

Article

The Change Rate of the *Fbxl21* Gene and the Amino Acid Composition of Its Protein Correlate with the Species-Specific Lifespan in Placental Mammals

Vassily A. Lyubetsky ¹, Gregory A. Shilovsky ¹, Jian-Rong Yang ^{2,3}, Alexandr V. Seliverstov ¹
and Oleg A. Zverkov ^{1,*}

¹ Institute for Information Transmission Problems of the Russian Academy of Sciences (Kharkevich Institute), 127051 Moscow, Russia; lyubetsk@iitp.ru (V.A.L.); gregory_sh@list.ru (G.A.S.); slvstv@iitp.ru (A.V.S.)

² Department of Genetics and Biomedical Informatics, Zhongshan School of Medicine, Sun Yat-sen University, Guangzhou 510080, China

³ Key Laboratory of Tropical Disease Control, Sun Yat-sen University, Ministry of Education, Guangzhou 510080, China

* Correspondence: zverkov@iitp.ru

Simple Summary: The relationship between genomic characteristics and species traits is of paramount importance for biology. For genomic characteristics, we considered characteristics close to the rate of gene change in the process of evolution, including such extreme cases as pseudogenization and gene loss. This can be conceptualized as selection at the gene level. Amino acid substitutions in the positions that functionally determine protein domains were also considered. For species traits, we considered the maximal reported lifespan, the body weight of an adult animal, and the related longevity quotient. The first and second traits are also related, but only at the stochastic level; therefore, their behavior is a priori not similar. We proposed a novel technique that allows one to determine the relationship between any genomic characteristic and species traits. This technique is exemplified in the physiologically significant genes involved in regulating circadian rhythms, which change quite rapidly during evolution. Regardless of devising this technique, the study of the genes that are critical for circadian rhythms is of interest on its own. For instance, we thoroughly examined the paralogous genes *Fbxl21* and *Fbxl3*, which are involved in the regulation of circadian rhythms. We found out that the above-mentioned characteristic of the *Fbxl21* gene correlates with the maximal reported lifespan and body weight only in two superorders of placental mammals, Euarchontoglires (the clades Euarchonta, Lagomorpha, and Rodentia) and Afrotheria. On the contrary, such a correlation is not observed in other superorders, such as Laurasiatheria and Xenarthra. The presence or absence of the correlation is confirmed statistically with a very high accuracy. The rate of change in the *Fbxl21* gene indicates, for example, the peculiarity of its evolution in apes.



Citation: Lyubetsky, V.A.; Shilovsky, G.A.; Yang, J.-R.; Seliverstov, A.V.; Zverkov, O.A. The Change Rate of the *Fbxl21* Gene and the Amino Acid Composition of Its Protein Correlate with the Species-Specific Lifespan in Placental Mammals. *Biology* **2024**, *13*, 792. <https://doi.org/10.3390/biology13100792>

Academic Editor: Xiaodong Li

Received: 21 August 2024

Revised: 24 September 2024

Accepted: 28 September 2024

Published: 2 October 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: This article proposes a methodology for establishing a relationship between the change rate of a given gene (relative to a given taxon) together with the amino acid composition of the proteins encoded by this gene and the traits of the species containing this gene. The methodology is illustrated based on the mammalian genes responsible for regulating the circadian rhythms that underlie a number of human disorders, particularly those associated with aging. The methods used are statistical and bioinformatic ones. A systematic search for orthologues, pseudogenes, and gene losses was performed using our previously developed methods. It is demonstrated that the least conserved *Fbxl21* gene in the Euarchontoglires superorder exhibits a statistically significant connection of genomic characteristics (the median of dN/dS for a gene relative to all the other orthologous genes of a taxon, as well as the preference or avoidance of certain amino acids in its protein) with species-specific lifespan and body weight. In contrast, no such connection is observed for *Fbxl21* in the Laurasiatheria superorder. This study goes beyond the protein-coding genes, since the accumulation of amino acid substitutions in the course of evolution leads to pseudogenization and even gene loss, although the relationship between the genomic characteristics and the species traits is still preserved. The proposed methodology is illustrated using the examples of circadian

rhythm genes and proteins in placental mammals, e.g., longevity is connected with the rate of *Fbxl21* gene change, pseudogenization or gene loss, and specific amino acid substitutions (e.g., asparagine at the 19th position of the CRY-binding domain) in the protein encoded by this gene.

Keywords: bioinformatics; mammals; circadian rhythms; longevity; rate of change; amino acid composition; pseudogenization; gene loss

1. Introduction

Biorhythms are crucial for adaptation and depend on many factors. Since circadian rhythm disruption typically occurs upon aging, maintaining the proper function of the circadian rhythm system, underlying biorhythm regulation, holds great promise for improving the quality of life and increasing lifespan [1]. Studying the changes in the proteins that regulate circadian rhythms induced by pharmaceuticals or other genes is one of the primary goals of antiaging medicine [2]. For instance, the chronobiotics targeted at the Cry1 cryptochrome protein influence circadian rhythms and lifespan [1,3,4].

Based on the previously developed algorithms and devised software [5–8], this article examines the relationship between the maximum lifespan of species together with the body weight of an adult animal and the accumulation of changes during evolution of the species-specific genes listed above and their proteins, containing the Cry-binding domain (CBD), which contribute to the stability of cryptochromes.

Circadian clock proteins. Cryptochrome proteins (Cry1, Cry2) and period proteins (Per1, Per2, Per3) are involved in the mammalian circadian clock regulation [9]. They interact with other proteins such as Fbxl3, Fbxl21, β -TrCP, Fbxw11, GSK3 α , GSK3 β , CK1 δ (Csnk1d), and CK1 ϵ (Csnk1e). The alternative names of the proteins commonly used in literature are given in brackets. The names of the respective genes are spelled in a traditional way: *Btrc* (or *Fbxw1*). The core circadian rhythm mechanism also involves the Clock:Bmal1 heterodimer [10]. Per and Cry proteins are translocated into the nucleus, forming stable complexes, which include Cry1 and Cry2, and suppress the transcription of *Per1* and *Per2* genes by interacting with the transcriptional activators Clock and Bmal1. This causes mRNA levels to alternately rise and fall, followed by oscillations of the Per1 and Per2 protein levels within a period of approximately 24 h. Per phosphorylation by the casein kinase 1 epsilon (CK1 ϵ) prevents its premature accumulation in the cytoplasm, inducing Per cytoplasmic degradation if it is not bound to Cry protein. The Per protein is less stable unless it is bound to Cry. Therefore, Per and Cry levels are controlled by ubiquitin-dependent proteasome degradation.

It is assumed that Cry protein mediates the adjustment of the circadian clock to the external lighting regime by inhibiting *Per* transcription and therefore regulating its level [11,12].

GSK3 α and GSK3 β kinases modify the products of almost all main genes responsible for circadian rhythms (*Bmal1*, *Clock*, *Per*, *Cry*, *Rev-erba*). Modifications of the positive regulators of biorhythms lead to their inactivation and degradation, whereas modification of the negative regulators induces their activation and translocation into the nucleus [13,14]. GSK3 β kinase, involved in the proteasome degradation system, also mediates cell aging and the development of many pathological processes, including sleep disorders and potentially leading to various neurodegenerative diseases.

Gsk3a, *Gsk3b*, and other genes contributing to circadian rhythm regulation, such as *Clock*, *Cry1*, and *Cry2*, are conserved in mammals [14]. For instance, the GSK3 β protein encoded by the *Gsk3b* gene is conservative even in invertebrates and is almost identical in humans and macaques. The *Gsk3a* gene has no orthologs in birds and some reptiles. In contrast, the *Clock*, *Cry1*, and *Cry2* genes have orthologs in most vertebrates.

Mutations in casein kinase 1 (CK1) alter circadian behavior in various placental mammals [15,16]. Inhibiting CK1 by pharmaceuticals can be used for therapy in aberrant

circadian rhythms [2,17]. CK1δ (Csnk1d) deficiency extends the circadian period, in contrast to the deficiency of CK1ε (Csnk1e) [18].

The system of regulating the mammalian circadian clock is complex and remains enigmatic. A possible scheme is shown in Figure 1, and questions concerning this regulation are discussed, among others, in [9]. However, in the present work, we touch upon only one aspect of the system at the end of this section. Of note, the regulation of circadian rhythms has been studied not only in mammals but also, for example, in *Culex pipiens* mosquitoes [19] and in *Drosophila* populations having different lifespans [20]. Meanwhile, these studies were performed within biological frameworks and did not address the question of statistically processing the relationship between species and genomic characteristics.

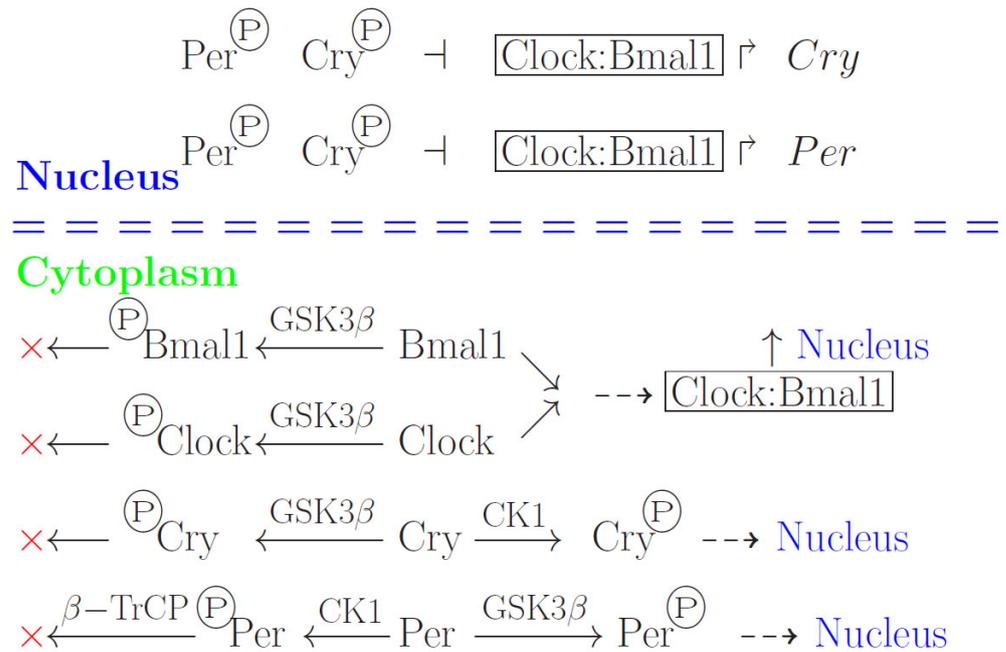


Figure 1. Molecular mechanism of circadian rhythms. The processes going on in the nucleus are shown at the top, and the processes in the cytoplasm are shown at the bottom. Dashed lines with arrows, translocation of the corresponding proteins into the nucleus; solid lines with arrows, direct effect, including catalysis; solid line with a blunt end, inhibition; circle with letter “P”, phosphate groups attached to the proteins; broken arrows, promoters; the diagonal crosses indicate protein degradation. Bmal1, basic helix-loop-helix ARNT like 1; CK1, casein kinase 1; Clock, circadian locomotor output cycles kaput protein; Cry, cryptochrome protein; GSK3, glycogen synthase kinase-3; Per, period protein; β-TrCP, beta-transducin repeat containing E3 ubiquitin–protein ligase.

Circadian rhythm protein degradation, the Fbx13 protein. The ubiquitin–proteasome protein degradation system is essential for various biological processes and is mediated by three enzymes working together: the ubiquitin-activating enzyme E1, the ubiquitin-conjugating enzyme E2, and the ubiquitin ligase E3 [21–24]. The latter makes ubiquitination specific by identifying the target substrates and mediating the transfer of ubiquitin from the ubiquitin-conjugating enzyme E2 to the substrate. The most well-known subfamily of the cullin-type E3 enzymes includes the Skp1-Cul1-F-box (SCF) complex [25]. It has four subunits: Skp1, Cul1, Rbx1, and the F-box protein with the characteristic spatial structure.

The Fbx13 protein (like Fbx121 below) is an F-box protein responsible for the ubiquitination and subsequent degradation of Cry cryptochromes, thus regulating the circadian rhythms. The formation of the SCF-Fbx13 complex is regulated by cryptochrome binding [26]. The domain repositioning analysis showed that the C-terminal leucine-rich-repeat domain of Fbx13 weakens the interaction with Skp1. This implies that a currently unknown protein interacts with the Fbx13 C-terminal domain of the SCF-Fbx13 complex formation.

To summarize, the SCF-Fbxl3 complex is responsible for the ubiquitin-dependent degradation of the Cry1 and Cry2 proteins, while Fbxl3 mutations extend the circadian clock cycle in mice [27–29]. Therefore, Fbxl3 is crucial for modulating Cry degradation during the circadian cycle.

Paralogous proteins β -TrCP and Fbxw11 (also known as β -TrCP1 and β -TrCP2) are targeted at Per and also cause its degradation, which depends on Per phosphorylation by casein kinases CK1 δ (Csnk1d) and CK1 ϵ (Csnk1e), [26].

The Fbxl21 protein (F-box and leucine-rich repeat protein 21) is closely related to circadian rhythm regulation [30–32]. Moreover, the expression of *Fbxl21* is tissue-specific. The highest *Fbxl21* expression is observed in the suprachiasmatic nucleus (SCN) of the hypothalamus [30,33,34], the Master Clock responsible for the biorhythms that govern the processes of aging and longevity [35,36].

The genomes of some species, including humans and gorillas, contain the pseudogenes related to the *Fbxl21* gene [6,7]. For instance, in humans, this pseudogene is transcribed and may be related to severe diseases, including schizophrenia development [37,38]. The *Fbxl21* gene is a paralog of *Fbxl3*, and their evolutionary divergence was relatively recent: unlike *Fbxl3*, *Fbxl21* has not been identified in fish genomes. In *Fbxl3* and *Fbxl21*, from *Mus musculus*, the identity of the amino acid sequences of the Cry-binding domain (CBD) is as high as 79%.

To summarize, *Fbxl21*, *Fbxl3*, β -TrCP, Fbxw11, GSK3 α , GSK3 β , CK1 δ , and CK1 ϵ interact with ubiquitin ligase adaptors, thus promoting the targeted proteasomal degradation. The concentration of the target proteins oscillates, establishing a circadian rhythm. Circadian clock disruption impairs sleep architecture and cognitive functions and can promote the development of neuropsychiatric disorders [14]. The GSK-3 β -FBXL21 functional axis controls Telethonin (also known as Titin cap or TCAP) degradation by forming the SCF complex and thus regulates skeletal muscle functions. The *Fbxl21* gene interacts with TCAP in a circadian manner, preventing its accumulation in skeletal muscle, while GSK3 β phosphorylates *Fbxl21* and TCAP, facilitating the *Fbxl21*-mediated, phosphodegron-dependent TCAP degradation. GSK3 β inhibition or knockdown reduces the formation of the *Fbxl21*-Cul1 complex and hampers *Fbxl21*-mediated TCAP degradation. *Fbxl21* hypomorphic mutation in *Psttm* mice disrupts circadian TCAP fluctuations. *Psttm* mice have substantial skeletal muscle defects, including abnormal fiber size, impaired exercise tolerance, reduced grip strength, and a response to glucocorticoid-induced muscle atrophy [39].

Although, in Laurasiatheria, we observe a considerable accumulation of the alterations in the *Fbxl21* gene, its protein product is also related to the species-specific lifespan (MRLS). For instance, the search for the genes related to longevity revealed four single-nucleotide polymorphisms associated with the extended lifespan of the Cane Corso dogs [40]. One of them, located in the third exon of *Fbxl21*, led to the substitution of arginine for tryptophan. This allele is associated with longevity in both homozygous and heterozygous forms.

Fbxl21 and Fbxl3 comparison. Although the amino acid sequences of mouse proteins *Fbxl21* and *Fbxl3* exhibit high similarity, they play nonidentical roles in circadian rhythm regulation [32]. *Fbxl21* performs a dual function: it protects Cry from *Fbxl3*-mediated degradation in the nucleus and, at the same time, it promotes Cry degradation in the cytoplasm, ensuring the balance and cell compartmentalization of E3 ligases that compete for Cry-binding and fine-tuning the circadian period in mammals [41].

Along with *Fbxl21*, there are other mechanisms underlying circadian regulation. For instance, *TARDBP* encodes the TAR DNA-binding protein 43 (TDP-43), which stabilizes the Cry cryptochrome and suppresses *Fbxl3* activity [32]. It is present in humans (GeneID 23435) and most mammals, which explains the existence of many species lacking *Fbxl21*. However, it is still unclear if *TARDBP* can fully compensate for the absence of *Fbxl21*.

This article proposes a methodology that establishes a relationship between the rate of change in a given gene (relative to a given taxon) together with the amino acid composition of its protein and the traits of the species harboring the gene. The study goes beyond the protein-coding genes, since the accumulation of amino acid substitutions during evolution

may lead to pseudogenization or even gene loss, although the relationship between genomic characteristics and the species-specific traits remains preserved. The methodology is illustrated by the mammalian genes involved in the regulation of circadian rhythms, which contribute to a number of human disorders, particularly those associated with aging. The approach is applied to the following mammalian genes: *Btrc*, *Fbxw11* (*Btrc2*), *Fbxl21*, *Fbxl3*, *Gsk3a*, *Gsk3b*, *Csnk1d*, *Csnk1e*.

Therefore, in this paper, we examine the evolution of a panel of the proteins responsible for circadian rhythm regulation in placental mammals. Specifically, we established the association of the accumulated substitutions in the sequences of these proteins and the genes encoding them, including the related evolutionary events with the species-specific lifespan, body weight, and longevity quotient (LQ). In particular, here, we test the hypothesis that the accumulation of substitutions and even pseudogenization and loss of the *Fbxl21* gene are selectively associated (or not) in terms of MRLS and body mass in large taxonomic groups (level of superorders). In many species with high MRLS and body mass values, including humans, gorillas, and elephants, a significant change in this gene is also associated with a significant change in closely related species. In other words, this gene in Euarchontoglires and Afrotheria is under positive selection pressure, and in Laurasiatheria and Xenarthra, it is under stabilizing selection. Also, the second hypothesis runs as follows: the role of a single amino acid substitution in the Cry-binding domain of *Fbxl21* is correlated with MRLS and body mass.

2. Materials and Methods

The data on the species-specific maximal reported lifespan (MRLS), body weight, longevity quotient (LQ), and reproductive performance were obtained from the AnAge database [42]; for *Galeopterus variegatus*, the estimated MRLS was 17.5 years [43]. Certain Chiroptera have high LQ and MRLS, as well as low body weight, which is particularly typical of *Myotis* (microbats) [44]. This characteristic is a distinctive feature of the order.

We studied 158 species of placental mammals from 18 orders. They comprise the existing representatives of Laurasiatheria, which include the following orders: Eulipotyphla, Chiroptera, Perissodactyla, Cetartiodactyla, Pholidota, and Carnivora. We also examined Euarchontoglires, which include the following orders: Lagomorpha, Rodentia, Scandentia, Dermoptera, and Primates. The Afrotheria superorder includes the following orders: Tubulidentata, Macroscelidea, Afrosoricida, Hyracoidea, Proboscidea, and Sirenia. The following number of species was considered in the above-mentioned orders: 1 species from Afrosoricida, 28 species from Artiodactyla, 36 species from Carnivora, 17 species from Chiroptera, 1 species from Cingulata, 1 species from Dermoptera, 3 species from Eulipotyphla, 2 species from Lagomorpha, 1 species from Macroscelidea, 4 species from Perissodactyla, 1 species from Pholidota, 1 species from Pilosa, 30 species from Primates, 1 species from Proboscidea, 28 species from Rodentia, 1 species from Scandentia, 1 species from Sirenia, and 1 species from Tubulidentata. The species were selected based on the quality of genome assembly and the availability of robust data on MRLS and body weight, which are unknown even for a number of species with a well-annotated genome. The sequences of proteins and transcripts have been obtained from the RefSeq database [45].

The search for orthologs, pseudogenes, and the genes lost during evolution was performed using the original methods described in [5–8]. These methods are based on homology and local synteny, along with computational parameters predicting orthology; the above-mentioned publications discuss the advantages (in our opinion) of these methods and the software used.

For each gene, we performed multiple sequence alignment (MSA) of the nucleotide sequences that encode the proteins under scrutiny, taking into account the codons in the open reading frames. We removed short sequences, low-quality proteins, and columns with deletions. Using the method and the software from the PAML package [46–48], we calculated the number of synonymous nucleotide substitutions dS (per synonymous site) and the number of nonsynonymous nucleotide substitutions dN (per nonsynonymous site).

For each gene and all its orthologues in the specific taxonomic group, we calculated the dN/dS ratio median. This median can be considered an estimate of the dN/dS value for a particular gene and its common ancestor for all species in a taxonomic group. We interpret it as the relative rate of accumulation of substitutions and changes in a gene compared with this taxon up to the possible loss of a protein-coding gene.

We analyzed the nonannotated genome [49] of *Lepus timidus* by alignment of protein sequences with the in silico translated DNA according to the codons, using the Muscle method and the MEGA 7 software [50]. Scatter plots were prepared using the Seaborn library [51]. The protein tree was built using the Fast Minimum Evolution method and the COBALT web tool (the constraint-based multiple alignment tool).

3. Results and Discussion

3.1. The Relationship between MRLS (and Body Weight) and the Substitution Accumulation Rate in the Genes Involved in Circadian Rhythm Regulation

In the introduction, we described the important function of the following mammalian genes: *Btrc*, *Fbxw11* (*Btrc2*), *Fbxl21*, *Fbxl3*, *Gsk3a*, *Gsk3b*, *Csnk1d*, *Csnk1e*. Particularly, they are involved in circadian rhythm regulation [1]. We examined a possible relationship between the substitution accumulation rate in these genes (including the formation of pseudogenes or even gene loss) and the MRLS and body weight. We used the median of the dN/dS values of a gene in relation to its orthologs in other species from the same mammalian taxon as the quantitative measure of the relative substitution accumulation rate in this gene. Further, we refer to it as “median (dN/dS)”, or simply “median”. Using the panel of the above-mentioned genes involved in circadian rhythm regulation, we developed an approach for studying the relationship between the characteristics of the genome and the species and further employed it to determine the relationship between the median and amino acid composition of a protein product with MRLS and body weight.

In the Euarchontoglires superorder which includes rodents and the Euarchonta grandorder, the variation interval of the *Fbxl3* gene median (depending on a gene) is much smaller than that of the *Fbxl21* gene: for *Fbxl3* and *Fbxl21*, it is [0.023, 0.054] and [0.074, 0.178], respectively (numbers in square brackets indicate the minimum and maximum of the median), which implies that the *Fbxl21* is far more sensitive to the changes in MRLS and typical adult body weight than its paralog *Fbxl3*, see Tables 1–4 (the “median interval” column). Similarly, *Fbxw11*, the most conservative gene among the eight genes under scrutiny, as well as *Csnk1d*, has narrow variation intervals of the median: [0, 0.012] and [0.004, 0.032], respectively.

For the Euarchontoglires superorder, or for the unified Euarchontoglires and Afrotheria superorders, and even for placental mammals, we determined the incremental statistically significant dependencies between the *Fbxl21* median and MRLS (linear regression) as well as between the *Fbxl21* median and the typical adult body weight (exponential regression). Calculated p -values are: 10^{-10} in the first case (for MRLS) and 10^{-7} (for body weight); 8×10^{-11} (MRLS) and 7×10^{-8} (body weight) in the second case; 10^{-6} (MRLS) and 10^{-8} (body weight) in the third case (Tables 1–4). It should be noted that 51 out of 61 Euarchontoglires species have the functional *Fbxl21* protein, although we have no data on MRLS for *Microtus oregoni* and body weight for *Grammomys surdaster*.

We observed a unique/uncommon dependency for *Fbxw11*. In Euarchontoglires, its p -value for MRLS is 10^{-3} despite the lack of correlation in the case of body weight, whereas both correlations were absent in the unified superorders. However, in placental mammals, the p -value is 10^{-6} both for MRLS and body weight. No such relationships were observed for *Fbxl3*. We also confirmed the absence of these correlations in Laurasiatheria: the p -value was about 1 for all eight genes, except for *Fbxw11* and body weight, with a p -value equal to 7×10^{-4} . These results are shown in Tables 1–4; for *Fbxl21* and *Fbxl3*, data are also presented in Figure 2. Figure 2 also shows the uncertainty in the position of the regression line. The data used for calculations for *Fbxl21* and *Fbxl3* in Euarchontoglires and Laurasiatheria are shown in scatter plots in Figures 3 and 4. In Euarchontoglires, the median

correlates with the maximal reported lifespan and body weight for the *Fbxl21* gene, but such a correlation is absent for other genes. A correlation is also absent in Laurasiatheria.

Table 1. In the *Fbxl21*, *Fbxl3*, *Fbxw11*, *Btrc*, *Gsk3a*, *Gsk3b*, *Csnk1d*, *Csnk1e* genes from the Euarchontoglires superorder, for the species-specific maximal reported lifespan (MRLS) and the common logarithm of the body weight, on the one hand, and the median, on the other hand, the following data are presented: the gene name; number *n* of the points used in linear regression; the linear correlation coefficient *r*; *p*-value (the probability of obtaining test results at least as extreme as the result actually observed, under the assumption that the null hypothesis is correct) assuming the lack of relationship between *y* and *x* (the null hypothesis); the linear regression equation $y = ax + b$ (as the mean-square approximation of a given point set, Figures 2–4), and the median variation interval (minimum and maximum values). Here, *y* indicates MRLS in years or the body weight logarithm in grams, while *x* indicates the median. Different scales of the coordinate axes explain the unusually high value of the coefficient of *x*. A highly significant correlation of the species characteristics, MRLS and body weight, with the proposed genomic characteristic—the median for the gene *Fbxl21* and a significant absence of such a correlation for other genes of the circadian rhythm are shown. The numerical characteristics of these correlations were also obtained.

Gene	<i>n</i>	<i>r</i>	<i>p</i>	Regression	Median Interval
median vs. lifespan					
<i>Fbxl21</i>	49	0.76	3×10^{-10}	$y = 439x - 36$	[0.074, 0.178]
<i>Fbxl3</i>	58	−0.15	0.3	$y = -463x + 42$	[0.023, 0.054]
<i>Fbxw11</i>	54	−0.37	0.006	$y = -1974x + 32$	[0.000, 0.012]
<i>Btrc</i>	56	0.12	0.4	$y = 34x + 23$	[0.008, 0.265]
<i>Gsk3a</i>	52	−0.11	0.4	$y = -135x + 30$	[0.022, 0.147]
<i>Gsk3b</i>	52	−0.17	0.2	$y = -102x + 28$	[0.006, 0.225]
<i>Csnk1d</i>	56	0.14	0.3	$y = 371x + 21$	[0.004, 0.032]
<i>Csnk1e</i>	47	0.02	0.9	$y = 10x + 25$	[0.003, 0.136]
median vs. lg(weight)					
<i>Fbxl21</i>	49	0.67	10^{-7}	$y = 25x - 1$	[0.074, 0.178]
<i>Fbxl3</i>	58	−0.17	0.2	$y = -28x + 4$	[0.023, 0.054]
<i>Fbxw11</i>	54	−0.35	0.01	$y = -99x + 3$	[0.000, 0.012]
<i>Btrc</i>	56	0.22	0.1	$y = 3x + 3$	[0.008, 0.265]
<i>Gsk3a</i>	52	−0.18	0.2	$y = -12x + 3$	[0.022, 0.147]
<i>Gsk3b</i>	52	−0.04	0.8	$y = -x + 3$	[0.006, 0.225]
<i>Csnk1d</i>	55	−0.04	0.8	$y = -6x + 3$	[0.004, 0.032]
<i>Csnk1e</i>	47	0.11	0.5	$y = 3x + 3$	[0.003, 0.136]

Table 2. The data and results for unified Euarchontoglires and Afrotheria superorders; for designations, see the description for Table 1.

Gene	<i>n</i>	<i>r</i>	<i>p</i>	Regression
median vs. lifespan				
<i>Fbxl21</i>	53	0.75	8×10^{-11}	$y = 463x - 39$
<i>Fbxl3</i>	63	−0.08	0.6	$y = -249x + 35$
<i>Fbxw11</i>	58	−0.35	0.007	$y = -2016x + 33$
<i>Btrc</i>	61	0.10	0.5	$y = 31x + 25$
<i>Gsk3a</i>	56	−0.05	0.7	$y = -78x + 30$
<i>Gsk3b</i>	57	−0.18	0.2	$y = -115x + 29$
<i>Csnk1d</i>	60	0.18	0.2	$y = 483x + 20$
<i>Csnk1e</i>	48	0.02	0.9	$y = 15x + 25$

Table 2. Cont.

Gene	<i>n</i>	<i>r</i>	<i>p</i>	Regression
median vs. lg (weight)				
<i>Fbxl21</i>	53	0.66	7×10^{-8}	$y = 27x - 1$
<i>Fbxl3</i>	63	-0.08	0.5	$y = -15x + 4$
<i>Fbxw11</i>	58	-0.34	0.01	$y = -111x + 3$
<i>Btrc</i>	61	0.16	0.2	$y = 3x + 3$
<i>Gsk3a</i>	56	-0.10	0.5	$y = -8x + 3$
<i>Gsk3b</i>	57	-0.07	0.6	$y = -2x + 3$
<i>Csnk1d</i>	59	0.03	0.8	$y = 5x + 3$
<i>Csnk1e</i>	48	0.11	0.4	$y = 4x + 3$

Table 3. The data and results for the Laurasiatheria superorder; for designations, see the description for Table 1.

Gene	<i>n</i>	<i>r</i>	<i>p</i>	Regression
median vs. lifespan				
<i>Fbxl21</i>	67	0.11	0.4	$y = 58x + 16$
<i>Fbxl3</i>	78	0.07	0.5	$y = 105x + 27$
<i>Fbxw11</i>	76	-0.29	0.01	$y = -4552x + 32$
<i>Btrc</i>	77	-0.08	0.5	$y = -40x + 31$
<i>Gsk3a</i>	75	0.33	0.004	$y = 339x + 17$
<i>Gsk3b</i>	77	0.38	0.0008	$y = 79x + 26$
<i>Csnk1d</i>	78	-0.10	0.4	$y = -138x + 34$
<i>Csnk1e</i>	42	0.29	0.06	$y = 211x + 8$
median vs. lg (weight)				
<i>Fbxl21</i>	67	0.26	0.03	$y = 13x + 1$
<i>Fbxl3</i>	78	-0.19	0.1	$y = -26x + 5$
<i>Fbxw11</i>	76	-0.38	0.0007	$y = -686x + 4$
<i>Btrc</i>	78	0.01	0.9	$y = x + 4$
<i>Gsk3a</i>	75	0.12	0.3	$y = 12x + 4$
<i>Gsk3b</i>	77	0.10	0.4	$y = 2x + 4$
<i>Csnk1d</i>	77	0.10	0.4	$y = 14x + 4$
<i>Csnk1e</i>	43	0.30	0.05	$y = 24x + 2$

Table 4. The data and results for all placental mammals; for designations, see the description for Table 1.

Gene	<i>n</i>	<i>r</i>	<i>p</i>	Regression
median vs. lifespan				
<i>Fbxl21</i>	120	0.42	2×10^{-6}	$y = 166x - 4$
<i>Fbxl3</i>	142	-0.00	1	$y = -6x + 29$
<i>Fbxw11</i>	135	-0.38	4×10^{-6}	$y = -2877x + 33$
<i>Btrc</i>	140	0.01	0.9	$y = 3x + 29$
<i>Gsk3a</i>	132	0.22	0.01	$y = 281x + 18$
<i>Gsk3b</i>	136	0.22	0.01	$y = 75x + 26$
<i>Csnk1d</i>	140	0.01	0.9	$y = 19x + 29$
<i>Csnk1e</i>	91	0.02	0.8	$y = 13x + 27$
median vs. lg (weight)				
<i>Fbxl21</i>	120	0.48	4×10^{-8}	$y = 17x + 0$
<i>Fbxl3</i>	142	-0.09	0.3	$y = -16x + 4$
<i>Fbxw11</i>	135	-0.40	2×10^{-6}	$y = -262x + 4$
<i>Btrc</i>	141	-0.02	0.8	$y = -x + 4$
<i>Gsk3a</i>	132	0.14	0.1	$y = 16x + 3$
<i>Gsk3b</i>	136	0.15	0.08	$y = 4x + 4$
<i>Csnk1d</i>	138	0.11	0.2	$y = 20x + 4$
<i>Csnk1e</i>	92	0.12	0.3	$y = 5x + 3$

We can conclude that for *Fbxl21* the stabilizing selection prevails in Laurasiatheria whereas positive selection dominates in Euarchontoglires. Without any doubt, MRLS and body weight correlate in most mammal species.

Among Euarchontoglires that have the functional *Fbxl21* protein, we revealed the statistically significant dependency between the relative substitution rate (the median) and MRLS (or body weight). On the contrary, we detected no dependency of the kind for the other examined genes, which are also involved in regulating circadian rhythms.

We would like to mention once again that Tables 1–4 show the data and results for Euarchontoglires, for unified Euarchontoglires and Afrotheria, for Laurasiatheria, and for all placental mammals, respectively.

Figure 2 shows the obtained linear regressions for the median and MRLS (A), as well as for the median and the common logarithm of body weight (B) in the *Fbxl21* and *Fbxl3* genes from the Euarchontoglires superorder.

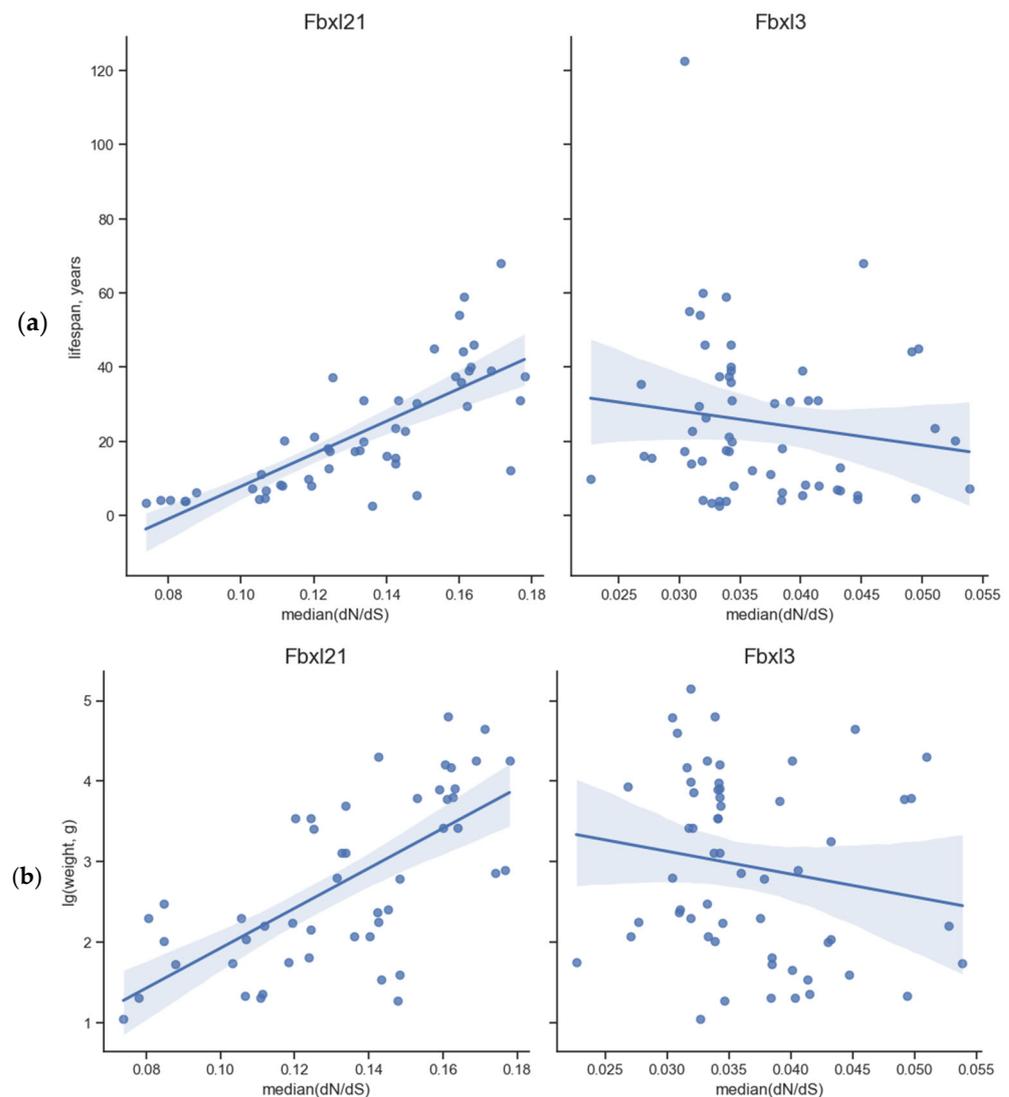


Figure 2. In the *Fbxl21* and *Fbxl3* genes from the Euarchontoglires superorder, for the species-specific maximal reported lifespan (MRLS) and the common logarithm of body weight, linear regressions (the lines with their estimated uncertainty) for the median and MRLS (a), as well as for the median and the common logarithm of body weight (b), are shown. Each dot represents an *Fbxl21* or *Fbxl3* gene from one animal species. The correlation between the considered species traits and the genomic characteristic—the median, which for *Fbxl3* does not seem reliable enough due to the ambiguity of the regression line and small interval of change in the median, is shown.

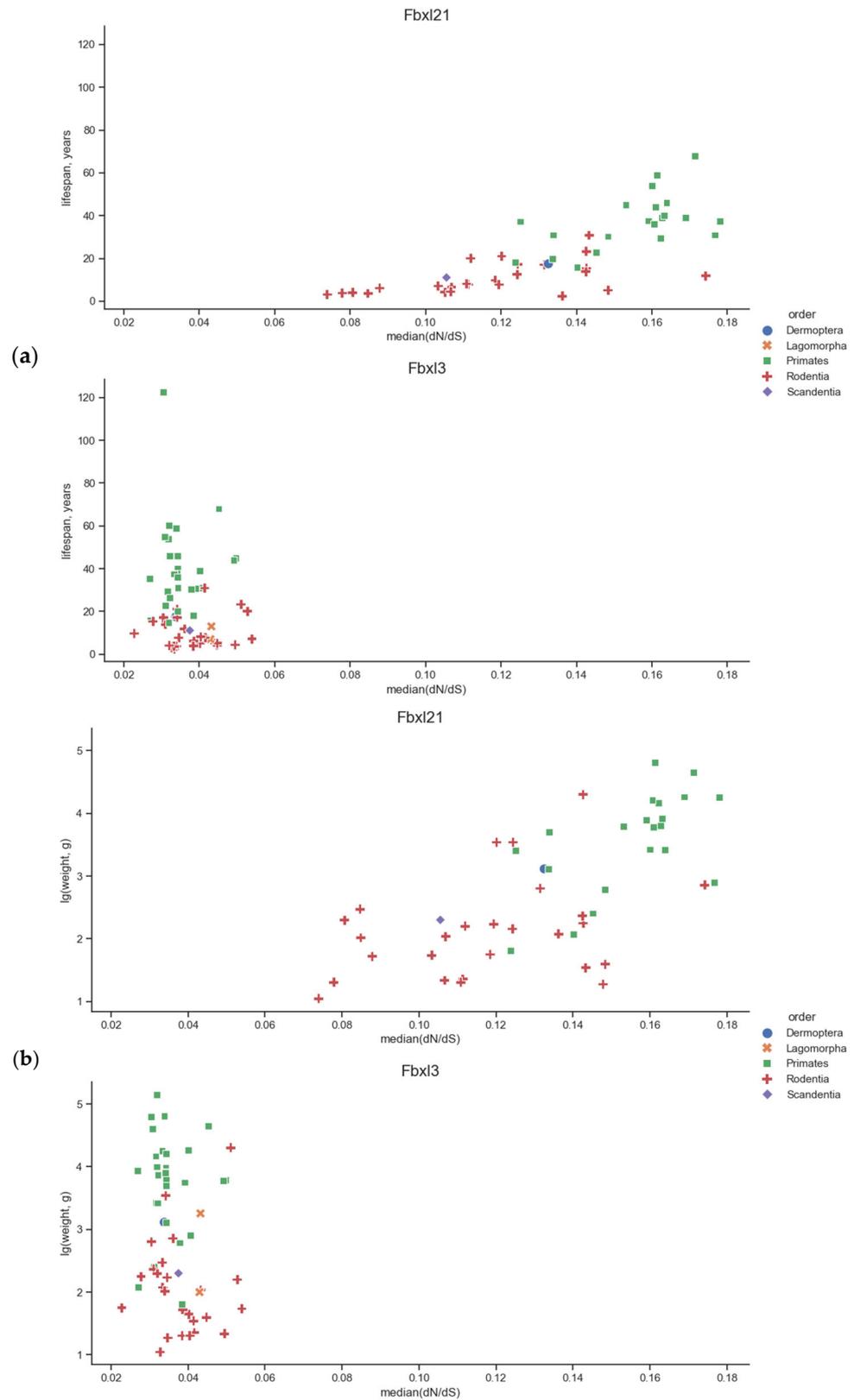


Figure 3. Scatter plots of the median and MRLS (a) and the common logarithm of body weight (b) for the *Fbxl21* and *Fbxl3* genes from the Euarchontoglires. Each dot represents an *Fbxl21* or *Fbxl3* gene from one animal species. The taxa are denoted by shape and color and are listed in the legend. The correlation between these species traits and the median is shown; see the description in Figure 2.

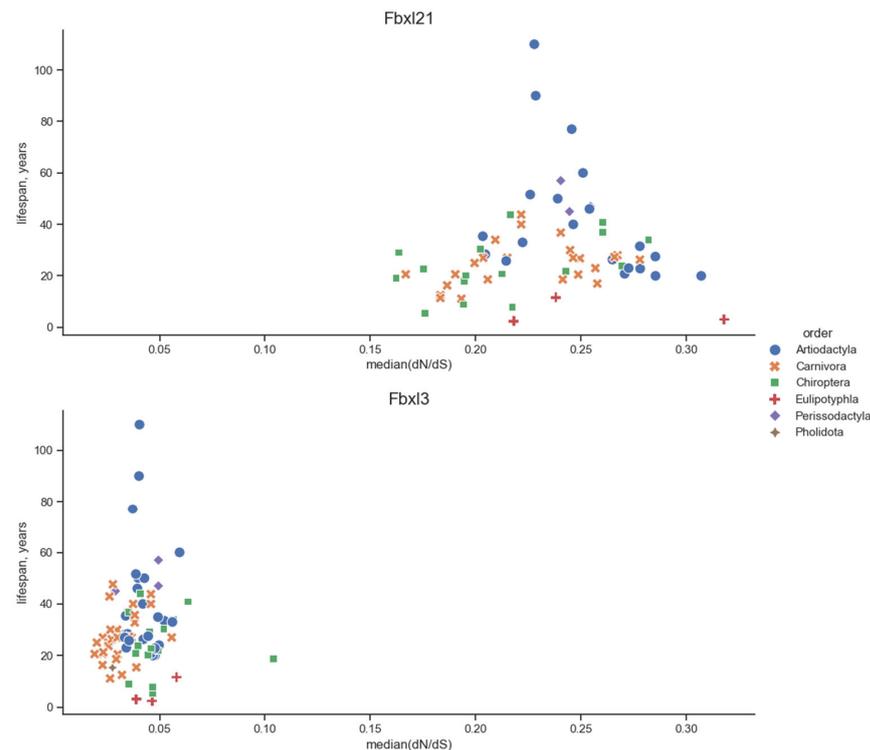


Figure 4. Scatter plots of the median and MRLS for the *Fbx121* and *Fbx13* genes in Cetartiodactyla (as a part of the Laurasiatheria superorder). Each dot represents an *Fbx121* or *Fbx13* gene from one animal species. The taxa are denoted with the shape and color and are listed in the legend. The correlation between the data on these species traits and the median is shown; see the description from Figure 2.

3.2. Pseudogenization and *Fbx121* Gene Loss

The relationship between the median and MRLS (or body weight) implies that in Euarchontoglires, with an increase in MRLS and body weight, as well as with changes in species-specific traits, important proteins (such as *Fbx121*) accumulate mutations that can alter their function which may lead to pseudogenization and a subsequent gene loss.

The human pseudogene *FBXL21P* is a counterpart of the *Fbx121* gene, which is confirmed by local synteny: conservative genes *IL9* and *LECT2* are closely located to the pseudogene [5,6,8]; *LECT2* is located on the complementary DNA strand and overlaps with *FBXL21P*. Similar findings are reported for the *IL9*, *FBXL21*, and *LECT2* genes in many other placental mammals.

In long-lived species, including humans, gorillas, and *Loxodonta africana* (Afrotheria, Proboscidea), *Fbx121* became a pseudogene. Interestingly, in *Trichechus manatus* (Afrotheria, Sirenia) with MRLS of 69 years and body weight of 322 kg, the cysteine in the Cry-binding domain was substituted by tyrosine (C364Y mutation), resulting in substantial changes in protein properties. In this regard, the data on *Fbx121* evolution in large Afrotheria (*Loxodonta africana* and *Trichechus manatus*) is in line with the results from human *Fbx121*. It should be noted that in *Gorilla gorilla*, this pseudogene corresponds to the human pseudogene *FBXL21P*. In *Loxodonta africana*, *Fbx121* contains a termination codon in the Cry-binding domain at the consensus position corresponding to the leucine residue L370, similar to its flanking genes *IL9* and *LECT2*. Proboscideans have the longest MRLS of all Afrotheria, being third only to humans and baleen whales among mammals.

The human pseudogene *FBXL21P* has a matching gene (GeneID 101436012) in *Dasypus novemcinctus* (Xenarthra, Cingulata) with two noncoding LincRNA type transcripts (however, we found only the ENSDN0G00000052722 protein in the Ensembl database and a low-quality protein XP_023439683 in RefSeq, as well as the flanking gene *LECT2*). The *FBXL21P* pseudogene in *Oryctolagus cuniculus* has a matching gene (ENSOCUG00000030764) with two noncoding LincRNA-type transcripts. The same human pseudogene has a matching

pseudogene with termination codons in *Lepus timidus*. The *Fbxl21* gene is absent in all known marsupials (the Marsupialia infraclass) and monotremes (the Monotremata order). It should be noted that the annotations in RefSeq sometimes include low-quality proteins such as Fbxl21p, in which the protein-coding region is terminated by a stop codon.

The lagomorphs do not follow the trend observed in Euarchontoglires: despite its MRLS and low body weight (the MRLS ranges between 7 and 18 years, and the body weight ranges between 100 g and 4.175 kg), lagomorph *Fbxl21* is a pseudogene. Typically, lagomorphs mostly follow the *r*-strategy [52,53], characterized by short life cycles and high fertility rates. However, there has been plentiful evidence of the animals being twice both the size and MRLS, such as rabbits [54]. In modern habitats, longer MRLS or higher body weight in lagomorphs is likely not adaptive and is not under positive selection pressure. Therefore, the deviation described above might be related to population factors.

The assumption formulated at the beginning of this section is in some ways true for anthropoid apes and the Cercopithecoidea that represent a small part of the Euarchontoglires, which is interesting. Pseudogenization or, on the contrary, few substitutions in *Fbxl21* are typical of most species in this group. Specifically, the *Fbxl21* gene underwent pseudogenization only in humans and gorillas (with long MRLS and high body weight). In contrast, other species belonging to the anthropoid apes and all Cercopithecoidea have *Fbxl21*-like proteins containing a few substitutions (Figure 5). Although this finding seems to contradict our hypothesis, the *Fbxl21* protein tree is not consistent with the phylogenetic tree, which divides the protein tree into clades with relatively high and low MRLS. The MRLS of chimpanzees is about 68 years (*Pan troglodytes* and *Pan paniscus* have identical proteins); *Pongo abelii*, which belongs to another clade, has an MRLS of 59 years. If we root the protein tree based on the root of the phylogenetic tree, we observe a high number of substitutions (distance) from the new root to *Pan* sp. (Figure 5). *Pan* sp. and, to a smaller extent, *Pongo*, represent outgroups in the phylogenetic tree; eliminating them from the protein tree makes it topologically close to the phylogenetic tree. In other words, in chimpanzees, the *Fbxl21* protein differs considerably from the orthologous proteins in *Pongo abelii*, *Nomascus leucogenys*, *Hylobates moloch*, and all Cercopithecoidea available for analysis (Figure 5). The common ancestor of chimpanzees has likely changed the direction of evolutionary changes, at least in the case of certain functions related to *Fbxl21*.

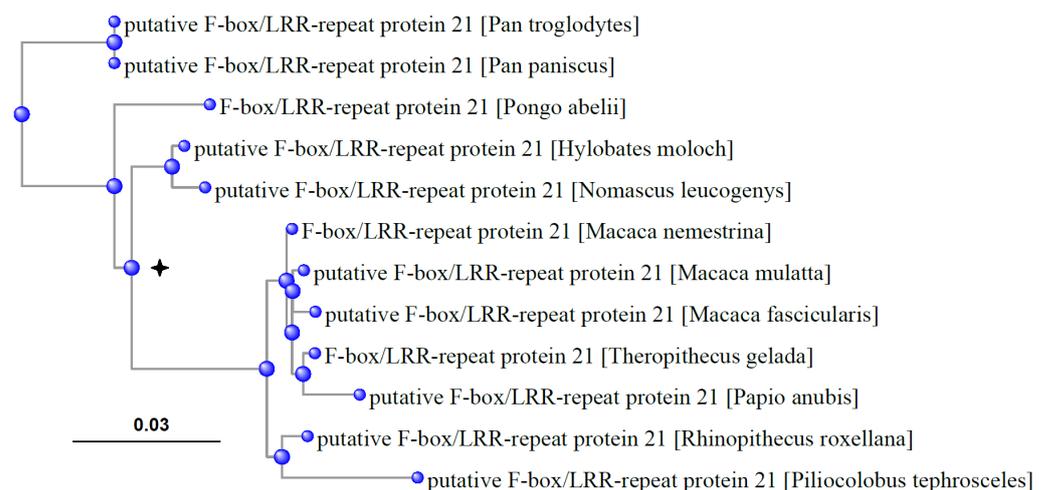


Figure 5. The protein tree of *Fbxl21* in the anthropoid apes and Cercopithecoidea (all species represented in RefSeq that have this protein). In the common chimpanzee and bonobo, the *Fbxl21* protein sequences are identical. The node corresponding to the root of the phylogenetic tree is marked with a cross. The “0.03” above the bar represents the scale of the branch lengths, which corresponds to the expected number of substitutions per site. Specifically, a branch length of 0.03 indicates an expected 3% difference in amino acid sequence between the nodes connected by that branch. The peculiarity of *Fbxl21* gene evolution in apes is shown.

3.3. Amino Acid Substitutions in the Fbxl21 Protein

The difference in the occurrence of amino acids in the Fbxl21 protein between three superorders (Euarchontoglires, Laurasiatheria, and all placental mammals) is less than 0.5%. The amino acid occurrence is shown in Figure 6a.

(a)																			
L	S	V	E	I	A	K	D	F	R	P	N	T	H	G	Q	C	Y	M	W
12	11	9	6	6	5	5	5	5	4	4	4	4	4	3	3	3	2	2	1

(b)							(c)						
	Euarchontoglires		Laurasiatheria		Placentalia			Euarchontoglires		Laurasiatheria		Placentalia	
	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>		<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>
L	0.62	3×10^{-7}	0.23	0.06	0.37	10^{-5}	L	0.51	6×10^{-5}	0.45	0.06	0.43	3×10^{-7}
S	0.23	0.09	-0.06	0.63	0.07	0.43	S	0.29	0.03	0.06	0.63	0.19	0.03
V	-0.38	4×10^{-3}	0.23	0.06	-0.05	0.6	V	-0.22	0.09	0.37	0.06	0.06	0.52
E	0.57	4×10^{-6}	-0.1	0.42	0.09	0.28	E	0.49	10^{-4}	-0.09	0.42	-0.03	0.72
I	0.17	0.21	-0.32	0.01	-0.13	0.15	I	0.08	0.54	-0.37	0.01	-0.19	0.03
A	-0.17	0.2	0.03	0.81	-0.02	0.85	A	-0.33	0.01	0.11	0.81	0.02	0.82
K	0.02	0.86	-0.01	0.91	0.03	0.7	K	0.06	0.67	-0.17	0.91	-0.03	0.7
D	-0.08	0.57	0.14	0.23	0.12	0.16	D	0.02	0.89	0.09	0.23	0.14	0.1
F	0.19	0.16	-0.19	0.11	-0.06	0.52	F	0.24	0.07	-0.56	0.11	-0.29	6×10^{-4}
R	-0.4	2×10^{-3}	0.15	0.21	-0.16	0.06	R	-0.44	7×10^{-4}	0.33	0.21	-0.05	0.59
P	-0.69	2×10^{-9}	0.04	0.73	-0.41	10^{-6}	P	-0.57	3×10^{-6}	-0.3	0.73	-0.49	3×10^{-9}
N	0.25	0.06	0.25	0.04	0.28	10^{-3}	N	0.39	3×10^{-3}	0.14	0.04	0.31	3×10^{-4}
T	-0.15	0.27	-0.11	0.36	-0.1	0.24	T	0.03	0.84	-0.12	0.36	-0.01	0.87
H	0.02	0.88	-0.44	10^{-4}	-0.3	6×10^{-4}	H	-0.22	0.1	-0.02	10^{-4}	-0.25	3×10^{-3}
G	-0.13	0.32	-0.04	0.77	0.03	0.75	G	-0.23	0.08	0.15	0.77	0.16	0.06
Q	-0.11	0.43	0.04	0.75	-0.01	0.87	Q	0.01	0.93	-0.19	0.75	-0.14	0.11
C	0.23	0.09	-0.15	0.22	0.01	0.86	C	0.24	0.07	0.06	0.22	0.09	0.29
Y	0.38	4×10^{-3}	0.23	0.06	0.36	2×10^{-5}	Y	0.37	4×10^{-3}	-0.03	0.06	0.23	0.01
M	0.01	0.95	-0.03	0.82	0.07	0.41	M	-0.17	0.2	-0.18	0.82	0.01	0.9
W	-0.26	0.05	0.03	0.79	-0.17	0.05	W	-0.23	0.09	0.04	0.79	-0.16	0.07

Figure 6. (a) The occurrence (%) of amino acids in the Fbxl21 protein, averaged by species. It is similar for Euarchontoglires, Laurasiatheria, and Placentalia. 20 amino acids are listed in the first row. (b) For the three superorders, Euarchontoglires, Laurasiatheria, and all placental mammals, the Pearson correlation coefficient *r* and the *p*-value of the amino acid occurrence in Fbxl21 with the species-specific lifespan are shown. (c) shows the common logarithm of the body weight; see (B) for a description. Correlation characteristics of the relationship between the frequency of each amino acid in the Fbxl21 protein in the indicated taxa and the species traits were obtained.

In the Euarchontoglires superorder, MRLS exhibited a significant correlation with the occurrence of certain amino acids in the Fbxl21 protein sequence. For amino acids, the Pearson correlation coefficient *r* and *p*-value of the dependency are given in Figure 6b (for MRLS). In Euarchontoglires, we observed the preference (the positive *r*, MRLS rises with growing occurrence) of leucine (*p*-value is 10^{-7}), glutamic acid (*p*-value is 10^{-6}), and tyrosine (*p*-value is 10^{-3}) and avoidance (the negative *r*, MRLS declines with growing occurrence) of valine (*p*-value is 10^{-3}), arginine (*p*-value is 10^{-3}), and proline (*p*-value is 10^{-9}). The most substantial preference and avoidance levels were observed for leucine and proline. This correlation does not depend on the amino acid occurrence, since we did not detect this correlation in the case of the amino acids with similar occurrence. The other amino acids in Euarchontoglires, as well as all amino acids in Laurasiatheria (the *p*-value is about 1), except histidine (with a *p*-value of 10^{-4}), did not reveal any correlation of the kind. We obtained similar results (Figure 6c) for the amino acid occurrence and the body

anthropoid apes (except humans and gorillas that have *Fbxl21* as a pseudogene), as well as most cetaceans (except for *Physeter catodon*). The deviation for the lagomorphs was discussed above in the “Pseudogenization and *Fbxl21* Gene Loss” section.

The nature of the amino acid that substitutes another amino acid in a functionally important domain is of high relevance. We assigned six amino acid groups [55]. The first group contained cysteine; the second one, serine, threonine, alanine, glycine, and proline; the third one, aspartic acid, glutamic acid, asparagine, and glutamine; the fourth one, histidine, arginine, and lysine; the fifth one, methionine, isoleucine, leucine, and valine; the sixth one, tryptophan, tyrosine, and phenylalanine. We use the term “radical substitution” for the substitution that alters an initial amino acid for another one from the other group. The lower part of Figure 7 shows the number of radical substitutions for each taxon listed in this part. The 19th domain position is the least conservative in these taxa, since radical substitutions in this position occurred in 52 species. In these species, the median LQ exceeds the median LQ for the remaining 94 species of placental mammals by 35%. This indicates the correlation between MRLS and the 19th position of the Cry-binding domain of the *Fbxl21* protein. Therefore, we can conclude that there is some relationship between MRLS and amino acid substitution in this domain and in *Fbxl21* in general.

4. Conclusions

- (1) The proposed method consists in determining the relationship between the species traits and a genomic characteristic, harnessing the methods of statistical analysis. The genomic characteristic is considered as the median dN/dS ratio.
- (2) For certain genes (such as *Fbxl21*), the accumulation of amino acid substitutions up to pseudogenization or gene loss, as well as the preference for certain amino acids in the encoded protein, is an effective way to achieve a significant phylogenetic change.
- (3) The *Fbxl21* gene and the species-specific maximal reported lifespan (MRLS), together with body weight, are examples of such a phylogenetic change in Euarchontoglires and Afrotheria, which is also observed in relatively small taxonomic groups, as, for example, in anthropoid apes and the Cercopithecidae.
- (4) In the *Fbxl21* protein, changes were identified not only in the gene change rate for orders and superorders of mammals but also in the nature of the accumulated amino acid substitutions. For this protein, the stabilizing selection and the positive selection clearly prevail in Laurasiatheria (a sufficiently large p -value indicates the absence of correlation) and Euarchontoglires together with Afrotheria, where the p -value approximately equals 10^{-10} , respectively.
- (5) In contrast, for proteins such as *Fbxl3*, β -TrCP, *Fbxw11*, *GSK3 α* and *GSK3 β* , *CK1 δ* , and *CK1 ϵ* , which are also closely related to circadian rhythm regulation in mammals, the stabilizing selection is characteristic of all mentioned superorders. The *Bmal1*, *Clock*, *Pers*, and *Cry*s proteins, like many other circadian rhythm proteins, are highly conserved. For example, when cryptochrome circadian regulator 1 proteins were aligned in humans (586 aa) and mice (606 aa), 96% of the human protein and 93% of the mouse protein matched. The authors focused on studying circadian rhythm proteins, which change markedly during evolution.
- (6) We proposed a methodology to study the relationship between any genomic characteristics and species traits and illustrated its application on a number of circadian rhythm proteins. The authors hope that this methodology may be useful in other circadian rhythm proteins and within a different context.

Author Contributions: Conceptualization, V.A.L.; data curation, A.V.S. and O.A.Z.; formal analysis, G.A.S., A.V.S., J.-R.Y. and O.A.Z.; funding acquisition, V.A.L.; Investigation, V.A.L.; Methodology, and V.A.L.; project administration, V.A.L.; supervision, V.A.L.; validation, G.A.S., J.-R.Y. and A.V.S.; visualization, A.V.S. and O.A.Z.; writing—original draft, V.A.L.; writing—review and editing, G.A.S., A.V.S., J.-R.Y. and O.A.Z. All authors have read and agreed to the published version of the manuscript.

Funding: The research was supported by the Russian Science Foundation grant No. 24-44-00099, <https://rscf.ru/en/project/24-44-00099/> (accessed on 1 September 2024).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

Acknowledgments: Computations were performed at the Joint Supercomputer Center of the Russian Academy of Sciences (JSCC RAS).

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Bonaconsa, M.; Malpeli, G.; Montaruli, A.; Carandente, F.; Grassi-Zucconi, G.; Bentivoglio, M. Differential modulation of clock gene expression in the suprachiasmatic nucleus, liver and heart of aged mice. *Exp. Gerontol.* **2014**, *5*, 7079. [[CrossRef](#)] [[PubMed](#)]
2. Kolarski, D.; Miró-Vinyals, C.; Sugiyama, A.; Srivastava, A.; Ono, D.; Nagai, Y.; Iida, M.; Itami, K.; Tama, F.; Szymanski, W.; et al. Reversible modulation of circadian time with chronopharmacology. *Nat. Commun.* **2021**, *12*, 3164. [[CrossRef](#)] [[PubMed](#)]
3. Solovev, I.A.; Shaposhnikov, M.V.; Moskalev, A.A. Chronobiotics KL001 and KS15 extend lifespan and modify circadian rhythms of *Drosophila melanogaster*. *Clocks Sleep* **2021**, *3*, 429–441. [[CrossRef](#)] [[PubMed](#)]
4. Gul, S.; Akyel, Y.K.; Gul, Z.M.; Isin, S.; Ozcan, O.; Korkmaz, T.; Selvi, S.; Danis, I.; Ipek, O.S.; Aygenli, F.; et al. Discovery of a small molecule that selectively destabilizes Cryptochrome 1 and enhances life span in p53 knockout mice. *Nat. Commun.* **2022**, *13*, 6742. [[CrossRef](#)]
5. Korotkova, D.D.; Lyubetsky, V.A.; Ivanova, A.S.; Rubanov, L.I.; Seliverstov, A.V.; Zverkov, O.A.; Martynova, N.Y.; Nesterenko, A.M.; Tereshina, M.B.; Peshkin, L.; et al. Bioinformatics screening of genes specific for well-regenerating vertebrates reveals c-answr, a regulator of brain development and regeneration. *Cell Rep.* **2019**, *29*, 1027–1040.e6. [[CrossRef](#)]
6. Rubanov, L.I.; Zaraisky, A.G.; Shilovsky, G.A.; Seliverstov, A.V.; Zverkov, O.A.; Lyubetsky, V.A. Screening for mouse genes lost in mammals with long lifespans. *BioData Min.* **2019**, *12*, 20. [[CrossRef](#)]
7. Rubanov, L.I.; Zverkov, O.A.; Shilovsky, G.A.; Seliverstov, A.V.; Lyubetsky, V.A. Protein-coding genes in Euarchontoglires with pseudogene homologs in humans. *Life* **2020**, *10*, 192. [[CrossRef](#)]
8. Lyubetsky, V.A.; Rubanov, L.I.; Tereshina, M.B.; Ivanova, A.S.; Araslanova, K.R.; Uroshlev, L.A.; Goremykina, G.I.; Yang, J.-R.; Kanovei, V.G.; Zverkov, O.A.; et al. Wide-scale identification of novel/eliminated genes responsible for evolutionary transformations. *Biol. Direct* **2023**, *18*, 45. [[CrossRef](#)]
9. von Schantz, M. Phenotypic effects of genetic variability in human clock genes on circadian and sleep parameters. *Rev. J. Genet.* **2008**, *87*, 513–519. [[CrossRef](#)]
10. Bunger, M.K.; Wilsbacher, L.D.; Moran, S.M.; Clendenin, C.; Radcliffe, L.A.; Hogenesch, J.B.; Simon, M.C.; Takahashi, J.S.; Bradfield, C.A. Mop3 is an essential component of the master circadian pacemaker in mammals. *Cell* **2000**, *103*, 1009–1017. [[CrossRef](#)]
11. Okamura, H. Integration of mammalian circadian clock signals: From molecule to behavior. *J. Endocrinol.* **2003**, *177*, 3–6. [[CrossRef](#)] [[PubMed](#)]
12. Young, M.W.; Kay, S.A. Time zones: A comparative genetics of circadian clocks. *Nat. Rev. Genet.* **2001**, *2*, 702–715. [[CrossRef](#)] [[PubMed](#)]
13. Souder, D.C.; Anderson, R.M. An expanding GSK3 network: Implications for aging research. *Geroscience* **2019**, *41*, 369–382. [[CrossRef](#)]
14. Shilovsky, G.A.; Putyatina, T.S.; Morgunova, G.V.; Seliverstov, A.V.; Ashapkin, V.V.; Sorokina, E.V.; Markov, A.V.; Skulachev, V.P. A Crosstalk between the biorhythms and gatekeepers of longevity: Dual role of glycogen synthase kinase-3. *Biochemistry* **2021**, *86*, 433–448. [[CrossRef](#)] [[PubMed](#)]
15. Ralph, M.R.; Menaker, M. A mutation of the circadian system in golden hamsters. *Science* **1988**, *241*, 1225–1227. [[CrossRef](#)] [[PubMed](#)]
16. Lowrey, P.L.; Shimomura, K.; Antoch, M.P.; Yamazaki, S.; Zemenides, P.D.; Ralph, M.R.; Menaker, M.; Takahashi, J.S. Positional syntenic cloning and functional characterization of the mammalian circadian mutation tau. *Science* **2000**, *288*, 483–492. [[CrossRef](#)]
17. Xu, Y.; Padiath, Q.S.; Shapiro, R.E.; Jones, C.R.; Wu, S.C.; Saigoh, N.; Saigoh, K.; Ptáček, L.J.; Fu, Y.H. Functional consequences of a CKIdelta mutation causing familial advanced sleep phase syndrome. *Nature* **2005**, *434*, 640–644. [[CrossRef](#)]
18. Etchegaray, J.-P.; Machida, K.K.; Noton, E.; Constance, C.M.; Dallmann, R.; Di Napoli, M.N.; DeBruyne, J.P.; Lambert, C.M.; Yu, E.A.; Reppert, S.M.; et al. Casein kinase 1 delta regulates the pace of the mammalian circadian clock. *Mol. Cell. Biol.* **2009**, *29*, 3853–3866. [[CrossRef](#)]
19. Shaikevich, E.V.; Karan, L.S.; Fyodorova, M.V. Comparative analysis of the circadian rhythm genes period and timeless in *Culex pipiens* Linnaeus, 1758 (Diptera, Culicidae). *Comp. Cytogenet.* **2016**, *10*, 483–504. [[CrossRef](#)]

20. Hunt, L.C.; Jiao, J.; Wang, Y.-D.; Finkelstein, D.; Rao, D.; Curley, M.; Robles-Murguia, M.; Shirinifard, A.; Pagala, V.R.; Peng, J.; et al. Circadian gene variants and the skeletal muscle circadian clock contribute to the evolutionary divergence in longevity across *Drosophila* populations. *Genome Res.* **2019**, *29*, 1262–1276. [[CrossRef](#)]
21. Pickart, C.M.; Eddins, M.J. Ubiquitin: Structures, functions, mechanisms. *Biochim. Biophys. Acta* **2004**, *1695*, 55–72. [[CrossRef](#)] [[PubMed](#)]
22. Stojkovic, K.; Wing, S.S.; Cermakian, N. A central role for ubiquitination within a circadian clock protein modification code. *Front. Mol. Neurosci.* **2014**, *7*, 69. [[CrossRef](#)] [[PubMed](#)]
23. Abdalla, O.H.M.H.; Mascarenhas, B.; Cheng, H.M. Death of a protein: The role of E3 ubiquitin ligases in circadian rhythms of mice and flies. *Int. J. Mol. Sci.* **2022**, *23*, 10569. [[CrossRef](#)] [[PubMed](#)]
24. Srikanta, S.B.; Cermakian, N. To Ub or not to Ub: Regulation of circadian clocks by ubiquitination and deubiquitination. *J. Neurochem.* **2021**, *157*, 11–30. [[CrossRef](#)] [[PubMed](#)]
25. Frescas, D.; Pagano, M. Deregulated proteolysis by the F-box proteins SKP2 and beta-TrCP: Tipping the scales of cancer. *Nat. Rev. Cancer* **2008**, *8*, 438–449. [[CrossRef](#)]
26. Shirogane, T.; Jin, J.; Ang, X.L.; Harper, J.W. SCF- β TRCP controls clock-dependent transcription via casein kinase 1-dependent degradation of the mammalian period1 (Per1) protein. *J. Biol. Chem.* **2005**, *280*, 26863–26872. [[CrossRef](#)]
27. Busino, L.; Bassermann, F.; Maiolica, A.; Lee, C.; Nolan, P.M.; Godinho, S.I.; Draetta, G.F.; Pagano, M. SCF/Fbxl3 controls the oscillation of the circadian clock by directing the degradation of cryptochrome proteins. *Science* **2007**, *316*, 900–904. [[CrossRef](#)]
28. Godinho, S.I.H.; Maywood, E.S.; Shaw, L.; Tucci, V.; Barnard, A.R.; Busino, L.; Pagano, M.; Kendall, R.; Quwailid, M.M.; Romero, M.R.; et al. The afterhours mutant reveals a role for Fbxl3 in determining mammalian circadian period. *Science* **2007**, *316*, 897900. [[CrossRef](#)]
29. Siepkka, S.M.; Yoo, S.-H.; Park, J.; Song, W.; Kumar, V.; Hu, Y.; Lee, C.; Takahashi, J.S. Circadian mutant overtime reveals Fbox protein FBXL3 regulation of cryptochrome and period gene expression. *Cell* **2007**, *129*, 1011–1023. [[CrossRef](#)]
30. Dardente, H.; Mendoza, J.; Fustin, J.M.; Challet, E.; Hazlerigg, D.G. Implication of the F-Box Protein FBXL21 in circadian pacemaker function in mammals. *PLoS ONE* **2008**, *3*, e3530. [[CrossRef](#)]
31. Hirano, A.; Yumimoto, K.; Tsunematsu, R.; Matsumoto, M.; Oyama, M.; Kozuka-Hata, H.; Nakagawa, T.; Lanjakornsiripan, D.; Nakayama, K.I.; Fukada, Y. FBXL21 regulates oscillation of the circadian clock through ubiquitination and stabilization of cryptochromes. *Cell* **2013**, *152*, 1106–1118. [[CrossRef](#)] [[PubMed](#)]
32. Hirano, A.; Nakagawa, T.; Yoshitane, H.; Oyama, M.; Kozuka-Hata, H.; Lanjakornsiripan, D.; Fukada, Y. USP7 and TDP-43: Pleiotropic regulation of cryptochrome protein stability paces the oscillation of the mammalian circadian clock. *PLoS ONE* **2016**, *11*, e0154263. [[CrossRef](#)] [[PubMed](#)]
33. Bonthuis, P.J.; Huang, W.-C.; Stacher Hörndli, C.N.; Ferris, E.; Cheng, T.; Gregg, C. Noncanonical genomic imprinting effects in offspring. *Cell Rep.* **2015**, *12*, 979–991. [[CrossRef](#)]
34. Gregg, C.; Zhang, J.; Weissbourd, B.; Luo, S.; Schroth, G.P.; Haig, D.; Dulac, C. High-resolution analysis of parent-of-origin allelic expression in the mouse brain. *Science* **2010**, *329*, 643–648. [[CrossRef](#)] [[PubMed](#)]
35. Dilman, V.M. Ontogenetic model of ageing and disease formation and mechanisms of natural selection. *J. Theor. Biol.* **1986**, *118*, 73–81. [[CrossRef](#)] [[PubMed](#)]
36. Skulachev, V.P.; Shilovsky, G.A.; Putyatina, T.S.; Popov, N.A.; Markov, A.V.; Skulachev, M.V.; Sadovnichii, V.A. Perspectives of *Homo sapiens* lifespan extension: Focus on external or internal resources? *Aging* **2020**, *12*, 5566–5584. [[CrossRef](#)]
37. Chen, X.; Wang, X.; Sun, C.; Chen, Q.; O'Neill, F.A.; Walsh, D.; Fanous, A.; Kendler, K.S. FBXL21 association with schizophrenia in Irish family and case-control samples. *Am. J. Med. Genet. Part B Neuropsychiatr. Genet.* **2008**, *147*, 1231–1237. [[CrossRef](#)]
38. Andrews, J.L.; Goodfellow, F.J.; Matosin, N.; Snelling, M.K.; Newell, K.A.; Huang, X.F.; Fernandez-Enright, F. Alterations of ubiquitin related proteins in the pathology and development of schizophrenia: Evidence from human and animal studies. *J. Psychiatr. Res.* **2017**, *90*, 31–39. [[CrossRef](#)]
39. Gunz, P.; Neubauer, S.; Falk, D.; Tafforeau, P.; Le Cabec, A.; Smith, T.M.; Kimbel, W.H.; Spoor, F.; Alemseged, Z. *Australopithecus afarensis* endocasts suggest ape-like brain organization and prolonged brain growth. *Sci. Adv.* **2020**, *6*, eaaz4729. [[CrossRef](#)]
40. Kaplan, H.; Hill, K.; Lancaster, J.; Hurtado, A.M. A theory of human life history evolution: Diet, intelligence, and longevity. *Evol. Anthropol.* **2000**, *9*, 156–185. [[CrossRef](#)]
41. Montgomery, S.H.; Capellini, I.; Venditti, C.; Barton, R.A.; Mundy, N.I. Adaptive evolution of four microcephaly genes and the evolution of brain size in anthropoid primates. *Mol. Biol. Evol.* **2010**, *28*, 625–638. [[CrossRef](#)] [[PubMed](#)]
42. de Magalhães, J.P.; Costa, J. A database of vertebrate longevity records and their relation to other life-history traits. *J. Evol. Biol.* **2009**, *22*, 1770–1774. [[CrossRef](#)] [[PubMed](#)]
43. Lekagul, B.; McNeely, J.A. *Mammals of Thailand*; Association for the Conservation of Wildlife: Bangkok, Thailand, 1977.
44. Brunet-Rossini, A.K.; Austad, S.N. Ageing studies on bats: A review. *Biogerontology* **2004**, *5*, 211–222. [[CrossRef](#)] [[PubMed](#)]
45. O'Leary, N.A.; Wright, M.W.; Brister, J.R.; Ciufu, S.; Haddad, D.; McVeigh, R.; Rajput, B.; Robbertse, B.; Smith-White, B.; Ako-Adjei, D.; et al. Reference sequence (RefSeq) database at NCBI: Current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res.* **2016**, *44*, D733–D745. [[CrossRef](#)] [[PubMed](#)]
46. Xu, B.; Yang, Z. PAMLX: A graphical user interface for PAML. *Mol. Biol. Evol.* **2013**, *30*, 2723–2724. [[CrossRef](#)]
47. Yang, Z. PAML 4: Phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **2007**, *24*, 1586–1591. [[CrossRef](#)]

48. Yang, Z.; Nielsen, R. Estimating synonymous and nonsynonymous substitution rates under realistic evolutionary models. *Mol. Biol. Evol.* **2000**, *17*, 32–43. [[CrossRef](#)]
49. Marques, J.P.; Seixas, F.A.; Farelo, L.; Callahan, C.M.; Good, J.M.; Montgomery, W.I.; Reid, N.; Alves, P.C.; Boursot, P.; Melo-Ferreira, J. An annotated draft genome of the mountain hare (*Lepus timidus*). *Genome Biol. Evol.* **2020**, *12*, 3656–3662. [[CrossRef](#)]
50. Kumar, S.; Stecher, G.; Li, M.; Knyaz, C.; Tamura, K. MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. *Mol. Biol. Evol.* **2018**, *35*, 1547–1549. [[CrossRef](#)]
51. Waskom, M.L. Seaborn: Statistical data visualization. *J. Open Source Softw.* **2021**, *6*, 3021. [[CrossRef](#)]
52. Pianka, E.R. On *r*- and *K*-selection. *Am. Nat.* **1970**, *104*, 592–597. [[CrossRef](#)]
53. Reznick, D.; Bryant, M.; Bashey, F. *R*- and *K*-selection revisited: The role of population regulation in life-history evolution. *Ecology* **2002**, *83*, 1509–1520. [[CrossRef](#)]
54. Comfort, A. *The Biology of Senescence*; Churchill Livingstone: Edinburgh, UK; London, UK; Melbourne, VIC, Australia, 1979.
55. Dayhoff, M.O.; Schwartz, R.M.; Orcutt, B.C. A model of evolutionary change in proteins. In *Atlas of Protein Sequence and Structure*; Dayhoff, M.O., Ed.; National Biomedical Research Foundation: Washington, DC, USA, 1978; Volume 5, pp. 345–352.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.