**GENETICS**

Growing old, yet staying young: The role of telomeres in bats’ exceptional longevity

Nicole M. Foley,1 Graham M. Hughes,1 Zixia Huang,1 Michael Clarke,1 David Jebb,1 Conor Whelan,1 Eric J. Petit,2 Frédéric Touzalin,3 Olivier Farcy,4 Gareth Jones,5 Roger D. Ransome,5 Joanna Kacprzyk,1 Mary J. O’Connell,6 Gerald Kerth,7 Hugo Rebelo,8,9 Luísa Rodrigues,10 Sébastien J. Puechmaille,11 Emma C. Teeling1*

Understanding aging is a grand challenge in biology. Exceptionally long-lived animals have mechanisms that underpin extreme longevity. Telomeres are protective nucleotide repeats on chromosome tips that shorten with cell division, potentially limiting life span. Bats are the longest-lived mammals for their size, but it is unknown whether their telomeres shorten. Using >60 years of cumulative mark-recapture field data, we show that telomeres shorten with age in *Rhinolophus ferrumequinum* and *Miniopterus schreibersii*, but not in the bat genus with greatest longevity, *Myotis*. As in humans, telomerase is not expressed in *Myotis myotis* blood or fibroblasts. Selection tests on telomere maintenance genes show that ATM and DETX, which repair and prevent DNA damage, potentially mediate telomere dynamics in *Myotis* bats. Twenty-one telomere maintenance genes are differentially expressed in *Myotis*, of which 14 are enriched for DNA repair, and 5 for alternative telomere-lengthening mechanisms. We demonstrate how telomeres, telomerase, and DNA repair genes have contributed to the evolution of exceptional longevity in *Myotis* bats, advancing our understanding of healthy aging.

**INTRODUCTION**

Aging is the gradual and irreversible breakdown of living systems due to the progressive deterioration of physiological function that causes an increase in susceptibility to disease and eventual mortality (1). Although the mean human life span in the European Union and United States is currently 78 years, and increasing by ~2 years per decade, the age of onset of age-related degenerative diseases appears stationary (for example, 77% of age-related cancers occur over 50 years of age) (2). Therefore, we urgently need to better understand the mechanisms of the aging process, with a view to improving the future quality of life of our aging populations (2, 3). Most aging studies have been carried out in shorter-lived laboratory model species, given the ease of manipulation, housing, and length of life span (4, 5). Although they are excellent study species, it is difficult to extrapolate experimental findings in these short-lived laboratory species to long-lived, outbred species such as humans (6). Therefore, it has been argued that long-lived, outbred species such as bats may be better models to investigate the aging processes of relevance to people (4).

Only 19 species of mammal are longer-lived than humans given their body size, and 18 of these species are bats (7). Bats are the longest-lived mammals relative to their body size (Fig. 1), with the oldest bat captured (*Myotis brandtii*) being >41 years old, weighing ~7 g, and living ~9.8 times longer than predicted for its size (4, 8). Most bat species are longer-lived than similar-sized mammals (6, 9) and exhibit little or no evidence of senescence (10, 11). Flight provides bats a means to escape predation and avoid adverse conditions that ultimately has driven the evolution of longer life spans (9). Of the species examined to date, the genus *Myotis* holds the longevity record with at least 3 species living beyond 20 years and 4 species exceeding 30 years (6, 12) despite their small body size (2 to 45 g). Although an excellent model species to study extended “health span,” logistically, it is difficult to study aging in bats because they are not easily maintained in captivity (5). Field-based aging studies require individuals to be marked, easily caught, and nonlethally sampled every year throughout their long life span. Here, uniquely drawing on more than 60 years of cumulative long-term, mark-recapture studies from four wild populations of long-lived bats, we determine whether telomeres, a driving factor of the aging process (1), shorten with age in *Myotis myotis* (*n* = 239; age, 0 to 6+ years), *Rhinolophus ferrumequinum* (*n* = 160; age, 0 to 24 years), *Myotis beechsteinii* (*n* = 49; age, 1 to 16 years), and *Miniopterus schreibersii* (*n* = 45; age, 0 to 17 years) (Table 1).

Telomeres are tandem 5′-TTAGGG-3′ repeats that form protective structures at both ends of chromosomes (13). Through the action of six telomere-associated proteins, collectively termed shelterin, telomeres are maintained by preventing the tips of chromosomes from being recognized as double-stranded breaks (14). In human somatic cells, telomeres shorten with successive rounds of cell division resulting from the “end replication problem” in which the cellular replicative machinery cannot replicate to the end of the chromosome (13). When telomeres become critically shortened, the cell elicits a DNA damage signal and undergoes cellular senescence (13, 15), ultimately driving the aging process. Although telomere shortening with age has been observed in a wide diversity of animal populations (16), no bats have been assessed to date. In germ, stem, and cancer cells, telomeres are maintained through the action of telomerase, a reverse transcriptase capable of restoring repeats to telomere ends (15). Among mammals, repressed telomerase and telomere shortening has evolved as a cancer-suppressing mechanism (17). Despite reportedly having long telomeres (9 to 30 kb) and some species potentially expressing telomerase (17), cancer is rarely reported in bats (18). These factors, coupled with positive selection acting on telomere maintenance genes in *Myotis lucifugus* (19),
suggest that the telomeres of long-lived bats may be unusual compared to other mammals. However, nothing is currently known about telomere dynamics in bats and its potential role in their longevity.

RESULTS
To address these questions, we captured four different species of wild bats (n = 493) of known age, at field sites (2013–2016) across Europe (see Materials and Methods and tables S1 and S2). Once captured, 3-mm wing biopsies were taken from each bat before release (see Materials and Methods). Samples were flash-frozen in liquid nitrogen (LN₂) or desiccated using silica beads, and high–molecular weight DNA was extracted (see Materials and Methods). Quantitative polymerase chain reaction (qPCR) assays of telomere length were repeated in triplicate on each plate. Each plate was assayed three times, and the relative telomere length (rTL) was calculated (see Materials and Methods). For each species, we modeled the relationship between rTL and age using a suite of linear, parabolic, piecewise-linear, and linear mixed models (LMMs) (table S2). For all data sets, except M. schreibersii, for which a simple linear model was the best fit, an LMM controlling for factors relating to technical variation (for example, interplate variation between DNA extraction plates or qPCR runs) was the best fit (see Materials and Methods and table S2).

Our results show that telomeres shorten with age in R. ferrumequinum and M. schreibersii (P < 0.001 in both cases) (Fig. 2, A and B), typical of most mammals (16). In the genus Myotis, we did not detect a relationship...
between rTL and age in either *M. myotis* (*P* = 0.16) or *M. bechsteinii* (*P* = 0.27) (Fig. 2, C and D). This genus includes the longest-lived bat species that have been studied to date (6). The difference in effect size (slope of the line) does not explain the difference in significance, which is potentially explained by sources of variation unrelated to age in *Myotis* (fig. S1, table S8, and Supplementary Text). Jackknife analyses confirmed the robustness of results and direction of the slope in all data sets (figs. S2 and S3 and Supplementary Text). To investigate whether differing results between species were simply artifacts relating to differences in sample sizes or age ranges, we compared *M. myotis* and *R. ferrumequinum* data sets of identical composition (same number of individuals per age cohort, *n* = 91), and this analysis reca-pitulated our initial results (fig. S4 and Supplementary Text). To explore the selective disappearance of individuals with short telomeres as a potential confounding factor in our cross-sectional analyses, we performed an upper quartile regression using the *M. myotis* and *R. ferrumequinum* data sets. Significant telomere shortening with age is observed among the *R. ferrumequinum* individuals with the longest telomeres (upper quartile), but this effect is not detected in *M. myotis* (fig. S5 and Supplementary Text). Although increased sampling or longitudinal data will be necessary to fully elucidate the degree to which age is a driver of telomere dynamics in *Myotis* species, it is clear from our study that the association between rTL and age observed in *R. ferrumequinum, M. schreibersii*, and many other mammal species.

![Fig. 2. The relationship between rTL and age in four long lived bat species.](http://advances.sciencemag.org/)

Plot showing the relationship between Box-Cox transformed rTL and age modeled in an LMM statistical framework for (A) *R. ferrumequinum*, (C) *M. myotis*, and (D) *M. bechsteinii*, and (B) *M. schreibersii*, using a linear model. Corresponding *P* values indicate the relationship between rTL and age for each species and are shown in the top right-hand corner of each plot. Models are plotted where significant. Note that scales differ between plots.
including humans (16, 20) is not detected in either of the *Myotis* species studied.

To better understand how telomeres may be maintained in *Myotis* species, we investigated telomerase expression in *M. myotis*. Using a publicly available blood transcriptome (table S3) derived from the same study population, we conducted an analysis of telomerase expression in *M. myotis* and compared it with data from mouse, pig, and human blood transcriptomes (table S3 and Supplementary Text). We also determined telomerase expression in primary fibroblast cell culture transcriptomes, generated from wing tissue of *M. myotis* and from a mouse ear biopsy (Supplementary Text). The expression of associated shelterin genes was used as an internal control within species, given that expression levels measured in fragments per kilobase of transcript per million mapped reads (FPKM) are not readily comparable across species. Our analysis determined the presence or absence of expression, rather than relative amounts. Our results showed that telomerase was expressed in the blood in mice and pigs but that *M. myotis*, like humans, showed no signs of telomerase expression (Fig. 3, A to D). This is further supported by analysis of the wing-derived fibroblast transcriptome, which also shows telomerase expression in mice but not in *M. myotis* (Fig. 3, E and F). Our results suggest that the maintenance of telomeres in *M. myotis* is not mediated by telomerase.

**Blood transcriptomes**

**A Myotis myotis**
- Libraries: *n* = 4
- Max. life span: 37.1 years

**B Mouse**
- Libraries: *n* = 4
- Max. life span: 4 years

**C Pig**
- Libraries: *n* = 3
- Max. life span: 27 years

**D Human**
- Libraries: *n* = 6
- Max. life span: 122.5 years

**Cell Culture Fibroblast Transcriptomes**

**E Myotis myotis**
- Libraries: *n* = 1

**F Mouse**
- Libraries: *n* = 1

---

**Fig. 3. Comparative analysis of telomerase expression.** Analysis of telomerase expression (*TERT*) in blood transcriptomes of (A) *M. myotis*, (B) mouse, and (C) pig, and (D) human and fibroblast cell cultures in (E) *M. myotis* and (F) mouse. Expression of other shelterin genes (*POT1, TERF2IP, TINF2, TPP1, TERF1*, and *TERF2*) is provided as a comparison to *TERT* expression levels within species. Note that scales differ between plots.
To further explore the genetic basis of telomere maintenance in *M. myotis*, we identified 225 target genes using gene ontology (GO) terms associated with telomere maintenance and literature searches (table S4). Coding sequences from our targeted gene set were mined from 52 publicly available eutherian mammal genomes, including 12 bat species, spanning ~98 million years of mammalian evolution (tables S4 and S5) (21). Where possible, data were mined using the Reference Sequence (RefSeq) database. A custom approach incorporating the annotation pipeline MAKER was designed to optimize data mining from poorly annotated genomes (see Materials and Methods). This enabled the inclusion of additional genome data from five bat species and three eutherian species (table S5 and Supplementary Text). Two data sets were prepared for downstream selection analyses, one containing only RefSeq data and one combining RefSeq and data mined using our custom MAKER pipeline. Tests of positive and divergent selection were carried out independently along five branches of the tree (fig. S6): the branches leading to (i) all bats, (ii) the genus *Myotis*, (iii) *R. ferrumequinum*, (iv) *Miniopterus natalensis* (*M. schreibersii* genome data were not available), and (v) the naked mole rat, an additional exceptionally long-lived mammal given its small size (Fig. 1) (4).

CodeML selective pressure variation tests were automated using a custom-built workflow, OH-SNAP (Optimised High-throughput SNakemake Automation of PAML; available at https://github.com/batlabucd/OHSNAP) (see Materials and Methods and fig. S7). To ensure that our results were robust to the addition or removal of data, only results recovered as significant in both data sets after false discovery rate (FDR) correction are further considered (tables S6 and S7).

Our analyses of selective pressure heterogeneity revealed a paucity of positive selection in all branches tested (table S6). This is unsurprising given the role that telomere maintenance genes play in controlling the cell cycle and maintaining genomic integrity. For example, deletion of all shelterin components except TERF2IP are lethal in mice (14), suggesting that conservative evolutionary forces have shaped the evolution of these genes in mammals. However, after FDR correction, divergent selection was detected in *SETX* and *ATM* along the branch leading to the genus *Myotis* (Fig. 4C and table S7). *ATM* is the chief mobilizer of the DNA damage response, and the helicase *SETX* is a known interactor. Advancing a known role for enhanced DNA repair in long-lived species (22), *ATM* and *SETX* may contribute to genomic stability in bats through transcriptional silencing at DNA damage sites and by preventing DNA damage at replication forks.
Defects in either of these genes cause ataxia, a disorder that is associated with accelerated telomere shortening and reduced life span (26, 27). Hence, adaptations in these genes in Myotis bats could serve as new targets for anti-aging research.

A comparative multi-tissue transcriptome analysis was conducted using publicly available data to investigate whether the same set of telomere maintenance genes are differentially expressed in Myotis bats compared to other mammals (for example, cow, pig, bowhead whale, human, mouse, rat, and naked mole rat; Materials and Methods and table S3). After FDR correction, our analysis showed significant differential expression of 21 telomere maintenance genes in Myotis bats compared to all other mammals. This list includes 20 up-regulated genes, including ATM and SETX that we showed to be evolving under divergent selection in the genus Myotis, as well as ABL1, CCT4, DCLRE1A, DOT1L, GNL3L, MLH3, MRE11A, PARP1, RAD50, RBP1, RFC3, RPA1, SDE2, SSB, TERF21P, WRAP53, WRN, XRCC5, and one down-regulated gene, TEN1 (table S8 and Supplementary Text). Because these genes were chosen with a focus on telomere maintenance a priori, a STRING (Search Tool for Recurring Instances of Neighbouring Genes) network analysis was conducted to determine whether these were enriched for any secondary biological process or pathway (see Supplementary Text). After FDR correction, our analysis showed significant functional enrichment (P < 0.05) for GO terms and KEGG (Kyoto Encyclopedia of Genes and Genomes) pathways associated with DNA repair (fig. S8 and tables S10 and S11). Divergent genomic and transcriptomic changes in these DNA repair genes, positioned at the axes of telomere maintenance, DNA damage prevention, and repair, may represent the molecular adaptations that have underpinned the evolution of exceptional longevity in the genus Myotis.

**DISCUSSION**

While we do not detect a significant relationship between rTL and age in Myotis and do not detect telomerase expression in M. myotis blood or fibroblasts, the specific mechanism through which these bats may maintain their telomers is not clear. In human cancers, in the absence of telomerase, telomeres can be maintained by homology-directed telomere synthesis or alternative lengthening of telomeres (ALT) (28, 29). Results from our comparative transcriptomic analyses show that genes that have roles in pathways associated with ALT, ATM, MRE11a, RAD50, and WRN (28, 29) are significantly differentially expressed in all tissues of Myotis bats compared to all other mammals examined. Furthermore, genes that play a role in resolving replication forks and recombination-mediated repair are essential for telomere maintenance in ALT cells (28). SETX performs these functions (24, 30) and is differentially expressed and divergently evolving in Myotis species; however, it has not yet been investigated in an ALT context. Given the observed divergent selection in ATM and SETX as well as differential expression of these genes and MRE11a, RAD50, and WRN, our results are suggestive of a potential role for ALT in the maintenance of telomeres in the genus Myotis with age. However, to date, ALT activity has only been reported in anomalous situations (for example, cancer), and so, hypotheses arising from our transcriptomic and genomic analyses relating to ALT-mediated telomere maintenance should be subject to future experimental validation to fully elucidate telomere dynamics in Myotis species.

Despite also living much longer than expected given their body size (Fig. 1), there was no evidence of divergent selection on the branches leading to R. ferrumequinum or the naked mole rat (Fig. 4D and table S7). This suggests that adaptation of telomere maintenance may be specific to the Myotis lineage and that other longevity mechanisms may have evolved in Rhinolophus and the naked mole rat (see Supplementary Text). Divergent selection was detected in CCT5, HNRNPU, and LIG1 along the branch leading to M. natalensis, suggesting the evolution of enhanced proteostasis mechanisms (Fig. 4B and Supplementary Text) rather than telomere maintenance in Miniopterus. This could explain the association of telomere shortening with age in this lineage and highlights the potential evolution of different longevity strategies in bats (see Supplementary Text). MYC was shown to be evolving under divergent selection along the bat ancestral branch (Fig. 4A). Although widely known as an oncogene, this pleiotropic transcription factor can also induce tumor suppression mechanisms, such as apoptosis, cellular senescence and the DNA damage response (31), which may contribute to longer life span and lower cancer incidence observed in bats (18). Because these results are based on in silico analyses, future functional assays are needed to confirm the role of these suggested pathways in bat longevity.

**CONCLUSION**

Among mammals, bats are unparalleled in their ubiquitous longevity in relation to their body size. The genus Myotis, in particular, includes the longest-lived species known to date (8) with field-based evidence for negligible senescence in M. bechsteinii (11). For the first time, we show that, unlike other long-lived bat species, we do not detect a relationship between rTL and age in either Myotis species studied. A comparative analysis of telomerase expression in blood transcriptomes revealed that, like in human blood, telomerase is not expressed in M. myotis blood, suggesting that telomere maintenance in this species is unlikely to be mediated by telomerase. Instead, tests of selective pressure heterogeneity in telomere maintenance genes suggest that ATM and SETX, which function to repair and prevent DNA damage, may contribute to telomere maintenance in this long-lived genus. These results, coupled with differential expression of ATM, SETX, MRE11a, RAD50, and WRN across all tissues in the genus Myotis compared to other mammals, suggest a potential role for ALT mechanisms in the maintenance of telomeres in these species. If telomeres are maintained by ALT mechanisms in Myotis species, then these genes may represent excellent therapeutic targets given that cancer incidence in bats is rare (18). In addition, from a human perspective, telomere maintenance in the absence of telomerase is highly desirable given that telomerase expression is present in ~90% of human cancers (13). Our analyses suggest that DNA repair genes, ATM and SETX, affecting DNA repair and telomere maintenance have contributed to the evolution of exceptional longevity in these Myotis species and represent excellent future study targets to better understand the aging process.

**MATERIALS AND METHODS**

**LQ estimation**

The data used to determine the mammalian LQ included all mammal species (n = 779) from the data set of Healy et al. (9). For each species, expected longevity was calculated using a linear regression fitted to logged values of maximum longevity and mass for all nonflying eutherian species (slope = 0.186, intercept = 0.546), following Austad and Fischer (7). As per Austad and Fischer (7), the LQ was calculated as follows: LQ = observed longevity/expected longevity.
Telomere length analysis

Study species and wing biopsy sampling

Full ethical approval and permission (AREC-13-38-Teeling) for capture and field sampling were awarded by the University College Dublin ethics committee to E.C.T. Sampling was carried out in accordance with the ethical guidelines and sampling permits issued in each country and are detailed in table S1. M. myotis, between ages 0 and 6+ years (that is, first fitted with transponders as adults 6 years before subsequent recapture, therefore true age unknown), were captured from Brittany, France, as they left the roost, using modified harp traps (Table 1). R. ferrumequinum, between ages 0 and 24 years from Woodchester, near Bristol, UK, were captured using hand nets in the roost. M. bechsteinii, between ages 1 and 16 years, were captured from bat boxes in forests close to Wurzburg, Germany. M. schreibersii, between ages 0 and 17+ years, were captured as they left the roost, using a harp trap from Tomar and Grândola, Portugal.

Once captured, all bats were placed in individual cloth bags or boxes before processing. Wing biopsies were taken using a 3-mm biopsy punch and were either flash-frozen in LN2 and transferred to a −150°C freezer for long-term storage (M. myotis), or immediately preserved in silica beads (32) at room temperature and then transferred to a −20°C freezer for long-term storage (R. ferrumequinum, M. bechsteinii, and M. schreibersii). To explore the viability of silica beads as an alternative preservation medium to LN2, we carried out an experiment using a subset of individual bats (n = 6). Two wing biopsies were taken from each individual; one biopsy was flash-frozen in LN2 and then stored at −150°C, whereas the other was stored in silica in the field and then transferred to −20°C for long-term storage. After 5 months, all 12 samples were extracted and assayed as described below. Tests indicated our data were homogeneous with respect to variance and non-normally distributed. The nonparametric Kruskal-Wallis and Wilcoxon rank sum tests showed that there was no significant difference (P = 0.423 and P = 0.48, respectively) in rTL obtained from the samples preserved with LN2 or silica.

DNA extraction and qPCR assay

DNA was extracted from biopsies using membrane filter-based extraction kits [Promega Wizard SV DNA Extraction kit (catalog no. A2371, Promega Corporation) or the Qiagen DNeasy Blood and Tissue kit (Qiagen)]. The Promega 96-well extraction protocol was partially automated using a custom protocol for the Hamilton STAR deck liquid handling robot but otherwise followed the manufacturer’s instructions. Qiagen DNeasy extractions were carried out following the manufacturer’s instructions. The integrity of DNA from subsets of extracted samples was confirmed using electrophoresis on a 1% agarose gel. DNA concentration and purity were quantified using a NanoDrop 8000 Spectrophotometer (Thermo Fisher Scientific).

rTL was measured using a modified version of Cawthon’s (33) method to measure telomere length in humans. Telomeres were amplified using primers described by Cawthon (33). A novel single copy gene (SCG) primer pair was designed for the mammalian brain-derived neurotrophic factor (BDNF) gene (BDNF_F1: AGCTGAGCG-TATGTGACAGT; BDNF_R1: TGGGATTACACTTGGTCTCGT). These primers were designed to amplify across diverse mammalian taxa and were optimized here for use in bats. Because of differences in optimum annealing temperatures, the telomere and SCG reactions were carried out on separate plates. All primer stocks were used at 50 μM concentration. Telomere reactions were carried out in 10 μl reaction volumes containing 2.5 μl of DNA (−2 to 5 ng/μl), 5 μl of Takara SYBR Green (Clontech), 0.027 μl of Tel2 primer, 0.09 μl of Tel2 primer, 0.2 μl of ROX (Takara SYBR Green, Clontech), and the remaining volume with PCR grade water (Invitrogen). The compositions of telomere and SCG reactions were identical except for the primers: 0.12 μl of BDNF_F1 and 0.2 μl of BDNF_R1. qPCR amplifications were carried out using the Stratagene Mx3000TM real-time PCR machine. Telomere reactions were amplified using the following cycling conditions: an initial step of 30 s at 95°C, then 25 cycles of 15 s at 95°C, and 2 min at 54°C, followed by a melt curve. SCG reactions were amplified as follows: an initial step of 30 s at 95°C, then 40 cycles of 15 s at 95°C, and 1 min at 58°C, followed by a melt curve. On each plate, samples were assayed in triplicate, and each plate was run three times to account for machine variation between runs. A “golden sample” was prepared from wing tissue and stored in identical 50 μl aliquots at −80°C for use throughout the experiment to account for among-plate variation (33, 34). A golden sample and a negative control were run on each plate in triplicate.

Data analysis and rTL

Raw data were exported from the Stratagene Mx3000TM software. Baseline correction and per-well amplification efficiencies were calculated using LinRegPCR (35). Data were normalized across plates for both telomere and SCG reactions using the golden sample as a reference. The coefficient of variation (CoV) was calculated for each sample. Where the CoV exceeded arbitrary cutoff values chosen to maximize the data (5% for telomeres and 2.5% for SCG), results from individual reaction replicates were either removed or the sample was excluded from the analysis. The average Ct (cycle threshold) value was calculated across nine replicates for each sample. The Pfaffl (36) method, which takes into account the difference in amplification efficiencies between primer sets, was used to calculate rTL. The average amplification efficiency for the telomere primer set was 90% and 93% for the SCG primer pair.

Statistical analysis

All statistical analyses were carried out in R v3.3.1 (37). Data sets from individual species were analyzed separately. A Box-Cox transformation (38) was used to normalize each data set. A suite of linear, parabolic, piecewise, and linear mixed-effect models were constructed using the R packages MASS (39) and lme4 (40). Models were constructed with rTL as the response variable, age as the fixed variable, and with additional random variables listed in table S2. The best-fitting model was selected using the Akaike information criteria (AIC). In addition to the AIC, model residuals were examined to ensure a good fit of the model to the data. Analysis of variance (ANOVA) was used to determine the significance of the fixed effect, age. Jackknife analyses were used to estimate the robustness of our results. The number of individuals subsampled without replacement varied with each data set (M. myotis, n = 160; R. ferrumequinum, n = 100; M. bechsteinii, n = 30; and M. schreibersii, n = 30). The P values from 100 resulting models were compared, as were slopes where underlying models were significant, for example, R. ferrumequinum and M. schreibersii (figs. S1 and S2). To determine whether differences in sample size or sampled age ranges affected our results, we directly compared the two largest data sets, M. myotis and R. ferrumequinum, by randomly subsampling data sets to have equal age cohorts and equal numbers of samples per age cohort (1 to 6+ years; n = 91). To explore selective disappearance as a confounding factor in our analyses, we performed an upper quartile regression using the M. myotis and R. ferrumequinum data sets. The upper quartile range of rTL for each age cohort was subsampled from the original data set, and the best-fitting model (Table 1) was refitted to the upper quartile data set. Significance was assessed as described above. The rate of telomere loss can be compared between species if...
the control sample used to calculate telomere length is consistent, as it is in this study (16). Previous analyses have calculated the rate of telomere loss after the removal of the youngest age cohorts to control for developmental differences, which may accelerate the rate of shortening (16). Hence, the data sets described in table S2 were reanalyzed with the 0-age cohort removed to facilitate a comparison of slopes (effect sizes). LMMs were used to analyze all data sets to ease comparison. \( R^2 \) values (conditional and marginal) were calculated to investigate the proportion of variance explained by each model.

Telomere maintenance genes were identified using associated GO terms through the Bioconductor package (52) in R. Gene targets used in previous analyses (19) were incorporated into our data set, resulting in a total of 225 genes for downstream analysis (table S4). In total, genomes from 52 eutherian mammals were mined to include a representative from each mammalian order where possible (table S5). This included 12 bat genomes, of which 5 are representatives from Yinpterochiroptera, which includes \( R. \) ferrumequinum; the remaining 7 are representatives from Yangochiroptera, which includes 4 from the genus \( M. \) myotis and 1 from \( M. \) natalensis, a species closely related to \( M. \) schreibersii. Where genome data existed for multiple taxa in a given mammalian order, taxa with data covering the highest number of telomere maintenance genes were chosen for inclusion in the final data set.

Gene sequences from annotated mammalian genomes were downloaded using their gene identification numbers in the Entrez RefSeq database. For poorly annotated genomes, such as the 17× unannotated \( R. \) ferrumequinum genome, a custom MAKER (46) pipeline was applied, using gene sets from closely related sister taxa as a reference (see table S5 for details). Because of the fragmented nature of the poorly annotated genomes, exonic gene regions can span multiple contigs/scaffolds, rendering annotation extremely difficult. To overcome this, the first pre-MAKER modification step involved locating all these exon-containing genome sequences in poorly annotated genomes using BLAST (50, 51) and a phylogenetically related reference gene. These sequences are concatenated end-to-end into a single larger “pseudo-scaffold,” which is then passed to MAKER. Within MAKER, putative coding regions were identified with blastn, spliced alignments with intron-exon boundaries resolved using Exonerate (53), and Hidden Markov Model (HMM) gene models trained on each individual gene within the telomere maintenance data set using SNAP (48). Output transcripts from MAKER were aligned to their respective reference genes using a pairwise aligner, matcher in the European Molecular Biology Open Software Suite (EMBOSS) software (54), to determine sequence coverage. A minimum threshold of 70% coverage was required before a gene could be considered “useable.” Any sequences containing an internal stop codon were removed because these likely represented poor-quality data, pseudogenes, or alternative isoforms. Where multiple transcript variants were available for a target gene, the longest variant was chosen to maximize coverage.

Given that significantly less gene information could be mined from the unannotated genomes, often because of poor quality, two data sets were prepared for downstream analyses. One contained RefSeq data
only, whereas the other included RefSeq data and the species mined using MAKER (table S4). This approach ensured that our results would be robust to the trade-off between the quality (RefSeq) and quantity (RefSeq and MAKER) of data. Both data sets were initially aligned as nucleotides with default settings using the Aqua (55) alignment optimization workflow, which compares MUSCLE (multiple sequence comparison by log-expectation) (56) and MAFFT (multiple alignment using fast fourier transform) (57, 58) alignments, applies a refinement step using RASCAL (59), and provides a norMD (60) score as a measure of alignment quality. Alignments were further optimized using the phylogeny-aware aligner Prank (61) by aligning as codons, using six iterations of the alignment and the -F option to specify that insertions should be trusted. Gblocks (62) was used to remove any large indels from the alignment by specifying the data as codons, a minimum block length of 2 and allowing positions that were half gaps. Alignments were visually inspected to remove any obvious poor-quality data, and the data were realigned as described above.

**OH-SNAP: Automation of CodeML using Snakemake**

An automated workflow called OH-SNAP was constructed to run CodeML as implemented in the PAML package (63). The workflow was designed using the Snakemake workflow management system (64) and was executed on a high-performance compute cluster. The workflow requires a species tree (described below), an alignment, a list of taxa contained within the branch to be labeled, and a template control file specifying the model(s) to run (fig. S7). To account for heterogeneity in missing data across the data set, the workflow begins by pruning the species tree to match the taxa present in the alignment. Using the aforementioned taxa list, the target branch to be tested is then labeled. The template control file, which details the model, is then populated, specifying the alignment, labeled tree, and output file names. The workflow then runs CodeML across a specified number of nodes, parallelizing this single threaded program. The species tree for all taxa contained within the data set was constructed using the topology of Meredith et al. (21) and Ruedi et al. (65) for the branching pattern within the genus Myotis (fig. S7). The following branches were labeled to detect selection in the branches leading to (i) bats, (ii) the genus Myotis, (iii) R. ferrumequinum, (iv) M. natalensis, and (v) Heterocephalus glaber, the naked mole rat (fig. S6).

We used two tests to detect signatures of adaptation in telomere maintenance genes: the branch site test (model A versus model A null) (66), which is used to detect positive selection, and clade model C versus M2a_rel, which is used to detect divergent selection (67). In the branch site test, omega (\( \omega \)) was calculated on a site-wise basis for the labeled branch designated as the foreground before being compared with estimates obtained for the background branch of the tree (model A; model = 2, NSsites = 2, fix_omega = 0, omega = 0.4, and clean_data = 0). This alternative model was then compared with a null model which assumes no positive selection (model A null; model = 2, NSsites = 2, fix_omega = 1, omega = 1, and clean_data = 0). A likelihood ratio test (LRT) was used to compare the fit of the two models, and a \( \chi^2 \) test with one degree of freedom was used to assess the significance of the model fit. To detect divergent selection, we used the clade model C, which estimates the average \( \omega \) of the labeled foreground clade and compared it with the estimated average \( \omega \) of the background clade (model = 3, NSsites = 2, and clean_data = 0). This model was compared with the null model M2a_rel (model = 0, NSsites = 22, and clean_data = 0) following Weadick and Chang (67), using an LRT and a \( \chi^2 \) test with one degree of freedom to determine the significance. Resulting \( P \) values from both tests were corrected using the FDR method to account for multiple testing with a significance level of 0.05. Alignments of significant results were once more visually inspected to confirm data quality for target branches and rule out alignment error, because poorly aligned positions can lead to false positive results. Where data did not meet this standard, the gene was removed from our final list of results (tables S6 and S7).

**Comparative transcriptomics of telomere maintenance genes**

**Data selection**

To determine whether our data set of 225 telomere maintenance genes was differentially expressed in bats compared to other mammals, we carried out a comparative transcriptomic analysis. Data were chosen under two criteria: (i) The study individual must be an adult and (ii) the tissue must be derived from a healthy individual. Using M. brandii and M. myotis as representatives of the long-lived bat genus Myotis, we compared the expression of these genes with their homologous counterparts in diverse mammalian taxa—cow (Bos taurus), human (Homo sapiens), mouse (Mus musculus), naked mole rat (H. glaber), pig (Sus scrofa), rat (Rattus norvegicus) and bowhead whale (Balaena mysticetus) across four tissue types: blood, kidney, liver and brain (table S3).

**1:1 ortholog identification**

For genome-wide, cross-species comparison of gene expression, 1:1 orthologs were identified between Myotis and other species for quantification. CDs were obtained from their respective genomes (table S3). The M. lucifugus genome was used as the common reference to standardize the analysis for both M. myotis and M. brandii because it is better annotated. Where multiple transcripts were present, the longest transcript was chosen for analysis. The genomic telomere maintenance gene data set was added to the CDS data sets for each species. Reciprocal BLAST was performed on the filtered CDS data set between Myotis and other species, with the e value cutoff of \( 1 \times 10^{-10} \). Resulting pair-wise orthologs were extracted for differential gene expression analysis.

**Analysis of differentially expressed genes**

Adaptor trimming and base quality filtering of raw RNA-seq data were carried out following Huang et al. (49). Using Sailfish v.0.7.6 (68), clean reads were mapped to 1:1 orthologous transcript sets that were identified between Myotis and other species respectively and the orthologs were quantified. Sailfish k-mer parameter settings were determined on the basis of the length of RNA-seq reads. k-mer values were set to 31 when RNA-seq reads were longer than 75 bp and set to 25 when reads were shorter than 75 bp. Because library type varied across data sets (single-end or paired-end), the single-end mode (−SR) was applied to all libraries for transcript quantification. For each tissue type, a pair-wise differential gene expression analysis was carried out between Myotis and other species using the Bioconductor package DESeq2 in R v.3.0.3 (69). Expression values for our target 225 telomere maintenance genes were extracted from this data set for further investigation. After FDR correction, genes with a significance value of \( <0.05 \) were considered significantly differentially expressed. To investigate whether differentially expressed genes (DEGs) were consistently differentially expressed across multiple tissues, we selected DEGs that showed up- or down-regulation in Myotis compared to at least three other species, for further analysis. Given that our data set was chosen a priori to explore telomere maintenance genes, a protein-protein interaction network was constructed to determine whether Myotis DEGs showed signs of secondary enrichment for any other process besides telomere maintenance. To explore this, the STRING database v.10.0 (70) was queried using significant DEGs as input. H. sapiens was chosen as the query species because it has the best characterized protein interactions. Default parameters were chosen to construct a protein-protein interaction network.


Acknowledgments: For assistance in the field, we would like to thank the members of Bretagne Vivante, S. Dool, local volunteers in Bristol, students from University College Dublin, and all owners/local authorities in each country for allowing access to the sites. We thank L. A. J. O’Neill for providing the mouse tissue for the fibroblast analysis. We would also like to thank J. A. Finarelli for helpful discussions of the analyses. Funding: This project was funded by a European Research Council Research Grant (ERC-2012-StG311000) awarded to E.C.T. The French field study is supported by a Contrat Nature grant awarded to Bretagne Vivante, G.K. was supported by the German (Deutsche Forschungsgemeinschaft: KE 746/2-1/3-1/4-1/5-1/6-1 and GRK 2010) and the Swiss (Schweizerische Nationalfonds: 31-59556.99) national science foundations. H.R. was funded by a Fundação para a Ciência e a Tecnologia contract (IF/00497/2013). Author contributions: E.C.T. and N.M.F. devised the study. M. myotis samples were collected by F.T., O.F., S.J.P., E.C.T., E.J.P., N.M.F., D.J., Z.H., and C.V.W. R. ferrumequinum samples were collected by G.J. and R.D.R. M. bechsteinii samples were collected by G.K. M. schreibersii samples were collected by H.R., N.M.F., and C.V.W. All telomere length data were generated by N.M.F. Genomic data sets were constructed by G.M.H., M.C., and N.M.F. All genomic and telomere length data sets were analyzed by N.M.F. Transcriptomic analyses were performed by Z.H. The M. myotis and mouse fibroblast cell culture was grown and extracted by J.K., and M.C. analyzed the resulting transcriptome. N.M.F. is responsible for the figures presented throughout. The manuscript was written by E.C.T. and N.M.F. with input from all authors. Competing interests: The authors declare that they have no competing interests. Data and materials availability: All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data or bioinformatic pipelines related to this paper may be requested from the authors. OH-SNAP is available for download from https://github.com/batlabucd/OHSNAP. The mouse and M. myotis fibroblast transcriptomes generated as part of this analysis are available through the National Center for Biotechnology Information Sequence Reads Archive under accession no. PRJNA415775.

Submitted 13 June 2017
Accepted 10 January 2018
Published 7 February 2018
10.1126/sciadv.aao0926

Growing old, yet staying young: The role of telomeres in bats' exceptional longevity

Nicole M. Foley, Graham M. Hughes, Zixia Huang, Michael Clarke, David Jebb, Conor V. Whelan, Eric J. Petit, Frédéric Touzalin, Olivier Farcy, Gareth Jones, Roger D. Ransome, Joanna Kacprzyk, Mary J. O'Connell, Gerald Kerth, Hugo Rebelo, Luisa Rodrigues, Sébastien J. Puechmaille and Emma C. Teeling

DOI: 10.1126/sciadv.aao0926