Positive selection and enhancer evolution shaped lifespan and body mass in great apes

Daniela Tejada-Martinez[1,2,*,+], Roberto A. Avelar[2,+], , Inês Lopes[2], Bruce Zhang[3], Guy Novoa[4], João Pedro de Magalhães[2] and Marco Trizzino[1,*]

[1] Department of Biochemistry and Molecular Biology, Thomas Jefferson University, Philadelphia, PA, USA

[2] Integrative Genomics of Ageing Group, Institute of Ageing and Chronic Disease, University of Liverpool, Liverpool, UK

[3] Institute of Healthy Ageing, and Research Department of Genetics, Evolution and Environment, University College London, London WC1E 6BT, UK.

[4] Centro Nacional de Biotecnología - CSIC, Spain

+ Co-first authors
* **Corresponding authors**: E-mails: marco.trizzino@jefferson.edu; dtejadamartinez@gmail.com

Abstract

Within primates, the great apes are outliers both in terms of body size and lifespan, since they include the largest and longest-lived species in the order. Yet, the molecular bases underlying such features are poorly understood. Here, we leveraged an integrated approach to investigate multiple sources of molecular variation across primates, focusing on ~1,550 genes previously described as tumor suppressors, oncogenes, ageing genes in addition to a novel Build of the CellAge database of cell-senescence genes (version 2), herein presented for the first time. Specifically, we analyzed dN/dS rates, positive selection, gene expression (RNA-seq) and gene regulation (ChIP-seq). By analyzing the correlation between dN/dS, maximum lifespan and body mass we identified 67 genes that in primates co-evolved with those traits. Further, we identified 6 genes, important for immunity, neurodevelopment and telomere maintenance (including *TERF2*), under positive selection in the great ape ancestor. RNA-seq data, generated from the liver of six species representing all the primate lineages, revealed that ~8% of the longevity genes are differentially expressed in apes relative to other primates. Importantly, by integrating RNA-seq with ChIP-seq for H3K27ac (which marks active enhancers), we show that the differentially expressed longevity genes are significantly more likely than expected to be located near a novel "ape-specific" enhancer. Moreover, these particular ape-specific enhancers are enriched for young transposable elements, and specifically SINE-*Vntr*-Alus (SVAs). In summary, we demonstrate that multiple evolutionary forces have contributed to the evolution of lifespan and body size in primates.

Keywords: Ageing, cancer, evolutionary genomics, cis-regulatory evolution, natural selections, apes, transposable elements, cell senescence.

Introduction

Uncovering the molecular and evolutionary bases of ageing in the tree of life is important to increase our understanding of the natural mechanisms of disease resistance (Galis and Metz 2003; Caulin and Maley 2011; Ganten and Nesse 2012; Nesse et al. 2012; Thomas et al. 2013; Petralia et al. 2014; Abegglen et al. 2015; Harris et al. 2017; Ciccarelli and DeGregori 2020). The evolution of lifespan in animals has been shaped by selective pressures arising from different environments, ecological niches, habitats, and diets (Healy et al. 2014; Tollis et al. 2017; Kacprzyk et al. 2021). Ageing and senescence are the result of gradual declines in biological functions, which lead to increased vulnerability, disease predisposition, and ultimately death (López-Otín et al. 2013). With the exception of a few species that show no signs of ageing (e.g. some jellyfishes and hydras; (Petralia et al. 2014), most animals undergo this process. Nevertheless, maximum lifespan is a highly variable trait across the tree of life, suggesting that long-lived species may have adopted several evolutionary strategies which may have favored longevity by reducing disease susceptibility (Gorbunova et al. 2014; Kacprzyk et al. 2021; Yu et al. 2021). Yet, most of these strategies remain unknown.

One of the main consequences of ageing is the increased risk of developing cancer. Cancer has been reported in almost all multicellular organisms (Aktipis et al. 2013; Albuquerque et al. 2018), and it is often associated with somatic genetic mutations that inactivate tumor suppressor genes or activate oncogenes. These mutations ultimately affect cell metabolism, promoting uncontrolled cell division (Stratton et al. 2009; Tomasetti and Vogelstein 2015). Statistically, animal species with larger body sizes and longer lifespans accumulate more cell divisions during their lives and are therefore expected to accumulate more deleterious genetic mutations. Nonetheless, large and long-lived species overall do not develop cancer at higher frequency than smaller species (Nunney 2018; Seluanov et al. 2018). This scientific paradox was first noted by Richard Peto (the "Peto's Paradox"; Peto et al. 1977; Nunney 1999).

Specific anti-cancer mechanisms have been discovered in long-lived species. These mechanisms include early contact inhibition in naked mole rats (Seluanov et al. 2009; Tian et al. 2013), positive selection in DNA repair & inflammatory genes in the giant tortoises from Galapagos (Quesada et al. 2019), and immunity pathways, telomere maintenance and cellular senescence (CS) genes in bats (Zhang et al. 2013; Foley et al. 2018; Tian et al. 2018; Wilkinson and Adams 2019). Particularly, CS - a process in which otherwise replicating cells reach the maximum number of divisions and cease proliferating - is another important mechanism of cancer suppression; previous studies indicate that genes regulating the senescent phenotype are strongly conserved amongst vertebrates compared to other

72 protein-coding genes (Avelar et al. 2020). Despite the importance of senescence processes in inhibiting
73 cancer, CS itself is detrimental for health; senescent cells produce pro-inflammatory cytokines that can
74 paradoxically promote cancer (Rao and Jackson 2016). A previous study found that senolytics - drugs
75 that specifically target and eliminate senescent cells - can enhance life- and health-span in old mice (Xu
76 et al. 2018). On another hand, in larger mammals such as elephants and whales, molecular substitutions
77 in cell senescence genes and increases in the copy number of important tumor suppressor genes have
78 contributed to significantly reduce the risk of developing cancer (Abegglen et al. 2015; Caulin et al.
79 2015; Keane et al. 2015; Tollis et al. 2019). Specific changes in the regulation and expression of specific
80 genes have also been associated with ageing (de Magalhães et al. 2009; Donlon et al. 2017; Morris et
81 al. 2019; Chatsirisupachai et al. 2021). Yet, the extent of the contribution of cis-regulatory evolution to
82 ageing is still poorly understood.

83 In primates, maximum lifespan (MLS) and body mass (BM) are correlated, despite being highly
84 variable across species, with great apes (gorillas, orangutans, humans and chimpanzees) representing
85 outliers (Finch and Austad 2012). The BM ranges from 5 kg in the grey mouse lemur (*Microcebus*
86 *murinus*) to ~140kg in the gorilla (*Gorilla gorilla*). Similarly, the MLS varies from 13 years in the
87 calabar angwantibo (*Arctocebus calabarensis*) to 55–122 years in great apes (AnAge database:
88 https://genomics.senescence.info/; Tacutu et al. 2018). Outside great apes, just a few primate species
89 can reach 50 years (e.g. the capuchin monkeys; Muntané et al. 2018; Orkin et al. 2021).

90 Consistent with Peto's Paradox, great apes develop cancer at lower rates than other primates
91 (Fowler et al. 1980; Lowenstine 1986; Cho et al. 2007; Lowenstine et al. 2016). Humans represent an
92 exception, possibly as a consequence of their life-style (Finch 2010; Hochberg and Noble 2017;
93 Albuquerque et al. 2018). In fact, in some human populations cancer rates are ~25% (Ferlay et al. 2015),
94 and some cancer types seem to be unique to our species (e.g prostate cancer and lung cancer). Several
95 studies have recently investigated the link between longevity, body mass and disease resistance in
96 several mammalian species (Tollis et al. 2017; Muntané et al. 2018; Boddy et al. 2020; Huang et al.
97 2021; Kacprzyk et al. 2021). On the other hand, the mechanisms of ageing in great apes are largely
98 uncharacterized, and several questions remain unanswered.

99 Here, we aimed at investigating the relative contribution of different sources of molecular
100 variation (molecular evolution in coding genes, positive selection, cis-regulatory and gene expression
101 evolution) to the evolution of longevity in great apes. First, we performed a literature search for genes
102 that are involved in the CS process. This led to the generation of a novel Build (version 2) of the CellAge

database, which includes 934 genes that induce or inhibit cellular senescence *in vitro* in human cell lines following genetic manipulation (Table S2).

Next, we interrogated 19 mammalian genomes, focusing on ~1,550 longevity-related genes, including the genes in the novel CellAge database, plus hundreds of genes previously associated with tumor suppression, oncogenic activity and ageing-related genes processes. We evaluated: 1) potential correlations between MLS, BM and coding gene evolution [non-synonymous/synonymous mutations]; 2) positive selection signal in the great-ape ancestor; 3) species-specific gene expression patterns (RNA-seq); 4) species-specific gene-regulation patterns (ChIP-seq for H3 lysine 27 acetylation [H3K27ac], which marks active cis-regulatory elements).

We identified 67 longevity genes that co-evolved with lifespan and/or body mass in primates, and 6 longevity genes under positive selection in the great-ape ancestor. These genes are important for immunity, neurodevelopment and telomere maintenance. In addition, RNA-seq and ChIP-seq data, generated from the liver of six species representing all the primate lineages (Trizzino et al. 2017), revealed that ~8% of the longevity genes investigated in this study are differentially expressed in apes relative to other primates, and that these lineage-specific gene expression patterns are associated with the rise of novel ape-specific enhancers, most of which derived from the insertion of young transposable elements.

Results

Ape-specific patterns in the evolution of lifespan and body mass

The great apes are the primates with the largest body size, and are also the most long-lived species in the entire order (fig. 1a-b). Given the high diversity of phenotypes across primates, we used Phylogenetic Generalized Least Squares (PGLS) models (Orme et al. 2012; Revell 2012; Pennell et al. 2014) to evaluate if body mass (BM) and maximum lifespan (MLS) evolved independently in great apes. The PGLS estimate the covariance among traits, MLS and BM in this case, taking into account the effect of the phylogenetic signal across the tree. We found that the allometric expectations are maintained across primates, with a positive correlation between BM and MLS (fig 1a). Our model revealed that the BM predicts ~6% of the variation in longevity across primate species (adjusted R-squared: 0.05904, 131 degrees of freedom, p = 0.0028). Nonetheless, the relation between MLS and BM in great apes is significantly different from the other primates (adjusted R-squared: 0.04568, degrees of freedom, p = 0.02888). This suggests that lineage specific molecular evolution may have shaped lifespan and body

134 mass in great apes.

135

136 *Build 2 of the CellAge Database of Cell Senescence Genes*

137 We compiled a Build 2 of the CellAge dataset by means of a scientific literature search of gene
138 manipulation experiments in primary, immortalized, or cancer human cell lines that caused cells to
139 induce or inhibit CS. The novel CellAge build comprises 934 distinct CS genes, of which 608 genes
140 affect replicative CS (our default annotation), 180 genes affect stress-induced CS, and 214 genes are
141 related to oncogene-induced CS. Of the 934 total genes, 393 genes induce CS (~ 42.1%), 518 inhibit it
142 (~ 55.5%), and 23 genes have unclear effects, both inducing and inhibiting CS depending on
143 experimental conditions (~ 2.5%). The genes in the dataset are also classified according to the
144 experimental context used to determine these associations (Table S2).

145

146 *The rate of evolution (dN/dS) in 67 genes positively correlates with the evolution of longevity and body*
147 *mass across primates*

148 The ratio between synonymous (dS) and nonsynonymous substitutions (dN) is considered a reliable
149 measure of natural selection in the evolution of the phenotypic diversity. In line with this, several studies
150 have reported a positive relationship between dN/dS and either body mass or maximum lifespan
151 (Romiguier et al. 2010; Weber et al. 2014; Figuet et al. 2016). Nonetheless, the contribution of individual
152 genes to the evolution of body size and lifespan remains largely unexplored, especially in primates. Even
153 if MLS and BM are positively correlated in most mammalian lineages, their molecular evolution is likely
154 largely independent, as has been previously reported in rodents (Seluanov et al. 2018). To investigate
155 the molecular evolution of longevity and body mass in primates, we focused on a set of 2,268 genes
156 previously implicated with lifespan, body mass and cancer (hereafter "longevity genes"). This list
157 includes genes that have been previously associated with tumor suppression, with oncogenic functions,
158 and with ageing and senescence-related processes (Table S1.2). Of these 2,268 genes, 1,553 have a six-
159 way ortholog in six species (human, chimpanzee, rhesus macaque, marmoset, bushbaby, mouse lemur),
160 representative of the main primate lineages (Catarrhini, Platyrrhini and Strepsirrhine), for which we
161 have also analyzed publicly available next generation sequencing data (see below). These 1,553 genes
162 were retained for downstream analyses. Additionally, 986/1,553 (~63.5%) of these orthologous genes
163 are also present in 11 additional primate species that we included in our molecular evolution analysis
164 (fig. 1b-c, Table S1.3). In general, the number of orthologous genes was similar across primates, with a

variation between 1,905 genes in the bushbaby to 2,183 in the chimpanzee. The species with the smallest number of orthologous genes was the tarsier (*Carlito syrichta*), likely reflecting the low sequencing coverage for the genome of this species (Table S1.1).

We correlated the rates of evolution (dN/dS) in the 986 genes with the two life history traits (body mass [BM], and maximum lifespan [MLS]) together and independently (Table S1.4-1.6). As a first step, we removed 89 genes with $\omega>2$ due to the possible overestimation of ω in the branch due to saturation or miss calculation of synonymous (dN) or non-synonymous (dN) substitutions. Next, we used the variance of the residuals from the first PGLS model (MLS ~ BM), in order to account for the expected covariance structure between the two variables and evaluate the contribution of each gene (dN/dS) in the evolution of MLS and BM simultaneously. As a result, we found 9 genes for which the dN/dS is positively correlated with both MLS and BM (figure 2a, table 1): five are tumor suppressor genes (TSGs: *PRDM2 , ADAMTS8 , NPRL2 , SGMS1 , LRIG1*), one is an TSG-oncogene (*NCOA4*), one is categorized as TSG-oncogene-senescence gene (*ATF3*) and one is an ageing gene (*TRAP1*). These genes are involved in important biological pathways. For instance, *PRDM2* is involved in cell differentiation (Fog et al. 2012) and *NCOA4* in iron signaling, which is a key autophagic cell death process (Bellelli et al. 2016). Notably, *ADAMTS8* contributes to ~30% of the variation in MLS and BM across primates. This gene encodes for a protein involved in several types of cancer (Zhao et al. 2018), it is important for the immune response (Zhang et al. 2020) and has been associated with multiple diseases such as rheumatoid arthritis (Baloğlu et al. 2020). Importantly, *ADAMTS8* has also been reported as a positively selected gene in whales, the longest-lived/most cancer-resistant group of species across mammals (Tejada-Martinez et al. 2021). Gene enrichment analysis on the nine genes ($p<0.05$, Table 1, Table S1.7; complete list of longevity genes used as background) revealed an overrepresentation in ontology process related with response to amino acid starvation (*NPRL2* and *ATF3*), which could reflect the important role that the diet plays in adaptation, longevity and species diversification (Yabu et al. 2012; Gallinetti et al. 2013; Slade and Staveley 2016).

In the second PGLS model, we only focused exclusively on lifespan. We identified 58 genes that positively coevolved with longevity across primates (fig. 2b). Most of those genes encode for transcriptional regulators (Table S1.7) and are predominantly tumor suppressor or senescence genes. The genes *SALL1*, *SALL4* and *ARID4B* contribute from 37% to 53% of the variation in longevity seen across species (table 1b) and are important developmental regulators (Hirsch et al. 2015; Buttgereit et al. 2016; Wu et al. 2019; Bon-Baret et al. 2021; GÜven and Terzİ ÇİzmecİoĞlu 2021). Among the other

196 genes, *WIPI1* is important for autophagy (Grimmel et al. 2015), *HIVEP3* for immunity and inflammation
197 processes (Hicar et al. 2001; Krovi et al. 2020), and *FXR1* for the regulation of memory and emotions
198 (Kalueff 2007; Sopova et al. 2019; Brown and Török 2021), important traits that have shaped primate
199 evolution.

200 Finally, in the third PGLS model, we focused exclusively on body mass. This analysis revealed
201 that *FAT1*, a highly conserved tumor suppressor gene involved in Wnt signaling pathway (Nusse and
202 Varmus 1992), and associated with multiple cancers (Morris et al. 2013), contributes to ~22% of the
203 variation in BM found across primates (table 1c).

204 In summary, we found that, across primates, the dN/dS ratio correlates with either longevity or
205 body mass (or both) in a total of 67 genes. These genes are involved in multiple biological functions
206 directly associated with ageing regulation, but are also related to particular features important for
207 primate diversification such as diet, metabolism and neurological development.

208

209 *Longevity genes under positive selection in the ancestor of great apes are involved in neurological*
210 *pathways, immunity and telomere length*

211 The integration of an evolutionary perspective in the study of genes related to human health could reveal
212 the discovery of novel molecular variants that underlie important biological questions, such as how to
213 live longer and healthier lives, or how to be more resistant to diseases. Here, we studied the signal of
214 positive selection in the evolution of 986 longevity genes in the ancestors of great apes, which is the
215 lineage including the species with largest body sizes and longest lifespan across primates. We found six
216 longevity genes with positive selection in the ancestors of great apes (*PAK1, ULK3, KIF1B, STAT1,*
217 *ABCC3, TERF2;* fig. 3a-g; Table S1.8-1.9). These genes are involved in important biological functions
218 that characterized the evolution of long-lived species, such as immunity, neurodevelopment, and
219 telomere maintenance (Zhang et al. 2013; Gorbunova et al. 2014; Orkin et al. 2021). For example, *PAK1*
220 is an oncogene involved in cytoskeletal reorganization and cortical development (Pan et al. 2015). The
221 codons under positive selection are located in a protein kinase domain, indicating that these changes
222 could have modified the structure or function of the gene in the ancestor of the great apes. Mutations in
223 *PAK1* have been associated with macrocephaly in humans (Harms et al. 2018), and memory deficits in
224 mice (Meng et al. 2005). In fact, this gene plays an important role in hippocampal synaptic plasticity
225 (Asrar et al. 2009). Like *PAK1*, *ULK3* is also an important component of several neurological pathways,
226 whose activity is significantly decreased in patients with Parkinson's disease (Miki et al. 2018). *KIF1B*

227 is an axonal precursor which encodes for an important motor protein that transports mitochondria. In
228 humans, *KIF1B* mutations have been associated with autoimmune and neurological diseases, including
229 multiple sclerosis (Aulchenko et al. 2008) and Charcot-Marie-Tooth (Zhao et al. 2001). *KIF1B*-
230 knockout mice do not survive after birth due to atrophies in the nervous system, resembling what is
231 observed in Charcot-Marie-Tooth patients (Zhao et al. 2001). *STAT1* regulates innate immunity and
232 protection against viral infection (Karst et al. 2003; Hossain et al. 2019). *ABCC3* has been implicated in
233 drug resistance, including chemotherapy (Balaji et al. 2016). Finally, *TERF2* is very important for
234 ageing-related processes, since it regulates telomere length (Ilyenko et al. 2011; Tian et al. 2019). For
235 this gene, the great apes show a specific substitution in the HTH myb-type Domain, a highly conserved
236 amino-terminal region (Peters et al. 1987), suggesting that the substitution could potentially have
237 affected the myb-type/DNA interaction important for TERF2 protein function.
238
239 *Longevity genes with human-specific expression patterns*
240 We reasoned that different strategies may have contributed to the evolution of longevity genes in great
241 apes. On the one hand, positive selection may lead to important changes in the amino acid composition,
242 and thus in the structure and function of the encoded protein. On the other hand, the coding sequence,
243 and hence the structure of the protein, may remain unaltered across species but the expression level of
244 the gene may change. Changes in gene expression would be reflected in the amount of translated protein,
245 which could have a significant impact on cell metabolism. Our data so far demonstrated that a number
246 of longevity genes are under positive selection in great apes. We surmised that changes in the expression
247 of longevity genes may provide an additional contribution to the differences in lifespan and size between
248 the great apes and the other primates. To test this hypothesis, we leveraged a publicly available RNA-
249 seq dataset generated from the liver of the six primate species analyzed in the present study (humans,
250 chimpanzee, rhesus macaque, marmoset, mouse lemur and bushbaby; Trizzino et al. 2017). The liver is
251 particularly relevant for this study because most genes associated with longevity and body mass are
252 highly expressed in the liver, and also because liver gene expression displays high variation across
253 species, likely reflecting adaptation to different diets and environments, in spite of a conserved core
254 function (Trizzino et al. 2017). Consistent with this premise, the liver has been employed in studies that
255 investigated biomarkers of aging (Lee et al. 2012; Bochkis et al. 2014; White et al. 2015).
256 As a first step, we examined the expression patterns of all the 1,553 six-way ortholog longevity
257 genes, and found that genes involved in both ageing & senescence are significantly more expressed in

258 humans relative to all the other primates grouped together (Wilcoxon rank-sum test, $p < 0.001$; fig. 4a).
259 Similarly, genes that are involved in both ageing and tumor suppression are more expressed in humans
260 than in other primates grouped together, although the *p*-value was only marginally significant, possibly
261 as a consequence of the small sample size (N=3; Wilcoxon rank-sum test, $p = 0.057$; fig. 4b).
262
263 *Longevity genes with ape-specific expression patterns*
264 Next, we compared the expression of the longevity genes between apes and other primates. Overall,
265 apes have a higher number of longevity genes expressed in the liver relative to other primates (fig. 4c;
266 Spearman's correlation coefficient R=0.53, *p*-value = 2.2e-16), and the expression level of the longevity
267 genes is strongly correlated between humans and chimpanzees (fig. 4d). We decomposed the longevity
268 genes across the different categories (fig. 4e), and observed that the oncogenes are significantly more
269 expressed in apes relative to other primates grouped together (Wilcoxon rank-sum test, $p = 0.027$), while
270 the senescence genes are significantly less expressed in apes (Wilcoxon rank-sum test, $p = 7.9 \times 10^{-10}$).
271 This pattern could reflect the findings from recent studies that reported that senescence genes are
272 beneficial in the younger ages, during which they increase reproductive fitness, while they might have
273 a negative impact later in life (Campisi 2003; Blagosklonny 2010; Di Micco et al. 2021). Accordingly,
274 gene expression meta-analysis across human tissues showed that cancer genes promote longevity,
275 whereas senescence genes showed anti-longevity patterns (Chatsirisupachai et al. 2019).
276
277 *Ape-specific enhancers drive the differential expression of longevity genes in apes*
278 We performed a differential gene expression analysis, comparing the expression of the longevity genes
279 in apes (human, chimpanzee) vs the other 4 primate species grouped together (rhesus macaque,
280 marmoset, bushbaby, mouse lemur). Overall, 122/1,553 (~7.9%) longevity genes were found to be
281 differentially expressed between apes and the other primates grouped together (DESeq2 FDR < 5%;
282 Supplemental Table S3). Of those, 61 genes were upregulated and 61 downregulated. We investigated
283 if cis-regulatory evolution could underlie the changes in expression of the 122 genes identified as
284 differentially expressed in apes compared to other primates. To this purpose, we leveraged H3K27ac
285 ChIP-seq data generated in the same study from the same liver samples of the same individuals (Trizzino
286 et al. 2017). Namely, this specific histone modifications marks active cis-regulatory elements (enhancers
287 and promoters). Importantly, in the original study, Trizzino and collaborators (2017) identified a set of
288 ape-specific enhancers. We thus scanned the surrounding regions of the 122 differentially expressed

genes found in the apes vs other primates comparison (fig. 5a). We focused on the 50 kilo-bases (kbs) up- down-stream of the transcription start site (TSS) of the 122 genes, and observed that 27 of the 122 differentially expressed genes (22.1%) have an ape-specific enhancer in the 50 kbs surrounding the TSS (38 total enhancers, median distance from TSS = 19.8 kb; fig. 5a; Table S3). To test if this number is higher than expected by chance, we performed a permutation test (1000 permutations), by randomly extracting (for 1000 times) 122 genes from a list including all the annotated human genes (Ensembl), and found that, on average, only 4/122 random genes (3.3%) were located within 50 kbs from an ape-specific enhancer. Together, these data indicate that the longevity genes identified as differentially expressed in apes vs other primates are ~six times more likely than expected to be found near an ape-specific enhancer (22.1% vs 3.3%; Permutation Test $p < 2.2 \times 10^{-16}$; fig. 5a).

Ape-specific enhancers associated with differentially expressed longevity genes are derived from recent Transposable Element insertions

Overall, a total of 38 ape-specific enhancers were found in the 50 kbs surrounding 27 of the 122 differentially expressed longevity genes (Table S1.10). Since multiple studies have demonstrated that Transposable Elements (TEs) are an important source of gene regulatory novelty in the primate gene regulation (Jacques et al. 2013; Chuong et al. 2016; Trizzino et al. 2017; Trizzino et al. 2018), we tested for potential association between the 38 ape-specific enhancers and TE insertions. We found that 20/38 (~52.6%) ape-specific enhancers are derived from TE insertions (fig. 5b). This was not higher than expected by chance, considering that TEs represent ~50% of the human genome. Nonetheless, TE-derived ape-specific enhancers were significantly more associated with upregulated (65%) than to downregulated (35%) longevity genes (fig. 5c, one-tailed Fisher's Exact Test $p = 0.0425$). Notably, *ADAMTS8,* a gene reported as under positive selection in other long-lived species (whales; Tejada-Martinez et al. 2021), and which in this study we identified as correlated with lifespan (fig. 2a), is upregulated in apes and is located near a TE-derived ape-specific enhancer.

Finally, we investigated if specific TE families were overrepresented in the TE-derived ape-specific enhancers associated with differentially expressed longevity genes (fig. 5d). Notably, 15% of those were derived from SINE-Vntr-Alu (SVA) insertions. SVAs are the youngest TE family, with most copies being either ape- or human-specific, and represent only 0.1% of all human (and 0.3% of chimpanzee's) annotated TEs. Overall, our data suggest that SVAs are significantly overrepresented in our set of ape-specific enhancers associated with differentially expressed longevity genes (expected:

320 0.1% - observed 15%; Fisher's Exact Test $p < 2.2 \times 10^{-16}$). These findings are consistent with recent
321 studies which demonstrated that SVA insertions contribute to human gene-regulatory networks
322 (Trizzino et al. 2017; Trizzino et al. 2018; Pontis et al. 2019). Similarly, the LTR transposons were also
323 overrepresented in our set of ape-specific enhancers (Fisher's Exact Test $p < 0.0001$). These findings
324 are also consistent with recent literature on the contribution of LTRs to human gene regulation (Wang
325 et al. 2007; Cohen et al. 2009; Sundaram et al. 2014; Chuong et al. 2016; Janoušek et al. 2016; Fuentes
326 2018).
327 In summary our analysis revealed that lineage specific enhancers may have contributed to the
328 evolution of the expression of longevity genes in great apes and that young transposon insertions may
329 have had a significant role in this process.
330
331 **Discussion**
332 Natural selection and gene regulatory evolution likely underlay the evolution of most phenotypic traits.
333 Here, we specifically focused on the evolution of lifespan and body mass in great apes. This primate
334 lineage includes species which evolved with a lifespan and body mass significantly different from the
335 rest of primates. We carried out comparative genomic analyses, examining different sources of
336 molecular variation, including both coding genes and non-coding genomic regions (i.e. cis-regulatory
337 elements: enhancers, and promoters). Cis-regulatory evolution plays an essential role in phenotypic
338 diversification (Trizzino et al. 2017; Berthelot et al. 2018; Farré et al. 2019; Feigin et al. 2019; Sundaram
339 and Wysocka 2020; Marand et al. 2021), and can lead to evolutionary innovations in disease resistance
340 across species (Gorbunova et al. 2007; MacRae et al. 2015; Tollis et al. 2020).
341 Our comparative genomic analysis revealed that 67 genes positively correlated with the
342 evolution of maximum lifespan and body mass in primates. Out of those, 85% co-evolved exclusively
343 with maximum lifespan, while the remaining 15% co-evolved with lifespan and body mass together, or
344 with body mass alone. These genes are mostly involved in tumor suppression and senescence-related
345 processes. Since the increase in longevity elicits a greater risk to develop cancer, the rise and fixation of
346 new molecular variants (non-synonymous substitutions) could have positively contributed to improve
347 the functionality of the tumor suppressor genes and thus be beneficial for the evolution of the most
348 recent primate lineage. Similar outcomes have been recently reported by a study on mammals focused
349 on copy number variation in tumor suppressor genes (Sulak et al. 2016; Tollis et al. 2020). Consistent
350 with these lines of evidence, previous studies have found that the rate of protein evolution in anti-cancer

351 and DNA damage response genes is accelerated in long-lived mammals (Li and de Magalhães 2013;
352 Tollis et al. 2019; Tejada-Martinez et al. 2021).

353 A greater lifespan correlates with a higher likelihood of being exposed to pathogens and
354 diseases. Consequently, the immune system of long-lived species is expected to be under particularly
355 strong pressure to sustain the prolonged arm-race with the pathogens (Quesada et al. 2019; Singh et al.
356 2019). Consistent with this premise, we found a significant number of immune, inflammation, and DNA
357 repair genes that coevolved with longevity in primates. Notably, several of those genes were also under
358 positive selection in the ancestors of great apes. This points towards an evolutionary strategy in which
359 disease resistance and lifespan evolved together in primates, and particularly in apes, as it has been
360 reported for other long-lived species (Harris et al. 2017; Muntané et al. 2018; Vazquez et al. 2018).

361 In the great ape ancestor, we identified genes under positive selection involved in both the
362 immune system and in the regulation of telomere length. Importantly, telomeres play a key role in
363 ageing-related processes, and the regulation of telomere length is an important mechanism to control
364 cancer in species with increases in body mass (Seluanov et al. 2008). Namely, telomeres shorten with
365 age, and this leads to increased frequency of chromosomal damage, cell senescence and apoptosis (Tian
366 et al. 2018). It has been proposed that in large-size and long-lived species the rate of telomere shortening
367 is slower than in smaller and shorter-lived species (Whittemore et al. 2019). Based on our data, we
368 speculate that genes involved in telomere function and length may have evolved to help great apes fight
369 diseases, ultimately contributing to their longer lifespan.

370 Among the genes with positive selections in apes, we identified several associated with diet,
371 metabolism, as well as with the regulation of memory and emotions, all of which are important great
372 ape features. This finding is consistent with several studies that have proposed a link between increased
373 life span and increased cognitive functions (Ghirlanda et al. 2014; Barton and Venditti 2017; Orkin et
374 al. 2021). Characterizing their molecular evolution could help us to better understand the rise of genetic
375 neurological disorders in humans, such as macrocephaly, cognitive disorders and Parkinson's disease.
376 Different molecular strategies may have contributed to the evolution of longevity genes across primates.
377 Positive selection can lead to changes in the structure and function of a protein through amino acid
378 substitutions, without affecting gene expression, and thus the amount of protein produced by a specific
379 cell-type. In parallel, the expression level of genes important for longevity and body mass could be
380 affected by evolutionary changes in the associated gene regulatory network. For example, a novel,
381 species-specific, enhancer could influence the expression of the associated gene, ultimately leading to

382 higher levels of protein production without any modification to the structure of the protein itself. Here,
383 using RNA-seq we have investigated how the expression of the longevity genes varies across primates,
384 using the liver as a proof of principle. We demonstrated that oncogenes are significantly more expressed,
385 and senescence genes less expressed, in great apes relative to other primates, suggesting a tradeoff
386 between living longer (which leads to increased likelihood to produce offspring) and disease
387 susceptibility later in life (Campisi 2003; Blagosklonny 2010; Rodríguez et al. 2017; Di Micco et al.
388 2021).

389 Overall, we report that ~8% of the longevity genes tested in this work were differentially
390 expressed in the liver of great apes relative to higher primates, and we demonstrate that lineage specific
391 (i.e. ape-specific) enhancers have significantly contributed to this process. In fact, longevity genes
392 differentially expressed between great apes and other primates are ~six times more likely than expected
393 to be located near an ape-specific enhancer. Consistent with recent studies that suggested an important
394 contribution of young transposable elements (TEs) in the evolution of primate gene regulation (Jacques
395 et al. 2013; Trizzino et al. 2017; Pontis et al. 2019), we demonstrate that SVA transposons are
396 significantly overrepresented in the DNA sequence of ape-specific enhancers located near differentially
397 expressed longevity genes. The SVAs are the youngest family of TEs, and include ape- and/or human
398 specific copies.

399 In summary, our work has shed light on the evolution of thousands of longevity genes in great
400 apes, highlighting a significant contribution to this process for both positive selection on coding genes,
401 as well as evolutionary changes in the non-coding regulatory regions.

402

403 **Materials and Methods**

404 *Relationship between body mass and longevity across primates*

405 To determine if the body mass contributes to the variation in longevity independently in primates relative
406 to other mammals, and to assess ape-specific patterns, we examined the evolution of the life history
407 traits maximum lifespan (MLS) and adult body mass (BM) across 932 mammals. We fit a linear
408 regression through the phylogenetic generalized least squares (PGLS; (Orme et al. 2012), while
409 controlling for potential phylogenetic signals. Two independent regression models were implemented:
410 the first one across primates, and in the second model, the regression was performed between great apes
411 vs other primates (log10 MLS ~ log10 BM * great_apes). The phylogenetic tree used was derived from
412 Uyeda et al. (2017) (Uyeda et al. 2017) and the MLS and BM data were gathered from AnAge database

413 (Tacutu et al. 2018).

414

415 *Genomic sampling, longevity associated genes and build 2 of CellAge Database of Cell Senescence*
416 *Genes*

417 To study the evolution of longevity in primates we carried out a phylogenetic design that included
418 genomes from 19 mammalian species. Our taxonomic sampling included 3 species from the family
419 Hominidae (*Homo sapiens* - human, *Pan troglodytes* - chimpanzee and *Pongo abelii* - sumatran
420 orangutan), one representant from Hylobatidae (*Nomascus leucogenys* - white-cheeked gibbon), three
421 Cercopithecidae (*Macaca mulatta* - Rhesus macaque, *Rhinopithecus roxellana* - Golden snub-nosed
422 monkey and *Chlorocebus sabaeus* - vervet), four Platyrrhini (*Callithrix jacchus* - white tufted ear
423 marmoset, *Aotus nancymaae* - night monkey, *Saimiri boliviensis* - bolivian squirrel monkey, and *Cebus*
424 *imitator* - white-faced capuchin), two Strepsirrhine (*Otolemur garnettii* - bushbaby and *Microcebus*
425 *murinus* - Gray mouse lemur) and finally 5 other mammalian species outside of primates (*Tupaia*
426 *chinensis* - Chinese tree shrew, *Mus musculus* - common mice, *Canis lupus familiaris* - dog, *Bos taurus*
427 - cow and *Sus scrofa* - pig).

428 The coding sequences of each species were downloaded from Ensembl v.96 and NCBI
429 databases (Table S1). To remove low quality records, sequences were clustered using CD-HIT-est
430 v.4.6(Fu et al. 2012) with a sequence identity threshold of 90% and an alignment coverage control of
431 80%. The longest open reading frame was kept using TransDecoder LongOrfs and TransDecoder-
432 predicted in TransDecoder v3.0.1 (https://github.com/TransDecoder/TransDecoder/).

433 The longevity associated coding genes, 2,268 in total, were gathered in four different categories:
434 ageing genes from GenAge Database (build 20 - 307 protein coding genes,
435 https://genomics.senescence.info/genes/index.html); TSGs from the Tumor Suppressor gene database
436 (TSGene 2.0 - 1018 protein coding genes, https://bioinfo.uth.edu/TSGene/; (Zhao et al. 2013);
437 oncogenes from the Oncogene database (698 protein coding genes, http://ongene.bioinfo-minzhao.org/;
438 (Liu et al. 2017); and finally, we compiled build 2 of the CellAge Database of Cell Senescence Genes
439 with an additional 655 genes, complementary to the 279 genes previously reported;
440 https://genomics.senescence.info/cells/, (Avelar et al. 2020).

441 Build 2 of CellAge was compiled in the same way as build 1. A robust scientific literature search
442 was performed for relevant papers before curation and annotation; genes were appended to the database
443 if they met the following criteria:

444 · Only gene manipulation experiments were used to identify the role of the genes in CS to
445 ensure objectivity in the selection process.
446 · The genetic manipulation caused cells to induce or inhibit the CS process in wet lab
447 experiments. CS was detected by growth arrest, increased SA-β-galactosidase activity, SA-
448 heterochromatin foci, a decrease in BrdU incorporation, changes in morphology, and/or
449 specific gene expression signatures such as p21 and p16.
450 · The experiments were performed in primary, immortalized, or cancer human cell lines.
451

452 The resulting list comprised 934 genes that in human cell lines can induce or inhibit the CS process
453 (Table S2). The full CellAge database, including all annotations regarding cell types and cell lines is
454 available at https://genomics.senescence.info/cells/.

455

456 *Homology inference*

457 We inferred homology relationships among the 2,268 longevity genes and the 19 mammalian species
458 included in our study using the program OMA standalone *v*.2.3.1 (Roth et al. 2008; Altenhoff et al.
459 2019). We inferred the OMA Groups (OG), containing the sets of orthologous genes, for which we
460 performed natural selection analyses. The amino acid sequences were aligned using the L-INS-i
461 algorithm from MAFFT v.7 (Katoh and Standley 2013) and the codon alignments using the function
462 pxaa2cdn in phyx (Brown et al. 2017). Finally, to reduce the chance of false positives given for low
463 quality alignment regions we used the codon.clean.msa algorithm of the rphast package (Hubisz et al.
464 2011) with the associated longevity gene from the OG as reference sequence.

465

466 *Evolution of longevity genes and life history traits*

467 To evaluate the possible co-evolution between the longevity genes and the life history traits MLS and
468 BM in primates, the rate of evolution (ω=dN/dS) was calculated. The ω ratio per gene was estimated for
469 each tip in the 14 primates species using the branch model (Yang 2007) as is implemented in ETE-
470 toolkit with the *ete-evol* function (Huerta-Cepas et al. 2016). To calculate the ω ratios the treeshrew,
471 dog, cow and pig were used as outgroups. Three independent PGLS models were performed per OG: 1)
472 between the ω and the MLS; 2) between the ω and the BM; and 3) between the ω and the PGLS residuals
473 between the MLS ~ BM. Previous to the analysis, all variables were transformed to log10. The genes
474 with a dN/dS>2 were removed from the analysis.

475

476 *Positive selection analysis*

477 To evaluate the role of natural selection in the evolution of longevity genes in the ancestor of great apes,
478 we used codon-based models through a maximum likelihood approach using the program PAML v4.9
479 (Yang 2007), as is implemented in ETE-toolkit with the *ete-evol* function (Huerta-Cepas et al. 2016).
480 We calculated the branch-site model in order to estimate changes in the ω value of individual sites. We
481 compared the null model, where the ω value in the foreground branch was set to 1, with the model in
482 which the ω value was estimated from the data (Zhang 2005). The comparisons between models were
483 made using likelihood ratio tests (LRT) and the p-values from the LRT were corrected with FDR
484 (Benjamini and Hochberg 1995). All the genes with positive selection signal but with a LRT > 1 were
485 manually checked for false positives.

486

487 *RNA-seq analysis*

488 To evaluate the contribution of the gene regulation to the evolution of longevity genes in great apes, we
489 took advantage of recently published RNAseq and Chip-seq data for histone H3 lysine 27 acetylation
490 (H3K27ac) from liver of six species representatives of the main groups of primates (human, chimpanzee,
491 rhesus macaque, marmoset, bushbaby and gray mouse lemur; (Trizzino et al. 2017). The list of the genes
492 differentially expressed between Apes and other primates was downloaded from the Supplementary
493 Materials of the same paper (Trizzino et al. 2017). To generate box-plots comparing Gene Expression
494 (GE) across species, we leveraged the transcripts per million (TPM), also available from the
495 Supplementary Materials of Trizzino et al (2017). The TPMs were quantile normalized using R version
496 3.6.3 (R Core Team 2020).

497

498 *Enhancer evolution analysis*

499 From the publicly available Chip-seq data (Trizzino et al. 2017), we downloaded the list and coordinates
500 all the enhancers previously identified as "ape-specific" by Trizzino and collaborators. We then selected
501 all the ape-specific enhancers overlapping a region encompassing +/- 50 kbs from the transcription start
502 site (TSS) of each differentially expressed longevity gene using BEDTools v2.29.2 (Quinlan and Hall
503 2010). Based on this association, we then calculated how many longevity genes were both differentially
504 expressed in the ape vs other primates comparison AND also associated with an ape-specific enhancer.
505 To evaluate if this number was higher than expected by chance, we selected 122 random genes in the
506 human genome and examined how many of them had an ape-specific enhancer within 50 kbs of the TSS

507 and assessed statistical significance by means of Fisher's exact test. Finally, using BEDTools, we looked
508 for overlap between transposable elements and ape-specific enhancers associated with genes
509 differentially expressed in ape vs primate comparison. The list and coordinates of the transposable
510 elements were downloaded from the UCSC Genome Browser (RepeatMasker; hg38 assembly).
511
512 *Enrichment analysis*
513 To gain insight into the particular functions of the longevity genes of interest, the enrichment analysis
514 was performed using WebGestaltR v0.4.4 (Liao et al. 2019). We tested for significant pathway
515 associations using the hypergeometric test for Over-Representation Analysis (ORA) (Khatri et al. 2012).
516 We selected the categories of gene ontology, biological processes, molecular pathways, diseases OMIM,
517 and human phenotype and we considered overrepresented categories to be those with a significance
518 level above that of an FDR of 0.01 after correction with the Benjamini-Hochberg multiple test. An
519 independent enrichment analysis was made using as a background the protein-coding genes and relative
520 to the 2,268 longevity genes.
521
522 *Libraries in R associated with the data treatment, statistical analysis and graphics*
523 Bioconductor, msa, ape, geiger, nlme, pythools, caper, ggpubr, gridExtra, tidyverse, grid, corrplot,
524 ggplot2, ggpubr, Hmisc, gmodels, car, DescTools, qqplotr, dplyr, RColorBrewer, calecopal.
525
526 **Acknowledgements**
527 This study was funded by NIH R35 GM138344-01, awarded to M.T. AnAge, GenAge and CellAge are
528 supported by funding from the Biotechnology and Biological Sciences Research Council
529 (BB/R014949/1) to J.P.M. The authors are grateful to Alejandra Tejada-Martinez for the graphical
530 support.
531
532 **Contributions**
533 D.T.M. and M.T. conceived the study and performed the genomic analyses. R.A., J.P.M., B.Z. and G.N.
534 compiled the Build 2 of the CellAge database. I.L. contributed to the Cellage website development.
535 D.T.M., M.T. and R.A. wrote the manuscript. M.T. and D.T.M. directed the research.
536
537

538 **References**

539 Abegglen LM, Caulin AF, Chan A, Lee K, Robinson R, Campbell MS, Kiso WK, Schmitt DL,
540 Waddell PJ, Bhaskara S, et al. 2015. Potential Mechanisms for Cancer Resistance in Elephants
541 and Comparative Cellular Response to DNA Damage in Humans. *JAMA* 314:1850–1860.

542 Aktipis CA, Boddy AM, Gatenby RA, Brown JS, Maley CC. 2013. Life history trade-offs in cancer
543 evolution. *Nat. Rev. Cancer* 13:883–892.

544 Albuquerque TAF, Drummond do Val L, Doherty A, de Magalhães JP. 2018. From humans to hydra:
545 patterns of cancer across the tree of life. *Biol. Rev. Camb. Philos. Soc.* 93:1715–1734.

546 Altenhoff AM, Levy J, Zarowiecki M, Tomiczek B, Warwick Vesztrocy A, Dalquen DA, Müller S,
547 Telford MJ, Glover NM, Dylus D, et al. 2019. OMA standalone: orthology inference among
548 public and custom genomes and transcriptomes. *Genome Res.* 29:1152–1163.

549 Asrar S, Meng Y, Zhou Z, Todorovski Z, Huang WW, Jia Z. 2009. Regulation of hippocampal long-
550 term potentiation by p21-activated protein kinase 1 (PAK1). *Neuropharmacology* 56:73–80.

551 Aulchenko YS, Hoppenbrouwers IA, Ramagopalan SV, Broer L, Jafari N, Hillert J, Link J, Lundström
552 W, Greiner E, Dessa Sadovnick A, et al. 2008. Genetic variation in the KIF1B locus influences
553 susceptibility to multiple sclerosis. *Nat. Genet.* 40:1402–1403.

554 Avelar RA, Ortega JG, Tacutu R, Tyler EJ, Bennett D, Binetti P, Budovsky A, Chatsirisupachai K,
555 Johnson E, Murray A, et al. 2020. A multidimensional systems biology analysis of cellular
556 senescence in aging and disease. *Genome Biol.* 21:91.

557 Balaji SA, Udupa N, Chamallamudi MR, Gupta V, Rangarajan A. 2016. Role of the Drug Transporter
558 ABCC3 in Breast Cancer Chemoresistance. *PLoS One* 11:e0155013.

559 Baloğlu M, Deveci E, Deveci Ş. 2020. ADAMTS-8 activity in synovial cells in patients with
560 rheumatoid arthritis. *Open J. Rheumatol. Autoimmune Dis.* 10:125–131.

561 Barton RA, Venditti C. 2017. Rapid Evolution of the Cerebellum in Humans and Other Great Apes.
562 *Curr. Biol.* 27:1249–1250.

563 Bellelli R, Federico G, Matte' A, Colecchia D, Iolascon A, Chiariello M, Santoro M, De Franceschi L,
564 Carlomagno F. 2016. NCOA4 Deficiency Impairs Systemic Iron Homeostasis. *Cell Rep.* 14:411–
565 421.

566 Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: A practical and powerful
567 approach to multiple testing. *J. R. Stat. Soc.* 57:289–300.

568 Berthelot C, Villar D, Horvath JE, Odom DT, Flicek P. 2018. Complexity and conservation of
569 regulatory landscapes underlie evolutionary resilience of mammalian gene expression. *Nat Ecol*
570 *Evol* 2:152–163.

571 Blagosklonny MV. 2010. Revisiting the antagonistic pleiotropy theory of aging: TOR-driven program
572 and quasi-program. *Cell Cycle* 9:3151–3156.

573 Bochkis IM, Przybylski D, Chen J, Regev A. 2014. Changes in nucleosome occupancy associated with
574 metabolic alterations in aged mammalian liver. *Cell Rep.* 9:996–1006.

575 Boddy AM, Abegglen LM, Pessier AP, Aktipis A, Schiffman JD, Maley CC, Witte C. 2020. Lifetime
576 cancer prevalence and life history traits in mammals. *Evol Med Public Health* 2020:187–195.

577 Bon-Baret V, Chignon A, Boulanger M-C, Li Z, Argaud D, Arsenault BJ, Thériault S, Bossé Y,
578 Mathieu P. 2021. System Genetics Including Causal Inference Identify Immune Targets for
579 Coronary Artery Disease and the Lifespan. *Circ Genom Precis Med* 14:e003196.

580 Brown A, Török M. 2021. Functional amyloids in the human body. *Bioorg. Med. Chem. Lett.*
581 40:127914.

582 Brown JW, Walker JF, Smith SA. 2017. Phyx: phylogenetic tools for unix. *Bioinformatics* [Internet]
583 33:1886–1888. Available from: http://dx.doi.org/10.1093/bioinformatics/btx063

584 Buttgereit A, Lelios I, Yu X, Vrohlings M, Krakoski NR, Gautier EL, Nishinakamura R, Becher B,
585 Greter M. 2016. Sall1 is a transcriptional regulator defining microglia identity and function. *Nat.*
586 *Immunol.* 17:1397–1406.

587 Campisi J. 2003. Cellular senescence and apoptosis: how cellular responses might influence aging
588 phenotypes. *Exp. Gerontol.* 38:5–11.

589 Caulin AF, Graham TA, Wang L-S, Maley CC. 2015. Solutions to Peto's paradox revealed by
590 mathematical modelling and cross-species cancer gene analysis. *Philos. Trans. R. Soc. Lond. B*
591 *Biol. Sci.* [Internet] 370. Available from: http://dx.doi.org/10.1098/rstb.2014.0222

592 Caulin AF, Maley CC. 2011. Peto's Paradox: evolution's prescription for cancer prevention. *Trends*
593 *Ecol. Evol.* 26:175–182.

594 Chatsirisupachai K, Lesluyes T, Paraoan L, Van Loo P, de Magalhães JP. 2021. An integrative
595 analysis of the age-associated multi-omic landscape across cancers. *Nat. Commun.* 12:2345.

596 Chatsirisupachai K, Palmer D, Ferreira S, de Magalhães JP. 2019. A human tissue-specific
597 transcriptomic analysis reveals a complex relationship between aging, cancer, and cellular
598 senescence. *Aging Cell* 18:e13041.

599 Cho H-S, Kim Y-S, Choi C, Lee J-H, Kurkure NV, Subramanian M, Park N-Y. 2007. Neuroendocrine
600 tumor in the lung of a captive black spider monkey (Ateles paniscus). *J. Vet. Med. Sci.* 69:783–
601 785.

602 Chuong EB, Elde NC, Feschotte C. 2016. Regulatory evolution of innate immunity through co-option
603 of endogenous retroviruses. *Science* 351:1083–1087.

604 Ciccarelli FD, DeGregori J. 2020. Approaching Cancer Evolution from Different Angles. *iScience*
605 23:101661.

606 Cohen CJ, Lock WM, Mager DL. 2009. Endogenous retroviral LTRs as promoters for human genes: a
607 critical assessment. *Gene* 448:105–114.

608 Di Micco R, Krizhanovsky V, Baker D, d'Adda di Fagagna F. 2021. Cellular senescence in ageing:
609 from mechanisms to therapeutic opportunities. *Nat. Rev. Mol. Cell Biol.* 22:75–95.

610 Donlon TA, Morris BJ, Chen R, Masaki KH, Allsopp RC, Willcox DC, Elliott A, Willcox BJ. 2017.
611 FOXO3 longevity interactome on chromosome 6. *Aging Cell* 16:1016–1025.

612 Farré M, Kim J, Proskuryakova AA, Zhang Y, Kulemzina AI, Li Q, Zhou Y, Xiong Y, Johnson JL,
613 Perelman PL, et al. 2019. Evolution of gene regulation in ruminants differs between evolutionary
614 breakpoint regions and homologous synteny blocks. *Genome Res.* 29:576–589.

615 Feigin CY, Newton AH, Pask AJ. 2019. Widespread cis-regulatory convergence between the extinct
616 Tasmanian tiger and gray wolf. *Genome Res.* 29:1648–1658.

617 Figuet E, Nabholz B, Bonneau M, Mas Carrio E, Nadachowska-Brzyska K, Ellegren H, Galtier N.
618 2016. Life History Traits, Protein Evolution, and the Nearly Neutral Theory in Amniotes. *Mol.*
619 *Biol. Evol.* 33:1517–1527.

620 Finch CE. 2010. Evolution of the human lifespan and diseases of aging: Roles of infection,
621 inflammation, and nutrition. *Proc. Natl. Acad. Sci. U. S. A.* 107:1718–1724.

622 Finch CE, Austad SN. 2012. Primate aging in the mammalian scheme: the puzzle of extreme variation
623 in brain aging. *Age* 34:1075–1091.

624 Fog CK, Galli GG, Lund AH. 2012. PRDM proteins: important players in differentiation and disease.
625 *Bioessays* 34:50–60.

626 Foley NM, Hughes GM, Huang Z, Clarke M, Jebb D, Whelan CV, Petit EJ, Touzalin F, Farcy O,
627 Jones G, et al. 2018. Growing old, yet staying young: The role of telomeres in bats' exceptional
628 longevity. *Sci Adv* 4:eaao0926.

629 Fowler ME, Benirschke K, Garner DM, Jones TC. 1980. Pathology of Laboratory Animals. *The*
630 *Journal of Zoo Animal Medicine* [Internet] 11:132. Available from:
631 http://dx.doi.org/10.2307/20094494

632 Fuentes A. 2018. How Humans and Apes Are Different, and Why It Matters. *J. Anthropol. Res.*
633 74:151–167.

634 Fu L, Niu B, Zhu Z, Wu S, Li W. 2012. CD-HIT: accelerated for clustering the next-generation
635 sequencing data. *Bioinformatics* [Internet] 28:3150–3152. Available from:
636 http://dx.doi.org/10.1093/bioinformatics/bts565

637 Galis F, Metz JAJ. 2003. Anti-cancer selection as a source of developmental and evolutionary
638 constraints. *Bioessays* 25:1035–1039.

639 Gallinetti J, Harputlugil E, Mitchell JR. 2013. Amino acid sensing in dietary-restriction-mediated
640 longevity: roles of signal-transducing kinases GCN2 and TOR. *Biochem. J* 449:1–10.

641 Ganten D, Nesse R. 2012. The evolution of evolutionary molecular medicine. *J. Mol. Med.* 90:467–
642 470.

643 Ghirlanda S, Enquist M, Lind J. 2014. Coevolution of intelligence, behavioral repertoire, and lifespan.
644 *Theor. Popul. Biol.* 91:44–49.

645 Gorbunova V, Seluanov A, Mao Z, Hine C. 2007. Changes in DNA repair during aging. *Nucleic Acids*
646 *Res.* 35:7466–7474.

647 Gorbunova V, Seluanov A, Zhang Z, Gladyshev VN, Vijg J. 2014. Comparative genetics of longevity
648 and cancer: insights from long-lived rodents. *Nat. Rev. Genet.* 15:531–540.

649 Grimmel M, Backhaus C, Proikas-Cezanne T. 2015. WIPI-Mediated Autophagy and Longevity. *Cells*
650 4:202–217.

651 GÜven G, Terzİ ÇİzmecİoĞlu N. 2021. Arid4b alters cell cycle and cell death dynamics during mouse
652 embryonic stem cell differentiation. *Turk. J. Biol.* 45:56–64.

653 Harms FL, Kloth K, Bley A, Denecke J, Santer R, Lessel D, Hempel M, Kutsche K. 2018. Activating
654 Mutations in PAK1, Encoding p21-Activated Kinase 1, Cause a Neurodevelopmental Disorder.
655 *Am. J. Hum. Genet.* 103:579–591.

656 Harris VK, Schiffman JD, Boddy AM. 2017. Chapter 7 - Evolution of Cancer Defense Mechanisms

657 Across Species. In: Ujvari B, Roche B, Thomas F, editors. Ecology and Evolution of Cancer.
658 Academic Press. p. 99–110.

659 Healy K, Guillerme T, Finlay S, Kane A, Kelly SBA, McClean D, Kelly DJ, Donohue I, Jackson AL,
660 Cooper N. 2014. Ecology and mode-of-life explain lifespan variation in birds and mammals.
661 *Proc. Biol. Sci.* 281:20140298.

662 Hicar MD, Liu Y, Allen CE, Wu L-C. 2001. Structure of the Human Zinc Finger Protein HIVEP3:
663 Molecular Cloning, Expression, Exon–Intron Structure, and Comparison with Paralogous Genes
664 HIVEP1 and HIVEP2. *Genomics* 71:89–100.

665 Hirsch S, El-Achkar T, Robbins L, Basta J, Heitmeier M, Nishinakamura R, Rauchman M. 2015. A
666 mouse model of Townes-Brocks syndrome expressing a truncated mutant Sall1 protein is
667 protected from acute kidney injury. *Am. J. Physiol. Renal Physiol.* 309:F852–F863.

668 Hochberg ME, Noble RJ. 2017. A framework for how environment contributes to cancer risk. *Ecol.*
669 *Lett.* 20:117–134.

670 Hossain MA, Larrous F, Rawlinson SM, Zhan J, Sethi A, Ibrahim Y, Aloi M, Lieu KG, Mok Y-F,
671 Griffin MDW, et al. 2019. Structural Elucidation of Viral Antagonism of Innate Immunity at the
672 STAT1 Interface. *Cell Rep.* 29:1934–1945.e8.

673 Huang X, Sun D, Wu T, Liu X, Xu S, Yang G. 2021. Genomic insights into body size evolution in
674 Carnivora support Peto's paradox. *BMC Genomics* [Internet] 22. Available from:
675 http://dx.doi.org/10.1186/s12864-021-07732-w

676 Hubisz MJ, Pollard KS, Siepel A. 2011. PHAST and RPHAST: phylogenetic analysis with space/time
677 models. *Brief. Bioinform.* 12:41–51.

678 Huerta-Cepas J, Serra F, Bork P. 2016. ETE 3: Reconstruction, Analysis, and Visualization of
679 Phylogenomic Data. *Molecular Biology and Evolution* [Internet] 33:1635–1638. Available from:
680 http://dx.doi.org/10.1093/molbev/msw046

681 Ilyenko I, Bazyka D, Lyaskivska O, Dyagil I, Kravchenko S. 2011. 74 Expression of TERT, TERF1
682 and TERF2 genes and telomere length in patients with myelodysplastic syndrome and acute
683 myeloid leukemia. *Leukemia Research* [Internet] 35:S28. Available from:
684 http://dx.doi.org/10.1016/s0145-2126(11)70076-5

685 Jacques P-É, Jeyakani J, Bourque G. 2013. The majority of primate-specific regulatory sequences are
686 derived from transposable elements. *PLoS Genet.* 9:e1003504.

687 Janoušek V, Laukaitis CM, Yanchukov A, Karn RC. 2016. The Role of Retrotransposons in Gene
688 Family Expansions in the Human and Mouse Genomes. *Genome Biol. Evol.* 8:2632–2650.

689 Kacprzyk J, Locatelli AG, Hughes GM, Huang Z, Clarke M, Gorbunova V, Sacchi C, Stewart GS,
690 Teeling EC. 2021. Evolution of mammalian longevity: age-related increase in autophagy in bats
691 compared to other mammals. *Aging* 13:7998–8025.

692 Kalueff AV. 2007. Neurobiology of memory and anxiety: from genes to behavior. *Neural Plast.*
693 2007:78171.

694 Karst SM, Wobus CE, Lay M, Davidson J, Virgin HW 4th. 2003. STAT1-dependent innate immunity
695 to a Norwalk-like virus. *Science* 299:1575–1578.

696 Katoh K, Standley DM. 2013. MAFFT Multiple Sequence Alignment Software Version 7:
697 Improvements in Performance and Usability. *Molecular Biology and Evolution* [Internet]
698 30:772–780. Available from: http://dx.doi.org/10.1093/molbev/mst010

699 Keane M, Semeiks J, Webb AE, Li YI, Quesada V, Craig T, Madsen LB, van Dam S, Brawand D,
700 Marques PI, et al. 2015. Insights into the evolution of longevity from the bowhead whale
701 genome. *Cell Rep.* 10:112–122.

702 Khatri P, Sirota M, Butte AJ. 2012. Ten years of pathway analysis: current approaches and
703 outstanding challenges. *PLoS Comput. Biol.* 8:e1002375.

704 Krovi SH, Zhang J, Michaels-Foster MJ, Brunetti T. 2020. Single-cell RNA transcriptomics identifies
705 Hivep3 as essential in regulating the development of innate-like T lymphocytes. *bioRxiv*
706 [Internet]. Available from:
707 https://www.biorxiv.org/content/10.1101/2020.06.08.135129v1.abstract

708 Lee JS, Ward WO, Ren H, Vallanat B, Darlington GJ, Han E-S, Laguna JC, DeFord JH,
709 Papaconstantinou J, Selman C, et al. 2012. Meta-analysis of gene expression in the mouse liver
710 reveals biomarkers associated with inflammation increased early during aging. *Mechanisms of*
711 *Ageing and Development* [Internet] 133:467–478. Available from:
712 http://dx.doi.org/10.1016/j.mad.2012.05.006

713 Liao Y, Wang J, Jaehnig EJ, Shi Z, Zhang B. 2019. WebGestalt 2019: gene set analysis toolkit with
714 revamped UIs and APIs. *Nucleic Acids Res.* 47:W199–W205.

715 Liu Y, Sun J, Zhao M. 2017. ONGene: A literature-based database for human oncogenes. *J. Genet.*
716 *Genomics* 44:119–121.

717 Li Y, de Magalhães JP. 2013. Accelerated protein evolution analysis reveals genes and pathways
718 associated with the evolution of mammalian longevity. *Age* 35:301–314.

719 López-Otín C, Blasco MA, Partridge L, Serrano M, Kroemer G. 2013. The hallmarks of aging. *Cell*
720 153:1194–1217.

721 Lowenstine LJ. 1986. Neoplasms and Proliferative Disorders in Nonhuman Primates. In: Primates.
722 Proceedings in Life Sciences. Springer, New York, NY. p. 781–814.

723 Lowenstine LJ, McManamon R, Terio KA. 2016. Comparative Pathology of Aging Great Apes:
724 Bonobos, Chimpanzees, Gorillas, and Orangutans. *Vet. Pathol.* 53:250–276.

725 MacRae SL, Zhang Q, Lemetre C, Seim I, Calder RB, Hoeijmakers J, Suh Y, Gladyshev VN,
726 Seluanov A, Gorbunova V, et al. 2015. Comparative analysis of genome maintenance genes in
727 naked mole rat, mouse, and human. *Aging Cell* 14:288–291.

728 de Magalhães JP, Curado J, Church GM. 2009. Meta-analysis of age-related gene expression profiles
729 identifies common signatures of aging. *Bioinformatics* 25:875–881.

730 Marand AP, Chen Z, Gallavotti A, Schmitz RJ. 2021. A cis-regulatory atlas in maize at single-cell
731 resolution. *Cell* 184:3041–3055.e21.

732 Meng J, Meng Y, Hanna A, Janus C, Jia Z. 2005. Abnormal long-lasting synaptic plasticity and
733 cognition in mice lacking the mental retardation gene Pak3. *J. Neurosci.* 25:6641–6650.

734 Miki Y, Shimoyama S, Kon T, Ueno T, Hayakari R, Tanji K, Matsumiya T, Tsushima E, Mori F,
735 Wakabayashi K, et al. 2018. Alteration of autophagy-related proteins in peripheral blood
736 mononuclear cells of patients with Parkinson's disease. *Neurobiol. Aging* 63:33–43.

737 Morris BJ, Willcox BJ, Donlon TA. 2019. Genetic and epigenetic regulation of human aging and
738 longevity. *Biochim. Biophys. Acta Mol. Basis Dis.* 1865:1718–1744.

739 Morris LGT, Kaufman AM, Gong Y, Ramaswami D, Walsh LA, Turcan Ş, Eng S, Kannan K, Zou Y,

740 Peng L, et al. 2013. Recurrent somatic mutation of FAT1 in multiple human cancers leads to
741 aberrant Wnt activation. *Nat. Genet.* 45:253–261.

742 Muntané G, Farré X, Rodríguez JA, Pegueroles C, Hughes DA, de Magalhães JP, Gabaldón T,
743 Navarro A. 2018. Biological Processes Modulating Longevity across Primates: A Phylogenetic
744 Genome-Phenome Analysis. *Mol. Biol. Evol.* 35:1990–2004.

745 Nesse RM, Ganten D, Gregory TR, Omenn GS. 2012. Evolutionary molecular medicine. *J. Mol. Med.*
746 90:509–522.

747 Nunney L. 1999. Lineage selection and the evolution of multistage carcinogenesis. *Proc. Biol. Sci.*
748 266:493–498.

749 Nunney L. 2018. Size matters: height, cell number and a person's risk of cancer. *Proc. Biol. Sci.*
750 [Internet] 285. Available from: http://dx.doi.org/10.1098/rspb.2018.1743

751 Nusse R, Varmus HE. 1992. Wnt genes. *Cell* 69:1073–1087.

752 Orkin JD, Montague MJ, Tejada-Martinez D, de Manuel M, Del Campo J, Cheves Hernandez S, Di
753 Fiore A, Fontsere C, Hodgson JA, Janiak MC, et al. 2021. The genomics of ecological flexibility,
754 large brains, and long lives in capuchin monkeys revealed with fecalFACS. *Proc. Natl. Acad. Sci.*
755 *U. S. A.* [Internet] 118. Available from: http://dx.doi.org/10.1073/pnas.2010632118

756 Orme D, Freckleton R, Thomas G, Petzoldt T, Fritz S, Isaac N, Pearse W, Others. 2012. Caper:
757 comparative analyses of phylogenetics and evolution in R. *R package version 0. 5* 2:458.

758 Pan X, Chang X, Leung C, Zhou Z, Cao F, Xie W, Jia Z. 2015. PAK1 regulates cortical development
759 via promoting neuronal migration and progenitor cell proliferation. *Mol. Brain* 8:36.

760 Pennell MW, Eastman JM, Slater GJ, Brown JW, Uyeda JC, FitzJohn RG, Alfaro ME, Harmon LJ.
761 2014. geiger v2.0: an expanded suite of methods for fitting macroevolutionary models to
762 phylogenetic trees. *Bioinformatics* 30:2216–2218.

763 Peters CW, Sippel AE, Vingron M, Klempnauer KH. 1987. Drosophila and vertebrate myb proteins
764 share two conserved regions, one of which functions as a DNA-binding domain. *EMBO J.*
765 6:3085–3090.

766 Peto R, Epidemiology MM, Tests STM. 1977. in The Origins of Human Cancer, ed. by HH Hiatt, JD
767 Watson and JA Winsten.

768 Petralia RS, Mattson MP, Yao PJ. 2014. Aging and longevity in the simplest animals and the quest for
769 immortality. *Ageing Res. Rev.* 16:66–82.

770 Pontis J, Planet E, Offner S, Turelli P, Duc J, Coudray A, Theunissen TW, Jaenisch R, Trono D. 2019.
771 Hominoid-Specific Transposable Elements and KZFPs Facilitate Human Embryonic Genome
772 Activation and Control Transcription in Naive Human ESCs. *Cell Stem Cell* 24:724–735.e5.

773 Quesada V, Freitas-Rodríguez S, Miller J, Pérez-Silva JG, Jiang Z-F, Tapia W, Santiago-Fernández O,
774 Campos-Iglesias D, Kuderna LFK, Quinzin M, et al. 2019. Giant tortoise genomes provide
775 insights into longevity and age-related disease. *Nat Ecol Evol* 3:87–95.

776 Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features.
777 *Bioinformatics* 26:841–842.

778 Rao SG, Jackson JG. 2016. SASP: Tumor Suppressor or Promoter? Yes! *Trends Cancer Res.*
779 [Internet] 2. Available from: https://pubmed.ncbi.nlm.nih.gov/28741506/

780 Revell LJ. 2012. phytools: an R package for phylogenetic comparative biology (and other things).
781 *Methods Ecol. Evol.* 3:217–223.

782 Rodríguez JA, Marigorta UM, Hughes DA, Spataro N, Bosch E, Navarro A. 2017. Antagonistic
783 pleiotropy and mutation accumulation influence human senescence and disease. *Nat Ecol Evol*
784 1:55.

785 Romiguier J, Ranwez V, Douzery EJP, Galtier N. 2010. Contrasting GC-content dynamics across 33
786 mammalian genomes: Relationship with life-history traits and chromosome sizes. *Genome*
787 *Research* [Internet] 20:1001–1009. Available from: http://dx.doi.org/10.1101/gr.104372.109

788 Roth ACJ, Gonnet GH, Dessimoz C. 2008. Algorithm of OMA for large-scale orthology inference.
789 *BMC Bioinformatics* [Internet] 9. Available from: http://dx.doi.org/10.1186/1471-2105-9-518

790 Seluanov A, Gladyshev VN, Vijg J, Gorbunova V. 2018. Mechanisms of cancer resistance in long-
791 lived mammals. *Nat. Rev. Cancer* 18:433–441.

792 Seluanov A, Hine C, Azpurua J, Feigenson M, Bozzella M, Mao Z, Catania KC, Gorbunova V. 2009.
793 Hypersensitivity to contact inhibition provides a clue to cancer resistance of naked mole-rat.
794 *Proc. Natl. Acad. Sci. U. S. A.* 106:19352–19357.

795 Seluanov A, Hine C, Bozzella M, Hall A, Sasahara THC, Ribeiro AACM, Catania KC, Presgraves
796 DC, Gorbunova V. 2008. Distinct tumor suppressor mechanisms evolve in rodent species that
797 differ in size and lifespan. *Aging Cell* 7:813–823.

798 Singh PP, Demmitt BA, Nath RD, Brunet A. 2019. The Genetics of Aging: A Vertebrate Perspective.
799 *Cell* 177:200–220.

800 Slade JD, Staveley BE. 2016. Extended longevity and survivorship during amino-acid starvation in a
801 Drosophila Sir2 mutant heterozygote. *Genome* 59:311–318.

802 Sopova JV, Koshel EI, Belashova TA, Zadorsky SP, Sergeeva AV, Siniukova VA, Shenfeld AA,
803 Velizhanina ME, Volkov KV, Nizhnikov AA, et al. 2019. RNA-binding protein FXR1 is
804 presented in rat brain in amyloid form. *Sci. Rep.* 9:18983.

805 Stratton MR, Campbell PJ, Futreal PA. 2009. The cancer genome. *Nature* 458:719–724.

806 Sulak M, Fong L, Mika K, Chigurupati S, Yon L, Mongan NP, Emes RD, Lynch VJ. 2016.
807 Correction: TP53 copy number expansion is associated with the evolution of increased body size
808 and an enhanced DNA damage response in elephants. *Elife* [Internet] 5. Available from:
809 http://dx.doi.org/10.7554/eLife.24307

810 Sundaram V, Cheng Y, Ma Z, Li D, Xing X, Edge P, Snyder MP, Wang T. 2014. Widespread
811 contribution of transposable elements to the innovation of gene regulatory networks. *Genome*
812 *Res.* 24:1963–1976.

813 Sundaram V, Wysocka J. 2020. Transposable elements as a potent source of diverse cis-regulatory
814 sequences in mammalian genomes. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 375:20190347.

815 Tacutu R, Thornton D, Johnson E, Budovsky A, Barardo D, Craig T, Diana E, Lehmann G, Toren D,
816 Wang J, et al. 2018. Human Ageing Genomic Resources: new and updated databases. *Nucleic*
817 *Acids Res.* 46:D1083–D1090.

818 Tejada-Martinez D, de Magalhães JP, Opazo JC. 2021. Positive selection and gene duplications in
819 tumour suppressor genes reveal clues about how cetaceans resist cancer. *Proc. Biol. Sci.*
820 288:20202592.

821 Thomas F, Fisher D, Fort P, Marie J-P, Daoust S, Roche B, Grunau C, Cosseau C, Mitta G,
822 Baghdiguian S, et al. 2013. Applying ecological and evolutionary theory to cancer: a long and
823 winding road. *Evol. Appl.* 6:1–10.

824 Tian X, Azpurua J, Hine C, Vaidya A, Myakishev-Rempel M, Ablaeva J, Mao Z, Nevo E, Gorbunova
825 V, Seluanov A. 2013. High-molecular-mass hyaluronan mediates the cancer resistance of the
826 naked mole rat. *Nature* 499:346–349.

827 Tian X, Doerig K, Park R, Can Ran Qin A, Hwang C, Neary A, Gilbert M, Seluanov A, Gorbunova V.
828 2018. Evolution of telomere maintenance and tumour suppressor mechanisms across mammals.
829 *Philos. Trans. R. Soc. Lond. B Biol. Sci.* [Internet] 373. Available from:
830 http://dx.doi.org/10.1098/rstb.2016.0443

831 Tian Y, Wang S, Jiao F, Kong Q, Liu C, Wu Y. 2019. Telomere Length: A Potential Biomarker for the
832 Risk and Prognosis of Stroke. *Front. Neurol.* 10:624.

833 Tollis M, Robbins J, Webb AE, Kuderna LFK, Caulin AF, Garcia JD, Bèrubè M, Pourmand N,
834 Marques-Bonet T, O'Connell MJ, et al. 2019. Return to the Sea, Get Huge, Beat Cancer: An
835 Analysis of Cetacean Genomes Including an Assembly for the Humpback Whale (Megaptera
836 novaeangliae). *Mol. Biol. Evol.* 36:1746–1763.

837 Tollis M, Schiffman JD, Boddy AM. 2017. Evolution of cancer suppression as revealed by
838 mammalian comparative genomics. *Curr. Opin. Genet. Dev.* 42:40–47.

839 Tollis M, Schneider-Utaka AK, Maley CC. 2020. The Evolution of Human Cancer Gene Duplications
840 across Mammals. *Mol. Biol. Evol.* 37:2875–2886.

841 Tomasetti C, Vogelstein B. 2015. Cancer etiology. Variation in cancer risk among tissues can be
842 explained by the number of stem cell divisions. *Science* 347:78–81.

843 Trizzino M, Kapusta A, Brown CD. 2018. Transposable elements generate regulatory novelty in a
844 tissue-specific fashion. *BMC Genomics* 19:468.

845 Trizzino M, Park Y, Holsbach-Beltrame M, Aracena K, Mika K, Caliskan M, Perry GH, Lynch VJ,
846 Brown CD. 2017. Transposable elements are the primary source of novelty in primate gene
847 regulation. *Genome Res.* 27:1623–1633.

848 Uyeda JC, Pennell MW, Miller ET, Maia R, McClain CR. 2017. The Evolution of Energetic Scaling
849 across the Vertebrate Tree of Life. *Am. Nat.* 190:185–199.

850 Vazquez JM, Sulak M, Chigurupati S, Lynch VJ. 2018. A Zombie LIF Gene in Elephants Is
851 Upregulated by TP53 to Induce Apoptosis in Response to DNA Damage. *Cell Rep.* 24:1765–
852 1776.

853 Wang T, Zeng J, Lowe CB, Sellers RG, Salama SR, Yang M, Burgess SM, Brachmann RK, Haussler
854 D. 2007. Species-specific endogenous retroviruses shape the transcriptional network of the
855 human tumor suppressor protein p53. *Proc. Natl. Acad. Sci. U. S. A.* 104:18613–18618.

856 Weber CC, Nabholz B, Romiguier J, Ellegren H. 2014. Kr/Kc but not dN/dS correlates positively with
857 body mass in birds, raising implications for inferring lineage-specific selection. *Genome Biol.*
858 15:542.

859 White RR, Milholland B, MacRae SL, Lin M, Zheng D, Vijg J. 2015. Comprehensive transcriptional
860 landscape of aging mouse liver. *BMC Genomics* 16:899.

861 Whittemore K, Vera E, Martínez-Nevado E, Sanpera C, Blasco MA. 2019. Telomere shortening rate
862 predicts species life span. *Proc. Natl. Acad. Sci. U. S. A.* 116:15122–15127.

863 Wilkinson GS, Adams DM. 2019. Recurrent evolution of extreme longevity in bats. *Biol. Lett.*
864 15:20180860.

865 Wu R-C, Young I-C, Chen Y-F, Chuang S-T, Toubaji A, Wu M-Y. 2019. Identification of the PTEN-
866 ARID4B-PI3K pathway reveals the dependency on ARID4B by PTEN-deficient prostate cancer.
867 *Nat. Commun.* 10:4332.

868 Xu M, Pirtskhalava T, Farr JN, Weigand BM, Palmer AK, Weivoda MM, Inman CL, Ogrodnik MB,
869 Hachfeld CM, Fraser DG, et al. 2018. Senolytics Improve Physical Function and Increase
870 Lifespan in Old Age. *Nat. Med.* 24:1246.

871 Yabu T, Imamura S, Mizusawa N, Touhata K, Yamashita M. 2012. Induction of autophagy by amino
872 acid starvation in fish cells. *Mar. Biotechnol.* 14:491–501.

873 Yang Z. 2007. PAML 4: Phylogenetic Analysis by Maximum Likelihood. *Molecular Biology and*
874 *Evolution* [Internet] 24:1586–1591. Available from: http://dx.doi.org/10.1093/molbev/msm088

875 Yu Z, Seim I, Yin M, Tian R, Sun D, Ren W, Yang G, Xu S. 2021. Comparative analyses of aging-
876 related genes in long-lived mammals provide insights into natural longevity. *Innov. J.* 2:100108.

877 Zhang G, Cowled C, Shi Z, Huang Z, Bishop-Lilly KA, Fang X, Wynne JW, Xiong Z, Baker ML,
878 Zhao W, et al. 2013. Comparative analysis of bat genomes provides insight into the evolution of
879 flight and immunity. *Science* 339:456–460.

880 Zhang J. 2005. Evaluation of an Improved Branch-Site Likelihood Method for Detecting Positive
881 Selection at the Molecular Level. *Molecular Biology and Evolution* [Internet] 22:2472–2479.
882 Available from: http://dx.doi.org/10.1093/molbev/msi237

883 Zhang Z, Li J, He T, Ding J. 2020. Bioinformatics Identified 17 Immune Genes as Prognostic
884 Biomarkers for Breast Cancer: Application Study Based on Artificial Intelligence Algorithms.
885 *Front. Oncol.* 10:330.

886 Zhao C, Takita J, Tanaka Y, Setou M, Nakagawa T, Takeda S, Yang HW, Terada S, Nakata T, Takei
887 Y, et al. 2001. Charcot-Marie-Tooth Disease Type 2A Caused by Mutation in a Microtubule
888 Motor KIF1Bβ. *Cell* 105:587–597.

889 Zhao M, Sun J, Zhao Z. 2013. TSGene: a web resource for tumor suppressor genes. *Nucleic Acids Res.*
890 41:D970–D976.

891 Zhao X, Yang C, Wu J, Nan Y. 2018. ADAMTS8 targets ERK to suppress cell proliferation, invasion,
892 and metastasis of hepatocellular carcinoma. *Onco. Targets. Ther.* 11:7569–7578.

893

Figures

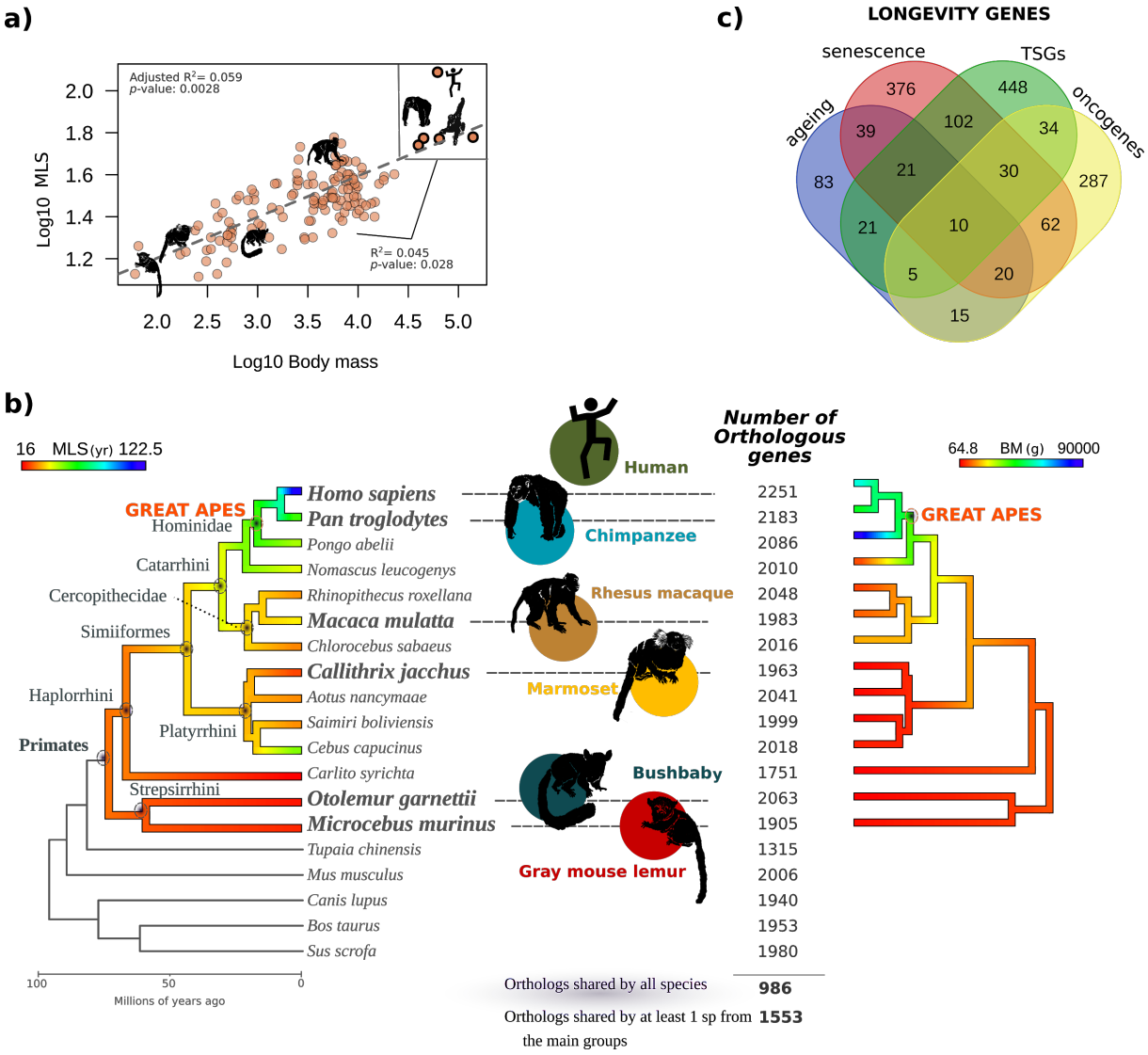

Figure 1. The evolution of body mass (BM) and maximum lifespan (MLS) across primates. a) Phylogenetic generalized least-squares models (PGLS) correlating MLS ~ BM across primates. The dashed lines represent a positive correlation between log10-transformed BM and MLS. The continuous line displays the correlation between great apes and other primates. b) Phylogenomic design. The molecular evolution analysis included 19 mammalian species. The six species highlighted are representatives from the main primates lineages (Catarrhini, Platyrrhini and Strepsirrhini). For these six species, RNA-seq and ChIP-seq data were publicly available (Trizzino et al. 2017). The colours in the phylogenetic tree reflect the values of MLS and BM in each primate branch respectively, from lowest (red) to highest (blue). c) Venn diagram representing the total number of orthologous longevity genes (1,553) shared by at least one species in each of main primates lineages, broken down based on the different categories (tumor suppressors, oncogenes, cell senescence and ageing genes).

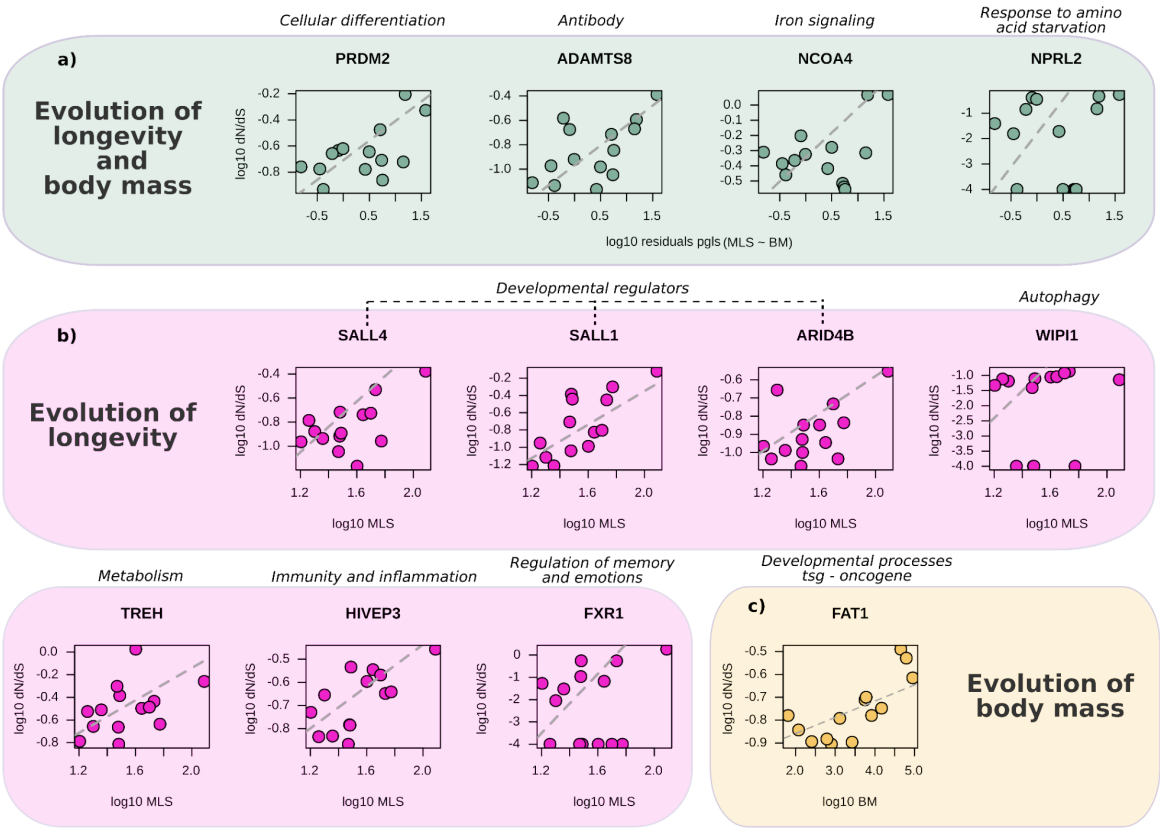

Figure 2. Examples of genes whose evolution correlated with the evolution of maximum lifespan (MLS) and/or body mass (BM). Three independent phylogenetic generalized least-squares models (PGLS) were performed in order to evaluate the independent contribution of each gene in the evolution of MLS and/or BM across primates. The dN/dS values were obtained from the branch model in PAML. a) PGLS model between dn/ds ~ residuals from the PGLS (MLS ~ BM). b) PGLS model between dn/ds ~ MLS and c) PGLS model between dn/ds ~ BM. In all the regression models the *p*-value was < 0.05 and the contribution of each gene (R^2) > 22% (Table 1).

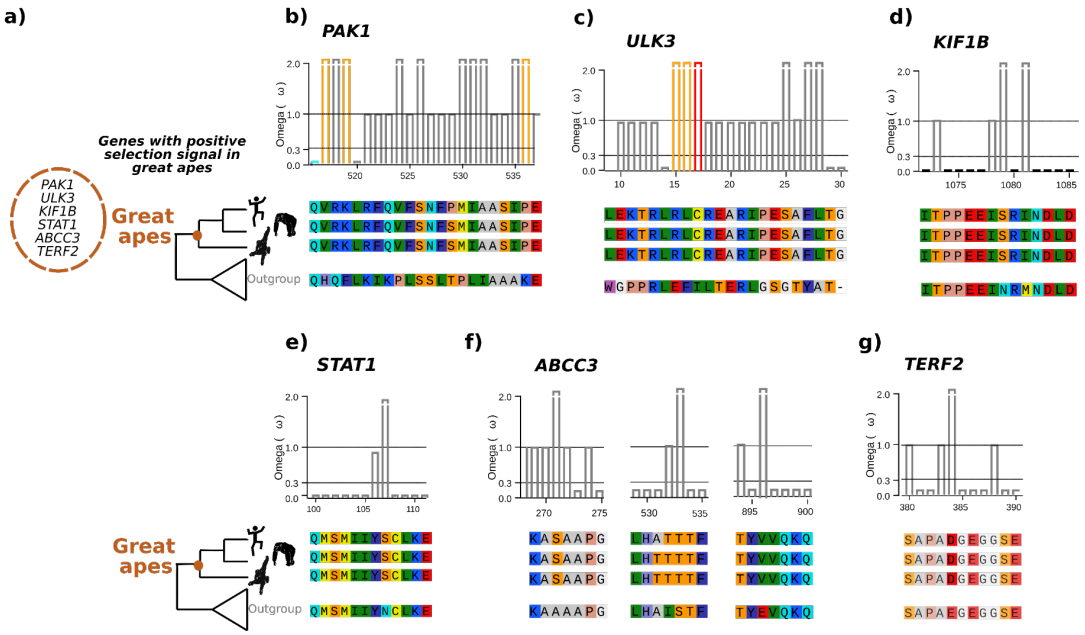

Figure 3. Genes under positive selection in the ancestor of great apes (a-g): *PAK1, ULK3, KIF1B, STAT1, ABCC3, TERF2.* The three first codon alignments on each gene represent fragments with an omega ω value >1. The colour gradients are used to display the degree and typology of amino acid change: the grey colour corresponds to the most conservative change, while red indicates the most radical. The alignments underneath display the consensus sequences across the outgroups (non-ape primates and other mammals). The amino acid colours reflect the different physicochemical properties.

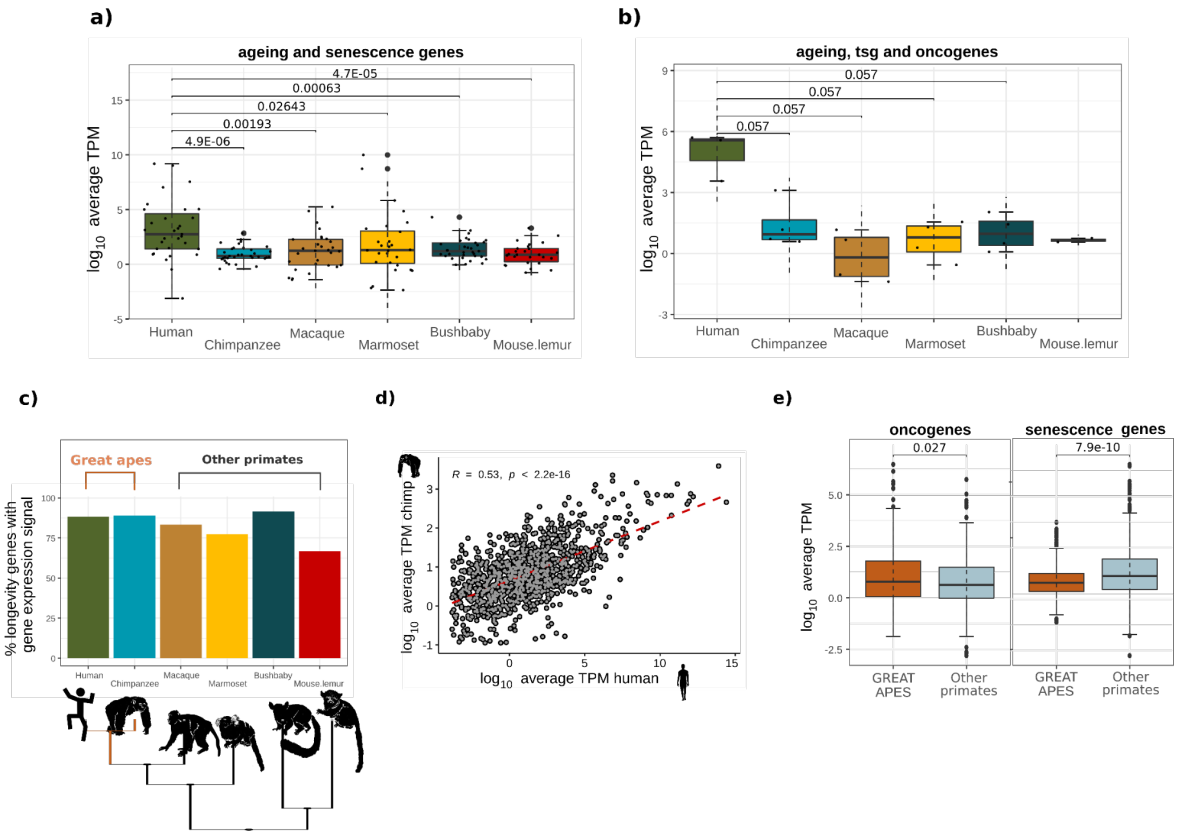

Figure 4. Evolution of longevity gene expression in the primate liver. Expression levels across species for the 1,553 longevity genes are divided into categories: ageing, senescence related genes, tumor suppressor genes (TSGs) and oncogenes; (a) ageing related genes and senescence genes (b) ageing related genes, tumor suppressors and oncogenes. In both groups, humans display higher expression compared to other primates (Wilcoxon rank-sum test, $p <$ or near 0.05). c) The percentage of longevity genes expressed in the liver is similar across the six primate species. d) The expression of the 1,553 longevity genes in the liver is positively correlated between humans and chimpanzee (Spearman's correlation coefficient, R=0.53, $p < 0.001$). e) Oncogenes are significantly more expressed in great apes compared to other primates. Senescence genes are significantly less expressed in great apes relative to the other primates (Wilcoxon rank-sum test, $p < 0.05$).

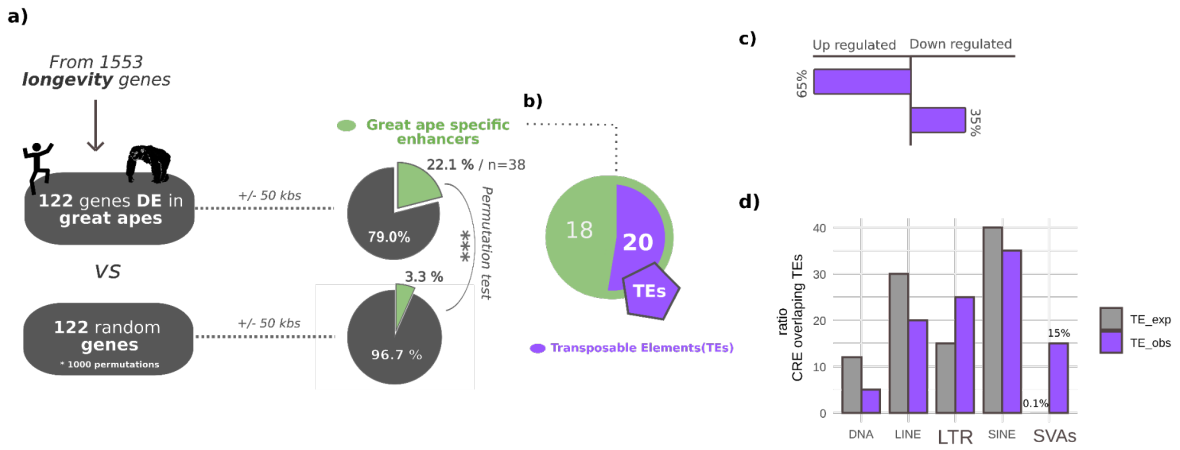

Figure 5. Ape-specific enhancers are enriched near differentially expressed longevity genes. a) 27/122 differentially expressed (DE) longevity genes (22.1%) are located within 50 kbs of an ape-specific enhancer (a total of 38 ape-specific enhancers for 27 DE genes). In comparison, only 4/122 (3.3%) randomly selected genes (1000 permutations) are located within 50 kbs of an ape-specific enhancer (Permutation Test $p < 2.2$ x 10-16).). b) 20/38 ape-specific enhancers (52%) located within 50 kbs of the DE longevity genes overlap annotated Transposable Elements (TEs). c) Of the DE longevity genes found near TE-derived ape-specific enhancers, 65% were identified as upregulated, 35% as downregulated (Fisher's Exact Test $p = 0.0425$). e) LTR and SVA transposons are overrepresented in the sequence of the ape-specific enhancers located within 50 kbs of a DE longevity gene (Fisher's Exact Test $p < 2.2 \times 10^{-16}$ for SVAs; $p < 0.0001$ for LTRs).