# Evolutionary transcriptomics reveals longevity mostly driven by polygenic and indirect selection in mammals

Weiqiang Liu[1,2#], Pingfen Zhu[1#], Meng Li[1], Zihao Li[1,2], Yang Yu[3], Gaoming Liu[1], Juan Du[1,2], Xiao Wang[1], Jing Yang[1,2], Ran Tian[4], Inge Seim[4,5], Alaattin Kaya[6], Mingzhou Li[7], Ming Li[1], Vadim N. Gladyshev[8], Xuming Zhou[1,3*]

[1]Key Laboratory of Animal Ecology and Conservation Biology, Institute of Zoology, Chinese Academy of Sciences, Beijing 100101, China

[2]University of Chinese Academy of Sciences, Beijing 100049, China

[3]School of Life Sciences, University of Science and Technology of China, Anhui, 230027, China

[4]Integrative Biology Laboratory, College of Life Sciences, Nanjing Normal University, Nanjing, Jiangsu, 210046, China

[5]School of Biology and Environmental Science, Queensland University of Technology, Brisbane, Queensland, 4000, Australia

[6]Department of Biology, Virginia Commonwealth University, Richmond, VA, 23284, USA

[7]Institute of Animal Genetics and Breeding, College of Animal Science and Technology, Sichuan Agricultural University, Chengdu, 611130, China

[8]Division of Genetics, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts, 02115 USA

[#]Authors contributed equally.

*Correspondence and requests for materials should be addressed to X.Z. (E-mail: zhouxuming@ioz.ac.cn).

## Abstract

The maximum lifespan varies more than 100-fold in mammals. This experiment of nature may uncover of the evolutionary forces and molecular features that define longevity. To understand the relationship between gene expression variation and maximum lifespan, we carried out a comparative transcriptomics analysis of liver, kidney, and brain tissues of 106 mammalian species. We found that expression is largely conserved and very limited genes exhibit common expression patterns with longevity in all the three organs analyzed. However, many pathways, e.g., “Insulin signaling pathway”, and "FoxO signaling pathway”, show accumulated correlations with maximum lifespan across mammals. Analyses of selection features further reveal that methionine restriction related genes whose expressions associated with longevity, are under strong selection in long-lived mammals, suggesting that a common approach could be utilized by natural selection and artificial intervention to control lifespan. These results suggest that natural lifespan regulation via gene expression is likely to be driven through polygenic model and indirect selection.

**Keywords:** Mammals, Longevity, Comparative transcriptomics, Natural selection

## Introduction

Over ~150 million years of evolution, mammals have diversified dramatically (over 100-fold) in terms of maximum lifespan (hereafter 'lifespan' and often used as a proxy for longevity). This experiment of nature has attracted much interest from biologists (*1*). The identification of species lifespan-related genetic variation has been a key approach to resolve this question, with focus primarily on exceptionally long-lived species. For example, a survey of the genome of the bowhead whale (the longest-lived mammal known with a lifespan exceeding 200 years) revealed specific sequence changes in genes associated with DNA repair, cell cycle, and aging (*2, 3*). Naked mole rats, which are the longest-lived rodents (lifespan > 30 years), were reported to harbor unique variations in genes related to macromolecular degradation, mitochondrial function, and telomere maintenance, as well as tumor suppression (*4*). Similarly, substitutions in genes related to the GH/IGF-1 axis are found in the Brandt's bat, which is the longest-lived flying mammal known (*5*). Recent studies have also shown that elephants could be an attractive model organism to study aging, as they exhibit a long lifespan (> 50 years), a low cancer rate, and present an unexpected expansion of potentially functional TRP53 pseudogenes (*6, 7*). These studies suggest that there is a diversity in genetic factors that supports molecular mechanisms of longevity in mammals.

In addition to genetic variation, the lifespan of mammals is also likely to be modulated by the expression level of genes (*8*). For example, it was shown that IGF1R knockout leads to a lifespan increase of 33% and 15.9% in female and male mice, respectively (*9-11*). Similarly, mTOR inhibition in mice increased the median lifespan of female and male mice by ~25% (*12-14*). In addition, *SIRT6* overexpression increased the median lifespan of male mice by 14.5% (*15*). Comparing gene expression across species is challenging

because variables such as developmental stages and environmental factors can mask or distort genuine expression differences. By assuming that gene expression is primarily shaped by stable selection, comparative transcriptomics analyses have been conducted to investigate gene expression patterns across species from an evolutionary perspective (*16*-*23*). For example, previous studies compared gene expression in the liver, kidney and brain tissues of 34 mammalian species (*16*) and cultured fibroblast cells of 16 mammals (13 rodents, two bats, and a shrew), revealing a number of genes and pathways showing association with maximum longevity (*17*). These studies found that the expression of genes related to central energy metabolism, DNA damage repair, sugar metabolism, and DNA repair was positively associated with longevity, whereas gene expression associated with mitochondrial metabolism, transcriptional regulation, calcium-mediated signaling pathways, protein ubiquitination, and protein localization was negatively associated (*16, 17*). At the same time, in the comparison of age-related transcriptomic changes in *Myotis*, human, mouse and wolf. *Myotis* exhibits unique molecular mechanisms for lifespan extension in functions related to DNA repair, autophagy, immunity, and tumor suppression (*24*). These studies provide many insights into the relationship between the variation of gene expression and longevity traits across species. Nevertheless, the number of species analyzed in previous studies is relatively small and may not fully represent the diversity of gene expression in mammals.

To better characterize the expression profile of protein-coding genes across the mammalian phylogeny, we generated or obtained transcriptome data from the brain, kidney and liver tissues of 106 mammals, covering 16 orders and 45 families. First, gene expression in different organs and species-specific expression patterns were assessed; then genes whose expression levels was significantly correlated with longevity traits were identified within a

phylogenetic framework. Pathways that showed signatures of expression changes associated with longevity were identified using a modified summary approach. Finally, an integrated analysis of gene expression and selection pressure was carried out to measure the intensity of selection of the associated genes. The data and analyses presented in this study currently represent the most comprehensive characterization of gene expression in mammalian organs and contribute to our understanding of how lifespan is regulated at the level of gene expression.

## Results and Discussion

### Data generation and species-specific gene expression

To capture the diversity of gene expression across mammals, RNA-seq data (~5.2 billion Illumina NovaSeq 6000 reads) were generated from polyadenylated RNA fraction of liver and kidney tissues of 56 species (**Table S1**, **Fig. 1A**, **Fig. S1**, and **Methods**). Previously published transcriptomes of liver, kidney and brain of 50 additional species were also used (*2, 16, 18, 25-35*) (**Fig. 1A**, **Fig. S1**, **Table S1**, and **Methods**). After filtering and orthologs calling, a comprehensive expression dataset was obtained for 13,508 protein-coding genes in three organs of 106 species, which covered the following orders: Artiodactyla (*n* = 9), Carnivora (*n* = 12), Chiroptera (*n* = 36), Cingulata (*n* = 1), Eulipotyphla (*n* = 5), Hyracoidea (*n* = 1), Lagomorpha (*n* = 1), Perissodactyla (*n* = 1), Pilosa (*n* = 1), Primates (*n* = 16), Rodentia (*n* = 18), and Scandentia (*n* = 1). In addition to 102 placental mammals, our dataset included the platypus (Monotremata), the Tasmanian devil (Dasyuromorphia), an opossum (Didelphimorphia), and the sugar glider (Diprotodontia) (**Fig. 1A** and **Table S2**). Information on adult weight (AW) and longevity-related traits—including maximum lifespan (ML), female time to maturity (FTM), adult-weight-adjusted residuals (i.e., MLres and FTMres), and other life-history traits (habitats, feeding habit, etc.) were also collected and analyzed (see **Methods**,

**Fig. 1A** and **Table S2**). Among these traits, ML and FTM reflect changes in absolute longevity, while the residuals indicate changes in relative longevity (**Fig. 1A**). Three algorithms (i.e., mice, missForest, and Phylopars) were used to impute and estimate missing life-history data for the species analyzed (**Figs. 1B-C**, **Fig. S2**, **Table S2**, and **Methods**).

Principal component analysis of the expression data metrics was performed to assess the gene expression patterns across species and tissues. Gene expression was tissue-specific rather than lineage-specific (**Fig. 1D**), consistent with previous reports (*16, 18-20*). To characterize genes with species-specific expression patterns, the specificity index ($\tau$; Tau) for gene expression was calculated for each tissue (**Methods**). Tau ranges from 0 to 1 and indicates how broadly (0) or specific (1) a gene is expressed (*36*). None of the 13,508 genes were broadly expressed across species ($\tau < 0.2$) (**Table S3**), which could be due to the large-scale sampling conducted in this study (**Fig. 1E**). Compared to the liver and kidney, the brain presented the highest number of genes showing species-specific expression (6,716, 2,519 and 1,186 genes for the brain, kidney, and liver, respectively) (**Fig. S3**). The gene with the highest species-specific index in the brain was *GRM1* (glutamate metabotropic receptor 1). The primary role of this gene is to protect neurons from apoptosis (*37, 38*). Studies have shown that motor coordination and context-specific associative learning is impeded in mice lacking GRM1 (*39*). It is interesting that *GRM1* is also the most highly expressed gene in bats, because species that specifically express this gene have excellent spatial memory and the ability to recognize individuals (*40*). In the kidney, a detoxification organ, the most highly species-specific genes were *ZNF518A* (zinc finger protein 518A) and *UBE2N* (ubiquitin-conjugating enzyme E2 N). These genes regulate the maintenance of cell types (*41*) and DNA repair (*42, 43*), respectively, suggesting that they play a role in kidney function. *UBE2N* is

also highly expressed among long-lived species—such as vervets, naked mole-rats, and greater short-nosed fruit bats—which is consistent with previous reports that DNA repair-related genes are commonly highly expressed in long-lived species (*3, 17, 44*). In the liver, the most species-specific gene was *PTPRG* (protein-tyrosine phosphatase gamma), a marker of oxidative stress associated with inflammation and aging. It has been reported that inflammation caused by obesity can promote *PTPRG* expression in the liver, and an excessive expression of this gene can cause severe liver and insulin resistance (*45*).

To characterize pathways showing expression specificity across organs and species, pathway enrichment analysis was performed under a polygenic model, in which we used the sum of the Tau index of genes in a pathway (*46, 47*) (**Fig. 1F**, **Tables S4-5**, and **Methods**). The brain showed enrichment for "ABCA transporters in lipid homeostasis" (Score = 10.79, $P = 2.68 \times 10^{-3}$), "semaphorin interactions" (Score = 43.99, $P = 2.86 \times 10^{-3}$), "PCP/CE pathway" (Score = 23.69, $P = 8.67 \times 10^{-3}$), and "interaction between L1 and ankyrins" (Score = 13.26, $P = 1.64 \times 10^{-3}$). Semaphorin interactions and the PCP/CE pathway regulate synaptic formation and help to determine neuronal polarity (*48-51*). They are also important molecular players in the aftermath of nervous system damage events (*52, 53*). In the kidney, specifically expressed genes were enriched in pathways related to detoxification, such as "drug metabolism-cytochrome P450" (Score = 18.80, $P = 8.03 \times 10^{-3}$), "*Staphylococcus aureus* infection" (Score = 22.77, $P = 2.24 \times 10^{-3}$), and "cytokine-cytokine receptor interaction" (Score = 90.39, $P = 3.92 \times 10^{-3}$). This is consistent with the finding above and represents the organ function. In the liver, pathways related to oxidative stress, including "nuclear receptors in lipid metabolism and toxicity" (Score = 18.38, $P = 1.59 \times 10^{-3}$), "synthesis of glycosylphosphatidylinositol (GPI)" (Score = 8.53, $P = 3.65 \times 10^{-3}$), and

"valine, leucine and isoleucine degradation" (Score = 31.07, $P = 3.63 \times 10^{-3}$) were significantly enriched. Interestingly, the "valine, leucine and isoleucine degradation" pathway is involved in fatty acid metabolism and immune cell proliferation and is indirectly involved in protein synthesis by affecting the mTOR signaling pathway (*54*). Limiting the intake of these amino acids can extend the lifespan of fruit flies and mice (*55, 56*). These results indicate that mammalian expression profiles are related to organ function.

**Genes with expression level associated with maximum lifespan**

Phylogenetic generalized least squares (PGLS) regression analysis was performed to identify genes showing a correlation with longevity between gene expression and four longevity traits (ML, FTM, MLres, and FTMres) and AW (adult weight) within the phylogenetic framework. The ordinary least squares (OLS), Brownian model, and Ornstein-Uhlenbeck model were used for each gene to determine the best correlation (*17, 19, 57*). Based on resampling, the robustness of the correlation was further evaluated through a two-step verification process (see **Methods** for details), to avoid effects by outliers ($P_{robust}$) or single species ($P_{max}$) (*17, 19, 58*). Genes that met both a $P_{robust} < 0.01$ and $P_{max} < 0.05$ threshold were considered significant (**Table 1** and **Table S6**). Based on the obtained results, the total life-history and longevity variation of mammals explained approximately 3.8–4.4% and 2.7–3.3% of the total variation in gene expression, respectively. This is likely due to the higher number (106) of species analyzed here, compared to a previous report that showed 11%–18% of the inter-species differences explained based on data from 34 mammal species (*16*).

We further used the sum of regression coefficient of each gene to identify pathways showing an enrichment with lifespan (*46, 47*) (**Fig. 2A**, **Fig. S4**, **Tables S7**, and **Table S8**). Significantly enriched pathways related to the

immune system and inflammation were detected in the liver. These included the “inflammatory response pathway” (FTM: Score = 5.54, $P = 8.00 \times 10^{-3}$; FTMres: Score = 4.36, $P = 3.02 \times 10^{-2}$), “regulation of IFNG signaling” (ML: Score = 8.62, $P = 1.90 \times 10^{-3}$; MLres: Score = 7.76, $P = 2.13 \times 10^{-3}$; FTMres: Score = 4.57, $P = 2.22 \times 10^{-2}$), “synthesis of leukotrienes (LT) and eoxins (EX)” (MLres: Score = 5.93, $P = 3.07 \times 10^{-2}$; FTMres: Score = 5.19, $P = 1.32 \times 10^{-2}$), and “prion diseases” (ML: Score = 10.87, $P = 7.98 \times 10^{-3}$; FTM: Score = 6.59, $P = 8.92 \times 10^{-3}$; FTMres: Score = 3.84, $P = 4.32 \times 10^{-2}$) (**Fig. 2A**, **Fig. S4** and **Fig. S5**). This enrichment is consistent with previous studies reporting that individuals with longer lifespans show an improved ability to resist inflammation (*59-62*). Interestingly, some well-known aging processes were enriched by genes whose expression showed a significant correlation with longevity. These included “cellular senescence” (ML: Score = −34.10, $P = 1.52 \times 10^{-3}$; FTM: Score = −23.26, $P = 8.14 \times 10^{-3}$; MLres: Score = −42.47, $P = 2.70 \times 10^{-4}$) (**Figs. 2A-B**, and **Fig. S4**), “direct p53 effectors” (ML: Score = 38.00, $P = 1.05 \times 10^{-2}$; FTM: Score = 24.03, $P = 1.94 \times 10^{-2}$) (**Figs. 2A-C**, and **Fig. S4**), which also found the same trend in two other tissues. And “Regulation of CDC42 activity” (ML: Score = −11.24, $P = 3.17 \times 10^{-3}$; FTM: Score = −4.51, $P = 3.25 \times 10^{-4}$) has the opposite trend in liver and brain (**Fig. 2A, Fig. S4**, and **Fig. S5**). In parallel, pathways related to mitochondrial function were also detected, such as “alpha-linolenic (Ω3) and linoleic (Ω6) acid metabolism” (ML: Score = −6.27, $P = 6.80 \times 10^{-3}$; FTM: Score = −4.00, $P = 2.59 \times 10^{-2}$), and “mitochondrial protein import” (ML: Score = −15.53, $P = 2.95 \times 10^{-3}$; FTM: Score = −10.51, $P = 1.13 \times 10^{-2}$), which were enriched by genes that showed a significant negative correlation with longevity (**Fig. 2A, Fig. S4**, and **Fig. S5**). In line with this observation, the expression of genes related to central energy metabolism was reported to be downregulated in long-lived animals in a previous study of 34 mammals (*16*). The top genes

positively associated with longevity in the liver were *CEP152* (MLres: $P_{robust}$ = $1.84 \times 10^{-3}$; FTMres: $P_{robust} = 1.91 \times 10^{-4}$), *CACYBP* (ML: $P_{robust} = 3.52 \times 10^{-4}$), *LRP8* (FTM: $P_{robust} = 4.27 \times 10^{-4}$), *RASL12* (FTMres: $P_{robust} = 5.53 \times 10^{-4}$) and *SLC25A12* (ML: $P_{robust} = 2.56 \times 10^{-4}$; FTM: $P_{robust} = 7.21 \times 10^{-4}$) (**Table S6**). Some of them are of interest in relation to aging. For example, the loss of *CEP152* function results in increased DNA damage and delayed entry of the S phase (*63*). *CACYBP* promotes autophagy under starvation conditions (*64*), and counteracts oxidative stress, and maintain nucleolus stability (*65*). *SLC25A12* (encodes aspartate-glutamate carrier isoform 1; AGC1) is highly expressed in liver cancer cells, where it promotes cell division (*66*). Among the top genes negatively associated with longevity in the liver (namely, *NRROS, SCAMP4, RBBP7, MRPL2,* and *F2R*) (**Fig. S8A**), NRROS (FTM: $P_{robust} = 1.35 \times 10^{-4}$) is acetylated after *SIRT6* expression is depleted, and the activity of this gene was found to be correlated with lifespan across mammals (*67*). High expression of *SCAMP4* (FTM: $P_{robust} = 4.52 \times 10^{-4}$) can support senescence-associated secretory phenotype and accelerate aging (*68*). *RBBP7* (ML: $P_{robust} = 1.05 \times 10^{-3}$) was found to be downregulated in senescent cells and is one of the causes of increased DNA damage (*69*). *MRPL2* (ML: $P_{robust} = 1.31 \times 10^{-3}$) activates the mitochondrial unfolded protein response and extends lifespan in worms (*70*). Inhibition of *F2R* (ML: $P_{robust}$ = $2.14 \times 10^{-3}$) can also delay aging in worms (*71*).

In the kidney, pathways involved in transcription and translation regulation, including “eukaryotic translation elongation” (FTM: Score = 8.26, $P = 1.67 \times 10^{-2}$; FTMres: Score = 11.09, $P = 1.44\times 10^{-3}$), and “deadenylation of mRNA” (ML: Score = 8.58, $P = 9.26 \times 10^{-3}$; FTM: Score = 6.36, $P = 5.55 \times 10^{-3}$; MLres: Score = 8.95, $P = 1.70 \times 10^{-2}$; FTMres: Score = 6.19, $P = 2.46 \times 10^{-2}$), were enriched by genes that were positively correlated with longevity-related traits (**Fig. 2A**, **Fig S4**, and **Fig. S6**). Interestingly, the “mRNA surveillance

pathway" (MLres: Score = −27.32, $P = 1.60 \times 10^{-3}$) and "tRNA aminoacylation" (FTM: Score = −11.18, $P = 4.89 \times 10^{-3}$), were enriched by genes negatively correlated with longevity (**Fig. 2A**, **Fig. S4**, and **Fig. S6**). It has been reported that the regulation of translation fidelity is one of the factors controlling lifespan in a large range of organisms (*72, 73*). For example, naked mole-rat has higher translational fidelity than mouse (*74*). Transfer RNA aminoacylation is inhibited in senescent cells to limit protein synthesis errors (*75*). In addition, seryl-tRNA can directly bind to telomere repeat sequences, leading to telomere shortening and cell senescence (*76*). However, it is unclear why those genes are highly expressed in short-lived species. The "heme biosynthesis" pathway (ML: Score = −4.41, $P = 3.96 \times 10^{-2}$) was negatively correlated with longevity in the kidney (**Fig. 2A**, **Fig. S4**, and **Fig. S6**). In contrast, the "porphyrin and chlorophyll metabolism" pathway (MLres: Score = 14.34, $P = 1.05 \times 10^{-3}$; FTMres: Score = 8.94, $P = 8.93 \times 10^{-3}$) was positively correlated with longevity in this organ (**Fig. 2A**, **Fig. S4**, and **Fig. S6**). Heme is an iron porphyrin compound that acts as a precursor for hemoglobin, cytochrome and catalase (*77*). The accumulation of cytochrome promotes cell senescence (*78*), and timely metabolized heme can maintain iron homeostasis and reduce ferroptosis stress in the kidney (*79*). The top genes positively associated with longevity in the kidney included *RHOT2* (ML: $P_{robust} = 3.49 \times 10^{-4}$; FTM: $P_{robust} = 1.72 \times 10^{-4}$), *PARN* (ML: $P_{robust} = 4.00 \times 10^{-4}$), and *DACH1* (MLres: $P_{robust} = 4.25 \times 10^{-4}$); while the negatively associated genes included *BHLHE40* (FTM: $P_{robust} = 1.11 \times 10^{-4}$), *DLAT* (FTMres: $P_{robust} = 4.22 \times 10^{-4}$), and *DIXDC1* (MLres: $P_{robust} = 7.11 \times 10^{-4}$) (**Fig. S8B**). *RHOT2* belongs to the Rho family of GTP enzymes, which are involved in mitochondrial transport and autophagy (*80*), and this gene was positively correlated with longevity in a previous cross-species study (*17*). *DACH1* is involved in suppressing tumor metastasis (*81*), and mutations of this gene

were identified in a survey of centenarians (*82*). Among the genes negatively associated with longevity, the overexpression of *BHLHE40* was shown to induce cell senescence, while its knockdown reduced p53-mediated senescence caused by DNA damage (*83, 84*). In addition, BHLHE40 can inhibit fat synthesis mediated by HIF1$\alpha$ under hypoxic conditions (*85*). DLAT can be hydrolyzed by SIRT4, thereby reducing pyruvate dehydrogenase activity and delaying aging (*86, 87*). A study conducted to identify biomarkers of aging, based on RNA-seq and microarray data derived from rats, mice, and humans, showed that the *DLAT* was downregulated during the aging process (*88*).

In the brain, the pathways "metabolism of porphyrins" (MLres: Score = 8.96, $P = 1.52 \times 10^{-3}$; FTMres: Score = 4.33, $P = 3.86 \times 10^{-2}$), "glycosaminoglycan biosynthesis–keratan sulfate" (MLres: Score = 7.57, $P = 2.02 \times 10^{-3}$), "voltage-gated potassium channels" (FTM: Score = 7.34, $P = 4.01 \times 10^{-2}$; FTMres: Score = 7.33, $P = 1.31 \times 10^{-2}$), and "syndecan-4 mediated signaling events" (FTM: Score = 5.53, $P = 1.05 \times 10^{-2}$; MLres: Score = 6.50, $P = 4.86 \times 10^{-2}$; FTMres: Score = 5.71, $P = 4.11 \times 10^{-3}$) were enriched by genes whose expression positively correlated with longevity traits (**Fig. 2A**, **Fig. S4**, and **Fig. S7**). Both "metabolism of porphyrins" and "glycosaminoglycan biosynthesis – keratan sulfate," are associated with clearance of damage in the central nervous system (*89, 90*). Genes in the "voltage-gated potassium channels" pathway have been suggested as potential markers of longevity (*91*), and a deficiency of this pathway results in circadian rhythm disruptions (*92*), shortened lifespan (*93*) and, possibly, obesity (*93, 94*). Among the pathways that were negatively correlated with longevity traits (**Fig. 2A**, **Fig. S4**, and **Fig. S7**), downregulation of the α2,6-linked sialic acid—which functions within the "sialic acid metabolism" pathway (ML: Score = −8.98, $P = 3.08 \times 10^{-2}$; MLres: Score = −13.19, $P = 4.82 \times 10^{-2}$)

has been shown to improve cognitive function in mice (*95*), whereas the knockdown of genes in the "PLK1 signaling events" pathway (ML: Score = −15.78, $P = 9.78 \times 10^{-3}$) induce autophagy, contributing to clearing proteins associated with Alzheimer's and Parkinson's diseases (*96*). Among the genes positively related to longevity in the brain, many respond to oxidative stress, including *AMBRA1* (FTMres: $P_{robust} = 4.77 \times 10^{-4}$), *ATG2A* (ML: $P_{robust} = 3.86 \times 10^{-3}$; FTM: $P_{robust} = 4.72 \times 10^{-4}$), and *MCAT* (MLres: $P_{robust} = 9.89 \times 10^{-4}$) (**Fig. S8C**). Both *AMBRA1* and *ATG2A* regulate autophagy and nervous system development (*97*). In addition, *ATG2A* overexpression prolongs the average lifespan in *Drosophila* (*98*). *MCAT* is able to scavenge reactive oxygen species, and its overexpression is neuroprotective (*99*), and reduces age-related oxidative stress in mitochondria (*100*).

**Comparison of longevity-related genes among tissues and gene sets**

Unexpectedly, only *SRSF4* (serine/arginine-rich splicing actor 4) showed strong correlation with longevity in all three examined tissues (Liver: ML-$P_{robust} = 4.00 \times 10^{-4}$, FTM-$P_{robust} = 1.80 \times 10^{-3}$; Kidney: ML-$P_{robust} = 7.20 \times 10^{-2}$, FTM-$P_{robust} = 2.50 \times 10^{-3}$; Brain: ML-$P_{robust} = 5.80 \times 10^{-3}$, FTM-$P_{robust} = 4.90 \times 10^{-3}$) (**Figs. 2D-G**). This gene is an essential components of spliceosomes, and thus functions in alternative splicing (*101*), and is downregulated during cellular senescence (*102*). Abnormal SRSF4 function is associated with heart disease and reproductive defects (*103, 104*). Expression changes of splicing factors with age have been described in human and other animal models (*105*). Hence, the strong correlation between longevity and *SRSF4* expression across species indicating that stable alternative splicing could be a 'long-lived' feature and that expression level of this gene may be a marker of lifespan. Nevertheless, it is worth noting that the expression of *SRSF4* positively correlated with longevity in the liver and brain, but negatively correlated with

longevity in the kidney, suggesting that this gene serves different roles in lifespan control across tissues. It has also been reported that *SRSF4* is negatively regulated by male sex hormones in mice Sertoli cells (*106*) and the kidney is one of the main target organs of these hormones. However, whether the expression pattern of *SRSF4* across tissue is regulated by male hormones, remains unknown to date. Genes (n = 974) with known effects on longevity of model organisms were also examined, revealing that only a few (n = 76, such as *AAK1*, *KL, NFKB1*,*STRN*, and *TERT* etc.) showed significant correlation with longevity across phylogeny (*1*) (**Fig. 2D**). This suggests that most of these genes do not serve as a basis for the evolution of longevity across species, although they have been shown to directly contribute to lifespan control in one or more model species.

Nevertheless, several pathways showed correlation with longevity traits in three tissues. The positively related pathways include “Insulin signaling pathway” (**Fig. 2H**), "FoxO signaling pathway" (**Fig. 2I**), "Galactose metabolism", and "FAS (CD95) signaling pathway". And the negatively related pathways include "Negative regulation of FGFR signaling pathway", “Aminoacyl-tRNA biosynthesis”, “Cytosolic tRNA aminoacylation”, “tRNA Aminoacylation”, and "Sema3A PAK-dependent axonal rejection". Interestingly, “Insulin signaling pathway” (Liver-MLres: Score = 34.45, $P$ = $3.46 \times 10^{-2}$; Liver-FTMres: Score = 26.58, $P = 4.23 \times 10^{-2}$; Kidney-FTM: Score = 23.40, $P = 3.86 \times 10^{-2}$; Brain-FTM: Score = 27.23, $P = 4.79 \times 10^{-2}$) was positively correlated with lifespan in all three tissues. We also performed PGLS analyses with longevity traits within Primates, Chiroptera, and Rodentia (**Tables S9**-**S11**). The "Insulin signaling pathway" was positively correlated with longevity in the kidneys of primates (Kidney-ML: Score = 142.62, $P$ = $8.03 \times 10^{-3}$; Kidney-FTM: Score = 105.60, $P = 1.86 \times 10^{-2}$; Kidney-FTMres: Score = 255.45, $P = 2.76 \times 10^{-3}$), the livers of bats (Liver-ML: Score = 142.62,

386 $P = 1.38 \times 10^{-3}$; Liver-MLres: Score = 153.41, $P = 2.58\times 10^{-4}$), and the brains
387 of rodents (Brain-FTM: Score = 69.21, $P = 3.86 \times 10^{-2}$; Brain-FTMres: Score
388 = 75.28, $P = 1.74\times 10^{-2}$), but negatively correlated in the kidneys of bats
389 (Kidney-ML: Score = -277.67, $P = 1.16 \times 10^{-2}$; Kidney-MLres: Score = -160.28,
390 $P = 4.99\times 10^{-3}$). Though it is reported that inhibiting insulin signaling at the
391 individual level can prolong lifespan (*9-11*), we found that the expression of
392 insulin signaling pathway was higher in animals with longer lifespan,
393 suggesting maintaining of this pathway is critical for regulating aging process.
394 Pathways, such as “Transport of glucose and other sugars, bile salts and
395 organic acids, metal ions and amine compounds” (Liver-MLres: Score = -
396 29.60, $P = 1.23 \times 10^{-2}$, Liver-FTMres: Score = -28.74, $P = 4.00 \times 10^{-4}$;
397 Kidney-FTMres: Score = 21.43, $P = 1.48 \times 10^{-2}$; Brain-MLres: Score = 23.32,
398 $P = 3.26 \times 10^{-2}$; Brain-FTMres: Score = 17.72, $P = 3.65 \times 10^{-2}$) and
399 “Hemostasis” (Liver-MLres: Score = -107.57, $P = 4.05 \times 10^{-2}$; Kidney-FTM:
400 Score = 75.38, $P = 2.30 \times 10^{-2}$; Brain-FTM: Score = 90.11, $P = 2.03\times 10^{-2}$)
401 positively correlated with longevity traits in the kidney and brain, but shown
402 negative correlation in the liver. “G alpha (s) signalling events” (Liver-ML:
403 Score = -21.35, $P = 1.42 \times 10^{-2}$, Kidney-ML: Score = 21.42, $P = 3.05 \times 10^{-2}$;
404 Kidney-FTM: Score = 15.37, $P = 3.70 \times 10^{-2}$; Brain-FTM: Score = 13.43, $P =$
405 $1.74 \times 10^{-2}$; Brain-FTMres: Score = 21.70, $P = 6.40 \times 10^{-3}$). This is interesting
406 as neither of them showing consistent direction of correlation with longevity
407 traits in three organs. The possible explanation could be the role of pathway in
408 organ aging are different or even opposite. For example, calorie restriction in
409 the kidneys reduces damage from oxidative stress by reducing the activity of
410 certain heme proteins (e.g., Myeloperoxidase, MPO) through anti-
411 inflammatory effects (*107*). However, caloric restriction activates FOXO1 in
412 the liver, which further disrupts mitochondrial function and lipid metabolism via
413 heme (*108*).

To understand the variation of longevity-related genes derived from different genomic sources, we next considered longevity-related genes within metrics that assess the degree of mutation tolerance (essential vs. non-essential genes, and disease-harboring vs. non-disease genes") and reflect fitness (haplo-sufficient vs. haplo-insufficient genes, and new vs. old genes) (see **Methods** for details). The results (**Table S12**) showed that the direction of the coefficients of absolute (ML and FTM) and relative (MLres and FTMres) longevity-related genes was inconsistent in each gene metric, indicating that most correlations were confounded by body weight. For example, among the mutation tolerance metrics, it was observed that essential genes showed a positive correlation in gene expression with the gradient of ML compared to non-essential genes (average coefficients for the essential/non-essential group: liver, 1.25/0.68; kidney, 0.36/0.17; brain, 0.92/0.31) (**Fig. 3**). However, most essential genes whose expression significantly correlated with MLres showed the opposite trend to those that correlated with ML (average coefficients for essential/non-essential group: liver, −0.31/0.02; kidney, −0.43/0.27; brain, −0.60/−0.91) (**Fig. 3**). The coefficients of young genes and disease-harboring genes had a similar pattern compared to their counterparts. In particular, the coefficient direction of the haplo-sufficient gene in the kidney was always opposite to the other two organs. This observation suggests that more attention should be paid to the selection of organs in aging experiments. The results also showed a considerable variation in the coefficient of haplo-sufficient genes among tissues (**Fig. 3**), which indicates that these genes play diverse roles in lifespan across organs.

**Relationship between selection pressure and gene expression**

To uncover the relationship between the intensity of selection and gene expression, we used RELAX in Hyphy (*109*) to assess if selection pressure

relaxed by inferring with the parameter (*k*) (**Table S13**). The relaxation parameter *k* was estimated for each gene with long-lived mammals set as the foreground branch ( $\omega_{\text{background branch}}^{k} = \omega_{\text{foreground branch}}$). A $k > 1$ indicates that a gene in the foreground branch is under intensified selection, while $k < 1$ indicates relaxed selection. The genes whose expression was correlated with longevity were divided into four categories (**Fig. 4A**): positively correlated genes under intensified selection (IU), positively correlated genes under relaxed selection (RU), negatively correlated genes under intensified selection (ID), and negatively correlated genes under relaxed selection (RD). The results showed that approximately 62% of total genes are under intensified selection in long-lived mammals (defined as having an ML > 30 yrs.) (**Figs. 4B-C** and **Table S14**). This is interesting, as a previous study reported that genes tend to be under relaxed selection in shorter-lived killifish (*110*). We also tested what type of selection was present at different maximum-lifespan intervals (ML < 12; 12 <= ML <= 26; 26 < ML; 50 < ML), and the results consistently showed that species with longer lifespans tend to have a greater number of genes under intensive selection (**Fig. S9** and **Tables S15-17**). We then found that the pattern of selection intensity related to the direction of longevity-associated genes is different across organs ( **Figs. 4D-I** and **Fig. S10**). Particularly, in the kidney, a similar proportion of both positively and negatively associated genes was under intensified selection. However, a comparatively higher proportion of positively associated genes—rather than negatively associated genes—was found to be under intensified selection in the liver, while a higher proportion of negatively associated genes was under intensified selection in the brain. It was also observed that several longevity-related genes are under strong intensification of selection in long-lived mammals (**Fig. 4** and **Fig. S10**). For example, among the genes in the IU category, *ICMT* is upregulated in *Drosophila* with extended lifespan (*111*), and

*KCNC4*, which mediates the voltage-dependent potassium ion permeability of excitable membranes, is a marker of longevity (*91, 93*). Within the ID category, the downregulation of *CCND2* promotes cell cycle arrest and apoptosis and leads to cell senescence (*112-114*), while *ZNRF2* is one of the top genes negatively correlated with longevity in the liver tissue and is involved in the mTOR signaling pathway (*115*). In addition, we observed that the longevity correlated genes are characterized by either intensified or relaxed positive selection. In parallel, the strength of purifying selection of those genes has been less changed (**Table S14**).

Pathway enrichment analysis of longevity correlated genes was performed using the relaxation parameter (*k*) as a statistic to further characterize the cumulative effect of relaxed and intensified selection in long-lived mammals (**Fig. S11** and **Table S18-19**). Pathways related to methionine metabolism were detected, such as the “methionine salvage pathway” (Score = 8.41, $P = 7.27 \times 10^{-4}$) and “glycerolipid metabolism” (Score = 21.31, $P = 2.82 \times 10^{-3}$), which are enriched by genes under intensified selection. Methionine is one of the essential amino acids, and similarly to calorie restriction, methionine restriction has been reported to extend lifespan (*116-118*), reverse inflammation, and reduce DNA damage (*118, 119*). Interestingly, our analysis revealed that several pathways associated with acceleration of the aging process are under relaxed selection in long-lived mammals. These pathways included “CD28 dependent PI3K/Akt signaling” (Score = -8.31, $P = 8.69 \times 10^{-3}$), “WNT ligand biogenesis and trafficking” (Score = -10.42, $P = 9.69 \times 10^{-3}$), and the “VEGF signaling pathway” (Score = -7.24, $P = 4.43 \times 10^{-2}$). Inhibition of these pathways has been shown to extend lifespan (*120-123*).

## Conclusions

In this study, a comparative transcriptomics analysis of 106 species representative of diverse families was performed to describe gene expression diversity in mammals. It was found that gene expression was conserved, with the strongest specificity in the brain, compared to the liver and kidney, and that the expression of species-specific genes may reflect adaptive traits (e.g., bats' requirement for long-term memory). In our study, the effect of adult weight on the robustness of longevity-related genes is very limited, and the intersection of longevity-related genes and adult weight-related genes is very small (**Fig. S12A**), which is consistent with previous study (*16, 17*). In addition to very few known aging-related genes (e.g., *TP53*, *LRP8*, *RBBP7*, and *SCAMP4*) that showed correlation with longevity, many new genes whose expression levels significantly correlated with longevity across mammalian phylogeny were identified (e.g., *SRSF4* in all tissues, *SLC26A6*, *ELFN1* and *NEURL1* in the liver, *SCARA3* in the kidney, and *STX5* in the brain) (**Figs. 2D-G** and **Figs. S12B-G**). Enrichment of many well-known aging-related pathways were also observed. These included "Insulin signaling pathway", "FoxO signaling pathway", "inflammatory response pathway," "cellular senescence," and "p53 signaling pathway". Importantly, our study also detected other pathways—such as "eukaryotic translation extension," "non-classical Wnt signaling pathway," "mRNA polyadenylate," "mRNA detection pathway," and "tRNA aminoxylation". This supports reports that a stable protein synthesis (proteostasis) is important for lifespan control (*73, 124*). We found that longer-lived animals always had more genes that under intensified selection than shorter-lived animals. Although it is generally believed that there is a certain correlation between selection pressure and gene expression (*110*), this was not significantly observed in relation to the gradient of longevity, suggesting lifespan is not favored directly by natural selection (*125*).

We still found that methionine restriction is under Intensified selection in long-live mammals. However, given most of our samples are matured males, the variation of gene expression among different ages and genders have not been quantified in our study. And, because of the high-quality genome assembly are not available for many species that analyzed, it is difficult to disentangle the role of the expression changes of paralogous and multi-copy genes in longevity evolution. Nevertheless, the data and results presented in this study would aid future investigations with inclusion of samples at different ages of both genders. Overall, our study suggests that the evolutionary correlation between gene expression and longevity is organ-specific and characterized by polygenic selection. The longevity-associated genes identified could serve as candidate targets for further exploration of healthy aging.

## Materials and Methods

### Tissue collection

The 331 organ samples analyzed in the present study were newly obtained from 56 species (**Tables S1** and **S2**). Liver, kidney and brain tissues were mostly sampled from adult and male individuals, if possible, and were freshly frozen in liquid $N_2$ and stored at −80°C. To maximize sample compatibility, each major part of each organ was dissected and homogenized. To objectively detect the biological variation of gene expression, three biological replicates were obtained when possible. All the experimental protocols were approved by the Animal Care and Use Committee of the Institute of Zoology, Chinese Academy of Sciences (No. IOZ-IACUC-2021-129).

### Transcriptome library preparation and sequencing

Total RNA was isolated from frozen tissue using TRIzol® Reagent (Invitrogen). To protect RNA as much as possible during homogenization, we first added

555 0.2 mL TRIzol® Reagent directly to a tube containing 100 mg of frozen tissue
556 and homogenized using a motorized homogenizer. After homogenization, we
557 added another 0.8 mL of TRIzol® to the tube. The resulting lysate was phase
558 separated with 0.2 ml chloroform and total RNA precipitated with 0.5 ml
559 isopropanol. The RNA was washed twice with 1 ml 75% ethanol and
560 resuspended in DEPC treated ddH2O. The resuspended RNA was assessed
561 for quality (260/280 nm absorbance ratio) and integrity (formaldehyde agarose
562 gel electrophoresis). The sequencing libraries were prepared using the
563 NEBNext Ultra RNA Library Prep Kit for Illumina (NEB, USA), and the
564 transcriptome libraries were sequenced on an Illumina NovaSeq 6000 system
565 (Novogene Co. Ltd) with paired-end reads of 150 bp. NGS QC Toolkit v2.3.3
566 (*126*) was used to remove reads containing adapters and filter low quality
567 reads (<Q20).
568
569 **Orthologous gene sets**
570 Genome annotations (GTF) for 39 mammals with sequenced genomes were
571 obtained from Ensembl, release 99. For the minke whale (*Balaenoptera*
572 *acutorostrata*), Indian muntjac (*Muntiacus muntjac*), great roundleaf bat
573 (*Hipposideros armiger*), Chinese rufous horseshoe bat (*Rhinolophus sinicus*),
574 Brandt's bat (*Myotis brandtii*), François's leaf Monkey (*Trachypithecus*
575 *francoisi*), and white-footed mouse (*Peromyscus leucopus*) we used GTF
576 annotations and genomes downloaded from NCBI database. For the bowhead
577 whale (*Balaena mysticetus*) we used genomes downloaded from 'The
578 Bowhead Whale Genome Resource' (*127*) (**Table S2**). The GTF of bowhead
579 whale was generated using augustus v 2.5.5 (*128*). Draft transcriptome of 59
580 mammals were *de novo* assembled using trinity v2.11.0 (*129*). First, the RNA-
581 seq reads from same species and tissues were assembled together. Because
582 the Trinity assembler filters low-coverage k-mers, we did not perform quality

583 trim of the reads before assembly. Trinity was run on 150 bp paired-end
584 sequences with default parameters k-mer size of 25 (fixed), minimum contig
585 length of 200, maximum paired fragment length of 500, and adjusted butterfly
586 maximum heap space setting to 30G. To remove redundancy, we then used
587 cd-hit (*130, 131*) to process the assembled transcripts from different tissues
588 of the same species, cluster the sequences with 90% similarity, and leave the
589 longest transcript in each cluster. We used augustus to perform gene
590 prediction on the de-redundant transcripts and obtain GTF annotation files.
591 We used gffread in the cufflink package v2.2.1 (*132*) to extract the CDS
592 sequence, filtered out incomplete ORF transcripts and pseudogene transcripts,
593 and extracted the longest transcript of each gene. Given the genome
594 assembly for most of species are scarce or not well-annotated, multi-copy
595 genes and transcripts were not considered in the analyses. To reduce the
596 effects of paralogs on the ortholog identification, we constructed the human
597 reference sequence using BLAST v2.9.0+ (*133*) to remove highly repetitive
598 and highly similar genes, with e-value < $10^{-6}$ and Identity > 90% as the filtering
599 threshold. In the end, 18,553 unique protein coding genes were obtained as
600 reference sequences. For other mammals, the longest transcript of each gene
601 was extracted and reciprocal BLAST was performed with the protein
602 sequences from human. The filtering threshold was $10^{-6}$ for e-value and 30%
603 for identity. Two genes that were best aligned with each other were defined as
604 orthologous genes. When a gene exists in fewer species, it indicates that the
605 gene is not highly conserved and cannot be representative of mammals.
606 However, as the number of species increases, the number of orthologous
607 genes that coexist in all species decreases (only 989 genes are present in all
608 species). To balance the number of species and genes, we filtered out genes
609 that exist in less than 70 species. The final dataset of orthologous gene
610 accounted for 13,916 individual groups of sequences. In downstream

611 analysis, each gene is analyzed individually, and only the species in which the
612 gene is present were considered.
613
614 **RNA-seq reads mapping and normalization**
615 Because the complete genome and the *de novo* genome are quite different
616 when compared, we used the CDS sequence of orthologous genes as the
617 reference genome, and generate annotation files in GTF format for RNA-seq
618 data mapping. STAR v2.7.1a (*134*) was used to construct an index. Because
619 of the specificity of the orthologous genomes, the parameters '--
620 genomeSAindexNbases' and '--genomeChrBinNbits' were calculated from the
621 sequence size of the homologous gene set and the read length of different
622 samples. And we used the default parameters to align the RNA-seq data with
623 the orthologous genomes. We used featureCounts v2.0.0 (*135*) to count reads,
624 and eliminate multiple-matched reads (**Table S20)**. Generating gene
625 expression profiles for all species based on pairwise orthologous relationships.
626 Finally for 18,553 genes, abnormally low-expressed genes that is, genes
627 whose expression levels were less than 10 in 4 or more samples were filtered
628 before normalization (2,564 genes were removed). And, abnormally high
629 expression genes, that is, genes whose total expression of all samples
630 accounted for 5% of the expression of the entire data set was also removed (1
631 gene). The function comBat_seq in the R package sva (*136*) was applied to
632 read counts to remove the batch effect, including the two factors most likely to
633 affect the data: different sources of data (Bioproject and sequencing batches
634 of our data) and the deviation caused by the sequencing platform (*137*). Two
635 factors were adjusted separately. And the covariates were tissue and species.
636 Genes with orthologous in more than 70 species were used for downstream
637 analysis to reduce the false positive rate in the analysis (13,827 genes in
638 total). We calculated the library size of each sample as a normalization factor.

The R software package edgeR (*138*) was used to normalize the library size and gene length (based on humans) by $\log_2$(TMM-RPKM + 1). For paired-end data, featureCounts counts fragments, so calculating RPKM for paired-end data is equivalent to FPKM.

**PCA analysis and species specificity of gene expression.**

We calculated the variance of each gene on the normalized expression matrix, and selected the top 5,000 genes with the largest variance to perform principal component analysis (PCA) using the R package 'FactoMineR' (*139*). In order to define gene sets that are widely expressed by species and species-specifically expressed genes. We calculated the mean value of log2(RPKM-TMM + 1) in each organ for each species. We calculated the species-specific expression index Tau, $\tau = \frac{\sum_{i=0}^{N}(1-x_i)}{N-1}$, which is used to quantify the tissue specificity of gene expression (*36*). Among them, N is the number of species, $x_i$ is the expression level of the i-th species. Tau > 0.8 is defined as a species-specific gene.

**Life-history data collection and imputation.**

To accurately estimate the species for which life history data were missing in this study. We collected data on highly correlated life-history traits (AW: adult weight, ML: maximum lifespan, and FTM: female mature time) for a total of 1,250 species from the online databases AnAge (*1*), Animal Diversity Web (https://animaldiversity.org/) and PanTHERIA (*140*), and from the literature. And the phylogenetic tree was retrieved from TimeTree (http://www.timetree.org/) (*141*). Three life-history traits from 816 species were complete and used as a training set and, we employed three imputation methods to estimate the missing data: (i) Based on the Markov chain Monte Carlo method, *mice* introduces the random process into the interpolation

process, uses other variables as predictors, and specifies a conditional model for each variable (*142*). We used the predictive mean matching (pmm) as the conditional model in the multiple regression model, or used mean matching (mean) instead if the first run did not converge. We selected the case where the predicted regression score was closest to the missing value. (ii) missForest first uses the mean to interpolate a column of data (*143*) and then uses the remaining variables of the data set to fit a random forest model to estimate missing values by applying trained random forest predictions. This process was looped for all variables that need to be interpolated, and the whole process was repeated until the stop criterion was reached. (iii) Phylopars estimates missing values based on restricted maximum likelihood (*144*). This method calculates the covariance matrix based on phylogenetic and phenotypic components (when multiple trait measures are given). It builds a multivariate normal model that combines the best phylogenetic and phenotypic covariance with the tree to calculate the covariance between the observed and missing values. Therefore, the estimated value was determined by phylogenetic distance (correlation between species) and ectopic relationship (correlation between features).

The percentages of missing values in the missing set were 6% (ML) and 30% (FTM). We tested three types of missing data: (i) Completely missing at random (MCAR); (ii) Missing at random based on weight (MAR.AW), and it was divided into two types of species according to the median weight. Because low-weight species may have more missing values; (iii) Missing at random based on the genetic distance between human (MAR.HD), and it was divided into two types of species according to the half of the farthest genetic distance. Because species with a greater genetic distance from humans may receive less attention from scientists, the life-history is also opted to be missed.

We performed chi-square tests on the two types of MARs in the missing set. In addition, missing values in large-weight species accounted for 17.57% of the total missing values in maximum lifespan (ML), and 82.43% in small-weight species. In FTM, the missing values of the large-weight species accounted for 37.66% of the total missing values, and the small-weight species 62.34%. We introduced multiple missing value ratios (5%, 10%, 15%, 20%, 25%, 30%, 40% and 50%) to the training set to simulate the distribution and pattern of missing values. We used the above three methods for 10 interpolations and imputation. In order to account phylogenetic relationships in the imputation process (Phylopars are only applicable to imputations that include phylogeny), R package PVR was used (*145*) to perform principal coordinate analysis (PCoA) on the genetic distance matrix of 816 species to obtain phylogenetic feature vectors. The phylogenetic relationship after dimensionality reduction is used as other predictor variables in the imputation process. At the same time, we obtained the optimal number for interpolation by adding phylogenetic vectors in the interpolation process incrementally. We evaluated the accuracy of the interpolation based on the normalized root mean square error (NMRSE) as $\mathrm{NRMSE} = \frac{\sqrt{\mathrm{mean}\left((Ximp - Xtrue)^2\right)}}{\max(Ytrue) - \min(Ytrue)}$.

And, in order to ensure that the estimated value retains biological significance. We also calculated the bias of the slope ($\mathrm{Bias} = |\mathrm{Slope}^{original} - \mathrm{Slope}^{imputed\ (or\ missing)}|$ ) between adult weight and maximum lifespan in the data set after interpolation.

Finally, we selected the best Phylopars based on the evaluation results for the imputation of the complete life history data set (**Table 2**). To identify confounding factors of maximum lifespan, we collected multiple complex effects such as society, diet, habitat, activity, body mass, basal metabolism rate, and offspring per year. We have used MCMCglmm to examine the

722 correlation between longevity and other factors. The result showed that body
723 mass and offspring per year were significantly associated with maximum
724 lifespan, and a weak association between diet and maximum lifespan (**Table**
725 **S22**). And we also detected a strong correlation between body mass and
726 offspring per year. Previous studies also have shown that the longevity (or
727 female time to maturity) was mainly correlated with body mass (*16*).
728 Nevertheless, many species show with small weight and long maximum
729 lifespan (or female time to maturity). Therefore, we calculated the residuals to
730 correct the confounding effects caused by weight (i.e., MLres and FTMres).
731 Both residual equations are obtained based on linear regression model using
732 the data from the AnAge database (*1*).
733
734 **Phylogenetic regression analysis.**
735 To identify genes with the expression related to longevity, three evolution
736 models were tested for gene expression in each tissue (mean value of $\log_2$-
737 scaled TMM-RPKM) and each $\log_2$-scaled longevity-related trait (ML, FTM,
738 MLres, FTMres), including regression models that do not consider
739 phylogenetic relationships (OLS), and regression models that consider
740 phylogenetic relationships (BM and OU). And the optimal model was selected
741 according to the maximum likelihood methodology. The phylogenetic tree was
742 retrieved from TimeTree (*141*). The unit of branches length of the
743 phylogenetic tree is million years. To avoid randomness, we took a resampling
744 approach (*58*) instead of using conventional P-value corrections (e.g., BH) . A
745 two-step method is used to correct the *P* value (*17, 19*). In the first step, the
746 species that has the greatest impact on the slope (i.e., potential outliers) is
747 removed by the residuals (the largest absolute value of the residual is
748 removed), and then the regression is performed again. At this time, the *P*
749 value obtained is defined as $P_{robust}$ to remove the influence of the outliers on

the regression. The second step is to repeat the regression process for the remaining species and remove one of remaining species each time until all remaining species are removed once, and take the largest (least significant) *P* value in the process as $P_{max}$ to remove the impact of species on regression. The cutoff for identifying longevity-related genes was $P_{max}$<0.05, $P_{robust}$<0.01.

To reduce the noise caused by missing data, for each gene, we only consider the species in which the gene exists, and did not add all species to the model, that is, set the expression value of a gene that does not exist in a species to the missing value instead 0.

**Multiple sequence alignment and selection pressure analysis.**

For each group of orthologous genes, the Perl script 'translatorX.pl' (*146*) is used for multiple sequence processing and comparison. This pipeline selects the default parameters of the 'MAFFT' (*147*) to first translate the nucleic acid sequence into a protein sequence for multiple sequence alignment and then translate it back into a nucleic acid sequence. We then used 'GBlock' (*148, 149*) to select the conservative blocks, with the number of conservative sites in the gene sequence after alignment is at least 75% of the total length of the gene, and the shortest flanking sequence is greater than 85% of the length of the gene after the alignment.

In order to test whether genes are under relaxed selection, we used the minimal model of RELAX (*109*) in 'Hyphy', with long-lived animals (ML > 30 years) were set as foreground branches. Because the number of non-long-lived mammals (n=82) is far more than that of long-lived mammals (n=24). We selected 24 representative species with good genome quality from non-long-lived mammals as background branches to eliminate the noise caused by the excessive number of background branch species (**Table S21**). This model uses the likelihood ratio test to compare the two models with the same

778 evolution rate (k = 1) and different rates (k ≠ 1) between the foreground
779 branch and the background branch. The parameters are set to estimate 3
780 types of ω (ω1: purification selection; ω2: neutral selection; ω3: positive
781 selection). The relaxation parameter k is an index of the selection strength
782 ($\omega_{\text{background branch}}^{k} = \omega_{\text{foreground branch}}$), with k > 1 indicates that the genes in the
783 foreground branch are under intensified selection and k < 1 indicates a
784 relaxed selection. And we also test relaxed selection at different interval (ML
785 < 12; 12 <= ML <= 26; 26 < ML; 50 < ML) of ML by use 57 mammals which
786 have real life-history traits data (**Table S15-17**).
787
788 **Gene set enrichment analysis**
789 We use 'Polysel' (*46, 47*) for gene set enrichment analysis which is possible
790 to detect pathways containing pleiotropic signals. In addition, other variables
791 (such as gene length, number of species, and genetic distance) can also be
792 used to adjust statistical variables. For the species-specific expression, we
793 used the species-specific expression index (Tau) as the gene score
794 (SUMSTAT) to detect species-specific expression pathways and ubiquitous
795 expression (1 - Tau) pathways. For gene expression variation, we used the
796 coefficients in the PGLS regression as SUMSTAT to enrich genes that are
797 positively related to longevity (SUMSTAT of negatively related genes is set to
798 0) and negatively related genes (SUMSTAT of positively related genes is set
799 to 0 and converted to Absolute value) to detect longevity-related pathways
800 with genetic minor effects. Since gene expression is mostly related to gene
801 length or species number, we used the function 'RescaleBins' to adjust
802 SUMSTAT. We used 'ks.test' in R to check whether the gene score
803 (SUMSTAT) is normally distributed or not. If not, a random data set was
804 generated to construct an empirical distribution.
805

## Gene category collection

We collected different types of gene sets from various sources for comparison. The essential genes were constructed based on the probability of intolerance to loss of function, which is the pLI score (*150*). The score data comes from ExAC version 0.3.1 (https://gnomad.broadinstitute.org/). Genes with pLI> 0.9 are defined as essential genes. The list of genes associated with human inherited disease was obtained from the manually curated HGMD (PRO 17.1) (*151*). Aging genes were obtained from the GenAge database (*1*) and determined based on experimental evidence from humans and model organisms. They included genes related to the basic human aging process as well as genes related to lifespan. According to the homology relationship, the respective gene ID numbers were converted into human gene ID numbers. The Haploid Insufficiency (HI) score from previous studies (*152*) was used to quantify the degree of haploid deficiency in human genes. After sorting in descending order, we defined genes greater than the first quartile as haplo-insufficient genes, and genes less than the fourth quartile as haplo-sufficient genes. Finally, the phylogenetic age of mammalian genes was retrieved from the GenTree database (http://gentree.ioz.ac.cn/) (*153*). We divided genes into two groups based on genetic age: (i) those genes that appear after therian, mammalian, vertebrate, or quadrupedal ancestors (genes are defined as relatively young) and (ii) those that appear earlier than bone vertebrates Genes (defined as relatively old genes).

## Data and code availability

Raw RNA-seq data for liver and kidney of this study are available from Sciencedb with doi number 10.11922/sciencedb.01196. Raw RNA-seq data for brain are available from Sciencedb with doi number 10.11922/sciencedb.01197. The SRA ids of the data retrieved from NCBI are

834 listed in **Table S1**. The code of this study is available from github
835 （https://github.com/liu-wq/expressionML）.
836
837 **Acknowledgements**
838 We thank Alice Hughes, Xing Chen and Yanhua Chen for the experience and
839 technology of sample collection. We thank Pengcheng Wang, Xiaoxiao Zhang,
840 Zhan Zhang, and Qi Pan for supporting sample collection and preparation.
841 This project was supported by National Natural Science Foundation of China
842 (82050002) and the Beijing Natural Sciences Foundation (JQ19022). VNG is
843 supported by grants from the National Institute on Aging.
844
845 **Author contributions**
846 X.Z. conceived the study and designed the project. W.L. and P.Z. managed
847 the project. W.L., P.Z., M.L., Z.L., Y.Y., G.L., X.J. and X.W. collected samples.
848 W.L., P.Z., M.L., Z.L., Y.Y., J.D. and J.Y. prepared samples and performed
849 RNA extraction. W.L. performed transcriptome assembly, annotation and
850 bioinformatics analysis. W.L. and P.Z. performed life-history collection and
851 imputation. W.L. and P.Z. discussed the data. W.L. wrote the manuscript with
852 contributions from X.Z., M.L., V.N.G., I.S., L.W. and A.K.. All authors
853 contributed to data interpretation.
854
855 **Competing interests**
856 The authors declare no competing interests.
857
858
859
860
861
862
863
864

# Reference


1. R. Tacutu, T. Craig, A. Budovsky, D. Wuttke, G. Lehmann, D. Taranukha, J. Costa, V. E. Fraifeld, J. P. de Magalhaes, Human Ageing Genomic Resources: integrated databases and tools for the biology and genetics of ageing. *Nucleic Acids Reseach* **41**, D1027-1033 (2013).
2. I. Seim, S. Ma, X. Zhou, M. V. Gerashchenko, S. G. Lee, R. Suydam, J. C. George, J. W. Bickham, V. N. Gladyshev, The transcriptome of the bowhead whale Balaena mysticetus reveals adaptations of the longest-lived mammal. *Aging (Albany NY)* **6**, 879-899 (2014).
3. M. Keane, J. Semeiks, A. E. Webb, Y. I. Li, V. Quesada, T. Craig, L. B. Madsen, S. van Dam, D. Brawand, P. I. Marques, P. Michalak, L. Kang, J. Bhak, H.-S. Yim, N. V. Grishin, N. H. Nielsen, M. P. Heide-Jørgensen, E. M. Oziolor, C. W. Matson, G. M. Church, G. W. Stuart, J. C. Patton, J. C. George, R. Suydam, K. Larsen, C. López-Otín, M. J. O'Connell, J. W. Bickham, B. Thomsen, J. P. de Magalhães, Insights into the evolution of longevity from the bowhead whale genome. *Cell reports* **10**, 112-122 (2015).
4. E. B. Kim, X. Fang, A. A. Fushan, Z. Huang, A. V. Lobanov, L. Han, S. M. Marino, X. Sun, A. A. Turanov, P. Yang, S. H. Yim, X. Zhao, M. V. Kasaikina, N. Stoletzki, C. Peng, P. Polak, Z. Xiong, A. Kiezun, Y. Zhu, Y. Chen, G. V. Kryukov, Q. Zhang, L. Peshkin, L. Yang, R. T. Bronson, R. Buffenstein, B. Wang, C. Han, Q. Li, L. Chen, W. Zhao, S. R. Sunyaev, T. J. Park, G. Zhang, J. Wang, V. N. Gladyshev, Genome sequencing reveals insights into physiology and longevity of the naked mole rat. *Nature* **479**, 223-227 (2011).
5. I. Seim, X. Fang, Z. Xiong, A. V. Lobanov, Z. Huang, S. Ma, Y. Feng, A. A. Turanov, Y. Zhu, T. L. Lenz, M. V. Gerashchenko, D. Fan, S. Hee Yim, X. Yao, D. Jordan, Y. Xiong, Y. Ma, A. N. Lyapunov, G. Chen, O. I. Kulakova, Y. Sun, S.-G. Lee, R. T. Bronson, A. A. Moskalev, S. R. Sunyaev, G. Zhang, A. Krogh, J. Wang, V. N. Gladyshev, Genome analysis reveals insights into physiology and longevity of the Brandt's bat Myotis brandtii. *Nature communications* **4**, 2212-2212 (2013).
6. M. Sulak, L. Fong, K. Mika, S. Chigurupati, L. Yon, N. P. Mongan, R. D. Emes, V. J. Lynch, TP53 copy number expansion is associated with the evolution of increased body size and an enhanced DNA damage response in elephants. *Elife* **5**, e11994 (2016).
7. R. P. Perez, T. Komiya, TP53 Gene and Cancer Resistance in Elephants. *JAMA* **315**, 1789-1790 (2016).
8. D. L. Stern, V. Orgogozo, The loci of evolution: how predictable is genetic evolution? *Evolution; international journal of organic evolution* **62**, 2155-2177 (2008).
9. M. Tatar, A. Kopelman, D. Epstein, M. P. Tu, C. M. Yin, R. S. Garofalo, A mutant Drosophila insulin receptor homolog that extends life-span and impairs neuroendocrine function. *Science* **292**, 107-110 (2001).
10. D. B. Friedman, T. E. Johnson, A mutation in the age-1 gene in Caenorhabditis elegans lengthens life and reduces hermaphrodite fertility. *Genetics* **118**, 75-86 (1988).
11. M. Holzenberger, J. Dupont, B. Ducos, P. Leneuve, A. Géloën, P. C. Even, P. Cervera, Y. Le Bouc, IGF-1 receptor regulates lifespan and resistance to oxidative stress in mice. *Nature* **421**, 182-187 (2003).
12. R. A. Miller, D. E. Harrison, C. M. Astle, E. Fernandez, K. Flurkey, M. Han, M. A. Javors, X. Li, N. L. Nadon, J. F. Nelson, S. Pletcher, A. B. Salmon, Z. D. Sharp, S. Van Roekel, L. Winkleman, R.

908 Strong, Rapamycin-mediated lifespan increase in mice is dose and sex dependent and
909 metabolically distinct from dietary restriction. *Aging Cell* **13**, 468-477 (2014).
910 13. D. E. Harrison, R. Strong, Z. D. Sharp, J. F. Nelson, C. M. Astle, K. Flurkey, N. L. Nadon, J. E.
911 Wilkinson, K. Frenkel, C. S. Carter, M. Pahor, M. A. Javors, E. Fernandez, R. A. Miller, Rapamycin
912 fed late in life extends lifespan in genetically heterogeneous mice. *Nature* **460**, 392-395 (2009).
913 14. C. J. Kenyon, The genetics of ageing. *Nature* **464**, 504-512 (2010).
914 15. Y. Kanfi, S. Naiman, G. Amir, V. Peshti, G. Zinman, L. Nahum, Z. Bar-Joseph, H. Y. Cohen, The
915 sirtuin SIRT6 regulates lifespan in male mice. *Nature* **483**, 218-221 (2012).
916 16. A. A. Fushan, A. A. Turanov, S. G. Lee, E. B. Kim, A. V. Lobanov, S. H. Yim, R. Buffenstein, S. R.
917 Lee, K. T. Chang, H. Rhee, J. S. Kim, K. S. Yang, V. N. Gladyshev, Gene expression defines
918 natural changes in mammalian lifespan. *Aging Cell* **14**, 352-365 (2015).
919 17. S. Ma, A. Upneja, A. Galecki, Y. M. Tsai, C. F. Burant, S. Raskind, Q. Zhang, Z. D. Zhang, A.
920 Seluanov, V. Gorbunova, C. B. Clish, R. A. Miller, V. N. Gladyshev, Cell culture-based profiling
921 across mammals reveals DNA repair and metabolism as determinants of species longevity. *Elife* **5**,
922 (2016).
923 18. D. Brawand, M. Soumillon, A. Necsulea, P. Julien, G. Csardi, P. Harrigan, M. Weier, A. Liechti, A.
924 Aximu-Petri, M. Kircher, F. W. Albert, U. Zeller, P. Khaitovich, F. Grutzner, S. Bergmann, R. Nielsen,
925 S. Paabo, H. Kaessmann, The evolution of gene expression levels in mammalian organs. *Nature*
926 **478**, 343-348 (2011).
927 19. S. Ma, S. H. Yim, S. G. Lee, E. B. Kim, S. R. Lee, K. T. Chang, R. Buffenstein, K. N. Lewis, T. J.
928 Park, R. A. Miller, C. B. Clish, V. N. Gladyshev, Organization of the Mammalian Metabolome
929 according to Organ Function, Lineage Specialization, and Longevity. *Cell Metab* **22**, 332-343 (2015).
930 20. M. Cardoso-Moreira, J. Halbert, D. Valloton, B. Velten, C. Chen, Y. Shao, A. Liechti, K. Ascencao,
931 C. Rummel, S. Ovchinnikova, P. V. Mazin, I. Xenarios, K. Harshman, M. Mort, D. N. Cooper, C.
932 Sandi, M. J. Soares, P. G. Ferreira, S. Afonso, M. Carneiro, J. M. A. Turner, J. L. VandeBerg, A.
933 Fallahshahroudi, P. Jensen, R. Behr, S. Lisgo, S. Lindsay, P. Khaitovich, W. Huber, J. Baker, S.
934 Anders, Y. E. Zhang, H. Kaessmann, Gene expression across mammalian organ development.
935 *Nature*, (2019).
936 21. K. Guschanski, M. Warnefors, H. Kaessmann, The evolution of duplicate gene expression in
937 mammalian organs. *Genome Res* **27**, 1461-1474 (2017).
938 22. I. Sarropoulos, R. Marin, M. Cardoso-Moreira, H. Kaessmann, Developmental dynamics of
939 lncRNAs across mammalian organs and species. *Nature* **571**, 510-514 (2019).
940 23. Z. Y. Wang, E. Leushkin, A. Liechti, S. Ovchinnikova, K. Mossinger, T. Bruning, C. Rummel, F.
941 Grutzner, M. Cardoso-Moreira, P. Janich, D. Gatfield, B. Diagouraga, B. de Massy, M. E. Gill, A.
942 Peters, S. Anders, H. Kaessmann, Transcriptome and translatome co-evolution in mammals.
943 *Nature*, (2020).
944 24. Z. Huang, C. V. Whelan, N. M. Foley, D. Jebb, F. Touzalin, E. J. Petit, S. J. Puechmaille, E. C.
945 Teeling, Longitudinal comparative transcriptomics reveals unique mechanisms underlying extended
946 healthspan in bats. *Nat Ecol Evol*, (2019).
947 25. P. Zhu, W. Liu, M. Li, G. Liu, Y. Yu, Z. Li, X. Li, J. Du, X. Wang, C. C. Grueter, M. Li, X. Zhou,
948 Transcriptome signatures of co-evolution between social organization and lifespan in mammals. **(In**
949 **preparation)**, (2021).

26. X. Fang, I. Seim, Z. Huang, M. V. Gerashchenko, Z. Xiong, A. A. Turanov, Y. Zhu, A. V. Lobanov, D. Fan, S. H. Yim, X. Yao, S. Ma, L. Yang, S. G. Lee, E. B. Kim, R. T. Bronson, R. Šumbera, R. Buffenstein, X. Zhou, A. Krogh, T. J. Park, G. Zhang, J. Wang, V. N. Gladyshev, Adaptations to a subterranean environment and longevity revealed by the analysis of mole rat genomes. *Cell Rep* **8**, 1354-1364 (2014).
27. M. Martínez-Pacheco, M. Tenorio, L. Almonte, V. Fajardo, A. Godínez, D. Fernández, P. Cornejo-Páramo, K. Díaz-Barba, J. Halbert, A. Liechti, T. Székely, A. O. Urrutia, D. Cortez, Expression Evolution of Ancestral XY Gametologs across All Major Groups of Placental Mammals. *Genome Biol Evol* **12**, 2015-2028 (2020).
28. J. Chen, R. Swofford, J. Johnson, B. B. Cummings, N. Rogel, K. Lindblad-Toh, W. Haerty, F. D. Palma, A. Regev, A quantitative framework for characterizing the evolutionary history of mammalian gene expression. *Genome Res* **29**, 53-63 (2019).
29. Q. Tang, Y. Gu, X. Zhou, L. Jin, J. Guan, R. Liu, J. Li, K. Long, S. Tian, T. Che, S. Hu, Y. Liang, X. Yang, X. Tao, Z. Zhong, G. Wang, X. Chen, D. Li, J. Ma, X. Wang, M. Mai, A. Jiang, X. Luo, X. Lv, V. N. Gladyshev, X. Li, M. Li, Comparative transcriptomics of 5 high-altitude vertebrates and their low-altitude relatives. *Gigascience* **6**, 1-9 (2017).
30. G. Yan, G. Zhang, X. Fang, Y. Zhang, C. Li, F. Ling, D. N. Cooper, Q. Li, Y. Li, A. J. van Gool, H. Du, J. Chen, R. Chen, P. Zhang, Z. Huang, J. R. Thompson, Y. Meng, Y. Bai, J. Wang, M. Zhuo, T. Wang, Y. Huang, L. Wei, J. Li, Z. Wang, H. Hu, P. Yang, L. Le, P. D. Stenson, B. Li, X. Liu, E. V. Ball, N. An, Q. Huang, Y. Zhang, W. Fan, X. Zhang, Y. Li, W. Wang, M. G. Katze, B. Su, R. Nielsen, H. Yang, J. Wang, X. Wang, J. Wang, Genome sequencing and comparison of two nonhuman primate animal models, the cynomolgus and Chinese rhesus macaques. *Nat Biotechnol* **29**, 1019-1023 (2011).
31. F. N. Carelli, A. Liechti, J. Halbert, M. Warnefors, H. Kaessmann, Repurposing of promoters and enhancers during mammalian evolution. *Nat Commun* **9**, 4066 (2018).
32. Y. Fan, Z. Y. Huang, C. C. Cao, C. S. Chen, Y. X. Chen, D. D. Fan, J. He, H. L. Hou, L. Hu, X. T. Hu, X. T. Jiang, R. Lai, Y. S. Lang, B. Liang, S. G. Liao, D. Mu, Y. Y. Ma, Y. Y. Niu, X. Q. Sun, J. Q. Xia, J. Xiao, Z. Q. Xiong, L. Xu, L. Yang, Y. Zhang, W. Zhao, X. D. Zhao, Y. T. Zheng, J. M. Zhou, Y. B. Zhu, G. J. Zhang, J. Wang, Y. G. Yao, Genome of the Chinese tree shrew. *Nat Commun* **4**, 1426 (2013).
33. Q. Qiu, G. Zhang, T. Ma, W. Qian, J. Wang, Z. Ye, C. Cao, Q. Hu, J. Kim, D. M. Larkin, L. Auvil, B. Capitanu, J. Ma, H. A. Lewin, X. Qian, Y. Lang, R. Zhou, L. Wang, K. Wang, J. Xia, S. Liao, S. Pan, X. Lu, H. Hou, Y. Wang, X. Zang, Y. Yin, H. Ma, J. Zhang, Z. Wang, Y. Zhang, D. Zhang, T. Yonezawa, M. Hasegawa, Y. Zhong, W. Liu, Y. Zhang, Z. Huang, S. Zhang, R. Long, H. Yang, J. Wang, J. A. Lenstra, D. N. Cooper, Y. Wu, J. Wang, P. Shi, J. Wang, J. Liu, The yak genome and adaptation to life at high altitude. *Nat Genet* **44**, 946-949 (2012).
34. M. V. Westbury, B. Petersen, E. D. Lorenzen, Genomic analyses reveal an absence of contemporary introgressive admixture between fin whales and blue whales, despite known hybrids. *PLoS One* **14**, e0222004 (2019).
35. X. Peng, J. Thierry-Mieg, D. Thierry-Mieg, A. Nishida, L. Pipes, M. Bozinoski, M. J. Thomas, S. Kelly, J. M. Weiss, M. Raveendran, D. Muzny, R. A. Gibbs, J. Rogers, G. P. Schroth, M. G. Katze, C. E. Mason, Tissue-specific transcriptome sequencing analysis expands the non-human primate reference transcriptome resource (NHPRTR). *Nucleic Acids Res* **43**, D737-742 (2015).

36. I. Yanai, H. Benjamin, M. Shmoish, V. Chalifa-Caspi, M. Shklar, R. Ophir, A. Bar-Even, S. Horn-Saban, M. Safran, E. Domany, D. Lancet, O. Shmueli, Genome-wide midrange transcription profiles reveal expression level relationships in human tissue specification. *Bioinformatics* **21**, 650-659 (2005).
37. A. Copani, V. M. G. Bruno, V. Barresi, G. Battaglia, D. F. Condorelli, F. Nicoletti, Activation of Metabotropic Glutamate Receptors Prevents Neuronal Apoptosis in Culture. *Journal of Neurochemistry* **64**, 101-108 (1995).
38. K. Maiese, A. Vincent, S.-H. Lin, T. Shaw, Group I and Group III metabotropic glutamate receptor subtypes provide enhanced neuroprotection. *Journal of Neuroscience Research* **62**, 257-272 (2000).
39. A. Aiba, C. Chen, K. Herrup, C. Rosenmund, C. F. Stevens, S. Tonegawa, Reduced hippocampal long-term potentiation and context-specific deficit in associative learning in mGluR1 mutant mice. *Cell* **79**, 365-375 (1994).
40. G. G. Carter, J. M. Ratcliffe, B. G. Galef, Flower bats (Glossophaga soricina) and fruit bats (Carollia perspicillata) rely on spatial cues over shapes and scents when relocating food. *PLoS One* **5**, e10808 (2010).
41. V. K. Maier, C. M. Feeney, J. E. Taylor, A. L. Creech, J. W. Qiao, A. Szanto, P. P. Das, N. Chevrier, C. Cifuentes-Rojas, S. H. Orkin, S. A. Carr, J. D. Jaffe, P. Mertins, J. T. Lee, Functional Proteomic Analysis of Repressive Histone Methyltransferase Complexes Reveals ZNF518B as a G9A Regulator. *Mol Cell Proteomics* **14**, 1435-1446 (2015).
42. R. M. Hofmann, C. M. Pickart, Noncanonical MMS2-encoded ubiquitin-conjugating enzyme functions in assembly of novel polyubiquitin chains for DNA repair. *Cell* **96**, 645-653 (1999).
43. C. Hoege, B. Pfander, G. L. Moldovan, G. Pyrowolakis, S. Jentsch, RAD6-dependent DNA repair is linked to modification of PCNA by ubiquitin and SUMO. *Nature* **419**, 135-141 (2002).
44. N. M. Foley, G. M. Hughes, Z. Huang, M. Clarke, D. Jebb, C. V. Whelan, E. J. Petit, F. Touzalin, O. Farcy, G. Jones, R. D. Ransome, J. Kacprzyk, M. J. O'Connell, G. Kerth, H. Rebelo, L. Rodrigues, S. J. Puechmaille, E. C. Teeling, Growing old, yet staying young: The role of telomeres in bats' exceptional longevity. *Sci Adv* **4**, eaao0926 (2018).
45. X. Brenachot, G. Ramadori, R. M. Ioris, C. Veyrat-Durebex, J. Altirriba, E. Aras, S. Ljubicic, D. Kohno, S. Fabbiano, S. Clement, N. Goossens, M. Trajkovski, S. Harroch, F. Negro, R. Coppari, Hepatic protein tyrosine phosphatase receptor gamma links obesity-induced inflammation to insulin resistance. *Nature Communications* **8**, 1820 (2017).
46. J. T. Daub, S. Moretti, Davydov, II, L. Excoffier, M. Robinson-Rechavi, Detection of Pathways Affected by Positive Selection in Primate Lineages Ancestral to Humans. *Mol Biol Evol* **34**, 1391-1402 (2017).
47. J. T. Daub, T. Hofer, E. Cutivet, I. Dupanloup, L. Quintana-Murci, M. Robinson-Rechavi, L. Excoffier, Evidence for polygenic adaptation to pathogens in the human genome. *Mol Biol Evol* **30**, 1544-1558 (2013).
48. T. S. Tran, A. L. Kolodkin, R. Bharadwaj, Semaphorin regulation of cellular morphology. *Annu Rev Cell Dev Biol* **23**, 263-292 (2007).
49. R. J. Pasterkamp, R. J. Giger, Semaphorin function in neural plasticity and disease. *Curr Opin Neurobiol* **19**, 263-274 (2009).
50. B. Gao, Wnt regulation of planar cell polarity (PCP). *Curr Top Dev Biol* **101**, 263-295 (2012).

51. L. G. Fradkin, J. M. Dura, J. N. Noordermeer, Ryks: new partners for Wnts in the developing and regenerating nervous system. *Trends Neurosci* **33**, 84-92 (2010).
52. R. J. Giger, E. R. Hollis, 2nd, M. H. Tuszynski, Guidance molecules in axon regeneration. *Cold Spring Harb Perspect Biol* **2**, a001867 (2010).
53. J. W. Fawcett, M. E. Schwab, L. Montani, N. Brazda, H. W. Müller, Defeating inhibition of regeneration by scar and myelin components. *Handb Clin Neurol* **109**, 503-522 (2012).
54. M. Monirujjaman, A. Ferdouse, Metabolic and Physiological Roles of Branched-Chain Amino Acids. *Advances in Molecular Biology* **2014**, 364976 (2014).
55. P. Juricic, S. Grönke, L. Partridge, Branched-Chain Amino Acids Have Equivalent Effects to Other Essential Amino Acids on Lifespan and Aging-Related Traits in Drosophila. *J Gerontol A Biol Sci Med Sci* **75**, 24-31 (2020).
56. N. E. Richardson, E. N. Konon, H. S. Schuster, A. T. Mitchell, C. Boyle, A. C. Rodgers, M. Finke, L. R. Haider, D. Yu, V. Flores, H. H. Pak, S. Ahmad, S. Ahmed, A. Radcliff, J. Wu, E. M. Williams, L. Abdi, D. S. Sherman, T. A. Hacker, D. W. Lamming, Lifelong restriction of dietary branched-chain amino acids has sex-specific benefits for frailty and life span in mice. *Nature Aging* **1**, 73-86 (2021).
57. S. R. Lavin, W. H. Karasov, A. R. Ives, K. M. Middleton, T. Garland, Jr., Morphometrics of the avian small intestine compared with that of nonflying mammals: a phylogenetic approach. *Physiol Biochem Zool* **81**, 526-550 (2008).
58. P. Westfall, S. Young, *Resampling-Based Multiple Testing: Examples and Methods for P-Value Adjustment* (Wiley, 1993).
59. L. Cheynel, J. F. Lemaître, J. M. Gaillard, B. Rey, G. Bourgoin, H. Ferté, M. Jégo, F. Débias, M. Pellerin, L. Jacob, E. Gilot-Fromont, Immunosenescence patterns differ between populations but not between sexes in a long-lived mammal. *Scientific Reports* **7**, 13700 (2017).
60. H. G. Hilton, N. D. Rubinstein, P. Janki, A. T. Ireland, N. Bernstein, N. L. Fong, K. M. Wright, M. Smith, D. Finkle, B. Martin-McNulty, M. Roy, D. M. Imai, V. Jojic, R. Buffenstein, Single-cell transcriptomics of the naked mole-rat reveals unexpected features of mammalian immunity. *PLOS Biology* **17**, e3000528 (2019).
61. M. J. Youngman, Z. N. Rogers, D. H. Kim, A decline in p38 MAPK signaling underlies immunosenescence in Caenorhabditis elegans. *PLoS Genet* **7**, e1002082 (2011).
62. H. Chen, X. Zheng, Y. Zheng, Age-associated loss of lamin-B leads to systemic inflammation and gut hyperplasia. *Cell* **159**, 829-843 (2014).
63. E. Kalay, G. Yigit, Y. Aslan, K. E. Brown, E. Pohl, L. S. Bicknell, H. Kayserili, Y. Li, B. Tüysüz, G. Nürnberg, W. Kiess, M. Koegl, I. Baessmann, K. Buruk, B. Toraman, S. Kayipmaz, S. Kul, M. Ikbal, D. J. Turner, M. S. Taylor, J. Aerts, C. Scott, K. Milstein, H. Dollfus, D. Wieczorek, H. G. Brunner, M. Hurles, A. P. Jackson, A. Rauch, P. Nürnberg, A. Karagüzel, B. Wollnik, CEP152 is a genome maintenance protein disrupted in Seckel syndrome. *Nat Genet* **43**, 23-26 (2011).
64. T. X. Jiang, J. B. Zou, Q. Q. Zhu, C. H. Liu, G. F. Wang, T. T. Du, Z. Y. Luo, F. Guo, L. M. Zhou, J. J. Liu, W. Zhang, Y. S. Shu, L. Yu, P. Li, Z. A. Ronai, S. I. Matsuzawa, A. L. Goldberg, X. B. Qiu, SIP/CacyBP promotes autophagy by regulating levels of BRUCE/Apollon, which stimulates LC3-I degradation. *Proc Natl Acad Sci U S A* **116**, 13404-13413 (2019).
65. S. Rosińska, A. Filipek, Interaction of CacyBP/SIP with NPM1 and its influence on NPM1 localization and function in oxidative stress. *J Cell Physiol* **233**, 8826-8838 (2018).

66. V. Infantino, F. Dituri, P. Convertini, A. Santarsiero, F. Palmieri, S. Todisco, S. Mancarella, G. Giannelli, V. Iacobazzi, Epigenetic upregulation and functional role of the mitochondrial aspartate/glutamate carrier isoform 1 in hepatocellular carcinoma. *Biochim Biophys Acta Mol Basis Dis* **1865**, 38-47 (2019).
67. X. Tian, D. Firsanov, Z. Zhang, Y. Cheng, L. Luo, G. Tombline, R. Tan, M. Simon, S. Henderson, J. Steffan, A. Goldfarb, J. Tam, K. Zheng, A. Cornwell, A. Johnson, J.-N. Yang, Z. Mao, B. Manta, W. Dang, Z. Zhang, J. Vijg, A. Wolfe, K. Moody, B. K. Kennedy, D. Bohmann, V. N. Gladyshev, A. Seluanov, V. Gorbunova, SIRT6 Is Responsible for More Efficient DNA Double-Strand Break Repair in Long-Lived Species. *Cell* **177**, 622-638.e622 (2019).
68. K. M. Kim, J. H. Noh, M. Bodogai, J. L. Martindale, P. R. Pandey, X. Yang, A. Biragyn, K. Abdelmohsen, M. Gorospe, SCAMP4 enhances the senescent cell secretome. *Genes Dev* **32**, 909-914 (2018).
69. G. Pegoraro, N. Kubben, U. Wickert, H. Göhler, K. Hoffmann, T. Misteli, Ageing-related chromatin defects through loss of the NURD complex. *Nat Cell Biol* **11**, 1261-1267 (2009).
70. M. R. Amin, S. A. Mahmud, J. L. Dowgielewicz, M. Sapkota, M. W. Pellegrino, A novel gene-diet interaction promotes organismal lifespan and host protection during infection via the mitochondrial UPR. *PLOS Genetics* **16**, e1009234 (2020).
71. D. Wu, W. Cai, X. Zhang, J. Lan, L. Zou, S. J. Chen, Z. Wu, D. Chen, Inhibition of PAR-1 delays aging via activating AMPK in C. elegans. *Aging (Albany NY)* **12**, 25700-25717 (2020).
72. T. von der Haar, J. E. Leadsham, A. Sauvadet, D. Tarrant, I. S. Adam, K. Saromi, P. Laun, M. Rinnerthaler, H. Breitenbach-Koller, M. Breitenbach, M. F. Tuite, C. W. Gourlay, The control of translational accuracy is a determinant of healthy ageing in yeast. *Open Biol* **7**, (2017).
73. V. E. Martinez-Miguel, C. Lujan, T. Espie--Caullet, D. Martinez-Martinez, S. Moore, C. Backes, S. Gonzalez, E. R. Galimov, A. E. X. Brown, M. Halic, K. Tomita, C. Rallis, T. von der Haar, F. Cabreiro, I. Bjedov, Increased fidelity of protein synthesis extends lifespan. *Cell Metabolism* **33**, 2288-2300.e2212 (2021).
74. J. Azpurua, Z. Ke, I. X. Chen, Q. Zhang, D. N. Ermolenko, Z. D. Zhang, V. Gorbunova, A. Seluanov, Naked mole-rat has increased translational fidelity compared with the mouse, as well as a unique 28S ribosomal RNA cleavage. *Proc Natl Acad Sci U S A* **110**, 17350-17355 (2013).
75. A. S. Anisimova, M. B. Meerson, M. V. Gerashchenko, I. V. Kulakovskiy, S. E. Dmitriev, V. N. Gladyshev, Multifaceted deregulation of gene expression and protein synthesis with age. *Proc Natl Acad Sci U S A* **117**, 15581-15590 (2020).
76. Y. Li, X. Li, M. Cao, Y. Jiang, J. Yan, Z. Liu, R. Yang, X. Chen, P. Sun, R. Xiang, L. Wang, Y. Shi, Seryl tRNA synthetase cooperates with POT1 to regulate telomere length and cellular senescence. *Signal Transduction and Targeted Therapy* **4**, (2019).
77. T. Shimizu, A. Lengalova, V. Martinek, M. Martinkova, Heme: emergent roles of heme in signal transduction, functional regulation and as catalytic centres. *Chem Soc Rev* **48**, 5624-5657 (2019).
78. R. S. Balaban, S. Nemoto, T. Finkel, Mitochondria, oxidants, and aging. *Cell* **120**, 483-495 (2005).
79. O. Adedoyin, R. Boddu, A. Traylor, J. M. Lever, S. Bolisetty, J. F. George, A. Agarwal, Heme oxygenase-1 mitigates ferroptosis in renal proximal tubule cells. *Am J Physiol Renal Physiol* **314**, F702-F714 (2018).
80. A. Fransson, A. Ruusala, P. Aspenström, Atypical Rho GTPases have roles in mitochondrial homeostasis and apoptosis. *J Biol Chem* **278**, 6495-6502 (2003).

81. K. Wu, S. Katiyar, A. Li, M. Liu, X. Ju, V. M. Popov, X. Jiao, M. P. Lisanti, A. Casola, R. G. Pestell, Dachshund inhibits oncogene-induced breast cancer cellular migration and invasion through suppression of interleukin-8. *Proc Natl Acad Sci U S A* **105**, 6924-6929 (2008).
82. J. Gambini, L. Gimeno-Mallench, M. Inglés, G. Olaso, K. M. Abdelaziz, J. A. Avellana, Á. Belenguer, R. Cruz, C. Mas-Bargues, C. Borras, J. Viña, Identification of single nucleotide polymorphisms in centenarians. *Rev Esp Geriatr Gerontol* **51**, 146-149 (2016).
83. Y. Wang, V. K. Rao, W. K. Kok, D. N. Roy, S. Sethi, B. M. Ling, M. B. Lee, R. Taneja, SUMO modification of Stra13 is required for repression of cyclin D1 expression and cellular growth arrest. *PLoS One* **7**, e43137 (2012).
84. Y. Qian, J. Zhang, B. Yan, X. Chen, DEC1, a basic helix-loop-helix transcription factor and a novel target gene of the p53 family, mediates p53-dependent premature senescence. *J Biol Chem* **283**, 2896-2905 (2008).
85. Z. Yun, H. L. Maecker, R. S. Johnson, A. J. Giaccia, Inhibition of PPAR gamma 2 gene expression by the HIF-1-regulated gene DEC1/Stra13: a mechanism for regulation of adipogenesis by hypoxia. *Dev Cell* **2**, 331-341 (2002).
86. Rommel A. Mathias, Todd M. Greco, A. Oberstein, Hanna G. Budayeva, R. Chakrabarti, Elizabeth A. Rowland, Y. Kang, T. Shenk, Ileana M. Cristea, Sirtuin 4 Is a Lipoamidase Regulating Pyruvate Dehydrogenase Complex Activity. *Cell* **159**, 1615-1625 (2014).
87. J. Kaplon, L. Zheng, K. Meissl, B. Chaneton, V. A. Selivanov, G. Mackay, S. H. van der Burg, E. M. Verdegaal, M. Cascante, T. Shlomi, E. Gottlieb, D. S. Peeper, A key role for mitochondrial gatekeeper pyruvate dehydrogenase in oncogene-induced senescence. *Nature* **498**, 109-112 (2013).
88. E. Bang, B. Lee, S.-G. Noh, D. H. Kim, H. J. Jung, S. Ha, B. P. Yu, H. Y. Chung, Modulation of senoinflammation by calorie restriction based on biochemical and Omics big data analysis. *BMB Rep* **52**, 56-63 (2019).
89. C. Vasavda, R. Kothari, A. P. Malla, R. Tokhunts, A. Lin, M. Ji, C. Ricco, R. Xu, H. G. Saavedra, J. I. Sbodio, A. M. Snowman, L. Albacarys, L. Hester, T. W. Sedlak, B. D. Paul, S. H. Snyder, Bilirubin Links Heme Metabolism to Neuroprotection by Scavenging Superoxide. *Cell Chem Biol* **26**, 1450-1460.e1457 (2019).
90. J. Melrose, Keratan sulfate (KS)-proteoglycans and neuronal regulation in health and disease: the importance of KS-glycodynamics and interactive capability with neuroregulatory ligands. *J Neurochem* **149**, 170-194 (2019).
91. I. Zironi, P. Gaibani, D. Remondini, S. Salvioli, S. Altilia, M. Pierini, G. Aicardi, E. Verondini, L. Milanesi, F. Bersani, S. Gravina, I. B. Roninson, C. Franceschi, G. Castellani, Molecular remodeling of potassium channels in fibroblasts from centenarians: a marker of longevity? *Mech Ageing Dev* **131**, 674-681 (2010).
92. P. Smith, E. Buhl, K. Tsaneva-Atanasova, J. J. L. Hodge, Shaw and Shal voltage-gated potassium channels mediate circadian changes in Drosophila clock neuron excitability. *J Physiol* **597**, 5707-5722 (2019).
93. K. Tucker, J. M. Overton, D. A. Fadool, Kv1.3 gene-targeted deletion alters longevity and reduces adiposity by increasing locomotion and metabolism in melanocortin-4 receptor-null mice. *Int J Obes (Lond)* **32**, 1222-1232 (2008).

94. C. Cirelli, D. Bushey, S. Hill, R. Huber, R. Kreber, B. Ganetzky, G. Tononi, Reduced sleep in Drosophila Shaker mutants. *Nature* **434**, 1087-1092 (2005).
95. J. V. Pluvinage, M. S. Haney, B. A. H. Smith, J. Sun, T. Iram, L. Bonanno, L. Li, D. P. Lee, D. W. Morgens, A. C. Yang, S. R. Shuken, D. Gate, M. Scott, P. Khatri, J. Luo, C. R. Bertozzi, M. C. Bassik, T. Wyss-Coray, CD22 blockade restores homeostatic microglial phagocytosis in ageing brains. *Nature* **568**, 187-192 (2019).
96. L. L. Chen, Y. B. Wang, J. X. Song, W. K. Deng, J. H. Lu, L. L. Ma, C. B. Yang, M. Li, Y. Xue, Phosphoproteome-based kinase activity profiling reveals the critical role of MAP2K2 and PLK1 in neuronal autophagy. *Autophagy* **13**, 1969-1980 (2017).
97. G. Maria Fimia, A. Stoykova, A. Romagnoli, L. Giunta, S. Di Bartolomeo, R. Nardacci, M. Corazzari, C. Fuoco, A. Ucar, P. Schwartz, P. Gruss, M. Piacentini, K. Chowdhury, F. Cecconi, Ambra1 regulates autophagy and development of the nervous system. *Nature* **447**, 1121-1125 (2007).
98. A. Simonsen, R. C. Cumming, A. Brech, P. Isakson, D. R. Schubert, K. D. Finley, Promoting basal levels of autophagy in the nervous system enhances longevity and oxidant resistance in adult Drosophila. *Autophagy* **4**, 176-184 (2008).
99. V. K. Parihar, B. D. Allen, K. K. Tran, N. N. Chmielewski, B. M. Craver, V. Martirosian, J. M. Morganti, S. Rosi, R. Vlkolinsky, M. M. Acharya, G. A. Nelson, A. R. Allen, C. L. Limoli, Targeted overexpression of mitochondrial catalase prevents radiation-induced cognitive dysfunction. *Antioxid Redox Signal* **22**, 78-91 (2015).
100. A. Csiszar, A. Yabluchanskiy, A. Ungvari, Z. Ungvari, S. Tarantini, Overexpression of catalase targeted to mitochondria improves neurovascular coupling responses in aged mice. *Geroscience* **41**, 609-617 (2019).
101. W. Tan, W. Wang, Q. Ma, Physiological and Pathological Function of Serine/Arginine-Rich Splicing Factor 4 and Related Diseases. *Biomed Res Int* **2018**, 3819719 (2018).
102. S. M. Kwon, S. Min, U. W. Jeoun, M. S. Sim, G. H. Jung, S. M. Hong, B. A. Jee, H. G. Woo, C. Lee, G. Yoon, Global spliceosome activity regulates entry into cellular senescence. *Faseb j* **35**, e21204 (2021).
103. J. Larrasa-Alonso, M. Villalba, F. Sanchez-Cabo, M. Lopez-Olaneta, P. Ortiz-Sanchez, P. Garcia-Pavia, E. Lara-Pezzi, in *Cardiovascular Research*. (Oxford Univ. Press, Oxford, 2016), vol. 111, pp. S36-S36.
104. S. Hassock, D. Vetrie, F. Giannelli, Mapping and characterization of the X-linked dyskeratosis congenita (DKC) gene. *Genomics* **55**, 21-27 (1999).
105. E. Latorre, L. W. Harries, Splicing regulatory factors, ageing and age-related disease. *Ageing Res Rev* **36**, 165-170 (2017).
106. H. Guo, Y. Li, M. Luo, S. Lin, J. Chen, Q. Ma, Y. Gu, Z. Jiang, Y. Gui, Androgen receptor binding to an androgen-responsive element in the promoter of the Srsf4 gene inhibits its expression in mouse Sertoli cells. *Mol Reprod Dev* **82**, 976-985 (2015).
107. T. G. Son, Y. Zou, B. P. Yu, J. Lee, H. Y. Chung, Aging effect on myeloperoxidase in rat kidney and its modulation by calorie restriction. *Free Radic Res* **39**, 283-289 (2005).
108. Z. Cheng, M. F. White, Targeting Forkhead box O1 from the concept to metabolic diseases: lessons from mouse models. *Antioxid Redox Signal* **14**, 649-661 (2011).
109. J. O. Wertheim, B. Murrell, M. D. Smith, S. L. Kosakovsky Pond, K. Scheffler, RELAX: detecting relaxed selection in a phylogenetic framework. *Mol Biol Evol* **32**, 820-832 (2015).

110. R. Cui, T. Medeiros, D. Willemsen, L. N. M. Iasi, G. E. Collier, M. Graef, M. Reichard, D. R. Valenzano, Relaxed Selection Limits Lifespan by Increasing Mutation Load. *Cell*, (2019).
111. D. A. Chavous, F. R. Jackson, C. M. O'Connor, Extension of the Drosophila lifespan by overexpression of a protein repair methyltransferase. *Proc Natl Acad Sci U S A* **98**, 14814-14818 (2001).
112. J. Kim, K. J. Lee, J. S. Kim, J. G. Rho, J. J. Shin, W. K. Song, E. K. Lee, J. M. Egan, W. Kim, Cannabinoids Regulate Bcl-2 and Cyclin D2 Expression in Pancreatic β Cells. *PLoS One* **11**, e0150981 (2016).
113. S. Singh, N. Dhaliwal, R. Crawford, Y. Xiao, Cellular senescence and longevity of osteophyte-derived mesenchymal stem cells compared to patient-matched bone marrow stromal cells. *J Cell Biochem* **108**, 839-850 (2009).
114. T. J. Hale, B. C. Richardson, L. I. Sweet, D. L. McElligott, J. E. Riggs, E. B. Chu, J. M. Glynn, D. LaFrenz, D. N. Ernst, R. Rochford, M. V. Hobbs, Age-related changes in mature CD4+ T cells: cell cycle analysis. *Cell Immunol* **220**, 51-62 (2002).
115. G. Hoxhaj, E. Caddye, A. Najafov, V. P. Houde, C. Johnson, K. Dissanayake, R. Toth, D. G. Campbell, A. R. Prescott, C. MacKintosh, The E3 ubiquitin ligase ZNRF2 is a substrate of mTORC1 and regulates its activation by amino acids. *Elife* **5**, (2016).
116. I. Pradas, M. Jové, R. Cabré, V. Ayala, N. Mota-Martorell, R. Pamplona, Effects of Aging and Methionine Restriction on Rat Kidney Metabolome. *Metabolites* **9**, (2019).
117. S. M. Lam, Z. Wang, J. Li, X. Huang, G. Shui, Sequestration of polyunsaturated fatty acids in membrane phospholipids of Caenorhabditis elegans dauer larva attenuates eicosanoid biosynthesis for prolonged survival. *Redox Biol* **12**, 967-977 (2017).
118. C. Bárcena, P. M. Quirós, S. Durand, P. Mayoral, F. Rodríguez, X. M. Caravia, G. Mariño, C. Garabaya, M. T. Fernández-García, G. Kroemer, J. M. P. Freije, C. López-Otín, Methionine Restriction Extends Lifespan in Progeroid Mice and Alters Lipid and Bile Acid Metabolism. *Cell Rep* **24**, 2392-2403 (2018).
119. I. Sanchez-Roman, G. Barja, Regulation of longevity and oxidative stress by nutritional interventions: role of methionine restriction. *Exp Gerontol* **48**, 1030-1042 (2013).
120. M. Lezzerini, Y. Budovskaya, A dual role of the Wnt signaling pathway during aging in Caenorhabditis elegans. *Aging Cell* **13**, 8-18 (2014).
121. V. A. Rafalski, A. Brunet, Energy metabolism in adult neural stem cell fate. *Prog Neurobiol* **93**, 182-203 (2011).
122. G. Laurent, F. Solari, B. Mateescu, M. Karaca, J. Castel, B. Bourachot, C. Magnan, M. Billaud, F. Mechta-Grigoriou, Oxidative stress contributes to aging by enhancing pancreatic angiogenesis and insulin signaling. *Cell Metab* **7**, 113-124 (2008).
123. N. Boucher, T. Dufeu-Duchesne, E. Vicaut, D. Farge, R. B. Effros, F. Schächter, CD28 expression in T cell aging and human longevity. *Exp Gerontol* **33**, 267-282 (1998).
124. J. Azpurua, Z. Ke, I. X. Chen, Q. Zhang, D. N. Ermolenko, Z. D. Zhang, V. Gorbunova, A. Seluanov, Naked mole-rat has increased translational fidelity compared with the mouse, as well as a unique 28S ribosomal RNA cleavage. *Proceedings of the National Academy of Sciences* **110**, 17350 (2013).
125. O. Omotoso, V. N. Gladyshev, X. Zhou, Lifespan Extension in Long-Lived Vertebrates Rooted in Ecological Adaptation. *Front Cell Dev Biol* **9**, 704966 (2021).

126. R. K. Patel, M. Jain, NGS QC Toolkit: a toolkit for quality control of next generation sequencing data. *PLoS One* **7**, e30619 (2012).
127. M. Keane, J. Semeiks, A. E. Webb, Y. I. Li, V. Quesada, T. Craig, L. B. Madsen, S. van Dam, D. Brawand, P. I. Marques, P. Michalak, L. Kang, J. Bhak, H. S. Yim, N. V. Grishin, N. H. Nielsen, M. P. Heide-Jørgensen, E. M. Oziolor, C. W. Matson, G. M. Church, G. W. Stuart, J. C. Patton, J. C. George, R. Suydam, K. Larsen, C. López-Otín, M. J. O'Connell, J. W. Bickham, B. Thomsen, J. P. de Magalhães, Insights into the evolution of longevity from the bowhead whale genome. *Cell Rep* **10**, 112-122 (2015).
128. M. Stanke, M. Diekhans, R. Baertsch, D. Haussler, Using native and syntenically mapped cDNA alignments to improve de novo gene finding. *Bioinformatics* **24**, 637-644 (2008).
129. M. G. Grabherr, B. J. Haas, M. Yassour, J. Z. Levin, D. A. Thompson, I. Amit, X. Adiconis, L. Fan, R. Raychowdhury, Q. Zeng, Z. Chen, E. Mauceli, N. Hacohen, A. Gnirke, N. Rhind, F. di Palma, B. W. Birren, C. Nusbaum, K. Lindblad-Toh, N. Friedman, A. Regev, Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol* **29**, 644-652 (2011).
130. W. Li, A. Godzik, Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **22**, 1658-1659 (2006).
131. L. Fu, B. Niu, Z. Zhu, S. Wu, W. Li, CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* **28**, 3150-3152 (2012).
132. C. Trapnell, B. A. Williams, G. Pertea, A. Mortazavi, G. Kwan, M. J. van Baren, S. L. Salzberg, B. J. Wold, L. Pachter, Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol* **28**, 511-515 (2010).
133. G. M. Boratyn, C. Camacho, P. S. Cooper, G. Coulouris, A. Fong, N. Ma, T. L. Madden, W. T. Matten, S. D. McGinnis, Y. Merezhuk, Y. Raytselis, E. W. Sayers, T. Tao, J. Ye, I. Zaretskaya, BLAST: a more efficient report with usability improvements. *Nucleic Acids Res* **41**, W29-33 (2013).
134. A. Dobin, C. A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski, S. Jha, P. Batut, M. Chaisson, T. R. Gingeras, STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21 (2013).
135. Y. Liao, G. K. Smyth, W. Shi, featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923-930 (2014).
136. J. T. Leek, W. E. Johnson, H. S. Parker, A. E. Jaffe, J. D. Storey, The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics* **28**, 882-883 (2012).
137. K. Fukushima, D. D. Pollock, Amalgamated cross-species transcriptomes reveal organ-specific propensity in gene expression evolution. *Nat Commun* **11**, 4459 (2020).
138. M. D. Robinson, D. J. McCarthy, G. K. Smyth, edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139-140 (2010).
139. S. Lê, J. Josse, F. Husson, FactoMineR: An R Package for Multivariate Analysis. *Journal of Statistical Software* **1**, (2008).
140. K. E. Jones, J. Bielby, M. Cardillo, S. A. Fritz, J. O'Dell, C. D. L. Orme, K. Safi, W. Sechrest, E. H. Boakes, C. Carbone, PanTHERIA: a species ‐ level database of life history, ecology, and geography of extant and recently extinct mammals: Ecological Archives E090 ‐ 184. *Ecology* **90**, 2648-2648 (2009).
141. S. Kumar, G. Stecher, M. Suleski, S. B. Hedges, TimeTree: A Resource for Timelines, Timetrees, and Divergence Times. *Mol Biol Evol* **34**, 1812-1819 (2017).

142. S. van Buuren, K. Groothuis-Oudshoorn, mice: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software* **1**, (2011).
143. D. J. Stekhoven, P. Bühlmann, MissForest—non-parametric missing value imputation for mixed-type data. *Bioinformatics* **28**, 112-118 (2012).
144. J. Bruggeman, J. Heringa, B. W. Brandt, PhyloPars: estimation of missing parameter values using phylogeny. *Nucleic Acids Research* **37**, W179-W184 (2009).
145. T. Santos. (R package version 0.3, 2018).
146. F. Abascal, R. Zardoya, M. J. Telford, TranslatorX: multiple alignment of nucleotide sequences guided by amino acid translations. *Nucleic Acids Research* **38**, W7-W13 (2010).
147. K. Katoh, K. Misawa, K. Kuma, T. Miyata, MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* **30**, 3059-3066 (2002).
148. J. Castresana, Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol* **17**, 540-552 (2000).
149. G. Talavera, J. Castresana, Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Syst Biol* **56**, 564-577 (2007).
150. M. Lek, K. J. Karczewski, E. V. Minikel, K. E. Samocha, E. Banks, T. Fennell, A. H. O'Donnell-Luria, J. S. Ware, A. J. Hill, B. B. Cummings, T. Tukiainen, D. P. Birnbaum, J. A. Kosmicki, L. E. Duncan, K. Estrada, F. Zhao, J. Zou, E. Pierce-Hoffman, J. Berghout, D. N. Cooper, N. Deflaux, M. DePristo, R. Do, J. Flannick, M. Fromer, L. Gauthier, J. Goldstein, N. Gupta, D. Howrigan, A. Kiezun, M. I. Kurki, A. L. Moonshine, P. Natarajan, L. Orozco, G. M. Peloso, R. Poplin, M. A. Rivas, V. Ruano-Rubio, S. A. Rose, D. M. Ruderfer, K. Shakir, P. D. Stenson, C. Stevens, B. P. Thomas, G. Tiao, M. T. Tusie-Luna, B. Weisburd, H. H. Won, D. Yu, D. M. Altshuler, D. Ardissino, M. Boehnke, J. Danesh, S. Donnelly, R. Elosua, J. C. Florez, S. B. Gabriel, G. Getz, S. J. Glatt, C. M. Hultman, S. Kathiresan, M. Laakso, S. McCarroll, M. I. McCarthy, D. McGovern, R. McPherson, B. M. Neale, A. Palotie, S. M. Purcell, D. Saleheen, J. M. Scharf, P. Sklar, P. F. Sullivan, J. Tuomilehto, M. T. Tsuang, H. C. Watkins, J. G. Wilson, M. J. Daly, D. G. MacArthur, Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **536**, 285-291 (2016).
151. P. D. Stenson, M. Mort, E. V. Ball, K. Evans, M. Hayden, S. Heywood, M. Hussain, A. D. Phillips, D. N. Cooper, The Human Gene Mutation Database: towards a comprehensive repository of inherited mutation data for medical research, genetic diagnosis and next-generation sequencing studies. *Human Genetics* **136**, 665-677 (2017).
152. H. A. Shihab, M. F. Rogers, C. Campbell, T. R. Gaunt, HIPred: an integrative approach to predicting haploinsufficient genes. *Bioinformatics* **33**, 1751-1757 (2017).
153. Y. Shao, C. Chen, H. Shen, B. Z. He, D. Yu, S. Jiang, S. Zhao, Z. Gao, Z. Zhu, X. Chen, Y. Fu, H. Chen, G. Gao, M. Long, Y. E. Zhang, GenTree, an integrated resource for analyzing the evolution and function of primate-specific coding genes. *Genome Res* **29**, 682-696 (2019).

## Table, Figures and Legends

**Table 1 Statistics on genes whose expression variation is associated with life-history variation**

| Variable* | Liver (n = 13, 508)† | | Kidney (n = 13, 183) | | Brain (n = 12, 950) | | |
|---|---|---|---|---|---|---|---|
| | Nb. of genes‡ | % from total | Nb. of genes | % from total | Nb. of genes | % from total | Combined¶ |
| Adult weight | 122 (101) | 0.90 (0.75) | 144 (119) | 1.09 (0.90) | 179 (127) | 1.38 (0.98) | 440 (5) |
| Maximum lifespan | 148 (81) | 1.10 (0.60) | 83 (47) | 0.63 (0.36) | 91 (59) | 0.70 (0.37) | 311 (11) |
| Female time to maturity | 155 (72) | 1.15 (0.53) | 131 (94) | 0.99 (0.71) | 137 (101) | 1.06 (0.69) | 413 (8) |
| Maximum lifespan residual | 96 (51) | 0.71 (0.38) | 96 (56) | 0.73 (0.42) | 131 (81) | 1.01 (0.66) | 323 (1) |
| Female time to maturity residual | 195 (133) | 1.44 (0.98) | 155 (102) | 1.18 (0.77) | 172 (104) | 1.33 (0.80) | 518 (4) |
| Combined§ | 563 (114) | 4.17 (0.84) | 497 (79) | 3.77 (0.60) | 569 (88) | 4.39 (0.68) | 1557 (275)# |

* The adjusted PGLS P-value cutoff is $P_{robust} < 0.01$ and $P_{max} < 0.05$.

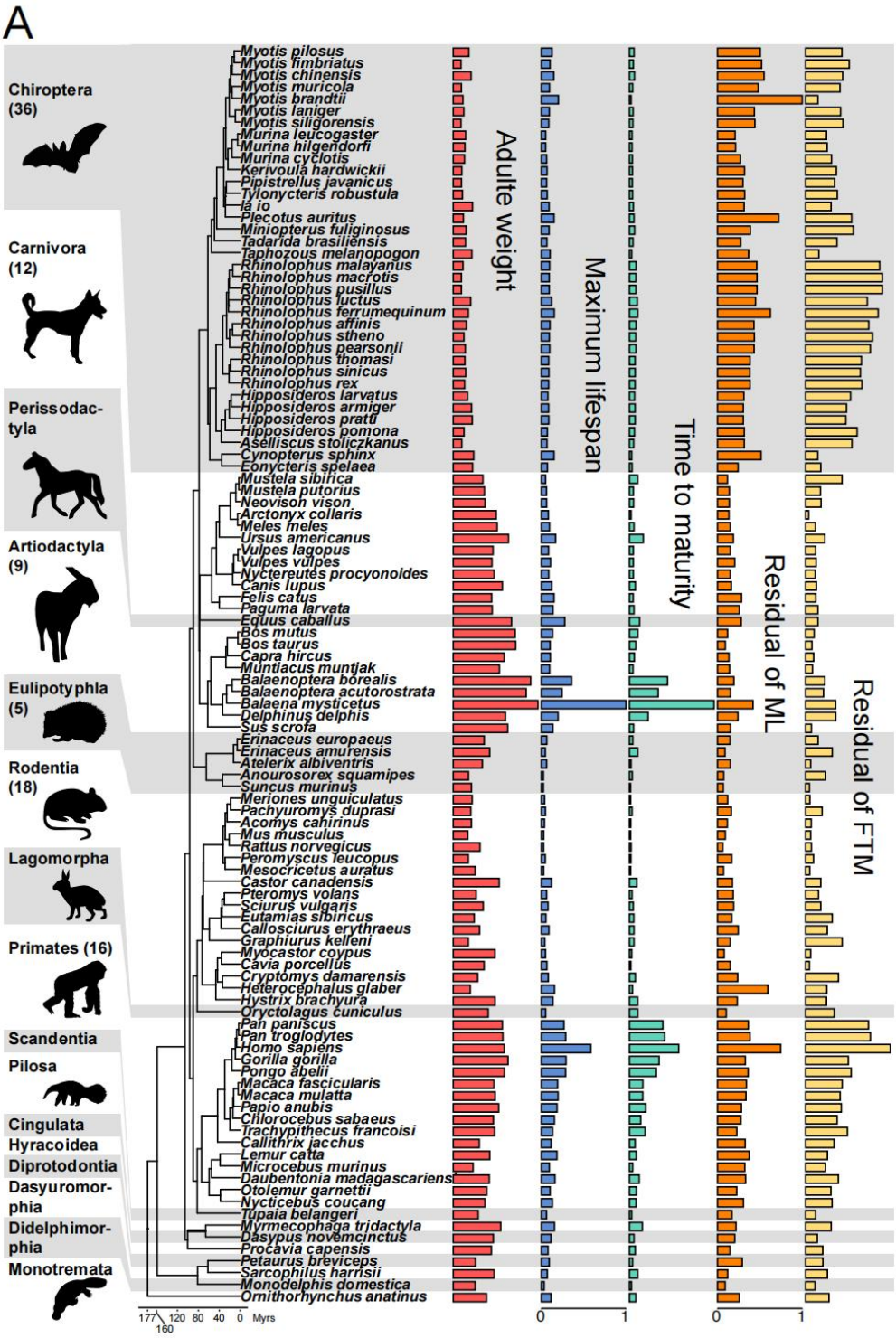
† n denotes total number of orthologous groups assayed in the analysis.

‡ Number of unique genes associated with trait variation and number of genes specific for a trait (in brackets).

§ Number of unique genes identified in the organ and number of core genes in the organ (in brackets).

¶ Number of unique genes identified in three organs for a specific trait and overlap in at least two organs (in brackets).

# Number of unique genes identified in three organs for all traits and number of core genes in all organs (in brackets).




1345

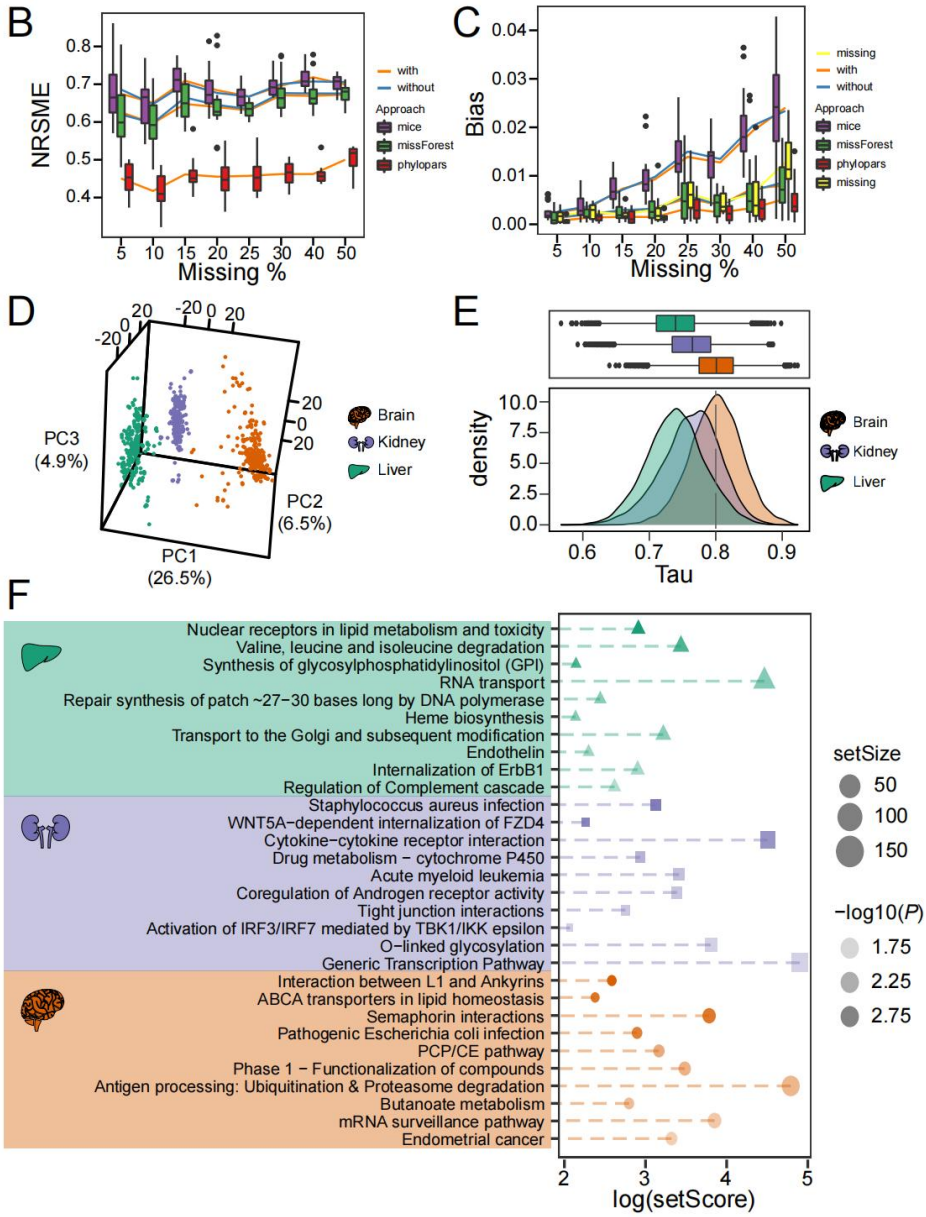
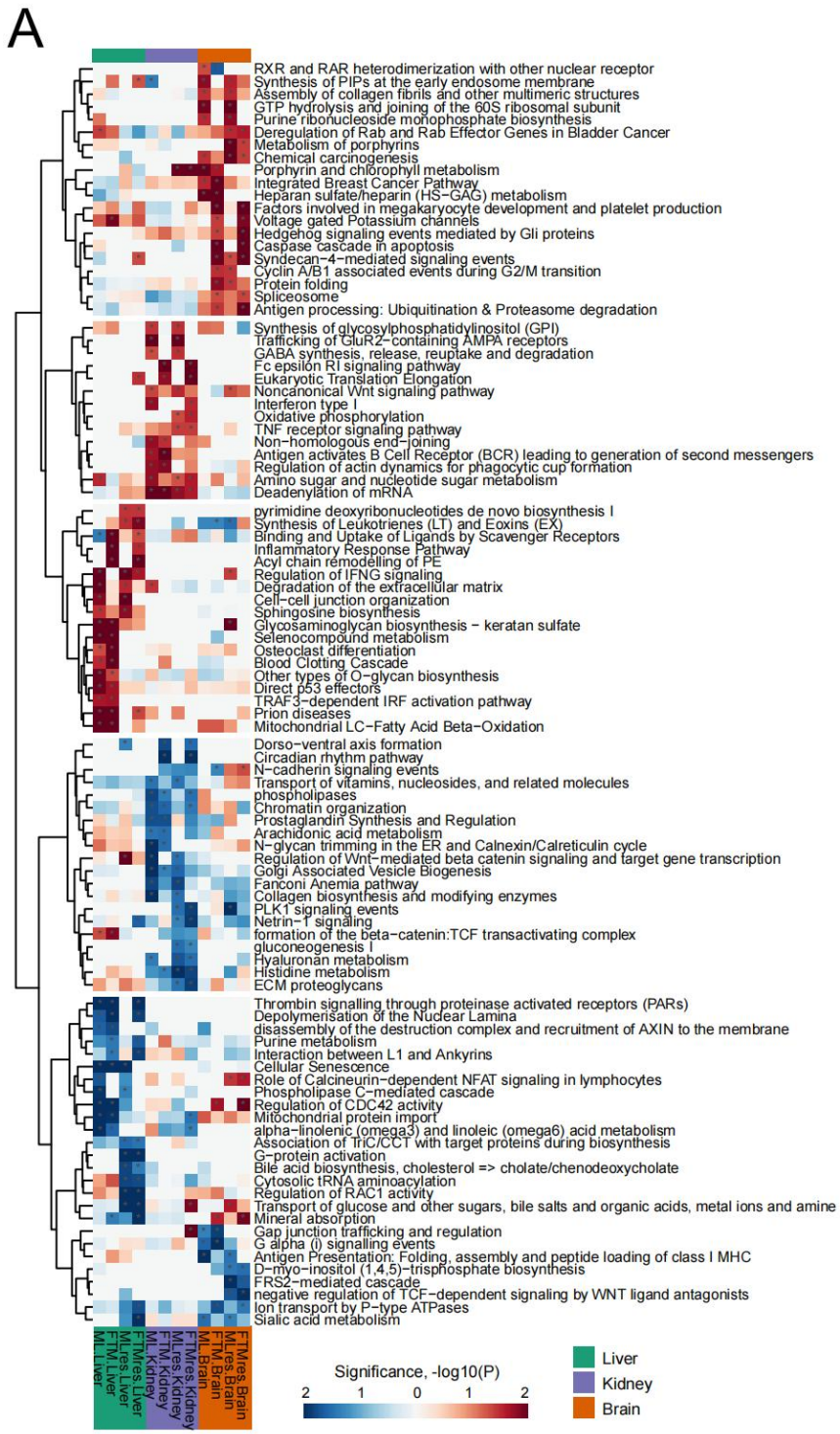



**Fig. 1. Life-history traits and gene expression profile of mammals.** (*A*) Mammalian phylogenetic tree with corresponding life-history traits. From left to right: Adult weight ($log_{10}$-transform), maximum lifespan, female sexual maturity time, and residual of the maximum lifespan and female sexual maturity time relative to the adult weight. Each bar denotes a value of life-history variable for a particular organism in standard scale. The animals image retrieved from PhyloPic (http://www.phylopic.org/) (*B*) Estimation accuracy of missing values of life history traits: the x-axis is the proportion of missing values, and the y-axis is the standard root mean square error (NRSME). (*C*) Estimation bias of missing values: the x-axis is the proportion of missing values, and the y-axis is the bias of biological significance. (*D*) Principal component analysis of gene expression across tissues. The first three principal components (PCs) and their variance explanation percentages are shown. Each repetition is treated as a point. (*E*) Distribution of species-

specific expression index (Tau) in the three tissues. Different organs are shown by different colors. The x-axis is the species-specific expression index; the y-axis represents the frequency (below). The dotted line represents the threshold of the species-specific expression index. (*F*) Enrichment analysis of species-specifically expressed genes for each tissue (Liver: green; Kidney: blue; Brain: orange). The x-axis is the log-transformed gene set score. The depth of the color represents the degree of significance of pathway enrichment, and dot size represents the size of the gene set.



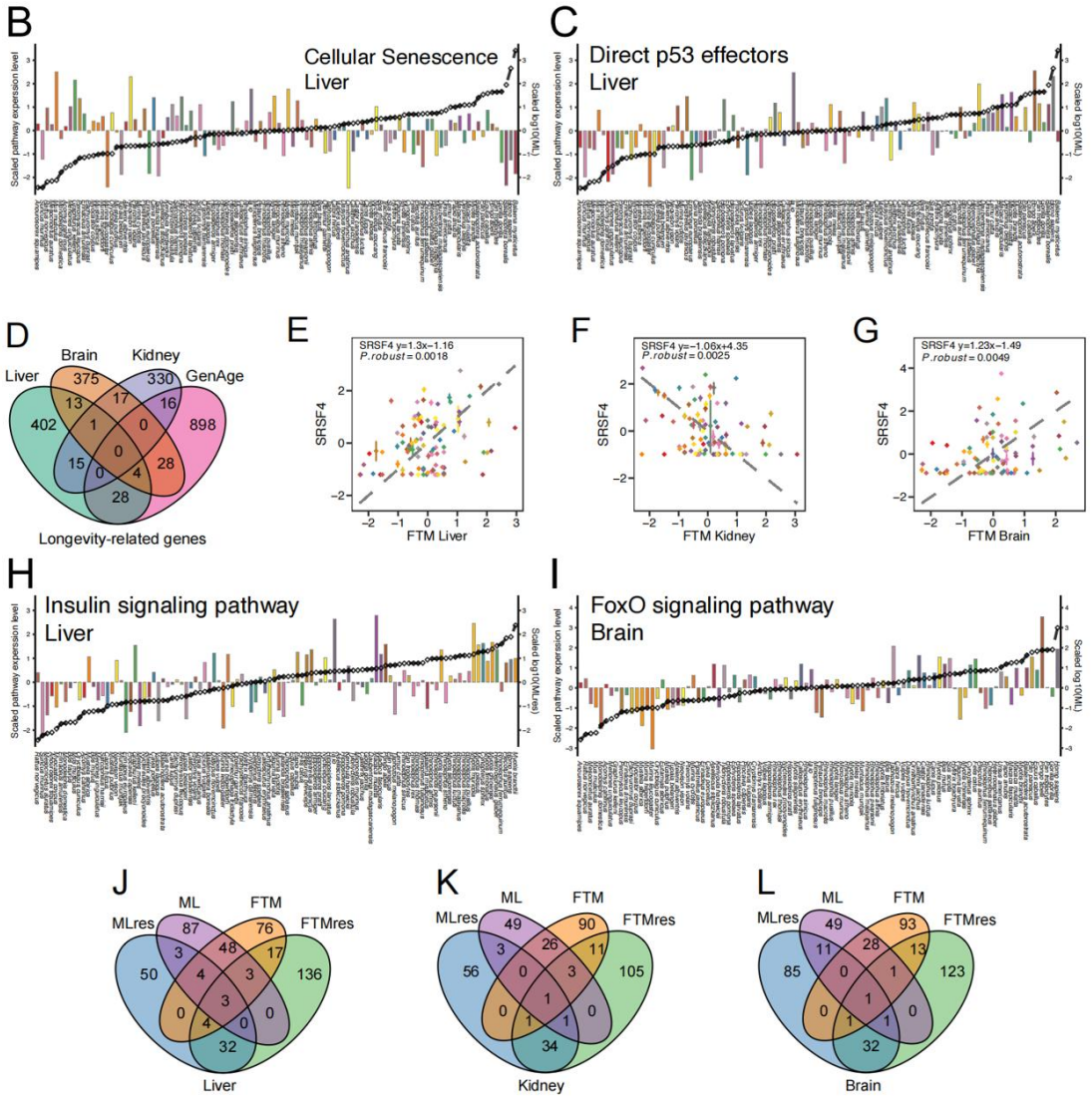



**Fig. 2. Relationship between gene expression variation and longevity.** (*A*) Heat map of the pathways enriched by longevity-related genes. The color intensity indicates the degree of significance, and the *P* value is transformed (-log10). Each row represents a different pathway, and each column represents traits related to longevity (marked at the bottom). Among them, red and blue colors show positive and negative correlations, respectively. The bar graph is the total expression of (*B*) Cellular Senescence negative longevity-related genes and (*C*) Direct p53 effectors positive longevity-related genes in the liver (y-axis on the left). Black line is the relative value of life-history variable (y-axis on the right). Species are shown at the bottom. All values are in standard scale. (*D*) Venn diagrams of the three tissues for longevity-related genes and aging genes of model organisms, obtained from the GenAge database. (*E*, *F*, *G*) *SRSF4* is positively correlated with longevity traits in the liver and brain, and negatively correlated in the kidney. In each figure, the y-axis is the scaled expression level of each gene with 0 as the center, and the x-axis represents the longevity traits (ML: maximum lifespan; FTM: female time to maturity; MLres and FTMres: ML and FTM residuals adjusted for adult weight). Potential

outliers have been removed. The optimal phylogenetic regression equation, *P* value is included in the figure. The bar graph is the total expression of (*H*) Insulin signaling pathway and (*I*) FoxO signaling pathway positive longevity-related genes in the liver and brain, respectively (y-axis on the left). Black line is the relative value of life-history variable (y-axis on the right). Species are shown at the bottom. All values are in standard scale. (*J, K, L*) Venn diagrams of each tissue for longevity-related genes.

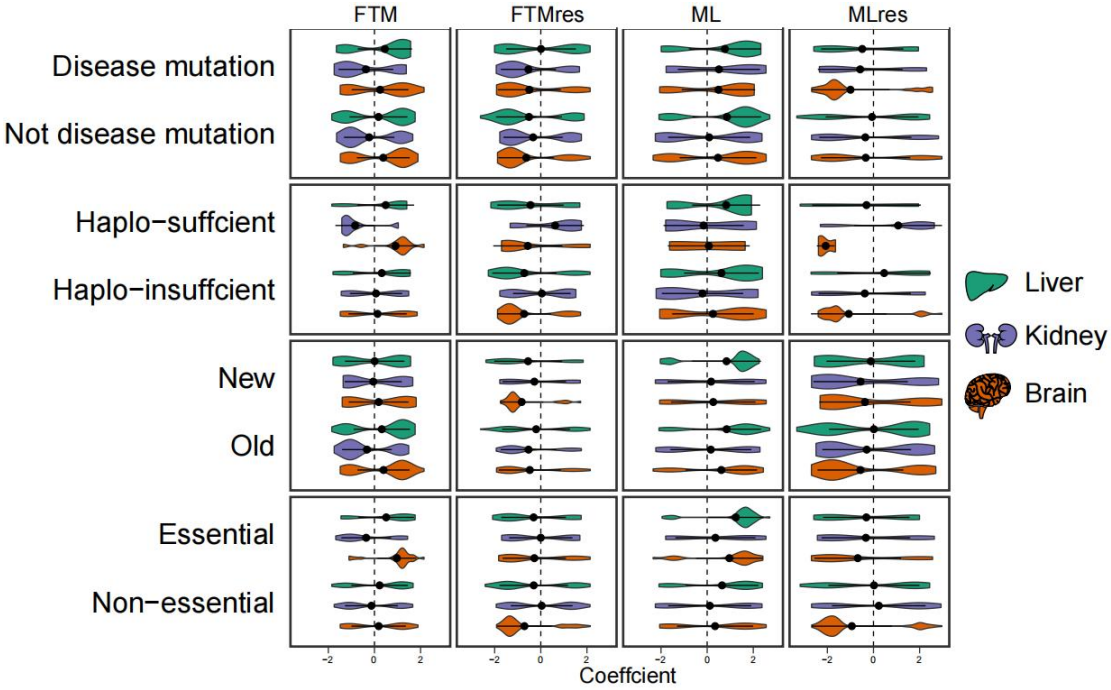



**Fig. 3. Comparison between the variation of gene expression and the gradient of longevity traits among different gene categories.** On the x-axis, the coefficient represents the rate of gene expression variation with life-history gradient. The violin plot and black dots represent the mean of the variation rate in a specific functional gene category, respectively.

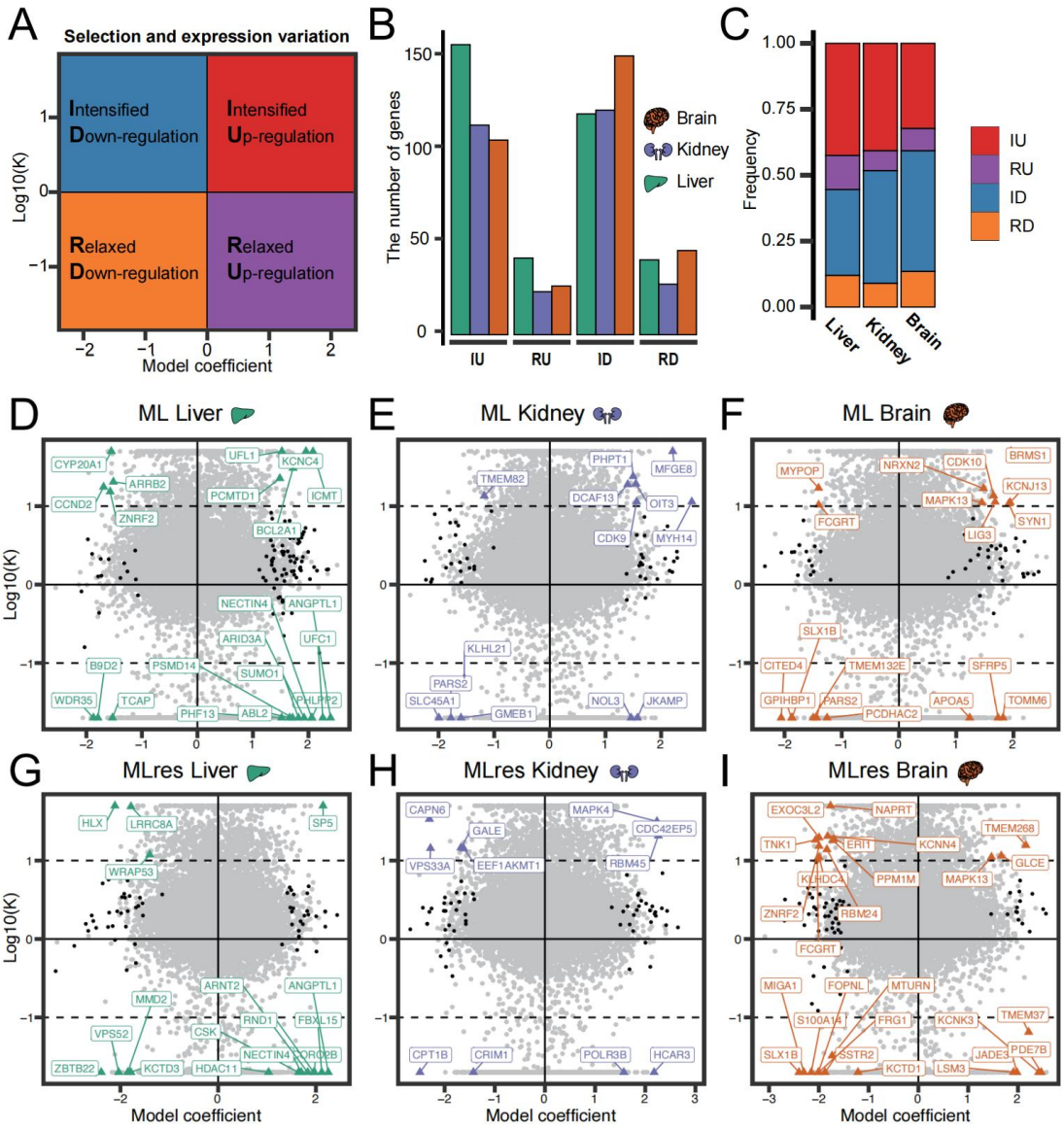



**Fig. 4. Pattern of association between selection index (*k*) and coefficients, and longevity.** (*A*) Colored picture illustrating the annotation of gene classification. (*B*) Number of genes in the three organs assigned to different gene classifications. IU stands for genes that are positively related to longevity and are under intensified selection; RU stands for genes that are positively related to longevity and are under relaxed choice; ID represents a gene that is negatively related to longevity and is under intensified selection; RD stands for genes that are negatively related to longevity and are under relaxed selection. (*C*) Cumulative frequency histogram showing the distribution of gene types in different organs. (*D-I*) Scatter plot showing the log10-transformed relaxation parameter (*k*) on the y-axis, and the variation rate of gene expression along the longevity trait gradient on the x-axis. The colored points represent longevity-related genes with strong selection signals (the color corresponds to the tissue type shown in panel A), the black points represent longevity-related genes with weak selection signals, and the grey points represent genes that are not significant.