Biomedical Journal 47 (2024) 100654

Contents lists available at ScienceDirect

Biomedical Journal

journal homepage: www.elsevier.com/locate/bj

Review Article

Peto's paradox: Nature has used multiple strategies to keep cancer at bay while evolving long lifespans and large body masses. A systematic mini-review

Matteo Perillo [a,*], Alessia Silla [b], Angela Punzo [a,c], Cristiana Caliceti [a,c], Andres Kriete [d], Christian Sell [e], Antonello Lorenzini [a,c]

[a] Department of Biomedical and Neuromotor Sciences, University of Bologna, Italy
[b] Department for Life Quality Studies, University of Bologna, Italy
[c] National Institute of Biosystems and Biostructures INBB, Rome, Italy
[d] School of Biomedical Engineering, Science and Health Systems, Drexel University, Philadelphia, PA, USA
[e] Department of Biochemistry & Molecular Biology, Drexel University College of Medicine, Philadelphia, PA, USA

ARTICLE INFO

Keywords:
Peto's paradox
Longevity
Body mass
Hallmarks of cancer
Evolution
Mini-review

ABSTRACT

Comparative oncology is an understudied field of science. We are far from understanding the key mechanisms behind Peto's paradox, i.e., understanding how long-lived and large animals are not subject to a higher cancer burden despite the longer exposure time to mutations and the larger number of cells exposed.

In this work, we investigated the scientific evidence on such mechanisms through a systematic mini-review of the literature about the relation of longevity and/or large body mass with physiological, genetic, or environmental traits among mammalian species. More than forty thousand articles were retrieved from three repositories, and 383 of them were screened using an active-learning-based tool. Of those, 36 articles on longevity and 37 on body mass were selected for the review. Such articles were examined focusing on: number and type of species considered, statistical methods used, traits investigated, and observed relationship with longevity and/or body mass. Where applicable, the traits investigated were matched with one or more hallmarks of cancer.

We obtained a list of potential candidate traits to explain Peto's paradox related to replicative immortality, cell senescence, genome instability and mutations, proliferative signaling, growth suppression evasion, and cell resistance to death.

Our investigation suggests that different strategies have been followed to prevent cancer in large and long-lived species. The large number of papers retrieved emphasizes that more studies can be launched in the future, using more efficient analytical approaches to comprehensively evaluate the convergent biological mechanisms essential for acquiring longevity and large body mass without increasing cancer risk.

Introduction

The fact that larger and long-lived animals do not develop more cancer than smaller and short-lived species has been termed Peto's paradox based on the scientist who first described this discrepancy [1]. In fact, assuming that cancer is a cell-intrinsic phenomenon, considering cancer frequency in mice as the norm and hypothesizing that cell-specific probability of malignant transformation is equal between species, it would be impossible to evolve large sizes and long lifespans. For example, humans would not even reach adulthood without displaying multiple cancers [2]. Among mice, which have an average lifespan of approximately two years, cancer is indeed a major cause of death. On the other hand, it is extremely rare to observe cancer in a 2-year-old human despite the significantly larger number of cells compared to a mouse of the same age.

A recent large and detailed analysis of pathological records from multiple zoos has observed that cancer incidence among mammals is neither related to longevity nor body mass. This suggests that long-lived and large animals are protecting themself from the expected higher burden of cancer, confirming Peto's paradox [3]. Still, there is a lack of

Peer review under responsibility of Chang Gung University.
* Corresponding author. Department of Biomedical and Neuromotor Sciences, University of Bologna, Italy.
E-mail address: matteo.perillo2@unibo.it (M. Perillo).

https://doi.org/10.1016/j.bj.2023.100654
Received 2 July 2023; Accepted 10 August 2023
Available online 19 August 2023

Table 1
Query design of the 2 sub-reviews for the 3 repositories in which the search has been performed.

		Query	Other filters
Longevity	PubMed	((("specie*"[Title/Abstract] OR "mammals"[Title/Abstract]) AND ("longevity"[Title/Abstract] OR "lifespan"[Title/Abstract] OR "life span"[Title/Abstract] OR "life-span"[Title/Abstract])) NOT ("review"[Publication Type] OR "meta analysis"[Publication Type] OR "systematic review"[Publication Type] OR "letter"[Publication Type] OR "editorial"[Publication Type])) AND ((2005/8/1:2023/4/30[pdat]) AND (english[Filter]))	
	Scopus	TITLE-ABS ("longevity" OR "lifespan" OR "life span" OR "life-span") AND TITLE-ABS (mammals OR species) AND NOT TITLE-ABS-KEY ("systematic review" OR "meta-analysis" OR review) AND (EXCLUDE (DOCTYPE,"ch") OR EXCLUDE (DOCTYPE,"re") OR EXCLUDE (DOCTYPE,"no") OR EXCLUDE (DOCTYPE,"le") OR EXCLUDE (DOCTYPE,"dp") OR EXCLUDE (DOCTYPE,"er") OR EXCLUDE (DOCTYPE,"sh") OR EXCLUDE (DOCTYPE,"bk"))	DOCTYPE exclude: "ch", "re", "no", "le", "dp", "er", "sh", "bk"; PUBYEAR limit to: 2005–2023; LANGUAGE limit to: "English"
	Web of science (core collection)	((TI=(mammals OR species) OR AB=(mammals OR species)) AND (TI=(longevity OR lifespan OR "life span" OR "life-span") OR AB=(longevity OR lifespan OR "life span" OR "life-span"))) NOT (TI=("systematic review" OR "meta-analysis" OR review) OR AB=("systematic review" OR "meta-analysis" OR review) OR KP=("systematic review" OR "meta-analysis" OR review))	PUBLICATION TYPE: Article, Proceeding paper, Early access, Meeting abstract, Correction, Data paper, Reprint, Retracted publication; PUBLICATION YEAR: 2005–2023; LANGUAGES: English
Body mass	PubMed	((("specie*"[Title/Abstract] OR "mammals"[Title/Abstract]) AND ("body weight"[Title/Abstract] OR "body mass"[Title/Abstract])) NOT ("review"[Publication Type] OR "meta analysis"[Publication Type] OR "systematic review"[Publication Type] OR "letter"[Publication Type] OR "editorial"[Publication Type])) AND ((2005/8/1:2023/4/30[pdat]) AND (english[Filter]))	
	Scopus	TITLE-ABS ("body weight" OR "body mass") AND TITLE-ABS (mammals OR species) AND NOT TITLE-ABS-KEY ("systematic review" OR "meta-analysis" OR review) AND (EXCLUDE (DOCTYPE,"ch") OR EXCLUDE (DOCTYPE,"re") OR EXCLUDE (DOCTYPE,"no") OR EXCLUDE (DOCTYPE,"le") OR EXCLUDE (DOCTYPE,"dp") OR EXCLUDE (DOCTYPE,"er") OR EXCLUDE (DOCTYPE,"sh") OR EXCLUDE (DOCTYPE,"bk"))	DOCTYPE exclude: "ch", "re", "no", "le", "dp", "er", "sh", "bk"; PUBYEAR limit to: 2005–2023; LANGUAGE limit to: "English"
	Web of science (core collection)	((TI=(mammals OR species) OR AB=(mammals OR species)) AND (TI=("body weight" OR "body mass") OR AB=("body weight" OR "body mass"))) NOT (TI=("systematic review" OR "meta-analysis" OR review) OR AB=("systematic review" OR "meta-analysis" OR review) OR KP=("systematic review" OR "meta-analysis" OR review))	PUBLICATION TYPE: Article, Proceeding paper, Early access, Meeting abstract, Correction, Data paper, Reprint, Retracted publication; PUBLICATION YEAR: 2005–2023; LANGUAGES: English

studies investigating the reasons behind this observed pattern.

In mammals, body mass varies about 100-million-fold and lifespan 100-fold (e.g., *Suncus etruscus* ≈ 2.1 g, 3.2 years; *Balaenoptera musculus* ≈ 136 tons, 110 years). It appears that both large body mass and significant longevity are traits that have evolved independently multiple times during evolution (convergent evolution) [4], and it is reasonable that selective pressure has focused on similar molecular mechanisms during these processes in order to prevent cancer. These mechanisms could, in some cases, converge on related pathways and, in other cases, impinge on different ones. For example, the accumulation of cells with DNA mutations may be reduced by enhanced DNA repair or through the elimination of cells with DNA damage through apoptosis.

This mini-review offers a glimpse into the complexity of the cancer preventing strategies "chosen" by the evolutionary processes. The work is aimed at collecting and summarizing scientific knowledge concerning the traits which have evolved synergically with longevity and/or body mass among mammals. Given that Peto's paradox relates to cancer, we have evaluated the mechanisms involved in the context of the "Hallmarks of Cancer" as defined by Hanahan and Weinberg [5–7]. Using the wealth of published studies, this work provides an organized overview of the molecular mechanisms associated with the evolution of cancer resistance.

It is important to bear in mind that correlation does not mean causation, so the observed associations are not necessarily due to a causal relation between the trait of interest and cancer suppression. Nevertheless, having a good knowledge of these associations certainly represents a fundamental element to plan therapeutic strategies. Given the immensity of the evolutionary timescale, it is certainly wise to tap the accumulated "knowledge" of the evolutionary processes.

Methodology

This work consists of two "sub-reviews" conducted in parallel, one focusing on the traits related to longevity and one focusing on the traits related to body mass. Despite the small portion of the available literature considered, the work has been performed following the methodology of a systematic review. The compliance to the investigation and reporting methodology of a systematic review have been assessed using the

Table 2
Inclusion and exclusion criteria of the literature selection process of the two sub-reviews.

	Inclusion criteria	Exclusion criteria
Longevity	• Comparative biology study • Original study • At least 33% of the species under study are mammals • Correlation measured between longevity and any other determinant	• Study without inter-species comparison • Review and/or meta-analysis • Less than 33% of species under study are mammals • Longevity not considered as primary variable (i.e., not considered at all, or used as control variable)
Body mass	• Comparative biology study • Original study • At least 33% of the species under study are mammals • Correlation measured between body mass and any other determinant	• Study without inter-species comparison • Review and/or meta-analysis • Less than 33% of species under study are mammals • Body mass not considered as primary variable (i.e., not considered at all, or used as control variable)

PRISMA checklist (shown in Appendix, file 1) [8].

Search strategy and query design

The literature of interest has been searched for in PubMed, Scopus, and Web of Science (Core collection). For each repository, two ad-hoc queries have been designed, one for each sub-review. The design of the queries was aimed at finding papers with the following characteristics:

- The word "mammals" and/or the word "species" appear in the title and/or in the abstract;
- For the review about longevity: the word "longevity" and/or the word "lifespan" (or "life span", or "life-span") appear in the title and/or in the abstract;
- For the review about body mass: the expression "body weight" and/or "body mass" appear in the title and/or in the abstract;
- The paper has been published between August 2005 and April 2023;
- The paper is not a review, systematic review, or meta-analysis;
- The paper is available in English.

Table 1 shows the six built queries. The search was conducted on the 1st of May 2023, and for both sub-reviews a merging and deduplication procedure has been performed on the results of the three parallel queries, using the reference manager software EndNote (version X9.2).

Studies selection

A two-phase screening process was set up to identify the papers coherent with the research question. To pass such screening, the works had to be comparative studies about mammals measuring the interspecies correlation between species longevity/body mass and any other trait. The exact inclusion and exclusion criteria can be found in Table 2.

The first phase of the screening was based on titles and abstracts, and was performed using the software ASReview (v1.2, [9]), a tool which exploits active learning and natural language processing (NLP) to learn about the features of the articles complying with the inclusion criteria of a review. This allows to speed up the screening process, as the tool iteratively suggests to the reviewer the next article to screen, based on the features learned [10]. In this way, the majority of the papers of interest are identified by the reviewer in the first part of the title and abstract screening process, reducing the time needed to identify a large number of papers satisfying the inclusion criteria.

Two reviewers worked separately on the screening of articles about longevity and body mass, consulting each other in case of uncertainty regarding the selection of a specific study. Following the example of Brouwer et al. [11], a two-fold stopping rule for this phase of screening was set: screening is stopped if more than 80% of irrelevant papers in a batch of 100 papers are found, or when the screeners have performed 8 h of screening each.

In the second phase, a third screener examined the full text of the papers which passed the first phase to assess their compliance with the inclusion criteria. Moreover, if a study from one pool (e.g., longevity) had also analyzed the relation between some trait and the other characteristic of interest (e.g., body mass), the study was selected for both sub-reviews. As the automatic deduplication step was not totally effective, it was necessary to perform manual deduplication of the papers during the full text screening.

Data extraction & analysis

Those studies found to be compliant with all the inclusion criteria were selected for the review. Such studies contained one or more analyses, for example, comparing more traits with body mass/longevity, or comparing both body mass and longevity with the same trait(s), or performing the same analysis on different subsets of mammalian species.

The full text of each study was carefully analyzed, extracting the following information: type and number of species analyzed, type of data collected, trait(s) measured and compared with the longevity/body mass, metric used to measure the longevity/body mass of the species, statistical methodology used to quantify the relation between the trait and longevity/body mass, the resulting estimate, and its statistical significance. Moreover, when possible, the traits analyzed in the studies were matched with one or more hallmarks of cancer, as defined by Hanahan (2022) [5], in order to identify the traits whose evolution could explain the resistance to cancer of large and long-lived species, and to have a better understanding of the "evolutionary strategies" behind Peto's paradox. This task was independently carried out by three researchers.

Risk of bias assessment

As discussed by Speakman et al. [12], comparative studies on longevity and body mass can be seriously impaired by two issues:

- The well-known correlation between longevity and body mass, if not accounted for, can generate spurious correlations between one of the two quantities and a third variable of interest.
- The species under analysis are usually not independent in terms of biological features, as they have a phylogenetic relationship. For this reason, the standard statistical techniques to measure correlation among two or more variables could not be suitable when analyzing comparative data, and the phylogenetic relations among the species under analysis should always be taken into account in order to obtain valid results.

In this present work, the methodologies put into practice in the studies to deal with these issues were carefully reviewed in order to have an idea of the reliability of the observed results.

Clustering of original analyses

With the number of data points rapidly growing, visualizations become increasingly important to quickly survey trends in the data. Here, the findings reported in the reviewed studies were visually clustered through a hierarchical clustering algorithm, utilizing Pearson correlation as a distance measure and complete linkage criterion. The algorithm evaluates the statistical similarity among each pair of analyses (based on the Pearson r value) and constructs a dendrogram based on the linkage function, defining the compactness of clusters. Analyses cluster together when the indicators "Hallmark of Cancer", "Phylogeny

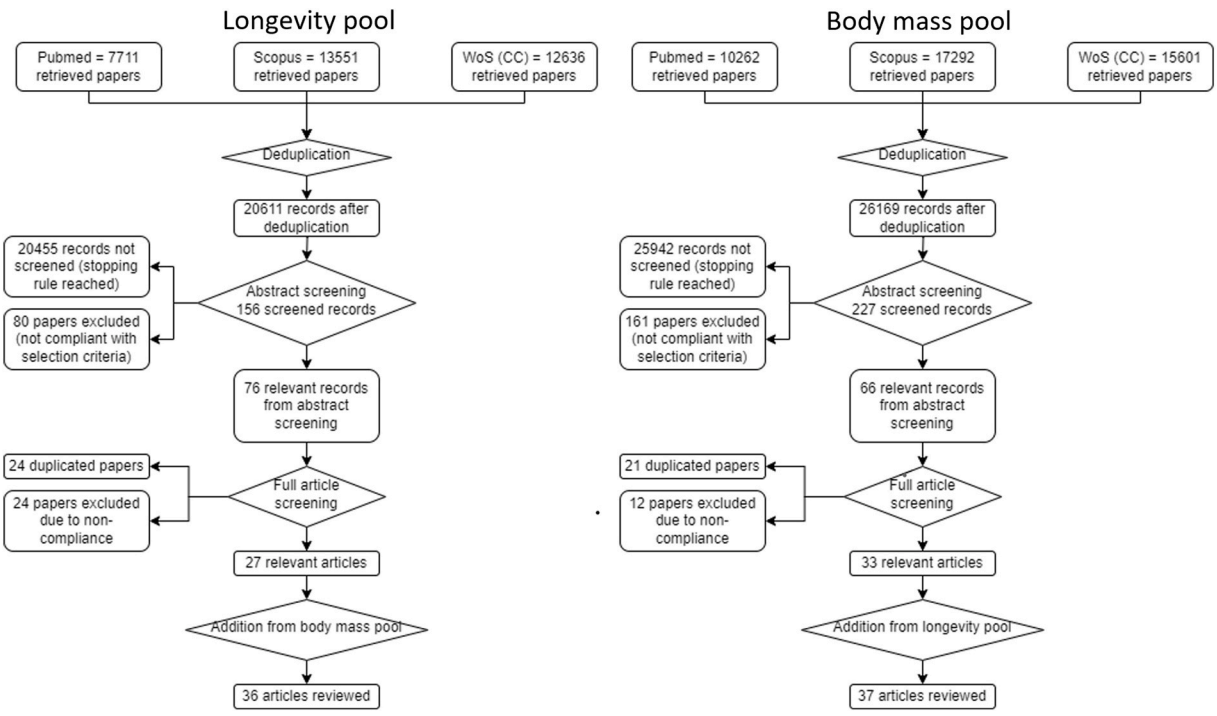

Fig. 1. PRISMA flow diagrams of the two sub-reviews, with details about the number of papers identified, screened, eligible, and included [WoS (CC) = web of science (core collection)].

Control", "Body Mass/Longevity Control" and "Result" have similar scores, so allowing to quickly identify similar literature findings in terms of topic, statistical methodology, and outcome, and enhancing the identification of studies with respect to the trait. The analyses were visualized using a color-coded spectrum [Figs. 2 and 3] generated with a MeV viewer (Multi Experiment Viewer) [21].

Results

Studies search and selection

The results of the literature search and selection procedure described in Section 2 are summarized by the PRISMA flow diagrams [Fig. 1]. In both sub-reviews, the search in Scopus and WoS identified almost twice as many papers as the search in PubMed. Moreover, there was considerable overlap among the sets of papers retrieved from different repositories.

For the pool of studies examining longevity, the deduplication performed through EndNote left 20661 supposedly unique records, but ASReview estimated that 1681 duplicates more were present, suggesting the actual number of unique records to be 18980. 13551 of those papers were retrieved through Scopus, meaning that 71% of the total papers of this pool were found in this repository. For the pool about body mass, the deduplication step performed through EndNote left 26169 records, but ASReview estimated that 654 duplicates more were present, suggesting the actual number of unique records to be 25515. Papers retrieved through Scopus are 17292 and make up for 68% of the total papers found through the search.

For the purpose of testing the feasibility of our approach and in the framework of a mini-review, for both sub-reviews the title and abstract screening was limited to 8 hours of work, whereby the paper acceptance rate did not show substantial decrease before stopping the task. During this time, 156 papers about longevity were screened, and 76 were considered potentially relevant. The analysis of the full texts and a further deduplication brought to a total of 27 studies selected from this pool. The main reason for papers exclusion in the phase of full text screening was the lack of a reported measure of correlation between longevity and the trait(s) under study, usually due to the limited number of species included (less than 4 species).

For what concerns body mass, a total of 227 titles and abstracts were screened: 66 papers were considered potentially interesting and 33 of them were included in the review after a full text screening and a further deduplication.

A final step of cross-addition of studies from one pool to the other led to a total of 36 studies about longevity and 37 studies about body mass. Moreover, the study of Lorenzini et al. (2005) [13] was excluded by the review and substituted with the study of Perillo et al., published in May 2023 [14], which consists in an extension of the analyses of the former.

As mentioned in Section 2.2, AsReview allows to identify most of the papers of interest in the first part of the screening, obtaining a high acceptance rate at the beginning of the procedure, and lower rates as the screening goes on. The observed average acceptance rate of the abstract screening procedure is about 49% for the longevity pool and 29% for the body mass pool, and it did not decrease substantially during the 8 h of screening, meaning that we were still in the highly relevant portion of papers. Considering the whole screening procedure (abstract and full text), the acceptance rate is 17% for longevity and 15% for body mass.

As the acceptance rate did not show any hint of decrease during the procedure, it is not possible to forecast the evolution that it would have had if we had kept screening the whole pools of papers, making in turn impossible to give a meaningful estimate of the number of unseen papers of interest.

In the hypothetical (and highly unlikely) case of AsReview not working properly, the acceptance rate would be constant throughout the whole pools of papers, and we could estimate that the 17% of the papers in the pool of longevity and 15% of the papers in the pool of body mass are of interest for the review. This would result in a total of 3576 and 3804 papers, respectively. Since we have no reason to suspect that AsReview has not worked properly, it is reasonable to claim that the number of unseen papers of interest is much lower than 3500 for each pool. Nevertheless, it is fair to argue that a lot of papers of interest are being ignored in this review, as we have screened only about 1% of the

Table 3
Overview of the analyses which found a statistically significant relation between mammals' longevity and any physiological trait. The column "phylo" has a value of 0 if the phylogenetic relation among the species was not considered, 1 if it was modeled without using phylogenetic trees, and 2 if it was modeled using phylogenetic trees. The column BM has a value of 1 if the relation between longevity and the trait of interest was measured, adjusting for the correlation between longevity and body mass, and 0 if such control was not performed. The Result column has the value "−" if a significant negative correlation was found between the trait and longevity and "+" if a significant positive correlation was found. The p-value is missing if the exact value is not reported in the original study. The traits are grouped by the hallmark (s) of cancer to which they are matched.

Paper (authors, year)	Trait	Type of sample	Phylo	BM	Result	*p*-value
Sustaining proliferative signaling, evading growth suppressors						
Seluanov et al., 2008	Cell proliferation rate for the first 100 days in culture	Skin fibroblasts	2	1	−	0.038
Attaallah et al., 2020	Propensity to stress induced cellular senescence	Skin fibroblasts	0	1	+	<0.01
Enabling replicative immortality, senescent cells						
Pepke et al., 2022	Telomere length	Lung fibroblasts	2	1	−	0.002
	Telomerase activity	Lung fibroblasts	2	1	+	0.029
Resisting cell death						
Brown et al., 2007	Relative manganese superoxide dismutase protein level	Skin fibroblast (O_2 3%)	0	0	+	
	Citrate synthase activity (nmol/min/mg)	Skin fibroblast (O_2 3%)	0	0	+	
Ma et al., 2016	Resistance to cadmium	Skin fibroblasts	2	0	+	0.009
	Resistance to paraquat	Skin fibroblasts	2	0	+	0.014
Pickering et al., 2014	Resistance to protein oxidative stress due to H_2O_2 – protein carbonyl change	Skin fibroblasts	0	1	−	0.004
	Resistance to protein oxidative stress due to H_2O_2 – protein carbonyl change	Skin fibroblasts	2	0	−	0.003
	Resistance to protein oxidative stress due to H_2O_2 – detergent insoluble protein change	Skin fibroblasts	0	0	−	0.04
	Resistance to protein oxidative stress due to H_2O_2 – disulphide bonds change	Skin fibroblasts	0	0	−	0.007
	Baseline protein carbonyl levels	Skin fibroblasts	2	0	−	0.04
Elbourkadi et al., 2014	Phosphorylation speed onset after exposure to H_2O_2	Skin fibroblasts	0	1	−	0.003
	Phosphorylation speed onset after exposure to cadmium	Skin fibroblasts	0	1	−	0.001
	Phosphorylation maintenance after exposure to H_2O_2	Skin fibroblasts	0	1	+	<0.0005
	Phosphorylation maintenance after exposure to cadmium	Skin fibroblasts	0	1	+	<0.0005
	Phosphorylation maintenance after exposure to H_2O_2	Skin fibroblasts	2	0	+	<0.001
	Phosphorylation maintenance after exposure to cadmium	Skin fibroblasts	2	0	+	<0.002
Harper et al., 2007	Cadmium resistance	Skin fibroblasts	2	0	+	0.01
	Cadmium resistance	Skin fibroblasts	0	1	+	0.01
	H_2O_2 resistance	Skin fibroblasts	2	0	+	0.04
	H_2O_2 resistance	Skin fibroblasts	0	1	+	0.02
	Heat resistance	Skin fibroblasts	0	1	+	0.01
	Rotenone resistance	Skin fibroblasts	0	1	+	0.001
Genome instability & mutations						
Lorenzini et al., 2009	DNA end-binding activity	Skin fibroblasts	0	1	+	
	DNA end-binding activity	Skin fibroblasts	2	0	+	0.025
Croco et al., 2015	Long-lasting 53BP1 foci	Fibroblast	0	0	+	0.021
Fink et al., 2011	Micronuclei abundance after damage	Lung fibroblast	0	0	−	0.0006
Brown et al., 2007	Polymerase beta activity	Skin fibroblast (O_2 3%)	0	0	+	0.01 < p < 0.1
Gredilla et al., 2020	Incision of 5-hydroxycytosine containing oligo	Liver mitochondrial fractions	0	0	+	<0.01
	Incision of THF containing oligo	Liver mitochondrial fractions	0	0	+	<0.0001
	Incision of THF containing oligo	Heart mitochondrial fractions	0	0	+	<0.0001
Tian et al., 2019	DNA DSB repair – non–homologous end joining	Lung and skin fibroblasts	2	0	+	0.001
	DNA DSB repair – non–homologous end joining	Lung and skin fibroblasts	2	0	+	0.033
	DNA DSB repair – homologous recombination	Lung and skin fibroblasts	2	0	+	0.0008
	DNA DSB repair – homologous recombination	Lung and skin fibroblasts	2	0	+	0.021
Zhang et al., 2021	Bleomycin-induced somatic mutation frequency – SNV	Lung fibroblasts	1	0	−	0.01
	Bleomycin-induced somatic mutation frequency – INDEL	Lung fibroblasts	1	0	−	0.005
	Dose sensitivity to bleomycin-induced som. mut. freq. – SNV	Lung fibroblasts	0	0	−	<0.0005
	Dose sensitivity to bleomycin-induced som. mut. freq. – INDEL	Lung fibroblasts	0	0	−	0.0035
Cagan et al., 2022	Somatic mutation rate	Colon tissues	0	1	−	<0.0001
	Somatic mutation rate	Colon tissues	2	0	−	
Indirectly related with cancer						
Rodriguez et al., 2016	Proteolytic degradation – HSF1	Liver tissues	2	1	+	<0.001
	Proteolytic degradation – HSF1	Muscle tissues	2	1	+	0.001
	Proteolytic degradation – HSP25	Liver tissues	2	1	+	<0.001
	Proteolytic degradation – HSP25	Muscle tissues	2	1	+	0.005
	Proteolytic degradation – ChTL activity	Muscle tissues	2	1	+	0.007
	Proteolytic degradation – PGPH activity	Muscle tissues	2	1	+	0.007
	Proteolytic degradation – Beclin–1	Muscle tissues	2	1	−	0.02
	Proteolytic degradation – ATG12	Muscle tissues	2	1	+	0.03
	Proteolytic degradation – p-HSF1/HSF1	Liver tissues	2	1	−	0.04
Mota-Martorell et al., 2022	Methionine metabolism – methionine	Heart tissues	2	0	−	<0.005
	Methionine metabolism – PLP	Heart tissues	2	0	−	0.001
	Amino acids – proline	Heart tissues	2	0	−	0.019
	Amino acids – tyrosine	Heart tissues	2	0	−	0.041
Mota-Martorell et al., 2021	Methionine metabolism – pyridoxamine	Plasma	2	0	−	0.045
	TCA cycle metabolites – succinate	Plasma	2	0	−	0.045

(*continued on next page*)

Table 3 (*continued*)

Paper (authors, year)	Trait	Type of sample	Phylo	BM	Result	*p*-value
Aledo et al., 2011	Methionine abundance	Mitochondrial genome and proteome	2	1	–	0.017
	Methionine abundance – AUA–coded methionines	Mitochondrial genome and proteome	2	0	–	0.006
	Number of methionine adding events	Mitochondrial genome and proteome	2	0	–	0.006
Peron et al., 2019	Elasticity of prime-age adult mortality hazard	Demographic data	2	0	–	
	Elasticity of rate of actuarial senescence	Demographic data	2	0	+	
Azpurua et al., 2013	IGF1R levels	Brain tissues	0	1	–	0.0009
	IGF1R levels	Brain tissues	2	0	–	0.0261

papers in each pool.

The selection through full texts screening has taken around 34 h in total (approx. 18 h for the pool about longevity and approx. 16 for the one about body mass), with an average of 1 minute spent to identify each duplicate, and 20 minutes to decide on the inclusion or exclusion of the other papers.

Characteristics of the selected studies

Almost all the studies selected had analyzed more than one trait in relation to longevity and/or body mass, up to a maximum of 35 traits in the study of Mota-Martorell N. et al. (2022) [15]. Usually, the traits analyzed in the same study are strictly related among them (e.g., telomere length and telomere shortening). Moreover, some studies performed multiple analyses on the same trait, focusing on different species or on different types of samples collected from the same species. The studies from the pool about body mass are generally bigger in terms of number of species considered: the median number of species analyzed in a single paper is 11 for the "non-genetic" studies from the pool about longevity (first quartile 6.5, third quartile 16.5); 35 for the "genetic" studies from the same pool (first quartile 26.5, third quartile 48.5); 49 for the "non-genetic" studies from the pool about body mass (first quartile 15, third quartile 123); and 28 for the "genetic" studies from the same pool (three studies in total, with 20, 28 and 36 species). There is some heterogeneity among the selected studies in terms of type of species analyzed: some studies are focused on animals of the same order (e.g., rodents), while other, larger studies compare animals of different orders, sometimes also living in different habitats (e.g., aquatic and terrestrial); some studies even include non-mammals (usually birds).

Most studies, both about longevity and body mass, have focused on physiologic traits (such as telomeres maintenance, cellular senescence and the like). The sub-review about body mass includes also several studies assessing the correlation of species' body mass with factors such as their diet or the characteristics of the environment they live in. Moreover, we found 7 genetic studies about longevity and 3 about body mass; these studies were analyzed separately due to their different study design. As mentioned before, where possible the traits were matched with one or more hallmarks of cancer. In both sub-reviews we found a number of studies analyzing traits related to replicative immortality, cells senescence and genome instability & mutations, and a few studies related to proliferative signaling and growth suppression evasion. Moreover, cell death resistance was strongly represented in the pool about longevity. Many studies (7 for longevity and 24 for body mass) investigated traits that were not traceable back to a specific hallmark of cancer.

As a measure of species longevity, the majority of the studies has considered the maximum lifespan (MLSP), as provided by the database AnAge [16]. The most common measure for the body mass was the average body weight of an adult, taken from AnAge or similar databases (e.g., PanTHERIA [17]). Some studies have used other measures, such as a quantile (e.g., 80th percentile) of the lifespan observed from one or more specific populations in captivity, or the sex-specific average body mass.

Many of the reviewed studies have applied one or more statistical methodologies aimed at overcoming the issues described in Section 2.3. Multiple linear regression and partial correlation coefficient were the most used strategy to control for the correlation between body mass and longevity, while the most used methodologies to account for the evolutionary relatedness among species were: analysis of phylogenetically independent contrasts [18], phylogenetic generalized least squares (PGLS) regression models [19], and hierarchical random-effect regression models accounting for phylogenetic correlation. In total, 12 out of 28 "non-genetic" reviewed studies about longevity have applied methodologies to control for both for phylogenetic relatedness and body mass, 10 have applied a methodology to control just for one of the two (7 for phylogenesis, 3 for body mass) and 6 have not applied any control methodologies. On the other hand, 5 out of 34 "non-genetic" studies about body mass have applied methodologies to control for both for phylogenetic relatedness and longevity, 20 have applied methodology to control just for one of the two (18 for phylogenesis, 2 for longevity) and 9 have not applied any control methodologies. Some of the studies tackling both issues contain two parallel analyses: one adjusting for the correlation between body mass and longevity and the other one adjusting for the phylogenetic structure. One example is the study from Cagan et al. (2020), which estimated the correlation between mutation rate and both longevity and body mass using a multiple regression model to control for the correlation between the two quantities [20]. In the same study, also a PGLS regression including only the longevity as explanatory variable was performed, resulting in a model controlling for the phylogenetic relation among species but not for the correlation between longevity and body mass. On the other hand, one example of analysis controlling for both the issues simultaneously is the multiple PGLS regression model used by Perillo et al. (2023) to analyze the relation between cells immortalization rates and both body mass and longevity across species [14].

The following subsections summarize the main findings of the studies selected. The details of the analyses founding a significant relationship between longevity or body mass and any traits are reported in Tables 3 and 4 respectively. The details about the "genetic" studies are reported in Table 5. The full list of analyses is available in Appendix (file 2), and it includes more details about their design, methodology and results, along with some miscellaneous notes.

The information extraction from the full texts has taken around 96 h in total (approx. 47 for the pool about longevity and more than 49 for the one about body mass), with an average of 90 min spent for each paper, which were almost doubled if the paper was to be included in both sub-reviews. In total, we have spent more than 145 person-hours in the activities of literature selection (16 h for the abstracts and 33.5 h for the full texts) and data extraction (96 h). Our experience provides an estimate for future, more comprehensive efforts.

Key traits appearing related to longevity

Genome instability and mutations

It has been shown that the capacity to recognize DNA double strand breaks (DSB) [4,21,22] and to promote DNA DSB repair pathways – homologous recombination (HR) and non-homologous end joining

Table 4
Overview of the analyses which found a relation between mammals' body mass and any physiological or environmental traits. The column "phylo" has a value of 0 if the phylogenetic relation among the species was not considered, 1 if it was modeled without using phylogenetic trees, and 2 if it was modeled using phylogenetic trees. The column LG has a value of 1 if the relation between body mass and the trait of interest was measured, adjusting for the correlation between longevity and body mass, and 0 if such control was not performed. The Result column has value "−" if a significant negative correlation was found between the trait and the body mass, and "+" if a significant positive correlation was found, while "Q+" means that a quadratic relation has been found, with a positive coefficient for the second-order regressor. The p-value is missing if the exact value is not reported in the original study. The traits are grouped by the hallmark(s) of cancer to which they are matched.

Paper (authors, year)	Trait	Type of sample	Phylo	LG	Result	*p*-value
Sustaining proliferative signaling						
Gillooly et al., 2012	Replicative capacity in culture	Red blood cells	0	0	+	<0.001
Enabling replicative immortality, senescent cells						
Perillo et al., 2023	Immortalization probability	Skin fibroblasts	2	1	−	0.021
Seluanov et al., 2008	Presence of replicative senescence	Skin fibroblasts	2	0	+	0.0398
Seluanov et al., 2007	Telomerase activity coefficient	Heart, liver, spleen, lung, skin, kidney and testis tissues	2	1	−	<0.0001
Pepke et al., 2022	Telomere length	Lung fibroblasts	2	0	−	0.007
Pepke et al., 2020	Telomere shortening rates	Adult individuals	2	0	−	0.011
Perillo et al., 2023	Proliferative capacity	Skin fibroblasts	2	1	+	0.002
Resisting cell death						
Gillooly et al., 2012	Temperature-corrected cell lifespan in vivo	Red blood cells	0	0	+	<0.0001
Genome instability & mutation						
Lorenzini et al., 2011	SAC tolerance	Skin fibroblasts	0	0	+	
	SAC tolerance after colcemid treatment	Skin fibroblasts	0	0	+	
Page et al., 2011	PolBeta activity – [32P]dCTP incorporation	Brain and liver tissues	2	0	−	0.012
Indirectly related with cancer						
Lemaitre et al., 2020	Onset of reproductive senescence	Demographic data	2	0	+	
	Onset of reproductive senescence	Demographic data	2	0	+	
	rate of reproductive senescence	Demographic data	2	0	−	
Tidiere et al., 2014	Onset of male actuarial senescence	Data on survival and seasonal timing of the ruts	2	0	+	
	Actuarial senescence rate – 6 to 9 years old	Data on survival and seasonal timing of the ruts	2	0	−	
Czarnoleski et al., 2018	Cell size	Erythrocytes, enterocytes, chondrocytes, skin epithelial cells, kidney proximal tubule	0	0	+	0.04
Peron et al., 2019	Prime-age adult mortality	Demographic data	2	0	−	
	Onset of actuarial senescence	Demographic data	2	0	−	
	Prime-age stage frequency detection	Demographic data	2	0	−	
Gaillard et al., 2015	Tooth wear index	Values from literature	1	0	−	
	Actuarial senescence rate	Demographic data	1	0	−	
Kozlowski et al., 2010	Cell volume	Values from literature	2	0	+	<0.02
	Erythrocytes size	Values from literature	2	0	+	0.01
Carranza et al., 2007	Cheek-teeth size	Lower occlusal surface area	1	1	+	<0.001
Sibly et al., 2011	Female group size	Values from literature	0	0	+	<0.001
	Female group size	Values from literature	0	0	+	<0.001
Jimenez-Arenas, 2013	Superior post canine tooth occlusal area	Values from literature	2	0	+	<0.0001
	Basal metabolic rate	Values from literature	2	0	+	<0.0001
Cid et al., 2020	Activity range metric 1	Camera traps records	2	0	+	0.003
Hudson et al., 2013	Field metabolic rate (kJ/day)	Adult individuals	1	0	+	
Famoso et al., 2018	Reproductive strategy index	Female life history variables	2	0	−	<0.001
Vallejo-Vargas et al., 2022	Night activity	Camera-trap photos of carnivores	0	0	−	
	Night activity	Camera-trap photos of omnivores	0	0	−	
	Night activity	Camera-trap photos of herbivores	0	0	+	
	Night activity	Camera-trap photos of insectivores (neotropics)	0	0	−	
	Night activity	Camera-trap photos of insectivores (Indo-Malayan tropics and afrotropics)	0	0	+	
Pineda-Munoz et al., 2016	Diet category	Stomach contents	0	0	D	<0.001
Rodriguez et al., 2006	Duration of availability to animals	Range maps from an atlas	0	0	−	
	Environmental temperature	Range maps from an atlas	0	0	Q+	
	Environmental temperature	Range maps from an atlas – previously glaciated areas	0	0	−	
	Plant production seasonality	Range maps from an atlas – previously non-glaciated areas	0	0	Q+	
Tokolyi et al., 2013	Mean temperature	Climate data – chiroptera	2	0	−	<0.01
	Precipitation variance	Climate data – chiroptera	2	0	+	<0.05
	Mean temperature	Climate data – Primates	2	0	−	<0.05
	Precipitation seasonality	Climate data – Rodentia	2	0	+	<0.05
	Annual fecundity	Life history data about chiroptera	2	0	−	<0.01
	Litter mass	Life history data about chiroptera	2	0	+	<0.001
	Total biomass	Life history data about chiroptera	2	0	+	<0.001
	Annual fecundity	Life history data about rodentia	2	0	−	<0.01
	Litter mass	Life history data about rodentia	2	0	+	<0.001
	Total biomass	Life history data about rodentia	2	0	+	<0.001
Carbone et al., 2007	Daily energy intake	Values from literature	0	0	+	
	Daily energy expenditure	Values from literature	0	0	+	
Tucker et al., 2014	Prey minimum mass	Values from PanTHERIA database – terrestrial mammals	2	0	+	
	Prey maximum mass	Values from PanTHERIA database – terrestrial mammals	2	0	+	

(*continued on next page*)

Table 4 (*continued*)

Paper (authors, year)	Trait	Type of sample	Phylo	LG	Result	*p*-value
	Prey mass range	Values from PanTHERIA database – terrestrial mammals	2	0	+	
Isaac et al., 2011	N. of individuals per km^2	Values from literature	1	0	−	<0.05
Rodriguez et al., 2016	Proteolytic degradation – TL activity	Liver tissues	2	1	+	0.008
	Proteolytic degradation – RPT5	Liver tissues	2	1	+	0.005

(NHEJ) pathways [23] – or base excision repair (BER) [24] are positively associated with longevity.

In this regard, Fink et al. have shown that fibroblast cultures from longer-lived species display a reduction in the accumulation of micronuclei when compared with fibroblasts from shorter-lived species, suggesting that unresolved DNA damage persists differently between these species [25]. In confirmation of this, Zhang et al. have observed that a fixed dose of bleomycin induces more mutations, both single-nucleotide variants (SNVs) and small insertions and deletions (INDELs), in mouse and guinea pig cells than in the same cells from long-lived rodents or humans [26]. Cagan et al. report that somatic mutation rate per year varies greatly across species and has a strong inverse relationship with species longevity [20].

Taken together, these data suggest that long-lived species may be capable of processing DNA damage more accurately than short-lived species, although the differences between species may involve multiple aspects of DNA repair and cell cycle controls.

Resisting cell death

An association between species longevity and resistance of proteins in primary fibroblasts to oxidative stress after exposure to H_2O_2 or paraquat has been reported, as well as an increase in protein carbonyl levels in short-lived species [27,28]. In addition, fibroblasts resistance to death induced by H_2O_2, cadmium, heat stress, and rotenone was found correlated with lifespan among mammals [28,29]. In the same context, Elbourkadi et al. report that fibroblasts from the shorter-lived species of rodents show rapid induction of ERK phosphorylation, as a response of cells to stress, while cells from longer-lived species show slower and more prolonged activation of this kinase, suggesting that fibroblasts from long-lived species may be less susceptible to the early phases of damage from stress inductors (e.g. cadmium or H_2O_2) and indicating that altered kinetics of ERK activity may contribute to their stress resistance properties [30]. See Section 4 for more details regarding the interpretation of this hallmark.

Key traits appearing related to large body mass

Enabling replicative immortality/senescence

We have found an inverse relationship between the spontaneous immortalization probability of mammalian skin fibroblasts and body mass, suggesting a need to evolve stringent mechanisms to prevent the unrestrained proliferation of somatic cells during the evolution of large body mass [14]. We and others also have observed that the proliferative capacity of normal cells is related to body mass, suggesting that large species need to be equipped with enough proliferative capacity of karyotypically stable cells in order to guarantee the construction of large bodies [13,14]. Seluanov et al. also report that telomerase activity negatively co-evolves with body mass, not lifespan [31]. Pepke and Eisenberg report that telomere length negatively co-evolves with body mass, and they find a negative correlation between telomere shortening rates and body mass [32,33]. Overall, these data support the notion that telomere shortening is an evolved anticancer mechanism.

Genome instability and mutations

Enhanced genome stability, which is associated with the evolution of longevity, also appears to be associated with the evolution of body mass, although the specific mechanisms involved may differ. For example, it seems that longer-lived species have acquired a more efficient capacity to segregate the genetic materials in daughter cells during cell division, as suggested by our observation of a more efficient spindle assembly checkpoint in large mammals [34]. This consideration is supported by a meta-analysis that shows an inverse correlation of the spontaneous micronucleated erythrocyte frequency with body mass [35]. Page and Stuart, in addition, report that the activities of DNA base excision repair enzymes in liver and brain correlate with body mass but not lifespan [36].

Genetic analyses

Several genomic analyses and transcriptomic studies have been carried out to gain insight into genetic mechanisms related to mammalian lifespan and body mass. An overview of those studies is provided in Table 5.

For example, Ma et al. have shown that fibroblasts from longer-lived species have high expression of genes related to DNA repair and maintenance and low expression levels of genes involved in proteolysis, autophagy, and apoptosis. Specifically, the genes that coded for the tumor suppressor TP53, apoptosis regulator BAX, and several growth and proliferation signaling pathways were downregulated in the longer-lived species' fibroblasts. On the other hand, genes involved in DNA repair (e.g., Msh6, Pms2, Pnkp, Ercc1, C17orf70, Fancg, Rif1, Terf1, Tinf2) and in glucose metabolism were up-regulated [29]. This strategy appears in contrast with the TP53 copy number expansion observed in elephants [37], suggesting that the evolution of longevity and large body mass may require the co-evolution of different cancer reduction strategies.

Sahm et al. have identified some enrichment for genes known to be related to aging in 17 rodent species with different lifespans [38]. Among these, they found enrichments of oxidoreductase activity, metal ion homeostasis, transport, cellular respiration, as well as processes regulated by the mTOR pathway: translation, autophagy, and inflammation.

Regarding body mass studies, Vedelek et al. have found that the elimination of the GABPA transcription factor (TF) site in large rodents leads to the loss of TERT promoter activity; thus, it could determine whether the replicative senescence plays a role in tumor suppressor in these species, which may be directly related to body mass [39]. Comparative genomics studies on body size-associated genes (BSAGs) have shown 100 BSAGs statistically significantly enriched in cancer control in carnivores, 15 of which were found to be under rapid evolution in extremely large carnivores, suggesting that large carnivores might have evolved an effective mechanism to resist cancer, which could be regarded as molecular evidence to explain Peto's paradox [40].

In addition, Caulin et al. have observed no positive correlation of tumor-suppressor genes with increasing body mass and longevity, but evidence of amplification of some genes among large or long-lived species [41].

Results of clustering

An overview of the characteristics of the analyses presented in the reviewed studies is provided by the dendrogram based on the complete linkage clustering provided in Figs. 2 and 3. Traits cluster together when the 4 variables considered have similar scores. For what concerns the analyses on longevity, 4 major clusters are observed, corresponding almost perfectly with the 4 possible levels of statistical robustness.

Table 5
Overview of the papers which have investigated the genetic mechanisms related to mammalian lifespan and body mass. The studies are grouped based on the sub-review to which they belong (longevity, body mass, or both).

Authors, year	Trait	Type of sample	Results
Longevity			
Ma et al., 2016 [29]	Genes expression	Skin fibroblasts	Longer-lived species cells up-regulated genes involved in DNA repair and glucose metabolism and down-regulated proteolysis and protein transport.
Davies et al., 2014 [42]	Growth hormone receptor gene (GHR) and Insulin-like growth factor 1 receptor gene (IGF1R) nucleotide sequences	Genomic and transcriptomic data	Longer-lived rodents showed little amino acid variation in the transmembrane domains of either GHR and IGF1R compared to much shorter-lived ones.
Kowalczyk et al., 2020 [43]	Genes related to cell cycle, DNA repair, cell death, IGF1 pathway and immunity	Genomic data	Long-lived species showed an increased constraint in inflammation, DNA repair, and NFKB-related pathways.
Sahm et al., 2018 [38]	Genes related to defense against free radicals, iron homeostasis, cellular respiration and translation	Genomic and transcriptomic data	Long-lived rodents reveal signatures of positive selection in genes related to aging; among enriched functional terms were many of the processes that are regulated by the mTOR pathway, e.g. translation, autophagy and cellular respiration.
Yu et al., 2021 [44]	Insulin/IGF-1 signaling and Immune-response-related pathway	Genomic data	In long-lived species were identified 16 unique positively selected genes and 23 rapidly evolving genes, including 9 genes involved in regulating lifespan through the insulin/IGF-1 signaling (IIS) pathway and 11 genes highly enriched in immune-response-related pathways.
Bozek et al., 2017 [45]	Lipid concentration and enzyme conservation	Liver, muscle, kidney, heart, cortex and cerebellum tissues	Lifespan is associated with distinct lipidome features shared across three mammalian clades.
Jobson et al., 2009 [46]	Amino acid conservation	Genomic data	Genes involved in lipid composition and vitamin C binding have collectively undergone increased selective pressure in long-lived species, whereas genes involved in DNA replication/repair or antioxidation have not.
Li and deMagalhães , 2011 [47]	Proteins under accelerated evolution – GOs: Actin cytoskeleton, 1-phosphatidylinositol-3-kinase activity, Phosphoinositide 3-kinase complex, Response to food and Circadian rhythm, Phospholipid metabolic and Cholesterol catabolic processes, Cellular responses to damage (DNA repair & others), Proteasome-ubiquitin system	Genomic data	Several proteins with longevity-specific selection patterns, previously related to aging and DNA damage repair and response were identified.
Body mass			
Vedelek et al., 2020 [39]	Telomerase promoter activity and transcription factor binding	Genomic and transcriptomic data	The elimination of GABPA transcription factor site(s) in large rodents leads to the loss of TERT promoter activity.
Huang et al., 2021 [40]	Body size associated genes (including REGs) and fixed aminoacid changes	Genomic data	337 genes were related to body size; among these, 100 genes were enriched in cancer control in carnivores, 15 of which were found to be under rapid evolution in extremely large carnivores. For small carnivores, 15 rapidly evolving genes were identified and 6 genes with fixed amino acid changes were reported to reduce body size.
Body mass and longevity			
Caulin et al., 2015 [41]	Number of tumor-suppressor genes copies (such as TP53, MAL, FBXO31), mutation rate and number of required mutations for carcinogenesis	Genomic data	No positive correlation between tumor-suppressor genes with increasing body mass and longevity. Evidence of the amplification of TP53 in elephants, MAL in horses and FBXO31 in microbats.

Reading the dendrogram from the top, we find studies with both controls, only body mass control, no control, and only phylogenetic control. Within those clusters, we observe sub-clusters based on the hallmark(s) of cancer matched with the trait and on the type of association between the trait and the longevity measure. The dendrogram on body mass studies shows 4 major clusters, too, but their structure is less straightforward. Starting from the top of the dendrogram, we can observe: one small cluster composed only of analyses about environmental traits performed without any control and observing a negative correlation of the traits with species' body mass; one big cluster of analyses performed without any control and observing positive or non-significant correlation; one big cluster of analyses performed with only phylogenetic control; one small cluster of analyses performed controlling for longevity. The results of the clustering give an idea of which traits have been studied with a more robust statistical methodology in the reviewed works.

Discussion and conclusions

This work shows that the process of evolution has utilized multiple avenues to improve cancer resistance in long-lived and large animals. Although the analysis presented here has limitations, it suggests that many pathways have been exploited to solve the intertwined problem of preventing cancer while allowing cellular longevity and the high number of cell divisions required to evolve long lifespan and large body mass. Thus, what is observed as Peto's paradox is the outcome of the repeated evolution of diverse mechanisms.

Based on the subset of the hallmarks of cancer, which are related to lifespan and body mass, it appears that genomic stability & mutation, resisting cell death, enabling replicative immortality, and possibly evading growth suppressors are common pathways which have been modified by evolutionary pressures to support these two phenotypes.

Regarding the hallmark "resisting cell death", we note that interpreting these findings in light of Peto's paradox needs consideration that an evolved resistance to cell death is a trait directly and positively contributing to the development of cancer, thus this hallmark does not qualify as a possible cancer-reducing evolutionary strategy for long-lived species, since it appears positively related to longevity.

As explained in Section 3, less than 1% of the potentially eligible papers have undergone the abstract screening procedure, resulting in the exclusion of a large proportion of papers and the absence of 8 out of 14 known hallmarks. Thus, our work is far from a complete analysis of

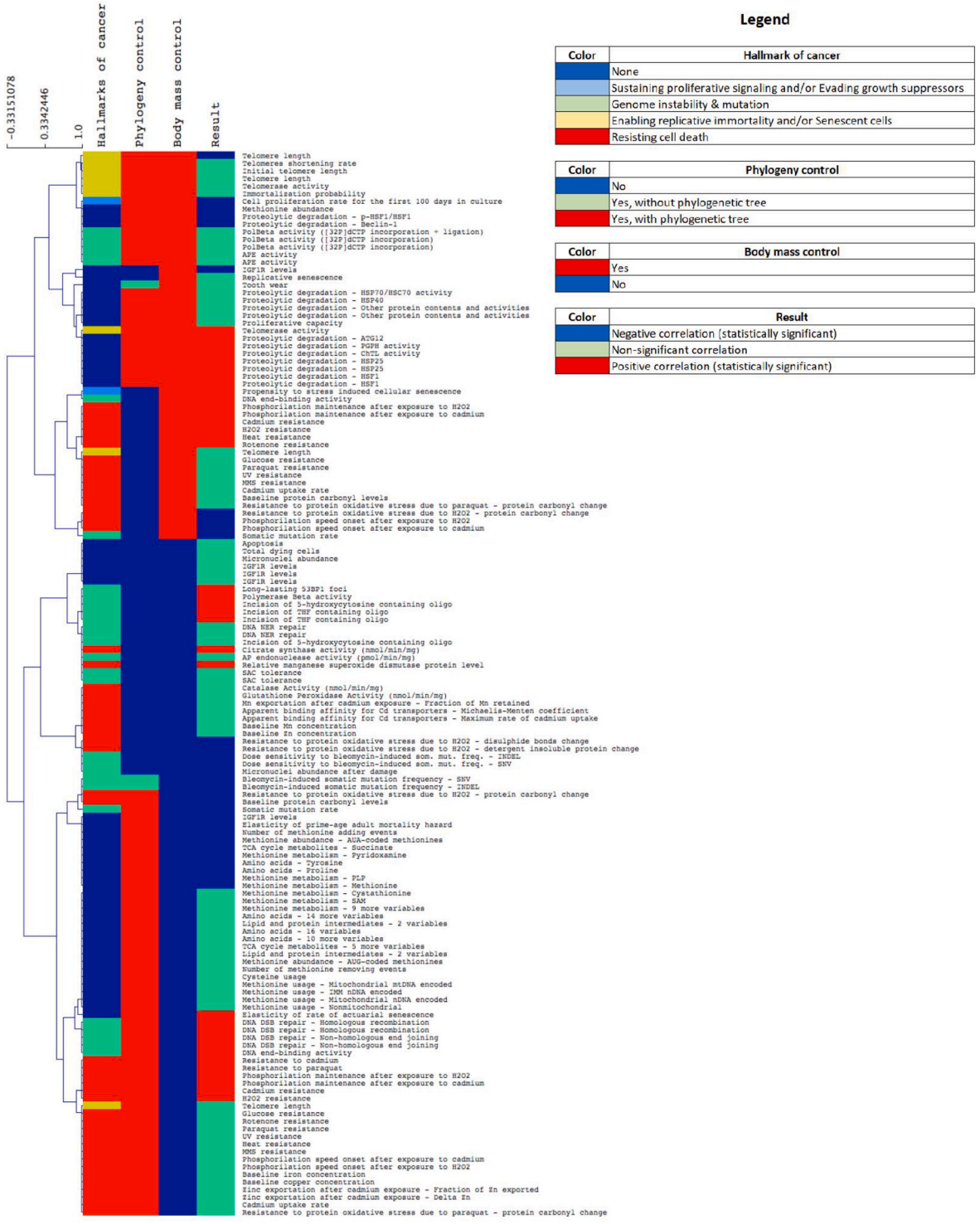

Fig. 2. Hierarchical clustering of curated literature data by trait – Longevity. The characteristics of each analysis present in the reviewed studies are statistically evaluated by their degree of correlation (Pearson) and clustered together in a hierarchical fashion, forming groups of traits. The characteristics include: hallmarks of cancer (sustaining proliferative signaling, evading growth suppressors, enabling replicative immortality, senescent cells, resisting cell death, genome instability & mutation); phylogeny control (with phylogenetic tree, without phylogenetic tree, no control at all); body mass control (yes or not) and results (positively correlated, negatively correlated, not statistically significant).

the present human knowledge of the traits that could explain the evolution of significant longevity and body mass.

This can be addressed providing as "prior knowledge" to AsReview papers investigating traits related to all the hallmarks, since those papers drive the screening procedure, especially in the very first phases. Still, this would likely not be enough to obtain a global picture of all the biological strategies against cancer. We can roughly estimate that by screening between 20% and 40% of the abstracts, we would have spent at least 2000 person-hours in the activities of studies selection and data extraction. This suggests that, to scale up the scope of this review, it would be crucial to use tools allowing to streamline the phases of full text screening, data extraction, and data visualization, as we have

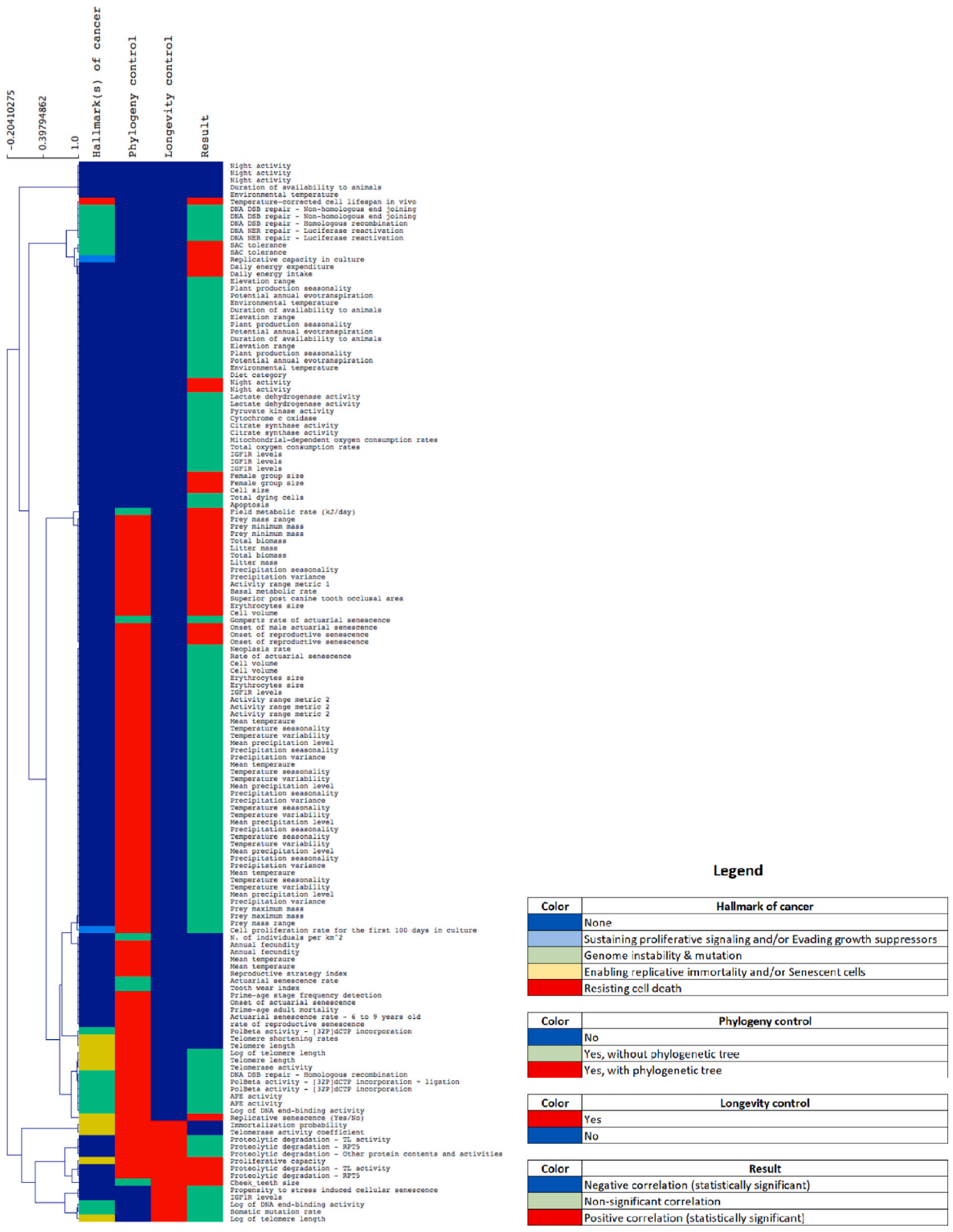

Fig. 3. Hierarchical clustering of curated literature data by trait – Body mass. The characteristics of each analysis present in the reviewed studies are statistically evaluated by their degree of correlation (Pearson) and clustered together in a hierarchical fashion, forming groups of traits. The characteristics include: hallmarks of cancer (sustaining proliferative signaling, evading growth suppressors, enabling replicative immortality, senescent cells, resisting cell death, genome instability & mutation); phylogeny control (with phylogenetic tree, without phylogenetic tree, no control at all); longevity control (yes or not) and results (positively correlated, negatively correlated, not statistically significant).

already done using AsReview for the abstract screening procedure.

In addition, the limitation inherent in a comparative approach may limit the types of questions examined by the available studies. For example, immune surveillance is a hallmark that bears a relationship to both cancer and aging but is difficult to address it in a comparative study due to both the limited understanding of the process in the human setting and the lack of tools (antibodies, etc.) to examine the process in multiple species. For these reasons, relatively few studies have investigated these relationships. These limitations help explain why we observe mainly hallmarks related to cellular mechanisms without observing

supracellular systems such as immune surveillance, angiogenesis, inflammation, etc. The key assumption in Peto's paradox, in fact, is that cancer is a cellular disease caused by random mutations in DNA. The existence of the paradox could itself be an indication that this assumption may limit our understanding of a very complex phenomenon.

A merit of our work is the big effort performed in the organization of this scientific knowledge, constructing detailed and curated tables summarizing the main characteristics and results of the reviewed studies. Moreover, we used some innovative methodologies, such as the active learning-based screening tool and a statistical clustering approach not previously applied for the visualization of literature data from this field. As the application of these tools is of interest to compiling literature reviews in general, the proposed approach can be applied also by researchers from different study fields. Still, in order to perform reviews involving a high number of papers, the process needs to be streamlined, especially in the steps related to full text screening and data curation.

An important element that emerges from the present work is the methodological heterogeneity among comparative biology studies. For example, there is heterogeneity in the quantity and types of species analyzed, the types of data collected, the metrics used to measure longevity and/or body mass, the statistical methodology applied, and the estimator reported to present the results. This makes it impractical to interpret the results of each study in detail, and almost impossible to quantitatively synthetize the results from multiple studies, even when only considering studies designed to examine similar questions. This challenges the extraction of meaningful information from the existing literature and highlights the need for novel approaches to effectively collect, appraise, and interpret the existing research in this field. Moreover, this suggests that a standard procedure of statistical analysis is needed for comparative biology studies, as already stressed by previous works [12].

However, the existence of traits within the same hallmark of cancer related to longevity and body mass indicates that these traits represent convergent strategies for cancer resistance which are not simply "chosen" randomly, but are more frequently exploited because of their efficacy. One may infer that the underlying biological mechanisms of such strategies are productive targets to beneficially engineer a cancer-resistant phenotype. As such, these traits may represent opportunities to develop novel treatments or preventive options for cancer in humans.

Appendix. Supplementary Materials

(1) The PRISMA checklist, (2) the data extracted from the studies and (3) the data used for the clustering procedure can be found online at https://doi.org/10.1016/j.bj.2023.100654.

References

[1] Peto R, Roe FJ, Lee PN, Levy L, Clack J. Cancer and ageing in mice and men. Br J Cancer 1975;32(4). 411–26.
[2] Callier V. Core concept solving Peto's Paradox to better understand cancer. Proc Natl Acad Sci USA 2019;116(6):1825–8.
[3] Vincze O, Colchero F, Lemaître JF, Conde DA, Pavard S, Bieuville M, et al. Cancer risk across mammals. Nature 2022;601(7892). 263–7.
[4] Croco E, Marchionni S, Storci G, Bonafè M, Franceschi C, Stamato TD, et al. Convergent adaptation of cellular machineries in the evolution of large body masses and long life spans. Biogerontology 2017;18(4):485–97.
[5] Hanahan D. Hallmarks of cancer: new dimensions. Cancer Discov 2022;12(1): 31–46.
[6] Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell 2011;144 (5). 646–74.
[7] Hanahan D, Weinberg RA. The hallmarks of cancer. Cell 2000;100(1):57–70.
[8] Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021:372: n71.
[9] ASReview LAB developers. ASReview LAB - A tool for AI-assisted systematic reviews. Zenodo; 2023.
[10] van de Schoot R, de Bruin J, Schram R, Zahedi P, de Boer J, Weijdema F, et al. An open source machine learning framework for efficient and transparent systematic reviews. Nat Mach Intell 2021;3:125–33.
[11] Brouwer M, Hofstee L, van de Brand S, Teijema J, Ferdinands G, de Boer J, et al. AI-aided systematic review to create a database with potentially relevant papers on depression, anxiety, and addiction. PsyArXiv preprint. 2022. https://doi.org/10.31234/osf.io/j6nqz.
[12] Speakman JR. Body size, energy metabolism and lifespan. J Exp Biol 2005;208(Pt 9):1717–30.
[13] Lorenzini A, Tresini M, Austad SN, Cristofalo VJ. Cellular replicative capacity correlates primarily with species body mass not longevity. Mech Ageing Dev 2005; 126(10):1130–3.
[14] Perillo M, Punzo A, Caliceti C, Sell C, Lorenzini A. The spontaneous immortalization probability of mammalian cell culture strains, as their proliferative capacity, correlates with species body mass, not longevity. Biomed J 2023;46(3):100596.
[15] Mota-Martorell N, Jové M, Berdún R, Òbis È, Barja G, Pamplona R. Methionine metabolism is down-regulated in heart of long-lived mammals. Biology 2022;11 (12):1821.
[16] de Magalhães JP, Costa J. A database of vertebrate longevity records and their relation to other life-history traits. J Evol Biol 2009;22(8). 1770–4.
[17] Jones KE, Bielby J, Cardillo M, Fritz SA, O'Dell J, Orme CDL, et al. PanTHERIA: a species-level database of life history, ecology, and geography of extant and recently extinct mammals. Ecology 2009;90. 2648.
[18] Felsenstein J. Phylogenies and the comparative method. Am Nat 1985;125:1–15.
[19] Grafen A. The phylogenetic regression. Philos Trans R Soc Lond B Biol Sci 1989; 326(1233). 119–57.
[20] Cagan A, Baez-Ortega A, Brzozowska N, Abascal F, Coorens THH, Sanders MA, et al. Somatic mutation rates scale with lifespan across mammals. Nature 2022;604 (7906):517–24.
[21] Croco E, Marchionni S, Bocchini M, Angeloni C, Stamato T, Stefanelli C, et al. DNA damage detection by 53BP1: relationship to species longevity. J Gerontol A Biol Sci Med Sci 2017;72(6). 763–70.
[22] Lorenzini A, Johnson FB, Oliver A, Tresini M, Smith JS, Hdeib M, et al. Significant correlation of species longevity with DNA double strand break recognition but not with telomere length. Mech Ageing Dev 2009;130(11–12). 784–92.
[23] Tian X, Firsanov D, Zhang Z, Cheng Y, Luo L, Tombline G, et al. SIRT6 is responsible for more efficient DNA double-strand break repair in long-lived species. Cell 2019;177(3):622–38. e22.
[24] Gredilla R, Sánchez-Román I, Gómez A, López-Torres M, Barja G. Mitochondrial base excision repair positively correlates with longevity in the liver and heart of mammals. Geroscience 2020;42(2). 653–65.
[25] Fink LS, Roell M, Caiazza E, Lerner C, Stamato T, Hrelia S, et al. 53BP1 contributes to a robust genomic stability in human fibroblasts. Aging 2011;3(9). 836–45.
[26] Zhang L, Dong X, Tian X, Lee M, Ablaeva J, Firsanov D, et al. Maintenance of genome sequence integrity in long- and short-lived rodent species. Sci Adv 2021;7 (44). eabj3284.
[27] Pickering AM, Lehr M, Kohler WJ, Han ML, Miller RA. Fibroblasts from longer-lived species of primates, rodents, bats, carnivores, and birds resist protein damage. J Gerontol Ser ABiol Sci Med Sci 2015;70(7). 791–9.
[28] Harper JM, Salmon AB, Leiser SF, Galecki AT, Miller RA. Skin-derived fibroblasts from long-lived species are resistant to some, but not all, lethal stresses and to the mitochondrial inhibitor rotenone. Aging Cell 2007;6(1):1–13.
[29] Ma S, Upneja A, Galecki A, Tsai YM, Burant CF, Raskind S, et al. Cell culture-based profiling across mammals reveals DNA repair and metabolism as determinants of species longevity. Elife 2016;5:e19130.
[30] Elbourkadi N, Austad SN, Miller RA. Fibroblasts from long-lived species of mammals and birds show delayed, but prolonged, phosphorylation of ERK. Aging Cell 2014;13(2). 283–91.
[31] Seluanov A, Chen Z, Hine C, Sasahara TH, Ribeiro AA, Catania KC, et al. Telomerase activity coevolves with body mass not lifespan. Aging Cell 2007;6(1): 45–52.
[32] Pepke ML, Eisenberg DTA. Accounting for phylogenetic relatedness in cross-species analyses of telomere shortening rates. Exp Results 2020;1:e11.
[33] Pepke ML, Eisenberg DTA. On the comparative biology of mammalian telomeres: telomere length co-evolves with body mass, lifespan and cancer risk. Mol Ecol 2022;31(23). 6286–96.
[34] Lorenzini A, Fink LS, Stamato T, Torres C, Sell C. Relationship of spindle assembly checkpoint fidelity to species body mass, lifespan, and developmental rate. Aging 2011;3(12):1206–12.
[35] Croco E, Marchionni S, Lorenzini A. Genetic instability and aging under the scrutiny of comparative biology: a meta-analysis of spontaneous micronuclei frequency. Mech Ageing Dev 2016;156:34–41.
[36] Page MM, Stuart JA. Activities of DNA base excision repair enzymes in liver and brain correlate with body mass, but not lifespan. Age 2012;34(5). 1195–209.
[37] Sulak M, Fong L, Mika K, Chigurupati S, Yon L, Mongan NP, et al. TP53 copy number expansion is associated with the evolution of increased body size and an enhanced DNA damage response in elephants. Elife 2016;5: e11994.
[38] Sahm A, Bens M, Szafranski K, Holtze S, Groth M, Görlach M, et al. Long-lived rodents reveal signatures of positive selection in genes associated with lifespan. PLoS Genet 2018;14(3):e1007272.
[39] Vedelek B, Maddali AK, Davenova N, Vedelek V, Boros IM. TERT promoter alterations could provide a solution for Peto's paradox in rodents. Sci Rep 2020;10 (1):20815.
[40] Huang X, Sun D, Wu T, Liu X, Xu S, Yang G. Genomic insights into body size evolution in Carnivora support Peto's paradox. BMC Genom 2021;22(1):429.
[41] Caulin AF, Graham TA, Wang LS, Maley CC. Solutions to Peto's paradox revealed by mathematical modelling and cross-species cancer gene analysis. Philos Trans R Soc B Biol Sci 2015;370(1673):20140222.

[42] Davies KT, Tsagkogeorga G, Bennett NC, Dávalos LM, Faulkes CG, Rossiter SJ. Molecular evolution of growth hormone and insulin-like growth factor 1 receptors in long-lived, small-bodied mammals. Gene 2014;549(2). 228–36.
[43] Kowalczyk A, Partha R, Clark NL, Chikina M. Pan-mammalian analysis of molecular constraints underlying extended lifespan. Elife 2020;9:e51089.
[44] Yu Z, Seim I, Yin M, Tian R, Sun D, Ren W, et al. Comparative analyses of aging-related genes in long-lived mammals provide insights into natural longevity. Innovation 2021;2(2):100108.
[45] Bozek K, Khrameeva EE, Reznick J, Omerbašić D, Bennett NC, Lewin GR, et al. Lipidome determinants of maximal lifespan in mammals. Sci Rep 2017;7(1):5.
[46] Jobson RW, Nabholz B, Galtier N. An evolutionary genome scan for longevity-related natural selection in mammals. Mol Biol Evol 2010;27(4). 840–7.
[47] Li Y, de Magalhães JP. Accelerated protein evolution analysis reveals genes and pathways associated with the evolution of mammalian longevity. Age 2013;35(2). 301–14.