bioRxiv preprint doi: https://doi.org/10.1101/2020.05.06.080200; this version posted June 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

The changes in the p53 protein across the animal kingdom pointing to its involvement in longevity

Václav Brázda[1*#], Martin Bartas[2#], Adriana Volná[3], Jiří Červeň[2], Petr Pečinka[2] and Joanna Zawacka-Pankau[4*]

[1]Institute of Biophysics, Czech Academy of Sciences, Brno, 612 65, Czech Republic.

[2]Department of Biology and Ecology, University of Ostrava, Ostrava, 710 00, Czech Republic

[3]Department of Physics, Faculty of Science, University of Ostrava, Chittussiho 10, Ostrava, CZ, 71000, Czech Republic

[4]Faculty of Chemistry, Pasteura 1, 02-093 Warsaw, University of Warsaw, Warsaw, Poland; Department of Oncology-Pathology, Karolinska Institute, Stockholm, SE 171-74 Sweden

[*]Correspondence and requests for materials should be addressed to Joanna Zawacka-Pankau (email: j.zawackapan@uw.edu.pl) and Václav Brázda (email: vaclav@ibp.cz)

[#]These authors contributed equally to this work.

Keywords: p53, aging, longevity

Abstract:

Recently, the quest for the mythical fountain of youth has turned into specific research programs aiming to extend the healthy lifespan of humans. Despite advances in our understanding of the molecular processes underlying aging, the surprisingly extended lifespan of some animals remains unexplained. In this respect, the p53 protein plays a crucial role not only in tumor suppression but also in tissue homeostasis and healthy aging. However, the mechanism through which p53 maintains the function as a gatekeeper of healthy aging is not fully understood. Thus, we inspected *TP*53 gene sequences in individual species of phylogenetically related organisms that show different aging patterns. We discovered novel correlations between specific amino acid variations in p53 and lifespan across different animal species. In particular, we found that species with extended lifespan have characteristic amino acid substitutions mainly in the p53 DNA binding domain that change its function. Our findings show a direct association of specific amino acid residues in p53 protein with extended organismal lifespan and the importance of p53 protein in aging.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.06.080200; this version posted June 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Introduction

Stories about the fountain of youth promising eternal health have inspired research into this topic across many civilizations through the millennia since Herodotus mentioned it first in writing 2,500 years ago. Although the average human lifespan is increasing, our health span appears to be lagging. Several researchers argue that human lifespan is physiologically and genetically limited (Whittemore et al., 2019), while recently published studies have proposed a potentially unlimited increase of lifespan in the future (Hughes & Hekimi, 2017) and demographical data show that human death rates increase exponentially up to about age 80, then decelerate and plateau after age 105 (Barbi et al., 2018). There are several theories to explain aging, for example, the free radical theory, stating that organisms age because of the accumulation of damage inflicted by reactive oxygen species (Harraan, 1955; Sohal & Weindruch, 1996). There is also a common agreement that fidelity of DNA repair favors longevity (Storci et al., 2019).

From the published studies, it becomes clear that the longevity of organisms undergoes complex regulatory challenges and that lifespan is a multi-nodal characteristic (Campisi, 2005). To date, several proteins have been described to play critical roles in aging. One crucial protein involved in human longevity is the p53 tumor suppressor, which is coded by the most often mutated gene in human cancers (Bennett et al., 1999; Kandoth et al., 2013; Levine & Oren, 2009; Petitjean et al., 2007). Cancer, together with heart diseases, is the main cause of human deaths worldwide, therefore it can be concluded that overall the *TP53* gene is important for human lifespan (Bray et al., 2018; Timmis et al., 2018). On the other hand, when we consider the "lifespan" of tumors and tumor cell lines, it is apparent that cancer cells often gain new functions, including "immortality", at least partially attributed to mutation in the *TP53* gene and/or in its pathway (Schmidt-Kastner et al., 1998). As reviewed by Stiewe and Haran (2018), cancer-associated mutations alter p53 in three ways: loss of wild-type (wt) p53-DNA binding, dominant-negative inhibition by the mutant p53 allele, and gain of new DNA interactions specific to the mutant protein. Loss of binding to canonical p53 target genes can be complete, but mutants can also show a variable degree of loss of binding resulting in attenuated or target-selective binding patterns (Stiewe and Haran, 2018). Multiple functions of p53 have been described and extensively reviewed (Levine, 1997; Rufini et al., 2013; Sabapathy & Lane, 2018). For example, the p53 protein plays roles in metabolism (Vousden & Lane, 2007), cell cycle arrest (Chen, 2016; Hafner et al., 2019), apoptosis (Aubrey et al., 2018), angiogenesis (Pfaff et al., 2018), DNA repair (Nicolai et al., 2015) and cell senescence (Itahana et al., 2001; Rufini et al., 2013). As a DNA-binding protein, p53 functions as a transcription factor and recognizes and binds to multiple target genes (Brázda & Fojta, 2019; El-Deiry et al., 1992; Vyas et al., 2017). Owing to its crucial role in protection against DNA damage, p53 is called "the guardian of the genome" (Lane, 1992; Toufektchan & Toledo, 2018).

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.06.080200; this version posted June 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

From the evolutionary point of view, the *TP53* gene is specific for the Holozoa branch and its ancestral p63/p73-like gene emerged approximately one billion years ago (Bartas et al., 2020; Belyi & Levine, 2009). The p53/p63/p73 family plays key roles in several major molecular and biological processes including tumor suppression, fertility, mammalian embryonic development and aging (Vousden & Lane, 2007). Some mutations in the *TP*53 gene drive tumor development and tumor progression and it has been demonstrated that all p53 family members take part in regulating aging (Nicolai et al., 2015; Rufini et al., 2013). The study by Tyner et al., using heterozygous mice having one p53 allele with deletion of the first six exons ($p53^{+/m}$, Δ exon 1-6) showed for the first time the role of constitutively expressed (hyper-stable) p53 in aging. The mutant mice exhibited enhanced resistance to spontaneous tumors, but also displayed accelerated aging compared to $p53^{+/+}$ mice (Tyner et al., 2002). On the other hand, the resulting idea that the constitutive expression of p53 accelerates aging was not confirmed in a follow-up study (García-Cao et al., 2002), in which the pro-aging phenotype was not seen in p53 "super-mice", indicating that over-activated p53 *per se* is not the driver of accelerated aging nor of a replicative senescence, a process known to contribute to aging. It has been demonstrated that replicative senescence is facilitated by p53 mainly through activation of CDKN1A/p21. However, there are several other factors related to this process, including activation of E2F and mTOR as described elsewhere (Rufini et al., 2013). It can be concluded that p53 prevents cancer and protects from aging under physiological conditions, however, stress-activated p53 has a detrimental effect on healthy aging despite retaining its tumor suppression function. Thus, p53 can be either a pro-aging or a pro-longevity factor, depending on the physiological context (Keizer et al., 2010).

In addition to full length p53, p53 isoforms may also play an important role in longevity modulation as it has been demonstrated that expression of certain short and long p53 isoforms might maintain a balance between tumor suppression and tissue regeneration (Maier et al., 2004).

Considering the limited and conflicting information on the role of p53 in aging, in the present work we have employed currently available tools and analyzed p53 protein variations across the animal kingdom in correlation with the organisms' lifespan. We found that, when compared to the majority of closely related organisms within their phylogenetic groups, animals with unusually long lifespans share atypical p53 sequence features, pointing to the important contribution of p53 in regulation of life expectancy.

Results

We inspected all currently available sequence data of long-lived animals to explore a link between longevity (maximal lifespan) and p53 protein sequences. For this, we used the longevity data from the AnAge Database (De Magalhaes & Costa, 2009). We downloaded all available p53 sequences from

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.06.080200; this version posted June 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

RefSeq database and merged them with AnAge Database (for more detail refer to Materials and methods). The p53 sequence from 118 species and their lifespan data were catalogued and sorted according to their phylogenetic group (Supplementary Material 1).

The longest living animal in our dataset is the bowhead whale (*Balaena mysticetus*) from **Artiodactyla** (subgroup Cetacea) with a maximal lifespan of 211 ± 35 years (Keane et al., 2015). Bowhead whales have significantly longer lifespan (about four times longer) compared with other whales. The comparison of p53 protein sequences showed that, in contrast to other Cetacea, *Balaena mysticetus* has a unique leucine substitution in the proline rich region, corresponding to amino acid residue 77 in human p53 (Figure 1). All other accessible p53 sequences of whales have the identical amino acid residue in this position as human p53.

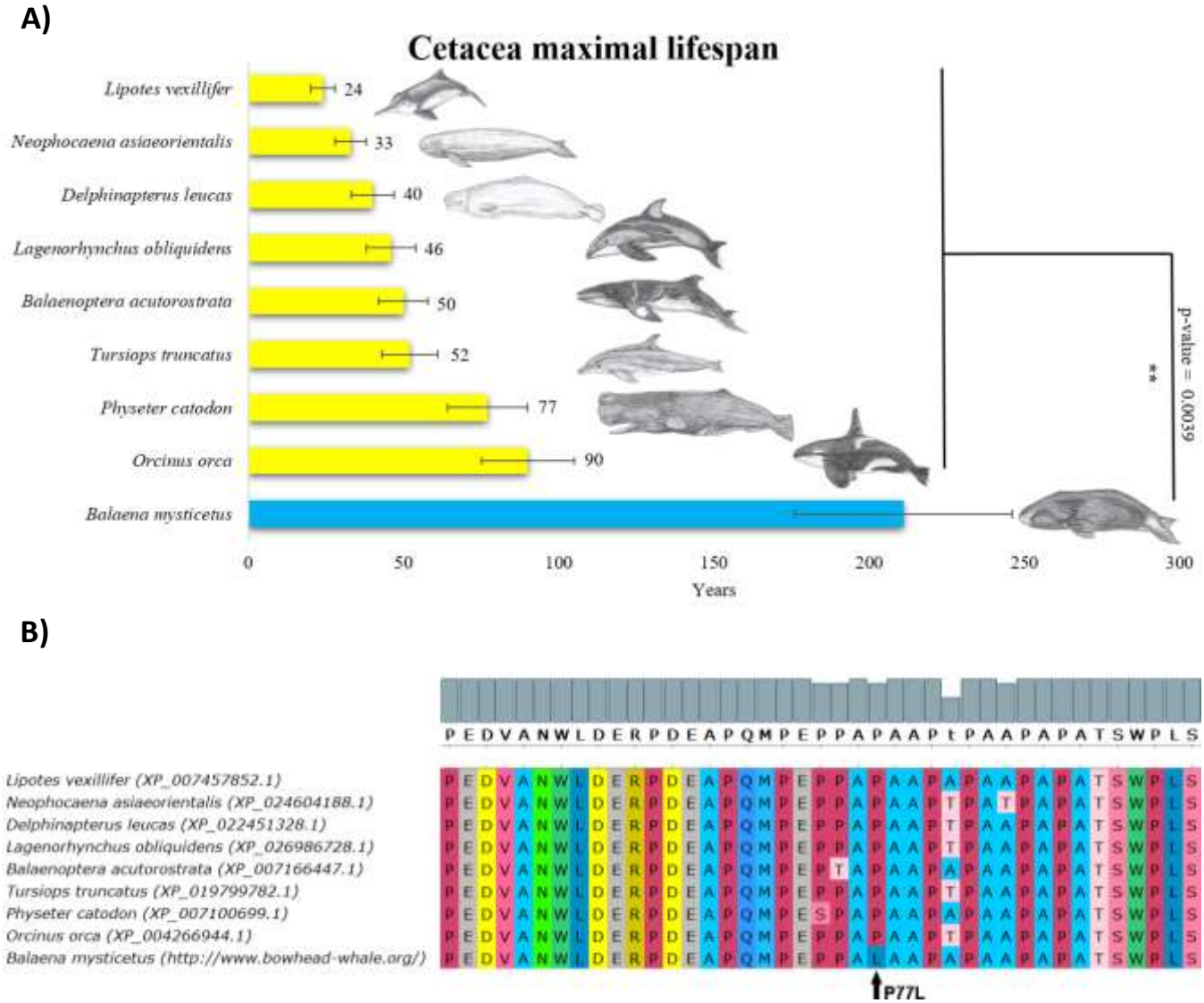


Figure 1. (A) Comparison of maximal cetacea lifespan in years. Bowhead whale (*Baleana mysticetus*) maximal lifespan is more than twice the maximal lifespan of the rest of cetacea (Wilcoxon one-sided signed rank test was used, p-value < 0.05). (B) Multiple sequence protein alignments of p53 proline-rich region, performed in MUSCLE with default parameters (Edgar, 2004), colors in "UGENE" style.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.06.080200; this version posted June 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Most amphibian species live for less than 30 years (De Magalhaes & Costa, 2009), however, olm (*Proteus anguinus*, Batrachia, Amphibians), the only exclusively cave-dwelling chordate, has a maximal documented lifespan of 102 years. Comparison of the p53 protein sequences in amphibians showed a previously unrecognized insertion in *Proteus anguinus*. The p53 protein from this specie has additional serine and arginine residues in its core domain (corresponding to insertion after amino acid 188L in human p53) (Figure 2).

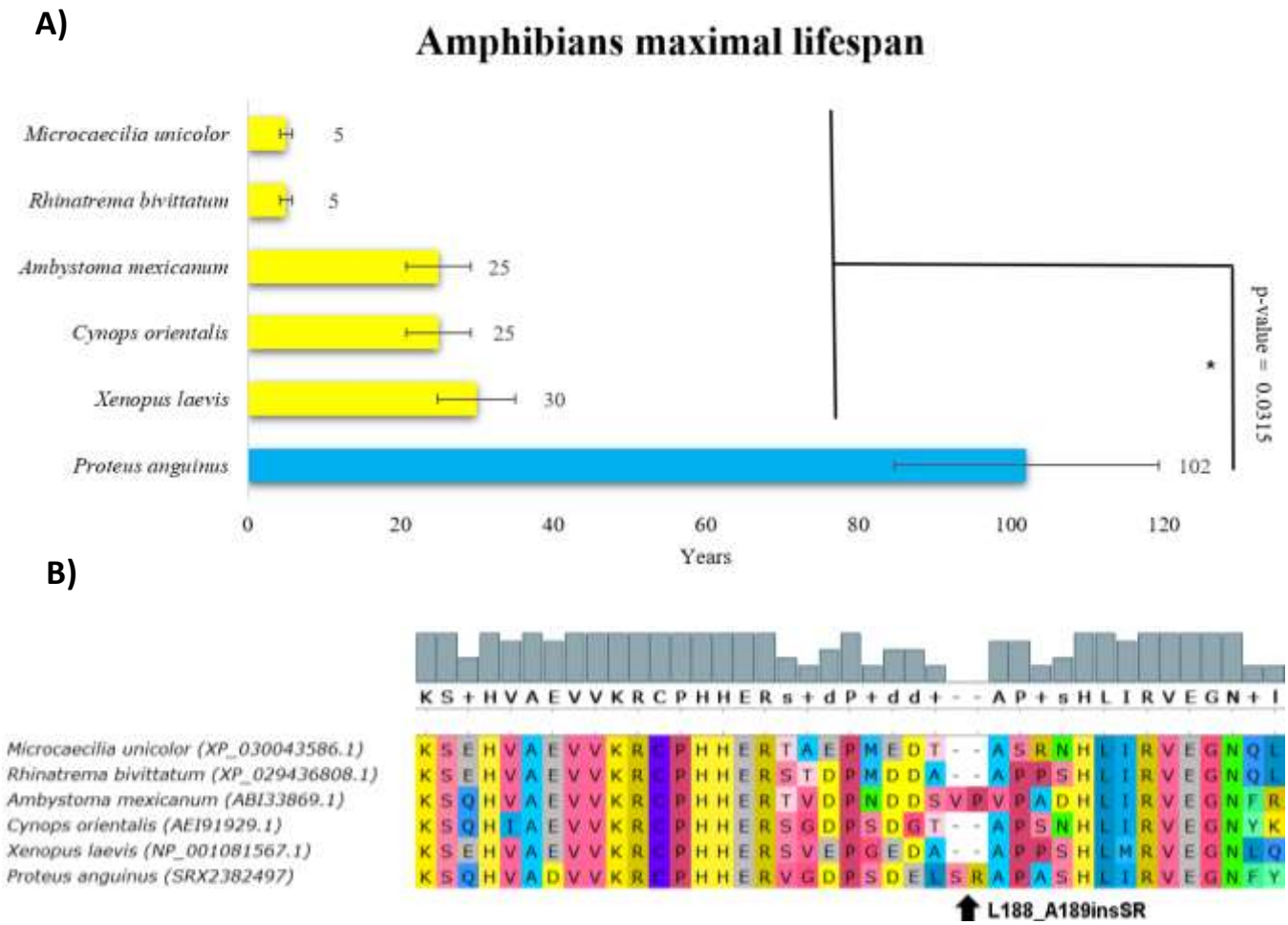


Figure 2: (A) Comparison of amphibian lifespans in years. Olm´s (*Proteus anguinus*) maximal lifespan is more than three times higher than the maximal lifespan of other amphibians (Wilcoxon one-sided signed rank test, p-value < 0.05). (B) Multiple protein alignments of the p53 dimerization region. Olm (*Proteus anguinus*) has a two-amino acid-residue long insertion following amino acid residue 188 (related to human p53 canonical sequence). The sequence of the p53 homolog from *Proteus anguinus* was determined using transcriptomic data from the SRA Archive (SRX2382497). Methods and color schemes are the same as in Figure 1B.

Kakapo (*Strigops habroptila*) is a long-lived, large, flightless, nocturnal, ground-dwelling parrot endemic to New Zealand with a lifespan of around 95 years (Figure 3A, blue bar). Comparison of p53

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.06.080200; this version posted June 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

protein sequence with other related species shows a change at positions 128 and 131, corresponding to human P128V and N131H (Figure 3B). Interestingly, N131H mutations in human p53 are found in pancreatic and colon cancers (Deuter & Müller, 1998; Pellegata et al., 1994). This mutation most probably changes the structure of the p53 core domain and decreases the ability of p53 to bind DNA. According to PHANTM classifier, the N131H mutation decreases p53 transcriptional activity by 47.19% (Giacomelli et al., 2018). In addition, according to the PROVEAN tool, substitutions at position 128 are deleterious with the score -4.45.

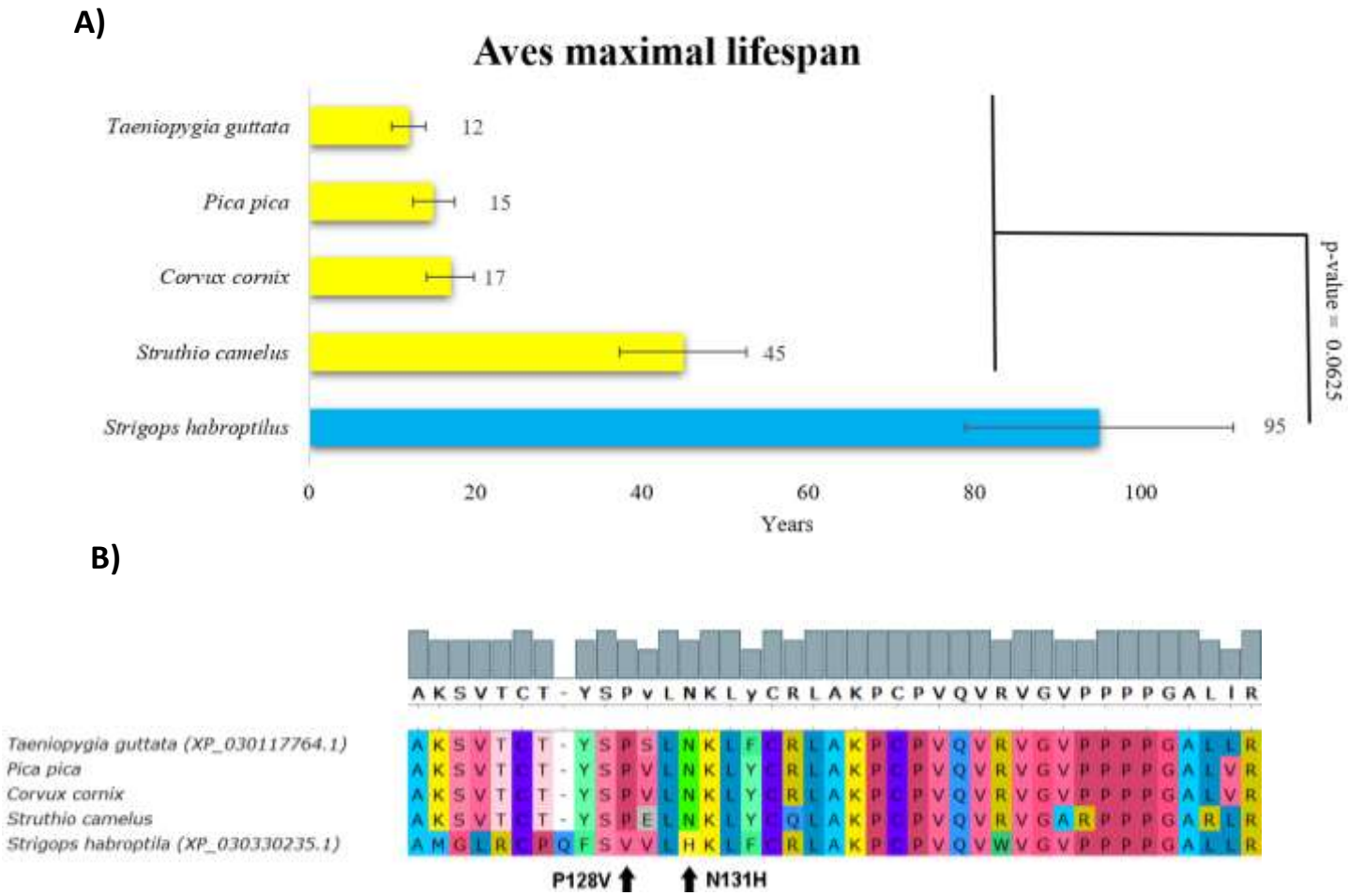


Figure 3: (A) Comparison of maximal Aves lifespan in years. Kakapo´s (*Strigops habroptila*) maximal lifespan is more than twice the maximal lifespan of other Aves (Wilcoxon one-sided signed rank test, p-value < 0.05). (B) Multiple protein alignments representing partial p53 core domain of the accessible Aves sequences. Sequences of all avian p53 homologs were determined using transcriptomic data from the SRA Archive, except from *Strigops habroptilus*, where the p53 sequence was known (XP_030330235.1. Methods and color schemes are the same as in Figure 1B.

Next, our analysis identified alterations associated with long lifespan in the **Chiroptera** group. The Brandt's bat (*Myotis Brandtii*) is an extremely long-lived bat with a documented lifespan of 41 years (Wilkinson & South, 2002). Together with its close relative *Myotis lucifugus* they have significantly

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.06.080200; this version posted June 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

longer lifespans than other bats (Figure 4 A, blue bars). These two species share a unique arrangement in the DNA binding region, with a seven amino acid residue insertion in the central DNA-binding region (following aa 295 in the human p53 canonical sequence) (Figure 4B). To test how this rearrangement in the DNA binding region changes the interaction of p53 with DNA, we modelled the p53 tetramer using SWISS-MODEL workflow. The insertion in the DNA binding domain of bats with long lifespan is present inside the DNA interaction cavity, suggesting decreased affinity of p53 for binding to DNA (Supplementary Material 2). *Myotis Brandtii* and *Myotis lucifugus* are very small bats (max 8 g bodyweight) and form a significant exception from the Max Kleiber's law (Kleiber, 1932) (mouse-to-elephant curve; their lifespan is extremely long in relation to their small body size).

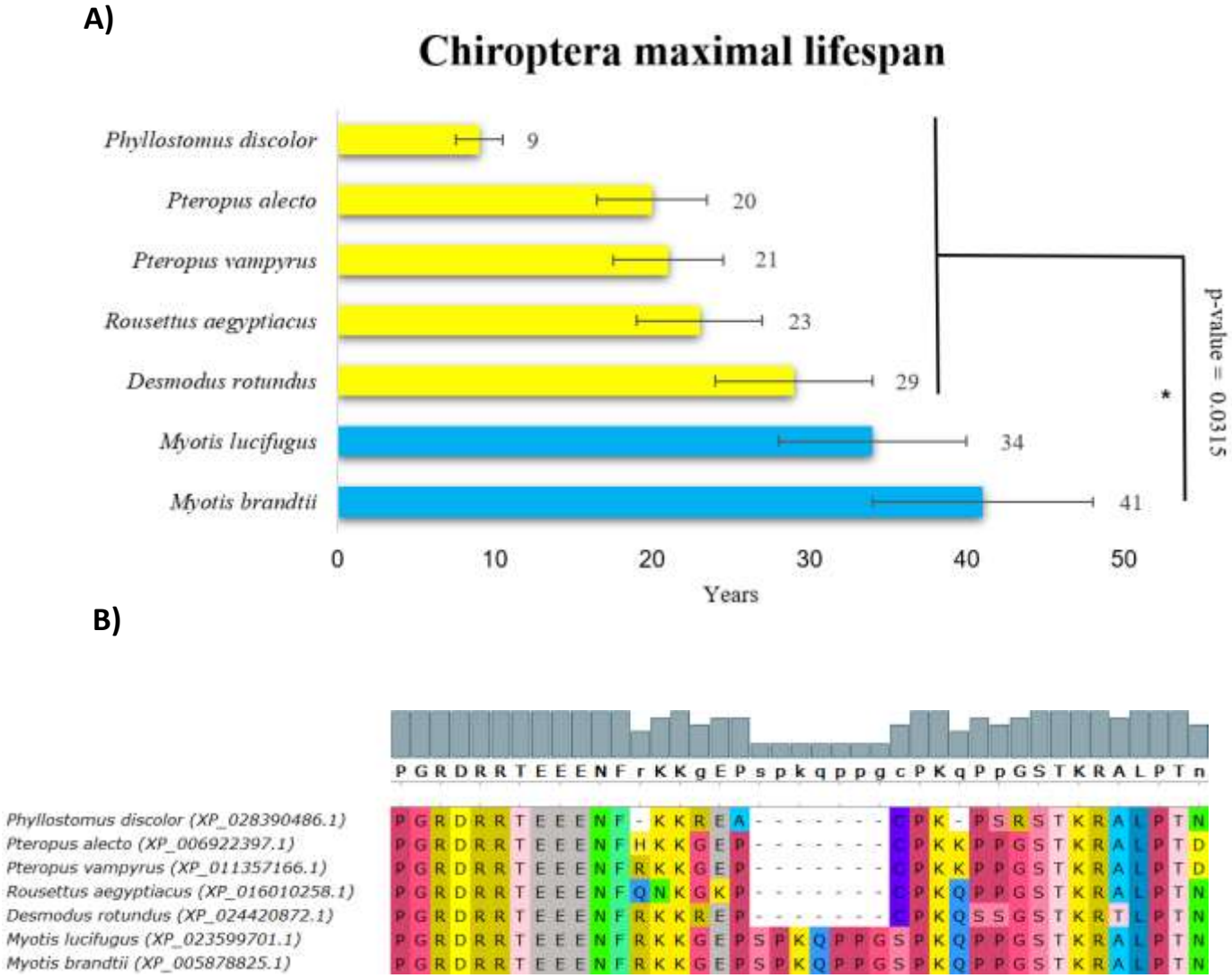
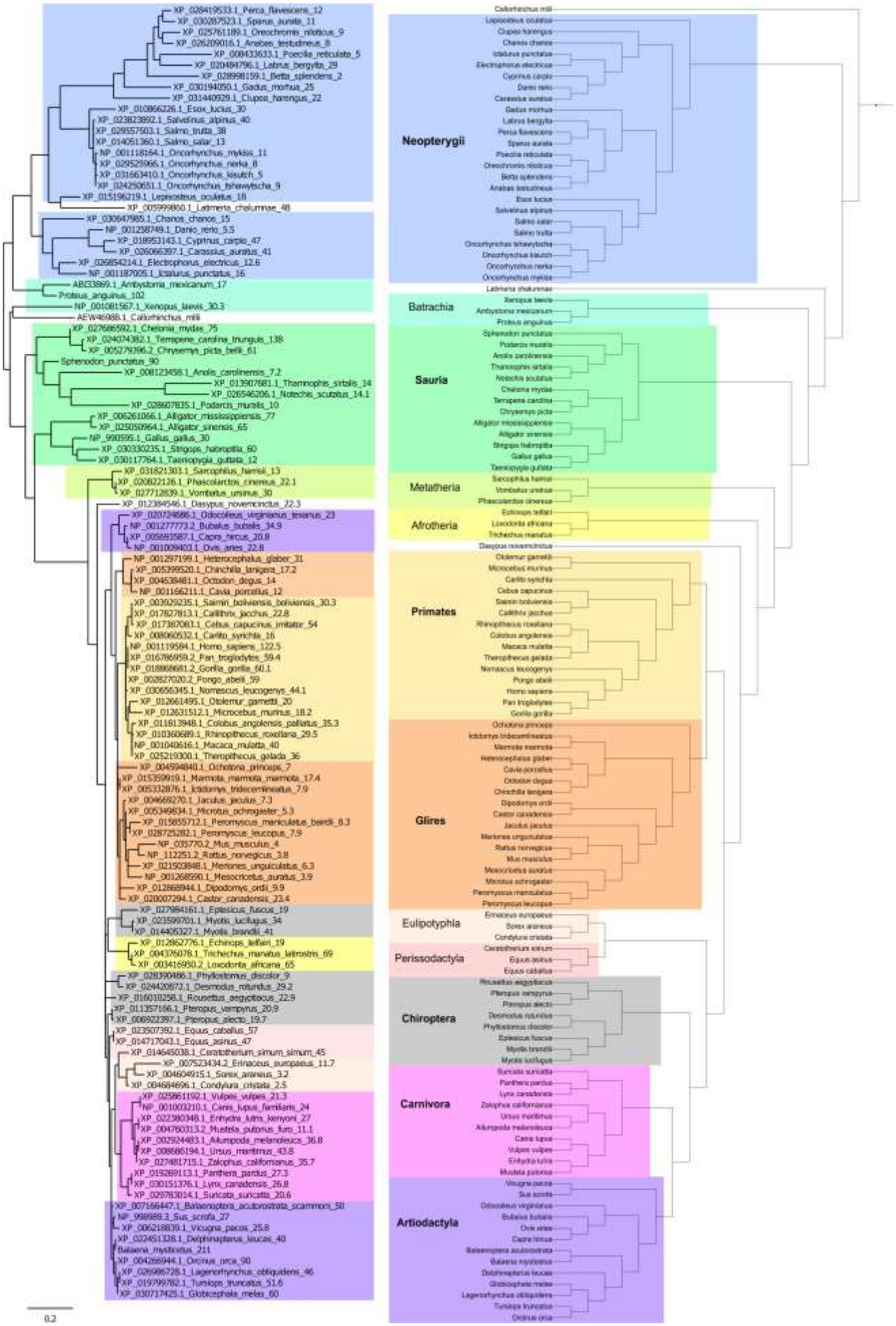


Figure 4: (A) Comparison of maximal Chiroptera lifespan in years. Brandt´s bat (*Myotis brandtii*) and *Myotis lucifugus* maximal lifespans are significantly longer compared with other sequenced bats (Wilcoxon one-sided signed rank test, p-value < 0.05). (B). Multiple protein alignments of the C-terminal part of the p53 core domain of accessible Chiroptera sequences. Methods and color schemes are the same as in Figure 1B.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.06.080200; this version posted June 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

The above-mentioned analysis of long-lived organisms in various animal groups led us to hypothesize that the sequence of p53 domains is associated with organismal lifespan. Therefore, we continued the analysis by further correlating the p53 amino acid sequence with the lifespan across the animal kingdom. Due to low homology between p53 N-terminal and C-terminal domains across species and a significant role of mutations in the p53 DNA binding domain in cancer, we focused on the most conserved core domain of p53 and constructed the p53-based tree (Bartas et al., 2020). We then compared contemporary phylogenetic tree with the tree based on p53 protein sequence (Figure 5). Then, the dataset with p53 sequences and animal lifespans were divided into 12 groups based on their phylogenetic relationships. Interestingly, some p53 variations are not closely associated with the phylogenetic tree but point to several parallel evolutionary processes (Figure 5). Even closely related species in various groups have significantly different lifespans (Supplementary Material 1) and therefore, are suitable for correlation analyses according to the method introduced by Jensen and colleagues (Jessen et al., 2013).

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.06.080200; this version posted June 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.



bioRxiv preprint doi: https://doi.org/10.1101/2020.05.06.080200; this version posted June 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figure 5: Comparison of p53 protein tree (left) and the real phylogenetic tree (right). The protein tree was built using Phylogeny.fr platform. The real phylogenetic tree was reconstructed using PhyloT and visualized in iTOL (see Methods for details). The color background represents the same phylogenetic groups.

The summary of the data for each group is visualized in Figure 6 and the total number of analyzed animals for each group with minimal and maximal values are shown in Supplementary Material 3. Only datasets with more than 5 members in the group were used in the correlation analyses.

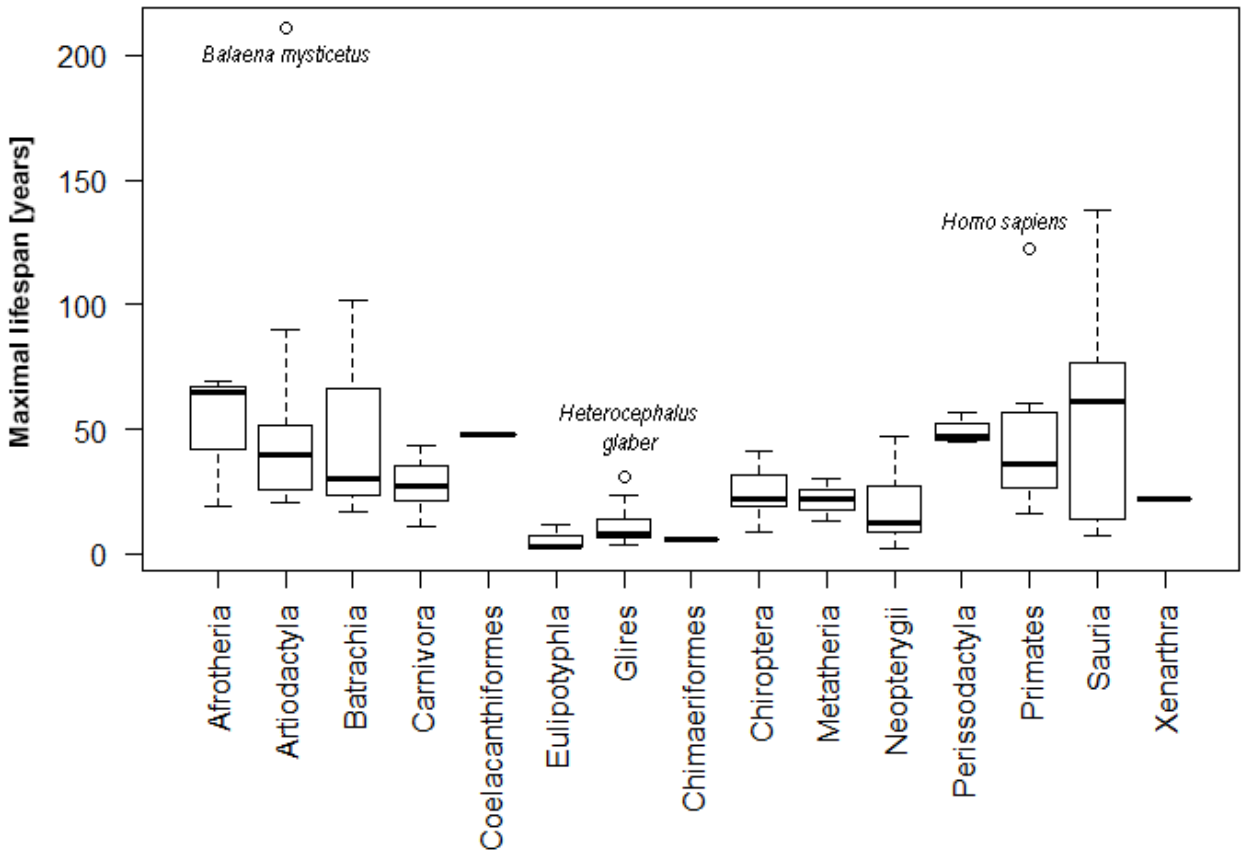


Figure 6: Box-plot representation of lifespans for all tested phylogenetic groups.

The organisms with the longest lifespan in the **Neopterygii** dataset are the carp (*Cyprinus carpio* (47 years)), followed by the goldfish (*Carassius auratus* (41 years)). Siamese fighting fish (*Betta splendens)* has the shortest lifespan in the group (2 years). The correlation analyses show that fifteen-amino acid-residues in the p53 core domain are significantly associated with prolonged lifespan (Figure 7). We found that the most common variation in the long-lived Neopterygii is the presence of serine (at positions corresponding to 98, 128 and 211 of human p53) and the presence of valine (at positions 128, 150, 217, 232). On other hand, in the short-lived organisms in Neoropterygii, we identified threonine (at positions 98, 100, 141, 217, 260), glutamic acid (at positions 110, 128, 150 and 291) and serine (at positions - 141,

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.06.080200; this version posted June 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

203, 235). It seems that the abundance of glutamic acid could relate to worse p53-DNA binding affinity due to the local change in the ionic charge at the site of the amino acid p53 variant. Indeed, the PROVEAN tool predicted a deleterious effect on p53 function for glutamic acid at position 128. In addition, according to PHANTM classifier, C141S substitution leads to a decrease of p53 transcription activity by 41.08% compared to wt-p53.

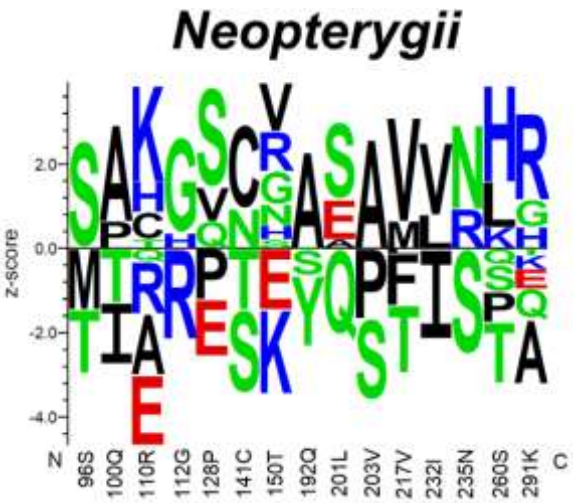


Figure 7: Logo quantifying the strength of p53 core domain residue association (related to the human aa 94 – 293 according to p53 canonical sequence) with the maximal lifespan in years in the Neopterygii group. Amino acid residues on the positive y-axis are significantly associated with the prolonged lifespan phenotype and residues on the negative y-axis are significantly associated with the shorter lifespan phenotype (significance threshold p-value ≤ 0.05). The height of each letter representing the strength of the statistical association between the residue and the data set-phenotype. The amino acids are colored according to their chemical properties as follows: Acidic [DE]: red, Basic [HKR]: blue, Hydrophobic [ACFILMPVW]: black and Neutral [GNQSTY]: green.

The lifespan of species in **Sauria** is significantly variable. The organisms with the longest lifespan in this group are three-toed box turtle (*Terrapene carolina triunguis* (138 years)) and kakapo (*Strigops habroptila* (95 years)). Green anole (*Anolis carolinensis)* has the shortest lifespan in the group (7.2 years). The correlation analyses show that similarly to **Neopterygii** a specific fifteen amino acid residue fragment in the p53 core domain is associated with the prolonged lifespan (Figure 8). The most common p53 variation for long-lived Sauria is similar to Neopterygii and it is the presence of serine (at positions 94, 95, 149 and 227 – corresponding to human p53) and the presence of valine (at positions 97 and 232, identical to Neopterygii). When compared to human p53, in short-lived organisms we identified threonine at positions 94, 149, 159 and 227 and glutamic acid at positions 114, 192 and 228. In addition, deletions in the p53 sequence were found at positions 94-97 and 114 (Figure 7). Similar to Neopterygii, the most common p53 variation for short-lived Sauria is the presence of threonines and glutamic acid residues.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.06.080200; this version posted June 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

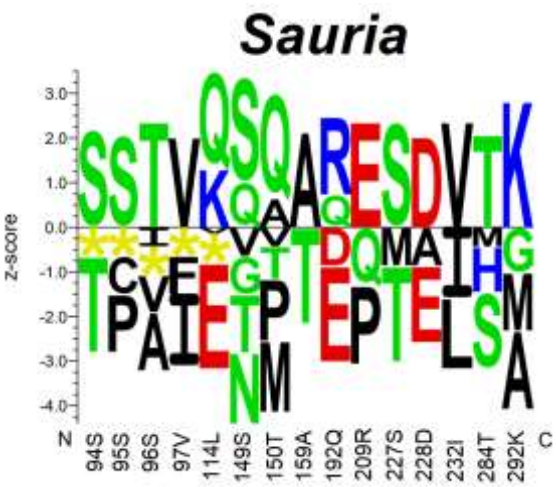


Figure 8: Logo quantifying the strength of p53 core domain residue association (related to the human aa 94 – 293 according to p53 canonical sequence) with the maximal lifespan in years in the Sauria group. For details see Figure 7.

The organisms with the longest lifespan in **Primates** group are human (*Homo sapiens* (122 years)) and western gorilla (*Gorilla gorilla* (60 years)). Tarsier (*Carlito syrichta)* has the shortest lifespan in the group (16 years). The correlation analyses show that the specific amino acid triad is significantly associated with a prolonged lifespan (Figure 9). Beside a serine residue at position 106, two others; glutamine at position 104 and leucine at position 289 are both hydrophobic. On the contrary, proline or histidine at position 104, asparagine at position 106 and phenylalanine, serine or tyrosine at position 289 are associated with short-living primates. While studying human longevity, one needs to consider that prolonged lifespan of *Homo sapiens* is associated with the cultural and socio-economical advantages. Therefore, we have additionally performed analyses upon excluding *Homo sapiens* from the dataset. The same variations were observed in the correlation analyses. Taken together, our results show that the amino acid variations shown in Figure 9 are conserved in the following closely related species (*Homo sapiens*, *Pan troglodytes* and *Gorilla gorilla*).

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.06.080200; this version posted June 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

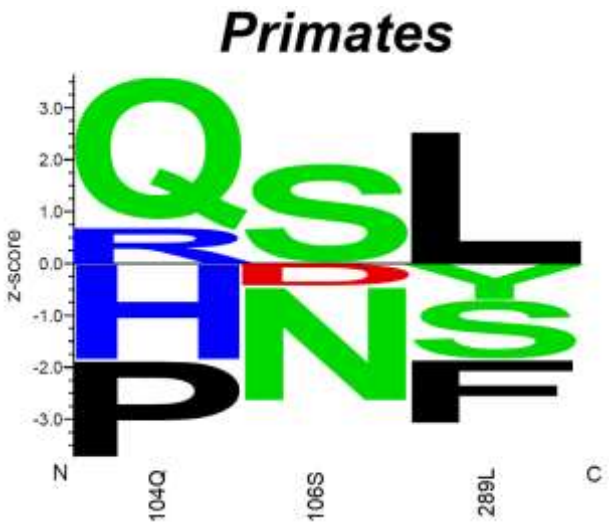


Figure 9: Logo quantifying the strength of p53 core domain residue association (related to the human aa 94 – 293 according to p53 canonical sequence) with the maximal lifespan in years in the Primates group. For details see Figure 7.

The dataset of **Glires** contains seventeen species with lifespans ranging from 3.8 to 31 years. The organisms with the longest lifespan in this group are *Heterocephalus glaber* (31 years) and *Castor canadensis* (23). The shortest lifespan in the group is *Rattus norvegicus* (3.8 years). The correlation analyses show that ten amino acid residues are significantly associated with prolonged lifespan (Figure 10). Two threonine residue variations (positions 123 and 210) are present in long-lived Glires. Other animo acid changes occur only once. Interestingly, in short-lived Glires, there is a significant presence of threonine also at two other locations (positions 148 and 150). Similar variations are also observed in the methionine residues (at positions 123 and 201), tyrosine (positions 202 and 229) and proline (positions 185 and 201).

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.06.080200; this version posted June 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

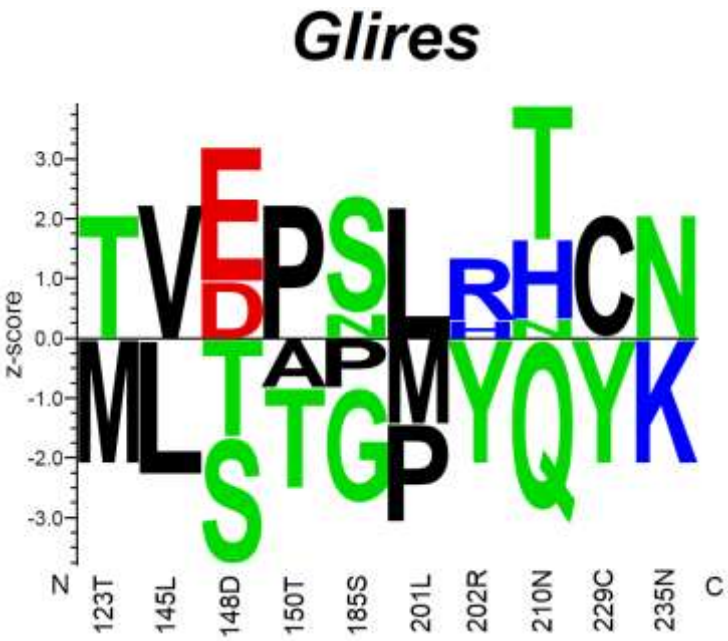


Figure 10: Logo quantifying the strength of p53 core domain residue association (related to human aa 94 – 293 according to p53 canonical sequence) with maximal lifespan in the Glires group. For details see Figure 7.

The organisms with the longest lifespan in the dataset of **Chiroptera** are Brandt's bat (*Myotis brandtii* (41 years)) and little brown bat (*Myotis lucifugus* (29 years)). Pale spear-nosed bat (*Phyllostomus discolour)* has the shortest lifespan (9 years). The correlation analyses show that nine amino acid residues are associated with significant prolonged lifespan association (Figure 11).

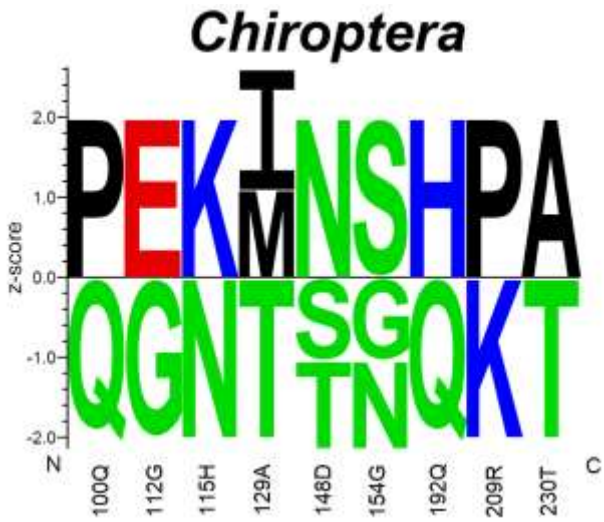


Figure 11: Logo quantifying the strength of p53 core domain residue association (related to the human aa 94 – 293 according to p53 canonical sequence) with the maximal lifespan in years in the Chiroptera group. For details see Figure 7.

The organisms with the longest lifespan in **Carnivora** group are polar bear (*Ursus maritimus)* and panda (*Ailuropoda melanoleuca)*. The shortest lifespan in the group is ferret (*Mustela putorius furo*). The correlation analyses show that two amino acids in the p53 core domain (positions 148 and 232) are

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.06.080200; this version posted June 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

significantly associated with lifespan (Figure 12). While the presence of asparagine (at position148) and valine (at position 232) is associated with long lifespan, the presence of serine (at position 148) and isoleucine (at position 232) is associated with short lifespan.

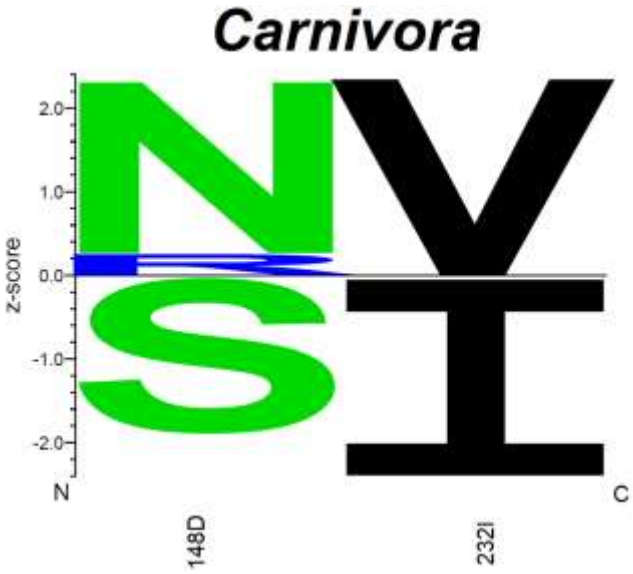


Figure 12: Logo quantifying the strength of p53 core domain residue association (related to the human aa 94 – 293 according to p53 canonical sequence) with the maximal lifespan in years in the Carnivora group. For details see Figure 7.

The organism with the longest lifespan in **Artiodactyla** group is bowhead whale (*Balaena mysticetus* (211 years)) followed by orca (*Orcinus orca* (90 years)). Correlation analyses show that twelve amino acid residues in the p53 core domain are significantly associated with prolonged lifespan (Figure 13). Similar to Neopterygii and Sauria, the most common variation present in the long-lived organisms are associated with serine (at positions 106, 148 and 166 – corresponding to human p53). The variation of serine at positions 129, 182 and 222 is the most common variation for short-lived Artiodactyla, together with variation in the glutamic acid residues at positions 226 and 228.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.06.080200; this version posted June 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

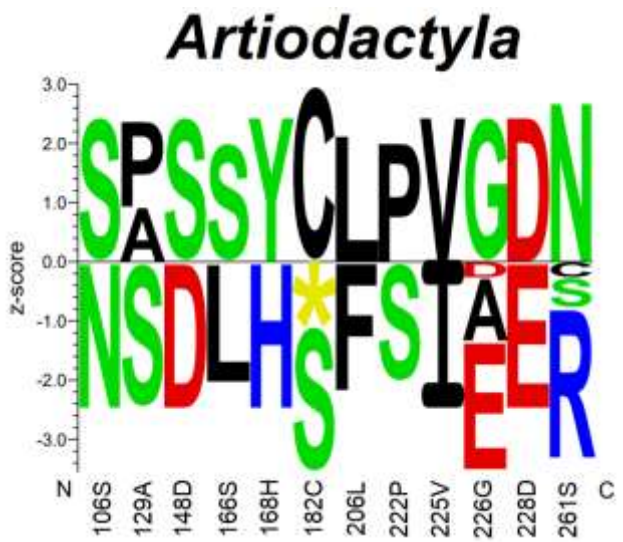


Figure 13: Logo quantifying the strength of p53 core domain residue association (related to the human aa 94 – 293 according to p53 canonical sequence) with the maximal lifespan in years in the Artiodactyla group. For details see Figure 7..

Next, we investigated all 118 RefSeq p53 sequences to evaluate associations between amino acid variations and maximal lifespan (Figure 14). When applying the Bonferroni correction (Figure 14A), only two significantly associated residues were revealed, corresponding to human serine 185 and asparagine 210. Organisms that have serine at position 185 live statistically longer than organisms with another amino acid in this position. On other hand, organisms that contain glutamine instead of asparagine at position 210 have significantly shorter maximal lifespan. Without Bonferroni correction, from the 200 analyzed positions of the aligned p53 core domains (related to human 94-293 aa), 64 positions were significantly associated with lifespan (Figure 14B). Positive correlations with longevity are shown by orange and red colors, green and blue show negative correlations.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.06.080200; this version posted June 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

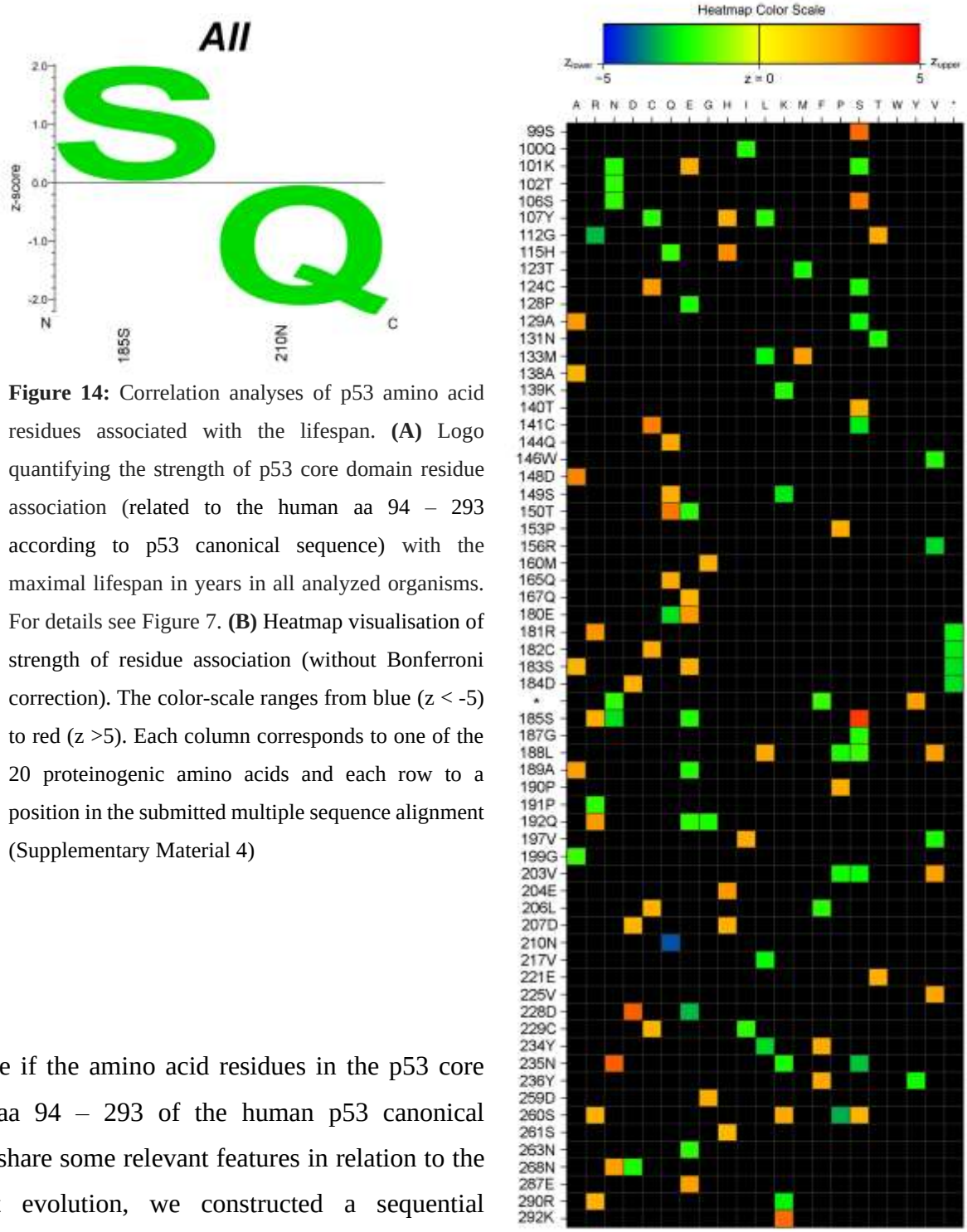


Figure 14: Correlation analyses of p53 amino acid residues associated with the lifespan. **(A)** Logo quantifying the strength of p53 core domain residue association (related to the human aa 94 – 293 according to p53 canonical sequence) with the maximal lifespan in years in all analyzed organisms. For details see Figure 7. **(B)** Heatmap visualisation of strength of residue association (without Bonferroni correction). The color-scale ranges from blue (z < -5) to red (z >5). Each column corresponds to one of the 20 proteinogenic amino acids and each row to a position in the submitted multiple sequence alignment (Supplementary Material 4)

To evaluate if the amino acid residues in the p53 core domains (aa 94 – 293 of the human p53 canonical sequence) share some relevant features in relation to the convergent evolution, we constructed a sequential circular representation of the multiple sequence alignments and the mutual information it contains (Figure 15). This figure shows that the amino acid residues significantly associated with aging (taken from Figure 14, heatmap, and highlighted in light green) very often coevolved together (represented by connected lines). This may be considered as a supportive evidence for the convergent evolution of p53 proteins in organisms with extreme longevity. According to Passow and colleagues, taxa with evidence

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.06.080200; this version posted June 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

of positive selection in the *TP53* gene are those with the lowest incidences of cancer reported in amniotes (elephants, snakes and lizards, crocodiles and turtles) (Passow et al., 2019).

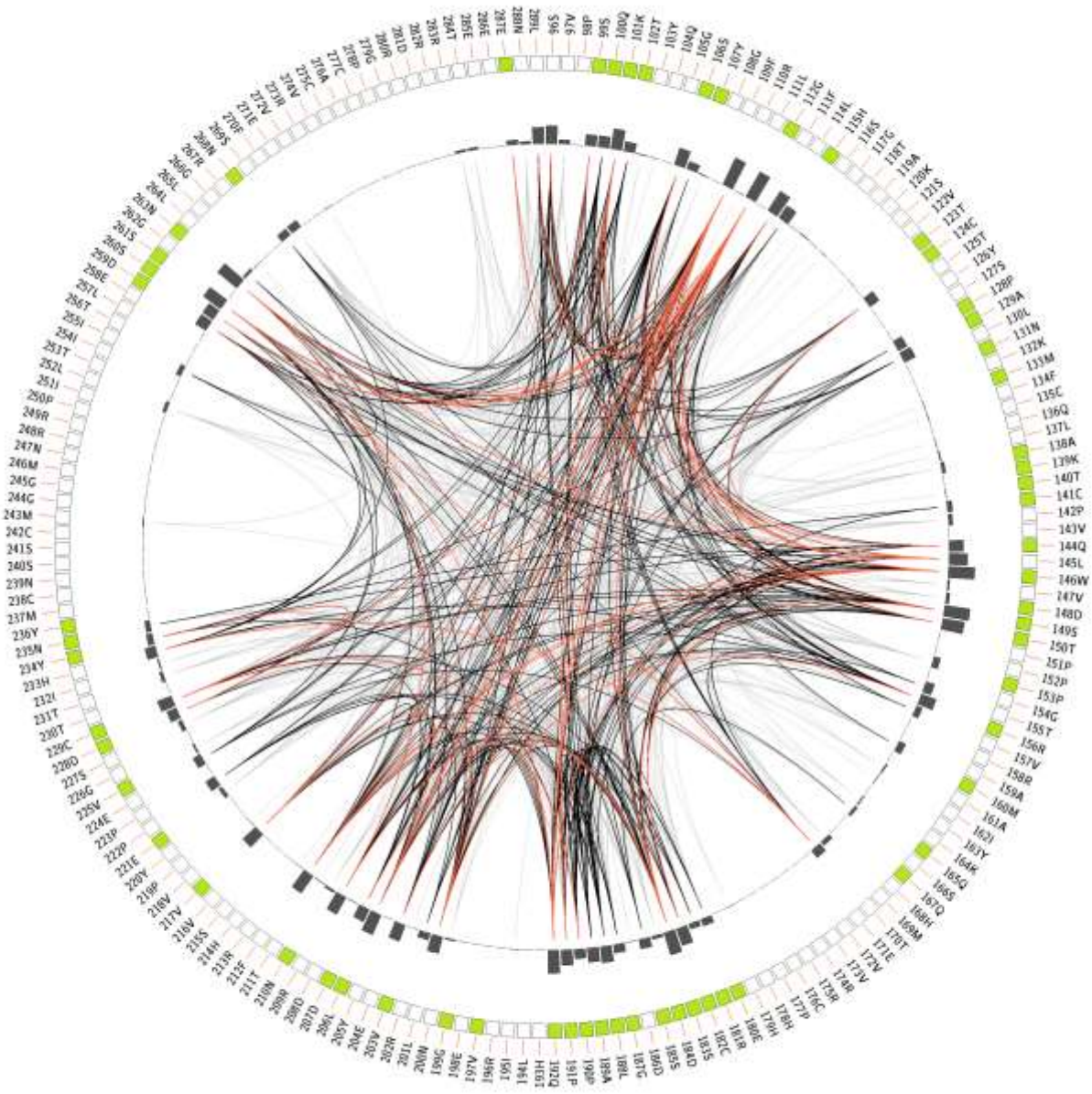


Figure 15: Mutual information to infer convergent evolution of p53 core domains. Circos plot is a sequential circular representation of the multiple sequence alignment and the information it contains. Green boxes in the outer circle indicate amino acid residues significantly correlated with maximal lifespan. Lines connect pairs of positions with mutual information greater than 6.5 (Buslje et al., 2009). Red edges represent the top 5%, black are between 70% and 95%, and gray edges account for the remaining 70%.

Table 1 summarizes the most relevant cases of p53 variations identified in our study. Apart from the unique substitutions (*Strigops habroptila*, *Balaena mysticetus*) and insertions (*Myotis Brandtii*, *Myotis lucifugus*, *Proteus anguinus*), a complete lack of *TP53* mRNA expression was found in *Turritopsis sp.*. In

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.06.080200; this version posted June 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

addition, a substitution in the nuclear localization signal (NLS) was observed in *Balaena mysticetus*, *Heterocephalus glaber*, *Loxodonta africana* and *Strigops habroptila* (Supplementary material 5). The NLS is required for transferring p53 from the cytoplasm into the nucleus, where p53 can bind DNA. It is well known that mutation of basic amino acid residues in the NLS compromises the transport of p53 to the nucleus (Addison et al., 1990; Shaulsky et al., 1990). Thus, organisms without a functional NLS will have only very limited or no p53 transcription activity.

Tab. 1: Comparison of animals characterized by extreme longevity and their atypical p53 features, where significance of particular changes were predicted. Default PROVEAN threshold -2.5 was used, insertions and deletions were submitted in respect to the human canonical protein sequence (NP_001119584.1) "*" indicates significant PROVEAN values (<-2.5).

Organism, classification	**Maximal lifespan [y]**	**Adult weight [kg]**	**P53 oddities**	**Effect by PROVEAN**
Balaena mysticetus Mammalia, Cetacea	211	100000	Unique substitution in proline rich region; weakest NLS within Cetacea	Neutral (P77L, Score = -0.993)
Heterocephalus glaber Mammalia, Rodentia	31	0.035	p53 sequence more similar to human p53 than mouse/rat; weak NLS	Not applicable
Loxodonta Africana Mammalia, Proboscidea	65	4800	Multiple copies of p53 genes, inhibitory effect via non-functional tetramers (?)	Not applicable
Myotis Brandtii, Myotis lucifugus Mammalia, Chiroptera	41	0.007	Insertion in DNA binding domain	Deleterious (P295_H296insPKQPPGS, Score = -2.526) *
Strigops habroptila Aves, Psittaciformes	95	111	Substitution in core domain; abolished NLS	Deleterious (N131H, Score = -3.162) *
Proteus anguinus Amphibia, Urodela	102	0.017	Insertion nearby dimerization region	Deleterious (L188_A189insSR, Score = -3.357) *
Turritopsis sp. Cnidaria	∞ rejuvenation	0.001	No p53/63/73 protein expressed (unprecedent phenomenon in the whole animal kingdom)	Not applicable

Taken together, our analyses have identified, for the first time, correlation between p53 sequence variations and longevity in the animal kingdom.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.06.080200; this version posted June 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Discussion

The p53 protein is a well-known tumor suppressor and *TP*53 is the most often mutated gene in human cancers. On the cellular level, decreased p53 functionality is essential for cellular immortalization and neoplastic transformation (Soussi & Wiman, 2015). However, the role of variations in the p53 amino acid sequence at the organism level has not been studied systematically. Here, we presented an in-depth correlation analysis manifesting the dependencies between p53 variations and organismal lifespan to address the role of p53 in longevity in the animal kingdom. To date, p53 expression has been detected in all sequenced animals from unicellular Holozoans to vertebrates (Bartas et al., 2020). The seminal work by Kubota provided important evidence demonstrating that immortality is not just a hypothetical phenomenon. In his work he demonstrated that Cnidarian species *Turritopsis* jellyfish is immortal and can repeatedly rejuvenate, reverse its life-cycle and, thus, was the first and only known "immortal" animal on earth (Kubota, 2011). Here, we have inspected recently published data from the whole-transcriptome data of "immortal" *Turritopsis sp.* (Hasegawa et al., 2016) and surprisingly found that no expression of any of the p53 family members was detected in the pooled data from all individuals at all developmental stages (polyp, dumpling with a short stolon, dumpling and medusa). This pointed to the possibility that the absence of p53 in *Turritopsis* might be directly related to its unique ability of life cycle reversal and "immortality".

The results from Protein Variation Effect Analyzer (Choi & Chan, 2015) show that the variability in lifespan among closely related species correlates with specific p53 variations. Long-lived organisms are characterized by in-frame deletions/changes, insertions or specific substitutions in the p53 sequence. It is likely that the changes imposed on p53 in long-lived species enable p53 to interact with different multiple protein partners to induce gene expression programmes varying from those induced in species with relatively normal lifespan. We can anticipate that these gene expression programmes would enable following changes (Figure 16): 1. more efficient tissue repair through autophagy, 2. loss of senescence, 3. enhanced clearance of senescent cells by the immune system, 4. enhanced regulation of intracellular ROS levels 5. improved resistance of mitochondria to ROS-induced damage or 6. loss of immune senescence that occurs in humans with age. All of the mentioned processes have been previously described as significantly contributing to longevity in humans (reviewed in (Rufini et al., 2013)). Intriguingly, a recent GWAS study on 1 million parent lifespans, identified only several variants influencing lifespan at genome-wide significance including *CDKN2B-AS1* and *IGF2R*. The *TP53* was not among the singled-out variants (Timmers et al., 2019) which, in accordance with our observations, indicate that no changes in human p53 might be attributing to longevity in humans. Thus, long-lived organisms apparently have a different mechanism of protection against cancer and their lifespan is not limited by somatic cell senescence caused by active p53 protein, which is the case for other species with shorter lifespan as mentioned above.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.06.080200; this version posted June 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

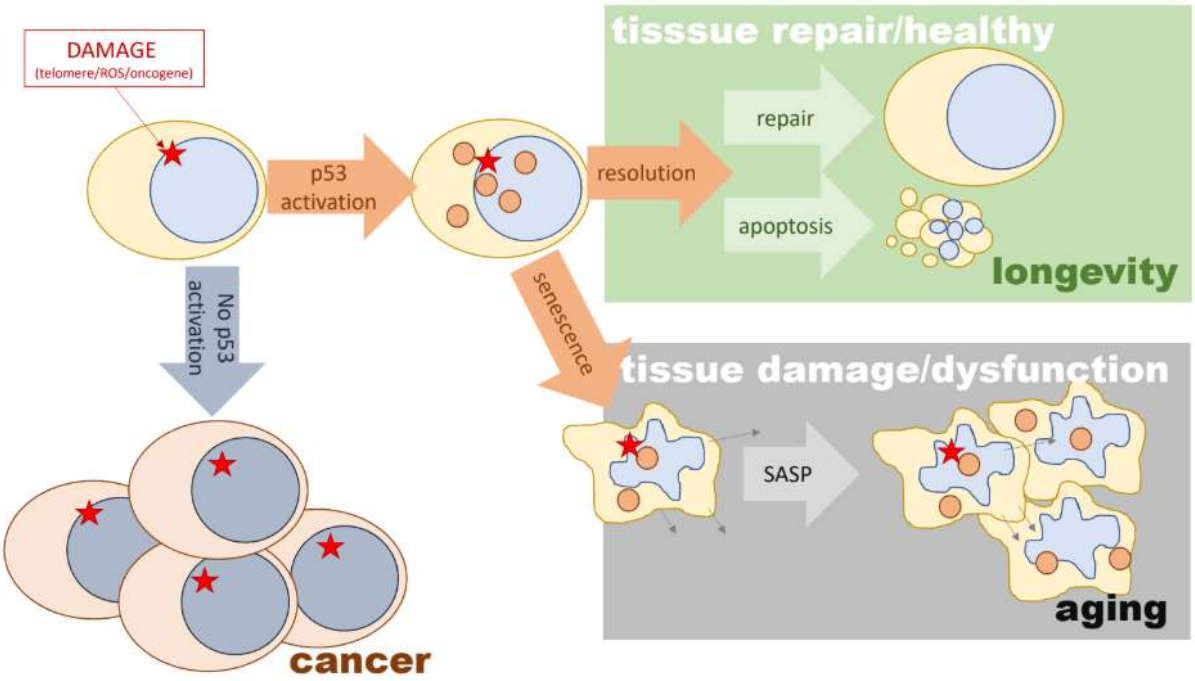


Figure 16. Proposed p53-centric theory of extreme longevity. Cell damage caused by ROS, telomere shortening, or oncogene activation activates p53 to enable repair, resolution (repair/apoptosis) supports tissue integrity, while sustained pausing promotes senescence and via senescence-associated secretory phenotype (SASP) leads to degeneration and premature aging

The maximal lifespan according to the AnAge database is attributed to Greenland shark with an estimated maximal life span of 300–500 years. Unfortunately, no transcriptomic nor genomic data for the Greenland shark (*Somniosus microcephalus*) are available. Compared to other sharks (with life expectancy of up to seventy years), its lifespan is extremely exceptional. It will be thus very interesting to know the sequence of their p53 protein.

Several experimental data support our hypothesis about specific p53 variation association with longevity. For example it has been found that the reduced expression of the *Caenorhabditis elegans* p53 ortholog cep-1 results in increased longevity (Arum & Johnson, 2007). It has also been demonstrated that neuronal expression of p53 dominant-negative proteins in adult *Drosophila melanogaster* extends lifespan (Bauer et al., 2005). The same principle is most probably present in humans, where, for example, p53 variants predisposing to cancer are present in healthy centenarians (Bonafè et al., 1999) and a meta-analyses showed that the codon 72 polymorphic variant of p53 with proline (compared to arginine) provides an increased cancer risk, but increased survival (van Heemst et al., 2005). In a recent study by Zhao et al., polymorphism at position 72 (P72 compared to R72) was reported to have a positive effect on lifespan and to delay the development of aging-related phenotypes in mice, supporting a role of p53 activity in longevity (Zhao et al., 2018). Another example of a long-lived vertebrate is the elephant, which has 20 copies of the *TP53* gene (Sulak et al., 2016). In this species, part of the DNA-binding region

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.06.080200; this version posted June 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

is deleted in all but one of the *TP*53 gene copies, which may result in the formation of dysfunctional p53 tetramers, thus, presumably, modulating p53 transcriptional activity in response to stress (Sulak et al., 2016).

Despite high complexity of the p53/63/73 protein family, modern methods of comparative genomics provide useful tools to explore protein variations in closely related species, and to correlate the extracted molecular information with lifespan. According to Sahin and DePinho, the increased activity of p53 in the presence of accumulated ROS is one of the main causes of aging (Sahin & DePinho, 2012). This observation is in congruence with our hypothesis that organisms with atypical p53 sequences leading to changed p53 activity are extremely long-lived. Even if several p53 differences have been found in various animal groups, some variations developed in convergent evolutions in different groups of species. For example, the presence of threonine and glutamic acid was observed in short-lived organisms of different groups and richness of serine residues was typical for long-lived organisms in several groups and serine residue at position 285 is significantly associated with prolonged lifespan across all analyzed species.

This study reveals a previously overlooked correlation between longevity and a potential change in p53 function due to p53 amino acid variations across the animal kingdom. Strikingly, several long-lived species, including *Myotis brandtii*, *Myotis lucifugus*, *Balaena mysticetus*, *Heterocephalus glaber*, *Strigops habroptila* and *Proteus anguinus* display unique p53 sequence properties not shared with their close relatives that have a shorter lifespan. Altogether, our evidence suggests convergent evolution of p53 sequences supporting a higher insensitivity to p53-mediated senescence in long-lived vertebrates. Our observations that specific variations of p53 protein are correlated with lifespan provide important grounds for further exploration of p53 sequences in species displaying extreme longevity. Most importantly, our data from a wide variety of vertebrates implies a general mechanism at work in all vertebrates leading to extended lifespan, which might be translated to studies on extension of the health span in humans.

Materials and methods

Searches of maximal lifespan

To access data of longevity and maximal lifespan, we used AnAge Database (https://genomics.senescence.info/species/, AnAge currently contains data on longevity of more than four thousand animals.) (De Magalhaes & Costa, 2009). We have downloaded the whole dataset and selected species presented in NCBI RefSeq database.

Protein homology searches

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.06.080200; this version posted June 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

For protein homology searches we have downloaded all available p53 sequences from RefSeq database (https://www.ncbi.nlm.nih.gov/refseq/) and merged them with AnAge Database. We received the p53 sequence information of 118 species with information about their lifespan and sorted them according to their phylogenetic group (Supplementary Material 1). In animals with extreme longevity, where the p53 homologs were not present in NCBI, local blast searches (tblastn) applied on *de novo* assembled transcriptomes were used together with the default "BLAST+ make database" command and searching parameters within UGENE standalone program (Okonechnikov et al., 2012).

Transcriptome assemblies

Transcriptomic data for bowhead whale was obtained from http://www.bowhead-whale.org/ (Keane et al., 2015). When there were only raw seq reads from the RNA-seq experiments available (deposited in the NCBI SRA), we performed the *de novo* assembly first, using Trinity tool (Grabherr et al., 2011) from the Galaxy webserver (https://usegalaxy.eu/)(Afgan et al., 2018) with default settings. This was done for *Proteus anguinus* (SRX2382497) and *Sphenodon punctatus* (SRX4014663), resulting assemblies are enclosed in Supplementary Material 6.

p53 protein tree and real phylogenetic tree construction

The protein tree was built using Phylogeny.fr platform (http://www.phylogeny.fr/alacarte.cgi) (Dereeper et al., 2008, 2010) and comprised the following steps. First, the sequences were aligned with MUSCLE (v3.8.31) (Edgar, 2004) configured for the highest accuracy (MUSCLE with default settings). After alignment, ambiguous regions (i.e. containing gaps and/or poorly aligned) were removed with Gblocks (v0.91b) (Castresana, 2000) using the following parameters: -minimum length of a block after gap cleaning: 10; -no gap positions were allowed in the final alignment; -all segments with contiguous non-conserved positions longer than 8 were rejected; -minimum number of sequences for a flank position: 85%. The phylogenetic tree was constructed using the maximum likelihood method implemented in the PhyML program (v3.1/3.0 aLRT) (Anisimova & Gascuel, 2006; Guindon & Gascuel, 2003). The JTT substitution model was selected assuming an estimated proportion of invariant sites (of 0.204) and 4 gamma-distributed rate categories to account for rate heterogeneity across sites. The gamma shape parameter was estimated directly from the data (gamma=0.657). Reliability for the internal branch was assessed using the bootstrapping method (100 bootstrap replicates). Graphical representation and edition of the phylogenetic tree were performed with TreeDyn (v198.3) (Chevenet et al., 2006). Real phylogenetic tree was reconstructed using PhyloT (https://phylot.biobyte.de/) and visualized in iTOL (https://itol.embl.de/) (Letunic & Bork, 2019).

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.06.080200; this version posted June 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Prediction and statistical evaluation by PROVEAN

The effect of the p53 variations in long-lived organisms was predicted and statistically evaluated by Protein Variation Effect Analyzer web-based tool (PROVEAN; http://provean.jcvi.org/index.php) (Choi et al., 2012; Choi & Chan, 2015). PROVEAN is a software tool which predicts whether an amino acid substitution or in/del has an impact on the biological function of a protein (Choi & Chan, 2015). All inspected p53 variations in selected animals were statistically evaluated and numbered according to the human canonical p53 sequence (NP_000537.3).

Modelling of 3D Protein Structures

We used SWISS-MODEL template-based approach (https://www.swissmodel.expasy.org/interactive) (Waterhouse et al., 2018) to predict 3D structures using individual FASTA sequences and reference PDB:4mzr as the crystal structure of p53 tetramer from *Homo sapiens* with bound DNA (Emamzadah et al., 2014). All resulting PDB files are enclosed in Supplementary Material 7. Predicted structures of p53 homologs were visualized in UCSF Chimera 1.12 (Pettersen et al., 2004).

Scanning for nuclear localization signals

The NucPred webserver in Batch mode was utilized to identify nuclear localization signals, (https://nucpred.bioinfo.se/nucpred/). NucPred is an ensemble of 100 sequence-based predictors. If the fraction of predictors giving a "yes" answer (also known as the NucPred score) exceeds nuclear localisation signal threshold, then the protein is predicted to have a nuclear role (Brameier et al., 2007). Complete predicted NLS scores for each species are in Supplementary Material 5.

Correlation of maximal lifespan and alterations within the p53 core domain in vertebrates

Residue level genotype/phenotype correlations in p53 multiple sequence alignment were performed using SigniSite 2.1 (http://www.cbs.dtu.dk/services/SigniSite/) (Jessen et al., 2013) with significance threshold p-value ≤ 0.05. Bonferroni single-step correction for multiple testing was applied for the global correlation of all sequences, no correction was applied for smaller groups of taxonomically related animals. The manually curated set of 118 high-quality p53 protein sequences obtained from the NCBI (https://www.ncbi.nlm.nih.gov/) was used as input file. These sequences were taken from the RefSeq database and only the canonical isoform 1 was filtered for each vertebrate species. The resulting set of these 118 p53 sequences was aligned within UGENE workflow (Okonechnikov et al., 2012), MUSCLE algorithm (Edgar, 2004) with default parameters. All sequences were then manually trimmed to preserve only the core domain, which corresponds to human 94 – 293 aa. Then the numerical values of maximal

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.06.080200; this version posted June 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

lifespan of each organism were added into the resulting FASTA file, based on the information in the reference AnAge database (http://genomics.senescence.info/species/) (De Magalhaes & Costa, 2009)

Convergent evolution

Multiple sequence alignment of p53 core domains from 118 species was uploaded to the MISTIC webserver (http://mistic.leloir.org.ar/index.php), with PDB 2ocj (A) as the reference and using default parameters (Simonetti et al., 2013)**.**

Gene gain and losses

*TP*53 gene gain or losses were inspected using Ensembl Comparative Genomics toolshed (Herrero et al., 2016) via Ensembl web pages and *TP53* gene query ENSG00000141510: https://www.ensembl.org/Homo_sapiens/Gene/SpeciesTree?db=core;g=ENSG00000141510;r=17:7661779-7687550

Acknowledgements
We thank Dr. Jean-Christophe Bourdon for valuable comments and discussion, Dr. Philip Coates for proofreading and editing. We would also like to express our gratitude to M. Sc. Alena Volná and Milan Bolek for their time spent on illustration preparation. This work was supported by The Czech Science Foundation (18-15548S) and by the SYMBIT project Reg. no. CZ.02.1.01/0.0/0.0/15_003/0000477 financed from the ERDF."

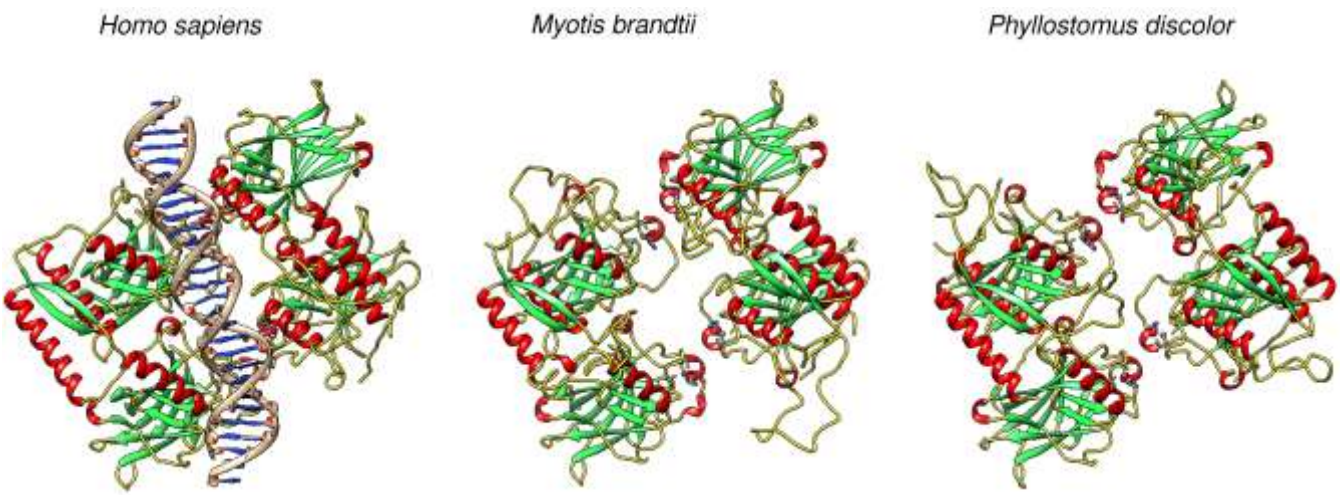
Conflict of interests
The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Supplementary materials:

Supplementary material 1: Table of all analyzed organisms, their maximal lifespan and reference p53 sequences

Supplementary material 2: 3D structures of p53 core domains. (Left) Crystal structure of human p53 mutant tetramer (S121F and V122G) bound to DNA (PDB: 4mzr) (Emamzadah et al., 2014). (Middle) Modelled structure of long-living bat (*Myotis brandtii*) p53 tetramer according to 4mzr template. (Right) Modelled structure of short-living bat (*Phyllostomus discolor*) p53 tetramer according to 4mzr template. Automatic modeling was done using SWISS-MODEL workflow (Waterhouse et al., 2018).

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.06.080200; this version posted June 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.



Supplementary material 3: The list of organisms with the longest and shortest lifespan in particular phylogenetic groups. Total number of analyzed animals for each group is depicted in column "Count". Maximal lifespans in years are in the brackets.

Supplementary material 4: Multiple sequence alignment of analyzed p53 sequences

Supplementary material 5: Nuclear localization signals prediction for all analyzed p53 sequences.

Supplementary material 6: Transcriptome assemblies of *Proteus anguinus* (SRX2382497) and *Sphenodon punctatus* (SRX4014663)

Supplementary material 7: Predicted structures of p53 homologs in PDB format

References

Afgan, E., Baker, D., Batut, B., Van Den Beek, M., Bouvier, D., Čech, M., Chilton, J., Clements, D., Coraor, N., & Grüning, B. A. (2018). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Research*, *46*(W1), W537–W544.

Anisimova, M., & Gascuel, O. (2006). Approximate likelihood-ratio test for branches: A fast, accurate, and powerful alternative. *Systematic Biology*, *55*(4), 539–552.

Arum, O., & Johnson, T. E. (2007). Reduced Expression of the Caenorhabditis elegans p53 Ortholog cep-1 Results in Increased Longevity. *The Journals of Gerontology: Series A*, *62*(9), 951–959. https://doi.org/10.1093/gerona/62.9.951

Aubrey, B. J., Kelly, G. L., Janic, A., Herold, M. J., & Strasser, A. (2018). How does p53 induce apoptosis and how does this relate to p53-mediated tumour suppression? *Cell Death & Differentiation*, *25*(1), 104–113.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.06.080200; this version posted June 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Barbi, E., Lagona, F., Marsili, M., Vaupel, J. W., & Wachter, K. W. (2018). The plateau of human mortality: Demography of longevity pioneers. *Science (New York, N.Y.)*, *360*(6396), 1459–1461. https://doi.org/10.1126/science.aat3119

Bartas, M., Brázda, V., Červeň, J., & Pečinka, P. (2020). Characterization of p53 Family Homologs in Evolutionary Remote Branches of Holozoa. *International Journal of Molecular Sciences*, *21*(1), 6. https://doi.org/10.3390/ijms21010006

Bauer, J. H., Poon, P. C., Glatt-Deeley, H., Abrams, J. M., & Helfand, S. L. (2005). Neuronal expression of p53 dominant-negative proteins in adult Drosophila melanogaster extends life span. *Current Biology*, *15*(22), 2063–2068.

Belyi, V. A., & Levine, A. J. (2009). One billion years of p53/p63/p73 evolution. *Proceedings of the National Academy of Sciences*, *106*(42), 17609–17610. https://doi.org/10.1073/pnas.0910634106

Bennett, W. P., Hussain, S. P., Vahakangas, K. H., Khan, M. A., Shields, P. G., & Harris, C. C. (1999). Molecular epidemiology of human cancer risk: Gene–environment interactions and p53 mutation spectrum in human lung cancer. *The Journal of Pathology*, *187*(1), 8–18.

Bonafè, M., Olivieri, F., Mari, D., Baggio, G., Mattace, R., Sansoni, P., De Benedictis, G., De Luca, M., Bertolini, S., & Barbi, C. (1999). P53 variants predisposing to cancer are present in healthy centenarians. *American Journal of Human Genetics*, *64*(1), 292.

Brameier, M., Krings, A., & MacCallum, R. M. (2007). NucPred—Predicting nuclear localization of proteins. *Bioinformatics (Oxford, England)*, *23*(9), 1159–1160. https://doi.org/10.1093/bioinformatics/btm066

Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., & Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*, *68*(6), 394–424.

Brázda, V., & Fojta, M. (2019). The Rich World of p53 DNA Binding Targets: The Role of DNA Structure. *International Journal of Molecular Sciences*, *20*(22), 5605. https://doi.org/10.3390/ijms20225605

Buslje, C. M., Santos, J., Delfino, J. M., & Nielsen, M. (2009). Correction for phylogeny, small number of observations and data redundancy improves the identification of coevolving amino acid pairs using mutual information. *Bioinformatics*, *25*(9), 1125–1131. https://doi.org/10.1093/bioinformatics/btp135

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.06.080200; this version posted June 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Campisi, J. (2005). Senescent cells, tumor suppression, and organismal aging: Good citizens, bad neighbors. *Cell*, *120*(4), 513–522.

Castresana, J. (2000). Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular Biology and Evolution*, *17*(4), 540–552.

Chen, J. (2016). The cell-cycle arrest and apoptotic functions of p53 in tumor initiation and progression. *Cold Spring Harbor Perspectives in Medicine*, *6*(3), a026104.

Chevenet, F., Brun, C., Bañuls, A.-L., Jacq, B., & Christen, R. (2006). TreeDyn: Towards dynamic graphics and annotations for analyses of trees. *BMC Bioinformatics*, *7*(1), 439.

Choi, Y., & Chan, A. P. (2015). PROVEAN web server: A tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics*, *31*(16), 2745–2747.

Choi, Y., Sims, G. E., Murphy, S., Miller, J. R., & Chan, A. P. (2012). Predicting the functional effect of amino acid substitutions and indels. *PloS One*, *7*(10), e46688.

de Keizer*, P. L. J., Laberge*, R.-M., & Campisi, J. (2010). p53: Pro-aging or pro-longevity? *Aging (Albany NY)*, *2*(7), 377–379.

De Magalhaes, J. P., & Costa, J. (2009). A database of vertebrate longevity records and their relation to other life-history traits. *Journal of Evolutionary Biology*, *22*(8), 1770–1774.

Dereeper, A., Audic, S., Claverie, J.-M., & Blanc, G. (2010). BLAST-EXPLORER helps you building datasets for phylogenetic analysis. *BMC Evolutionary Biology*, *10*(1), 8.

Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet, F., Dufayard, J.-F., Guindon, S., Lefort, V., & Lescot, M. (2008). Phylogeny. fr: Robust phylogenetic analysis for the non-specialist. *Nucleic Acids Research*, *36*(suppl_2), W465–W469.

Deuter, R., & Müller, O. (1998). Detection of APC mutations in stool DNA of patients with colorectal cancer by HD-PCR. *Human Mutation*, *11*(1), 84–89.

Edgar, R. C. (2004). MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, *32*(5), 1792–1797.

El-Deiry, W. S., Kern, S. E., Pietenpol, J. A., Kinzler, K. W., & Vogelstein, B. (1992). Definition of a consensus binding site for p53. *Nature Genetics*, *1*(1), 45–49. https://doi.org/10.1038/ng0492-45

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.06.080200; this version posted June 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Emamzadah, S., Tropia, L., Vincenti, I., Falquet, B., & Halazonetis, T. D. (2014). Reversal of the DNA-Binding-Induced Loop L1 Conformational Switch in an Engineered Human p53 Protein. *Journal of Molecular Biology*, *426*(4), 936–944. https://doi.org/10.1016/j.jmb.2013.12.020

García-Cao, I., García-Cao, M., Martín-Caballero, J., Criado, L. M., Klatt, P., Flores, J. M., Weill, J.-C., Blasco, M. A., & Serrano, M. (2002). 'Super p53'mice exhibit enhanced DNA damage response, are tumor resistant and age normally. *The EMBO Journal*, *21*(22), 6225–6235.

Giacomelli, A. O., Yang, X., Lintner, R. E., McFarland, J. M., Duby, M., Kim, J., Howard, T. P., Takeda, D. Y., Ly, S. H., & Kim, E. (2018). Mutational processes shape the landscape of TP53 mutations in human cancer. *Nature Genetics*, *50*(10), 1381.

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., & Zeng, Q. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, *29*(7), 644.

Guindon, S., & Gascuel, O. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology*, *52*(5), 696–704.

Hafner, A., Bulyk, M. L., Jambhekar, A., & Lahav, G. (2019). The multiple mechanisms that regulate p53 activity and cell fate. *Nature Reviews Molecular Cell Biology*, *20*(4), 199–210.

Harraan, D. (1955). *Aging: A theory based on free radical and radiation chemistry*.

Hasegawa, Y., Watanabe, T., Takazawa, M., Ohara, O., & Kubota, S. (2016). De novo assembly of the transcriptome of Turritopsis, a jellyfish that repeatedly rejuvenates. *Zoological Science*, *33*(4), 366–372.

Herrero, J., Muffato, M., Beal, K., Fitzgerald, S., Gordon, L., Pignatelli, M., Vilella, A. J., Searle, S. M., Amode, R., & Brent, S. (2016). Ensembl comparative genomics resources. *Database*, *2016*.

Hughes, B. G., & Hekimi, S. (2017). Many possible maximum lifespan trajectories. *Nature*, *546*(7660), E8.

Itahana, K., Dimri, G., & Campisi, J. (2001). Regulation of cellular senescence by p53. *European Journal of Biochemistry*, *268*(10), 2784–2791.

Jessen, L. E., Hoof, I., Lund, O., & Nielsen, M. (2013). SigniSite: Identification of residue-level genotype-phenotype correlations in protein multiple sequence alignments. *Nucleic Acids Research*, *41*(W1), W286–W291.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.06.080200; this version posted June 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Kandoth, C., McLellan, M. D., Vandin, F., Ye, K., Niu, B., Lu, C., Xie, M., Zhang, Q., McMichael, J. F., & Wyczalkowski, M. A. (2013). Mutational landscape and significance across 12 major cancer types. *Nature*, *502*(7471), 333–339.

Keane, M., Semeiks, J., Webb, A. E., Li, Y. I., Quesada, V., Craig, T., Madsen, L. B., van Dam, S., Brawand, D., Marques, P. I., Michalak, P., Kang, L., Bhak, J., Yim, H.-S., Grishin, N. V., Nielsen, N. H., Heide-Jørgensen, M. P., Oziolor, E. M., Matson, C. W., … de Magalhães, J. P. (2015). Insights into the Evolution of Longevity from the Bowhead Whale Genome. *Cell Reports*, *10*(1), 112–122. https://doi.org/10.1016/j.celrep.2014.12.008

Kleiber, M. (1932). Body size and metabolism. *Hilgardia*, *6*(11), 315–353.

Kubota, S. (2011). *Repeating rejuvenation in Turritopsis, an immortal hydrozoan (Cnidaria, Hydrozoa)*.

Lane, D. P. (1992). Cancer. P53, guardian of the genome. *Nature*, *358*, 15–16.

Letunic, I., & Bork, P. (2019). Interactive Tree Of Life (iTOL) v4: Recent updates and new developments. *Nucleic Acids Research*, *47*(W1), W256–W259.

Levine, A. J. (1997). P53, the Cellular Gatekeeper for Growth and Division. *Cell*, *88*(3), 323–331. https://doi.org/10.1016/S0092-8674(00)81871-1

Levine, A. J., & Oren, M. (2009). The first 30 years of p53: Growing ever more complex. *Nature Reviews Cancer*, *9*(10), 749–758.

Maier, B., Gluba, W., Bernier, B., Turner, T., Mohammad, K., Guise, T., Sutherland, A., Thorner, M., & Scrable, H. (2004). Modulation of mammalian life span by the short isoform of p53. *Genes & Development*, *18*(3), 306–319. https://doi.org/10.1101/gad.1162404

Nicolai, S., Rossi, A., Di Daniele, N., Melino, G., Annicchiarico-Petruzzelli, M., & Raschellà, G. (2015). DNA repair and aging: The impact of the p53 family. *Aging (Albany NY)*, *7*(12), 1050.

Okonechnikov, K., Golosova, O., Fursov, M., & Team, U. (2012). Unipro UGENE: A unified bioinformatics toolkit. *Bioinformatics*, *28*(8), 1166–1167.

Passow, C. N., Bronikowski, A. M., Blackmon, H., Parsai, S., Schwartz, T. S., & McGaugh, S. E. (2019). Contrasting Patterns of Rapid Molecular Evolution within the p53 Network across Mammal and Sauropsid Lineages. *Genome Biology and Evolution*, *11*(3), 629–643. https://doi.org/10.1093/gbe/evy273

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.06.080200; this version posted June 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Pellegata, N. S., Sessa, F., Renault, B., Bonato, M., Leone, B. E., Solcia, E., & Ranzani, G. N. (1994). K-ras and p53 gene mutations in pancreatic cancer: Ductal and nonductal tumors progress through different genetic lesions. *Cancer Research*, *54*(6), 1556–1560.

Petitjean, A., Mathe, E., Kato, S., Ishioka, C., Tavtigian, S. V., Hainaut, P., & Olivier, M. (2007). Impact of mutant p53 functional properties on TP53 mutation patterns and tumor phenotype: Lessons from recent developments in the IARC TP53 database. *Human Mutation*, *28*(6), 622–629.

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF Chimera—A visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, *25*(13), 1605–1612.

Pfaff, M. J., Mukhopadhyay, S., Hoofnagle, M., Chabasse, C., & Sarkar, R. (2018). Tumor suppressor protein p53 negatively regulates ischemia-induced angiogenesis and arteriogenesis. *Journal of Vascular Surgery*, *68*(6), 222S–233S.

Rufini, A., Tucci, P., Celardo, I., & Melino, G. (2013). Senescence and aging: The critical roles of p53. *Oncogene*, *32*(43), 5129.

Sabapathy, K., & Lane, D. P. (2018). Therapeutic targeting of p53: All mutants are equal, but some mutants are more equal than others. *Nature Reviews Clinical Oncology*, *15*(1), 13.

Sahin, E., & DePinho, R. A. (2012). Axis of ageing: Telomeres, p53 and mitochondria. *Nature Reviews Molecular Cell Biology*, *13*(6), 397–404. https://doi.org/10.1038/nrm3352

Schmidt-Kastner, P. K., Jardine, K., Cormier, M., & McBurney, M. W. (1998). Absence of p53-dependent cell cycle regulation in pluripotent mouse cell lines. *Oncogene*, *16*(23), 3003–3011.

Simonetti, F. L., Teppa, E., Chernomoretz, A., Nielsen, M., & Marino Buslje, C. (2013). MISTIC: Mutual information server to infer coevolution. *Nucleic Acids Research*, *41*(W1), W8–W14. https://doi.org/10.1093/nar/gkt427

Sohal, R. S., & Weindruch, R. (1996). Oxidative stress, caloric restriction, and aging. *Science*, *273*(5271), 59–63.

Soussi, T., & Wiman, K. G. (2015). TP53: An oncogene in disguise. *Cell Death & Differentiation*, *22*(8), 1239–1249.

Storci, G., Carolis, S. D., Papi, A., Bacalini, M. G., Gensous, N., Marasco, E., Tesei, A., Fabbri, F., Arienti, C., Zanoni, M., Sarnelli, A., Santi, S., Olivieri, F., Mensà, E., Latini, S., Ferracin, M., Salvioli, S., Garagnani, P.,

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.06.080200; this version posted June 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Franceschi, C., & Bonafè, M. (2019). Genomic stability, anti-inflammatory phenotype, and up-regulation of the RNAseH2 in cells from centenarians. *Cell Death & Differentiation*, *26*(9), 1845–1858. https://doi.org/10.1038/s41418-018-0255-8

Sulak, M., Fong, L., Mika, K., Chigurupati, S., Yon, L., Mongan, N. P., Emes, R. D., & Lynch, V. J. (2016). TP53 copy number expansion is associated with the evolution of increased body size and an enhanced DNA damage response in elephants. *Elife*, *5*, e11994.

Timmers, P.R., Mounier, N., Lall, K., Fischer, K., Ning, Z., Feng, X., Bretherick, A.D., Clark, D.W., eQTLGen Consortium, Shen, X., Esko, T., Kutalik, Z., Wilson, J.F., Joshi, P.K. (2019). Genomics of 1 million parent lifespans implicates novel pathways and common diseases and distinguishes survival chances. *Elife*, *8*(e39856). doi:10.7554/eLife.39856

Timmis, A., Townsend, N., Gale, C., Grobbee, R., Maniadakis, N., Flather, M., Wilkins, E., Wright, L., Vos, R., & Bax, J. (2018). European Society of Cardiology: Cardiovascular disease statistics 2017. *European Heart Journal*, *39*(7), 508–579.

Toufektchan, E., & Toledo, F. (2018). The guardian of the genome revisited: P53 downregulates genes required for telomere maintenance, DNA repair, and centromere structure. *Cancers*, *10*(5), 135.

Tyner, S. D., Venkatachalam, S., Choi, J., Jones, S., Ghebranious, N., Igelmann, H., Lu, X., Soron, G., Cooper, B., Brayton, C., Park, S. H., Thompson, T., Karsenty, G., Bradley, A., & Donehower, L. A. (2002). P53 mutant mice that display early ageing-associated phenotypes. *Nature*, *415*(6867), 45–53. https://doi.org/10.1038/415045a

van Heemst, D., Mooijaart, S. P., Beekman, M., Schreuder, J., de Craen, A. J. M., Brandt, B. W., Eline Slagboom, P., & Westendorp, R. G. J. (2005). Variation in the human TP53 gene affects old age survival and cancer mortality1. *Experimental Gerontology*, *40*(1), 11–15. https://doi.org/10.1016/j.exger.2004.10.001

Vousden, K. H., & Lane, D. P. (2007). P53 in health and disease. *Nature Reviews. Molecular Cell Biology*, *8*(4), 275–283. https://doi.org/10.1038/nrm2147

Vyas, P., Beno, I., Xi, Z., Stein, Y., Golovenko, D., Kessler, N., Rotter, V., Shakked, Z., & Haran, T. E. (2017). Diverse p53/DNA binding modes expand the repertoire of p53 response elements. *Proc Natl Acad Sci*, *114*(40), 10624–10629. https://doi.org/10.1073/pnas.1618005114

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.06.080200; this version posted June 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F. T., de Beer, T. A. P., Rempfer, C., Bordoli, L., Lepore, R., & Schwede, T. (2018). SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Research*, *46*(W1), W296–W303. https://doi.org/10.1093/nar/gky427

Whittemore, K., Vera, E., Martínez-Nevado, E., Sanpera, C., & Blasco, M. A. (2019). Telomere shortening rate predicts species life span. *Proceedings of the National Academy of Sciences*, *116*(30), 15122–15127.

Wilkinson, G. S., & South, J. M. (2002). Life history, ecology and longevity in bats. *Aging Cell*, *1*(2), 124–131.

Zhao, Y., Wu, L., Yue, X., Zhang, C., Wang, J., Li, J., Sun, X., Zhu, Y., Feng, Z., & Hu, W. (2018). A polymorphism in the tumor suppressor p53 affects aging and longevity in mouse models. *Elife*, *7*, e34701.