28% O; 37% A). The substantial orthogonal fraction across tissues (15-54%) indicates that SRSF1 also engages aging-independent biology. The strong liver rejuvenation is consistent with the reduced epigenetic age detected by the liver-specific intervention clock (Extended Data Fig. 12c). These results have been added as new Extended Data Fig. 12e, with Methods (Lines 698-713) and Results (Lines 498-500) updated.

7. The authors report significant extension of both median and maximum lifespan in mice following *Srsf1* overexpression, but they support these claims using only the log-rank test, which examines overall differences between survival curves rather than specifically testing maximum lifespan. If the authors wish to claim a statistically significant effect on maximum lifespan, they should perform a prespecified upper-tail survival test, such as the Wang-Allison test or another appropriate late-life survival test. Otherwise, the claim should be softened.

We thank the reviewer for this point. Following the recommendation, we added an upper-tail survival analysis. As these tests were run during review, they were not prespecified, so we present the maximal-lifespan result as supportive rather than confirmatory and have softened the corresponding claims in the manuscript.

We define maximal lifespan as the 90th survival percentile, the field-standard cutoff underlying the Wang-Allison test. As the continuous analogue of Wang-Allison, we estimated the 90th-percentile lifespan by quantile regression: *Srsf1* extended it by 95 days ($P = 0.030$). We also applied the Wang-Allison test itself at the same 90th-percentile (field-standard cutoff; $T = 933$ days): 2 of 10 *Srsf1* mice versus 0 of 9 GFP controls exceeded T (Fisher's exact $P = 0.47$). The two tests evaluate the same percentile, but the binary Wang-Allison version is underpowered at this cohort size ($n = 9-10$ per group), whereas quantile regression retains the magnitude of the difference. Late-life survival is further supported by the Fleming-Harrington G(0,1) tail-weighted log-rank test ($P = 5 \times 10^{-4}$) and restricted mean survival time at $\tau = 972$ days (+131 days, 95% CI 73-188, $P < 0.001$).

Accordingly, the revised manuscript attributes the maximal-lifespan claim to the 90th-percentile quantile regression, reports the Wang-Allison test alongside it, describes the Fleming-Harrington and RMST results as late-life survival extension, and notes that these analyses were performed post hoc. These changes are in the Methods (Lines 969-979), the main text (Lines 451-453), and the Fig. 6c legend (Lines 480-482).

Minor

1. The authors claim age reversal in response to FOXM1 overexpression in NHDFs (Fig. 1f), but this result is based on a single sample ($n = 1$) and is not statistically supported. It should therefore be described as a non-significant trend rather than as confirmed age reversal.

We agree with the reviewer. The *FOXM1* overexpression result is based on $n = 1$ and is not statistically supported ($P = 0.509$). We have softened the corresponding text in both the Results (Line 131) and the Fig. 1f legend (Line 112) to describe the *FOXM1* OE finding as a non-significant trend rather than confirmed age reversal.

2. Extended Data Fig. 2b: the statistical comparisons appear to compare different fibroblast lines under the same treatment rather than comparing treatment versus matched control conditions within the same cell line. If the goal is to show that metabolic or mitochondrial perturbations increase transcriptomic age, the more direct test would compare each treatment against the corresponding control condition within the same fibroblast line. As currently presented, the significant comparisons are difficult to interpret as evidence of treatment-induced transcriptomic ageing.

We thank the reviewer for this observation. The statistical comparisons in Extended Data Fig. 2b are paired t-tests of treatment vs. its line-matched control, with pairing by fibroblast line and passage (NT for all treatments except SURF1+DEX, which is paired against SURF1). The original figure design with brackets across the two lines within each treatment group confusingly suggested a between-line comparison. We have removed these brackets and updated the figure legend (Lines 1297-1301) to state explicitly that all comparisons are line-matched treatment vs. NT (or SURF1 for SURF1+DEX), paired by line and passage.

3. Extended Data Fig. 2c: the authors assess the effect of mitoNUTs and DEX on epigenetic age in human fibroblasts using the DNAm DunedinPACE clock, while using the Skin & Blood clock in other analyses (e.g., Extended Data Fig. 2d). The rationale for using DunedinPACE in fibroblasts is not clear, given that this model was trained on blood data. The Skin & Blood clock, which the authors use elsewhere, would be a more natural choice for this analysis. The authors should also indicate in the manuscript that the DNAm clock used in Extended Data Fig. 6e is the Skin & Blood clock. This is mentioned in their response to reviewers but appears to be missing from the manuscript itself.

We thank the reviewer and have clarified the rationale in the manuscript. For Extended Data Fig. 2c, the perturbation experiments were designed to test whether metabolic and mitochondrial stressors affect the rate of aging, which DunedinPACE (ref. 31) is calibrated to estimate and can serve as an orthogonal molecular readout to the transcriptomic age. For the SRSF1 and Rapa experiments (Extended Data Figs. 2d, 6e) the design is a cross-sectional endpoint comparison of chronological-age estimates, for which the tissue-matched Skin & Blood clock (ref. 27) is appropriate. DunedinPACE is also one of two pace-of-aging measures included in the standard clock panel of the dataset analyzed here (ref. 30). We have tightened the Results language to clarify that Extended Data Fig. 2c serves as an orthogonal molecular readout rather than a chronological-age replication (Lines 133-134), added this rationale to the Methods (Lines 993-998), and explicitly identified the Skin & Blood clock in the Extended Data Fig. 6e legend (Line 1377), as suggested.

4. Extended Data Fig. 5i: the changes in alternative splicing events induced by SRSF1 overexpression are intriguing, but they are compared with the NTg group rather than with BFP-expressing cells. The authors report that the screening platform itself induces substantial transcriptomic differences between NTg and BFP controls. Therefore, comparing SRSF1-

overexpressing cells with NTg cells may confound SRSF1-specific splicing effects with effects of the transgene system burden. A comparison of SRSF1 with BFP (and with OSKM as an alternative rejuvenative treatment) would better isolate SRSF1-specific splicing changes.

We thank the reviewer for raising this methodological point. We agree that comparing SRSF1-overexpressing cells with NTg controls may confound SRSF1-specific splicing effects with transcriptomic changes induced by the transgene system itself. Therefore, we performed the requested comparison using BFP-expressing cells as the appropriate baseline and included OSKM as an alternative treatment for context (Extended Data Fig. 5j).

The revised analysis shows that SRSF1 overexpression produces a splicing landscape dominated by alternative first exon (AF; ~45%) and skipped exon (SE; ~25%) events relative to BFP controls, while OSKM overexpression is characterized by a higher fraction of AF events (~70%) and fewer SE events (~10%). The remaining event types (e.g. Alternative Last Exon, Alternative 5' and 3' splice sites, Mutually Exclusive Exons and Retained Introns) are largely consistent across conditions and cell lines. We have included these results in the revised manuscript (Results, Lines 278-280; new Extended Data Fig. 5j, Lines 1351-1353) and updated Fig 3e and Methods (Line 721) to display the spliced genes from the BFP comparison. While some of the dual-regulated hub genes have changed (*NPLOC4*, *KDM5B* swapped with *CXXC1*, *OS9*, *UBE2G2*), they still consist of key epigenetic and ER pathway modifiers (Lines 281-283).

5. The manuscript should clearly state the exact age of mice and the time after treatment at which mouse DNAm and RNA-seq profiling were performed. This information seems to be currently lacking.

We thank the reviewer for catching this omission. Tissues for RNA-seq (n = 4/group; mean age 808.5 ± 39.5 days GFP, 865.8 ± 34.6 days SRSF1) and DNAm profiling (n = 5/group; mean age 791.0 ± 52.0 days GFP, 865.6 ± 70.6 days SRSF1) were collected at endpoint from the lifespan cohort. Per-animal ages and time since Dox initiation are now provided in Supplementary Table 9. These details have been added to the Methods (Lines 979-984) and Extended Data Fig. 12 legend (Lines 1452-1453).

6. Extended Data Fig. 12c: the mouse liver DNAm analysis is based on a mouse Liver Interventions Clock. It would be informative to also report results for more established chronological liver and pan-tissue mouse DNAm clocks to determine whether the findings are consistent across different clocks. This would improve the reliability and interpretability of the result.

We thank the reviewer for this suggestion. We have added the chronological liver and pan-tissue mouse DNAm clocks to Extended Data Fig. 12c and updated the figure legend (Lines 1457-1460) and results (Lines 490-493). The SRSF1-associated reduction in liver epigenetic age was significant with the intervention-optimized Liver Interventions Clock ($P = 0.02$), directionally concordant, but non-significant, with the chronological liver clock, and not detected by the pan-tissue clock.

Reviewer #4

We thank Reviewer #4 for suggesting we provide context for the mouse lifespan results.

Minor

1. The only remaining concern from my part is that the lifespan of the mice is relatively short compared to other studies. I think requesting a repeat of these studies at the current stage would be unfair, but I believe it would be in the interest of both the authors and the readers to point out that the lifespans of the mice are somewhat shorter than what is published in the general literature.

We have added a brief section in the discussion paragraph (Lines 572-576) noting that the median lifespan of GFP control mice in this study (778 days) is consistent with published median lifespans of C57BL/6 cohorts (sample-size-weighted median 740 days; unweighted median 801 days across 131 studies; Pabis *et al.*, Ageing Res. Rev. 2024, ref. 52), but it falls below cohorts maintained under optimal husbandry conditions, where median lifespans can approach 900 days. As all comparisons in this study are within-cohort under identical conditions, the relative SRSF1-mediated lifespan extension is robust to absolute baseline differences across published cohorts. We suspect contributing factors include AAV-associated burden, vivarium-specific environmental conditions, and the biweekly frailty assessments employed in this cohort, which exceed the handling frequency of most published longevity studies and may have introduced cumulative handling stress. However, we did not directly test these factors and acknowledge them as plausible rather than confirmed contributors.