Giant tortoise genomes provide insights into longevity and age-related disease

Víctor Quesada1,19, Sandra Freitas-Rodríguez1,19, Joshua Miller2,19, José G. Pérez-Silva1,19, Zi-Feng Jiang3, Washington Tapia4,5, Olaya Santiago-Fernández1, Diana Campos-Iglesias1, Lukas F.K. Kuderna6,7, Maud Quinzin2, Miguel G. Álvarez2, Dido Carrero1, Luciano B. Beheregaray8, James P. Gibbs9, Ylenia Chiari10, Scott Glaberman10, Claudio Ciofi11, Miguel Araujo-Voces1, Pablo Mayoral1, Javier R. Arango1, Isaac Tamargo-Gómez1, David Roiz-Valle1, María Pascual-Torner1, Benjamin R. Evans1, Danielle L. Edwards12, Ryan C. Garrick13, Michael A. Russello14, Nikos Poulakakis15,16, Stephen J. Gaughran2, Danny O. Rueda1, Gabriel Bretones1, Tomàs Marquès-Bonet6,7,17,18, Kevin P. White3, Adalgisa Caccone2* and Carlos López-Otín10,1*.

Giant tortoises are among the longest-lived vertebrate animals and, as such, provide an excellent model to study traits like longevity and age-related diseases. However, genomic and molecular evolutionary information on giant tortoises is scarce. Here, we describe a global analysis of the genomes of Lonesome George—the iconic last member of *Chelonoïdis abingdonii—and the Aldabra giant tortoise (*Aldabrachelys gigantea*). Comparison of these genomes with those of related species, using both unsupervised and supervised analyses, led us to detect lineage-specific variants affecting DNA repair genes, inflammatory mediators and genes related to cancer development. Our study also hints at specific evolutionary strategies linked to increased lifespan, and expands our understanding of the genomic determinants of ageing. These new genome sequences also provide important resources to help the efforts for restoration of giant tortoise populations.

Comparative genomic analyses leverage the mechanisms of natural selection to find genes and biochemical pathways related to complex traits and processes. Multiple works have used these techniques with the genomes of long-lived mammals to shed light on the signalling and metabolic networks that might play a role in regulating age-related conditions1–3. Similar studies on unrelated longevous organisms might unveil novel evolutionary strategies and genetic determinants of ageing in different environments. In this regard, giant tortoises constitute one of the few groups of vertebrates with an exceptional longevity: in excess of 100 years according to some estimates.

In this manuscript, we report the genomic sequencing and comparative genomic analysis of two long-lived giant tortoises: Lonesome George—the last representative of *Chelonomis abingdonii*, endemic to the island of Pinta (Galapagos Islands, Ecuador)—and an individual of *Aldabrachelys gigantea*, endemic to the Aldabra Atoll and the only extant species of giant tortoises in the Indian Ocean1 (Fig. 1a). Unsupervised and supervised comparative analyses of these genomic sequences add new genetic information on the evolution of turtles, and provide novel candidate genes that might underlie the extraordinary characteristics of giant tortoises, including their gigantism and longevity.

**Results and discussion**

The genome of Lonesome George was sequenced using a combination of Illumina and PacBio platforms (Supplementary Section 1.1). The assembled genome (CheloAbing 1.0) has a genomic size of 2.3 gigabases and contains 10,623 scaffolds with an N50 of 1.27 megabases (Supplementary Section 1.1 and Supplementary Tables 1–3). We also sequenced, with the Illumina platform, the closely related tortoise *A. gigantea* at an average read depth of 28×. These genomic sequences were aligned to CheloAbing 1.0.

TimeTree database estimations (http://www.timetree.org) indicate that Galapagos and Aldabra giant tortoises shared a last common ancestor about 40 million years ago, while both diverged from the human lineage more than 300 million years ago (Supplementary Section 1.4). A preliminary analysis of demographic history using the pairwise sequentially Markovian coalescent (PSMC)5 model...
Articles

Nature Ecology & Evolution

showed that while the effective population size of *C. abingdonii* has been steadily declining for the past million years, with a slight uptick about 90,000 years ago, the population of Aldabra giant tortoises experienced substantial fluctuations over this period (Fig. 1b). Effective population size reconstructions for *C. abingdonii* lose statistical power at the million-year time frame, probably due to complete coalescence. In turn, this suggests that overall diversity in these giant tortoises must have been low throughout many generations. Together, these results prompt us to propose that the populations of these insular giant tortoises were vulnerable at the time of human discovery of the Galapagos Islands, probably elevating their extinction risk.

Using homology searches with known gene sets from humans and *Pelodiscus sinensis* (the Chinese soft-shell turtle), along with RNA sequencing (RNA-Seq) data from *C. abingdonii* blood and an *A. gigantea* granuloma, we automatically predicted a primary set of 27,208 genes from the genome assembly using the MAKER2 algorithm. We then performed pairwise alignments between each of the primary predicted protein sequences and the UniProt databases for humans and *P. sinensis*, whose annotated sequences show relatively high quality when compared with data available for other turtles. Using alignments spanning at least 80% of the longest protein and showing more than 60% identity, we constructed sets of protein families shared among these species. This preliminary analysis singled out several protein families that seem to have undergone moderate expansion in a common ancestor of *C. abingdonii* and *A. gigantea*. Almost all of these expansions were also confirmed in the genome of the related, long-lived tortoise *Gopherus agassizii* (Supplementary Section 1.2 and Supplementary Table 4). Most of these genes have been linked to exosome formation, suggesting that this process may have been important in tortoise evolution.

We also interrogated the predicted gene set for evidence of positive selection in giant tortoises. This analysis singled out 43 genes with evidence of giant-tortoise-specific positive selection (Supplementary Section 1.2, Supplementary Table 5 and Supplementary Fig. 1). This list includes genes with known roles in the dynamics of the tubulin cytoskeleton (*TUBE1* and *TUBG1*) and intracellular vesicle trafficking (*VPS35*). Importantly, the analysis of genes showing evidence of positive selection also includes *AHSG* and *FGF19*, whose expression levels have been linked to successful ageing in humans. The role of both factors in metabolism regulation—another hallmark of ageing—suggests that the specific changes observed in these proteins may have arisen to accommodate the challenges that longevity poses on this system. The list of genes with signatures of positive selection also features *TDO2*, whose inhibition has been proposed to protect against age-related diseases through regulation of tryptophan-mediated proteostasis.

In addition, we found evidence for positive selection affecting several genes involved in immune system modulation, such as *MVK, IRAK1BP1* and *IL1R2*. Taken together, these results identify...
Proteostasis, metabolism regulation and immune response as key processes during the evolution of giant tortoises via effects on longevity and resistance to infection.

Parallel to this automatic analysis, we used manually supervised annotation on more than 3,000 genes selected a priori for a series of hypothesis-driven studies on development, physiology, immunity, metabolism, stress response, cancer susceptibility and longevity (Supplementary Section 1.3 and Supplementary Fig. 2). We searched for truncating variants, variants affecting known motifs and variants whose human counterparts are related to known genetic diseases (Supplementary Section 1.3 and Supplementary Table 6). These variants were first confirmed with the RNA-Seq data. Then, more than 100 of the most interesting variants in terms of putative functional relevance were also validated by PCR amplification followed by Sanger sequencing. To this end, we used a panel of genomic DNA samples of 11 different species of giant tortoises endemic to different islands from the Galapagos Archipelago (Supplementary Section 1, Supplementary Table 7 and Supplementary Fig. 3).

The manually supervised annotation of development-related genes showed the complete conservation of the Hox gene set among giant tortoises, with the exception of HOXC3, which seems to have been lost in the radiation of Archelosauria14,15 (Supplementary Section 2, Supplementary Table 8 and Supplementary Fig. 4). BMP and GDF gene families were also found to be conserved, although the duplication event that gave rise to GDF1 and GDF3 in mammals did not occur in turtles, birds and crocodiles. In contrast, we found a duplication of the ParaHox gene CDX4 in giant tortoises, also present in other reptiles as well as avian reptiles (birds). This annotation also showed the duplication of WNT11 in turtles and chickens (but not in the lizard Anolis carolinensis), and the specific duplication of WNT4 in turtles. Given the roles of these duplicated genes and their conservation in most vertebrate species, they could prove to be useful candidates to study the morphological development of turtles, particularly in relation to shell formation. Of note, KDSR—one of the genes possibly under positive selection in giant tortoises—has been linked to hyperkeratinization disorders16. Also, in this regard, we annotated 30 β-keratins in C. abingdonii, 26 of which seem to be functional. These numbers are lower than those previously reported for β-keratins in other turtles17. Finally, we did not find in C. abingdonii or A. gigantea any functional orthologues of genes specifically involved in tooth development (such as ENAM, AMEL, AMBN, DSPP, KLK4 and MMP20). This finding confirms a pattern in the evolutionary molecular mechanisms for tooth loss, which seems to have been followed consistently and independently across vertebrates. Taken together, these results offer multiple candidates to study developmental traits in tortoises (Supplementary Section 2 and Supplementary Figs. 5–8).

In most species, the immune function is an evolutionary driver that is under strong selective pressure and has important implications in ageing and disease18. The specific components and functionality of immune system components in Reptilia, however, have not been extensively characterized beyond the major histocompatibility complex (MHC)19,20. Our detailed analysis of 891 genes involved in immune function consistently found duplications affecting immunity genes in giant tortoises compared with mammals (Supplementary Section 3, Supplementary Table 9 and Supplementary Figs. 5–8). We found a genomic expansion of PRF1 (encoding perforin) in giant tortoises and other turtles, compared with chickens (one copy), A. carolinensis (two copies) and most mammals (one copy). Both C. abingdonii and A. gigantea possess 12 copies of this gene (validated by Sanger sequencing), although three of them have been pseudogenized in C. abingdonii. In addition, we detected and validated, by Sanger sequencing, an expansion of the chymase locus, containing granzymes, in giant tortoises (Supplementary Section 3.1 and Supplementary Fig. 10). Both expansions are expected to affect cytotoxic T lymphocyte and natural killer functions, which play important roles in defence against both pathogens and cancer21,22. Other concurrent expansions involve APOBEC1, CAMP, CHIA and NLRP genes, which participate in viral, microbial, fungal and parasite defence, respectively. These results suggest that the innate immune system in turtles, and especially in giant tortoises, may play a more relevant role than in mammals, consistent with the less important role that adaptive immunity seems to play23. We found that class I and II MHC genes probably underwent a duplication event in a common ancestor between giant tortoises and painted turtles (Chrysemys picta bellii). We also annotated 40 class III MHC genes, thus confirming the conservation of this cluster in giant tortoises. The large number of MHC genes in giant tortoises is consistent with the suggestion that ancestors of archosaurs and chelonians did not possess a minimal essential MHC as found in the chicken genome24 (Supplementary Section 3.3, Supplementary Table 10 and Supplementary Figs. 14–16).

Giant tortoises are at the upper end of the size scale for extant Chelonia, and have often been used as an example of gigantism25. We analysed a series of genes involved in size regulation in vertebrates, most notably dogs (Supplementary Section 2, Supplementary Table 8 and Supplementary Fig. 6). Our results on genes related to growth hormone, the insulin-like growth factor (IGF) system and stanniocalcins suggest that these genes are well conserved; therefore, additional size determinants may exist in giant tortoises. As a complex phenotype, gigantism in tortoises is expected to be caused by interactions between different genetic and environmental factors. An interesting finding in this regard is the presence of several gene variants in tortoises (including A. agassizii) probably affecting the activities of glucose metabolism genes, such as MIF (p.N111C; expected to yield a locked trimer) and GSK3A (p.R272Q in the activation loop). Given the roles of these positions in the mammalian orthologues of these genes, tortoise-specific changes could point to differences in the regulation of glucose intake and tolerance (Supplementary Section 4, Supplementary Table 11, and Supplementary Figs. 17 and 18). We also found expansions and inactivations in other genes involved in energy metabolism. Thus, glyceraldehyde-3-phosphate dehydrogenase (GAPDH)—a glycolytic enzyme with a key role in energy production, as well as in DNA repair and apoptosis26—is expanded in giant tortoises. Conversely, the NLN gene encoding neurolysin is pseudogenized in tortoises. The loss of this gene in mice has been related to improved glucose uptake and insulin sensitivity27. Taken together, these results led us to hypothesize that genomic variants affecting glucose metabolism may have been a factor in the development of tortoises.

The analysis of genes related to the stress response has also highlighted several putative variants in giant tortoises affecting globins and DNA repair factors (Supplementary Section 5, Supplementary Tables 12 and 13, and Supplementary Figs. 19–22, 32 and 33). We found that, despite living terrestrially, giant tortoises conserve the hypoxia-related globin Gbx28. Together with coelacanths, turtles, including giant tortoises, are the only organisms known to possess all eight different types of globins29. Consistent with this, we found in both giant tortoise genomes a variant in the transcription factor TP53 (p.S106E) that has been linked to hypoxia resistance in some mammals and fishes28. The presence of the same residue in Testudines strongly suggests a process of convergent evolution in the adaptation to hypoxia, probably driven by an ancestral aquatic environment, which left this footprint in the genomes of terrestrial giant tortoises.

An important trait of large, long-lived vertebrates is their need for tighter cancer protection mechanisms, as illustrated by Pető’s paradox28,30. In turn, this need for additional protection illustrates the deep relationship and interdependence between cancer and longevity (Fig. 2). Notably, tumours are believed to be very rare in turtles31. Therefore, we analysed more than 400 genes classified in...
Fig. 2 Genomic basis of longevity and cancer in giant tortoises. a, Genes potentially implicated in *C. abingdonii* and *A. gigantea* longevity extension and cancer resistance, classified according to their putative role in the different hallmarks. Tables indicate copy-number variations and relevant variants of age-related genes and tumour suppressors found in *C. abingdonii*, *A. gigantea* and other species. Within these tables, numbers indicate gene copy numbers, and asterisks represent pseudogenization events. Dots in colours relating to each hallmark represent presence of the variant. b, Venn diagrams showing the relationships between cancer-, ageing- and immunity-related genes, as classified before annotation. Top, all of the genes related to each category that have been manually annotated, including the number of genes in each group. Bottom, those genes showing potentially interesting variations after annotation.

a well-established census of cancer genes as oncogenes and tumour suppressors. Although most presented a highly conserved amino acid sequence when compared with the sequences of other organisms, we uncovered alterations in several tumourigenesis-related genes (Fig. 2a, Supplementary Section 6, Supplementary Table 14 and Supplementary Figs. 23–29). First, we found that several putative tumour suppressors are expanded in turtles compared with other vertebrates, including duplications in SMAD4, NF2, PML, PTPN11 and P2RY8. In addition, the aforementioned expansion of PRF1, together with the tortoise-specific duplication of PRDM1, suggests that immunosurveillance may be enhanced in turtles. Likewise, we found giant-tortoise-specific duplications affecting two putative proto-oncogenes—MYCN and SET. Notably, the SET complex mediates oxidative stress responses induced by mitochondrial damage through the action of PRF1 and GZMA in cytotoxic T lymphocyte- and natural killer-mediated cytotoxicity. Taken together, these results suggest that multiple gene copy-number alterations may have influenced the mechanisms of spontaneous tumour growth. Nevertheless, further studies are needed to evaluate the genomic determinants of putative giant-tortoise-specific cancer mechanisms.

Finally, we selected, for manually supervised annotation, a set of 500 genes that may be involved in ageing modulation (Supplementary Section 7 and Supplementary Table 15). The extreme longevity of giant tortoises is expected to involve multiple genes affecting different hallmarks of ageing. We found several alterations in the genomes of giant tortoises that may play a direct role in six of them, and impinge on other ageing hallmarks and processes, such as cancer progression (Fig. 2b). First, we identified changes in three candidate factors (NEIL1, RMI2 and XRCC6) related to the maintenance of genome integrity, a primary hallmark of ageing (Fig. 3a). Thus, we found and validated a duplication affecting NEIL1, a key protein involved in the base-excision repair process whose expression has been linked to extended lifespans in several species. Likewise, RMI2 is duplicated in tortoises, suggesting an enhanced ability to resolve homologous recombination intermediates to limit DNA crossover formation in cells. In a preliminary exploration of this hypothesis, we overexpressed NEIL1 and RMI2 in HEK-293T cells and exposed the infected cells to a sublethal dosage of H2O2 or ultraviolet light, monitoring DNA damage by western blot analysis at 24 and 48 h after treatment. As shown in Supplementary Figs. 22, 32 and 33, the expression of both genes results in reduced levels of phosphorylated histone H2AX and cleaved poly (ADP-ribose) polymerase (PARP), suggesting reduced levels of DNA damage. In turn, this result is consistent with the hypothesis that NEIL1 and RMI2 levels may regulate the strength of DNA repair mechanisms. Also in relation to DNA repair mechanisms, we identified and validated a variant affecting XRCC6—encoding a helicase involved in non-homologous end joining of double-strand DNA breaks—which may affect a known sumoylation site (p.K556R). This lysine is conserved in diverse vertebrates but, notably, is changed in giant tortoises, and also in the naked mole rat (p.K556N), the longest-lived rodent, which suggests a putative process of convergent evolution (Fig. 3b). Since
sumoylation is induced following DNA damage and plays a key role in DNA repair response and multiple regulatory processes\(^2\), this variant may reflect selective pressures acting on the regulation of the repair of double-strand DNA breaks in long-lived organisms (Supplementary Section 5.5). Regarding telomere attrition—another primary hallmark of ageing\(^1\)—we uncovered in giant tortoises one variant in DCLRE1B (p.R498C) potentially affecting its binding interface with telomeric repeat binding factor 2 (TERF2) (Fig. 3b and Supplementary Section 7.2). This change, together with the aforementioned variants affecting DNA repair genes that may also impinge on telomere maintenance as a regulatory mechanism of longevity in tortoises. Moreover, we found changes potentially affecting proteostasis (Fig. 2a). We independently found specific expansions of the elongation factor gene EEF1A1 in C. abingdonii, A. gigantea, and G. agassizii, as described with the automatic annotation. Importantly, overexpression of EEF1A1 homologues in Drosophila melanogaster has been linked to an increased lifespan in this species\(^6\). Over time, nutrient sensing deregulation—another hallmark of ageing—can result from alterations in metabolic control mechanisms and signalling pathways\(^2\). The aforementioned variant affecting the activation loop of GSK3A (Supplementary Section 4.1), which is present in C. abingdonii and all tested tortoises from the Galapagos Islands and Aldabra Atoll, as well as their continental outgroups, G. agassizii and C. picta bellii, may be involved in the maintenance of glucose homeostasis. Interestingly, the inhibition of GSK3 can extend lifespan in D. melanogaster\(^4\). Likewise, the identified alterations in other giant tortoise genes implicated in glucose metabolism, such as the aforementioned inactivation of NLN, may provide interesting candidates to study nutrient sensing in these long-lived species (Supplementary Section 7.4).

Regarding the mitochondrial function, we found two variants (p.Q366M and p.M487T) potentially affecting the function of ALDH2, a mitochondrial aldehyde dehydrogenase involved in alcohol metabolism and lipid peroxidation, among other detoxification processes\(^4\). Notably, the p.Q366M variant, which may alter the NAD-binding site of ALDH2, is exclusively found in Galapagos giant tortoises, but not in their continental close relative Chelonoidis chilensis, nor in the more distantly related Aldabra or Agassiz's tortoises. Thus, these changes could also alter the detoxification process and contribute to pro-longevity mechanisms. Together with the above described specific alterations in other genes of giant tortoises, such as NLN and GAPDH, which encode enzymes associated with mitochondrial functions\(^4\), these variants may also impinge on mitochondrial dysfunction, an antagonistic hallmark of ageing\(^1\) (Supplementary Section 7.5).

We have also found evidence in tortoises of some variants related to altered intercellular communication (Supplementary Section 7.6 and Supplementary Fig. 30), an integrative hallmark of ageing\(^1\). Thus, we have detected exclusively in C. abingdonii a premature stop codon affecting ITGA1 (p.R990*), an essential integrin involved in cell–matrix and cell–cell interactions. In addition, the aforementioned variant affecting MIF is also expected to cause the formation of inactivating interchain disulfide bonds, inhibiting intracellular signalling cascades\(^4\). Moreover, MIF deficiency reduces chronic inflammation in white adipose tissue and expands lifespan, especially in response to caloric restriction\(^6\). Finally, we have annotated a specific variant in IGF1R that is expected to affect the interaction between this receptor and the IGF1/2 growth

**Fig. 3 | DNA repair response in giant tortoises.** a. Copy-number variations and putative function-altering point variants found in C. abingdonii, A. gigantea and closely related species. b. Alignments showing the variants highlighted in XRCC6 and DCLRE1B.
IGF1R loss of two amino acids. The fact that variants have been associated with human longevity opens the possibility that the variant found in tortoises could also contribute to increasing the lifespan of these long-lived animals.

In summary, in this work, we report the preliminary characterization of giant tortoise genomes. We complemented the automatic annotation of genomes from two giant tortoise species with a hypothesis-driven strategy using manually supervised annotation of a large set of genes. The analysis of the resulting sequences offers candidate genes and pathways that may underlie the extraordinary characteristics of these iconic species, including their development, gigantism and longevity. A better understanding of the processes that we have studied may help to further elucidate the biology of these species and therefore aid the ongoing efforts to conserve these dwindling lineages. Lonesome George—the last representative of C. abingdonii, and a renowned emblem of the plight of endangered species—left a legacy including a story written in his genome whose unveiling has just started.

**Methods**

**Genome sequencing and assembly.** We obtained DNA from a blood sample from Lonesome George—the last member of C. abingdonii. This DNA was sequenced, using the Illumina HiSeq 2000 platform, from a 180-base pair-insert paired-end library, a 5-kilobase (kb)-insert mate-pair library and a 20-kb-insert mate-pair library. These libraries were assembled with the AllPaths algorithm for a draft genome containing 64,657 contigs with an N50 of 74 kb. Then, we scaffolded the
contigs with SSPACE version 3.0 (ref. 33) using the long-insert mate-pair libraries. Finally, we filled the gaps with PBjelly version 15.8.24 (ref. 25) using the reads obtained from 18 BioPac cells. This step yielded 10,623 scaffolds with an N50 of 1,272 megabases, for a final assembly 2.3 gigabases long. Then, we soft-masked repeated regions using RepeatMasker (http://www.repeatmasker.org) with a database containing chordate repeated elements (included in the software) as a reference. Additionally, we assessed the completeness of assembly by their estimated gene content, using Benchmarking Universal Single-Copy Orthologs (BUSCO) version 3.0 which tested the completeness of a set of 2,586 vertebrate genes from the comprehensive catalogue of orthologues 24. We also performed RNA-Seq from C. abingdoni blood and A. gigantea granuloma, and aligned the resulting reads to the assembled genome using TopHat ver. 2.0.14. Finally, we obtained whole-genome data from A. gigantea with one Illumina lane of a 180-base pair paired-end library. The resulting reads were aligned to the C. abingdoni genome with BWA (version 0.7.5a). Raw reads from C. abingdoni were also aligned to the genome for manual curation of the results. All work on field samples was conducted at Yale University under Institutional Animal Care and Use Committee permit number 2016-10825, Galapagos Park Permit PC-75-16 and Convention on International Trade in Endangered Species number 15US2091429.

Genome annotation. Using the genome assembly of C. abingdoni and the RNA-Seq reads from C. abingdoni and A. gigantea, we performed de novo annotation with MAKER2. The algorithm was also fed both human and P. sinensis reference sequences, and performed two runs in a Microsoft Azure virtual machine (Supplementary Table 16). In parallel, we used selected genes from the human protein database in EnsEMBL as a reference to manually predict the corresponding homologues in the genome of C. abingdoni using the BATIC algorithm (Blast, Annotate, Tune, Iterate) 25. Briefly, this algorithm allows a user to annotate the position and intron/exon boundaries of genes in novel genomes from blastn results. In addition, blastn results are integrated to search for novel homologues in the Ensembl reference database containing data from GenBank, EMBL and RefSeq (https://www.ncbi.nlm.nih.gov/raa), with comments showing which regions were filled with the BioPac reads and therefore may contain frequent errors.

Effective population size changes and diversity. We reconstructed changes in the effective population over time using the PSMC model 5 in the following way: Effective population size changes and diversity. We reconstructed changes in which regions were filled with the BioPac reads and therefore may contain.

References

1. Kim, E. B. et al. Genome sequencing reveals insights into physiology and longevity of the naked mole rat. Nature 479, 223–227 (2011).

2. Keane, M. et al. Insights into the evolution of longevity from the bowhead whale genome. Cell Rep. 10, 112–122 (2015).

3. Nicholls, H. The legacy of Lonesome George. Nature 487, 279–280 (2012).

4. Kehlmaier, C. et al. Tropical ancient DNA reveals relationships of the extinct Bahamian giant tortoise Chelonoidis albyorum. Proc. Roy. Soc. B 284, 20162235 (2017).

5. Li, H. & Durbin, R. Inference of human population history from individual genomes. Nat. Genet. 45, 701–706 (2013).

6. Campbell, M. S., Holt, C., Moore, B. & Yandell, M. Genome annotation and curation using MAKER and MAKER-P. Curr. Protoc. Bioinformatics 48, 1–39 (2014).

7. Wang, Z. et al. The draft genomes of soft-shell turtle and green sea turtle yield insights into the development and evolution of the turtle-specific body plan. Nat. Genet. 45, 701–706 (2013).

8. Sanchis-Gomar, F. et al. A preliminary candidate approach identifies the combination of chemerin, fetuin-A, and fibrinolysis growth factors 19 and 21 as a potential biomarker panel of successful aging. Age 37, 9776 (2015).

9. Pal, D. et al. Fetuin-A acts as an endogenous ligand of TLR4 to promote lipid-induced insulin resistance. Nat. Med. 18, 1279–1285 (2012).

10. Kir, S. et al. FGF19 as a postprandial, insulin-independent activator of hepatic p53 gene provides adaptive apoptotic prevention. Proc. Natl Acad. Sci. USA 106, 12430–12435 (2009).

11. Long, A. S. et al. ATM and ATR signaling regulate the recruitment of human telomerase to telomeres. Cell 13, 1633–1646 (2015).

12. Ribes-Zamora, A., Indiviglio, S. M., Mihalek, I., Williams, C. L. & Bertuch, A. A. TRF2 interaction with KU heterotetramerization interface gives insight into c-NHEJ prevention at human telomeres. Cell Rep. 5, 194–206 (2013).

13. Shikama, N., Ackermann, R. & Brack, C. Protein synthesis elongation factor EF-1 alpha expression and longevity in Drosophila melanogaster. Proc. Natl Acad. Sci. USA 91, 4199–4203 (1994).

14. Castillo-Quan, J. I. et al. Lithium promotes longevity through GSK3/NFR2-dependent hormesis. Cell Rep. 15, 638–650 (2016).

15. Verschueren, L. et al. MIF deficiency reduces chronic inflammation in white adipose tissue and impairs the development of insulin resistance, glucose intolerance, and associated atherosclerotic disease. Circ. Res. 105, 99–107 (2009).

16. Harper, J. M., Wilkinson, J. E. & Miller, R. A. Macrophage migration inhibitory factor-knockout mice are long lived and respond to caloric restriction. FASEB J. 24, 2436–2442 (2010).

17. Plos ONE 10, e0145587 (2015).

18. Englisch, A. C. et al. Mind the gap: upgrading genomes with Pacific SSPACE. Bioinformatics 27, 578–579 (2011).

19. Boecker, M., Henkel, C. V., Jansen, H. J., Butler, D. & Pirovano, W. Scaffolding pre-assembled contigs using SSPACE. Bioinformatics 27, 578–579 (2011).

20. English, A. C. et al. The mind gap: upgrading genomes with Pacific SSPACE. Bioinformatics 27, 578–579 (2011).

21. Fotedar, P. A. A. A census of human cancer genes. Nat. Rev. Cancer 4, 177–183 (2004).

22. Martinvalet, D., Zhu, P. & Lieberman, J. Granzyme A induces caspase-independent mitochondrial damage, a required first step for apoptosis. Immunity 22, 355–370 (2005).

23. Gorbunova, V., Seluanov, A., Zhang, Z., Gladyshev, V. N. & Vijg, J. Comparative genetics of longevity and cancer: insights from long-lived rodents. Nat. Rev. Genet. 15, 531–540 (2014).

24. MacRae, S. L. et al. DNA repair in species with extreme lifespan differences. Aging 7, 1117–1184 (2015).

25. McKee, M. S. & Elston, G. Comparative genomics of CD74 activation of CD74 by endogenous MIF trimer at physiologic concentrations. J. Biol. Chem. 289, 123–133 (2012).

26. Kim, E. B. et al. Genome sequencing reveals insights into physiology and longevity of the naked mole rat. Nature 479, 223–227 (2011).

27. Keane, M. et al. Insights into the evolution of longevity from the bowhead whale genome. Cell Rep. 10, 112–122 (2015).

28. Nicholls, H. The legacy of Lonesome George. Nature 487, 279–280 (2012).

29. Kehlmaier, C. et al. Tropical ancient DNA reveals relationships of the extinct Bahamian giant tortoise Chelonoidis albyorum. Proc. Roy. Soc. B 284, 20162235 (2017).

30. Li, H. & Durbin, R. Inference of human population history from individual genomes. Nat. Genet. 45, 701–706 (2013).

31. Zdobnov, E. M. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics 31, 3210–3212 (2015).

32. Zdobnov, E. M. et al. OrthoDB9.1: cataloging evolutionary and functional annotations for animal, fungal, plant, archael, bacterial and viral orthologs. Nucleic Acids Res. 45, D744–D749 (2017).

33. Martinvalet, D., Zhu, P. & Lieberman, J. Granzyme A induces caspase-independent mitochondrial damage, a required first step for apoptosis. Immunity 22, 355–370 (2005).

34. Gorbunova, V., Seluanov, A., Zhang, Z., Gladyshev, V. N. & Vijg, J. Comparative genetics of longevity and cancer: insights from long-lived rodents. Nat. Rev. Genet. 15, 531–540 (2014).

35. MacRae, S. L. et al. DNA repair in species with extreme lifespan differences. Aging 7, 1117–1184 (2015).

36. McKee, M. S. & Elston, G. Comparative genomics of CD74 activation of CD74 by endogenous MIF trimer at physiologic concentrations. J. Biol. Chem. 289, 123–133 (2012).

37. Simao, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V. & Zdobnov, E. M. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics 31, 3210–3212 (2015).

38. Simao, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V. & Zdobnov, E. M. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics 31, 3210–3212 (2015).
Author contributions

V.Q., and J.G.P.-S. performed the automatic analysis of genomes. S.F.-R. coordinated the manual genomic annotation, which was performed by J.G.P.-S., O.S.-F., D.C.-I., M.G.A., M.A.-V., D.C., P.M., J.R.A., I.T.-G., D.R.-V. and M.F.-T. S.F.-R. and D.C.-I. performed the validation of the identified genomic variants. G.B. coordinated the functional analyses of the identified genomic variants, which were carried out by O.S.-F., D.C.-I., M.G.A., M.A.-V., D.C., P.M., J.R.A. and I.T.-G. J.M. helped to screen the wild samples for SNP validation, and contributed to results interpretation. M.Q., L.B.B., J.P.G., Y.C., S.G., C.C., B.R.E., S.I.G., D.L.E., R.C.G., M.A.R. and N.F. contributed to early data collection and analyses. W.T., D.O.R. and J.P.G. helped to obtain material-securing permits and biological samples. K.P.W. partly supported data collection and supervised the initial analysis. Z.-F.J. prepared DNA and RNA samples for genomic analyses and conducted raw data quality checks. L.F.K. and T.M.-B. performed population history and diversity studies. V.Q., A.C. and C.L.-O. directed the research, analysed the data and wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41559-018-0733-x.

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

Correspondence and requests for materials should be addressed to A.C. or C.L.

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

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

© The Author(s), under exclusive licence to Springer Nature Limited 2018
Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

| n/a | Confirmed |
|-----|-----------|
| ☑   | The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement |
| ☑   | An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| ☑   | The statistical test(s) used AND whether they are one- or two-sided |
| ☑   | Only common tests should be described solely by name; describe more complex techniques in the Methods section. |
| ☑   | A description of all covariates tested |
| ☑   | A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| ☑   | A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| ☑   | For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted |
| ☑   | Give P values as exact values whenever suitable. |
| ☑   | For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| ☑   | For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| ☑   | Estimates of effect sizes (e.g. Cohen’s d, Pearson’s r), indicating how they were calculated |
| ☑   | Clearly defined error bars |
| ☑   | State explicitly what error bars represent (e.g. SD, SE, CI) |

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection | No software was used for data collection. 

Data analysis | BATI is available at http://degradome.uniovi.es/downloads/VBI_O-BATI-0.02.tar.gz 
Statistical comparisons were performed with GraphPad Prism v7.0 and R3.4.3 
Signatures of positive selection were studied with PAML v4 
Custom scripts are accessible from https://github.com/vqf/LG 
Other standard analysis programs are described in the manuscript and supplementary information 

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.
Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Data supporting the findings of this study are available within the paper and its supplementary information files. Sequencing data have been deposited at the Sequence Read Archive (SRA, https://www.ncbi.nlm.nih.gov/sra) with BioProject accession number PRJNA416050. The accession number of the assembled genomic sequence is PKMU00000000. MAKER2-predicted protein sequences can be downloaded from https://github.com/vqf/LG

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- [x] Life sciences
- [ ] Behavioural & social sciences
- [ ] Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size

The most important hypotheses in this work actually refer to a unique individual of C. abingdonii. For comparisons with other species, we rely on inter-species conservation, which is much more stringent than intra-species conservation. For biochemical experiments, no sample size calculation was performed.

Data exclusions

No data were excluded.

Replication

The experiment shown in Fig. 4 was independently replicated three times, as explained in the Methods section (independent infections per replicate). The experiment shown in Suppl. Fig. S22 was performed with one clone per construct. This experiment involved two different treatments and two time points.

Randomization

The main hypotheses in this work refer to species, and therefore randomization is not relevant to this study. For biochemical experiments, groups were established from the same cell line based on infection with different constructs. Therefore, group allocation was pre-established and no randomization was necessary.

Blinding

For the primary results in this work, blinding was not possible, as hypotheses were tested in a single genomic sequence. For biochemical experiments, investigators were not blinded to group allocation.

Reporting for specific materials, systems and methods

| Materials & experimental systems | Methods |
|---------------------------------|---------|
| n/a                             | n/a     |
| [x] Unique biological materials | [x] ChIP-seq |
| [x] Antibodies                  | [x] Flow cytometry |
| [x] Eukaryotic cell lines       | [x] MRI-based neuroimaging |
| [x] Palaeontology               |         |
| [x] Animals and other organisms |         |
| [x] Human research participants |         |

Unique biological materials

Policy information about availability of materials

Obtaining unique materials

Samples from all the species of the Giant Galapagos tortoises species complex are protected and require a CITES permit (see below)
### Antibodies

| Antibodies used | The primary antibodies used were: anti-phospho-Histone H2AX (Ser139) (EMD Millipore, 05-636, clone JBW301, lot 2854120), anti-PARP (Cell Signalling, 9542S, rabbit polyclonal, lot 15), anti-FLAG (Cell Signalling, 2368S, rabbit polyclonal, lot 12), anti-IGF1R (Abcam, ab182408, clone EPR19322, lot GR31678-8), anti-IGF1R (p Tyr1161) (Novus Biologicals, NB100-92555, rabbit polyclonal, lot CJ36131), anti-β-actin (Sigma, A5441, clone AC-15, lot 014M4759) and anti-α-tubulin (Sigma, T6074, clone B-5-1-2, lot 075M4823V). The secondary antibodies used were: |
| --- | --- |
| | LI-COR, IRDye 680RD, 926-68071, polyclonal goat-anti-rabbit, lot C41217-03 |
| | LI-COR, IRDye 680, 926-32220, polyclonal goat-anti-mouse, lot CO072-03 |
| | LI-COR, IRDye 800CW, 926-32211, polyclonal goat-anti-rabbit, lot C60113-05 |
| | LI-COR, IRDye 800CW, 926-32210, polyclonal goat-anti-mouse, lot C50316-03 |

### Validation

All antibodies used in this study were purchased from commercial companies, and they had been verified by the manufacturer. As stated in their websites:

- Anti-phospho-Histone H2AX (Ser139), clone JBW301 is a well published Mouse Monoclonal Antibody validated in ChIP, ICC, IF, WB. This purified mAb is highly specific for phospho-Histone H2AX (Ser139) also known as H2AXS139p.
- PARP Antibody detects endogenous levels of full length PARP1 (116 kDa), as well as the large fragment (89 kDa) of PARP1 resulting from caspase cleavage. The antibody does not cross-react with related proteins or other PARP isoforms.
- DYKDDDDK Tag Antibody (Anti-Flag) detects exogenously expressed DYKDDDDK proteins in cells. The antibody recognizes the DYKDDDDK peptide (the same epitope recognized by Sigma's Anti-FLAG® antibodies) fused to either the amino- or carboxy-terminus of targeted proteins. The binding specificity of this antibody is NOT dependent on the presence of divergent metal cations.
- Our Abpromise guarantee covers the use of ab182408 in the following tested applications: WB (1/1000 dilution). Detects a band of approximately 100,200 kDa (predicted molecular weight: 156 kDa)...
- Anti-IGF1R (p Tyr1161): Validated by Western blot (WB) analysis of p-IGF-1R (Y1161) pAb in extracts from Hela cells.
- Anti-β-actin western blot validation: 1.5 Rapid 1-10,000 using cultured human or chicken fibroblast cell extracts. Reacts against guinea pig, canine, Hirudo medicinalis, feline, pig, carp, mouse, chicken, rabbit, sheep, rat, human and bovine orthologs. Does not react against Dictyostelium discoideum.
- Anti-α-tubulin western blot validation: 0.25-0.5 μg/mL using total cell extract of human foreskin fibroblast cell line (FS11). Species reactivity: human, Chlamydomonas, African green monkey, chicken, kangaroo rat, bovine, mouse, rat, sea urchin.

### Eukaryotic cell lines

**Policy information about cell lines**

**Cell line source(s):** ATCC

**Authentication:** PCR-based microsatellite characterization was performed at the University of Oviedo.

**Mycoplasma contamination:** Cell lines were not tested for mycoplasma contamination

**Commonly misidentified lines (See [ICLAC register](#))**

HEK-293T cells are widely used for infection experiments. The identity of these cells was assessed by PCR-based microsatellite characterization

### Animals and other organisms

**Policy information about studies involving animals. ARRIVE guidelines recommended for reporting animal research**

**Laboratory animals**

The study did not involve laboratory animals.

**Wild animals**

The study did not involve observations but did involve temporary captures of wild animals to extract blood samples.

**Field-collected samples**

All work on field samples was conducted at Yale University under IACUC permit number 2016-10825, Galapagos Park Permit PC-75-16 and CITES number 15US209142/9