# scientific reports

# OPEN



# Maximum lifespan and brain size in mammals are associated with gene family size expansion related to immune system functions

Huseyin Kilili<sup>1,14</sup>, Benjamin Padilla-Morales<sup>1,14</sup>, Atahualpa Castillo-Morales<sup>3</sup>, Jimena Monzón-Sandoval<sup>3</sup>, Karina Díaz-Barba<sup>1,2,4</sup>, Paola Cornejo-Paramo<sup>1,4</sup>, Orsolya Vincze<sup>5,6,7</sup>, Mathieu Giraudeau<sup>5</sup>, Stephen J. Bush<sup>8</sup>, Zhidan Li<sup>9</sup>, Lu Chen<sup>9</sup>, Evangelos Mourkas<sup>10</sup>, Sergio Ancona<sup>2</sup>, Alejandro Gonzalez-Voyer<sup>2</sup>, Diego Cortez<sup>11</sup>, Humberto Gutierrez<sup>12</sup>, Tamás Székely<sup>1,13</sup>, Alín P. Acuña-Alonzo<sup>2</sup> & Araxi O. Urrutia<sup>1,2</sup>

Mammals exhibit an unusual variation in their maximum lifespan potential, measured as the longest recorded longevity of any individual in a species. Evidence suggests that lifespan increases follow expansion in brain size relative to body mass. Here, we found significant gene family size expansions associated with maximum lifespan potential and relative brain size but not in gestation time, age of sexual maturity, and body mass in 46 mammalian species. Extended lifespan is associated with expanding gene families enriched in immune system functions. Our results suggest an association between gene duplication in immune-related gene families and the evolution of longer lifespans in mammals. These findings explore the genomic features linked with the evolution of lifespan in mammals and its association with life story and morphological traits.

Keywords Longevity, Ageing, Immunity, Gene duplication, Gene expression, Gene family size

Mammals exhibit high diversity in their maximum lifespan potential (MLSP, the age at death (longevity) of the longest-lived individual ever recorded in a species), ranging from less than a year in some shrew species to over a hundred years in humans and up to two hundred in bowhead whales<sup>1</sup>. Even in captivity lifespan variation persists, pointing to intrinsic biological factors limiting an individual's longevity<sup>2,3</sup>. Unlike average lifespan, which reflects both intrinsic and extrinsic factors such as the risk of predation and resource availability, MLSP is assumed to reflect a species' inherent longevity limit and is widely available used in comparative studies focused on life history trade-offs and the genomic determinants of longevity<sup>4–10</sup>.

Chiroptera (bats), the second-largest mammalian order, represent an important model to study longevity and ageing<sup>11</sup>. Bats exhibit extended lifespans relative to body size<sup>12,13</sup>, disease resistance to ageing-related illnesses such as cancer<sup>14</sup>, viral infections<sup>15</sup>. Similarly, genomic analyses of the bowhead whale (MLSP larger than 200 years) have identified changes in genes related to DNA repair, cell-cycle regulation, cancer and ageing<sup>16</sup>. Other

<sup>1</sup>Milner Centre for Evolution, Department of Life Sciences, University of Bath, Bath BA2 7AY, UK. <sup>2</sup>Instituto de Ecología, Universidad Nacional Autónoma de México, Ciudad de México 04510, Mexico. <sup>3</sup>UK Dementia Research Institute, Cardiff University, CF24 4HQ Cardiff, UK. <sup>4</sup>Licenciatura en Ciencias Genómicas, Universidad Nacional Autónoma de México, CP62210 Cuernavaca, Mexico. <sup>5</sup>Littoral, Environnement et Sociétés (LIENSs), UMR 7266 CNRS-La Rochelle Université, 2 Rue Olympe de Gouges, FR-17000 La Rochelle, France. <sup>6</sup>Institute of Aquatic Ecology, Centre for Ecological Research, 4026 Debrecen, Hungary. <sup>7</sup>Evolutionary Ecology Group, Hungarian Department of Biology and Ecology, Babeş-Bolyai University, 400006 Cluj-Napoca, Romania. <sup>8</sup>School of Automation Science and Engineering, Xi'an Jiaotong University, Xi'an, China. <sup>9</sup>Key Laboratory of Birth Defects and Related Diseases of Women and Children of MOE, Department of Laboratory Medicine, State Key Laboratory of Biotherapy, West China Second Hospital, Sichuan University, 610041 Chengdu, China. <sup>10</sup>Zoonosis Science Centre, Department of Medical Sciences, Uppsala University, Uppsala, Sweden. <sup>11</sup>Centro de Ciencias Genómicas, Universidad Nacional Autónoma de México, CP62210 Cuernavaca, México. <sup>12</sup>Instituto Nacional de Medicina Genomica, 14610 Ciudad de Mexico, Mexico. <sup>13</sup>Department of Evolutionary Zoology and Human Biology, University of Debrecen, Debrecen, Hungary. <sup>14</sup>Huseyin Kilili and Benjamin Padilla-Morales contributed equally to this work. <sup>IM</sup>email: bpm29@bath.ac.uk; accuna@iecologia.unam.mx

long-lived mammals, such as elephants, show expansion in gene families associated with DNA repair and tumour suppression, including *TP53*<sup>17</sup>.

MLSP is thought to reflect intrinsic differences in the molecular machinery governing the ability of organisms to cope with age-related cognitive and/or physical decline and vulnerability to disease<sup>18–24</sup>. Studies on the molecular basis of longevity, suggest that non-dividing differentiated cells (e.g., neurons) may be a limiting factor for longevity<sup>25,26</sup>. In humans, genes linked to post-mitotic cell longevity are enriched in pathways like cytoskeleton-dependent transport, tRNA metabolism, cell morphogenesis, and ribosome biogenesis<sup>25,26</sup>. These genes show reduced expression in neurodegenerative diseases and progeria, hinting at a protective role against ageing<sup>25,26</sup>. However, how these mechanisms have influenced MLSP evolution across species remains unclear.

Identifying the overarching genomic signatures associated with the evolution of MLSP can provide insights into the evolution of key life history traits and variations in longevity between individuals in a species. Comparative studies have linked MLSP variations to changes in gene expression profiles and differences at the genome sequence level, including protein rates of evolution and gene family size differences<sup>27–29</sup>. Genes associated with MLSP in these studies were enriched in DNA repair, defence response cell cycle and immunological process related terms<sup>28</sup>. Genes such as *PMS2* (DNA repair), *PNMA1* (cell fate determination), and *OGDHL* (ROS regulation) show positive correlation with MLSP across mammalian tissues. *BCL7B*, which inhibits carcinogenesis through Wnt pathway regulation, and *GATM*, associated with oxidative stress protection, are prominently linked to increased lifespan. These molecular signatures collectively enhance cellular maintenance and stress resistance mechanisms that appear critical for extended longevity<sup>30</sup>.

Gene family size evolution, driven by duplication and deletion events, has been suggested to play a key role in phenotypic evolution<sup>31-34</sup>. While many gene duplication events in some cases can result in pseudogenization and reversal to pre-duplication copy number, preserved duplicates can increase gene dosage<sup>35,36</sup>, or expand transcript and protein diversity<sup>37</sup>. Although overall gene number has remained stable over 800 million years of metazoan evolution, gene family expansions and contractions have phenotypic diversity<sup>38–43</sup>. Genome sequencing projects for long-lived mammals have uncovered gene duplication events, likely linked to increased longevity, such as in the bowhead whale<sup>16</sup> and naked mole rat<sup>44</sup>.

Life history and morphological traits influencing MLSP have yet to be considered in comparative genomic and transcriptomic studies. Body mass presents a positive correlation with MLSP<sup>45</sup>, possibly because higher vulnerability of smaller species to predation compared to larger ones and the prioritisation of reproduction over self-maintenance to maximise fitness<sup>46</sup>. Another explanation is the negative correlation between metabolic rates with body mass in endothermic animals<sup>47</sup>, as accelerated metabolism in small species ultimately results in a faster accumulation of molecular damage<sup>48,49</sup>. Gestation time<sup>23</sup> and age at sexual maturity<sup>50</sup> also show a positive association with MLSP, reflecting potential developmental constraints linked to monotocous large-body mass mammals with long life spans<sup>51</sup>. Brain size relative to body mass is another key correlate, as large species tend to have large MLSP<sup>52–54</sup>, potentially due to higher behavioural complexity and survival strategies<sup>55,56</sup>.

It is also important considering that larger species have lower effective population sizes, which may weaken purifying selection across the genome<sup>57</sup>, influencing genomic features such as transposable element content, intron and intergenic region length<sup>57–59</sup>. A study examining over 100 bird species genomes found evidence of lower purifying selection in larger species<sup>60</sup>, though some evidence suggests it remains effective even in species with an small effective population<sup>6</sup>. Thus, any genomic feature associated with MLSP must be assessed to exclude the confounding effects of reduced purifying selection in long-lived species. No prior study examining genomic correlates of MSLP has considered this factor.

Despite the growing body of research on longevity and MLSP genomics and transcriptomics<sup>61</sup>, gene lists identified in individual studies are rarely directly compared with previous findings. Existing lists include genes associated with MLSP across species and those linked to longevity variation within species. Comparisons have shown little overlap, suggesting that molecular mechanisms underlying evolutionary MLSP changes may differ from those governing species longevity differences. Performing such comparisons would enhance the interpretation of the results obtained and assess how different aspects of genome evolution in line with MLSP, such as gene expression changes, protein rates of evolution, and gene family expansions, relate to one another and within species genome-wide association studies in human leading to a more integrated view of this key phenotype.

Furthermore, several biological processes have been proposed to contribute to ageing and longevity, including DNA repair<sup>62</sup>, apoptosis<sup>63,64</sup>, immunity, inflammation<sup>65</sup>, autophagy<sup>66</sup>, oxidative stress<sup>67</sup> and epigenetic markers<sup>68</sup>. Interventions such as caloric restriction and compounds known as senolytics can modulate longevity, highlighting the importance of these pathways. Directly testing the enrichment of genes related to these processes could magnify our understanding of MLSP evolution.

Here, we use a comparative genomics approach to identify genomic signatures associated with the evolution of MLSP across mammals. We examine whether MLSP variations correlate with gene family sizes (of proteincoding genes) in 46 fully sequenced mammalian species, accounting for potential confounders such as body mass, brain size, and effective population size. Specifically, we consider the potential confounding effect of a generalised genome-wide gene family size increase in line with MLSP resulting from diminished purifying selection among longer-lived species. Furthermore, We conduct a genome-wide analysis not restricted to specific functional gene categories and analyse the overlap between MLSP-associated genes identified in this study with gene sets previously associated with human longevity and molecular processes relevant to ageing and longevity evolution.

# Results

# Significant association between MLSP and gene family expansion

We conducted a phylogenetic generalised least-square (PGLS) analysis, corrected by Benjamini-Hochberg<sup>69</sup>, to analyse the association between gene family size (dependent variable) and MLSP (independent variable). A total of 4,136 gene families in 46 fully sequenced mammalian species were included in the analysis (Fig. 1; supplementary Table 1). Our analysis calculated 236 statistically significant MLSP-associated families under expansion (p < 0.05; effect sizes ranging from r = 0.43 to 0.60; Fig. 2a), while one presented significant contraction (p < 0.05; effect size of r = -0.49; Fig. 2a). It has been suggested that several gene features proliferate in genomes of species with smaller effective population sizes. To rule out a possible generalised increase in protein-coding



**Fig. 1**. Phylogenetic distribution of life history and morphological traits (maximum lifespan potential, relative brain size, body mass, gestation time and age at sexual maturity) in mammals. A phylogenetic tree for 46 mammal species with fully sequenced genomes is shown. Bars show the relative value of each variable, including log10 values of maximum lifespan potential (MLSP, green), relative brain size (orange), body mass (purple), gestation time (pink), and age at sexual maturity (blue). See Supplementary Table 1 for raw data and silhouette species names.



b



**Fig. 2.** Significantly associated and functional annotation overrepresentation among gene families with size variations in mammals. Panel (**a**) shows the number of gene families expanding and contracting associated with MLSP; relative brain size; and MLSP corrected by relative brain size. (**b**) Gene ontology (GO) term enrichment analysis among the families significantly associated with several PGLS models. Coloured cells in each column represent significantly enriched GO categories after correction for multiple testing (Benjamini Hochberg). Colour intensity (towards purple) increases for smaller adjusted *p*-values. Individual columns show significantly enriched categories for (A) MLSP-associated families from PGLS model which has 46 species ( $r_{mLSP-associated families > 0, p < 0.05, n = 236$ ), (B) relative brain size-associated families from PGLS with 59 species ( $r_{relative brain size-associated families > 0, p < 0.05, n = 360$ ), (C) MLSP-associated families > 0, p < 0.05, n = 267), (D) relative brain size-associated families from PGLS model that contains MLSP and relative brain size ( $r_{relative brain size-associated families > 0, p < 0.05, n = 184$ ).

.....

gene number among species with higher MLSP, we tested the association between total protein-coding genes and MLSP. Our findings indicate no significant association (p > 0.05).

# Relative brain size, not body mass nor life history traits, influences gene family expansion associated with MLSP

We then tested whether body mass and relative brain size could explain the observed associations between gene family size and MLSP. In a set of 46 species with data available for all three variables, we observed a significant association between MLSP with both body mass (r=0.60; p<0.0001) and relative brain size (r=0.70; p<0.0001,

Supplementary Fig. 1). Gene family size analysis found 360 expanding gene families and 165 contracting gene families associated with relative brain size after Benjamini-Hochberg correction (Fig. 2a). For body mass, our analysis did not find significant associations between MLSP and gene family size.

Subsequently, in another set of 42 and 31 species, we examined whether MLSP correlated with gestation time and age at sexual maturity, respectively. We detected positive correlations with MLSP (single-predictor PGLS models: MLSP vs. gestation time; r=0.70; p<0.0001 and MLSP vs. age at sexual maturity; r=0.80; p<0.0001; Supplementary Fig. 1). We did not find significant associations between the models analysing gene family size and gestation time or age at sexual maturity.

To identify gene families associated with MLSP after accounting for the effect of relative brain size, we ran a two-predictor PGLS regression model that included both MLSP and relative brain size. We found 267 gene families presenting expansion while five were undergoing contraction (Fig. 2a). The five gene families undergoing contraction include genes involved in biological processes, such as blood clotting<sup>70</sup>, transcription regulation<sup>71</sup>, cytoskeleton organisation<sup>71</sup>, cholesterol trafficking<sup>71</sup>, and protein hydrolysis<sup>71</sup> (Supplementary Table 2). A total of 184 expanding gene families were associated with relative brain size. Of these, 161 gene families were significantly associated with both traits, consistent with a shared evolutionary path. The 267 gene families that present expansion include a total of 2061 genes. Henceforth, these will be referred to as MLSP-associated families and MLSP-associated genes, respectively. Similarly, the 184 gene families associated with relative brain size comprise 1673 genes. These will be referred to as relative brain size-associated genes.

Sensitivity test using a Leave-One-Out approach, assessing effect size and significant changes in the correlation structure indicate that most species have a negligible effect size (Cohen's d < 0.2) and non-significant variations in *r* values (p > 0.05). However, *Heterocephalus glaber* (Cohen's d = -0.22, Wilcoxon: p < 0.002) and *Homo sapiens* (Cohen's d = -0.59, Wilcoxon: p > 0.05) exhibit larger effect size, while *Gorilla gorilla* (Cohen's d = 0.13, Wilcoxon: p < 0.002) and *Loxodonta africana* (Cohen's d = 0.13, Wilcoxon: p < 0.002) show small but significant effect (Supplementary data 1). This suggests that our observations might not be solely driven by one species but can still be influenced by taxa with extreme values.

# Immune system functional annotations are enriched among MLSP-associated and relative brain size-associated genes

MLSP-associated genes showed a significant overrepresentation of immune system-related functional annotations (Fig. 2b; supplementary Table 4). No functional enrichment analysis was conducted, given the small number of MLSP-associated contracting gene families (n = 5). Relative brain size-associated genes show similar functional enrichment patterns to MLSP-associated genes, with significantly overrepresented immune system functions. On the other hand, contracting relative brain size-associated gene families were not enriched in any functional category.

# MLSP-associated genes present higher gene expression and alternative splicing in the human

The observed gene family expansions associated with the evolution of MLSP may respond to selective pressures on gene dosage<sup>35,36</sup> or selective pressures related to transcript diversity<sup>37</sup>. Thus, we next examined gene expression patterns and alternative splicing in MLSP-associated genes in humans. We found that MLSP-associated genes have a higher gene expression level and produce a higher number of unique transcripts compared to background genes (Wilcoxon rank sum test: *p* FPKM < 0.0001 and *p* for Transcript Number = 0.0022, respectively). No significant results were obtained when examining relative brain size-associated genes (Wilcoxon rank sum test: *p* FPKM > 0.05, respectively). However, since this analysis is based solely on human data, these results should be interpreted with caution, as gene expression and splicing patterns may differ across species and tissue types. Future studies incorporating cross-species transcriptomic data will be necessary to determine the broader evolutionary significance of these findings.

Genes previously associated with ageing and longevity are also enriched among MLSP-associated gene families. We assessed whether MLSP-associated genes significantly overlap with genes previously shown to (a) be associated with related molecular processes, (b) have age-dependent expression, (c) be manually curated associated with ageing or longevity, (d) targets of longevity modifying interventions and (e) be associated with lifespan and longevity within and between individuals and across species (Fig. 3). Among genes with functional annotations of molecular processes previously suggested to play important roles in ageing and longevity, MLSPassociated genes were significantly enriched in genes with DNA repair ( $X^2$  test p = 0.0028)<sup>72</sup> and inflammation<sup>73</sup> ( $X^2$  test p = 0.0002) functions with a significant underrepresentation of autophagy<sup>74</sup> ( $X^2$  test p = 0.0075) related genes. No significant over or under-representation was observed for genes related to oxidative stress<sup>73</sup>, epigenetic markers<sup>75</sup> and apoptosis<sup>76</sup>. Genes with age-dependent expression that increase in activity with age were found to be underrepresented among MLSP-associated genes for one database but not for a second one ("agedependent cellular expression"<sup>77</sup> and "age-dependent expression")<sup>27</sup>. Genes that decrease their expression with age obtained from the same sources were not over or under-represented among MLSP-associated genes. Two sets of manually curated gene lists for cell senescence<sup>29</sup> ( $X^2$  test p = 0.0024) and longevity<sup>78,79</sup> were found to be significantly underrepresented among MLSP-associated genes. Similarly, significant underrepresentation among MLSP-associated genes was observed for genes that suppress the life-extending effects of caloric restriction<sup>80</sup>  $(X^2 \text{ test } p = 0.0058)$  and targets of life-extending drugs (senolytics)<sup>81-84</sup> ( $X^2 \text{ test } p = 0.0036$ ). Among gene lists correlating with cell longevity, individual longevity and species lifespan, significant over-representation among MLSP-associated genes ( $X^2$  test p < 0.0001) was observed for genes with human centenarians-associated genetic variants<sup>85</sup> and among genes with faster protein evolution in species with higher MLSP; no significant under



**Fig. 3.** Longevity-associated databases gene enrichment among MLSP-associated genes. Enrichment analysis of longevity-related databases among MLSP-associated genes. Bar lengths show the log2 values of the observed number of genes from individual datasets divided by the expected number of genes under the null hypothesis. Bars are colour-coded according to the type of gene lists collected. Stars denote a significant deviation from the null under a two-tail chi-square test.

of over-representation was observed for genes whose transcriptional profiles correlate with MLSP across mammalian species<sup>86</sup> or those associated with postmitotic cellular longevity<sup>25,26</sup>.

To further understand the functional relevance of MLSP-associated genes and how they compare to gene sets from previous studies, we conducted a GO enrichment analysis. Finding highly significant (p < 0.01) enrichment of immune system-related (e.g. innate immune response, immune response, adaptive immune response, inflammatory response) and longevity-related (e.g. DNA repair, negative regulation of apoptotic process, autophagy, DNA damage response, regulation of apoptotic process) functions among the MLSP-associated genes and genes presented in previous studies (Supplementary Fig. 2). Apoptosis, senescence, Ei-associated, human longevity associated variants, life extending drug targets and MLSP-associated genes share several biological functional annotations. In contrast, genes previously associated with age-dependent cellular expression, age dependent expression, cellular turnover, dietary restriction benefit suppressors, DNA repair and post-mitotic cell longevity show none or minimal overlap with the other gene sets. These results suggest that while MLSPassociated genes share common pathways with some immune and longevity-related processes discovered in previous studies, other longevity-related genes may represent distinct mechanisms of lifespan regulation not shared among the species analysed in this study.

## Discussion

We identified 236 gene families showing significant positive correlations with maximum lifespan potential across the mammalian phylogeny. Genes in these families are enriched in immune system functional annotations and among genes previously associated with ageing and longevity. Unlike earlier studies, our work accounts for potential confounding effects of morphology variables and life history traits when exploring genomic signatures associated with lifespan evolution.

Our findings are unlikely to be explained by smaller effective population sizes among species with the highest MLSP. Among mammals, species with the highest MLSP also tend to be larger, which correlates inversely with effective population size<sup>57–59</sup>. However, we found no significant correlation between MLSP and total protein-coding gene number. Furthermore, body mass, which correlates with effective population size<sup>87</sup>, was not associated with gene family size increases.

We uncovered evidence of a shared molecular machinery associated with the evolution of MLSP and relative brain size, with 161 gene families significantly related to both phenotypes when included in the same models. This common genomic signature aligns with the established evolutionary relationship between lifespan and relative brain size in vertebrates<sup>54,88</sup>. Conversely, no associations were found with gestation time, age at sexual maturity, or body mass.

Gene ontology analysis of MLSP-associated families revealed consistent enrichment of genes involved in "immune response", "defence response", "adaptive immune response" and "antigen processing and presentation via MHC class II" processes. The immune system has multiple mechanisms that could support longer lifespan: removing senescent cells (critical for neuron maintenance)<sup>89</sup> detecting and removing cells with uncontrolled cell division<sup>90</sup>infections<sup>91</sup>. These findings support the concept that the immune system plays an important role in determining lifespan and shaping its evolution across the animal kingdom, including humans<sup>92–95</sup>.

Interestingly, MLSP-associated families are significantly enriched in genes with human longevity-associated variants<sup>85</sup>. Functional enrichment analysis of these human longevity-associated genes<sup>85</sup> conducted in this study revealed immune system-related functions among the enriched gene ontology categories. While previous studies found little overlap between genes with MLSP-associated protein evolution rates and genes with longevity-associated variants in humans<sup>9</sup>, our enrichment comparing our MLSP-associated genes and multiple previous studies' gene sets shows functional convergence. This convergence is particularly evident in immune system regulations and apoptotic processes, indicating that despite identifying different gene sets, these studies highlight common underlying biological functions. However, a more recent study examining gene expression patterns across tissues for species with different MLSP did find some genes in common when examining age-associated genes within species<sup>10</sup>, suggesting a degree of commonality on specific aspects of MLSP and within species variation in longevity. These results suggest an overlap between molecular mechanisms associated with lifespan evolution across species and those involved in longevity differences within a species. Future studies directly testing overlap between genes identified within and between species comparisons should establish the extent of the common molecular pathways explaining evolutionary patterns and between individual differences.

Among biological processes previously associated with ageing and longevity, we found that genes associated with DNA repair and inflammation are significantly overrepresented among MLSP-associated genes, while autophagy-associated genes are underrepresented. This is noteworthy as autophagy has been found to remain stable in longevity model species like the naked mole rat<sup>96</sup>. Our functional enrichment analysis further reinforces these associations, revealing that MLSP-associated genes share significant functional enrichment with previously identified longevity-related processes, particularly in DNA damage response, inflammation regulation and apoptotic processes, despite limited overlap in the specific genes identified.

However, we did check the overlap between our MLSP-associated gene set and the candidate gene list published in a previous study<sup>16</sup>, and these genes were significantly underrepresented among MLSP-associated genes. This suggests that this manually curated gene list is potentially capturing a different aspect of longevity-associated molecular mechanisms not involved in driving the evolution of MLSP.

Alternative splicing has long been proposed as an alternative mechanism for expanding transcript diversity in addition to gene duplication<sup>37</sup>. Past studies have suggested an inverted correlation between gene family size and alternative splicing<sup>97–99</sup>, suggesting that the two mechanisms are, to some extent, equivalent<sup>100</sup>. We observed that MLSP-associated genes are more highly expressed and more alternatively spliced than background genes. These results may reflect an expansion of both gene expression and alternative splicing among MLSP-associated genes in line with MLSP evolution. However, in-depth comparative studies examining gene expression and alternative splicing across different species are needed to assess the selective pressures behind gene family expansions in line with MLSP evolution.

Our functional analyses reveal a meaningful picture, demonstrating that while direct gene list overlaps between studies may be limited, there is substantial convergence at the functional level. MLSP-associated genes share significant functional annotations with several gene sets (including apoptosis, senescence, human longevity-associated variants, and life-extending drug targets), particularly in immune system response and longevity-related processes. This functional convergence suggests that while different studies may identify distinct gene sets, they often highlight the same biological pathways, reinforcing the importance of these processes in longevity determination across phylogenetic lineages. The minimal functional overlap observed with certain gene sets (age-dependent expression, dietary restriction benefit suppressors, post-mitotic cell longevity) further suggests that some longevity mechanisms may represent distinct evolutionary pathways not universally shared among the species analysed.

In this study, we use MLSP as a measure of the intrinsic limit of longevity for individuals in each species. While widely used in comparative genomic and transcriptomic studies, MLSP is not a perfect index. The longestliving individual in a species might reflect an outlier with unique mutations absent from the reference genome, and in species with limited longevity records, MLSP may approximate average lifespan rather than longevity potential. Furthermore, whether records are taken from wild or captive individuals may influence evolution inferences. As demography-based indexes based on mortality risk by age become available, a clearer picture of the genomic basis of lifespan will emerge<sup>101</sup>. However, given the differences in MLSP among the species included in our study, we consider it is unlikely that the above-mentioned limitations would significantly impact our results.

Our findings are consistent with a scenario where gene family expansions have contributed to the evolution of longer lifespans in mammals. However, the associative nature of this study prevents establishing causality. We cannot rule out the possibility that gene family size variations are a by-product of longer lifespan evolution driven by unknown molecular mechanisms, or that other genomic changes unrelated to gene family size variations drive both phenotypic changes in MLSP.

#### Conclusions

In summary, our study conducts a comprehensive genome scan for signatures of gene family expansions in line with the evolution of longer maximum lifespans. Our results on gene family expansions associated with MLSP are robust, remaining significant even when accounting for potential confounding variables, including life story traits (gestation time, age at sexual maturity) and morphological traits (body mass). Notably, following

earlier evidence of MLSP being driven by the evolution of larger brains, we find evidence for a strong common genomic signal in gene family expansions for both phenotypes. Thus, we propose that relative brain size must be considered in future studies investigating genomic signatures of MLSP. Interestingly, we find significant and consistent enrichment of immune system genes among MLSP-associated genes. Additionally, we find that MLSP-associated families significantly overlap those associated with human longevity, suggesting shared molecular mechanisms between the evolution of MLSP across species and the variation in longevity in humans.

## Materials and methods

### Gene family size annotations and genome completeness

Annotated gene families and fasta sequences encompassing 92 fully sequenced mammalian genomes were obtained from Ensembl<sup>102</sup> using the *BiomaRt* R Package<sup>103</sup> (n = 17,722). To exclude lower-quality and incomplete genomes, which could bias the assessment of gene family sizes, we used the Benchmarking Universal Single-Copy Ortholog (BUSCO) tool<sup>104</sup> using the OrthoDB v10 database for mammals as benchmarking to assess genome completeness. A total of 46 species (supplementary Table 1) with genome completeness scores higher than 80% were considered further. Gene families with no variance in gene number across species were excluded from the analyses<sup>34</sup>. Gene families were also required to have at least three genes in one species. Finally, gene families must be present in at least 80% of species to rule out lineage-specific gene families. This resulted in 4,121 gene families<sup>34</sup>.

## Maximum lifespan potential, brain mass, body mass and relative brain size

Maximum lifespan potential (MLSP) and body mass (BM) estimates for 46 species (supplementary Table 1) with good-quality annotated genomes were obtained from the Animal Aging and Longevity Database build 14 <sup>79</sup>. Brain mass measurements were also obtained for 46 species from various sources (supplementary Table 1). Brain mass and body mass estimates were used to obtain relative brain size values, which quantify brain mass controlling for the allometric effect of body size, by calculating the residuals of a log-log least squares linear regression of brain mass against body mass as applied by<sup>34,54</sup> (see supplementary Tables 1 and references therein).

#### Correlation of maximum lifespan potential with other traits

Correlations between phenotype traits were assessed using independent contrast correlations, which allow assessing relatedness between two variables after correction for phylogenetic relatedness<sup>105</sup> using corrgram function in R.

#### Phylogenetic regressions of gene family size with life history and morphological traits

Ultrametric phylogeny of the 46 mammalian species with available good-quality genome and phenotypic data was obtained from Ensembl<sup>102</sup>. Phylogenetic tree and trait data were visualised using iTOL<sup>106</sup>. We used phylogenetic generalised least square regressions to assess the strength of associations between gene family sizes and life history or morphological traits across species. To rule out associations between phenotypic parameters and variations in gene family size being explained by shared ancestry, a phylogenetically generalised least squares regression (PGLS)<sup>107,108</sup> was used. Phylogenetically corrected regressions were performed using the "nlme" R package<sup>109</sup>, assuming a Brownian motion model of evolution and using the maximum likelihood method. Benjamini-Hochberg correction for multiple testing was implemented to identify gene families associated with the phenotypic traits examined<sup>69</sup>. Additionally, to evaluate the robustness of our findings and quantify the influence of individual species we conducted a Leave-One-Out sensitivity test by iteratively removing each species and re-estimating the two-predictor PGLS model. For each iteration we used statistically significant *t* values to calculate Cohen's d, using "effsize" R package<sup>110</sup>, to assess the effect size of each removed species. Later we conducted a Wilcoxon rank-sum test, using "stats" R package<sup>111</sup>, to detect significant deviations of *r* values from the original results.

#### Gene ontology term enrichment analysis

Biological process gene ontology (GO) functional terms annotations for each gene for each species were obtained from the Gene Ontology Consortium database<sup>112</sup>. GO terms were linked to a family whenever that term was assigned to any gene in the family in any of the 46 sequenced mammalian species available in Ensembl<sup>102</sup>. To prioritise GO terms with high association with gene families and increase the potential functional significance, we excluded GO terms annotated to less than 50 gene families from the analysis<sup>68</sup>. Enrichment of GO categories among the set of gene families associated with each trait of interest was carried out by measuring the proportion of families assigned to each GO term within the analysed set of gene families and comparing it with the proportion of gene families associated with each GO term in 1,000 equally sized samples of randomly chosen gene families from the background set. The mean and standard deviation of GO term representation measured in each 1000 random samples were taken to determine the corresponding p-values for each GO term using Z-scores with Benjamini-Hochberg correction for multiple testing, as implemented in<sup>34</sup>.

## Ageing and longevity database enrichment analysis

Lists of genes previously associated with longevity, post-mitotic cell maintenance, genes with age-dependent expression and genes associated with other molecular functions such as DNA repair, autophagy, immunity and inflammation, oxidative stress, epigenetic and apoptosis considered necessary for the regulation of ageing and longevity (Supplementary Table 3) were downloaded from various sources.

Gene numbers associated with each database are shown in Supplementary Table 3). DNA repair genes<sup>72</sup> were examined as DNA damage has been suggested to play an essential role in ageing<sup>113</sup>. In several studies, it has been found that long-lived organisms had increased DNA repair activity<sup>114-117</sup>. Inflammation is a natural immune

response to injury or infection. However, chronic inflammation can damage cells and tissues, which is thought to play a role in ageing<sup>65</sup>. Oxidative stress is caused by an imbalance between the production of reactive oxygen species (ROS) and the body's ability to detoxify them. ROS can damage biomolecules, including DNA, and they are thought to play a role in ageing<sup>67</sup>. Various factors, including ageing, environmental toxins, and lifestyle choices, can induce epigenetic markers in DNA and histone proteins. Epigenetic changes can contribute to ageing by altering gene expression in cell growth, metabolism, and repair<sup>68,118</sup>. Notably, specific categories of epigenetic data possess the capacity to exert a transgenerational impact on the longevity of progeny. These investigations yield several pivotal insights: instead of being genetically preordained, our life expectancy is predominantly under the sway of epigenetic determinants; dietary and other environmental factors have the potential to modify our life span by altering epigenetic data; and inhibitors targeting epigenetic enzymes can exert influence over the life expectancy of model organisms. These novel discoveries enhance our comprehension of the mechanisms underpinning the ageing process. Given the reversible nature of epigenetic data, these inquiries illuminate promising avenues for therapeutic interventions in ageing and age-related ailments, including cancer<sup>118</sup>. Apoptosis-related genes<sup>76</sup> were examined as apoptosis can delay ageing by removing damaged and senescent cells from tissue, but it can promote ageing when it eliminates irreplaceable post-mitotic cells<sup>95,119</sup>. Autophagy is a cellular process that degrades damaged or unnecessary cellular components. Autophagy decreases with age, contributing to the accumulation of cellular debris and the development of age-related diseases<sup>66</sup>.

Secondly, we examined the age-dependent expressed genes. Genes with age-dependent gene expression increased or decreased their activity in an age-dependent manner and were compiled from two different studies among MLSP-associated genes<sup>27,77</sup>. Genes with age-dependent transcription profiles may be markers of the ageing process and/or parts of the molecular mechanism to counteract the deleterious effects of the ageing process<sup>27,77</sup>.

Third, we examined two manually curated datasets—cell senescence-promoting genes and longevityassociated genes. Cell senescence-promoting genes<sup>29</sup> are thought to contribute to organism ageing. Longevityassociated genes are a database of ageing or longevity-related genes in humans and model organisms. This database is divided into pro and anti-longevity genes depending on whether they have been deemed to promote or hinder longevity<sup>78,79</sup>.

Fourth, we also examined longevity-modifying interventions. Dietary restriction benefit suppressor genes counteract the life-extending effects of caloric restriction<sup>80</sup>. It has been suggested that dietary restriction can boost longevity by reducing the intake of nutrients and delaying age-related degeneration<sup>120–122</sup>. Life-extending drug target genes were compiled from several studies<sup>81–84</sup>.

Finally, we examined genes associated with lifespan from several studies. "Human longevity-associated variants" is a dataset compiled from manually curated literature of cross-sectional and extreme longevity (centenarians) GWAS studies in healthy individuals<sup>85</sup>. Only genes with variants significantly associated with longevity in at least one human population were selected<sup>85</sup>. Longevity protein evolution genes were identified as those with a higher protein evolution rate associated with a longer species MLSP<sup>123</sup>). Maximum lifespan genes are an independent database of MLSP-associated genes identified from transcriptome comparative analysis in 26 mammalian species with varying MLSP<sup>86</sup>. Post-mitotic cell maintenance genes are those whose expression patterns correlate with the longevity of different cell types in a non-dividing state. These are assumed to be necessary for overall organism longevity<sup>25,26</sup>. Chi-squared tests were applied to assess if genes in each gene list were over or underrepresented among MLSP-associated genes. For those gene lists with an available background gene set, we restricted the analysis to genes present in both the latter background and our own background gene set of 4121 gene families encompassing 10,235 genes.

GO enrichment analysis was performed by comparing gene sets from previous studies at the functional level, using the background and focal gene sets specified in Supplementary Table 2. For datasets lacking background genes, we used the MLSP-associated genes background. The analysis followed the same framework as the gene ontology term enrichment analysis, except that GO terms annotated to fewer than 200 genes were excluded.

#### Alternative splicing analysis

Data for alternative splicing, comprising the number of unique transcripts and expression level for each human protein-coding gene, was obtained from the MeDAS database<sup>124</sup>. Non-parametric tests (Wilcoxon rank sum test) were then used to assess differences in alternative splicing and gene expression in trait associated genes.

#### Data availability

The datasets used and/or analysed during the current study available from the corresponding author on reasonable request.

Received: 17 September 2024; Accepted: 15 April 2025 Published online: 29 April 2025

#### References

- 1. Stearns, S. C. The Evolution of Life Histories249 (Oxford University Press, 1992).
- 2. Tidière, M. et al. Comparative analyses of longevity and senescence reveal variable survival benefits of living in zoos across mammals. *Sci. Rep.* **6**, 1–7 (2016).
- 3. Lemaître, J. F. et al. Sex differences in adult lifespan and aging rates of mortality across wild mammals. *Proc. Natl. Acad. Sci.* 117, 8546–8553 (2020).
- de Magalhães, J. P., Costa, J. & Church, G. M. An analysis of the relationship between metabolism, developmental schedules, and longevity using phylogenetic independent contrasts. J. Gerontol. Biol. Sci. Med. Sci. 62, 149–160 (2007).
- Galtier, N., Blier, P. U. & Nabholz, B. Inverse relationship between longevity and evolutionary rate of mitochondrial proteins in mammals and birds. *Mitochondrion* 9, 51–57 (2009).

- 6. Galtier, N. Adaptive protein evolution in animals and the effective population size hypothesis. PLoS Genet. 12, e1005774 (2016).
- 7. Kowalczyk, A., Partha, R., Clark, N. L. & Chikina, M. Pan-mammalian analysis of molecular constraints underlying extended lifespan. *Elife* 9, e51089 (2020).
- 8. Tejada-Martinez, D. et al. Positive selection and enhancer evolution shaped lifespan and body mass in great apes. *Mol. Biol. Evol.* **39**, msab369 (2021).
- Farré, X. et al. Comparative analysis of mammal genomes unveils key genomic variability for human life span. Mol. Biol. Evol. 38, 4948–4961 (2021).
- 10. Tyshkovskiy, A. et al. Distinct longevity mechanisms across and within species and their association with aging. *Cell* 186, 2929–2949e20 (2023).
- 11. Cooper, L. N. et al. Bats as instructive animal models for studying longevity and aging. Ann. N Y Acad. Sci. 1541, 10–23 (2024).
- 12. Davies, K. T. J. et al. Molecular evolution of growth hormone and insulin-like growth factor 1 receptors in long-lived, smallbodied mammals. *Gene* **549**, 228–236 (2014).
- 13. Wilkinson, G. S. & Adams, D. M. Recurrent evolution of extreme longevity in bats. Biol. Lett. 15, 20180860 (2019).
- 14. Hua, R. et al. Experimental evidence for cancer resistance in a Bat species. Nat. Commun. 15, 1401 (2024).
- Baid, K., Irving, A. T., Jouvenet, N. & Banerjee, A. The translational potential of studying Bat immunity. Trends Immunol. 45, 188–197 (2024).
- 16. Keane, M. et al. Insights into the evolution of longevity from the bowhead Whale genome. Cell. Rep. 10, 112-122 (2015).
- 17. Chusyd, D. E. et al. Aging: what we can learn from elephants. Front. Aging. 2, 726714 (2021).
- 18. Kirkwood, T. L., Kapahi, P. & Shanley, D. P. Evolution, stress, and longevity. J. Anat. 197 Pt 4, 587–590 (2000).
- Harper, J. M., Salmon, A. B., Leiser, S. F., Galecki, A. T. & Miller, R. A. Skin-derived fibroblasts from long-lived species are resistant to some, but not all, lethal stresses and to the mitochondrial inhibitor rotenone. *Aging Cell.* 6, 1–13 (2007).
- 20. Schumacher, B., Garinis, G. A. & Hoeijmakers, J. H. Age to survive: DNA damage and aging. Trends Genet. 24, 77-85 (2008).
- Perez, V. I. et al. Protein stability and resistance to oxidative stress are determinants of longevity in the longest-living rodent, the naked mole-rat. Proc. Natl. Acad. Sci. U. S. A. 106, 3059–3064 (2009).
- 22. Finch, C. E., Morgan, T. E., Longo, V. D. & de Magalhaes J. P. Cell resilience in species life spans: a link to inflammation? *Aging Cell.* 9, 519–526 (2010).
- Ricklefs, R. E. Life-history connections to rates of aging in terrestrial vertebrates. Proc. Natl. Acad. Sci. U. S. A. 107, 10314–10319 (2010).
- 24. Kourtis, N. & Tavernarakis, N. Cellular stress response pathways and ageing: intricate molecular relationships. *EMBO J.* 30, 2520–2531 (2011).
- Seim, I., Ma, S. & Gladyshev, V. N. Gene expression signatures of human cell and tissue longevity. Npj Aging Mech. Disease. 2, 1–8 (2016).
- Castillo-Morales, A., Monzón-Sandoval, J., Urrutia, A. O. & Gutiérrez, H. Postmitotic cell longevity-associated genes: a transcriptional signature of postmitotic maintenance in neural tissues. *Neurobiol. Aging.* 74, 147–160 (2019).
- de Magalhães, J. P., Curado, J. & Church, G. M. Meta-analysis of age-related gene expression profiles identifies common signatures of aging. *Bioinformatics* 25, 875–881 (2009).
- 28. Fushan, A. A. et al. Gene expression defines natural changes in mammalian lifespan. Aging Cell. 14, 352–365 (2015).
- Avelar, R. A. et al. A multidimensional systems biology analysis of cellular senescence in aging and disease. *Genome Biol.* 21, 1–22 (2020).
- Liu, W. et al. Large-scale across species transcriptomic analysis identifies genetic selection signatures associated with longevity in mammals. EMBO J. 42, e112740 (2023).
- Holland, P. W., Garcia-Fernàndez, J., Williams, N. A. & Sidow, A. Gene duplications and the origins of vertebrate development. Development 1994, 125–133 (1994).
- 32. Kaessmann, H. Origins, evolution, and phenotypic impact of new genes. Genome Res. 20, 1313-1326 (2010).
- Cavalier-Smith, T. Origin of animal multicellularity: precursors, causes, consequences—the choanoflagellate/sponge transition, neurogenesis and the cambrian explosion. *Philos. Trans. R Soc. Lond. B Biol. Sci.* 372, 20150476 (2017).
- 34. Padilla-Morales, B. et al. Sexual size dimorphism in mammals is associated with changes in the size of gene families related to brain development. *Nat. Commun.* **15**, 1–11 (2024).
- Rice, A. M. & McLysaght, A. Dosage sensitivity is a major determinant of human copy number variant pathogenicity. *Nat. Commun.* 8, 1–11 (2017).
- 36. Rice, A. M. & McLysaght, A. Dosage-sensitive genes in evolution and disease. BMC Biol. 15, 1-10 (2017).
- Bush, S. J., Chen, L., Tovar-Corona, J. M. & Urrutia, A. O. Alternative splicing and the evolution of phenotypic novelty. *Philos. Trans. R Soc. Lond. B Biol. Sci.* 372, 20150474 (2017).
- 38. Ashburner, M. et al. Gene ontology: tool for the unification of biology. The gene ontology consortium. *Nat. Genet.* **25**, 25–29 (2000).
- 39. Fortna, A. et al. Lineage-specific gene duplication and loss in human and great ape evolution. PLoS Biol. 2, E207 (2004).
- Hughes, A. L. & Friedman, R. Shedding genomic ballast: extensive parallel loss of ancestral gene families in animals. J. Mol. Evol. 59, 827–833 (2004).
- 41. Demuth, J. P., De Bie, T., Stajich, J. E., Cristianini, N. & Hahn, M. W. The evolution of mammalian gene families. *PLoS One.* 1, e85 (2006).
- 42. Hahn, M. W., Demuth, J. P. & Han, S. G. Accelerated rate of gene gain and loss in primates. Genetics 177, 1941–1949 (2007).
- 43. Hahn, M. W., Han, M. V. & Han, S. G. Gene family evolution across 12 Drosophila genomes. PLoS Genet. 3, e197 (2007).
- 44. Kim, E. B. et al. Genome sequencing reveals insights into physiology and longevity of the naked mole rat. *Nature* **479**, 223–227 (2011).
- 45. Speakman, J. R. Body size, energy metabolism and lifespan. J. Exp. Biol. 208, 1717-1730 (2005).
- Monarca, R. I., Speakman, J. R. & Mathias, M. L. Effects of predation risk on the body mass regulation of growing wood mice. J. Zool. 312, 122–132 (2020).
- Clarke, A., Rothery, P. & Isaac, N. J. B. Scaling of basal metabolic rate with body mass and temperature in mammals. *J. Anim. Ecol.* 79, 610–619 (2010).
- Bronikowski, A. & Vleck, D. Metabolism, body size and life span: a case study in evolutionarily divergent populations of the garter snake (Thamnophis elegans). Integr. Comp. Biol. 50, 880–887 (2010).
- 49. Flouris, A. D. & Piantoni, C. Links between thermoregulation and aging in endotherms and ectotherms. *Temperature* **2**, 73–85 (2015).
- 50. Millar, J. S. & Zammuto, R. M. Life histories of mammals: an analysis of life tables. Ecology 64, 631-635 (1983).
- 51. Isler, K. & van Schaik, C. P. The expensive brain: a framework for explaining evolutionary changes in brain size. *J. Hum. Evol.* 57, 392–400 (2009).
- Allen, J. S., Bruss, J. & Damasio, H. The aging brain: the cognitive reserve hypothesis and hominid evolution. Am. J. Hum. Biol. 17, 673–689 (2005).
- Barrickman, N. L., Bastian, M. L., Isler, K. & van Schaik, C. P. Life history costs and benefits of encephalization: a comparative test using data from long-term studies of primates in the wild. J. Hum. Evol. 54, 568–590 (2008).
- 54. Gonzalez-Lagos, C., Sol, D. & Reader, S. M. Large-brained mammals live longer. J. Evol. Biol. 23, 1064–1074 (2010).

- Sol, D., Sayol, F., Ducatez, S. & Lefebvre, L. The life-history basis of behavioural innovations. *Philos. Trans. R Soc. Lond. B Biol. Sci.* 371, 20150187 (2016).
- 56. Rigby Dames, B. A. et al. Evolutionary and genomic perspectives of brain aging and neurodegenerative diseases. In *Progress in Brain Research* (eds. Calvey, T., Ekhtiari, H., Lau, C. I., O'Mara, S. & Santos, F) (Elsevier, 2023).
- Popadin, K., Polishchuk, L. V., Mamirova, L., Knorre, D. & Gunbin, K. Accumulation of slightly deleterious mutations in mitochondrial protein-coding genes of large versus small mammals. *Proc. Natl. Acad. Sci. U S A.* 104, 13390–13395 (2007).
- 58. Lynch, M. & Conery, J. S. The origins of genome complexity. Science 302, 1401-1404 (2003).
- 59. Urrutia, A. O., Ocaña, L. B. & Hurst, L. D. Do Alu repeats drive the evolution of the primate transcriptome? *Genome Biol.* 9, R25 (2008).
- 60. Wanders, K. et al. Polygamy and purifying selection in birds. Evolution 77, 276-288 (2022).
- 61. Teulière, J. et al. Network analyses unveil ageing-associated pathways evolutionarily conserved from fungi to animals. *Geroscience* **45**, 1059–1080 (2023).
- 62. Nijnik, A. et al. DNA repair is limiting for Haematopoietic stem cells during ageing. Nature 447, 686-690 (2007).
- 63. Zhang, Y. & Herman, B. Ageing and apoptosis. Mech. Ageing Dev. 123, 245-260 (2002).
- Salminen, A., Ojala, J. & Kaarniranta, K. Apoptosis and aging: increased resistance to apoptosis enhances the aging process. *Cell. Mol. Life Sci.* 68, 1021–1031 (2011).
- 65. Finkel, T. & Holbrook, N. J. Oxidants, oxidative stress and the biology of ageing. Nature 408, 239-247 (2000).
- 66. He, L. Q., Lu, J. H. & Yue, Z. Y. Autophagy in ageing and ageing-associated diseases. Acta Pharmacol. Sin. 34, 605-611 (2013).
- 67. Beckman, K. B. & Ames, B. N. The free radical theory of aging matures. Physiol. Rev. 78, 547-581 (1998).
- Castillo-Fernandez, J. E., Spector, T. D. & Bell, J. T. Epigenetics of discordant monozygotic twins: implications for disease. Genome Med. 6, 60 (2014).
- Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. J. R Stat. Soc. 57, 289–300 (1995).
- Martins, I. M. & Choupina, A. Protein-Glutamine Gamma-Glutamyltransferase. In *Encyclopedia of Signaling Molecules* (ed. Choi, S) 4253–4257 (Springer, Cham, 2018).
- 71. Thomas, P. D. et al. Making genome-scale phylogenetics accessible to all. Protein Sci. 31. PANTHER, 8–22 (2022).
- 72. Milanowska, K. et al. REPAIRtoire—a database of DNA repair pathways. Nucleic Acids Res. 39, D788-D792 (2010).
- Subramanian, A. et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. U S A. 102, 15545–15550 (2005).
- 74. Wang, N. N. et al. HAMdb: a database of human autophagy modulators with specific pathway and disease information. J. Cheminform. 10, 34 (2018).
- 75. Marakulina, D. et al. EpiFactors 2022: expansion and enhancement of a curated database of human epigenetic factors and complexes. *Nucleic Acids Res.* 51, D564–D570 (2023).
- Diez, J., Walter, D., Munoz-Pinedo, C. & Gabaldón, T. DeathBase: a database on structure, evolution and function of proteins involved in apoptosis and other forms of cell death. *Cell. Death Differ.* 17, 735–736 (2010).
- 77. Palmer, D., Fabris, F., Doherty, A. & Freitas, A. A. De Magalhães, J. P. Ageing transcriptome meta-analysis reveals similarities and differences between key mammalian tissues. *Aging* 13, 3313–3341 (2021).
- Doherty, A. Magalhães, J. P. Has gene duplication impacted the evolution of eutherian longevity? Aging Cell. 15, 978–980 (2016). de.
- 79. Tacutu, R. et al. Human ageing genomic resources: new and updated databases. Nucleic Acids Res. 46, D1083–D1090 (2018).
- 80. Wuttke, D. et al. Dissecting the gene network of dietary restriction to identify evolutionarily conserved pathways and new functional genes. (2012).
- 81. Fuhrmann-Stroissnigg, H. et al. Identification of HSP90 inhibitors as a novel class of senolytics. Nat. Commun. 8, 1-14 (2017).
- 82. Lozano-Torres, B. et al. The chemistry of senescence. Nat. Reviews Chem. 3, 426-441 (2019).
- Nogueira-Recalde, U. et al. Fibrates as drugs with senolytic and autophagic activity for osteoarthritis therapy. *EBioMedicine* 45, 588–605 (2019).
- Proshkina, E., Shaposhnikov, M. & Moskalev, A. Genome-Protecting compounds as potential geroprotectors. Int. J. Mol. Sci. 21, 4484 (2020).
- Budovsky, A. et al. LongevityMap: a database of human genetic variants associated with longevity. Trends Genet. 29, 559–560 (2013).
- Lu, J. Y. et al. Comparative transcriptomics reveals circadian and pluripotency networks as two pillars of longevity regulation. *Cell. Metab.* 34, 836–856e5 (2022).
- Brevet, M. & Lartillot, N. Reconstructing the history of variation in effective population size along phylogenies. *Genome Biol. Evol.* 13, evab150 (2021).
- Jiménez-Ortega, D., Kolm, N., Immler, S., Maklakov, A. A. & Gonzalez-Voyer, A. Long life evolves in large-brained bird lineages. Evolution 74, 2617–2628 (2020).
- Golde, T. E. & Miller, V. M. Proteinopathy-induced neuronal senescence: a hypothesis for brain failure in Alzheimer's and other neurodegenerative diseases. Alzheimers Res. Ther. 1, 1–12 (2009).
- Egloff, A. M., Vella, L. A. & Finn, O. J. Cyclin B1 and other cyclins as tumor antigens in immunosurveillance and immunotherapy of cancer. Cancer Res. 66, 6–9 (2006).
- 91. Janeway, C. A. Jr. How the immune system protects the host from infection. Microbes Infect. 3, 1167-1171 (2001).
- 92. Aspinall, R. Longevity and the immune response. *Biogerontology* 1, 273–278 (2000).
- 93. Candore, G. et al. Biology of longevity: role of the innate immune system. *Rejuvenation Res.* 9, 143–148 (2006).
- 94. Larbi, A. et al. Aging of the immune system as a prognostic factor for human longevity. *Physiology* 23, 64–74 (2008).
- Salvioli, S. et al. Immune system, cell senescence, aging and longevity-inflamm-aging reappraised. Curr. Pharm. Des. 19, 1675– 1679 (2013).
- Zhao, S. et al. High autophagy in the naked mole rat May play a significant role in maintaining good health. *Cell. Physiol. Biochem.* 33, 321–332 (2014).
- Kopelman, N. M., Lancet, D. & Yanai, I. Alternative splicing and gene duplication are inversely correlated evolutionary mechanisms. *Nat. Genet.* 37, 588–589 (2005).
- 98. Su, Z., Wang, J., Yu, J., Huang, X. & Gu, X. Evolution of alternative splicing after gene duplication. Genome Res. 16, 182–189 (2006).
- 99. Talavera, D., Vogel, C., Orozco, M., Teichmann, S. A. & de la Cruz, X. The (in)dependence of alternative splicing and gene duplication. *PLoS Comput. Biol.* **3**, e33 (2007).
- Roux, J. & Robinson-Rechavi, M. Age-dependent gain of alternative splice forms and biased duplication explain the relation between splicing and duplication. *Genome Res.* 21, 357–363 (2011).
- Ronget, V. & Gaillard, J. Assessing ageing patterns for comparative analyses of mortality curves: going beyond the use of maximum longevity. *Funct. Ecol.* 34, 65–75 (2020).
- 102. Yates, A. D. et al. Ensembl 2020. Nucleic Acids Res. 48, D682-d688 (2020).
- Durinck, S., Spellman, P. T., Birney, E. & Huber, W. Mapping identifiers for the integration of genomic datasets with the R/ Bioconductor package biomart. Nat. Protoc. 4, 1184 (2009).

- 104. Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V. & Zdobnov, E. M. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **31**, 3210–3212 (2015).
- Felsenstein, J. Phylogenies and the comparative method. Am. Nat. 125, 1–15 (1985).
  Letunic, I. & Bork, P. Interactive tree of life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Res. 49, W293–W296 (2021).
- 107. Grafen, A. The phylogenetic regression. Philos. Trans. R Soc. Lond. B Biol. Sci. 326, 119–157 (1989).
- 108. Grafen, A. The uniqueness of the phylogenetic regression. J. Theor. Biol. **156**, 405–423 (1992).
- 109. Pinheiro, P. O. openaccess.thecvf.com, Unsupervised domain adaptation with similarity learning. in *Proceedings of the IEEE conference on computer vision and pattern recognition* 8004-8013 (2018).
- 110. Torchiano, M. Effsize a package for efficient effect size computation. Zenodo https://doi.org/10.5281/ZENODO.1480624 (2016).
- 111. R Core Team. R: A Language and Environment for Statistical Computing. Preprint at (2023). https://www.R-project.org/
- 112. Mi, H., Muruganujan, A., Ebert, D., Huang, X. & Thomas, P. D. PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. *Nucleic Acids Res.* **47**, D419–D426 (2019).
- 113. Vijg, J. & Suh, Y. Genome instability and aging. Annu. Rev. Physiol. 75, 645-668 (2013).
- 114. Chevanne, M. et al. Oxidative DNA damage repair and Parp 1 and Parp 2 expression in Epstein-Barr virus-immortalized B lymphocyte cells from young subjects, old subjects, and centenarians. *Rejuvenation Res.* **10**, 191–204 (2007).
- 115. Humphreys, V. et al. Age-related increases in DNA repair and antioxidant protection: a comparison of the Boyd Orr cohort of elderly subjects with a younger population sample. *Age Ageing.* **36**, 521–526 (2007).
- Franzke, B., Neubauer, O. & Wagner, K. H. Super DNAging-New insights into DNA integrity, genome stability and telomeres in the oldest old. *Mutat. Res. - Rev. Mut Res.* 766, 48–57 (2015).
- 117. Kim, Y. J., Kim, H. S. & Seo, Y. R. Genomic Approach to Understand the Association of DNA Repair with Longevity and Healthy Aging Using Genomic Databases of Oldest-Old Population. *Oxid. Med. Cell. Longev.* 2984730 (2018). (2018).
- 118. Pal, S. & Tyler, J. K. Epigenetics and aging. Sci. Adv. 2, e1600584 (2016).
- 119. Freitas, A. A. & De Magalhães, J. P. A review and appraisal of the DNA damage theory of ageing. *Mutat. Research/Reviews Mutat. Res.* **728**, 12–22 (2011).
- Bishop, N. A. & Guarente, L. Genetic links between diet and lifespan: shared mechanisms from yeast to humans. *Nat. Rev. Genet.* 8, 835–844 (2007).
- 121. Willcox, B. J. et al. Caloric restriction, the traditional Okinawan diet, and healthy aging: the diet of the world's longest-lived people and its potential impact on morbidity and life span. Ann. N Y Acad. Sci. 1114, 434–455 (2007).
- 122. Fontana, L., Partridge, L. & Longo, V. D. Extending healthy life span-from yeast to humans. Science 328, 321-326 (2010).
- 123. Li, Y. & de Magalhaes, J. P. Accelerated protein evolution analysis reveals genes and pathways associated with the evolution of mammalian longevity. Age 35, 301–314 (2013).
- 124. Li, Z. et al. McDAS: a metazoan developmental alternative splicing database. Nucleic Acids Res. 49, D144-D150 (2021).

#### Acknowledgements

This study was supported by a TRNC Department of Higher Education and Foreign Affairs State Scholarship for Doctoral Studies, The Korner Travelling Fellowship Fund 2020 and Santander Mobility Award 2020 to HK; CONACyT PhD scholarships to APAA, ACM and JMS; a University of Bath fee PhD scholarship to BPM; a travel grant SEP-UNAM-FUNAM Programa de Capacitación en Métodos de investigación, the Korner Travelling Fellowship Fund and TELMEX foundation fellowship to KDB; a travel grant SEP-UNAM-FUNAM Programa de Capacitación en Métodos de Investigación/2016 and a special grant from the Undergraduate Programme on Genomic Sciences – UNAM to PCP; by the Agence Nationale de la Recherche (COVER ANR-23-CE02-0019) to MG and OV; by the Région Nouvelle-Aquitaine—Chaire d'excellence "Cancer et Biodiversité" to MG and OV; by the National Scientific Research Fund (OTKA K143421) to OV a Frontiers in Science CONACyT grant (No. FC-2016 /1682) and a Royal Society Newton Advanced Fellowship (no. NA160564) to TS, DC and AOU; a Frontiers in Science CONACyT (FC-2020 / 682142) to SA, AGV, DC, HG and AOU; a NERC grant (NE/P004121/1) and PAPPIT-DGAPA-UNAM grant (IA204020) to AOU.

#### **Author contributions**

Authors' contributions. HK, AOU, APAA conceived the study. HK, BPM, AOU, ACM, JMS, OV designed the analyses. KDB, PCP, LC, ZL, SJB collected data, BPM, HK, PCP, KDB conducted analyses. EM, DC, SA, TS, OV, MG, AGV provided expert knowledge of evolution and genomics topics. All authors contributed to the paper writing.

## Declarations

#### **Competing interests**

The authors declare no competing interests.

#### Additional information

**Supplementary Information** The online version contains supplementary material available at https://doi.org/1 0.1038/s41598-025-98786-3.

Correspondence and requests for materials should be addressed to B.P.-M., A.P.A.-A. or A.O.U.

Reprints and permissions information is available at www.nature.com/reprints.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2025