Analysis of different strains of the turquoise killifish identify transcriptomic signatures associated with heritable lifespan differences

Dr. Mariateresa Mazzetto[1,2], Dr. Kathrin Reichwald[3], Dr. Philipp Koch[3], Dr. Marco Groth[3] and Prof. Alessandro Cellerino[1,3*]

[1]BIO@SNS, Scuola Normale Superiore, Piazza dei Cavalieri,7, Pisa, 56126, Italy.
[2] current Yale School of Medicine, 333 Cedar Street New Haven, CT 06510, USA
[3]Leibniz Institue on Aging, Fritz Lipmann Institute, Beutenbergstr. 11, Jena, D-07745, Germany.

*Corresponding Author: alessandro.cellerino@sns.it

Abstract

The African turquoise killifish *Nothobranchius furzeri* represents an emerging short-lived model for aging research. Captive strains of this species are characterized by large differences in lifespan. To identify the gene expression correlates of this lifespan differences, we analyzed a public transcriptomic dataset consisting of four different tissues in addition to embryos. We focused on the GRZ and the MZM0410 captive strains, which show a near twofold difference in lifespan, but similar growth and maturation and validated the results in a newly-generated dataset from a third longer-lived strain. The two strains show distinct transcriptome expression patterns already as embryos and the genotype has a larger effect than age on gene expression, both in terms of number of differentially expressed genes and magnitude of regulation. Network analysis detected RNA processing and histone modifications as the most prominent categories upregulated in GRZ that also showed idiosyncratic expression patterns such as high expression of DND is somatic tissues. The short-lived GRZ strain shows transcriptional aging signatures already at sexual maturity (anticipated aging) in all four tissues suggesting that short lifespan is the results of events that occur early in life rather than the progressive accumulation of strain-dependent differences. The GRZ strain is the most commonly used *N. furzeri* strain in intervention studies and our results warrant replication of at least key intervention studies in longer-lived strains.

Keywords

Animal model, killifish, systems biology, genetics, aging

1 | Introduction

Age-dependent transcript regulation has been investigated extensively in several models and in humans, yet the impact of genetic variability on age-dependent gene regulation is yet to be completely understood [1]. Differences in lifespan between individuals of the same species can be influenced by intrinsic- (random variations) and extrinsic-factors (such as nutrition) [2,3,4]. Disentangling the effects of genetic heterogeneity and stochasticity on lifespan of mammals is technically challenging and the majority of studies have been performed using invertebrates, such as the roundworm, *C. elegans* [4,5,3].

The annual African turquoise killifish (*Nothobranchius furzeri*) was recently established as a short-lived vertebrate model for aging research [6]. *N. furzeri* is particularly suited to investigate the effects of interventions on lifespan and aging-associated dysfunctions. Its maximum lifespan is 3-12 months, depending on the strain [7], the shortest lifespan ever recorded in captivity for a vertebrate [8]. Along with its short lifespan, the Turquoise killifish shows an accelerated aging phenotype, which shares many molecular and behavioral aging hallmarks that have been described in mammals, like accumulation of lipofuscin, gliosis, shortening of telomeres and cognitive decline [9].

A specific feature of this model is the large influence of the genetic background on lifespan. *N. furzeri* is indeed characterized by the presence of multiple laboratory strains whose founders originate from different spot habitats in Mozambique and Zimbabwe with large differences in captive lifespan [7] that can be as short as 3 months in the GRZ inbred strain [8] while recently wild-derived strains have lifespans in the order of at least 6-8 months and do not replicate GRZ extreme phenotype [10,11,12]. Interestingly, these strains do not differ in growth and maturation time, indicating that analysis of different strains can disentangle the covariation of lifespan with age at maturity. QTL mapping studies revealed a complex genetic architecture underlying this lifespan difference [13,14] that is therefore not the result of the fixation of a single deleterious mutation and must arise from a combination of naturally-occurring alleles fixed in the GRZ strain.

RNA sequencing (RNA-seq) studies have demonstrated that the patterns of genome-wide transcript regulation in the *N. furzeri* mimic those observed in mammals [15,16] and this species

was used in longitudinal experiment to investigate the correlation between individual lifespan and individual patterns of gene expression early in adult life [17,18].

In this study, we exploited *N. furzeri* public transcriptomic datasets (Supplementary Table 1) in order to identify gene expression signatures of genetically-determined lifespan differences and to compare these with patterns of age-dependent gene expression. We analyzed the impact of age and genotype on gene expression across different tissues. We then compared gene expression profiles of aging biomarkers in two strains and performed network analysis using weighted gene co-expression network analysis (WGCNA) [19], in order to correlate gene modules with external variables (strain in this case) and identified central hubs in the co-expression network of genes associated with heritable lifespan differences.

2 | Methods

2.1 | RNA sequencing and data processing

In general, sequencing of RNA samples was performed using Illumina's next- generation sequencing methodology [20]. In detail, total RNA was quantified and quality checked using Agilent 2100 Bioanalyzer Instrument (RNA 6000 Nano assay). Libraries were prepared from 500 ng of total RNA using TruSeq RNA v2 library preparation kit (Illumina) according the manufacturer's instructions and subsequently quantified and quality checked using Agilent 2100 Bioanalyzer Instrument (DNA 7500 assay). Libraries were sequenced using a HiSeq 2500 System running in 51 cycle/single-end/high output mode. Sequence information was converted to FASTQ format using bcl2fastq v1.8.4. The RNA sequencing reads were aligned to the *N. furzeri* genome assembly (assembly version Nfu 20150522, gene annotation version 150922, both downloaded from the *Nothobranchius furzeri* Information Network Genome Browser https://nfingb.leibniz-fli.de/) using STAR 2.5.1b [21] (parameters: –alignIntronMax 100000, – outSJfilterReads Unique). For each gene, all reads that map uniquely to one genomic position were counted with FeatureCounts 1.5.0 under default settings [22].

2.2 | Data normalization

Data were normalized with DESeq2 [24] using either the conversion to pseudocounts (for plot counts, which normalizes counts by the estimated size factors and adds a pseudocount of ½ to allow for log scale plotting), the variance stabilizing transformation (VST) method (for the Principal Component Analysis), or the regularized algorithm or *rlog* transformation (to visualize the expression profile of aging biomarkers); in the latter case normalized data were then converted to z-scores, i.e. centered to the mean and scaled on the standard deviation of all samples, before visualization with the ggplot2 package [25].

2.3. | Principal Component Analysis (PCA)

Differentially expressed genes for both the age and strain variables were selected for the analysis. Principal Components of the samples were obtained through the prcomp() function and then plotted with ggplot2; centroids, as means of the replicates for each group of samples, were also measured and included in the visualization.

2.4. | Differential expression analysis

Differentially expressed transcripts/miRNAs were obtained using the DESeq2 package using the default pipeline [24]: differential expression results were filtered for padj<0.05 and then used as input for visualization.

2.5. | Enrichment analysis

Overrepresentation analysis of Gene Ontology terms was performed for each tissue with the WebGestalt online tool using GO Process; enriched categories for each quadrant were filtered for FDR<0.05. To combine the results meta analysis was performed using the Fisher's Method, which uses this formula:

$$X_{2k}^2 = -2\sum_{i=1}^{k} \ln p_i,$$

and then p-values were adjusted using the p.adjust() function. Generally Applicable Gene-set/Pathway enrichment (GAGE) was performed using the gage package [26]; enriched categories were filtered for q-value<0.05 and used for visualization with the Revigo software. Entrez IDs

were used as identifiers for the analysis, from the human orthologues gene symbols: these were obtained with the biomart() package [27,28].

2.6 | Network analysis

Weighted gene correlation network analysis (WGCNA) analysis was performed using the WGCNA package in R [29] using the unsigned correlation as option and setting the minimum cluster size to 30 members. Input for WGCNA was a list of 4247 genes (with counts>100 for each sample): we obtained consensus modules across tissues using the default pipeline. The selected module was then visualized using Cytoscape filtering the network on the weight of connection (values>0.08).

2.7 | Cox-hazard model

Methods used for the collection and processing of the longitudinal dataset used for the analysis are reported in Baumgart et al. [17]; the method used to normalize the data and measure the cox-hazard coefficient is explained in Kelmer Sacramento, Mazzetto, Kirkpatrick et al. [18].

2.8 | External datasets and gene lists

We used the following gene lists for comparison with our data:

- genes with either negative or positive sign log(β) obtained with the Cox-Hazard model were taken from Kelmer Sacramento, Mazzetto, Kirkpatrick et al. [18];
- Aging biomarkers and Rotenone-induced DEGs were taken by Baumgart et al. [17];
- *APRT* mutant DEGs were taken by Astre et al. [40];
- *DND-1* mutant DEGs were taken by Moses et al. [41].

We also used the following external datasets:

- Transcriptomic data from *H. Glaber* (liver) by Sahm et al. [30] for the comparison between the killifish and naked mole rat transcriptomes;

2.9 | Analysis with Connectivity Map

The analysis was performed with the online ‘‘query’’ tool in https://clue.io/query [31]. The query was run against the dataset released on (Dec 17, 2020). In each query, the 150 genes with the highest fold change according to strain or age were used as input.

2.9 | Data and code availability

The accession numbers for the data reported in this paper are: GSE52462 (MZM aging for brain, liver and skin), GSE103816 (MZM aging, muscle), PRJEB5837 (diapause and non-diapause, GRZ and MZM, embryo), GSE103132 (GRZ aging, brain), GSE103137 (GRZ aging, liver), GSE103140 (GRZ aging, skin), GSE103815 (GRZ aging, muscle), GSE92854 (small-RNA-seq, both MZM and GRZ for brain, liver and skin), GSE207748 (validation dataset), and GSE98744 (transcriptome for *H.glaber*, liver) and are reporter also in **Supplementary Table 1**.

Codes are available upon request from the corresponding author.

2.9 | Animal procedures and Ethical statement

Fish were bred at the Leibniz Institute on Aging, Fritz Lipmann Institute, Jena. Procedures for fish breeding, husbandry and euthanasia were performed in accordance with the rules of the German Animal Welfare Law and approved by the Landesamt für Verbraucherschutz Thüringen, Germany.

Eggs were maintained on wet peat moss at room temperature in sealed Petri dishes. When embryos had developed, eggs were hatched by flushing the peat with tap water at 16–18°C. Embryos were scooped with a cut plastic pipette and transferred to the system tank. Fry were fed with newly hatched Artemia *nauplii* for the first 2 weeks and then weaned with finely chopped Chironomus *larvae*.

The system water temperature was set at a constant 27°C. At the desired age, fish were sacrificed via anesthetic overdose (Tricaine, MS-222) in accordance with the prescription of the European (Directive 2010/63/UE) and the rules of the German Animal Welfare Law.

3 | Results

3.1 | Genotype has a large effect on gene expression in *Nothobranchius furzeri*

In order to identify the transcriptional signature associated with heritable lifespan differences, we analyzed 211 public RNA-seq libraries obtained from two strains differing in lifespan (MZM-0410, with a median lifespan of 40 weeks in our colony at the time these data were generated [32], and GRZ-D, with a median lifespan of 15 weeks in our colony at the time these data were generated [33]), four tissues (brain, liver skin, muscle) and 5 time points. The time points collected correspond to sexual maturity (5 weeks post hatching, or wph) and (for MZM) to young adult (12 wph), adult (20 wph), old (27 wph), and geriatric (39 wph), with the exception of muscle for which only young (9 wph) and old (27 wph) samples were available. For GRZ, the time points collected were 5 wph, 7 wph, 10 wph, 12 wph, and 14 wph, which correspond to similar biological age if median lifespan is taken as a reference. In addition, we analyzed embryos at mid somitogenesis stage in two conditions (diapause and direct development) (**Figure 1A**). To investigate the global similarities in the regulation of the coding and non-coding transcriptome in this analysis , we included150 libraries of miRNA-seq [34] obtained from a subset of the same adult sample.

Principal component analysis (PCA) was used to visualize the transcriptional similarities among samples and revealed a clear separation of the samples based on the "strain" variable. In particular, it is evident that samples of different strains belonging to the same tissue are separated in a diagonal direction from the lower right to the upper left (**Figure 1B**). Surprisingly, this separation started already at the embryonic stage (**Figure 1B**, inset). The same effect of genotype on global expression patterns could be observed also for miRNAs (**Figure 1C**). Differential expression analysis was performed with DESeq2 [24] using generalized linear models to separate the effects of different independent variables on gene expression. To identify gene sets that are differentially expressed in the two strains, we performed generally applicable gene set enrichment (GAGE) analysis [25] excluding the embryo samples. GAGE detected significant enrichment for a large number of gene sets; to enable visualization these were further grouped into semantically similar clusters using REVIGO [35] (**Figure 1D**). We observed an upregulation of categories related to development, RNA processing, protein-targeting, -transport or -modification, autophagy and membrane organization in the short-lived strain (GRZ-D). Genes upregulated in the long-lived strain were enriched in the DNA replication/cell cycle and the immune response/protein activation

cascade pathways. These results highlight profound and unexpected differences in gene expression between these two strains.

Some genes showed extreme differences in expression between the two strains, such as the genes dead-end (DND), which is normally expressed during early embryonic development of the primordial germs cells [36] and in adult gonads, and the transporter SLC22A13, whose expression is normally restricted to the kidney [37]. These genes show negligible expression in MZM04010, but are strongly upregulated in all the tissues collected from animals belonging to the GRZ strain (**Figure S1**). Examples of the opposite pattern are represented by ACO2, coding for aconitase 2, and MACROD2, a gene whose deletion causes chromosome instability and represents a cancer risk [38] (**Figure S1**). In order to validate these findings, we generated a new independent RNA-seq dataset that contains samples related to two different strains, GRZ and MZM-0403, a longer-lived strain genetically distinct from MZM04010 [7] (**Figure 1I**). The selected genes show a similar expression profile during aging, with DND and SLC22A13 showing increased expression in the GRZ strain, and with ACO2 and MACROD2 being more expressed in the MZM0403 strain (**Figure S3 A-D**).

3.2 | Network analysis identifies genes related to splicing and protein modification as hubs of gene modules related to genotype differences

One method to identify differential signatures among different backgrounds is the identification of biomarkers and gene hubs using correlation networks. These methods allow the identification of correlation clusters among genes based on their expression profiles, and to relate these modules to external variables (like age or sex). In this work, we applied the weighted gene co-expression network analysis (WGCNA) method, which groups genes (or transcripts or proteins) into discrete modules based on their co-expression patterns quantified as correlation coefficients between the profiles of two nodes (the nodes being the gene pairs)[29]. WGCNA also computes the eigengene for each module as a summary vector of the coherent regulation of genes within a module, which can be then related to an external trait (in this case, strain or age). We performed a variant of the WGCNA method that, after clustering the genes in each dataset (i.e. tissue) in modules, creates a consensus matrix among all datasets to identify modules (and hub genes) conserved (i.e. with similar co-expression similarity values) across all tissues (brain, liver, skin and muscle). We

analyzed 4247 genes with medium to high expression in all tissues (counts per sample > 100); we obtained 10 modules, and we computed the association with the "strain" variable for each module. We focused our attention on the "turquoise" module (**Figure 1E**), which had the highest statistical association with strain (computed from Fisher's meta analysis among each tissue correlation value). The module was enriched in mRNA processing, transport-related and histone modification categories (**Figure 1F**). In order to validate these results, we selected the genes with the highest intramodular connectivity and gene-trait significance, which are highlighted as orange points in **Figure 1E** and are related to splicing, (such as SNW1, SF1 and WTAP), and to maintenance of organelle structure (such as GOLGA5). The differential expression of SPECC1L (involved in the organization of the actin cytoskeleton) and PPP2R2D (a phosphatase involved in the control of mitosis entry and exit) was validated by qPCR in a third independent biological sample set i.e. brains from MZM-0410 and GRZ fish not used for the previous RNAseq (**Figure 1G-H**). Differential expression of hub genes was also validated by exploiting the independent RNA-seq dataset described above (**Figure 1J**).

3.3 | GRZ strain shows an anticipated aging expression profile

To investigate the relationships between the effects of age and strain on global gene expression, we analyzed the transcripts that were differentially expressed (FDR< 0.1) both during aging in MZM0410 (i.e. in the contrast 5 wph vs. 39 wph) and between MZM0410 and GRZ at the first time point analyzed (9 weeks for muscle, 5 weeks for the other three tissues) and plotted the log2(fold change) during aging on the X axis and the log2(fold change) GRZ vs. MZM0410 on the Y axis for each tissue separately. In all four tissues, the DEGs showed preferentially the same direction of regulation in the two conditions and the data points were therefore concentrated in quadrants I and III (i.e. genes down-regulated with age are also lower-expressed in GRZ and *vice versa*, Fisher's exact test: p-value < 2.2e-16 for all the 4 tissues) (**Figure 2A-2D).** Similar positive correlations between the effects of age and strain was obtained also when only differential expression according to age in MZM0410 was used as filtering criterion (**Figure S2A-S2D**).

In order to identify the biological processes that are differentially expressed according to strain and age, we performed GO over-representation analysis for the genes in the four quadrants of **Figure 2A-D**. Results of the three individual tissues were combined using Fisher's metanalysis

where the p-value of each GO represents a summary of the p-value for over-representation in the individual tests. Results are displayed in **Figure 2E**. Genes consistently down-regulated in the short-lived strain and during aging were enriched in terms related to cell cycle, DNA replication, cell-proliferation and -differentiation, indicating that age-dependent decrease in mitotic activity is anticipated in GRZ. Genes up-regulated in GRZ and during aging are enriched in terms related to autophagy, protein localization and I-kB/NF-kB signaling. Interestingly, genes up-regulated during aging and down-regulated in GRZ show a highly significant enrichment for the terms ncRNA processing and defense response. Genes down-regulated during aging and up-regulated in GRZ are enriched in terms related to cell cycle. The same trend of positive correlation was observed also when differentially expressed miRNAs in the same conditions were investigated, showing that this phenomenon is conserved across different types of transcripts (**Figure 2F-H**).

In order to validate these findings, we exploited the independent RNA-seq dataset described above and repeated the analysis shown in **Figure 2A-D**. In this case, we analyzed the transcripts that were detected as differentially expressed both during aging in MZCS0403 (comparing 287 days-old animals to 21 days-old animals) and between MZCS0403 and GRZ at the first time point analyzed (21 dph/3 wph) and plotted the log2(fold change) for aging on the X axis and for strain on the Y axis (**Figure S3E**). In both the comparisons, the DEGs showed preferentially the same direction of regulation in the two conditions (as observed for the GRZ vs. MZM0410 comparison) and the data points were therefore concentrated in quadrants I and III, with a strong positive correlation (r= 0.5127 for GRZ vs. MZCS0403; r = 0.604 for GRZ vs. MZCS0403)**.** The genes down-regulated in the GRZ strain and in old animals (for both comparisons) were enriched in categories related to genome maintenance and development (reproducing the results shown in **Figure 2**), while genes up-regulated for both the variables were enriched in terms different from those identified in the previous comparison.

3.4. Cox-Hazard model reveals genes with antagonistic effects on lifespan that show a strain/age-dependent expression profile

To assess whether age- and strain-dependent transcript regulation are correlated with individual lifespan variations, we re-analyzed a longitudinal RNA-seq dataset, where transcripts from fin biopsies of male MZM0410 fish were quantified at 10 and 20 wph, and individual lifespan

information is available for each subject [17] (**Figure 3A**). We then applied the Cox-proportional-hazard model to this dataset. Cox analysis describes the association of an independent variable (i.e. transcript levels in our case) and mortality risk by an exponential coefficient β. Negative values of log(β) indicate that the higher the gene expression, the lower the mortality risk; positive values of log(β) indicate that the higher the gene expression, the higher the mortality risk. In the scatterplot in **Figure 3B**, we separately report the log(β) of individual genes at 10 wph and 20 wph. The majority of significant cases are concentrated in quadrant II and IV, with a log(β) of opposite signs at the two tested ages. Since factors with positive sign log(β) are risk factors and factors with negative sign log(β) are protective factors, this result implies that for the majority of significant cases expression levels at the two ages have an antagonistic action: detrimental at early time points and beneficial at later time points (quadrant II), or the opposite (quadrant IV).

We also analyzed the regulation (as 20wph/10wph ratio) of genes in Quadrant II or IV in individuals in different survivorship classes (**Figure 3C**): short-lived with age of death between 28 and 36 weeks, long-lived with age of death between 45 and 50 weeks and longest-lived with age of death between 57 and 71 weeks [15]. Genes contained in Quadrant IV of **Figure 3B** that are protective at early age and detrimental at late ages, increased on average their expression with age in short-lived fish and decreased on average their expression in longest-lived fish. On the other hand, genes contained in Quadrant II that are detrimental at early ages and protective at late ages, displayed on average decreased expression with age in short-lived fish and on average increased expression in longest-lived fish.

GO overrepresentation analysis of the genes contained in each quadrant of **Figure 3D** revealed that risk genes (quadrant I) were enriched in terms related to immune response, as expected given the know association of chronic inflammation with mortality risk [39]. On the other hand, protective genes (quadrant III) were highly enriched in developmental terms. Antagonistic genes with initial protective effect (quadrant IV) were enriched in response to ER stress, protein folding and homeostasis, and antagonistic genes with initial detrimental effect (quadrant II) were enriched for Ras protein signal transduction. Age-dependent regulations of individual significant genes of these categories in the three lifespan groups are displayed in **Figure 3E**.

To display the relationship between risk/protective factors from longitudinal datasets and cross-sectional datasets, we combined the two analysis: the genes shown in Figure 2A-D were displayed

for brain, liver, skin and muscle, and antagonistic genes were highlighted in different colors in these datasets (**Figure 3F-H**). From the plots, it is evident that antagonistic factors with a late detrimental effect are preferentially upregulated in brain, liver and skin both in old fish and in short-lived fish; this is reflected in the enrichment of these factors (displayed as red dots) in Quadrant I of **Figure 3F-H**. On the other hand, factors with the opposite risk profile (displayed as black dots) are preferentially downregulated both in old and short-lived fish and are enriched in Quadrant III. The results in the muscle were less consistent (**Figure 3I**).

3.5. Effects of life-extending interventions on gene expression

In order to further validate the biological relevance of the gene regulation patterns described, we analyzed killifish models of life-extension for which expression data are available: rotenone treatment, knock-out of the AMP biosynthetic enzyme APRT and transgenic germline ablation. From Baumgart et al. [17] we selected genes that were up-regulated (red) and down-regulated (blue) after Rotenone treatment in the skin and assessed their regulation in the same tissues in either old vs. young longer-lived MZM0410 or young shorter-lived GRZ vs. young longer-lived MZM0410. Cumulative probability distributions show that genes up-regulated in the skin after Rotenone treatment are more frequently down-regulated in old fish and down-regulated in GRZ than genes down-regulated by Rotenone (**Figure 4A**), confirming that Rotenone counteracts an aging-associated signature.

We analyzed two different datasets corresponding to genetic models of life extension in the killifish:

- a RNA-seq dataset obtained from the livers of *APRT* mutants and controls at young and old ages [40];

- a RNA-seq dataset obtained from the livers of *DND-1* mutants (to obtain a germline-depleted killifish model) and controls at young and old ages [41].

Both transgenic models were generated in the short-lived GRZ line. In the dataset from Astre et al. (**Figure 4B**), genes upregulated in the mutant (shown in red) are up-regulated both in old MZM04010 fish and in the short-lived strain as compared to genes down-regulated in the mutant. This observation is contrary to the expectations. In the dataset from Moses et al. (**Figure 4C**),

genes upregulated in dnd1 mutants are up-regulated in old MZM0410 fish as compared to genes down-regulated in the mutant while we detected no significant difference across strains.

It is difficult to identify the underlying reason for these results, but it possible the different conditions in different laboratories (feeding, water parameters etc …) are strong confounders in genome-wide expression data either mask of interact with smaller effects of interventions. It should also e noted that DND-1 is one of the genes with the highest expression difference between GRZ and MZM0410 (**Supplementary Figure S1**).

3.6. Small molecules and upstream drivers associated with aging-related gene signatures

To investigate whether there are specific manipulations that can replicate the observed differences in the killifish, we created a connectivity map using the top 150 genes which were upregulated or downregulated with aging or in the GRz vs. MZM0410 contrasts using tissues as covariate. From the results (**Figure 5**), we detected two interventions that were correlated negatively to both gene sets: mTOR inhibition and topisomerase inhibition. Some of the treatments negatively correlated with age target typical geroprotective pathways such as HDAC inhibiton, oxidative phosphorylation, GnRH signalling. Some treatments negatively correlated with strain also inhibit typical aging pathways such TGF-Beta, JAK-STAT and MAP Kinase (**Figure 5**). No perturbagen were correlated positively to both datasets. Treatments positively correlated with age target typical aging-related pathways such as WNT activation, mTOR activation via LKB1, apoptosis, MYC activity. Treatments positively correlated with strain also inhibits typical geroprotective pathways such as proteasome and telomeres (**Supplementary Figure 5**).

3.7. Aging biomarkers reveal an “anticipated” aging profile in GRZ strain

To compare expression of putative aging biomarkers between the two strains, we used a public list of genes with either negative or positive monotonic expression profiles in MZM0410 [17] and we plotted their expression profile through aging in both strains as shown in **Figure 4**. Transcript expression profiles show that MZM biomarkers in GRZ have an “anticipated aging” profile. In particular, the expression level in the first time point (5 wph) was lower in GRZ strain for down-regulated biomarkers (**Figure 6A-B**), but higher for upregulated biomarkers (Figure 3C-3D), whereas slopes are reduced in both cases. This data suggests that GRZ shows higher biological age at the earliest time point analyzed rather than a steeper slope of age-related change.

3.9 | Comparison with mammals

The previous results suggest that baseline transcript differences between short- and long-lived strains are correlated with age-dependent transcript regulation. We used two different approaches to test the relevance of the genes identified in killifish for mammalian aging.

A study that analyzed liver proteome of (*Heterocephalus glaber*), a notorious example of exceptional longevity [42] and guinea pig as a related shorter-lived species detected a positive correlation between age-dependent expression in naked mole rat and naked-mole rat vs. guinea pig differential expression as young age similar to what we detected in killifish. Here, we analyzed a public dataset of age-dependent RNA-seq data from the skin and liver of naked mole rats [30, 43] and compared these with RNA-seq data for the same organs of the guinea pig, which represents a shorter-lived close relative. Indeed, a recent paper analyzed liver proteome in these two species and revealed a correlation pattern between inter-species comparison and age-dependent expression similar to one we report in **Figure 2A-2D**. We intersected age-dependent DEGs in the naked mole rat with DEGs of the contrast between naked mole rat and guinea pig at young and displayed these as 2D plots of the log2(fold change) in either condition obtaining a significant positive correlation (**Figure 7A-B**). The same result was obtained by analyzing the union of the DEGs for skin (**Figure 7C**).

Further, we highlighted in Fig. 5A the genes that in *N.furzeri* were significantly downregulated both during aging and in GRZ strain (blue dots) (**Figure 2A-2D**). In the skin, these are concentrated in quadrant III as well, i.e. they are also downregulated during aging in the Naked mole rat and in the guinea pig (19 genes in quadrant III vs. 36 total, Fisher's test p-value = 2.45e-03, **Figure 2C**, **Figure 7B**). The mentioned conserved genes (quadrant III in Figure 4B) are related to development and growth (such as PROX1 and NPR2), cell division/cytokinesis (such as HMCN1), neural stem cells differentiation and quiescence (such as GPM6A and CDH2), nervous system structure and development (such as PLP1 and LRRN1), and signal transduction (such as ARHGAP28).

To identify the biological processes that are differentially expressed according to species and age in the Naked Mole Rat, we performed GO over-representation analysis for the genes in the four quadrants of **Figure 7C**. The results (**Figure 7D**) show that genes down-regulated in the Guinea Pig and in old animals are enriched in categories related to membrane potential and metabolism. Interestingly, genes that are up-regulated during aging and in Guinea Pig show a highly significant

enrichment for RNA processing and splicing (**Figure S4**). These results are consistent with the prominent enrichment for the same terms in the gene coexpression network of the killifish transcriptome.

In a second approach we compared our data with a dataset that analyzed gene expression in 26 different rodent species and identified genes whose expression is positively- or negatively-correlated with species' maximum lifespan in five different tissues: brain, skin, liver, kidney, heart and lung [44]. Specifically, we analyzed the genes whose expression is negatively or positively correlated with maximum lifespan (MLS) across tissues. As shown in Figure 7, genes positively correlated with lifespan (shown in red) are upregulated in old long-lived fish with respect to genes negatively correlated with lifespan (shown in blue) (**Figure 7E, left panel**). This is consistent with the observation of Tyshvovskiy et al. [45] who detected an age-dependent up-regulation in mice of genes whose expression is positively correlated with mammalian maximum lifespan. This counterintuitive result suggests that aging is associated with a compensatory response. Analysis in the different strains showed a similar effect of lesser magnitude: genes positively correlated with lifespan are upregulated in the long-lived strain as opposed to the short-lived strain (**Figure 7E, right panel**).

4 | Discussion

We have analyzed the impact of genetically-determined lifespan differences on gene expression in the short-lived teleost model *N. furzeri* by analyzing laboratory strains with large differences in lifespan. The sequencing of its genome [46,11] and the availability of multi-omics datasets [15,17] make this species a convenient model system for computational exploration of aging-related expression datasets.

Here, we specifically analyzed differences in gene expression between two strains of *N. furzeri:* the shorter-lived GRZ and the longer-lived MZM-0410 during embryonic life and at different stages of adulthood. Validation of these differences was performed in a third longer-lived strain. The first somewhat surprising result is that the number of DEGs between GRZ and these longer-lived strains is higher than the number of age-dependent DEGs within the strains. In addition, GRZ shows an idiosyncratic gene expression pattern with some genes expressed at levels a couple of

orders of magnitude higher than in the other two strains. A striking example is DND, a gene whose expression is normally restricted to the primordial germs cells and gonads [36] and accordingly is not detectable in other tissues of the longer-lived strains. A second main result is the similarity of genotype-related transcriptional signature across organs, apparent as the aligned directions by which samples belonging to different tissues separate according to strain in the PCA plot. Strikingly, this difference is detectable, albeit smaller, in embryonic samples. In addition, we show that at sexual maturity the GRZ strains anticipates in all four tissues the age-dependent regulation that is observed later and more progressively in longer lived strains, a phenomenon that we call "anticipated aging" related to enrichment in autophagy (referred mostly to genes related to mitochondrial degradation) and cell cycle-related categories. Both age-dependent decrease in cell cycle and development and upregulation of autophagy categories has been extensively reported [47,48]. Interestingly, reduced mitotic activity in the brain, must be due to reduced neurogenesis and suggests heterochronic brain development across strains. The GRZ strain is the most used strain for experimental studies in killifish. Our results indicate that its transcriptional profile is different both in baseline and age-dependent regulation from that of longer-lived strains and warrant replication of at least key intervention studies in longer-lived strains.

The relevance of these pathways for lifespan regulation is confirmed by reanalysis of the longitudinal RNAseq dataset presented in [17]. Here, we highlight the presence of factors that follow antagonistic pleiotropic profile i.e. factors that are protective at one age and detrimental at another age or *vice versa*. This pattern is consistent with the *pleiotropic antagonism* theory [49,50]: a single gene controls multiple traits in a heterochronic way. So, if an allele increases reproduction in early life and reduces fitness later in life, it would still be positively selected . The expression of these antagonistic genes depend on age and genetic differences highlighting the validity of this theory to explain aging and suggests candidate genes for interventions on lifespan.

Analysis of the DEGs we identified in external datasets from life-extending interventions in killifish provided mixed results. Treatment with rotenone resulted in the expected direction of regulation but analysis of longer-living mutants showed inconsistent results. Another interesting comparison was that with a dataset of gene whose expression is correlated with interspecific differences in maximum lifespan. In that case, genes associated to longevity showed a significant

tendency for up-regulation during aging. This seemingly contradictory result parallels, however, what observed in rodents where longevity-associated genes are up-regulated with age [45].

Another important result is the centrality of RNA splicing during aging across genetic differences: network analysis revealed a conserved module across brain, liver, skin and muscle, which is positively correlated with genetic differences and enriched in this category. Comparison of age and genetic-dependent DEGs in a long-lived species, the Naked Mole Rat (*Heterocephalous glabrus*), highlights again this category as particularly enriched both during aging (in contrast with the killifish, where the category is downregulated) and in the Guinea Pig (consistently with short-lived/long-lived results).

The importance of splicing was already reported in literature: in the longitudinal work performed on *Nothobranchius furzeri* and presented by [17] the spliceosome complex was found among the top enriched categories in longest-lived animals (in comparison to short-lived individuals). In monkeys, on the other hand, a direct link between nutrition and splicing was observed: caloric restriction (CR) showed to regulate RNA processing and to direct alternative splicing usage in the liver [51]; in the transcriptome of human fibroblasts spliceosome was found to be downregulated with aging [47], consistently with previous work on *Nothobranchius furzeri* aging [15].

5 | Conflict of Interests

The authors declare no conflict of interest

6 | Acknowledgements

The authors thank the fish facility of the Leibniz Institute on Ageing for support in the animal breeding and husbandry. The Core Facility Next Generation Sequencing of the FLI is gratefully acknowledged for their technological support in library preparation and sequencing. This work was partially supported by the Grant of the Deutsche Forschungs Gemeinschaft 511530279 (CE 257/8-1).

7 | References

[1] Yamamoto R, Chung R, Vazquez JM, Sheng H, Steinberg PL, Ioannidis NM, Sudmant PH. **Tissue-specific impacts of aging and genetics on gene expression patterns in humans**. Nat Commun. 2022 Oct 3;13(1):5803. doi: 10.1038/s41467-022-33509-0. PMID: 36192477; PMCID: PMC9530233.

[2] Liao CY, Rikke BA, Johnson TE, Diaz V, Nelson JF. **Genetic variation in the murine lifespan response to dietary restriction: from life extension to life shortening.** Aging Cell. 2010 Feb;9(1):92-5. doi: 10.1111/j.1474-9726.2009.00533.x.

[3] Pincus Z, Smith-Vikos T, Slack FJ. **MicroRNA predictors of longevity in *Caenorhabitis elegans*.** PLOS Genetics, 2011 Sep;7(9):e1002306. doi: 10.1371/journal.pgen.1002306.

[4] Rea SL, Wu D, Cypser JR, Vaupel JW, Jonhson TE. **A stress-sensitive reporter predicts longevity in isogenic populations of *Caenorhabitis elegans*.** Nature Genetics, 2005 Aug;37(8):894-8. doi: 10.1038/ng1608.

[5] Wu D, Rea SL, Yashin AI, Johnson TE. **Visualizing hidden heterogeneity in isogenic populations of *C.elegans*.** Exp Gerontol. 2006 Mar;41(3):261-70. doi: 10.1016/j.exger.2006.01.003.

[6] Dance A. **Live fast, die young.** Nature. 2016 Jul 21;535(7612):453-5. doi: 10.1038/535453a. PMID: 27443744.

[7] Terzibasi Tozzini E, Valenzano DR, Benedetti M, Roncaglia P, Cattaneo A, Domenici L, Cellerino A. **Large differences in Aging Phenotype between Strains of the Short-Lived Annual Fish *Nothobranchius furzeri*.** PLoS One. 2008;3(12):e3866. doi: 10.1371/journal.pone.0003866.

[8] Valdesalici S and Cellerino A. **Extremely short lifespan in the annual fish *Nothobranchius furzeri*.** Proc Biol Sci. 2003 Nov 7;270 Suppl 2(Suppl 2):S189-91. doi: 10.1098/rsbl.2003.0048.

[9] Cellerino A, Valenzano DR, Reichard M. **From the bush to the bench: the annual *Nothobranchius* fishes as a new model system in biology.** Biol Rev Camb Philos Soc. 2016 May;91(2):511-33. doi: 10.1111/brv.12183.

[10] Terzibasi Tozzini E, Dorn A, Ng'oma E, Polacik M, Blazek R, Reichwald K, Petzold A, Watters B, Reichard M, Cellerino A. **Parallel evolution of senescence in annual fishes in response to extrinsic mortality.** BMC Evol Biol. 2013 Apr 3;13:77. doi: 10.1186/1471-2148-13-77.

[11] Valenzano DR, Benayoun BA, Singh PP, Zhang E, Etter PD, Hu CK, Clément-Ziza M, Willemsen D, Cui R, Harel I, Machado BE, Yee MC, Sharp SC, Bustamante CD, Beyer A, Johnson EA, Brunet A. **The African Turquoise Killifish genome provides insights into Evolution and Genetic Architecture of Lifespan.** Cell. 2015 Dec 3;163(6):1539-54. doi: 10.1016/j.cell.2015.11.008.

[12] Blažek R, Polačik M, Kačer P, Cellerino A, Řežucha R, Methling C, Tomášek O, Syslová K, Terzibasi Tozzini E, Albrecht T, Vrtílek M, Reichard M. **Repeated intraspecific divergence in life span and aging of African annual fishes along an aridity gradient.** Evolution, 2017 Feb;71(2):386-402. doi: 10.1111/evo.13127

[13] Kirschner J, Weber D, Neuschl C, Franke A, Böttger M, Zielke L, Powalsky E, Groth M, Shagin D, Petzold A, Hartmann N, Englert C, Brockmann GA, Platzer M, Cellerino A, Reichwald K. **Mapping of**

quantitative trait loci controlling lifespan in the short-lived fish *Nothobranchius furzeri* – a new vertebrate model for age research. Aging Cell, 2012 Apr;11(2):252-61. doi: 10.1111/j.1474-9726.2011.00780.x.

[14] Ng'oma E, Groth M, Ripa R, Platzer M, Cellerino A. **Transcriptome profiling of natural dichromatism in the annual fishes *Nothobranchius furzeri* and *Nothobranchius kadleci*.** BMC Genomics, 2014 Sep 2;15(1):754. doi: 10.1186/1471-2164-15-754.

[15] Baumgart M, Groth M, Priebe S, Savino A, Testa G, Dix A, Ripa R, Spallotta F, Gaetano C, Ori M, Terzibasi Tozzini E, Guthke R, Platzer M, Cellerino A. **RNA-seq of the aging brain in the short-lived fish *N.furzeri* – conserved pathways and novel genes associated with neurogenesis.** Aging Cell, 2014 Dec;13(6):965-74. doi: 10.1111/acel.12257.

[16] Aramillo Irizar P, Schäuble S, Esser D, Groth M, Frahm C, Priebe S, Baumgart M, Hartmann N, Marthandan S, Menzel U, Müller J, Schmidt S, Ast V, Caliebe A, König R, Krawczak M, Ristow M, Schuster S, Cellerino A, Diekmann S, Englert C, Hemmerich P, Sühnel J, Guthke R, Witte OW, Platzer M, Ruppin E, Kaleta C. **Transcriptomic alterations during ageing reflect the shift from cancer to degenerative diseases in the elderly.** Nat Commun. 2018 Jan 30;9(1):327. doi: 10.1038/s41467-017-02395-2

[17] Baumgart M, Priebe S, Groth M, Hartmann N, Menzel U, Pandolfini L, Koch P, Felder M, Ristow M, Englert C, Guthke R, Platzer M, Cellerino A. **Longitudinal RNA-seq Analysis of Vertebrate Aging identifies Mitochondrial Complex I as a small-molecule-sensitive Modifier of Lifespan.** Cell Systems, 2016 Feb 24;2(2):122-32. doi: 10.1016/j.cels.2016.01.014.

[18] Kelmer Sacramento E, Kirkpatrick JM, Mazzetto M, Di Sanzo S, Caterino C, Sanguanini M, Papaevgeniou N, Lefaki M, Childs D, Bagnoli S, Terzibasi Tozzini E, Bartolome A, Romanov N, Baumgart M, Huber W, Chondrogianni N, Vendruscolo M, Cellerino A, Ori A. **Reduced proteasome activity in the aging brain results in ribosome stoichiometry loss and aggregation.** Mol Syst Biol. 2020 Jun;16(6):e9596. doi: 10.15252/msb.20209596.

[19] Zhang B, Horvath S. **A general framework for weighted gene co-expression network analysis.** Stat Appl Genet Mol Biol. 2005;4:Article17. Doi: 10.2202/1544-6115.1128.

[20] Bentley DR, Balasubramanian S, Swerdlow HP, Smith GP, Milton J, Brown CG, Hall KP, Evers DJ, Barnes CL, Bignell HR, et al. **Accurate whole human genome sequencing using reversible terminator chemistry.** Nature. 2008 Nov 6;456(7218):53-9. doi: 10.1038/nature07517. PMID: 18987734; PMCID: PMC2581791.

[21] Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. **STAR: ultrafast universal RNA-seq aligner. Bioinformatics.** 2013 Jan 1;29(1):15-21. doi: 10.1093/bioinformatics/bts635. Epub 2012 Oct 25. PMID: 23104886; PMCID: PMC3530905.

[22] Liao Y, Smyth GK, Shi W. **featureCounts: an efficient general purpose program for assigning sequence reads to genomic features.** Bioinformatics. 2014 Apr 1;30(7):923-30. doi: 10.1093/bioinformatics/btt656. Epub 2013 Nov 13. PMID: 24227677.

[23] Buczak K, Kirkpatrick JM, Truckenmueller F, Santinha D, Ferreira L, Roessler S, Singer S, Beck M, Ori A. **Spatially resolved analysis of FFPE tissue proteomes by quantitative mass spectrometry.** Nat Protoc. 2020 Sep;15(9):2956-2979. doi: 10.1038/s41596-020-0356-y. Epub 2020 Jul 31. PMID: 32737464.

[24] Love MI, Huber W, Anders S. **Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2**. Genome Biol. 2014;15(12):550. doi: 10.1186/s13059-014-0550-8.

[25] Wickham H (2016). **ggplot2: Elegant Graphics for Data Analysis**. Springer-Verlag New York. ISBN 978-3-319-24277-4

[26] Luo W, Friedman MS, Shedden K, Hankenson KD, Woolf PJ. **GAGE: generally applicable gene set enrichment for pathway analysis.** BMC Bioinformatics. 2009 May 27;10:161. doi: 10.1186/1471-2105-10-161.

[27] Durinck S, Moreau Y, Kasprzyk A, Davis S, De Moor B, Brazma A, Huber W. **BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis.** Bioinformatics. 2005 Aug 15;21(16):3439-40. doi: 10.1093/bioinformatics/bti525.

[28] Durinck S, Spellman PT, Birney E, Huber W. **Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt.** Nat Protoc. 2009;4(8):1184-91. doi: 10.1038/nprot.2009.97.

[29] Langfelder P, Horvath S. **WGCNA: an R package for weighted correlation network analysis.** BMC Bioinformatics. 2008 Dec 29;9:559. doi: 10.1186/1471-2105-9-559.

[30] Sahm A, Bens M, Szafranski K, Holtze S, Groth M, Görlach M, Calkhoven C, Müller C, Schwab M, Kraus J, Kestler HA, Cellerino A, Burda H, Hildebrandt T, Dammann P, Platzer M. **Long-lived rodents reveal signatures of positive selection in genes associated with lifespan.** PLoS Genet. 2018 Mar 23;14(3):e1007272. doi: 10.1371/journal.pgen.1007272. PMID: 29570707; PMCID: PMC5884551.

[31] Subramanian A, et al. **A Next Generation Connectivity Map: L1000 Platform And The First 1,000,000 Profiles.** (https://www.ncbi.nlm.nih.gov/pubmed/29195078) 171(6):1437–1452

[32] Tozzini ET, Baumgart M, Battistoni G, Cellerino A. **Adult neurogenesis in the short-lived teleost Nothobranchius furzeri: localization of neurogenic niches, molecular characterization and effects of aging.** Aging Cell. 2012 Apr;11(2):241-51. doi: 10.1111/j.1474-9726.2011.00781.x. Epub 2012 Jan 13. PMID: 22171971; PMCID: PMC3437507.

[33] Ng'oma E, Reichwald K, Dorn A, Wittig M, Balschun T, Franke A, Platzer M, Cellerino A. **The age related markers lipofuscin and apoptosis show different genetic architecture by QTL mapping in short-lived Nothobranchius fish.** Aging (Albany NY). 2014 Jun;6(6):468-80. doi: 10.18632/aging.100660. PMID: 25093339; PMCID: PMC4100809.

[34] Baumgart, M., Barth, E., Savino, A. *et al.* **A miRNA catalogue and ncRNA annotation of the short-living fish *Nothobranchius furzeri*.** BMC Genomics. 2017 Sep 5;18(1):693. doi: 10.1186/s12864-017-3951-8. Erratum in: BMC Genomics. 2019 Nov 27;20(1):898.

[35] Supek F, Bošnjak M, Škunca N, Šmuc T. **REVIGO summarizes and visualizes long lists of gene ontology terms.** PLoS One. 2011;6(7):e21800. doi: 10.1371/journal.pone.0021800.

[36] Gross-Thebing T, Yigit S, Pfeiffer J, Reichman-Fried M, Bandemer J, Ruckert C, Rathmer C, Goudarzi M, Stehling M, Tarbashevich K, Seggewiss J, Raz E. **The Vertebrate Protein Dead End Maintains Primordial Germ Cell Fate by Inhibiting Somatic Differentiation.** Dev Cell. 2017 Dec 18;43(6):704-715.e5. doi: 10.1016/j.devcel.2017.11.019.

[37] Bahn A, Hagos Y, Reuter S, Balen D, Brzica H, Krick W, Burckhardt BC, Sabolic I, Burckhardt G. **Identification of a new urate and high affinity nicotinate transporter, hOAT10 (SLC22A13)**. J Biol Chem. 2008 Jun 13;283(24):16332-41. doi: 10.1074/jbc.M800737200.

[38] Sakthianandeswaren A, Parsons MJ, Mouradov D, MacKinnon RN, Catimel B, Liu S, Palmieri M, Love C, Jorissen RN, Li S, Whitehead L, Putoczki TL, Preaudet A, Tsui C, Nowell CJ, Ward RL, Hawkins NJ, Desai J, Gibbs P, Ernst M, Street I, Buchert M, Sieber OM. ***MACROD2*** **Haploinsufficiency Impairs Catalytic Activity of PARP1 and Promotes Chromosome Instability and Growth of Intestinal Tumors.** Cancer Discov. 2018 Aug;8(8):988-1005 doi: 10.1158/2159-8290.CD-17-0909.

[39] Simon AK, Hollander GA, McMichael A. **Evolution of the immune system in humans from infancy to old age.** Proc Biol Sci. 2015 Dec 22;282(1821):20143085. doi: 10.1098/rspb.2014.3085.

[40] Astre G, Atlan T, Goshtchevsky U, Oron-Gottesman A, Smirnov M, Shapira K, Velan A, Deelen J, Levy T, Levanon EY, Harel I. **Genetic perturbation of AMP biosynthesis extends lifespan and restores metabolic health in a naturally short-lived vertebrate.** Dev Cell. 2023 Aug 7;58(15):1350-1364.e10. doi: 10.1016/j.devcel.2023.05.015. Epub 2023 Jun 14. PMID: 37321215.

[41] Moses E, Atlan T, Sun X, Franek R, Siddiqui A, Marinov GK, Shifman S, Zucker DM, Oron-Gottesman A, Greenleaf WJ, Cohen E, Ram O, Harel I. **The killifish germline regulates longevity and somatic repair in a sex-specific manner**. bioRxiv [Preprint]. 2024 May 12:2023.12.18.572041. doi: 10.1101/2023.12.18.572041. Update in: Nat Aging. 2024 Jun;4(6):791-813. doi: 10.1038/s43587-024-00632-0. PMID: 38187630; PMCID: PMC10769255.

[42] Buffenstein R, Amoroso V, Andziak B, Avdieiev S, Azpurua J, Barker AJ, Bennett NC, Brieño-Enríquez MA, Bronner GN, Coen C, Delaney MA, Dengler-Crish CM, Edrey YH, Faulkes CG, Frankel D, Friedlander G, Gibney PA, Gorbunova V, Hine C, Holmes MM, Jarvis JUM, Kawamura Y, Kutsukake N, Kenyon C, Khaled WT, Kikusui T, Kissil J, Lagestee S, Larson J, Lauer A, Lavrenchenko LA, Lee A, Levitt JB, Lewin GR, Lewis Hardell KN, Lin TD, Mason MJ, McCloskey D, McMahon M, Miura K, Mogi K, Narayan V, O'Connor TP, Okanoya K, O'Riain MJ, Park TJ, Place NJ, Podshivalova K, Pamenter ME, Pyott SJ, Reznick J, Ruby JG, Salmon AB, Santos-Sacchi J, Sarko DK, Seluanov A, Shepard A, Smith M, Storey KB, Tian X, Vice EN, Viltard M, Watarai A, Wywial E, Yamakawa M, Zemlemerova ED, Zions M, Smith ESJ. **The naked truth: a comprehensive clarification and classification of current 'myths' in naked mole-rat biology.** Biol Rev Camb Philos Soc. 2022 Feb;97(1):115-140. doi: 10.1111/brv.12791. Epub 2021 Sep 3. PMID: 34476892; PMCID: PMC9277573.

[43] Heinze I, Bens M, Calzia E, Holtze S, Dakhovnik O, Sahm A, Kirkpatrick JM, Szafranski K, Romanov N, Nagender Sama S, Holzer K, Singer S, Ermolaeva M, Platzer M, Hildebrandt T, Ori A. **Species comparison of liver proteomes reveals links to naked mole-rat longevity and human aging.** BMC Biology, 2018 BMC Biol. 2018 Aug 2;16(1):82. doi: 10.1186/s12915-018-0547-y.

[44] Lu JY, Simon M, Zhao Y, Ablaeva J, Corson N, Choi Y, Yamada KYH, Schork NJ, Hood WR, Hill GE, Miller RA, Seluanov A, Gorbunova V. **Comparative transcriptomics reveals circadian and pluripotency networks as two pillars of longevity regulation.** Cell Metab. 2022 Jun 7;34(6):836-856.e5. doi: 10.1016/j.cmet.2022.04.011. Epub 2022 May 16. PMID: 35580607; PMCID: PMC9364679.

[45] Tyshkovskiy A, Ma S, Shindyapina AV, Tikhonov S, Lee SG, Bozaykut P, Castro JP, Seluanov A, Schork NJ, Gorbunova V, Dmitriev SE, Miller RA, Gladyshev VN. **Distinct longevity mechanisms across and within species and their association with aging.** Cell. 2023 Jun 22;186(13):2929-2949.e20. doi: 10.1016/j.cell.2023.05.002. Epub 2023 Jun 3. PMID: 37269831; PMCID: PMC11192172.

[46] Reichwald K, Petzold A, Koch P, Downie BR, Hartmann N, Pietsch S, Baumgart M, Chalopin D, Felder M, Bens M, Sahm A, Szafranski K, Taudien S, Groth M, Arisi I, Weise A, Bhatt SS, Sharma V, Kraus JM, Schmid F, Priebe S, Liehr T, Görlach M, Than ME, Hiller M, Kestler HA, Volff JN, Schartl M, Cellerino A, Englert C, Platzer M. **Insights into Sex Chromosome Evolution and Aging from the Genome of a Short-Lived Fish.** Cell, 2015 Dec 3;163(6):1527-38. doi: 10.1016/j.cell.2015.10.071.

[47] Marthandan S, Baumgart M, Priebe S, Groth M, Schaer J, Kaether C, Guthke R, Cellerino A, Platzer M, Diekmann S, Hemmerich P. **Conserved Senescence-Associated Genes and Pathways in Primary Human Fibroblasts Detected by RNA-seq.** PLoS One, 2016 May 3;11(5):e0154531. doi: 10.1371/journal.pone.0154531.

[48] Hernandez-Segura A, De Jong TV, Melov S, Guryev V, Campisi J, Demaria M. **Unmasking Transcriptional Heterogeneity in Senescent Cells.** Curr Biol. 2017 Sep 11;27(17):2652-2660.e4. doi: 10.1016/j.cub.2017.07.033.

[49] Williams GC. **Pleiotropy, natural selection, and the evolution of senescence**. Evolution 1957;11:398–411

[50] Long E, Zhang J. **Evidence for the role of selection for reproductively advantageous alleles in human aging.** Sci Adv. 2023 Dec 8;9(49):eadh4990. doi: 10.1126/sciadv.adh4990

[51] Rhoads TW, Burhans MS, Chen VB, Hutchins PD, Rush MJP, Clark JP, Stark JL, McIlwain SJ, Eghbalnia HR, Pavelec DM, Ong IM, Denu JM, Markley JL, Coon JJ, Colman RJ, Anderson RM. **Caloric Restriction Engages Hepatic RNA Processing Mechanisms in Rhesus Monkeys**. Cell Metab. 2018 Mar 6;27(3):677-688.e5. doi: 10.1016/j.cmet.2018.01.014

8 | Tables

9 | Figure legends

Figure 1. (A) Description of the discovery datasets: an RNA-seq dataset comprehensive of 5 tissues (brain, liver, skin, muscle and embryo), 2 strains (MZM-0410 and GRZ-D), 5 different time points (for brain, liver and skin) and 2 different stages (for embryo), with 5 replicates per sample and a small RNA-seq dataset (Baumgart, 2017) containing 3 different tissues (brain, liver and skin) and up to 6 different time points. (B) Principal Component Analysis (PCA) of brain (green), liver (red), skin (blue), muscle (yellow), and embryos (violet) samples based on transcript expression. Circles represent MZM samples, triangles represent GRZ samples; intensity of hues code for the time points. Inset shows a magnification of the embryo samples. (C) Principal Component Analysis (PCA) of brain (green), liver (red), skin (blue), muscle (yellow), and embryos (violet) samples based on miRNAs expression. Circles represent MZM samples, triangles represent GRZ samples; intensity of hues code for the time points. (D) Gene set enrichment analysis for strain-

dependent transcripts. Enrichment analysis was performed with the gage() package and enriched categories were summarized and displayed using REVIGO [31]. Significance of the single categories is coded by the intensity of the hue; GO terms were clustered according to semantic similarity and manually grouped into functional clusters (light blue dashed lines). (E) Overview of the "turquoise" gene module obtained by WGCNA: only the centre of the module, containing nodes with connection weight > 0.08, is visualised. Nodes highlighted in orange are the genes whose differential expression was tested in the validation set. Embryo samples were excluded from the network analysis. (F) Over-representation analysis of GO terms in the "turquoise" module obtained from WebGestalt: all Biological Process categories with FDR < 0.05 are shown. (G-H) Validation by qPCR of the hub genes SPECC1L (G) and PPP2R2D (H): qPCR was performed on MZM0410 and GRZ brain samples (5-weeks old) and Cq values were normalized to expression of INSR. Statistica, significance was determined with the Student t-test. (I) Description of the validation dataset. (L) Expression of the genes highlighted in (E) in the validation RNA-seq dataset. Statistical significance was computed with the Wilcoxon test. Only brain samples were considered for the analysis.

Figure 2. (A-D) Scatterplots of strain- and age-dependent DEGs in the transcriptome from A) brain, B) liver, C) skin and D) muscle. For each DEG, Log_2 Fold change for Old vs. Young comparison is plotted on the X axis and Log2 fold-change Short-lived vs. Long-lived strain comparison on the Y axis. Only genes DEGs significant in both comparisons are shown. Pearson's correlation coefficient and total number of DEGs in the intersection as reported on the graphs. The density of genes is coded by the intensity of the blue hues, the black line represents a spline fit of the data. (E) GO summary analysis of the genes contained in the 4 quadrants of the plots shown in Fig. 3A-3D. Gene enrichment was performed for each tissue separately using Biological Process as a the GO term database and then a metanalysis was performed using Fisher's method. The length of the bars represents –log10 of the false discovery rate (FDR). (F-H) Comparison between strain and age-dependent differential expression in the miRNome of brain (F), liver (G) and skin (H). Gene regulation is plotted as Log_2FC for Old vs. Young comparison on the X axis and Short-lived vs. Long-lived strain comparison on the Y axis. MicroRNAs differentially expressed in either of the two conditions (union) are shown. Pearson's correlation coefficient, total number of DEGs and number of genes per quadrant are reported. The density of genes is coded by the intensity of the blue hues.

Figure 3. Cox regression analysis on a longitudinal dataset. (A) Workflow of the longitudinal experiment (modified from [15]). (B) Scatterplot beta coefficients obtained from Cox hazard regression analysis at 10 weeks (x-axis) and 20 weeks (y-axis). Only genes with significant beta coefficients for both the analyses (pval < 0.05) are displayed. Fisher's test results and total number of genes, as well as number of genes per quadrant are shown. The density of genes is coded by the intensity of the blue hues. (C) Age-dependent regulation (shown as log (20w/10w ratio)) of genes with antagonistic pleiotropic action (i.e. genes contained in Quadrant II in the right-side plot and factors contained in Quadrant IV in the left-side plot) in different classes of individuals (from [17]): short-lived, long-lived and longest-lived animals were taken into account, as individuals with lifespans respectively between 28-36 wph (short-lived), 45-50 wph (long-lived) and 57-71 wph (longest-lived). Statistical analysis was performed with ANOVA test. (D) GO analysis of the genes contained in the 4 quadrants of the plot showed in Fig. 3B. Gene enrichment was performed using GO Process as a functional database; length of the bars represents –log10(FDR). (E) Detail on ER stress and Ras signalling pathway as examples of specific enriched categories with antagonistic pleiotropic action. Age-dependent regulation (as ratio 20w/10w) in short-lived, long-lived and longest-lived animals is displayed as gradient from red (upregulated) to white (neutral) to blue (downregulated), while beta coefficients (from Cox analysis) are displayed as gradient from purple (beta>0) to green (beta<0). (F-I) Aging- and strain-dependence of factors with antagonistic pleiotropic behaviour. Scatterplot of strain and age-dependent DEGs in the (F) brain, (G) liver and (H) skin and (I) muscle transcriptome, displayed as in Fig. 2. Genes in quadrant II and IV of Fig. 3B are highlighted as red and black dots, respectively.

Figure 4. Effects of life-extending interventions on gene expression using the genes from Baumgart et al. (A), from Astre et al. (B), and from Moses et al. (C). We used the killifish DEGs from both our comparisons (age and strain) and plotted the variation in expression for the genes from each one of the papers. Genes were filtered for adjusted p-value< 0.05, and visualization was performed using a cumulative distribution.

Figure 5. List of perturbagens whose transcriptional signature is positively or negatively correlated with the effects of age or strain. (A) Alignment with the effects of age ranked based on connectivity scores. (B) Alignment with the effects of strain ranked based on connectivity scores (raw_cs). Positive values (orange) imply positive associations and negative values (blue) negative

associations. Only significant terms are displayed. TRT_SH.CGS indicates loss of function by shRNAs. TRT_CP indicates perturbation by compounds. TRT_OE indicates perturbation by overexpression. TRT_XPR indicates loss of function by CRISPR. MOA indicates mechanism of action class. PCL perturbagen class.

Figure 6. Comparison of expression of MZM-specific aging biomarkers in both strains (GRZ-D as red line and MZM0410 as blue line) for brain (A, C) and skin (B, D). The gene list for each cluster was obtained from [15]: clusters 1 (A-C) shows negative monotonic expression profile in all the analysed tissues, while clusters 2 (B-D) shows positive monotonic expression profile. Scaled expression values are plotted with the regression spline using the ggplot2 package, the shaded envelope represents confidence intervals of the mean expression.

Figure 7. Comparison of age-dependent and lifespan-dependent gene expression between *N.furzeri* and *Heterocephalus glaber* in the liver (A) and in the skin (B). (A) Scatterplot of age-dependent DEGs in the Naked mole rat on the X axis and *H. glaber* vs. *Cavia porcellus* on the Y axis for liver samples (from [37]); LogFC for Old vs. Young and Long-lived vs. Short-lived comparisons are displayed. Only genes with adjusted p-value< 0.01 for both comparisons are displayed. Pearson correlation and total number of significant genes are also shown in the plot. Coloured dots correspond to genes found in *N.furzeri* comparisons in the same tissue: in particular blue dots are genes downregulated in both the variables. (B) Scatterplot of age and strain DEGs in the Naked mole rat taken from skin samples; LogFC for Old vs. Young and Long-lived vs. Short-lived comparisons are displayed. Genes with adjusted p-value< 0.01 for both the comparisons were considered as significant. Pearson correlation and total number of significant genes are also shown in the plot. Colored dots correspond to genes found in *N.furzeri* comparisons in the same tissue: blue dots are genes downregulated in both the variables. (C). LogFC value for Old vs. Young and Short-lived vs. Long-lived samples are displayed: filtering was performed on the combined p-value between the two variables using meta-analysis. Pearson correlation and total number of genes, as well as number of genes per quadrant are shown. The density of genes is coded by the intensity of the blue hues. (D) GO analysis of the genes contained in the four quadrants of Fig. S7A. Gene enrichment was performed using GO Biological Process as a functional database. The categories are plotted as –log10 of the false discovery rate (FDR). (E) Visualization of the gene expression trajectory in killifish for genes negatively (left) and positively (right) correlated with

maximum lifespan, obtained from [44]. We used the killifish DEGs from both our comparisons (age and strain) and plotted the variation in expression for both genes with positive and negative correlation with maximum lifespan (MLS). Genes were filtered for adjusted p-value < 0.05, and visualization was performed using a cumulative distribution.

Figure 1

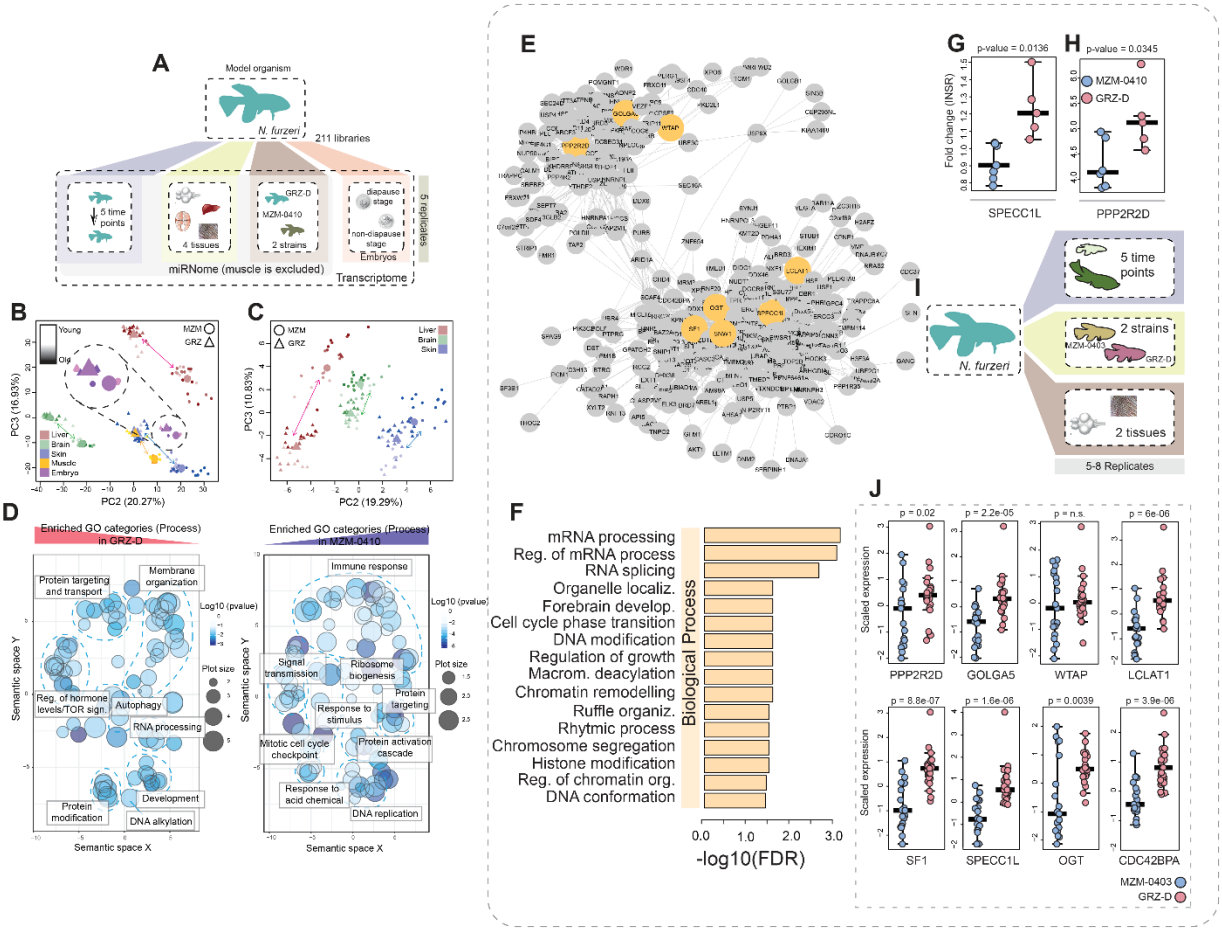

Figure 2

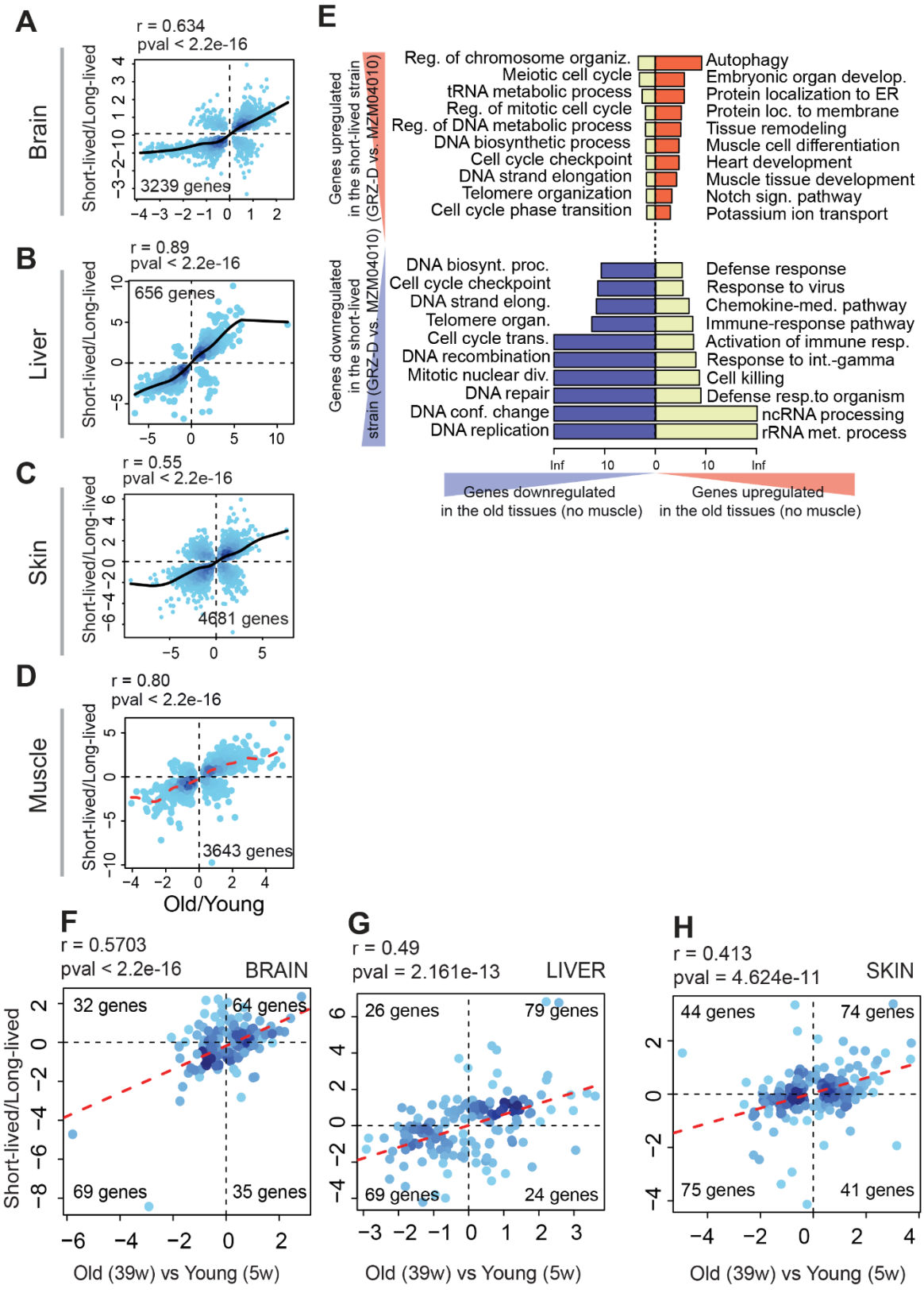

Figure 3

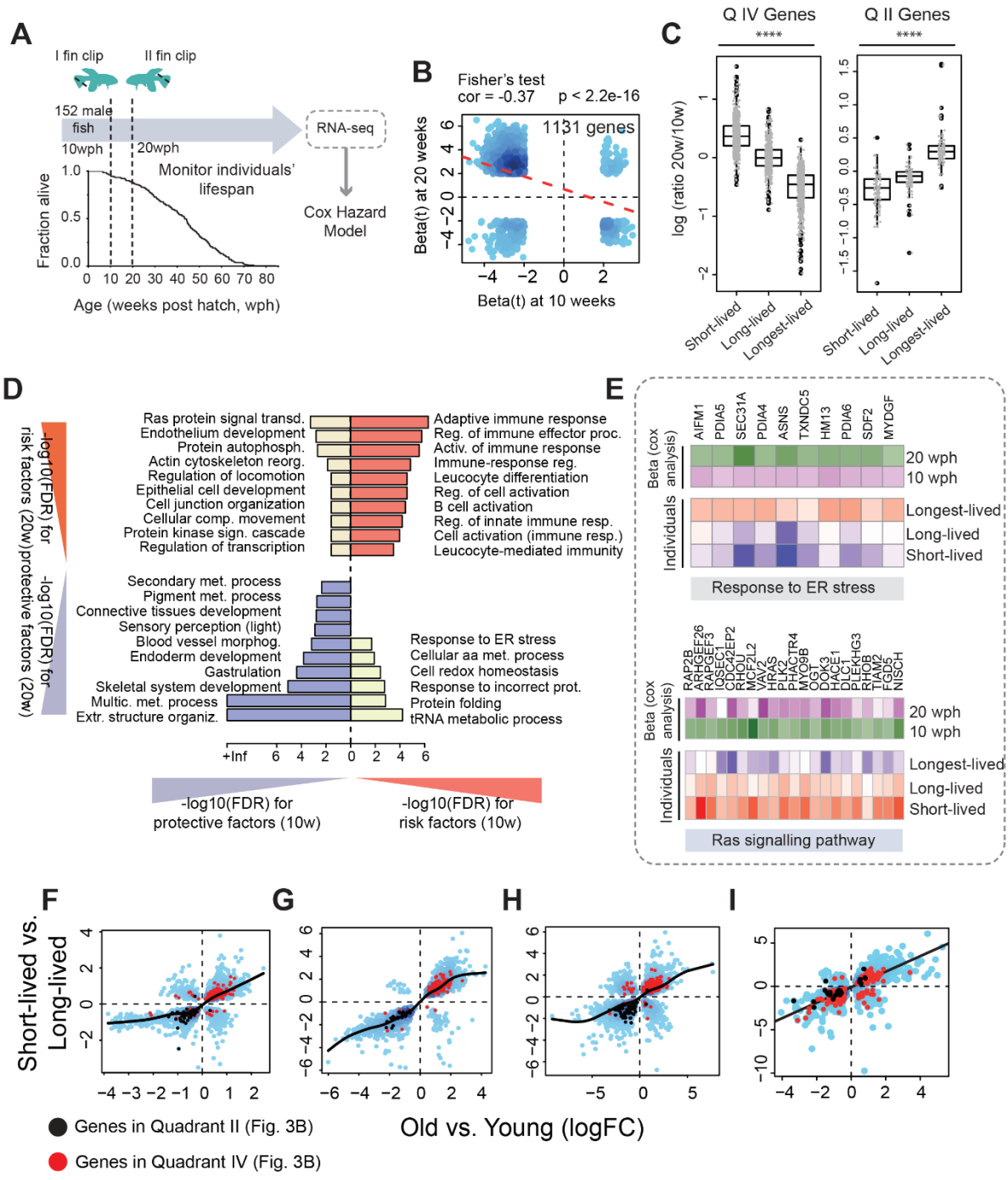

Figure 4

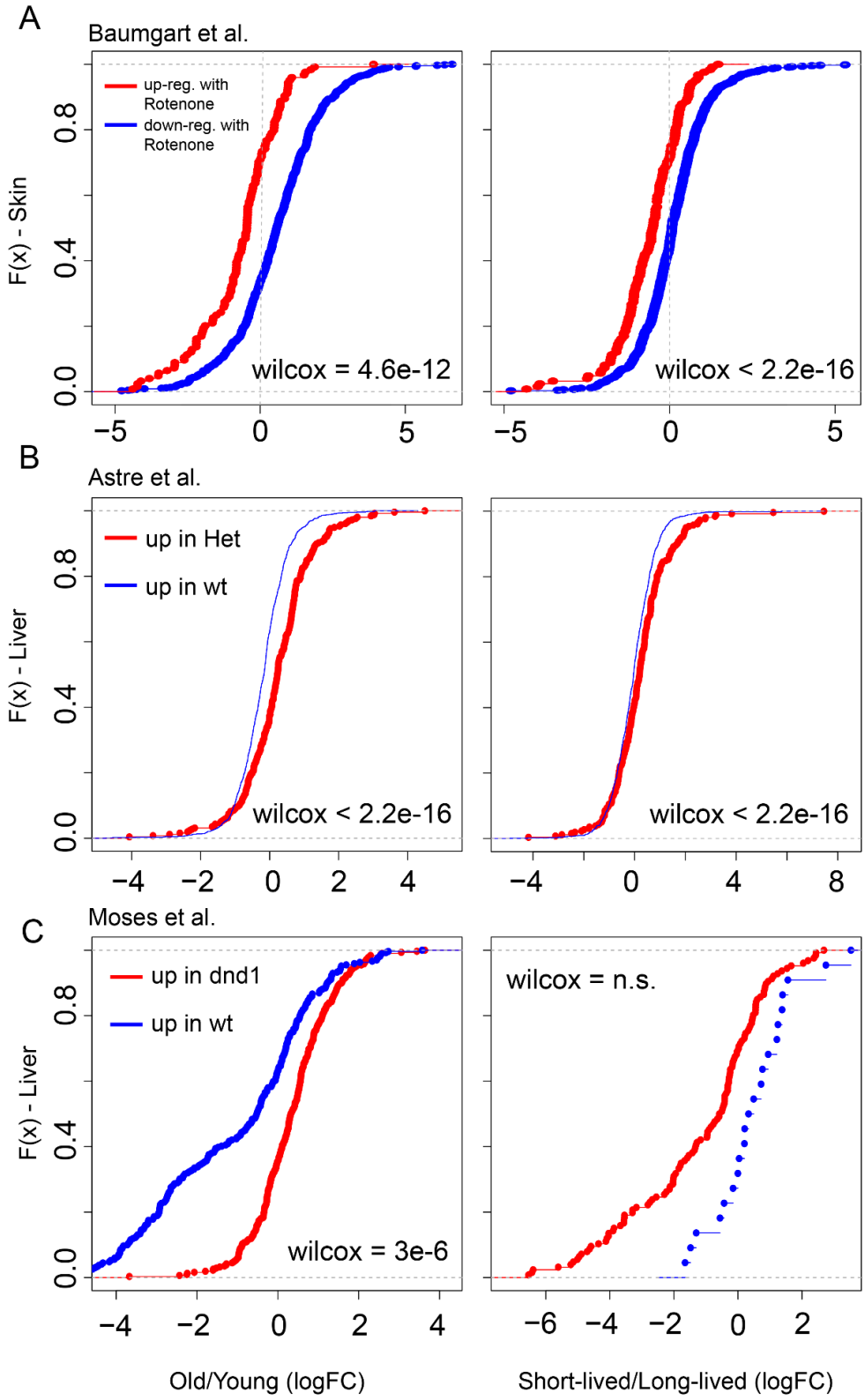

Figure 5

A

id	norm_cs	cell_iname	pert_type	raw_cs	fdr_q_nlog10	set_type	src_set_id
		HEPG2	TRT_SH.CGS	0.85	15.65	PATHWAY_SET	BIOCARTA_HIVNEF_PATHWAY
		VCAP	TRT_CP	0.92	2.54	MOA_CLASS	TUBULIN_INHIBITOR
		HCC515	TRT_SH.CGS	1.00	2.43	PATHWAY_SET	REACTOME_TCR_SIGNALING
		HA1E	TRT_CP	0.96	1.68	MOA_CLASS	ESTROGEN_RECEPTOR_ANTAGONIST
		VCAP	TRT_CP	0.91	1.66	PCL	CP_TUBULIN_INHIBITOR
		A549	TRT_OE	0.71	1.65	PATHWAY_SET	HALLMARK_WNT_BETA_CATENIN_SIGNALING
		A549	TRT_XPR	0.87	1.60	PATHWAY_SET	REACTOME_LIGAND_GATED_ION_CHANNEL_TRANSPORT
		HCC515	TRT_SH.CGS	1.00	1.59	PATHWAY_SET	REACTOME_DOWNSTREAM_TCR_SIGNALING
		-666	TRT_SH.CGS	0.87	1.52	PCL	KD_PI3K_SIGNALING
		A375	TRT_OE	0.86	1.52	PATHWAY_SET	REACTOME_CELL_DEATH_SIGNALLING_VIA_NRAGE_NRIF_AND_NADE
		A549	TRT_OE	0.69	1.52	PATHWAY_SET	PID_LKB1_PATHWAY
		HCC515	TRT_SH.CGS	1.00	1.52	PATHWAY_SET	BIOCARTA_TOLL_PATHWAY
		ES2	TRT_XPR	0.91	1.38	PATHWAY_SET	REACTOME_INCRETIN_SYNTHESIS_SECRETION_AND_INACTIVATION
		HEK293T	TRT_OE	0.93	1.18	PATHWAY_SET	KEGG_BLADDER_CANCER
		HA1E	TRT_OE	0.78	1.18	PATHWAY_SET	REACTOME_APOPTOTIC_EXECUTION_PHASE
		PC3	TRT_CP	-0.82	1.02	MOA_CLASS	MTOR_INHIBITOR
		MCF7	TRT_CP	-0.91	1.15	MOA_CLASS	FLT3_INHIBITOR
		PC3	TRT_CP	-0.82	1.47	MOA_CLASS	HDAC_INHIBITOR
		-666	TRT_SH.CGS	-0.96	15.65	PCL	KD_CANONICAL_HIGH_MOBILITY_GROUP_PROTEINS
		A549	TRT_CP	-0.89	15.65	PCL	CP_TOPOISOMERASE_INHIBITOR
		HT29	TRT_SH.CGS	-0.79	15.65	PATHWAY_SET	HALLMARK_OXIDATIVE_PHOSPHORYLATION
		VCAP	TRT_SH.CGS	-0.84	15.65	PATHWAY_SET	KEGG_GNRH_SIGNALING_PATHWAY
		A549	TRT_CP	-0.85	15.65	MOA_CLASS	TOPOISOMERASE_INHIBITOR
		-666	TRT_CP	-0.68	15.65	PCL	CP_TOPOISOMERASE_INHIBITOR
		A375	TRT_CP	-0.93	15.65	PCL	CP_TOPOISOMERASE_INHIBITOR
		MCF7	TRT_SH.CGS	-0.99	15.65	PATHWAY_SET	REACTOME_ANTIVIRAL_MECHANISM_BY_IFN_STIMULATED_GENES
		-666	TRT_XPR	-0.77	15.65	PCL	KD_PROTEASOME_PATHWAY
		-666	TRT_CP	-0.70	15.65	MOA_CLASS	HDAC_INHIBITOR
		A375	TRT_CP	-0.90	15.65	MOA_CLASS	TOPOISOMERASE_INHIBITOR
		-666	TRT_CP	-0.78	15.65	PCL	CP_HDAC_INHIBITOR

B

id	norm_cs	cell_iname	pert_type	raw_cs	fdr_q_nlog10	set_type	src_set_id
		-666	TRT_SH.CGS	0.90	15.65	PATHWAY_SET	BIOCARTA_PROTEASOME_PATHWAY
		-666	TRT_SH.CGS	0.90	1.43	PCL	KD_PROTEASOME_PATHWAY
		H1299	TRT_OE	0.87	1.18	PATHWAY_SET	BIOCARTA_ARENRF2_PATHWAY
		SALE	TRT_OE	0.87	1.15	PATHWAY_SET	BIOCARTA_ERK_PATHWAY
		A375	TRT_SH.CGS	0.94	1.15	PATHWAY_SET	PID_TELOMERASE_PATHWAY
		-666	TRT_OE	0.85	1.10	PCL	OE_NFKB_ACTIVATION
		-666	TRT_SH.CGS	0.85	1.06	PATHWAY_SET	PID_VEGF_VEGFR_PATHWAY
		MCF7	TRT_CP	0.99	1.04	MOA_CLASS	JAK_INHIBITOR
		SALE	TRT_OE	0.92	1.03	PATHWAY_SET	PID_CERAMIDE_PATHWAY
		A375	TRT_OE	0.91	1.03	PATHWAY_SET	BIOCARTA_CARDIACEGF_PATHWAY
		MCF7	TRT_SH.CGS	0.95	0.75	PATHWAY_SET	PID_AR_PATHWAY
		SALE	TRT_OE	0.95	0.75	PATHWAY_SET	BIOCARTA_IL6_PATHWAY
		MCF7	TRT_SH.CGS	0.94	0.75	PATHWAY_SET	PID_CASPASE_PATHWAY
		A375	TRT_OE	0.89	0.75	PATHWAY_SET	KEGG_DORSO_VENTRAL_AXIS_FORMATION
		VCAP	TRT_SH.CGS	0.96	0.75	PATHWAY_SET	PID_SHP2_PATHWAY
		HCC515	TRT_SH.CGS	-0.95	0.72	PATHWAY_SET	REACTOME_ACTIVATED_TLR4_SIGNALLING
		PC3	TRT_SH.CGS	-0.98	0.72	PATHWAY_SET	PID_ERBB4_PATHWAY
		HCC515	TRT_SH.CGS	-0.95	0.72	PATHWAY_SET	REACTOME_DOWNSTREAM_SIGNAL_TRANSDUCTION
		HCC515	TRT_SH.CGS	-0.95	0.72	PATHWAY_SET	REACTOME_SIGNALING_BY_ERBB2
		U251MG	TRT_XPR	-0.82	1.02	PATHWAY_SET	PID_IL2_PI3K_PATHWAY
		VCAP	TRT_SH.CGS	-0.90	1.02	PATHWAY_SET	BIOCARTA_MAPK_PATHWAY
		-666	TRT_CP	-0.69	1.02	MOA_CLASS	MTOR_INHIBITOR
		A549	TRT_SH.CGS	-1.00	1.02	PATHWAY_SET	REACTOME_SIGNALING_BY_TGF_BETA_RECEPTOR_COMPLEX
		MCF7	TRT_CP	-0.90	1.02	MOA_CLASS	TOPOISOMERASE_INHIBITOR
		VCAP	TRT_SH.CGS	-0.98	1.02	PATHWAY_SET	HALLMARK_IL6_JAK_STAT3_SIGNALING
		AGS	TRT_XPR	-0.83	1.02	PATHWAY_SET	PID_FOXO_PATHWAY
		A549	TRT_SH.CGS	-0.98	1.02	PATHWAY_SET	HALLMARK_TGF_BETA_SIGNALING
		VCAP	TRT_SH.CGS	-0.99	15.65	PATHWAY_SET	REACTOME_REGULATION_OF_MRNA_STABILITY_BY_PROTEINS_THAT_BIND_AU_RICH_ELEMENTS
		VCAP	TRT_SH.CGS	-0.99	15.65	PATHWAY_SET	REACTOME_ANTIGEN_PROCESSING_CROSS_PRESENTATION
		VCAP	TRT_SH.CGS	-0.94	15.65	PATHWAY_SET	KEGG_MELANOGENESIS

pert_type
TRT_CP
TRT_LIG
TRT_OE
TRT_SH.CGS
TRT_XPR

Figure 6

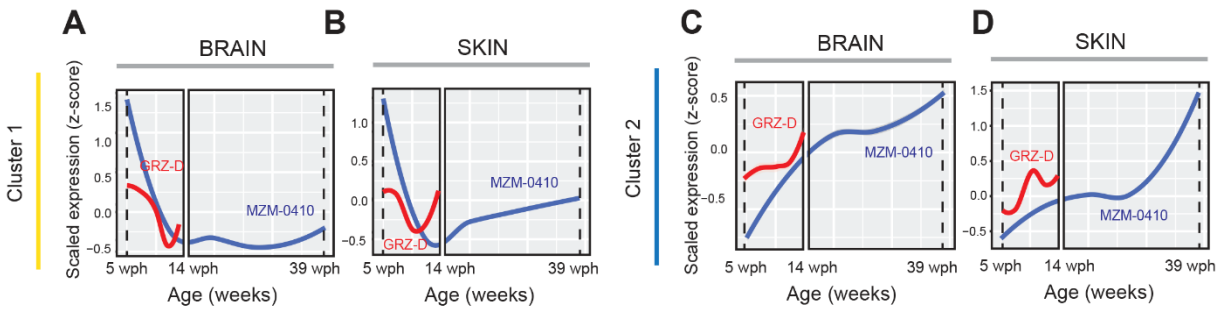

Figure 7