Universal transcriptomic signature of age reveals temporal scaling of *Caenorhabditis elegans* aging trajectories

Andrei E. Tarkhov[1,2], Ramani Alla[3,4], Srinivas Ayyadevara[3,4], Mikhail Pyatnitskiy[1,5], Leonid I. Menshikov[1,6], Robert Shmookler Reis[3,4,7] and Peter O. Fedichev[1,8]

[1] Gero LLC, Novokuznetskaya street 24/2, Moscow 119017, Russia

[2] Skolkovo Institute of Science and Technology, Skolkovo Innovation Center, Novaya street 100, Skolkovo 143025, Russia

[3] Central Arkansas Veterans Healthcare System, Research Service, Little Rock, Arkansas

[4] Department of Geriatrics, Reynolds Institute on Aging, University of Arkansas for Medical Sciences, Little Rock, Arkansas

[5] Institute of Biomedical Chemistry, 119121, Moscow, Russia

[6] National Research Center "Kurchatov Institute", 1, Akademika Kurchatova pl., Moscow 123182, Russia

[7] Bioinformatics Program, University of Arkansas for Medical Sciences, and University of Arkansas at Little Rock, Little Rock, Arkansas

[8] Moscow Institute of Physics and Technology, 141700, Institutskii per. 9, Dolgoprudny, Moscow Region, Russia

To detect overlap or convergence among the diverse genetic pathways controlling longevity, we collected an RNA-seq dataset of 60 *C. elegans* age-dependent transcriptomes for nematode strains with vastly different lifespans. We selected four exceptionally long-lived mutants (mean adult lifespan extended 2.2- to 9.4-fold) and three examples of some of the most effective life-extending RNAi treatments (that increased mean lifespan by $19-35\%$). Principal Component Analysis (PCA) in all the strains and treated groups reveals aging as a transcriptomic drift along a single direction, coinciding with the first principal component of the combined dataset, suggesting that the coarse-grained dynamics of the state of an organism is a manifestation of a single-mode dynamic instability. To enable a comparison of our data with other experiments, we introduced a version of a transcriptomic biological age from publicly available datasets. The biological age dynamics in our experiments reveal a universal "aging trajectory" modulated by a single parameter, having the meaning of the rate of aging, which is inversely proportional to lifespan. We hypothesize that the scaling of gene-expression kinetics underlies the recently observed scaling of survival curves in *C. elegans*. This observation is independently confirmed in our selected strains and treatments. The experimental survival data imply a stochastic nature of transcript levels that lead to a deceleration of mortality, reaching a plateau at advanced age. The limiting mortality value at the plateau agrees closely with the Gompertz exponent estimated at the cross-over age near the average lifespan.

1 Introduction

The strongest lifespan extension yet recorded has been achieved in *C. elegans*, and corresponds to an almost 10-fold increase with the *mg44* nonsense mutation in the *age-1* gene (Ayyadevara et al. 2008; Shmookler Reis et al. 2011). However, this hyperlongevity requires homozygosity of the mutation for two generations, resulting in total pre-embryonal genetic disruption. In human subjects, reasonable anti-aging therapies would instead be applied in adulthood, ideally at advanced age. Sadly, the best "therapeutic" models yield significant, but considerably smaller reported increases in life-span by interventions at embryonic and post-embryonic stages (e.g., up to roughly +90% by *let-363* RNAi (Vellai et al. 2003)). Pharmacological inhibition of the gene or its homologues yielded even smaller effects on lifespan in flies, nematodes, yeast (Moskalev et al. 2010, 2011; Ye et al. 2014) and mice (Anisimov et al. 2010; Harrison et al. 2009; Miller et al. 2011). It remains to be understood why a single nonsense mutation can dramatically extend animal lifespan, while an RNAi block of the same gene does not produce the same effect, especially when administered later in life. Perhaps the mutation dramatically changes the molecular machinery of the whole organism during development such that the course of aging of the super-long-living strains is qualitatively different both in terms of rates and form, and hence, could not be easily reproduced therapeutically. Alternatively, perhaps the gene regulatory network is sufficiently robust, so that the pace of aging can be reduced by a therapy without qualitative alterations of the relevant molecular mechanisms.

To address these alternatives, we compiled an RNA-seq dataset of age-dependent transcriptomes, produced from *C. elegans* strains that have vastly different lifespans. We

started by studying the most long-lived isogenic *C. elegans* strains carrying mutations: *daf-2*(*e1370*) [strain SR806], *age-1*(*mg44*) [SR808, at the first and second generations of homozygosity], and the longest-lived *daf-2(e1391);daf-12(m20)* double mutant [strain DR1694] (Ayyadevara et al. 2008; Shmookler Reis et al. 2011). The average lifespans in the series range from twofold to nearly tenfold longer than that of the wild type. To model the effects of therapeutic interventions in adult animals, we chose five target genes for RNAi treatments; *daf-4*, *che-3*, *cyc-1*, *cco-1*, and *eat-4*. Mutation or RNAi of each of them were reported to prolong life-spans: +40-120% by *daf-4(e1364)* (Shaw et al. 2007), +37% by *che-3(e1124)* and +100% by *che-3(p801)* (Apfeld et al. 1999), +87% (Dillin et al. 2002) and +60-100% (S.-J. Lee et al. 2010) by *cyc-1* RNAi, +61% (Dillin et al. 2002) and +57-80% by *cco-1* RNAi (S.-J. Lee et al. 2010), and by *eat-4* RNAi (Hamilton et al. 2005) (for which effect size was not stated). We confirmed longevity of worm strains subjected to these five RNAi interventions at 20°C, and selected three RNAi treatments (*daf-4*, *che-3* and *cyc-1*) for transcriptomic analyses representing knockdown of diverse pathways. The overall variation in lifespan among the strains across all experiments ranges from 17 to 160 days. Wildtype nematodes lived for 17, 18.5 and 25.5 days. For each of the mutants or interventions, we characterized transcriptomes over time, across their lifespans, collecting 60 transcriptomes in total (9 different biological conditions). The experimental data were augmented by a meta-analysis of publicly available gene expression measurements in *C. elegans* (roughly 4000 samples).

Principal Component Analysis (PCA) of gene expression reveals aging in all strains and treated groups is characterized by a highly coordinated drift across multiple transcripts along a single direction, coinciding with the first principal component of the combined dataset. The observation is consistent with our "aging at criticality" hypothesis (Podolskiy et al. 2015). This hypothesis suggests that the gene regulatory networks of most species operate near an order-disorder bifurcation point (Balleza et al. 2008) and are intrinsically unstable (Kogan et al. 2015). The coarse-grained dynamics of an organism's physiological state can be associated with a stochastic evolution of an emergent organism-level property, coinciding approximately with the first principal component score. Thus, this is a natural indicator of the organism's development and aging progress and is directly associated with mortality. The evolution of the biological age variable as a function of chronological age is completely determined by the underlying regulatory interactions. The curvature of the potential imposes a natural time scale, which is proportional to the mortality rate doubling time, the most important characteristic of the aging trajectory. The pattern of molecular changes associated with age appeared to be robust across diverse biological conditions tested in our experiments. We produced a transcriptomic signature of age using publicly available gene expression data from *C. elegans* and demonstrated that it works well across the datasets. If applied to all our data, the change of biological age over time is characterized by a universal "aging trajectory" modulated by a single parameter that indicates the rate of aging and is inversely proportional to life-span. We infer from this that life-extending mutations and knockdown therapies extend lifespan by stabilizing the regulatory network and thus reducing the rate of aging. Therefore, the universality of aging trajectories might indicate the molecular basis for the scaling universality of survival curves recently observed (Stroustrup et al. 2016) and independently confirmed in the survival data of all the strains and treatments in our experiments.

Finally, we addressed the role of stochasticity in aging dynamics and the saturation of mortality later in life. The survival data analysis from an independent experiment confirmed another important theoretical prediction on the equivalence between the Gompertz exponent (inferred at the cross-over age, corresponding to the average lifespan) and the limiting mortality value (corresponding to the mortality plateau). Therefore the "aging at criticality" hypothesis emerges as a powerful tool for the understanding of a broad range of aging-dynamic and survival properties. This theoretical framework will be helpful for future efforts to identify anti-aging targets in *C. elegans* and other species.

2 Results

2.1 Selection of long-lived strains and life-extending interventions

Several mutations leading to exceptional longevity of *C. elegans* have been identified (Friedman et al. 1988; Gottlieb et al. 1994; Kenyon et al. 1993; Klass 1983) and studied extensively for their remarkable elevations of both lifespan and stress-resistance traits (Ayyadevara et al. 2008; Shmookler Reis et al. 2011) (see Table 1). To model the effects of anti-aging interventions, we tested five RNAi treatments targeting genes known to affect aging (see Table 1 for the summary of our survival experiment). In some cases, the relative lifespan modification was somewhat smaller, than reported, which was likely due to differences between mutation and RNAi treatments, or differences from the original studies, in RNAi-exposure times and/or culture temperatures (see Table. S1 for comparison).

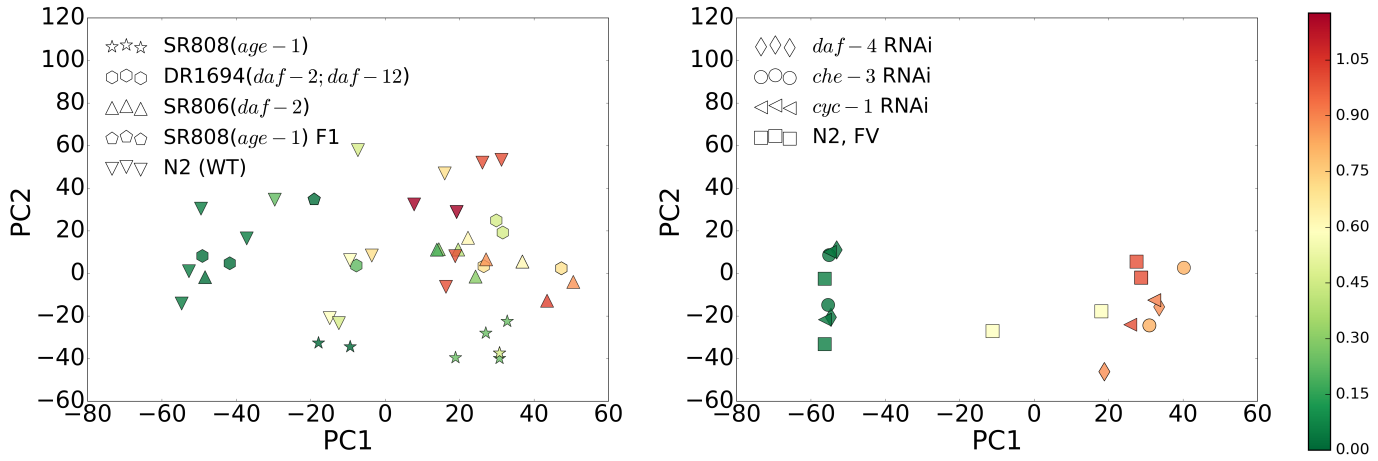
The strongest effect we obtained by RNAi corresponds to an increase in mean lifespan from 18.5 days to 25 days (+35%) by RNAi of *daf-4*. We selected three genes producing significant results in our survival experiment and representing apparently diverse modes of action, *che-3*, *daf-4* and *cyc-1*, for follow-up with transcriptomics (see Table 1, highlighted in bold).

Table 1
Summary of survivals for the mutant strains and the confirmatory RNAi interventions.

Mutant strain[a]	LS[*](d)	Rel. LS
SR806 [*daf-2(e1370)*], N2-DRM background[b]	39	2.3
SR808 [*age-1(mg44)*], second-generation homozygotes in N2-DRM background[b]	160	9.4
SR808 [*age-1(mg44)*], first-generation homozygotes ("F1") in N2-DRM background[b]	38	2.2
N2-DRM (WT)[b]	17	1.0
DR1694 [*daf-2(e1391);daf-12(m20)*][c]	60.5	2.4
N2-DRM (WT)[c]	25.5	1.0

Gene for RNAi[d]	Protein encoded	LS[*](d)	Rel.LS
N2, FV	(none)	18.5	1.00
***daf-4* RNAi**	**TGF-beta receptor ortholog**	**25**	**1.35**
***che-3* RNAi**	**Dynein H chain iso 1b**	**24.5**	**1.32**
***cyc-1* RNAi**	**Cytochrome C1**	**22**	**1.19**
cco-1 RNAi	Cytochrome C oxidase	24	1.30
eat-4 RNAi	BNPI glutamate transporter[**]	21	1.14

[a]The long-lived-mutant *C. elegans* strains (Ayyadevara et al. 2008; Shmookler Reis et al. 2011) used for mRNA preparation and analysis by RNA-seq technique.
[b,c]The survivals were done separately. The relative lifespan was calculated with respect to the wild-type control from the corresponding series.
[d]The survival analysis summary for the RNA interference experiments against the five selected *C. elegans* target genes. Three successful RNAi interventions (in bold) representing apparently different mechanisms of action were selected for further transcriptomic measurements.
Abbreviations: LS, mean adult lifespan; Rel. LS, relative adult lifespan; N2, the DRM stock of wild-type strain Bristol-N2; FV, empty feeding vector in place of RNAi.
[*]Mean adult lifespan, days post-L4/adult molt.
[**]Affecting pharyngeal pumping.



(a) (b)

Figure 1. Principal components analysis (PCA) of the experimental RNA-seq datasets for (a) four long-lived mutants and *C. elegans* wild-type (Bristol-N2, strain DRM), and (b) three life-prolonging RNAi-treated groups and *C. elegans* FV controls (fed bacteria harboring empty feeding-vector plasmid). The marker type denotes strains and RNAi groups (see Table 1 for lifespans). The markers in (a) and (b) are colored according to age rescaled to the average lifespan in the groups (see the color bar).

2.2 Age-dependent transcriptomes of long-lived *C. elegans* strains

To understand the nature of gene expression changes associated with aging and lifespan, we produced an age-dependent RNA-seq experimental dataset, consisting of the four mutant strains, the three examples of life-extending RNAi and two controls representing *C. elegans* wild-type (Bristol-N2, strain DRM) from Table 1. We utilized a standard pipeline for RNA-seq data processing (see Section S4 for details), yielding a dataset of 60 transcriptomes.

Principal Component Analysis (PCA) of gene expression is represented separately for the mutants (Fig. 1a) and the RNAi treated strains in Figs. (1a) and (1b), respectively. Aging emerges as a transcriptomic drift: a highly coordinated change of multiple transcripts in a single direction in the multidimensional space of all possible transcript levels, coinciding with the first principal component of the combined dataset. The data covariance matrix is profoundly singular (the first four principal components explain nearly 75% of the overall variance in the data). This observation is quite striking, considering that the number of samples in the data is dramatically smaller than the number of identified genes.

Remarkably, the nature of the gene expression variance and its association with the first principal component remained quite consistent across all the experiments, for samples ranging from youngest to oldest of each cohort, irrespective of the variation in lifespan. The first principal component score (PC1) is highly correlated with age rescaled by the average lifespan, $r = 0.75$ ($p < 10^{-10}$, accounting for $r^2 = 56\%$ of total variance). No other principal components scores have any appreciable correlation with age. The third principal component score is associated with the rescaled age variable, $r = 0.4$ ($p \approx 10^{-3}$), which would account for $r^2 = 16\%$ of total variance, just reaching significance after all appropriate adjustments. It therefore appears that, apart from the obvious difference in the rate of aging, the character of age-related changes in gene-expression in mutants that have large effects on lifespan (Fig. 1a), are approximately the same as in RNAi interventions that show smaller effects (Fig. 1b). Aside from PC1, the positions and aging dynamics of points representing the strains and the RNAi-treated samples differ in the directions of age-dependent drift. Transcript-level variation is a signature of altered biology and contains information about the effects of specific mutations or gene silencing, including transcriptomic variation associated with the apparent difference in the rate of aging. We are leaving such an analysis for a future work.

2.3 Meta-analysis of aging dynamics in *C. elegans* transcriptomes

Our gene expression data suggest that there may be a universal aging signature; a set of genes associated with aging at the transcription level under a reasonably broad range of experimental conditions and genetic backgrounds. Typically, however, any small-scale experiment, including ours, yields a huge number of measured transcripts' levels, albeit in relatively small number of samples. Any analysis of such small datasets is challenging (Gui et al. 2005; Johnstone et al. 2009) and appropriate only for illustrative purposes because it is prone to over-fitting and false positive errors even if all of the samples are collected in the same laboratory under the same conditions. A direct comparison of gene expression data obtained in different laboratories is further complicated by divergence in experimental procedures, leading to uncontrollable batch effects requiring extensive normalization (Johnson et al. 2007; Lê Cao et al. 2014; Rohart et al. 2017).

To address this dilemma, we predicted that a robust aging signature could be obtained from a sufficiently large collection of publicly deposited "shallow" datasets from independent experiments (a dozen samples each, on average) using *C. elegans*, since all the transcriptomes would have to share the same essential biology. In contrast, the specific experimental conditions and uncontrollable batch differences should be mostly uncorrelated, and thus would comprise "noise" in a combined dataset of sufficient overall size. To demonstrate such a possibility, we compiled a comprehensive transcriptomic collection for *C. elegans* by combining all the publicly-available gene-expression experiments into a single database. The resulting "MetaWorm" dataset contains, *in toto*, more than 400 different transcriptomic experiments with $N = 3724$ nematode samples characterized by $G = 4861$ of the most commonly expressed/detected genes in the samples. The composition of the dataset and the details of the normalization procedures are described in Section S5.

The gene-expression variance in the combined MetaWorm dataset is dominated by batch effects and hence we do not expect PCA to reveal aging in association with the first principal component in a totally unsupervised way. Instead, we attempted to identify the direction of aging by first searching for genes associated with age. Our MetaWorm dataset is sufficiently large to generate the cross-validation ensemble of single-gene association tests using random resampling. We further reduced the number of candidate genes by thresholding the transcripts by the frequency of significant associations in the resampling; we estimate that the chosen cross-validation threshold corresponds to $p < 10^{-6}$ uncorrected, or $p < 0.005$ after Bonferroni correction. The final list of genes robustly associated with aging in the MetaWorm study consists of 327 genes: 260 up- and 67 down-regulated with age, respectively. The genes are therefore markers of age and are surprisingly few in number: approximately 4000 out of 4861 genes were never observed in the top-200 gene-list after resampling (see Electronic Supplementary Materials for the full list of genes implicated in aging in our calculations).

Table 2
Transcription factors, RNA-binding proteins and miRNAs whose targets are significantly enriched in the identified MetaWorm aging genes (327 genes).

Transcription factors or RNA-binding proteins (WormExp)[a]	FDR
DAF-16 targets (Kumar et al. 2015)	4.6e-7
CEC-3 targets (*spr-5* mutant, generation 10) (Greer et al. 2014)	2.0e-4
small RNAs decreased by starvation at P0 (Rechavi et al. 2014)	1.3e-3
PQM-1 L3 targets (Wei Niu et al. 2011)	1.4e-3
Rb/E2F pathway (DPL-1,EFL-1,LIN-35), intestine (Kudron et al. 2013)	2.7e-3
PMK-1 targets down in Day 15 vs. Day 6 (Youngman et al. 2011)	3.8e-3
age-regulated **ELT-2** targets (Mann et al. 2016)	3.8e-3
up by CSR-1 and w/out CSR-1-bound 22G RNA (Gerson-Gurwitz et al. 2016)	7.1e-3
proteins interacting with CEY-1 (Arnold et al. 2014)	7.1e-3
MUT-14, SMUT-1 (DEAD box) (Phillips et al. 2014)	1.6e-2
Hox gene targets (LIN-39, MAB-5, EGL-5) (W. Niu et al. 2011)	3.3e-2

Transcription factors (Fuxman Bass et al. 2016)[a,b]
NHR-1, NHR-10, NHR-86 (Seah et al. 2016)
ELT-6 (Budovskaya et al. 2008)
ZTF-9 (Howe et al. 2016)
MAB-3

miRNA (TargetScan)[a,c]
miR-57
miR-244
miR-253
***let-7*/miR-48/84/241/795** (Hsin et al. 1999; Inukai et al. 2013)
miR-245

miRNA (MirTarBase)[a,d]
miR-59-3p
miR-256
miR-1819-3p
miR-60-3p (Kato et al. 2016)
miR-52-5p

[a]Longevity-related regulators are highlighted in bold.
[b]The top-six transcription-factor hits are reported.
[c,d]The top-five miRNA hits from each of the databases.

The MetaWorm aging-gene set may not be exhaustive, yet by design it was reproducible across independent experiments and hence should be useful for *C. elegans* aging studies. In our dataset, for example, 1200 genes are significantly associated with the age rescaled by the average lifespan ($p < 0.05$ after Bonferroni correction) out of 4861 genes most commonly detected in the MetaWorm samples. However, the selection by means of Bonferroni correction may not be stringent enough, since it is based on the assumption of independence among the genes considered. The prominence of transcription factors among the genes that are age-dependent would inevitably lead to a gene-set with high internal cross-correlation, and a far higher than predicted fraction of age-associated genes. The Mann-Whitney U test shows that 327 genes comprising the MetaWorm aging genes signature is highly enriched with the genes having the most significant correlation with rescaled age in our RNA-seq data (the corresponding area-under-curve (AUC) statistic for the receiver operating characteristic (ROC) curve is AUC = 0.610 ± 0.015, at the significance level $p < 10^{-30}$).

2.4 Biological interpretation of aging drift

Correlation with age does not necessarily imply a causal relation to aging, yet genes correlated with age are usually the primary target in aging studies. To decipher all causal relations among genes related to aging, one would have to directly reconstruct a substantial part of the gene-gene interactions, controlling the subspace responsible for the progression of aging, which is beyond the scope of this work.

We could check, however, whether the MetaWorm aging genes were enriched for the targets of gene expression regulators (see Table 2). We used four databases for enrichment analyses: a curated database for transcription factors and RNA-binding proteins from published high-throughput expression studies in *C. elegans* WormExp (Yang et al. 2016), a high-quality protein-DNA interaction network (Fuxman Bass et al. 2016), and two databases of miRNA-target interactions: the *in silico* predicted TargetScan (Jan et al. 2011) and the experimentally validated MirTarBase (Chou et al. 2016).

Remarkably, the enrichment analysis of the MetaWorm aging genes detected 10 of the previously-known longevity-related regulators: DAF-16 (Henderson et al. 2001), ELT-2 (Mann et al. 2016), ELT-6 (Budovskaya et al. 2008), PMK-1 (Troemel et al. 2006; Youngman et al. 2011), PQM-1 (Tepper et al. 2014), NHR-1, NHR-10, NHR-86 (Seah et al. 2016), *let-7* (Hsin et al. 1999; Inukai et al. 2013), and miR-60 (Kato et al. 2016).

2.5 Scaling universality of aging trajectories

Principal Component Analysis (PCA) from Section 2.2 shows that the progression of aging is dynamically restricted to a one-dimension manifold, a trajectory, and therefore the first PC score (or any other metric associated with the organism state vector position along the aging trajectory) may be used as a plausible biomarker of age. We used our MetaWorm dataset to produce a transcriptomic biomarker of biological age, based on linear regression of the gene-expression levels to chronological age, cf. DNA methylation age, which is a weighted sum of DNA methylation features, trained to "predict" chronological age in humans (Hannum et al. 2013; Horvath 2013; Marioni et al. 2015) and mice (Maegawa et al. 2010; Singhal et al. 1987). Such an algorithmic-optimization approach for the determination of biomarkers of biological age in highly-dimensional data poses a number of challenges, mostly due to a limited number of samples even in the larger MetaWorm database. The degeneracy of the data is not, however, a purely technical issue, but is rather a physical property of the underlying biological system, a direct consequence of the inherent low-dimensional nature of the aging dynamics.

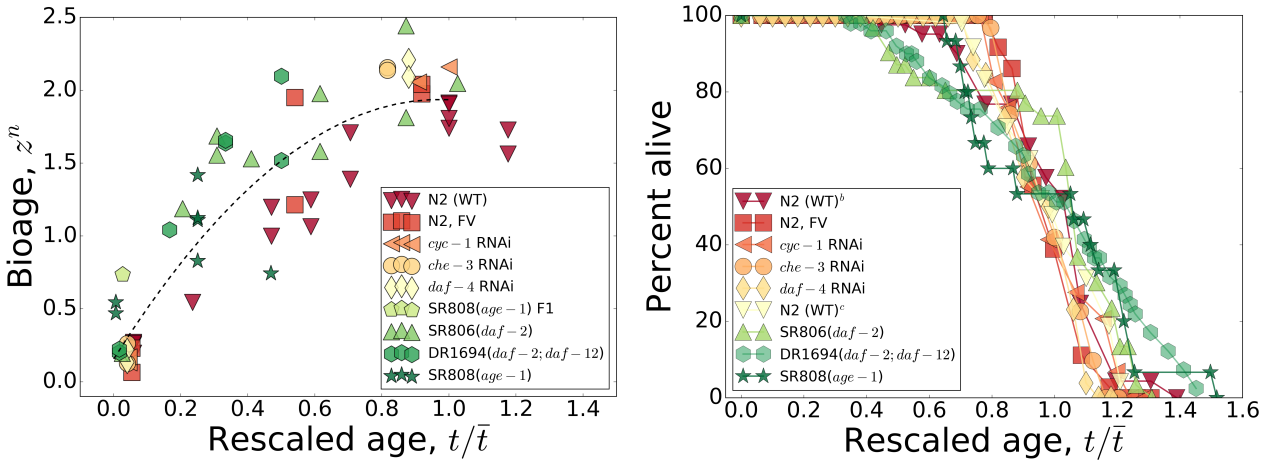
There is no unique solution to the regression problem, so any number of other convenient signatures can be obtained by applying additional requirements, such as, for example, sparsity. We choose to perform a cross-validated lasso regression of gene expression in the MetaWorm dataset on age re-scaled by the average lifespan, in this 327-gene subset comprising the robust direction of aging (see Electronic Supplementary Materials for the summary of the respective regression weights). The pipeline is a variant of the distance correlation sure independent screening (DC-SIS), which is a fairly intuitive proposition given the extremely high dimensionality of the data (Li et al. 2012).

To test if the regression model does not over-fit and hence retains its predictive power, we performed the biological age prediction from our RNA-seq data (not used during training). The biological-age dynamics yield a universal dependence of the transcriptomic bioage as a function of a rescaled variable, representing the chronological age divided by the average lifespan of a strain, for the mutants and the RNAi treated strains, see Fig. 2a (c.f. with Fig. S1a of the same data as a function of age without the rescaling). The scaling transformation works exceptionally well and brings together the aging trajectories of animals with average adult lifespans ranging from 17 days for wild-type N2-DRM control worms to 160 days for *age-1(mg44)* mutants (the Pearson correlation of the bioage with the rescaled age is $r = 0.86$ ($p = 2 \cdot 10^{-18}$), cf. $r = 0.54$ ($p = 8 \cdot 10^{-6}$) for the chronological age in Fig. S1a with no rescaling). In other words, in the coarse-grained sense, the transcriptome remodelling at any given age is pretty well characterized by a single quantity, the biological age. The biological age dynamics is controlled by yet another single variable, having a meaning of the rate of aging and depending on the biological conditions, such as genetic background or therapy.

The aging signature, such as that found in the MetaWorm list consisting of the transcripts of 327 genes universally associated with aging in multiple studies, is a well defined object with a clear biological meaning. On the other hand, the magnitude and the sign of the contribution of individual transcripts to the biological age model is not unique and hence has no direct physical or biological meaning other than being a convenient tool for experimental data analysis.

2.6 Temporal scaling of mortality law, mortality deceleration

Quite naturally, the scaling universality of the variation in gene transcript levels along the aging trajectory exemplified by Fig. 2a might be a molecular-level explanation behind the temporal scaling of survival curves, first reported in (Stroustrup et al. 2016). In that study, nematodes were subjected to various life-shortening stresses, and had their life-spans reduced in such a way that any two survival curves could be transformed one into another by a simple rescaling of age. Our survival data with life-extending mutations and RNA interference follow the same pattern after rescaling by age, see Fig. 2b for the survival fraction in all our experi-



(a) (b)

Figure 2. The universal temporal scaling of survival curves and transcriptomic bioage. (a) The universal scaling of transcriptomic bioage z^n (calculated from the transcriptomic signature obtained independently via training on the MetaWorm dataset) as a function of the rescaled chronological age $t^n/\bar{t}$ for the in-house collection of gene-expression data for four long-lived mutants and *C. elegans* wild-type (Bristol-N2, stock DRM), and three groups of N2 wild-type worms treated with life-prolonging RNAi or bacteria carrying only the empty feeding vector (FV) without an RNAi insert. The dashed line is a guide-to-eye for the apparent bioage plateauing. (b) The temporal scaling of the corresponding survival curves from (a), as a function of the rescaled chronological age $t^n/\bar{t}$. In both panels, all markers are colored according to the lifespans of the strains (see Table 1 for the lifespans): red for small and green for large lifespans. Overall, the scaling spans almost tenfold variation in median adult lifespan, ranging from 17 to 160 days.

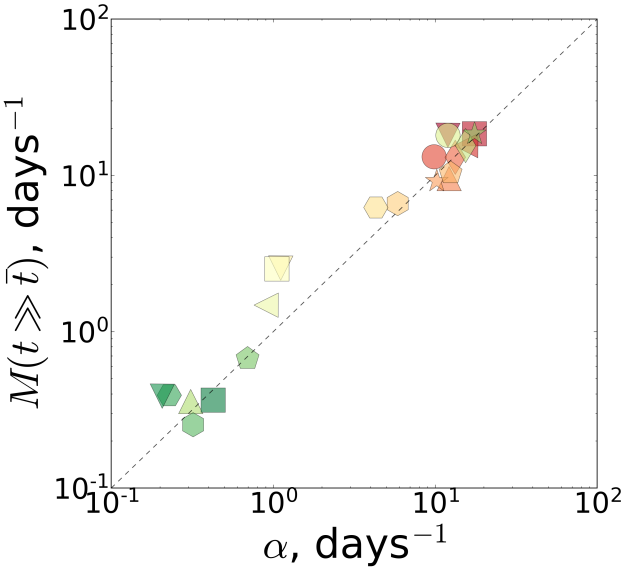


Figure 3. The plateau mortality $M(t \gg \bar{t})$ obtained from the tail of cumulative hazard $m(t)$ versus the Gompertz exponent α calculated from the experimental *C. elegans* survival curves (Stroustrup et al. 2016). The predicted behavior is shown by the dotted line. The values of the average lifespan $\bar{t}$ are depicted by the pseudo-color: red for small and green for large values.

ments as a function of the rescaled chronological age (c.f. with Fig. S1b of the same data as a function of age without the rescaling). The scaling transformation works exceptionally well and brings together the aging trajectories of animals with drastically different average adult lifespans: from mere 17 days for wild-type N2-DRM control worms and up to 160 days for *age-1(mg44)* mutants.

The scaling property of the aging trajectories and the survival functions can be naturally explained with the help of the aging-at-criticality model, providing a coarse-grained description of aging dynamics with the help of a simple stochastic Langevin equation and allowing for an analytic solution for mortality and survival fraction (Podolskiy et al. 2015). Early in life, up to approximately the average lifespan, mortality turns out to follow the Gompertz law $M(t) \approx M_0 \exp(\alpha t)$, where M_0 is the initial mortality rate (IMR) and α is the Gompertz exponent, related to the mortality rate doubling time. We predicted however, that the exponential rise in mortality rates would cease at late ages, approaching a plateau determined at the value of the Gompertz exponent (ibid.). The mortality records from (Stroustrup et al. 2016) can be used to test the theoretical prediction. We computed the approximate values of the Gompertz exponent using the data in mid-life and the mortality plateau estimates later in life for all the reported conditions, see Section S7 for the details of the calculations. The results are summarized

in Fig. 3 and demonstrate a fairly tight correlation between the parameters, in a remarkable semi-quantitative agreement with the theoretical calculation, across a life-span range of almost two orders of magnitude.

3 Discussion

The apparent reduction of the gene expression dynamics to a one-dimension manifold, the self-similarity of the aging trajectory and the survival curves can be naturally explained with the help of an aging-at-criticality conjecture (Podolskiy et al. 2015). The coarse-grained change in gene expression in an aging organism is then characterized by a one-dimensional drift in direction, coinciding with the singular eigenvector of the data covariance matrix, or the first principal component. The results of the present work support the concept of aging being an organism-level manifestation of critical dynamics of a single unstable mode of the underlying gene regulatory network.

A key observation, relevant to both theoretical and practical applications, is that there is a subset of genes reproducibly associated with aging in all publicly available transcriptomic data. The long-lived mutants and RNAi examples characterized by as much as 10-fold variation in lifespan share the same aging direction. The apparent stability of the aging direction is not surprising from a theoretical viewpoint since the lifespan and the aging signature are related to the smallest eigenvalue and the corresponding eigenvector of the kinetic gene-gene interaction matrix, respectively (see Section S6 for further details). Small perturbations, such as effects of mutations or RNAi, thus produce small variations in already small eigenvalue and hence may yield very large variations in the lifespan. At the same time, alterations of the aging direction by a weak perturbation remains small.

The robust nature of the aging signature let us perform a meta-analysis of published gene expression data from the MetaWorm dataset. Using the age of the animals, we could reproducibly identify the gene-signature associated with aging in most of the available gene-expression measurements, including ours. There are relatively few aging genes (less than 7% of all the available transcripts) and these are enriched with gene-expression regulators that promote longevity via disparate pathways, such as DAF-16, ELT-2, ELT-6, NHR-1, NHR-10, NHR-86, ZTF-9, *let-7*, and miR-60 (see Table 2 for the complete results of the over-representation analysis). It would be thus sensible to test experimentally some of the uncharacterized hits from our lists.

An opposite argument can be made: given the effectively one-dimensional character of the physiological state vector trajectory, we can regress the gene expression data to predict the chronological age and therefore obtain a universal transcriptomic biological age model, suitable for aging studies in *C. elegans*. The dependence of the biological age as a function of chronological age can thus be used to distinguish the progression of aging across strains. Early in life, up to approximately average lifespan, the age-dependent rise in transcriptomic indices of biological age is a universal function of a dimensionless age variable, obtained by rescaling the chronological age to the strain life expectancy. This is easy to interpret if the influence of stochastic forces is small in young animals and therefore the progression is almost deterministic with the same time scale defining the shape of the gene expression variations and the value of the average lifespan. In theory, this happens whenever the average lifespan greatly exceeds the mortality rate doubling time. In practice, the assumption is only qualitatively correct, but still provides a reasonable explanation of the experimental situation. According to the model, the temporal scale is defined by the underlying gene regulatory network stiffness and thereby mechanistically relates the organism level properties, such as lifespan, with potentially modifiable molecular-level interaction properties of the underlying regulatory network, such as the characteristic molecular and genetic damage and repair rates (Kogan et al. 2015).

The temporal scaling of aging trajectories and the survival curves is, of course, an approximate statement. The dynamics of biological age are stochastic and therefore depend on random environmental and endogenous factors. The intensity of random forces perturbing the regulatory network is the other major variable, determining statistical properties of the organism state vector dynamics and mortality statistics in experimental cohorts. This may be an explanation behind the deviations from temporal scaling in survival curves in different replicates of the same strain in (Stroustrup et al. 2016). Certain external conditions or therapeutic interventions, in principle, may produce a developmental delay or acceleration without a change in the rate of aging and thus produce deviations from the universal scaling behavior. We may see such an example in our version of the survival curve corresponding to DR1694 strain and demonstrating the most discordant survival fraction as a function of the rescaled age. Such small but significant deviations from the aging trajectory could be explained by different levels of stochastic noise in the gene expressions and would deserve an independent study.

Later in life, we expect the dynamics of gene expression and mortality should increasingly depend on non-linearity of the gradually disintegrating gene regulatory network, as the aging drift and stochastic forces perturb it. Large deviations from the youthful state are incompatible with survival and hence the biological age should plateau at roughly the average lifespan. Fig. 2a supports this interpretation in all our experimental cohorts across a 10-fold range of lifespan difference. As the biological age approaches the limiting value, mortality also decelerates and reaches a plateau at the level of the Gompertz exponent obtained from an exponential fit in the age range close to the mean lifespan (Podolskiy et al.

2015).

Using the survival data from (Stroustrup et al. 2016), we showed that the age-dependent mortality in *C. elegans* indeed decelerates and reaches a plateau at late ages near the level of the Gompertz exponent, measured in the vicinity of the average lifespan, thus fully confirming the theoretical prediction. This phenomenon is not limited to the experiments with nematodes and apparently underlies the plateau in mortality rates observed previously in very large populations of medflies and fruit flies (Carey 2011; Rose et al. 2012), which we have here extended to relatively small populations of *C. elegans* (Fig. 3). The results match expectations and, together with the scaling universality of the aging trajectory, both in transcriptomes (Fig. 2a) and corresponding survival curves (Fig. 2b), are the key results of our study and supports our theoretical model.

There can be multiple explanations for plateauing mortality at advanced ages (Gavrilov et al. 2001; Strehler et al. 1960; Weitz et al. 2001). Yet, all the proposed models are multi-parametric. The main advantage of our approach is that it the exponential increase of mortality earlier in life and the saturation of mortality late in life are explained within a simple reaction kinetics theory framework with the help of the single parameter, identifiable with the Gompertz exponent, measured at midlife.

Critical dynamics can be a useful metaphor for qualitative understanding of the nature of aging dynamics well beyond nematodes. For example, a recent analysis of the development and aging in a slice of the large-scale human 2003-2006 National Health and Nutrition Examination Survey (NHANES) representing physical activity metrics (Pyrkov et al. 2017) shares most of this work's findings. The data are extremely deep (ca 10000 samples), collected in an extremely uniform and controlled way. Therefore PCA helps to reveal the age-related changes in association with the direction of the most variance in the signal as well. The progress along the aging trajectory explains most of the variation in mortality. Therefore, the corresponding collective mode variable has a meaning of biological age. Across the population, the biological age increases approximately linearly with age and exhibits a growing variance, a signature of stochastic forces.

The increasing stochastic heterogeneity effects (including the leveling-off of mortality and bioage) help explain when an anti-aging treatment should be applied to obtain the largest possible effect. We speculate that at pre-embryonal and embryonal stages in the simplest animals, or early in life in humans, the growth of an organism is to a large degree determined by a developmental program. At more advanced ages, the stochasticity of the gene regulatory network kinetics takes its toll over time and leads to increasing phenotypic heterogeneity at every level. Accordingly, we expect that anti-aging interventions at the early stages have a broader and more generic effect on aging across diverse species. In contrast, interventions applied at late ages should be precisely selected to help treat specific conditions of an individual at a given point along the aging trajectory; a consequence of life history in the form of stochastically accumulated errors.

The experimental evidence for the relation between the transcriptomic aging trajectories and the survival curves are key conclusions from this study. The particular form of this dependence is the temporal scaling, implying a possibility of modifying the rate of aging and hence eventually dramatically increasing lifespan while avoiding detrimental effects. We therefore believe that these ideas of aging as a critical dynamic of the physiological state vector will provide a systematic approach to map longevity-controlling mechanisms and, eventually, bring about powerful lifespan-modifying interventions.

4 Acknowledgements

We thank Prof. Fontana and Nicholas Stroustrup for providing raw survival data and experimental lifetables from (Stroustrup et al. 2016), G. Getmantsev and O. Burmistrova for proof-reading of the work, I. Molodtsov and V. Kogan for invaluable help and advice on data analysis. The work was funded by Gero LLC, NIH Grant P01 AG012411-17A1 (WST Griffin, PI), and VA Merit Review Grant I01 BX001655 (RJ Shmookler Reis, PI).

5 Conflict of interest

None declared.

References

Anisimov, V. N., Egormin, P. A., Piskunova, T. S., Popovich, I. G., Tyndyk, M. L., Yurova, M. N., ... Romanyukha, A. A. (2010). Metformin extends life span of HER-2/neu transgenic mice and in combination with melatonin inhibits growth of transplantable tumors in vivo. *Cell Cycle*, *9*(1), 188–197.

Apfeld, J. & Kenyon, C. (1999). Regulation of lifespan by sensory perception in Caenorhabditis elegans. *Nature*, *402*(6763), 804–809.

Arnold, A., Rahman, M. M., Lee, M. C., Muehlhaeusser, S., Katic, I., Gaidatzis, D., ... Ciosk, R. (2014). Functional characterization of C. elegans Y-box-binding proteins reveals tissue-specific functions and a critical role in the formation of polysomes. *Nucleic Acids Res.* *42*(21), 13353–13369.

Ayyadevara, S., Alla, R., Thaden, J. J., & Shmookler Reis, R. J. (2008). Remarkable longevity and stress resistance of nematode PI3K-null mutants. *Aging Cell*, *7*(1), 13–22.

Balleza, E., Alvarez-Buylla, E. R., Chaos, A., Kauffman, S., Shmulevich, I., & Aldana, M. (2008). Critical dynamics in genetic regulatory networks: Examples from four kingdoms. *PLoS One*, *3*(6), e2456.

Budovskaya, Y. V., Wu, K., Southworth, L. K., Jiang, M., Tedesco, P., Johnson, T. E., & Kim, S. K. (2008). An elt-3/elt-5/elt-6 gata transcription circuit guides aging in C. elegans. *Cell*, *134*(2), 291–303.

Carey, J. R. (2011). Biodemography of the mediterranean fruit fly: Aging, longevity and adaptation in the wild. *Experimental Gerontology*, *46*(5), 404–411.

Chou, C. H., Chang, N. W., Shrestha, S., Hsu, S. D., Lin, Y. L., Lee, W. H., … Huang, H. D. (2016). miRTarBase 2016: updates to the experimentally validated miRNA-target interactions database. *Nucleic Acids Res. 44*(D1), D239–247.

Dillin, A., Hsu, A.-L., Arantes-Oliveira, N., Lehrer-Graiwer, J., Hsin, H., Fraser, A. G., … Kenyon, C. (2002). Rates of behavior and aging specified by mitochondrial function during development. *Science*, *298*(5602), 2398–2401.

Friedman, D. B. & Johnson, T. E. (1988). A mutation in the age-1 gene in Caenorhabditis elegans lengthens life and reduces hermaphrodite fertility. *Genetics*, *118*(1), 75–86.

Fuxman Bass, J. I., Pons, C., Kozlowski, L., Reece-Hoyes, J. S., Shrestha, S., Holdorf, A. D., … Walhout, A. J. (2016). A gene-centered C. elegans protein-DNA interaction network provides a framework for functional predictions. *Mol. Syst. Biol. 12*(10), 884.

Gavrilov, L. A. & Gavrilova, N. S. (2001). The reliability theory of aging and longevity. *Journal of theoretical Biology*, *213*(4), 527–545.

Gerson-Gurwitz, A., Wang, S., Sathe, S., Green, R., Yeo, G. W., Oegema, K., & Desai, A. (2016). A Small RNA-Catalytic Argonaute Pathway Tunes Germline Transcript Levels to Ensure Embryonic Divisions. *Cell*, *165*(2), 396–409.

Gottlieb, S. & Ruvkun, G. (1994). daf-2, daf-16 and daf-23: Genetically interacting genes controlling dauer formation in Caenorhabditis elegans. *Genetics*, *137*(1), 107–120.

Greer, E. L., Beese-Sims, S. E., Brookes, E., Spadafora, R., Zhu, Y., Rothbart, S. B., … Shi, Y. (2014). A histone methylation network regulates transgenerational epigenetic memory in C. elegans. *Cell Rep*, *7*(1), 113–126.

Gui, J. & Li, H. (2005). Penalized cox regression analysis in the high-dimensional and low-sample size settings, with applications to microarray gene expression data. *Bioinformatics*, *21*(13), 3001–3008.

Hamilton, B., Dong, Y., Shindo, M., Liu, W., Odell, I., Ruvkun, G., & Lee, S. S. (2005). A systematic RNAi screen for longevity genes in C. elegans. *Genes & development*, *19*(13), 1544–1555.

Hannum, G., Guinney, J., Zhao, L., Zhang, L., Hughes, G., Sadda, S., … Gao, Y. et al. (2013). Genome-wide methylation profiles reveal quantitative views of human aging rates. *Molecular Cell*, *49*(2), 359–367.

Harrison, D. E., Strong, R., Sharp, Z. D., Nelson, J. F., Astle, C. M., Flurkey, K., … Carter, C. S. et al. (2009). Rapamycin fed late in life extends lifespan in genetically heterogeneous mice. *Nature*, *460*(7253), 392–395.

Henderson, S. T. & Johnson, T. E. (2001). daf-16 integrates developmental and environmental inputs to mediate aging in the nematode Caenorhabditis elegans. *Curr. Biol. 11*(24), 1975–1980.

Horvath, S. (2013). DNA methylation age of human tissues and cell types. *Genome Biology*, *14*(10), R115.

Howe, K. L., Bolt, B. J., Cain, S., Chan, J., Chen, W. J., Davis, P., … Sternberg, P. W. (2016). WormBase 2016: expanding to enable helminth genomic research. *Nucleic Acids Res. 44*(D1), D774–780.

Hsin, H. & Kenyon, C. (1999). Signals from the reproductive system regulate the lifespan of C. elegans. *Nature*, *399*(6734), 362–366.

Inukai, S. & Slack, F. (2013). MicroRNAs and the genetic network in aging. *J. Mol. Biol. 425*(19), 3601–3608.

Jan, C. H., Friedman, R. C., Ruby, J. G., & Bartel, D. P. (2011). Formation, regulation and evolution of Caenorhabditis elegans 3'UTRs. *Nature*, *469*(7328), 97–101.

Johnson, W. E., Li, C., & Rabinovic, A. (2007). Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics*, *8*(1), 118–127.

Johnstone, I. M. & Lu, A. Y. (2009). On consistency and sparsity for principal components analysis in high dimensions. *Journal of the American Statistical Association*, *104*(486), 682–693.

Kato, M., Kashem, M. A., & Cheng, C. (2016). An intestinal microRNA modulates the homeostatic adaptation to chronic oxidative stress in C. elegans. *Aging (Albany NY)*, *8*(9), 1979–2005.

Kenyon, C., Chang, J., Gensch, E., Rudner, A., & Tabtiang, R. (1993). A C. elegans mutant that lives twice as long as wild type. *Nature*, *366*(6454), 461–464.

Klass, M. R. (1983). A method for the isolation of longevity mutants in the nematode Caenorhabditis elegans and initial results. *Mechanisms of Ageing and Development*, *22*(3), 279–286.

Kogan, V., Molodtsov, I., Menshikov, L. I., Shmookler Reis, R. J., & Fedichev, P. (2015). Stability analysis of a model gene network links aging, stress resistance, and negligible senescence. *Scientific Reports*, *5*, 13589.

Kudron, M., Niu, W., Lu, Z., Wang, G., Gerstein, M., Snyder, M., & Reinke, V. (2013). Tissue-specific direct targets

of Caenorhabditis elegans Rb/E2F dictate distinct somatic and germline programs. *Genome Biology*, *14*(1), R5.

Kumar, N., Jain, V., Singh, A., Jagtap, U., Verma, S., & Mukhopadhyay, A. (2015). Genome-wide endogenous DAF-16/FOXO recruitment dynamics during lowered insulin signalling in C. elegans. *Oncotarget*, *6*(39), 41418–41433.

Lê Cao, K.-A., Rohart, F., McHugh, L., Korn, O., & Wells, C. A. (2014). YuGene: A simple approach to scale gene expression data derived from different platforms for integrated analyses. *Genomics*, *103*(4), 239–251.

Lee, S.-J., Hwang, A. B., & Kenyon, C. (2010). Inhibition of respiration extends C. elegans life span via reactive oxygen species that increase HIF-1 activity. *Current Biology*, *20*(23), 2131–2136.

Li, R., Zhong, W., & Zhu, L. (2012). Feature screening via distance correlation learning. *Journal of the American Statistical Association*, *107*(499), 1129–1139.

Maegawa, S., Hinkal, G., Kim, H. S., Shen, L., Zhang, L., Zhang, J., ... Issa, J.-P. J. (2010). Widespread and tissue specific age-related DNA methylation changes in mice. *Genome Research*, *20*(3), 332–340.

Mann, F. G., Van Nostrand, E. L., Friedland, A. E., Liu, X., & Kim, S. K. (2016). Deactivation of the GATA Transcription Factor ELT-2 Is a Major Driver of Normal Aging in C. elegans. *PLoS Genet. 12*(4), e1005956.

Marioni, R. E., Shah, S., McRae, A. F., Chen, B. H., Colicino, E., Harris, S. E., ... Cox, S. R. et al. (2015). DNA methylation age of blood predicts all-cause mortality in later life. *Genome Biology*, *16*(1), 25.

Miller, R. A., Harrison, D. E., Astle, C. M., Baur, J. A., Boyd, A. R., De Cabo, R., ... Nelson, J. F. et al. (2011). Rapamycin, but not resveratrol or simvastatin, extends life span of genetically heterogeneous mice. *The Journals of Gerontology: Series A*, *66*(2), 191–201.

Moskalev, A. & Shaposhnikov, M. (2010). Pharmacological inhibition of phosphoinositide 3 and tor kinases improves survival of Drosophila melanogaster. *Rejuvenation Research*, *13*(2-3), 246–247.

Moskalev, A. & Shaposhnikov, M. (2011). Pharmacological inhibition of NF-κB prolongs lifespan of Drosophila melanogaster. *Aging (Albany NY)*, *3*(4), 391.

Niu, W. [W.], Lu, Z. J., Zhong, M., Sarov, M., Murray, J. I., Brdlik, C. M., ... Reinke, V. (2011). Diverse transcription factor binding features revealed by genome-wide ChIP-seq in C. elegans. *Genome Res. 21*(2), 245–254.

Niu, W. [Wei], Lu, Z. J., Zhong, M., Sarov, M., Murray, J. I., Brdlik, C. M., ... Preston, E. et al. (2011). Diverse transcription factor binding features revealed by genome-wide chip-seq in C. elegans. *Genome Research*, *21*(2), 245–254.

Phillips, C. M., Montgomery, B. E., Breen, P. C., Roovers, E. F., Rim, Y. S., Ohsumi, T. K., ... Montgomery, T. A. (2014). MUT-14 and SMUT-1 DEAD box RNA helicases have overlapping roles in germline RNAi and endogenous siRNA formation. *Curr. Biol. 24*(8), 839–844.

Podolskiy, D., Molodtcov, I., Zenin, A., Kogan, V., Menshikov, L., Shmookler Reis, R. J., & Fedichev, P. (2015). Critical dynamics of gene networks is a mechanism behind ageing and Gompertz law. *arXiv preprint arXiv:1502.04307*.

Pyrkov, T. V., Getmantsev, E., Zhurov, B., Avchaciov, K., Pyatnitskiy, M., Menshikov, L., ... Fedichev, P. O. (2017). Quantitative characterization of biological age and frailty based on locomotor activity records. *bioRxiv*, 186569.

Rechavi, O., Houri-Ze'evi, L., Anava, S., Goh, W. S. S., Kerk, S. Y., Hannon, G. J., & Hobert, O. (2014). Starvation-induced transgenerational inheritance of small RNAs in C. elegans. *Cell*, *158*(2), 277–287.

Rohart, F., Eslami, A., Matigian, N., Bougeard, S., & Le Cao, K. A. (2017). MINT: a multivariate integrative method to identify reproducible molecular signatures across independent experiments and platforms. *BMC Bioinformatics*, *18*(1), 128.

Rose, M. R., Flatt, T., Graves, J. L., Greer, L. F., Martinez, D. E., Matos, M., ... Shahrestani, P. (2012). What is aging? *Frontiers in Genetics*, *3*.

Seah, N. E., de Magalhaes Filho, C. D., Petrashen, A. P., Henderson, H. R., Laguer, J., Gonzalez, J., ... Lapierre, L. R. (2016). Autophagy-mediated longevity is modulated by lipoprotein biogenesis. *Autophagy*, *12*(2), 261–272.

Shaw, W. M., Luo, S., Landis, J., Ashraf, J., & Murphy, C. T. (2007). The C. elegans TGF-β dauer pathway regulates longevity via insulin signaling. *Current Biology*, *17*(19), 1635–1645.

Shmookler Reis, R. J., Xu, L., Lee, H., Chae, M., Thaden, J. J., Bharill, P., ... Zimniak, P. et al. (2011). Modulation of lipid biosynthesis contributes to stress resistance and longevity of C. elegans mutants. *Aging (Albany NY)*, *3*(2), 125–147.

Singhal, R. P., Mays-Hoopes, L. L., & Eichhorn, G. L. (1987). DNA methylation in aging of mice. *Mechanisms of Ageing and Development*, *41*(3), 199–210.

Strehler, B. L. & Mildvan, A. S. (1960). General theory of mortality and aging. a stochastic model relates observations on aging, physiologic decline, mortality, and radiation. *Science See Saiensu*, *132*.

Stroustrup, N., Anthony, W. E., Nash, Z. M., Gowda, V., Gomez, A., López-Moyado, I. F., ... Fontana, W. (2016). The temporal scaling of Caenorhabditis elegans ageing. *Nature*.

Tepper, R. G., Murphy, C. T., & Bussemaker, H. J. (2014). DAF-16 and PQM-1: partners in longevity. *Aging (Albany NY)*, *6*(1), 5–6.

Troemel, E. R., Chu, S. W., Reinke, V., Lee, S. S., Ausubel, F. M., & Kim, D. H. (2006). p38 MAPK regulates expression of immune response genes and contributes to longevity in C. elegans. *PLoS Genet. 2*(11), e183.

Vellai, T., Takacs-Vellai, K., Zhang, Y., Kovacs, A. L., Orosz, L., & Müller, F. (2003). Influence of TOR kinase on lifespan in C. elegans. *Nature*, *426*, 620.

Weitz, J. S. & Fraser, H. B. (2001). Explaining mortality rate plateaus. *Proc. Natl. Acad. Sci. U. S. A. 98*(26), 15383–6. doi:10.1073/pnas.261228098

Yang, W., Dierking, K., & Schulenburg, H. (2016). WormExp: a web-based application for a Caenorhabditis elegans-specific gene expression enrichment analysis. *Bioinformatics*, *32*(6), 943–945.

Ye, X., Linton, J. M., Schork, N. J., Buck, L. B., & Petrascheck, M. (2014). A pharmacological network for lifespan extension in Caenorhabditis elegans. *Aging Cell*, *13*(2), 206–215.

Youngman, M. J., Rogers, Z. N., & Kim, D. H. (2011). A decline in p38 MAPK signaling underlies immunosenescence in Caenorhabditis elegans. *PLoS Genet. 7*(5), e1002082.

Supplementary Materials

S1 Strains

The following *C. elegans* strains were used in this study: wild-type strain Bristol-N2, subline DRM (herein called "N2" or "N2-DRM"); SR806 [*daf-2(e1370)*]; DR1694 [*daf-2(e1391);daf-12(m20)*], and SR808 [*age-1(mg44)*] at the first ("F1") and second ("F2") generations of homozygosity. Strains SR806-SR808 were outcrossed 6 generations into N2-DRM; please see (Ayyadevara et al. 2008) for details. The above mutant strains, and N2-DRM, were grown in 35-mm Petri dishes, on the surface of NGM-agar (1% Bacto-Peptone, 2% agar in nematode growth medium) spotted with *E.coli* OP50 (a uracil-requiring mutant). Several RNAi treatments of genes (*daf-4*, *che-3*, *cyc-1*, *cco-1*, *eat-4*), mutation or RNAi of which were reported to prolong life-span, were also assessed. The animals were maintained on NGM-agar plates at 20°C, seeded with *E.coli* HT115 expressing double-stranded RNAs for target-gene knockdowns (Kamath et al. 2003) for both RNA-preparation and lifespan studies.

S2 Survivals

Lifespan assays were conducted at 20°C, as described previously (Ayyadevara et al. 2008). Briefly, synchronous cultures were initiated by lysis of gravid hermaphrodites in alkaline hypochlorite. Worms were selected at the L4 larval stage, placed 50 worms per plate, and transferred at 1- to 2-day intervals onto fresh plates during days 1-7, and at 2- to 3-day intervals thereafter. A worm was scored as dead if it failed to move, either spontaneously or in response to a mechanical stimulus; lost worms were excluded (censored) from the survival analysis.

Our survival study confirms longevity of the worm strains subjected to the treatments targeting genes known to affect aging in previous studies. The relative lifespan modification effects in some cases proved to be somewhat smaller, which can probably be attributed to the use of mutation instead of RNAi or a different developmental temperature in the original studies (see Table. S1 for comparison).

S3 RNA isolation

Synchronized strains of *C. elegans* were grown on 100-mm NGM plates, as above, and harvested for RNA extraction at the ages indicated. Worms were washed off plates and rinsed twice in survival buffer; after 30 min at 20°C (to allow digestion of enteral bacteria), they were flash frozen and stored at −80°C. Frozen worms were ground in a dry-ice-cooled mortar and pestle, and total RNA was extracted using RNeasy RNA extraction kits (Qiagen), followed by RNA purification for construction of transcript libraries using TruSeq RNA kits (Illumina, v.2). Sequences are generated as PE100 multiplexes, 100-bp paired-end reads from

Table S1
Summary of survivals for the mutant strains and the confirmatory RNAi interventions.

Treatment	Dev. Temp. (°C),	LS, control*(d)	LS, treatment*(d)	Effect on LS	Reference
daf-4 RNAi	20	18.5	25	+35%	in-house
daf-4(e1364)	20	15.8	22.1	+40%[b]	(Shaw et al. 2007)
daf-4(e1364)	20	14.2	31	+120%[b,**]	(Shaw et al. 2007)
che-3 RNAi	20	18.5	24.5	+32%	in-house
che-3(p801)	20	18.8	37.5	+100%	(Apfeld et al. 1999)
che-3(e1124)	20	18.8	25.7	+37%	(Apfeld et al. 1999)
cyc-1 RNAi	20	18.5	22	+19%	in-house
	25	13.6	25.4	+87%	(Dillin et al. 2002)
	20	19.6	31.3	+60%[c]	(Lee et al. 2010)
	20	16.8	31.9	+90%[c]	(Lee et al. 2010)
cco-1 RNAi	20	18.5	24	+30%	in-house
	25	15.2	24.5	+61%	(Dillin et al. 2002)
	20	19.6	30.8	+57%[d]	(Lee et al. 2010)
	20	16.8	30.3	+80%[d]	(Lee et al. 2010)
eat-4 RNAi	20	18.5	21	+14%	in-house
	25			prolonging[***]	(Hamilton et al. 2005)

[a]The survival experiment for the RNA interference of the five selected *C. elegans* target genes. Three successful RNAi interventions (in bold) were selected for further transcriptomic measurements.
[b,c,d]Both experiments were done in parallel.
Abbreviations: Dev.Temp., development temperature; LS, mean adult lifespan.
*Mean adult lifespan, days post-L4/adult molt.
**Quote from (Shaw et al. 2007): "*daf-4(e1364)* mutants lived more than twice as long as the wild-type in one trial."
***In (Hamilton et al. 2005), the quantity of effect is not stated, but it is life-prolonging.

an Illumina HiSeq2500 or NextSeq instrument, producing $40 - 50 \times 10^6$ reads per sample. Paired samples are analyzed with DESeq2 (v1.4.5), and combined sequences are mapped to the *C. elegans* genome using TopHat (Kim et al. 2013).

S4 Experimental RNA-seq dataset

RNA-seq reads were mapped to the *C. elegans* genome (WBcel235, Ensembl annotation) using TopHat 2.1.1 (with --b2-very-sensitive and --GTF options) (ibid.) and gene-level read counts were obtained using the htseq-count software (Anders et al. 2015). Low-expressed genes with at least one zero read count per sample were removed from subsequent analysis. Raw read counts were normalized using the upper quartile method (Bullard et al. 2010) and converted to RPKM values using the edgeR library (Robinson et al. 2010).

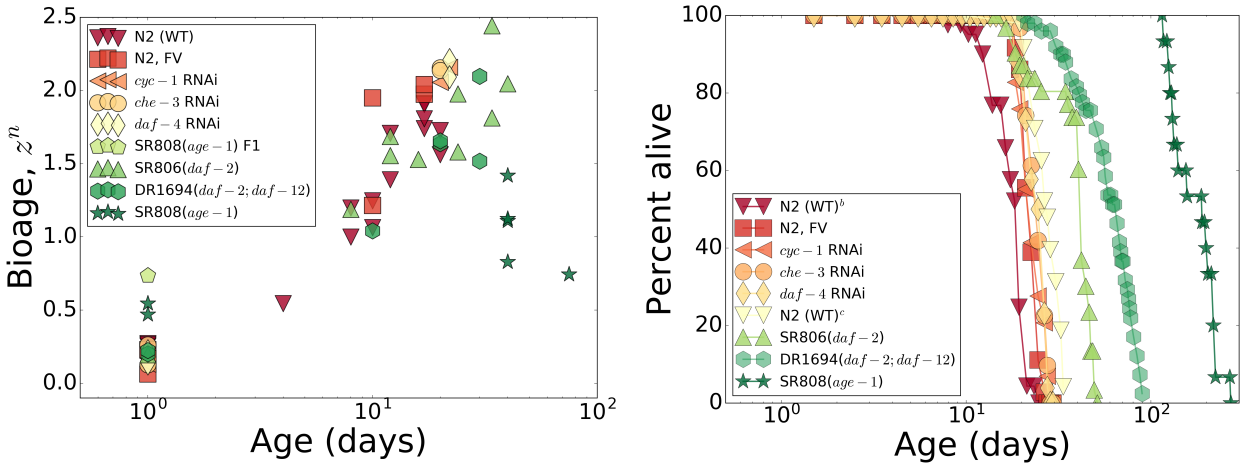
S5 MetaWorm transcriptome

The "MetaWorm" dataset was compiled from all publicly available information on expression profiles *C. elegans* from GEO database (Barrett et al. 2013) and annotated with the corresponding worms' ages. For individual genes represented by multiple probesets, the probeset with the largest signal was used. Gene expression in all datasets were normalized using the YuGene (Lê Cao et al. 2014) algorithm, which was specifically developed for gene expression comparisons between different platforms. The final "MetaWorm" dataset represents a 3724×4861 matrix (samples-x-genes) and incorporates more than 400 transcriptomic experiments (see Electronic Supplementary Materials).

S6 Gene regulatory networks at criticality

We focus on transcriptomic data and describe time-evolution of gene expression by a matrix x_i^n, where indices $n = 1 \dots N$ and $i = 1 \dots G$ enumerate samples and gene transcripts, respectively, G is the total number of genes and N is the total number of samples. The measurements are taken at successive instances of time/age, t^n. Following (Podolskiy et al. 2015) we describe the dynamics of a transcriptome by a differential equation, representing the gene-expression kinetics: $dx_i(t)/dt = f(x_i, t)$, where all important properties of an organism are encapsulated into the function $f(x_i, t)$. The equations are still complex and for small deviations from the initial state we use a linearized version, $\dot{x}_i = f_0 + \sum_i K_{ij} x_j$. Here $K_{ij} = df_i/dx_j$ is the matrix, representing the simplest form of interaction in the model.

The matrix consists of kinetic rates corresponding to time scale of elementary biological processes, which are normally very high compared to characteristic rates associated with



(a) (b)

Figure S1. The universal temporal scaling of survival curves and transcriptomic bioage with no rescaling of age on the average lifespan. (a) The transcriptomic bioage z^n (calculated from the transcriptomic signature obtained independently via training on the MetaWorm dataset) as a function of the chronological age for the in-house collection of gene-expression data for four long-lived mutants and *C. elegans* wild-type (Bristol-N2, stock DRM), and three groups of N2 wild-type worms treated with life-prolonging RNAi or bacteria carrying only the empty feeding vector (FV) without an RNAi insert. (b) The corresponding survival curves from (a) as a function of the chronological age. In both panels, the scale of x-axis is logarithmic, and all markers are colored according to the lifespans of the strains (see Table 1 for the lifespans): red for small and green for large lifespans. Overall, the lifespan ranges from 17 to 160 days.

aging. The lifelong dynamics associated with aging is only possible whenever the interaction matrix K_{ij} possesses a very small eigenvalue, α. Under the circumstances, the dynamics of the biological variables is said to be critical and can be accurately described with the help of the following one-factor model

$$x_i^n = \bar{x}_i + b_i z^n + \xi_i^n, \tag{1}$$

where $\bar{x}_i$ is the system state, corresponding to the animals early in life, b_i is the right eigenvector of the matrix K_{ij} corresponding to the singular eigenvalue α, and z^n is the order parameter, associated with the gene regulatory network instability, a stochastic variable, associated with aging, and following the Langevin equation:

$$\dot{z} = \alpha z + \eta,$$

where the stochastic variable η represents the effects of external and endogenous stress factors. In (Podolskiy et al. 2015) we solved the equation and demonstrated, that α is nothing else, but the Gompertz exponent, characterizing exponential increase of chances to die as a function of chronological age. The parameter defines a natural time scale, associated with the gene regulatory network instability, and depends on the balance between the damage accumulation and repair systems efficacy (Kogan et al. 2015).

The mode variable z^n, in contrast to the chronological age t^n, is a universal parameter describing the progress of aging and thus can be used as biological age. Mutations or anti-aging therapies modify interactions in the gene regulatory network. Formally, this corresponds to alterations in the matrix elements K_{ij} and change in its most relevant eigenvalue, α. The mortality in the model can be related to the effects of non-linearities: at sufficiently large values of z the higher order corrections can no longer be neglected and, starting at about average lifespan, the dynamics of the order parametre accelerates and the gene regulatory network disintegrates. Since normally the average lifespan exceeds the mortality rate doubling time, the effects of the non-linearities and stochastic forces can be neglected, at least early in life. Therefore, we expect that aging trajectories corresponding to different lifespans are self-similar and different by a single time scale factor α.

Practically, it would be interesting to derive an estimate of the biological age from the gene expression data using a linear model

$$z^n \approx \sum_{i=1}^{G} a_i \left(x_i^n - \bar{x}_i\right). \tag{2}$$

Since, however, the aging dynamics in biological signals is restricted to a one-dimensional manifold, the definition of biological age is not unique. This difficulty is not only a consequence of limited number of samples in a typical experiment. Any vector, orthogonal to b_i can be added to the

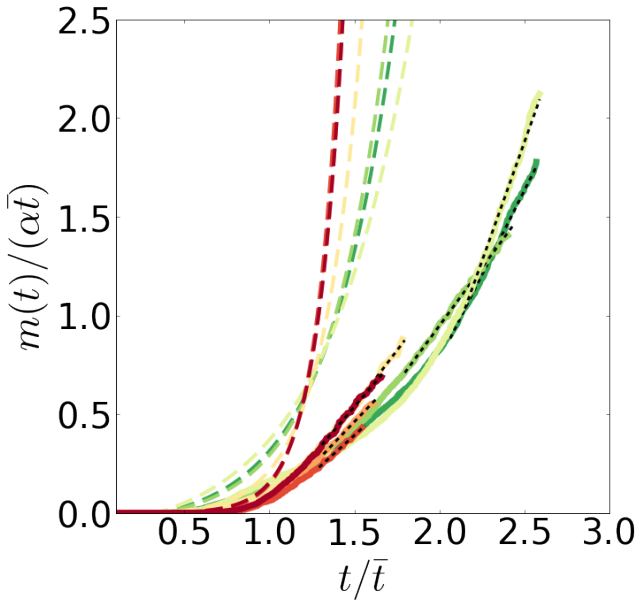


Figure S2. Evidence for the deceleration and plateauing of experimental mortality in *C. elegans*. The normalized cumulative hazard $m(t)/(\alpha\bar{t})$ calculated from the Gompertz equation (colored thick dashed) as derived from the experimental Kaplan-Meier plots (Stroustrup et al. 2016) (colored thick solid lines). The tail of the cumulative hazard (black thin dashed lines) was used for the calculation of the plateau mortality $M(t \gg \bar{t})$ by linear regression. The values of the Gompertz exponent α are indicated by the pseudocolor: red for large and green for small values.

projector a_i without changing the prediction results significantly, especially if the experimental noise (such as batch effects) is large. The best possible candidate for a_i is the left eigenvector of the matrix K_{ij} corresponding to the eigenvalue α.

S7 Mortality analysis

The discrepancy between the mortality behavior predicted by the Gompertz equation and experimental mortality for late ages in *C. elegans* is sufficiently large and thus can be used to test the aging theory predictions quantitatively with high-quality mortality data. Mortality data of appropriate quality were recently published in (Stroustrup et al. 2016), where a temporal scaling law of aging in *C. elegans* was observed, similar to that inferred for *D. melanogaster* (Helfand et al. 1995). This scaling law states that under the influence of some intervention, survival curves are stretched along the age axis by a dimensionless factor.

To extract the Gompertz exponent α from the mortality data (Stroustrup et al. 2016), we used the corresponding survival curves and fitted them to the prediction of the Gompertz equation. The procedure is only sensitive to the behavior of the survival curves in the neighborhood of the average lifespan. This is fortunate, since a gene regulatory network's stiffness parameter α coincides with the Gompertz exponent in this interval only. The value of the plateau mortality $M(t \gg \bar{t})$ was then calculated from the tail of the cumulative hazard $m(t)$, estimated from the raw mortality data by the well-defined Nelson-Aalen routine (Borgan 2005) with the help of the *Lifelines* package (Davidson-Pilon 2016).

Since the mortality rate reaches a plateau at late ages, the behavior of the cumulative hazard for these ages is linear and the value of $M(t \gg \bar{t})$ can be extracted by linear regression of the cumulative hazard on age. We calculated the cumulative hazard $m(t)$ from the experimental data and as the prediction of the Gompertz equation and compared them in Fig. S2, where the disagreement between the two is substantial and significant, both qualitatively (exponential growth for the Gompertz equation and linear growth for the plateau mortality) and quantitatively (the cumulative hazard for the Gompertz equation is several orders of magnitude larger for late ages).

References

Anders, S., Pyl, P. T., & Huber, W. (2015). HTSeq, a python framework to work with high-throughput sequencing data. *Bioinformatics*, *31*(2), 166–169.

Apfeld, J. & Kenyon, C. (1999). Regulation of lifespan by sensory perception in Caenorhabditis elegans. *Nature*, *402*(6763), 804–809.

Ayyadevara, S., Alla, R., Thaden, J. J., & Shmookler Reis, R. J. (2008). Remarkable longevity and stress resistance of nematode PI3K-null mutants. *Aging Cell*, *7*(1), 13–22.

Barrett, T., Wilhite, S. E., Ledoux, P., Evangelista, C., Kim, I. F., Tomashevsky, M., … Soboleva, A. (2013). NCBI GEO: archive for functional genomics data sets–update. *Nucleic Acids Res. 41*(Database issue), D991–995.

Borgan, Ø. (2005). Nelson–Aalen estimator. *Encyclopedia of Biostatistics*.

Bullard, J. H., Purdom, E., Hansen, K. D., & Dudoit, S. (2010). Evaluation of statistical methods for normalization and differential expression in mRNA-Seq experiments. *BMC Bioinformatics*, *11*, 94.

Davidson-Pilon, C. (2016). Lifelines package. https://github.com/camdavidsonpilon/lifelines. GitHub.

Dillin, A., Hsu, A.-L., Arantes-Oliveira, N., Lehrer-Graiwer, J., Hsin, H., Fraser, A. G., … Kenyon, C. (2002). Rates of behavior and aging specified by mitochondrial function during development. *Science*, *298*(5602), 2398–2401.

Hamilton, B., Dong, Y., Shindo, M., Liu, W., Odell, I., Ruvkun, G., & Lee, S. S. (2005). A systematic RNAi screen for longevity genes in C. elegans. *Genes & development*, *19*(13), 1544–1555.

Helfand, S. L., Blake, K. J., Rogina, B., Stracks, M. D., Centurion, A., & Naprta, B. (1995). Temporal patterns of gene expression in the antenna of the adult Drosophila melanogaster. *Genetics*, *140*(2), 549–555.

Kamath, R. S. & Ahringer, J. (2003). Genome-wide RNAi screening in Caenorhabditis elegans. *Methods*, *30*(4), 313–321.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., & Salzberg, S. L. (2013). TopHat2: Accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol. 14*(4), R36.

Kogan, V., Molodtsov, I., Menshikov, L. I., Shmookler Reis, R. J., & Fedichev, P. (2015). Stability analysis of a model gene network links aging, stress resistance, and negligible senescence. *Scientific Reports*, *5*, 13589.

Lê Cao, K.-A., Rohart, F., McHugh, L., Korn, O., & Wells, C. A. (2014). YuGene: A simple approach to scale gene expression data derived from different platforms for integrated analyses. *Genomics*, *103*(4), 239–251.

Lee, S.-J., Hwang, A. B., & Kenyon, C. (2010). Inhibition of respiration extends C. elegans life span via reactive oxygen species that increase HIF-1 activity. *Current Biology*, *20*(23), 2131–2136.

Podolskiy, D., Molodtcov, I., Zenin, A., Kogan, V., Menshikov, L., Shmookler Reis, R. J., & Fedichev, P. (2015). Critical dynamics of gene networks is a mechanism behind ageing and Gompertz law. *arXiv preprint arXiv:1502.04307*.

Robinson, M. D., McCarthy, D. J., & Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, *26*(1), 139–140.

Shaw, W. M., Luo, S., Landis, J., Ashraf, J., & Murphy, C. T. (2007). The C. elegans TGF-β dauer pathway regulates longevity via insulin signaling. *Current Biology*, *17*(19), 1635–1645.

Stroustrup, N., Anthony, W. E., Nash, Z. M., Gowda, V., Gomez, A., López-Moyado, I. F., … Fontana, W. (2016). The temporal scaling of Caenorhabditis elegans ageing. *Nature*.