Molecular evolution of growth hormone and insulin-like growth factor 1 receptors in long-lived, small-bodied mammals

Kalina T.J. Davies[a*], Georgia Tsagkogeorga[a], Nigel C. Bennett[b], Liliana M. Dávalos[c], Christopher G. Faulkes[a], Stephen J. Rossiter[a*]

(a) School of Biological & Chemical Sciences, Queen Mary University of London, Mile End Road, London, E1 4NS

(b) Mammal Research Institute, Department of Zoology & Entomology, University of Pretoria, Private Bag X20, 0028 South Africa

(c) Department of Ecology and Evolution, State University of New York at Stony Brook, 650 Life Sciences Building, Stony Brook, NY 11794-5245, USA

[*] Corresponding authors

Email addresses: kalinadavies@gmail.com (K.T.J. Davies); g.tsagkogeorga@qmul.ac.uk (G. Tsagkogeorga); ncbennett@zoology.up.ac.za (N.C. Bennett); liliana.davalos-alvarez@stonybrook.edu. (L.M. Dávalos); c.g.faulkes@qmul.ac.uk (C.G. Faulkes); s.j.rossiter@qmul.ac.uk (S.J. Rossiter)

Abstract

Mammals typically display a robust positive relationship between lifespan and body size. Two groups that deviate markedly from this pattern are bats and the African mole-rats;

with members of both groups being extremely long-lived given their body size, with the maximum documented lifespan for many species exceeding 20 years. A recent genomics study of the exceptionally long-lived Brandt's bat, *Myotis brandtii* (41 years), suggested its longevity and small body size may be at least partly attributed to key amino acid substitutions in the transmembrane domains of the receptors of growth hormone (GH) and insulin-like growth factor 1 (IGF1). However, whereas elevated longevity is likely to be common across all 19 bat families, the reported amino acid substitutions were only observed in two closely related bat families. To test the hypothesis that an altered GH/IGF1 axis relates to the longevity of African mole-rats and bats, we compared and analysed the homologous coding gene sequences in genomic and transcriptomic data from 26 bat species, five mole-rats and 38 outgroup species. Phylogenetic analyses of both genes recovered the majority of nodes in the currently accepted species tree with high support. Compared to other clades, such as primates and carnivores, the bats and rodents had longer branch lengths. The single 24 amino acid transmembrane domain of IGF1R was found to be more conserved across mammals compared to that of GHR. Within bats, considerable variation in the transmembrane domain of GHR was found, including a previously unreported deletion within the Emballonuridae. The transmembrane domains of rodents were found to be more conserved, with mole-rats lacking uniquely conserved amino acid substitutions. Molecular evolutionary analyses showed that both genes were under purifying selection in bats and mole-rats. Our findings suggest that while the previously documented mutations may confer some additional lifespan to *Myotis* bats, other, as yet unknown, genetic differences are likely to

account for the long lifespans observed in many bat and mole-rat species.

Keywords: mammals; bats; mole-rats; longevity; transmembrane domains.

Abbreviations: *GH*, growth hormone; *GHR,* growth hormone receptor; *IGF1*, insulin-like growth factor 1; *IGF1R*, insulin-like growth factor 1 receptor; SNP, single nucleotide polymorphism;

1. Introduction

Understanding the genetic basis of ageing and longevity is of exceptional interest. Typically two main sources of information have shed light on this field; first, the manipulation of specific genes in model organisms that can lead to increased lifespan [e.g. as reviewed in (Kenyon 2010)], and second, attempts to identify the genetic mutations that have led to increased longevity in natural populations [e.g. (Kim *et al.,* 2011)]. Due to their exceptionally long lifespans, high metabolic rates and small body sizes, bats have been proposed as potentially underexploited models for ageing studies [e.g. as reviewed in (Wilkinson and South 2002; Brunet-Rossinni and Austad 2004)]. Although little is known about senescence in bats, it appears they do not undergo the same typical ageing processes as humans (Brunet-Rossinni and Wilkinson 2009). For example, studies suggest that bats may be able to generate new hair cells within certain regions of the inner ear after birth (Kirkegaard and Jørgensen 2000); although the functional impact of this on their sensory perception remains unclear. The particular diets of bat species may be either high in fats or sugars, yet bats appear to avoid the associated

health implications such as artherosclerosis or hyperglycemia (Widmaier *et al.,* 1996; Brunet-Rossinni and Austad 2004; Mqokeli and Downs 2012) which are frequently seen in ageing human populations.

Hypotheses previously put forward to explain bats' long lives include several relating to hibernation, such as altered metabolism and increased predator avoidance [for review see (Wilkinson and South 2002; Brunet-Rossinni and Austad 2004)]. However, hibernation alone is unlikely to account for the increased lifespan of bats, since not all bats hibernate and hibernation is associated with increased survival across mammals generally (Turbill *et al.,* 2011). The ability to fly has been proposed as leading to greater predator avoidance in bats and volant birds, and thus may at least partly explain their longevity compared to similarly sized non-volant species; in this case extrinsic mortality is reduced which simultaneously drives an increase in lifespan (Healy *et al.,* 2014). Reductions in extrinsic mortality are expected to result in evolutionary adaptation to enhance survival at later life stages (Williams 1957).

Currently, little is known regarding the genetic basis behind the increased longevity displayed across the ~1,300 currently known species of bat. A recent study by Seim *et al*. (2013) carried out a genomic analysis of Brandt's bat, *Myotis brandtii*, which holds the longevity record for bats, with one male documented to live 41 years old (Podlutsky *et al.,* 2005). This exceptional longevity, coupled with a small body mass (7g) implies that this species represents the most extreme mammal species outlier in the proposed lifespan body mass relationship [see Figure 2 from (Podlutsky *et al.,* 2005)]. Seim *et al.* (2013)

found that members of two bat families (Vespertilionidae and Molossidae) shared unique mutations in the transmembrane domains of two genes thought to play crucial roles in growth and ageing: the growth hormone receptor gene (*GHR*) and insulin-like growth factor 1 receptor gene (*IGF1R*). The protein products of these genes are transmembrane receptors found on the surface of mammalian cells, with a single transmembrane region each. GHR regulates the cellular effects of growth hormone and IGF1R the effects of insulin-like growth factor 1.

Mutations in the genes related to these two hormones and their associated receptors have been linked to several clinical disorders including dwarfism in humans and a long-lived dwarf phenotype in mice (Godowski *et al.,* 1989; Flurkey *et al.,* 2001). Furthermore, genomic evidence from domestic dogs suggests that allelic variation in *IGF1* is responsible for nearly all of the variation in body size found across breeds (Sutter *et al.,* 2007), which in turn is inversely related to breed lifespan (Greer *et al.,* 2007). Additional evidence suggests that a non-synonymous single nucleotide polymorphism (SNP) in *IGF1R* may further contribute to the body size of 'tiny' dog breeds (Hoopes *et al.,* 2012). Single nucleotide polymorphisms in the coding sequence of *IGF1R* in Angus cattle have also been shown to be associated with significant body mass differences in calves (Szewczuk *et al.,* 2013). Such evidence led Seim *et al.,* (2013) to propose that the observed amino acid substitutions in the Vespertilionidae bat transmembrane regions of GHR and IGF1R, together with traits such as hibernation and low reproductive rate, may contribute to the unusually long lifespan of *Myotis* bats species.

On average bats have a maximum recorded lifespan that is 3.5 times longer than expected given their body size (Wilkinson and South 2002). Since great longevity is a trait shared by most bats, it is interesting that the documented amino acid changes in bat GHR and IGF1R were not found in all bat species examined. In particular, the transmembrane domains of larger-bodied fruit-eating bats from the Phyllostomidae and Pteropodidae families were not found to share the same amino acid substitutions seen in *Myotis* bats (Seim *et al.,* 2013). While there is good evidence that the mutations are conserved across *Myotis* bats and closely related species from the same suborder (Yangochiroptera), it is currently unknown whether other long-lived, small-bodied bat species from the other suborder (Yinpterochiroptera), e.g. *Rhinolophus ferrumequinum* –30.5 years [references within (Wilkinson and South 2002)], share these mutations.

In addition to bats, African mole-rats (Family Bathyergidae) have also been shown to be long-lived for their body size. This is classically illustrated in the naked mole-rat, *Heterocephalus glaber*, which can achieve a maximum lifespan of 31 years with a body mass of 35 grams [The AnAge Database: (Tacutu *et al.,* 2013)]. Naked mole-rats have been cited as an example of a mammal that display negligible senescence [as reviewed in (Buffenstein 2008)]. They do not undergo age-related mortality until very late in their lives, breeding females remain fertile and physiologically do not show the typical signs of ageing; for example, decline of vascular system function and increased tumorigenesis (Csiszar *et al.,* 2007; Buffenstein 2008; Liang *et al.,* 2010). Less is known about the maximum lifespan of many of the other mole-rat species; however, the recorded maximum age of captive *Georychus capensis* is ~5 years (Bennett *et al.,* 2006); while

reproductive queens of *Fukomys damarensis* may live >8.5 years (Schmidt *et al.,* 2013). Similarly to bats, reduced extrinsic mortality through predator avoidance has recently been suggested as one possible route to increased longevity in mole-rats; although in this case this is attributed to their fossorial lifestyles (Healy *et al.,* 2014). A high quality genome is available for *H. glaber* (Kim *et al.,* 2011), and through this and related resources, several possible underlying molecular mechanisms underpinning its exceptional lifespan have begun to be documented (Kim *et al.,* 2011; Yu *et al.,* 2011; Edrey *et al.,* 2012; Morgan *et al.,* 2013). However, to date little is known about the molecular evolution of the GH and IGF1 receptors in the naked mole-rat and closely related rodent species.

To gain further insights into the molecular evolution of these two receptors relating to the genetic control of longevity and body size in bats and mole-rats, we performed phylogenetic analysis of sequence data from each clade, combined with outgroup mammal species. We compared overall substitution rates and selection pressures acting on both of these genes in clades of interest compared to other mammals, and also examined the specific amino acid substitutions that have occurred in the transmembrane region of each gene. Finally, we examined levels of parallel sequence evolution across all pair-wise branch comparisons within the tree (excluding tips) for each gene to test for evidence of molecular convergence between mole-rats and bats.

2. Materials and methods

2.1. Species representation and datasets

We surveyed *GHR* and *IGF1R* nucleotide sequences in 75 mammals, including 26 bats and five mole-rats, generating a total of 64 and 49 sequences for *GHR* and *IGF1R*, respectively. Bat and mole-rat sequences were obtained from a range of sources; including 10 published genomes and five transcriptomes, as well as 17 RNA-seq assemblies of short-read Illumina data assembled with the default parameters of Trinity v. 2013-02-25 (Grabherr *et al.,* 2011). We obtained wide representation of bat species from both the Yinpterochiroptera and Yangochiroptera suborders, including non-echolocating Old World fruit bats and laryngeal echolocating species. From the mole-rats, we obtained sequences from five species which display a range of ecologies and diverse social structures. In addition, to increase the taxonomic sampling we obtained sequences for other divergent subterranean mammals for both genes from RNA-seq assemblies of short-read Illumina data from the East African root rat, *Tachyoryctes splendens*, and *IGF1R* from the golden mole, *Amblysomus hottentotus*.

2.2. Identification of homologous sequences

Complete and partial coding sequences were obtained from genomes and transcriptomes using a BLAST approach (Altschul *et al.,* 1997). For genomic datasets, scaffolds putatively containing genes of interest were initially identified with TBLASTX against the query transcript with an e-value cutoff of $1e^{-6}$ and only keeping hits recovered with >75% identity and with the query sequence in the correct reading frame. Identities were then confirmed by best reciprocal BLAST hits with the same parameters as above. Coding sequences were subsequently extracted from the genomic sequences using

BL2SEQ with the query coding sequence. In each case multiple queries were used; human coding sequences (Ensembl IDs: ENSG00000112964 and ENSG00000140443) were initially used in all cases. Additionally, sequences from *Pteropus vampyrus* (Ensembl IDs: ENSPVAG00000005609 and ENSPVAG00000003279), *Myotis lucifugus* (Ensembl IDs: ENSMLUG00000017190), and *M. brandtii* (GenBank accession: XM_005875995) were used as subjects for bat searches. For mole-rats, we searched using *H. glaber* coding sequences (GenBank accessions: XM_004848566.1 and XM_004879528). Extracted novel sequences were combined with coding sequences downloaded from Ensembl (Flicek *et al.,* 2013) for all mammalian one-to-one orthologs with the human gene. Ensembl sequences that contained more than 10% missing data were excluded from further analysis. The *Myotis lucifugus* annotation currently lists *GHR* as a one-to-many orthologue (Ensembl 73); therefore these sequences were excluded from the analysis. Additional mammal sequences were obtained from GenBank, and also the assemblies of short-read Illumina RNA-seq data following the procedure outlined above (see Appendix A: Supplementary Table S1 for species and source information). Novel *GHR* and *IGF1R* sequences have been deposited in GenBank (accession numbers: KM190081–KM190105).

2.3. Alignment and phylogenetic analyses of nucleotide sequences

Nucleotide sequences of each gene were aligned with GUIDANCE (Penn *et al.,* 2010) using the PRANK algorithm (Löytynoja and Goldman 2005), with codons enforced and 10 bootstraps. Low-quality sequences, those that obtained a quality score of below 0.6, were removed from the multi-fasta file and the alignment recalculated. A Perl script was

then used to remove all codon positions from the alignments that contained >50% missing data. This resulted in alignments consisting of 1,821 and 3,948 base pairs for *GHR* and *IGF1R* respectively. The GTRCAT model was implemented in RAxML v.7.2.8 (Stamatakis 2006) to produce phylogenies based on each gene alignment, and nodal support for the resultant phylogeny was estimated with 100 bootstraps.

2.4. Examination of amino acid variation in the transmembrane domains

Nucleotide alignments were translated in-frame using the standard genetic code. The 24 amino acids corresponding to the transmembrane domain of each gene (Ullrich *et al.,* 1986; Edens and Talamantes 1998) were then extracted and sequence variation examined across the phylogeny. Additionally residue conservation scores were calculated for each column across the entire amino acid alignment in trimAlvs1.4 (Capella-Gutierrez *et al.,* 2009).

2.5. Testing for divergent selection

To test for divergent selection acting on both genes in the two focal clades, bats and mole-rats, the clade model C (Bielawski and Yang 2004) was run with codeml in PAMLv4.4 (Yang 2007). In this case either the bat or mole-rat clade was set as the foreground clade, and the estimated averaged ω (the number of non-synonymous substitutions per non-synonymous site: the number of synonymous substitutions per synonymous site) of this clade was then compared to that of the averaged ω of the background clade consisting of Laurasiatheria and Euarchontoglires. For each gene, alignments were recalculated, using the above methods, on a pruned taxa set containing

only one clade of interest at a time, i.e. when bats were set as the foreground clade of interest, mole-rats were removed from the background dataset and *vice versa* (see Appendix B: Supplementary data for alignments. The topology of the species tree used was based on recent studies (Faulkes *et al.,* 2004; Blanga-Kanfi *et al.,* 2009; Agnarsson *et al.,* 2011; Meredith *et al.,* 2011; Tsagkogeorga *et al.,* 2013) (Appendix C: Supplementary figure S1). Values estimated by each clade model were then compared with model M1a (nearly neutral) via a likelihood ratio test (LRT) with three degrees of freedom (DF), with *P*-values <0.05 indicating the alternative model has a significantly better fit compared to the null.

Additionally, we used branch-site models (Zhang *et al.,* 2005) to test for evidence of positive selection acting on three ancestral branches; the common ancestral bat branch, the common ancestral mole-rat branch and the common ancestral Vespertilionidae + Molossidae branch. In this test, the single branch was set as the foreground and the estimates of site-wise ω values were compared with estimates across the remaining background branches in the phylogeny under model A. This model was compared with the null model A again using a LRT with one DF. The Vespertilionidae + Molossidae branch-site models were carried out on a reduced alignment containing only bat species.

2.6. Quantifying convergent evolution at the amino acid level

For each of the two genes, we characterised the distribution of sequence convergence between pairs of branches in the species phylogeny using the package codeml ancestral (Castoe *et al.,* 2009). We were particularly interested in levels of amino acid convergence

between the two focal long-lived clades, bats and mole-rats. The species tree was used to estimate branch lengths and model parameters under the Dayhoff model of amino-acid substitution in codeml in PAMLv4.4. These values were then used in codeml ancestral to estimate posterior probabilities of all possible amino-acid substitutions, convergent substitutions (same amino acid) and divergent substitutions, between pair-wise branch comparisons under a Dayhoff model of amino-acid substitution.

3. Results

3.1. Summarising phylogenetic signal in GHR and IGF1R

Gene sequence alignments of mammalian *GHR* spanned 1,821 base pairs (607 amino acids) and that of *IGF1R* consisted of 3,948 base pairs (1,316 amino acids). Phylogenetic analysis of these nucleotide alignments correctly recovered the majority of the accepted species relationships and major mammalian sub-divisions (Blanga-Kanfi *et al.,* 2009; Meredith *et al.,* 2011; Tsagkogeorga *et al.,* 2013). Trees based on *GHR* and *IGF1R* sequences recovered bats as monophyletic with high (100%) and moderate (79%) bootstrap support, respectively (see Figure 1). Within bats, the monophyly of the Yangochiroptera sub-division received higher support (100% and 96%) than that of the Yinpterochiroptera (72% and 18%) for *GHR* and *IGF1R*, respectively. Within each suborder the correct familial placements were recovered by *GHR*, but not *IGF1R*, where for example, Phyllostomidae were not recovered as monophyletic. Across taxa, primates typically had the shortest branch-lengths, corresponding to the lowest substitution rates, while rodents and Glires typically had the longest branch-lengths corresponding to a

greater number of substitutions. Within bats, Yangochiroptera had the longest branch-lengths; in particular in the *IGF1R* tree, the branch leading to the Vespertilionidae family (*Eptesicus fuscus* + *Myotis spp.*) was the longest branch across the placental mammals surveyed.

3.2. Examination of amino acid variation in the transmembrane domains

Examination of the amino acids corresponding to the transmembrane domains of GHR and IGF1R indicated that few residues are completely conserved across all mammals (Figure 1 and Appendix D: Supplementary Figure S2). Typically, IGF1R was seen to be less variable, although fewer sequences were recovered for this gene.

Within bats, the transmembrane region of GHR displayed little variation across the six species of Yinpterochiroptera examined (*Eidolon helvum–Megaderma lyra* in Figure 1A), which ranged in adult body mass from 23–872 g, and in maximum recorded lifespan from 10–30.5 years (see supplementary Table S1 and Figure 2). The transmembrane region of *Megaderma lyra* was the most variable, containing two unique amino acid substitutions not seen in any other bat species. Within the suborder Yangochiroptera, considerably more variation was observed with sequence variation supporting three main clades corresponding to the Emballonuroidea, Noctilionoidea and Vespertilionoidea (Teeling *et al.,* 2005). We confirmed that the reported deletion at Leu284 and substitution of Ile275Met (Seim *et al.,* 2013) in the transmembrane domain of GHR is shared by *Myotis davidii*, *M. ricketti*, *M. elegans* and *Rhogeessa aeneus* in the Vespertilionidae and also *Molossus sinaloae* (Molossidae). It is interesting to note that across all mammals

examined the only other group to display a deletion in the GHR transmembrane region were the two species of Emballonuridae examined (*Saccopteryx bilineata* and *Peropteryx kappleri*) at Phe270.

Examination of the rodent GHR transmembrane domain sequence showed little consistent variation between the relatively short-lived muroid rodents and the much longer-lived hystricomorph mole-rats. Only a single amino acid substitution in the transmembrane domain was found to be shared across all mole-rats with the exclusion of guinea pig and other rodents. However, other variable regions were observed outside of the transmembrane domain, including a conserved six amino acid deletion (corresponding to codons 522–527 in the human GHR transcript) observed in the guinea pig and mole-rat sequences.

Overall the amino acid sequence of the transmembrane domain of IGF1R was found to be more conserved across all mammals, including most bats, than that of GHR. For example, across Yinpterochiroptera, Mormoopidae and Phyllostomidae, there is little amino acid variation observed (see Figure 1B). The exception to this pattern is found within the Vespertilionidae (*Myotis spp.* and *Eptesicus fuscus*), which have a total of eight variable amino acid sites not seen in the other bat species examined. In contrast to bats, the transmembrane region of IGF1R in rodents was found to be highly conserved, with identical amino acids shared between mole-rats and the muroid species (Figure 1B). Some IGF1R sequence variation was seen in members of the Afrotheria, which also display a great range in both lifespan and body mass with the golden mole

transmembrane domain displaying considerable sequence variation compared to the remaining species (Figure 1B).

Amino acid conservation scores for GHR are highly variable across all three (extracellular, transmembrane and cytoplasmic) domains (see Appendix E: Supplementary figure S3A). The GHR transmembrane domain does not appear to be remarkable in terms of conservation scores compared to the surrounding non-transmembrane regions. The residues of the extracellular and cytoplasmic domains of IGF1R typically are much higher compared to those of GHR, indicating greater levels of conservation (Supplementary figure S3B). However, the IGF1R transmembrane region, together with the ~20 flanking residues, display much lower conservation levels indicating this region contains considerable variation across the species included in this study.

3.3. Selection pressures acting on GHR and IGF1R

Clade models of molecular evolution constructed for *GHR* revealed evidence of significant divergent selection in bats and mole-rats compared to the background clade of carnivores, ungulates, other rodents, Glires and Primates (LRT: bats: $P<0.0001$ and mole-rats: $P<0.0001$; see Appendix A: Supplementary Table S2A for full results). However, in both model comparisons the estimated ω (dN/dS) value on the foreground clade (FG) and background clade (BG) fell within the range of purifying selection (Bats: FGω = 0.289; BGω = 0.403; mole-rats: FGω = 0.489; BGω = 0.320). Clade models for *IGF1R* again indicated that the alternative model of divergent selection fit the data significantly better

than the null model of neutral evolution for bats and mole-rats (LRT: bats: $P<0.0001$; mole-rats: $P<0.0001$); similarly in both cases the estimated ω (dN/dS) on the foreground clade (FG) and background clade (BG) fell within the range of purifying selection (Bats: FGω = 0.219; BGω = 0.123; mole-rats: FGω = 0.102; BGω = 0.196).

Branch-site tests for positive selection acting on the ancestral bat or the ancestral mole-rat branches did not detect any significant positive selection in either *GHR* (LRT: ancestral bat: $P = 1.00$; ancestral mole-rat: $P = 0.17$) or *IGF1R* (LRT: ancestral bat: $P = 0.98$; ancestral mole-rat: $P = 1.00$). The majority of sites along each branch were found to be in site-class 0 and therefore, under purifying selection (*GHR*: ancestral bat: $\omega_0 = 0.106$ and ancestral mole-rat: $\omega_0 = 0.116$; *IGF1R*; ancestral bat: $\omega_0 = 0.023$ and ancestral mole-rat: $\omega_0 = 0.019$, see Appendix A: Supplementary Table S2B). Despite the levels of sequence variation shown by Vespertilionidae + Molossidae, branch-site models did not detect any significant positive selection acting on the ancestral branch in either gene (LRT: $P = 1.00$ in both *GHR* and *IGF1R*; see Appendix A: Supplementary Table S2B), with the majority of sites along each branch found to be under purifying selection ($\omega_0 = 0.101$ and $\omega_0 = 0.020$, in *GHR* and *IGF1R* respectively).

3.4. Levels of convergent sequence evolution

Tests for convergent and parallel amino acid substitutions based on branch-wise comparisons were conducted for both genes using the species tree topology. Plots of the summed posterior probability (PP) of total convergent (i.e. both convergent and parallel) substitutions versus summed posterior probability of divergent substitutions across all

placental mammal comparisons, excluding any comparisons with a terminal branch, revealed that the majority of substitutions along each gene were divergent (Figure 3).

In the case of GHR, the ancestral mole-rat branch versus the ancestral Vespertilionidae branch had the highest summed posterior probability of total convergent substitutions among all combinations of mole-rat versus bat branches compared (see Figure 3A). In total, three sites were detected with a posterior probability of undergoing convergent substitutions >0.20 (Tyr65: PP = 0.67; Cys223: PP = 0.95 and Val404: PP = 0.23). The second of these substitutions, Cys223, falls within the Fibronectin type 3 domains. The posterior probabilities suggest these substitutions are most likely to be parallel amino acid changes, i.e. arisen from the same ancestral state.

In comparison, the same analysis conducted for IGF1R revealed no such evidence of parallel substitutions between mole-rats and Vespertilionidae; summed posterior probability of total convergence = 0.002 (see Figure 3B). Out of all placental mammals sampled, the branch-pair comparison with the highest summed posterior probability of convergence was that between the ancestral Vespertilionidae and the ancestral Old World fruit bat branch, summed posterior probability of total convergence = 9.15. In total, 9 sites were detected with a posterior probability of undergoing convergent substitutions >0.20 (Ala185: PP = 0.97; Gly188: PP = 0.99; Phe678: PP = 1.00; Ile944: PP = 0.99; Ser963: PP = 0.87; Asp964: PP = 0.87; Asp1289: PP = 0.99; Arg1324: PP = 0.99 and Pro1358: PP = 1.00). Once again these were all parallel substitutions.

4. Discussion

4.1. Molecular evolution of GHR and IGF1R in bats and mole-rats

In this study we identified the coding sequence of two hormone receptor genes, *GHR* and *IGF1R*, in a comprehensive sampling of 26 bats and five mole-rats. By using a phylogenetic approach, as well as by focusing on individual species, we aimed to test for signatures of molecular adaptation linked to/associated with the increased longevity of these groups. We identified changes unique to members of both groups (for example, a deletion of Phe270 in the transmembrane domain of GHR in emballonurid bats), however, codon-based selection analyses did not identify evidence for positive selection in these focal groups compared to out-group species. Nevertheless, we found statistical support for three amino acids that may have undergone convergent substitutions along the ancestral mole-rat branch and the ancestral Vespertilionidae branch.

4.2. Reconciling GHR/IGF1R evolution with life-history traits in bats and mole-rats

Overall we found little evidence to suggest that the transmembrane domains of the GHR and IGF1R proteins have undergone evolutionary changes across bats that could be linked to their reduced body mass and increased longevity. Despite being the oldest and smallest-bodied Yinpterochiroptera bat examined, with a maximum lifespan of 30.5 years and body mass of 23 g (Wilkinson and South 2002), the greater horseshoe bat, *Rhinolophus ferrumequinum* was not seen to possess any unique amino acid substitutions compared to the remaining Yinpterochiroptera in either gene's transmembrane domain. Overall, the transmembrane domain of the IGF1R protein was found to be more highly

conserved across all bats and Laurasiatheria, when compared to that of GHR – with the exception of the *Myotis* bat branch. Although reliable longevity information was not available for all the *Myotis* bats examined in this study, as a group they are known to be small-bodied and typically long-lived; for example, *M. lucifugus* has a maximum recorded lifespan of 34 years and an adult body mass of around 10g (Wilkinson and South 2002). Given the previously documented roles of GH and IGF1 in regulating postnatal growth, it seems unlikely that these two genes alone are responsible for controlling the overall body size of a species. However, it is plausible that the previously documented amino acid substitutions in the transmembrane domains of both GHR and IGF1R in *Myotis* bats (Seim *et al.,* 2013) may confer a particular functional change in hormonal regulation in these species. In particular, this proposed functional change could relate to the metabolism of these species (especially as they are known to hibernate), although this remains experimentally untested. In general, it seems most likely that increased longevity, coupled with reduced body mass, evolved early in the evolutionary history of bats. Therefore, given that bats show no consistent variation in their GHR and IGF1R amino acid sequences compared to other mammals, and that most differences appear to have arisen among bat families, it seems doubtful that molecular evolution of these two hormone receptors has been the principal driving force behind longevity in bats.

Given the wide variation seen in longevity and body mass across rodents, the observed low number of amino acid differences in the two transmembrane domains across the clades is at first somewhat surprising. Nucleotide sequences of *GHR*, however, have been

frequently employed as a phylogenetic marker in rodents [for example (Adkins *et al.,* 2001; Galewski *et al.,* 2006)], and this gene is often chosen as it has a relatively low substitution rate thus reducing homoplasy and providing good resolution of taxa (Steppan *et al.,* 2004). This suggests that differences in body mass among rodents are related not to the gene sequence, but to expression or other mechanisms. It has been demonstrated that despite normal expression of GHR in guinea pig livers the animals appear to be resistant to the effects of GH, thus suggesting alternative regulatory pathways may be important (Hull *et al.,* 1996). Whether this is specific to guinea pigs or common to all hystricomorph rodents is currently unknown. In addition to being present as membrane bound dimers, the *GHR* mRNA can undergo proteolytical cleavage in humans and rabbits, or alternative processing in rodents, to generate growth hormone binding protein (GHBP) that circulates in the blood (Edens and Talamantes 1998; González *et al.,* 2007). In mice and other rodents, GHBP can also be present as a membrane-associated form (González *et al.,* 2007). All three forms bind with GH with a high affinity and so are all likely to play interrelated roles in the regulation of this hormone.

Conflicting evidence exists for the role that *IGF1R* plays in longevity in rodents; for example, female heterozygote knock-out mice live significantly longer than wild-types whereas, male heterozygote knock-out mice do not (Holzenberger *et al.,* 2003). However, no obvious phenotypic traits, such as dwarfism, were observed. A recent study demonstrated a negative correlation between expression levels of IGF1R in the brain and longevity across a number of diverse rodent species, including the naked mole-rat (Azpurua *et al.,* 2013). Moreover, because no such relationship was seen in tissue from

the heart, lung and kidney, the authors suggested that tissue-specific expression in the nervous tissue may be important to the evolution of longevity in mammals. Despite this, *IGF1R* was not one of the top 20 genes that displayed differential expression in the brain tissue of 2–3 year old naked mole-rats and 6.5 month old mice (Yu *et al.,* 2011). This finding suggests amino acid changes are not responsible for the increased longevity in African mole-rats, and instead other factors or mechanisms influencing gene regulation may be important, such as differential expression, copy number variation or epigenetic modifications. Alternatively, the increased longevity of mole-rats over muroid rodents may involve any number of alternative genes and/or pathways, potential candidates for which include those previously identified as displaying differential expression such as *EPCAM*, *SUCLG2* and *EIF4GL* (Yu *et al.,* 2011).

4.3. The wider roles of GHR and IGF1R in mammals

Growth hormone is the major regulator of postnatal growth (Yang *et al.,* 2007), such that sufferers diagnosed with Laron-type dwarfism are typically born with a normal mass and body size (Godowski *et al.,* 1989). This contrasts with insulin-like growth factor 1, which is principally involved in the regulation of growth during development, but also affects postnatal growth by interacting with GH. Aberrations in the receptors and associated pathways have been found to be associated with several forms of cancer and growth problems [e.g. (Adams *et al.,* 2000)]. Previous evidence, mainly from mutational studies, suggests that longevity can also be affected – with mutations in the genes relating to these two receptors seen to extend longevity [for example (Kaletsky and Murphy 2010; Junnila *et al.,* 2013)].

The structure of GHR and IGF1R has been established in several species – both proteins have a single transmembrane domain flanked by extracellular and cytoplasmic domains (Godowski *et al.,* 1989; Edens and Talamantes 1998; Adams *et al.,* 2000). In order to bind with their respective ligands, the receptors must be present as dimers. Experimental evidence has shown that the IGF1R transmembrane domain is important for activation and function of the receptor (Takahashi *et al.,* 1995). However the GHR transmembrane domain's role in receptor dimerization and activation is debated (Yang *et al.,* 2007). Experimental manipulation of the GHR transmembrane domain amino acid sequence found that its ability to pre-dimerize and therefore, potentially bind with GH, was not significantly affected (Yang *et al.,* 2007). Conversely, numerous mutations in the GHR extracellular domain – which is where the ligand binds – have been shown to result in Laron syndrome [e.g. (Amselem *et al.,* 1993; Pantel *et al.,* 2003)]. Therefore, although previous studies of these receptors in bats (Seim *et al.,* 2013), as well as this current one, have mainly focused on the amino acid variation in the transmembrane domain of GHR the significance of the detected substitutions in this domain remain far from clear. Furthermore we did not detect any sites under positive selection within the extracellular domain of either bats or mole-rats. However, as noted, experimental studies have not been performed on the previously detected bat amino acid substitutions (Seim *et al.,* 2013), so at present any structural and/or functional effects cannot be rules out.

5. Conclusions

Despite the dramatic variation in lifespan and body size seen across the mammals sampled by this study, both *GHR* and *IGF1R* were found to be under purifying selection in small-bodied, long-lived bats and mole-rats. However, we did detect several examples of family-specific amino acid substitutions in the transmembrane region of bat GHR, which could suggest that this gene may play a role in some aspect of the biology particular to Yangochiroptera. Little amino acid variation was found in the transmembrane domains of either GHR or IGF1R in long-lived mole-rats compared to much shorter-lived rodents. Therefore, evidence suggests that the sequence variation of *GHR* and *IGF1R* does not play a key role in the small body size and longevity seen in bats or mole-rats.

Acknowledgements

NCB Acknowledges funding to the DST-NRF SARChI Chair for Behavioural Ecology and Physiology. SJR, GT and KTJD were supported by the European Research Council. LMD was partially supported by the NSF (DEB-0949759). We also thank Burton K. Lim (Royal Ontario Museum) for donating tissue samples.

Supplementary files:

Appendix A: Supplementary data

Appendix B: Supplementary data

Alignments used for selection analyses in PAML.

(A) *GHR* – bats with 27 Laurasiatheria and Euarchontoglires

(B) *GHR* – mole-rats with 27 Laurasiatheria and Euarchontoglires

(C) *IGF1R* – bats with 23 Laurasiatheria and Euarchontoglires

(D) *IGF1R* – mole-rats with 23 Laurasiatheria and Euarchontoglires

Appendix C: Supplementary figure S1

Species tree topology used for selection analyses in PAML, and convergence analyses implemented in codeml-ancestral. Coloured branches represent foreground branches in clade models; bat clade – blue; mole-rat – pink. Numbered branches indicate branches set as foreground for branch-site models; common ancestral bat – 1; common ancestral mole-rat – 2; ancestral Vespertilionidae and Molossidae branch – 3. Key bat and mole-rat, suborders and families are labelled.

Appendix D: Supplementary figure S2

Amino acids alignments of the 24 amino acids that make up the transmembrane domains of (A) GHR and (B) IGF1R for non-bats and non-mole-rats species included in this study. In both cases, amino acid substitutions are shown relative to the reference sequence with conserved sites shown as empty coloured boxes. Dashed boxes indicate missing data.

Appendix E: Supplementary figure S3

Amino acid similarity scores calculated along (A) GHR and (B) IGF1R alignments including all taxa included in the study; with higher scores (~1.00) indicating high levels of conservation. Alignment position refers to amino acid column number following filtering of poorly aligned columns during the alignment process. The 24 amino acids that make up the transmembrane domains are indicated with black columns and labelled TM.

References:

AnAge Database of Animal Ageing and Longevity, Build 12, 16 September 2013

Adams TE, Epa VC, Garrett TP, Ward CW (2000). Structure and function of the type 1 insulin-like growth factor receptor. *Cell. Mol. Life Sci.* **57:** 1050–1093.

Adkins RM, Gelke EL, Rowe D, Honeycutt RL (2001). Molecular phylogeny and divergence time estimates for major rodent groups: evidence from multiple genes. *Mol. Biol. Evol.* **18:** 777–791.

Agnarsson I, Zambrana-Torrelio CM, Flores-Saldana NP, May-Collado LJ (2011). A time-calibrated species-level phylogeny of bats (Chiroptera, Mammalia). *PLoS Currents* **3:** RRN1212.

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W *et al* (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25:** 3389–3402.

Amselem S, Duquesnoy P, Duriez B, Dastot F, Sobrier ML, Valleix S *et al* (1993). Spectrum of growth hormone receptor mutations and associated haplotypes in Laron syndrome. *Hum. Mol. Genet.* **2:** 355–359.

Azpurua J, Yang J-N, Van Meter M, Liu Z, Kim J, Lobo Ladd AA *et al* (2013). IGF1R levels in the brain negatively correlate with longevity in 16 rodent species. *Aging (Albany NY)* **5** 304–314.

Bennett NC, Maree S, Faulkes CG (2006). *Georychus capensis*. *Mamm. Species* **799:** 1–4.

Bielawski JP, Yang ZH (2004). A maximum likelihood method for detecting functional divergence at individual codon sites, with application to gene family evolution. *J. Mol. Evol.* **59:** 121–132.

Blanga-Kanfi S, Miranda H, Penn O, Pupko T, DeBry RW, Huchon D (2009). Rodent phylogeny revised: analysis of six nuclear genes from all major rodent clades. *BMC Evol. Biol.* **9:** 71.

Brunet-Rossinni A, Wilkinson G (2009). Methods for age estimation and the study of senescence in bats. In: Kunz T and Parsons S (eds) *Ecological and behavioral methods for the study of bats*. The Johns Hopkins University Press: Baltimore, MD, USA, pp 315–325.

Brunet-Rossinni AK, Austad SN (2004). Ageing studies on bats: a review. *Biogerontology* **5:** 211–222.

Buffenstein R (2008). Negligible senescence in the longest living rodent, the naked mole-rat: insights from a successfully aging species. *J. Comp. Physiol. B* **178:** 439-445.

Capella-Gutierrez S, Silla-Martinez JM, Gabaldon T (2009). TrimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25:** 1972–1973.

Castoe TA, de Koning APJ, Kim HM, Gu WJ, Noonan BP, Naylor G *et al* (2009). Evidence for an ancient adaptive episode of convergent molecular evolution. *Proc. Natl. Acad. Sci. U. S. A.* **106:** 8986–8991.

Csiszar A, Labinskyy N, Orosz Z, Xiangmin Z, Buffenstein R, Ungvari Z (2007). Vascular aging in the longest-living rodent, the naked mole rat. *Am. J. Physiol. Heart Circ. Physiol.* **293:** H919–H927.

Edens A, Talamantes F (1998). Alternative processing of growth hormone receptor transcripts. *Endocr. Rev.* **19:** 559–582.

Edrey YH, Casper D, Huchon D, Mele J, Gelfond JA, Kristan DM *et al* (2012). Sustained high levels of neuregulin-1 in the longest-lived rodents; a key determinant of rodent longevity. *Aging Cell* **11:** 213–222.

Faulkes CG, Verheyen E, Verheyen W, Jarvis JUM, Bennett NC (2004). Phylogeographical patterns of genetic divergence and speciation in African mole-rats (Family: Bathyergidae). *Mol. Ecol.* **13:** 613-629.

Flicek P, Ahmed I, Amode MR, Barrell D, Beal K, Brent S *et al* (2013). Ensembl 2013. *Nucleic Acids Res.* **41:** D48–D55.

Flurkey K, Papaconstantinou J, Miller RA, Harrison DE (2001). Lifespan extension and delayed immune and collagen aging in mutant mice with defects in growth hormone production. *Proc. Natl. Acad. Sci. U. S. A.* **98:** 6736–6741.

Galewski T, Tilak M-k, Sanchez S, Chevret P, Paradis E, Douzery EJP (2006). The evolutionary radiation of Arvicolinae rodents (voles and lemmings): relative contribution of nuclear and mitochondrial DNA phylogenies. *BMC Evol. Biol.* **6:** 80.

Godowski PJ, Leung DW, Meacham LR, Galgani JP, Hellmiss R, Keret R *et al* (1989). Characterization of the human growth hormone receptor gene and demonstration of a partial gene deletion in two patients with Laron-type dwarfism. *Proc. Natl. Acad. Sci. U. S. A.* **86:** 8083–8087.

González L, Curto LM, Miquet JG, Bartke A, Turyn D, Sotelo AI (2007). Differential regulation of membrane associated-growth hormone binding protein (MA-GHBP)

and growth hormone receptor (GHR) expression by growth hormone (GH) in mouse liver. *Growth Horm. IGF Res.* **17:** 104–112.

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I *et al* (2011). Full-length transcriptome assembly from RNA-seq data without a reference genome. *Nat. Biotechnol.* **29:** 644–652.

Greer KA, Canterberry SC, Murphy KE (2007). Statistical analysis regarding the effects of height and weight on life span of the domestic dog. *Res. Vet. Sci.* **82:** 208–214.

Healy K, Guillerme T, Finlay S, Kane A, Kelly SBA, McClean D *et al* (2014). Ecology and mode-of-life explain lifespan variation in birds and mammals. *Proc. R. Soc. B* **281:** 20140298.

Holzenberger M, Dupont J, Ducos B, Leneuve P, Géloën A, Even PC *et al* (2003). IGF-1 receptor regulates lifespan and resistance to oxidative stress in mice. *Nature* **421:** 182–187.

Hoopes BC, Rimbault M, Liebers D, Ostrander EA, Sutter NB (2012). The insulin-like growth factor 1 receptor (*IGF1R*) contributes to reduced size in dogs. *Mamm. Genome* **23:** 780–790.

Hull KL, Janssens WC, Baumbach WR, Harvey S (1996). Growth hormone (GH)-binding proteins in GH-resistant guinea pigs. *Growth Regul.* **6:** 165–175.

Junnila RK, List EO, Berryman DE, Murrey JW, Kopchick JJ (2013). The GH/IGF-1 axis in ageing and longevity. *Nat. Rev. Endocrinol.* **9:** 366–376.

Kaletsky R, Murphy CT (2010). The role of insulin/IGF-like signaling in *C. elegans* longevity and aging. *Dis. Model. Mech.* **3:** 415–419.

Kenyon CJ (2010). The genetics of ageing. *Nature* **464:** 504–512.

Kim EB, Fang X, Fushan AA, Huang Z, Lobanov AV, Han L *et al* (2011). Genome sequencing reveals insights into physiology and longevity of the naked mole rat. *Nature* **479:** 223–227.

Kirkegaard M, Jørgensen J (2000). Continuous hair cell turnover in the inner ear vestibular organs of a mammal, the Daubenton's bat (*Myotis daubentonii*). *Naturwissenschaften* **87:** 83–86.

Liang S, Mele J, Wu Y, Buffenstein R, Hornsby PJ (2010). Resistance to experimental tumorigenesis in cells of a long-lived mammal, the naked mole-rat (*Heterocephalus glaber*). *Aging Cell* **9:** 626–635.

Löytynoja A, Goldman N (2005). An algorithm for progressive multiple alignment of sequences with insertions. *Proc. Natl. Acad. Sci. USA* **102:** 10557–10562.

Meredith RW, Janecka JE, Gatesy J, Ryder OA, Fisher CA, Teeling EC *et al* (2011). Impacts of the Cretaceous Terrestrial Revolution and KPg extinction on mammal diversification. *Science* **334:** 521–524.

Morgan CC, McCartney AM, Donoghue MT, Loughran NB, Spillane C, Teeling EC *et al* (2013). Molecular adaptation of telomere associated genes in mammals. *BMC Evol. Biol.* **13:** 251.

Mqokeli BR, Downs CT (2012). Blood plasma glucose regulation in Wahlberg's epauletted fruit bat. *Afr. Zool.* **47**.

Pantel J, Grulich-Henn J, Bettendorf M, Strasburger CJ, Heinrich U, Amselem S (2003). Heterozygous nonsense mutation in exon 3 of the growth hormone receptor (GHR) in severe GH insensitivity (Laron syndrome) and the issue of the origin and function of the GHRd3 isoform. *J. Clin. Endocrinol. Metab.* **88:** 1705–1710.

Penn O, Privman E, Landan G, Graur D, Pupko T (2010). An alignment confidence score capturing robustness to guide-tree uncertainty. *Mol. Biol. Evol.* **27:** 1759–1767.

Podlutsky AJ, Khritankov AM, Ovodov ND, Austad SN (2005). A new field record for bat longevity. *J. Gerontol. A Biol. Sci. Med. Sci.* **60:** 1366–1368.

Schmidt CM, Jarvis JUM, Bennett NC (2013). The long-lived queen: reproduction and longevity in female eusocial Damaraland mole-rats (*Fukomys damarensis*). *Afr. Zool.* **48:** 193–196.

Seim I, Fang X, Xiong Z, Lobanov AV, Huang Z, Ma S *et al* (2013). Genome analysis reveals insights into the evolution, physiology, and longevity of the Brandt's bat. *Nat. Commun.* **4:** 2212.

Stamatakis A (2006). RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22:** 2688–2690.

Steppan S, Adkins R, Anderson J (2004). Phylogeny and divergence-date estimates of rapid radiations in muroid rodents based on multiple nuclear genes. *Syst. Biol.* **53:** 533–553.

Sutter NB, Bustamante CD, Chase K, Gray MM, Zhao K, Zhu L *et al* (2007). A single IGF1 allele is a major determinant of small size in dogs. *Science* **316:** 112–115.

Szewczuk M, Zych S, Wójcik J, Czerniawska-Piątkowska E (2013). Association of two SNPs in the coding region of the insulin-like growth factor 1 receptor (IGF1R) gene with growth-related traits in Angus cattle. *J. Appl. Genet.* **54:** 305–308.

Tacutu R, Craig T, Budovsky A, Wuttke D, Lehmann G, Taranukha D *et al* (2013). Human Ageing Genomic Resources: Integrated databases and tools for the biology and genetics of ageing. *Nucleic Acids Res.* **41:** D102–D1033.

Takahashi K, Yonezawa K, Nishimoto I (1995). Insulin-like growth factor I receptor activated by a transmembrane mutation. *J. Biol. Chem.* **270:** 19041-19045.

Teeling EC, Springer MS, Madsen O, Bates P, O'Brien SJ, Murphy WJ (2005). A molecular phylogeny for bats illuminates biogeography and the fossil record. *Science* **307:** 580–584.

Tsagkogeorga G, Parker J, Stupka E, Cotton JA, Rossiter SJ (2013). Phylogenomic analyses elucidate the evolutionary relationships of bats. *Curr. Biol.* **23:** 2262–2267.

Turbill C, Bieber C, Ruf T (2011). Hibernation is associated with increased survival and the evolution of slow life histories among mammals. *Proc. R. Soc. B* **278:** 3355–3363.

Ullrich A, Gray A, Tam AW, Yang-Feng T, Tsubokawa M, Collins C *et al* (1986). Insulin-like growth factor I receptor primary structure: comparison with insulin receptor suggests structural determinants that define functional specificity. *Embo Journal* **5:** 2503–2512.

Widmaier E, Gornstein E, Hennessey J, Bloss J, Greenberg J, Kunz T (1996). High plasma cholesterol, but low triglycerides and plaque-free arteries, in Mexican free-tailed bats. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **40:** R1101–R1106.

Wilkinson GS, South JM (2002). Life history, ecology and longevity in bats. *Aging Cell* **1:** 124–131.

Williams GC (1957). Pleiotropy, natural selection, and the evolution of senescence. *Evolution* **11:** 398–411

Yang N, Wang X, Jiang J, Frank SJ (2007). Role of the growth hormone (GH) receptor transmembrane domain in receptor predimerization and GH-induced activation. *Mol. Endocrinol.* **21:** 1642–1655.

Yang ZH (2007). PAML 4: Phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24:** 1586–1591.

Yu C, Li Y, Holmes A, Szafranski K, Faulkes CG, Coen CW *et al* (2011). RNA sequencing reveals differential expression of mitochondrial and oxidation reduction genes in the long-lived naked mole-rat when compared to mice. *PLoS ONE* **6:** e26729.

Zhang JZ, Nielsen R, Yang ZH (2005). Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. *Mol. Biol. Evol.* **22:** 2472–2479.

Figure legends

Figure 1

Maximum likelihood phylogenetic trees based on the coding sequences of (A) *GHR* and (B) *IGF1R.* Bootstrap support values for each node are represented by shaded circles (support values: ≥95% – black; ≥50% – grey and <50% – white circles). Coloured panels represent alignments of the 24 amino acids that make up the transmembrane domain of each gene for key taxa numbered in the phylogeny, in each case amino acid substitutions are shown relative to the reference sequence with conserved sites shown as empty coloured boxes. Dashed boxes indicate missing data. Bat species belonging to the two bat suborders are indicated by the black bar – Yinpterochiroptera and grey bar – Yangochiroptera.

Figure 2

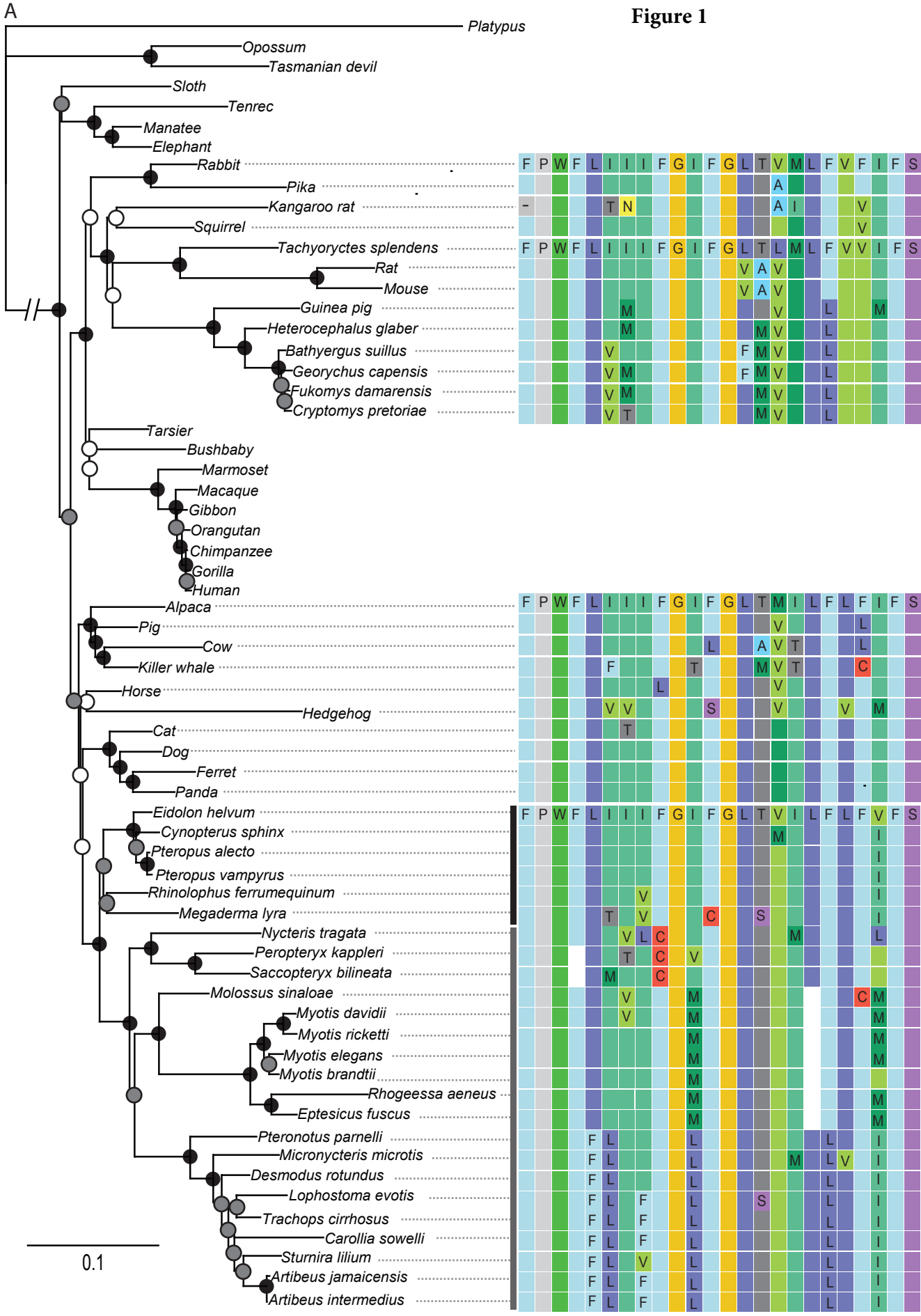
Scatter plot showing the relationship between maximum recorded lifespan (years) against adult body weight (grams) for the species included in this study, data from AnAge (Tacutu *et al.,* 2013); Yangochiroptera bats (grey points); Old World fruit bats (yellow); echolocating Yinpterochiroptera (blue points), mole-rats (black points) and remaining mammals (white points). GC – *Georychus capensis*; HG – *Heterocephalus glaber*; FD – *Fukomys damarensis*; MB – *Myotis brandtii* and ML – *Myotis lucifugus*.

Figure 3

Plots of the summed posterior probability of total convergent (i.e. both convergent and parallel) substitutions versus summed posterior probability of divergent substitutions

across all placental mammal comparisons, excluding any comparisons with a terminal branch for (A) *GHR* and (B) *IGF1R*. Placental mammal pair-wise comparisons (white); mole-rat vs. bat pair-wise comparisons (blue) and ancestral mole-rat vs. ancestral Vespertilionidae (black).

Figure 1


A
Platypus
Opossum
Tasmanian devil
Sloth
Tenrec
Manatee
Elephant
Rabbit
Pika
Kangaroo rat
Squirrel
Tachyoryctes splendens
Rat
Mouse
Guinea pig
Heterocephalus glaber
Bathyergus suillus
Georychus capensis
Fukomys damarensis
Cryptomys pretoriae
Tarsier
Bushbaby
Marmoset
Macaque
Gibbon
Orangutan
Chimpanzee
Gorilla
Human
Alpaca
Pig
Cow
Killer whale
Horse
Hedgehog
Cat
Dog
Ferret
Panda
Eidolon helvum
Cynopterus sphinx
Pteropus alecto
Pteropus vampyrus
Rhinolophus ferrumequinum
Megaderma lyra
Nycteris tragata
Peropteryx kappleri
Saccopteryx bilineata
Molossus sinaloae
Myotis davidii
Myotis ricketti
Myotis elegans
Myotis brandtii
Rhogeessa aeneus
Eptesicus fuscus
Pteronotus parnelli
Micronycteris microtis
Desmodus rotundus
Lophostoma evotis
Trachops cirrhosus
Carollia sowelli
Sturnira lilium
Artibeus jamaicensis
Artibeus intermedius
0.1

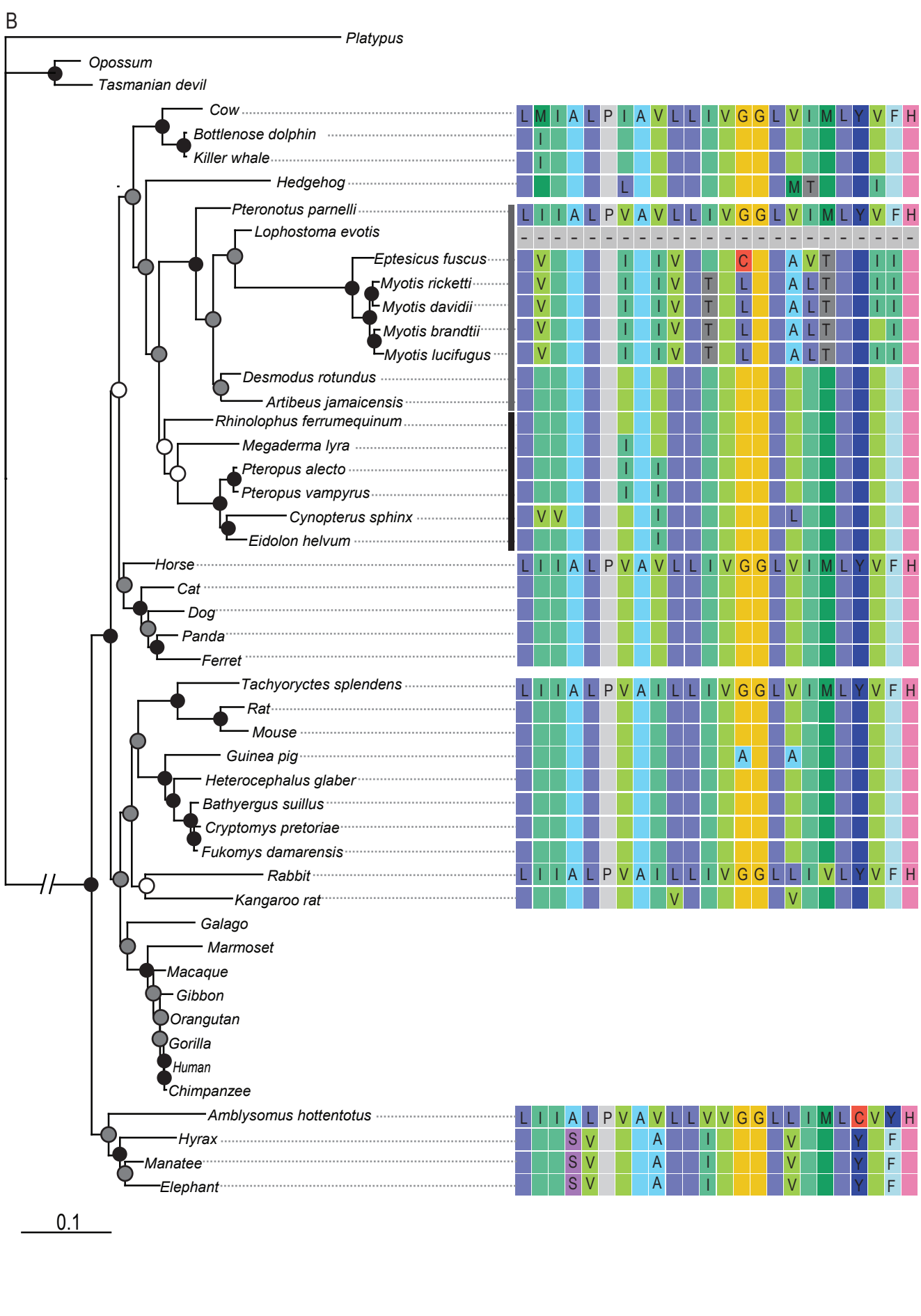

B
Platypus
Opossum
Tasmanian devil
Cow
Bottlenose dolphin
Killer whale
Hedgehog
Pteronotus parnelli
Lophostoma evotis
Eptesicus fuscus
Myotis ricketti
Myotis davidii
Myotis brandtii
Myotis lucifugus
Desmodus rotundus
Artibeus jamaicensis
Rhinolophus ferrumequinum
Megaderma lyra
Pteropus alecto
Pteropus vampyrus
Cynopterus sphinx
Eidolon helvum
Horse
Cat
Dog
Panda
Ferret
Tachyoryctes splendens
Rat
Mouse
Guinea pig
Heterocephalus glaber
Bathyergus suillus
Cryptomys pretoriae
Fukomys damarensis
Rabbit
Kangaroo rat
Galago
Marmoset
Macaque
Gibbon
Orangutan
Gorilla
Human
Chimpanzee
Amblysomus hottentotus
Hyrax
Manatee
Elephant
0.1

Figure 2

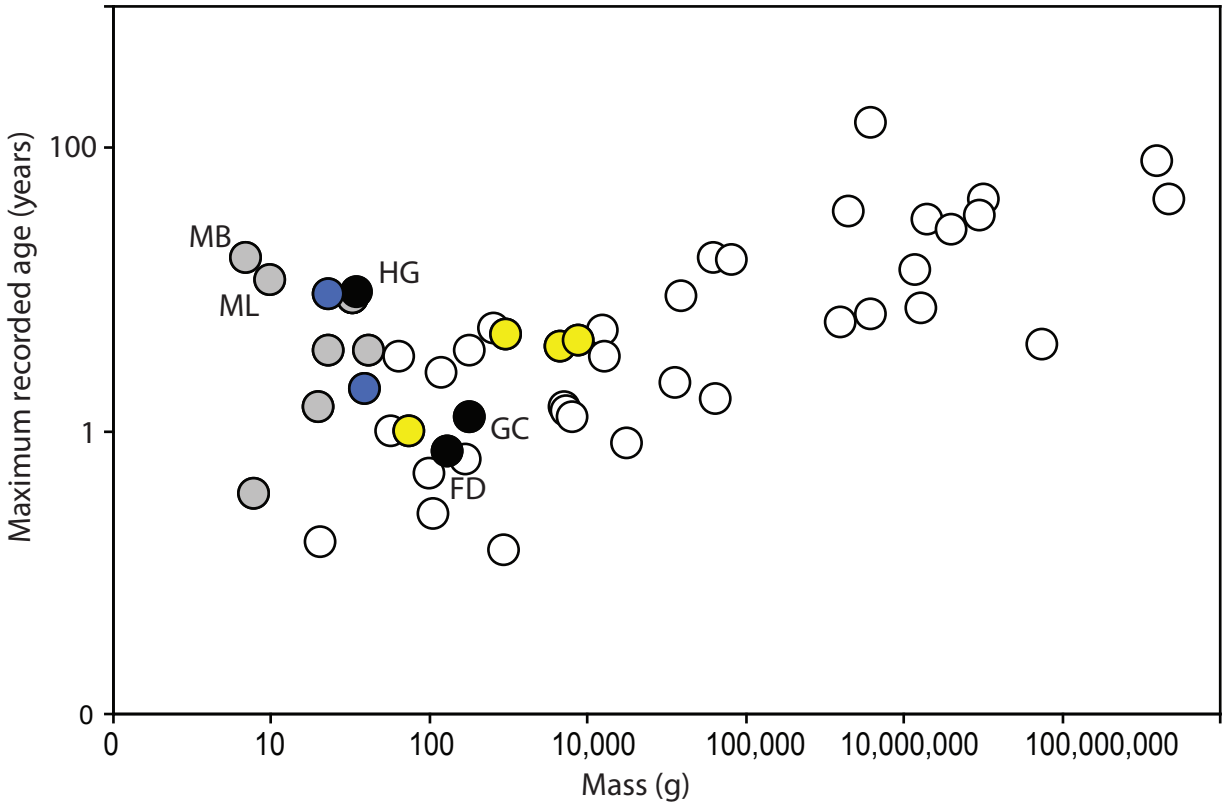
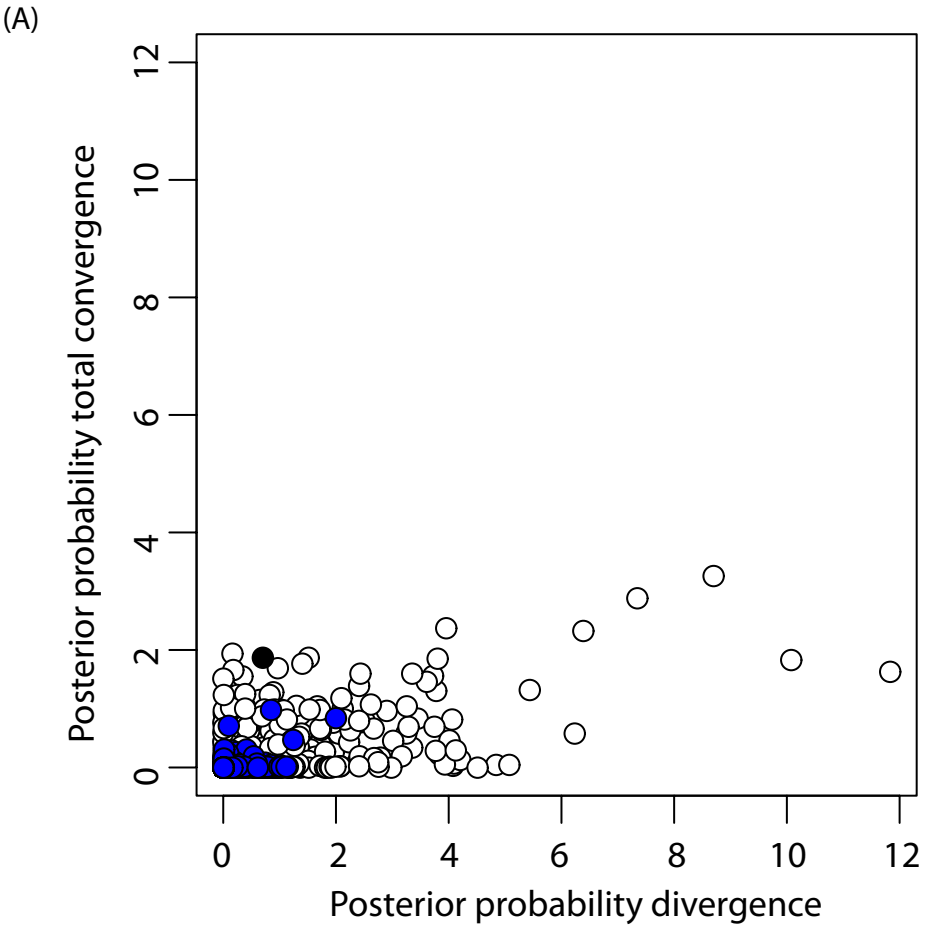
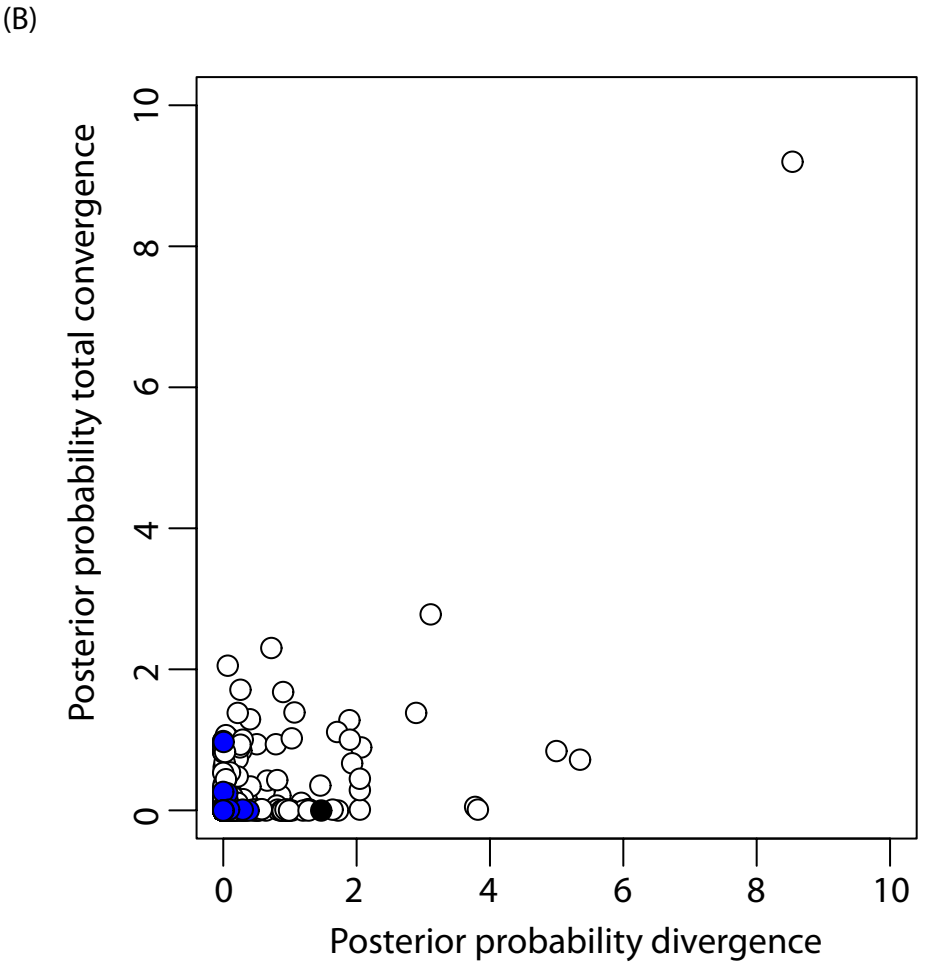

Maximum recorded age (years)
100
1
0
MB
ML
HG
GC
FD
0
10
100
10,000
100,000
10,000,000
100,000,000
Mass (g)

Figure 3


(A)
Posterior probability total convergence
0
2
4
6
8
10
12
Posterior probability divergence
0
2
4
6
8
10
12


(B)
Posterior probability total convergence
0
2
4
6
8
10
Posterior probability divergence
0
2
4
6
8
10

APPENDIX A: SUPPLEMENTARY INFORMATION

Molecular evolution of growth hormone and insulin-like growth factor 1 receptors in long-lived, small-bodied mammals

Kalina T.J. Davies, Georgia Tsagkogeorga, Nigel C. Bennett, Liliana M. Dávalos, C.G. Faulkes, Stephen J. Rossiter

Supplementary Table S1
Taxonomic and life history information for the species included in the study
The SRA identifier or assembly number and accompanying publication are provided for published genomes and transcriptomes, Ensembl short codes (Flicek *et al.,* 2013) and GenBank accession numbers are provided for the remaining sequences (novel sequences generated for this study are indicated with *). Maximum life-spans and body mass information was obtained from the AnAge database (Tacutu *et al.,* 2013) and (Schmidt *et al.,* 2013). Abbreviations: yrs – years; g – grams; NA – not available.

Order	Family	Species	Common name	Age (yrs)	Mass (g)	Dataset/gene identifier (*GHR, IGF1R*)
Monotremata	Ornithorhynchidae	*Ornithorhynchus anatinus*	Platypus	22.6	1,250	ENSOANG00000021687, ENSOANG00000000134
Didelphimorphia	Didelphidae	*Monodelphis domestica*	Gray short-tailed opossum	5.1	105	ENSMODG00000020284, ENSMODG00000012761
Dasyuromorphia	Dasyuridae	*Sarcophilus harrisii*	Tasmanian devil	13	6,500	ENSSHAG00000012829, ENSSHAG00000013808
Afrosoricida	Tenrecidae	*Echinops telfairi*	Lesser hedgehog tenrec	19	180	ENSETEG00000011018, NA
	Chrysochloridae	*Amblysomus hottentotus*	Hottentot golden mole	-	-	NA, KM190086*
Proboscidea	Elephantidae	*Loxodonta Africana*	African elephant	65	4,800,000	ENSLAFG00000025524, ENSLAFG00000011658
Procaviidae	Hyracoidea	*Procavia capensis*	Rock hyrax	14.8	3,600	NA, ENSPCAG00000016571
Pilosa	Megalonychidae	*Choloepus hoffmanni*	Hoffmann's two-toed sloth	41	6,250	ENSCHOG00000007988. NA
Sirenia	Trichechidae	*Trichechus manatus*	West Indian manatee	65	322,000	XM_004388441, XM_004372657.1
Primates	Hominidae	*Homo sapiens*	Human	122.5	62,035	ENSG00000112964, ENSG00000140443
		Pan troglodytes	Chimpanzee	59.4	44,984	ENSPTRG00000016836. ENSPTRG00000007489
		Gorilla gorilla gorilla	Gorilla	55.4	139,842	ENSGGOG00000012636, ENSGGOG00000012828

		Pongo abelii	Orangutan	-	-	ENSPPYG00000015431, ENSPPYG00000006808
	Hylobatidae	*Nomascus leucogenys*	Gibbon	-	-	ENSNLEG00000016678, ENSNLEG00000000814
	Cercopithecidae	*Macaca mulatta*	Macaque	40	8,235	ENSMMUG00000001336, ENSMMUG00000012305
	Callitrichidae	*Callithrix jacchus*	Marmoset	22.8	255.2	ENSCJAG00000001237, ENSCJAG00000016895
	Cheirogaleidae	*Microcebus murinus*	Mouse lemur	18.2	64.8	NA, ENSMICG00000010119
	Galagidae	*Otolemur garnettii*	Bushbaby	18.3	1,300	ENSOGAG00000012855, NA
	Tarsiidae	*Tarsius syrichta*	Tarsier	16	119.2	ENSTSYG00000011979, NA
Rodentia	Sciuridae	*Ictidomys tridecemlineatus*	Squirrel	7.9	172.7	ENSSTOG00000007190, NA
	Heteromyidae	*Dipodomys ordii*	Kangaroo rat	9.9	57	ENSDORG00000008027, ENSDORG00000004530
	Muridae	*Mus musculus*	Mouse	4	20.5	ENSMUSG00000055737, ENSMUSG00000005533
		Rattus norvegicus	Rat	3.8	300	ENSRNOG00000015654, ENSRNOG00000014187
	Spalacidae	*Tachyoryctes splendens*	East African root rat	-	220	KM190081*, KM190090*
	Caviidae	*Cavia porcellus*	Guinea pig	12	728	ENSCPOG00000014345, ENSCPOG00000005399
	Bathyergidae	*Bathyergus suillus*	Cape dune mole-rat	-	-	KM190082*, KM190088*
		Cryptomys pretoriae	Highveld mole-rat	-	-	KM190083*, KM190087*
		Fukomys damarensis	Damaraland mole-rat	8.5	>130	KM190084*, KM190089*
		Georychus capensis	Cape mole-rat	11.2	181	KM190085*, NA
		Heterocephalus glaber	Naked mole-rat	31	35	XM_004848566.1, XM_004852705.1
Lagomorpha	Ochotonidae	*Ochotona princeps*	Pika	7	100	ENSOPRG00000016585, NA
	Leporidae	*Oryctolagus cuniculus*	Rabbit	9	1,800	ENSOCUG00000008496, ENSOCUG00000014795
Erinaceidae	Erinaceidae	*Erinaceus europaeus*	Hedgehog	11.7	750	ENSEEUG00000011363, ENSEEUG00000013409

Chiroptera	Pteropodidae	*Pteropus alecto*	Black flying fox	19.7	672	SRR628071 and ASM32557v1 (Zhang *et al.,* 2013)
		Pteropus vampyrus	Large flying fox	20.9	872	ENSPVAG00000005609, ENSPVAG00000003279
		Eidolon helvum	Straw-coloured fruit bat	21.8	306	ASM46528v1 (Tsagkogeorga *et al.,* 2013)
		Cynopterus sphinx	Greater short-nosed fruit bat	10	75	SRR837385 (Dong *et al.,* 2013)
	Rhinolophidae	*Rhinolophus ferrumequinum*	Greater horseshoe bat	30.5	23	ASM46549v1 (Tsagkogeorga *et al.,* 2013)
	Megadermatidae	*Megaderma lyra*	Greater false vampire bat	14	39	ASM46534v1 (Tsagkogeorga *et al.,* 2013)
	Phyllostomidae	*Artibeus intermedius*	Great fruit-eating bat	-	-	KM190096*, NA
		Artibeus jamaicensis	Common fruit bat	19.2	42	SRP014960 (Shaw *et al.,* 2012)
		Carollia sowelli	Sowell's short-tailed bat	-	-	KM190100*, NA
		Lophostoma evotis	Davis's round-eared bat	-	-	KM190101*, KM190092*
		Micronycteris microtis	Common big-eared bat	-	-	KM190102 *, NA
		Sturnira lilium	Little yellow-shouldered bat	12	20	KM190103 *, NA
		Trachops cirrhosus	Fringe-lipped bat	-	-	KM190095*, NA
		Desmodus rotundus	Common vampire bat	29.2	33	KM190094*, KM190091*
	Vespertilionidae	*Rhogeessa aeneus*	Yucatan yellow bat	-	-	KM190098*, NA
		Myotis elegans	Elegant myotis	-	-	KM190097*, NA
		Myotis ricketti	Rickett's big-footed bat	-	-	SRR837386 (Dong *et al.,* 2013)
		Myotis davidii	David’s myotis	-	4	SRP014729 and ASM32734v1 (Zhang *et al.,* 2013)
		Myotis lucifugus	Little brown bat	34	10	NA, ENSMLUG00000017190
		Myotis brandtii	Brandt's bat	41	7	ASM41265v1 (Seim *et al.,* 2013)
		Eptesicus fuscus	Big brown bat	19	23	EptFus1.0
	Molossidae	*Molossus sinaloae*	Sinaloan mastiff bat	-	-	KM190099*, NA
	Mormoopidae	*Pteronotus parnellii*	Parnell's mustached bat	-	-	ASM46540v1 (Tsagkogeorga *et al.,* 2013)

	Emballonuridae	*Peropteryx kappleri*	Greater dog-like bat	-	-	KM190104 *, NA
		Saccopteryx bilineata	Greater sac-winged bat	6	7.9	KM190105 *, NA
	Nycteridae	*Nycteris tragata*	Malayan slit-faced bat	-	-	KM190093*, NA
Perissodactyla	Equidae	*Equus caballus*	Horse	57	300,000	ENSECAG00000002986, ENSECAG00000021238
Cetartiodactyla	Bovidae	*Bos taurus*	Cow	20	750,000	ENSBTAG00000001335, ENSBTAG00000021527
	Camelidae	*Vicugna pacos*	Alpaca	25.8	62,000	ENSVPAG00000002555. NA
	Suidae	*Sus scrofa*	Pig	27	130,000	ENSSSCG00000016866, NA
	Delphinidae	*Orcinus orca*	Killer whale	90	3,987,500	XM_004265958.1, XM_004271659.1
		Tursiops truncatus	Bottlenose dolphin	51.6	200,000	NA, ENSTTRG00000014670
Carnivora	Ursidae	*Ailuropoda melanoleuca*	Panda	36.8	117,500	ENSAMEG00000003826, ENSAMEG00000005572
	Felidae	*Felis catus*	Cat	30	3,900	ENSFCAG00000026499, ENSFCAG00000018164
	Canidae	*Canis lupus familiaris*	Dog	24	40,000	ENSCAFG00000018579, ENSCAFG00000010881
	Mustelidae	*Mustela putorius furo*	Ferret	11.1	809	ENSMPUG00000014445, ENSMPUG00000010753

Supplementary Table S2
(A) Results of clade-models.
All likelihood ratio tests significance levels determined with 3 degrees of freedom.
Abbreviations: *np* – number of parameters; lnL – log likelihood; BG – background; FG – foreground; *p* – proportion; ω – dN/dS; *P* – p-value.

Gene	Comparison	Model	*np*	-lnL	Model parameters:	2Δ lnL	*P*
GHR	Bats vs. 27 Laurasiatheria and Euarchontoglires	M1A	108	21285.37	$p_0 =$ 0.72 ($p_1 = 0.28$), $\omega_0 = 0.11$, $\omega_1 = 1.00$	204.23	<0.001
		C	105	21387.48	$p_0 =$ 0.50, $p_1 =$ 0.13, $p_2 = 0.37$ BG: $\omega_0 = 0.04$, $\omega_1 = 1.00$, $\omega_2 = 0.40$ FG: $\omega_0 = 0.04$, $\omega_1 = 1.00$, $\omega_2 = 0.29$		
GHR	Mole-rats vs. 27 Laurasiatheria and Euarchontoglires	M1A	65	15132.31	$p_0 =$ 0.72 ($p_1 = 0.28$), $\omega_0 = 0.12$, $\omega_1 = 1.00$	89.17	<0.001
		C	68	15087.72	$p_0 =$ 0.43, $p_1 =$ 0.16, $p_2 = 0.42$ BG: $\omega_0 = 0.02$, $\omega_1 = 1.00$, $\omega_2 = 0.32$ FG: $\omega_0 = 0.02$, $\omega_1 = 1.00$, $\omega_2 = 0.49$		
IGF1R	Bats vs. 23 Laurasiatheria and Euarchontoglires	M1A	77	27962.33	$p_0 =$ 0.97 ($p_1 = 0.03$), $\omega_0 = 0.02$, $\omega_1 = 1.00$	748.14	<0.001
		C	80	27588.26	$p_0 =$ 0.84, $p_1 =$ 0.01, $p_2 = 0.16$ BG: $\omega_0 = 0.01$, $\omega_1 = 1.00$, $\omega_2 = 0.12$ FG: $\omega_0 = 0.01$, $\omega_1 = 1.00$, $\omega_2 = 0.22$		
IGF1R	Mole-rats vs. 23 Laurasiatheria and Euarchontoglires	M1A	55	21150.77	$p_0 =$ 0.97 ($p_1 = 0.03$), $\omega_0 = 0.02$, $\omega_1 = 1.00$	240.31	<0.001
		C	58	21030.61	$p_0 =$ 0.88, $p_1 =$ 0.00, $p_2 = 0.11$ BG: $\omega_0 = 0.01$, $\omega_1 = 1.00$, $\omega_2 = 0.20$ FG: $\omega_0 = 0.01$, $\omega_1 = 1.00$, $\omega_2 = 0.10$		

(B) Branch-site model results

All likelihood ratio tests significance levels determined with 1 degree of freedom.

Abbreviations: *np* – number of parameters; lnL – log likelihood; BG – background; FG – foreground; *p* – proportion; ω – dN/dS; *P* – p-value.

Gene	Focal branch	Model A	*np*	-lnL	Model parameters:	2ΔlnL	*P*
GHR	Common ancestral Molossidae + Vespertilionidae	Null	52	9741.24	$p_0 = 0.83$, $p_1 = 0.17$, $p_{2a} = 0.00$, $p_{2b} = 0.00$ BG: $\omega_0 = 0.10$, $\omega_1 = 1.00$, $\omega_{2a} = 0.10$, $\omega_{2b} = 1.00$ FG: $\omega_0 = 0.10$, $\omega_1 = 1.00$, $\omega_{2a} = 1.00$, $\omega_{2b} = 1.00$	0.00	1.000
		Alternative	53	9741.24	$p_0 = 0.83$, $p_1 = 0.17$, $p_{2a} = 0.00$, $p_{2b} = 0.00$ BG: $\omega_0 = 0.10$, $\omega_1 = 1.00$, $\omega_{2a} = 0.10$, $\omega_{2b} = 1.00$ FG: $\omega_0 = 0.10$, $\omega_1 = 1.00$, $\omega_{2a} = 1.00$, $\omega_{2b} = 1.00$		
GHR	Common ancestral mole-rat	Null	66	15131.35	$p_0 = 0.61$, $p_1 = 0.23$, $p_{2a} = 0.12$, $p_{2b} = 0.04$ BG: $\omega_0 = 0.12$, $\omega_1 = 1.00$, $\omega_{2a} = 0.12$, $\omega_{2b} = 1.00$ FG: $\omega_0 = 0.12$, $\omega_1 = 1.00$, $\omega_{2a} = 1.00$, $\omega_{2b} = 1.00$	1.90	0.168
		Alternative	67	15130.40	$p_0 = 0.70$, $p_1 = 0.27$, $p_{2a} = 0.02$, $p_{2b} = 0.01$ BG: $\omega_0 = 0.12$, $\omega_1 = 1.00$, $\omega_{2a} = 0.12$, $\omega_{2b} = 1.00$ FG: $\omega_0 = 0.12$, $\omega_1 = 1.00$, $\omega_{2a} = 14.10$, $\omega_{2b} = 14.10$		
GHR	Common ancestral bat	Null	106	21387.48	$p_0 = 0.72$, $p_1 = 0.28$, $p_{2a} = 0.00$, $p_{2b} = 0.00$ BG: $\omega_0 = 0.11$, $\omega_1 = 1.00$, $\omega_{2a} = 0.11$, $\omega_{2b} = 1.00$ FG: $\omega_0 = 0.11$, $\omega_1 = 1.00$, $\omega_{2a} = 1.00$, $\omega_{2b} = 1.00$	0.00	1.000
		Alternative	107	21387.48	$p_0 = 0.72$, $p_1 = 0.28$, $p_{2a} = 0.00$, $p_{2b} = 0.00$ BG: $\omega_0 = 0.11$, $\omega_1 = 1.00$, $\omega_{2a} = 0.11$, $\omega_{2b} = 1.00$ FG: $\omega_0 = 0.11$, $\omega_1 = 1.00$, $\omega_{2a} = 1.00$, $\omega_{2b} = 1.00$		
IGF1R	Common ancestral Vespertilionidae	Null	32	12880.28	$p_0 = 0.91$, $p_1 = 0.03$, $p_{2a} = 0.06$, $p_{2b} = 0.00$ BG: $\omega_0 = 0.02$, $\omega_1 = 1.00$, $\omega_{2a} = 0.02$, $\omega_{2b} = 1.00$ FG: $\omega_0 = 0.02$, $\omega_1 = 1.00$, $\omega_{2a} = 1.00$, $\omega_{2b} = 1.00$	0.00	1.000
IGF1R	Common ancestral mole-rat	Alternative Null	33 56	12880.28 21150.77	$p_0 = 0.91$, $p_1 = 0.03$, $p_{2a} = 0.06$, $p_{2b} = 0.00$ BG: $\omega_0 = 0.02$, $\omega_1 = 1.00$, $\omega_{2a} = 0.02$, $\omega_{2b} = 1.00$ FG: $\omega_0 = 0.02$, $\omega_1 = 1.00$, $\omega_{2a} = 1.00$, $\omega_{2b} = 1.00$ $p_0 = 0.97$, $p_1 = 0.03$, $p_{2a} = 0.00$, $p_{2b} = 0.00$ BG: $\omega_0 = 0.02$, $\omega_1 = 1.00$, $\omega_{2a} = 0.02$, $\omega_{2b} = 1.00$ FG: $\omega_0 = 0.02$, $\omega_1 = 1.00$, $\omega_{2a} = 1.00$, $\omega_{2b} = 1.00$	0.00	0.999

		Alternative	57	21150.77	$p_0 = 0.97$, $p_1 = 0.03$, $p_{2a} = 0.00$, $p_{2b} = 0.00$ BG: $\omega_0 = 0.02$, $\omega_1 = 1.00$, $\omega_{2a} = 0.02$, $\omega_{2b} = 1.00$ FG: $\omega_0 = 0.02$, $\omega_1 = 1.00$, $\omega_{2a} = 1.00$, $\omega_{2b} = 1.00$		
IGF1R	Common ancestral bat	Null	78	27962.33	$p_0 = 0.97$, $p_1 = 0.03$, $p_{2a} = 0.00$, $p_{2b} = 0.00$ BG: $\omega_0 = 0.02$, $\omega_1 = 1.00$, $\omega_{2a} = 0.02$, $\omega_{2b} = 1.00$ FG: $\omega_0 = 0.02$, $\omega_1 = 1.00$, $\omega_{2a} = 1.00$, $\omega_{2b} = 1.00$	0.00	0.978
		Alternative	79	27962.33	$p_0 = 0.97$, $p_1 = 0.03$, $p_{2a} = 0.00$, $p_{2b} = 0.00$ BG: $\omega_0 = 0.02$, $\omega_1 = 1.00$, $\omega_{2a} = 0.02$, $\omega_{2b} = 1.00$ FG: $\omega_0 = 0.02$, $\omega_1 = 1.00$, $\omega_{2a} = 1.00$, $\omega_{2b} = 1.00$		

Supplementary references:

Dong D, Lei M, Liu Y, Zhang S (2013). Comparative inner ear transcriptome analysis between the Rickett's big-footed bats (*Myotis ricketti*) and the greater short-nosed fruit bats (*Cynopterus sphinx*). *BMC Genomics* **14** 916.

Flicek P, Ahmed I, Amode MR, Barrell D, Beal K, Brent S *et al* (2013). Ensembl 2013. *Nucleic Acids Res.* **41:** D48–D55.

Schmidt CM, Jarvis JUM, Bennett NC (2013). The long-lived queen: reproduction and longevity in female eusocial Damaraland mole-rats (*Fukomys damarensis*). *Afr. Zool.* **48:** 193–196.

Seim I, Fang X, Xiong Z, Lobanov AV, Huang Z, Ma S *et al* (2013). Genome analysis reveals insights into the evolution, physiology, and longevity of the Brandt's bat. *Nat. Commun.* **4:** 2212.

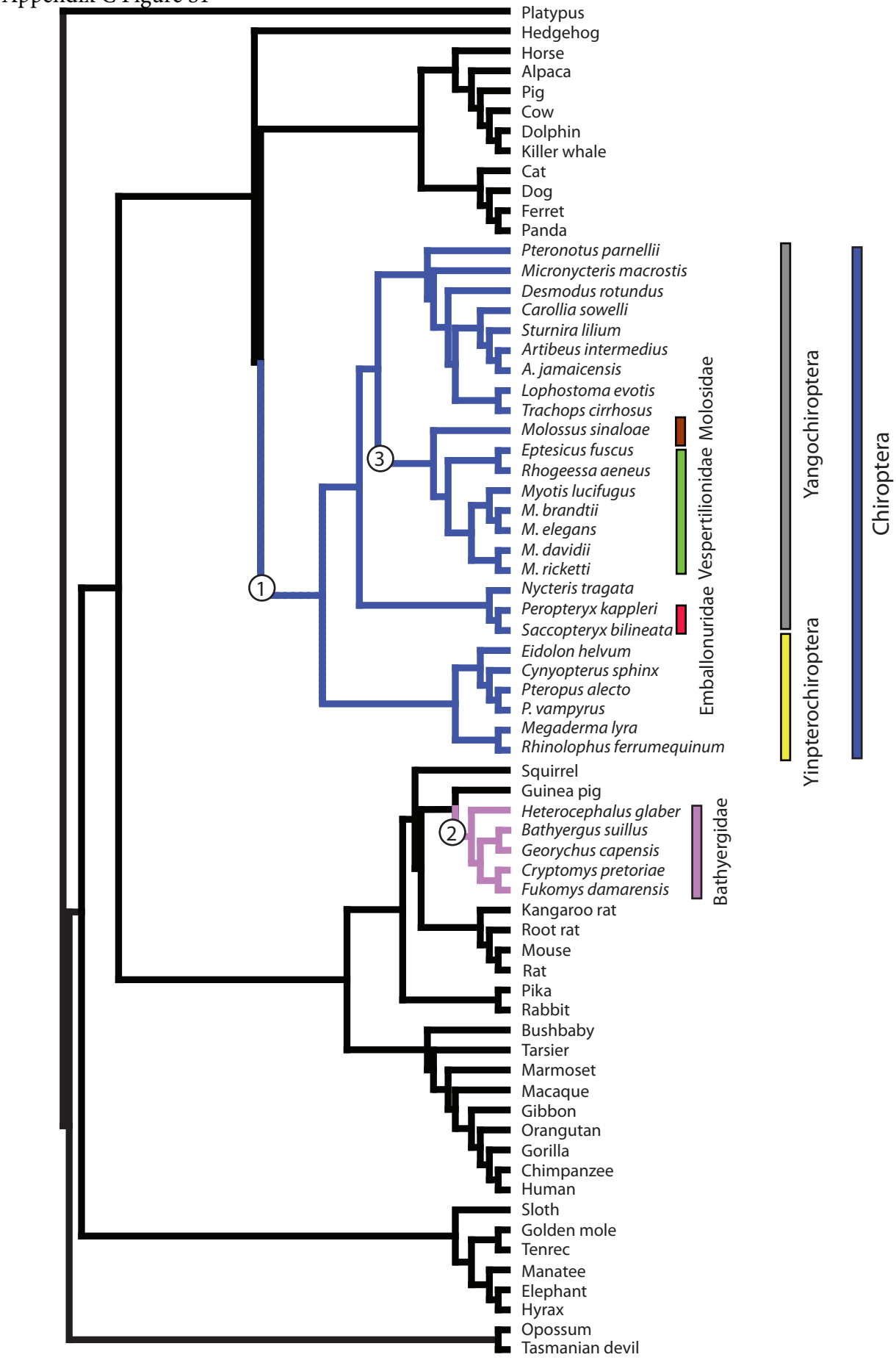
Shaw TI, Srivastava A, Chou W-C, Liu L, Hawkinson A, Glenn TC *et al* (2012). Transcriptome sequencing and annotation for the Jamaican fruit bat (*Artibeus jamaicensis*). *PLoS ONE* **7:** e48472.

Tacutu R, Craig T, Budovsky A, Wuttke D, Lehmann G, Taranukha D *et al* (2013). Human Ageing Genomic Resources: Integrated databases and tools for the biology and genetics of ageing. *Nucleic Acids Res.* **41:** D102–D1033.

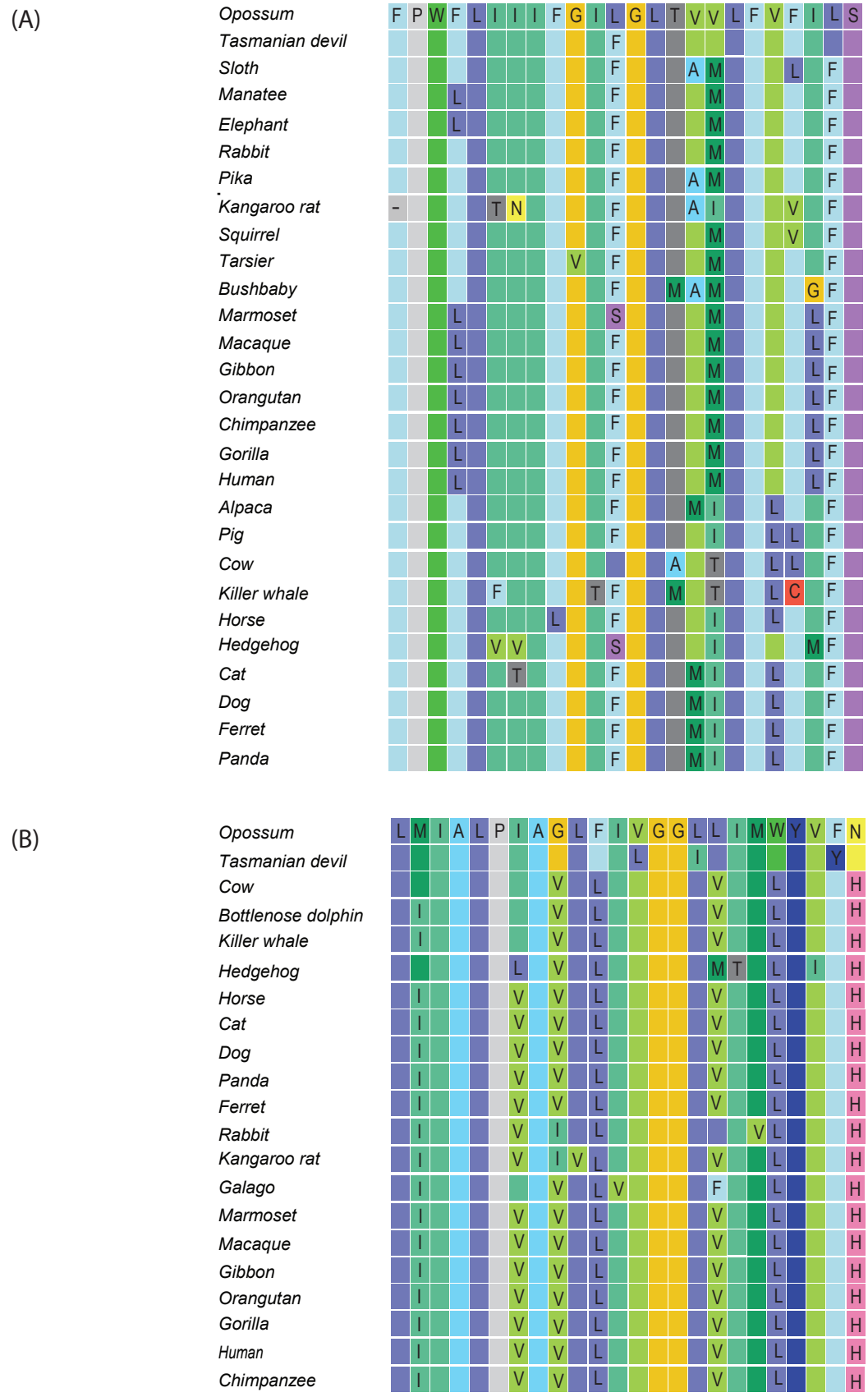
Tsagkogeorga G, Parker J, Stupka E, Cotton JA, Rossiter SJ (2013). Phylogenomic analyses elucidate the evolutionary relationships of bats. *Curr. Biol.* **23:** 2262–2267.

Zhang G, Cowled C, Shi Z, Huang Z, Bishop-Lilly KA, Xiaodong F *et al* (2013). Comparative analysis of bat genomes provides insight into the evolution of flight and immunity. *Science* **339:** 456–460.

Appendix C Figure S1

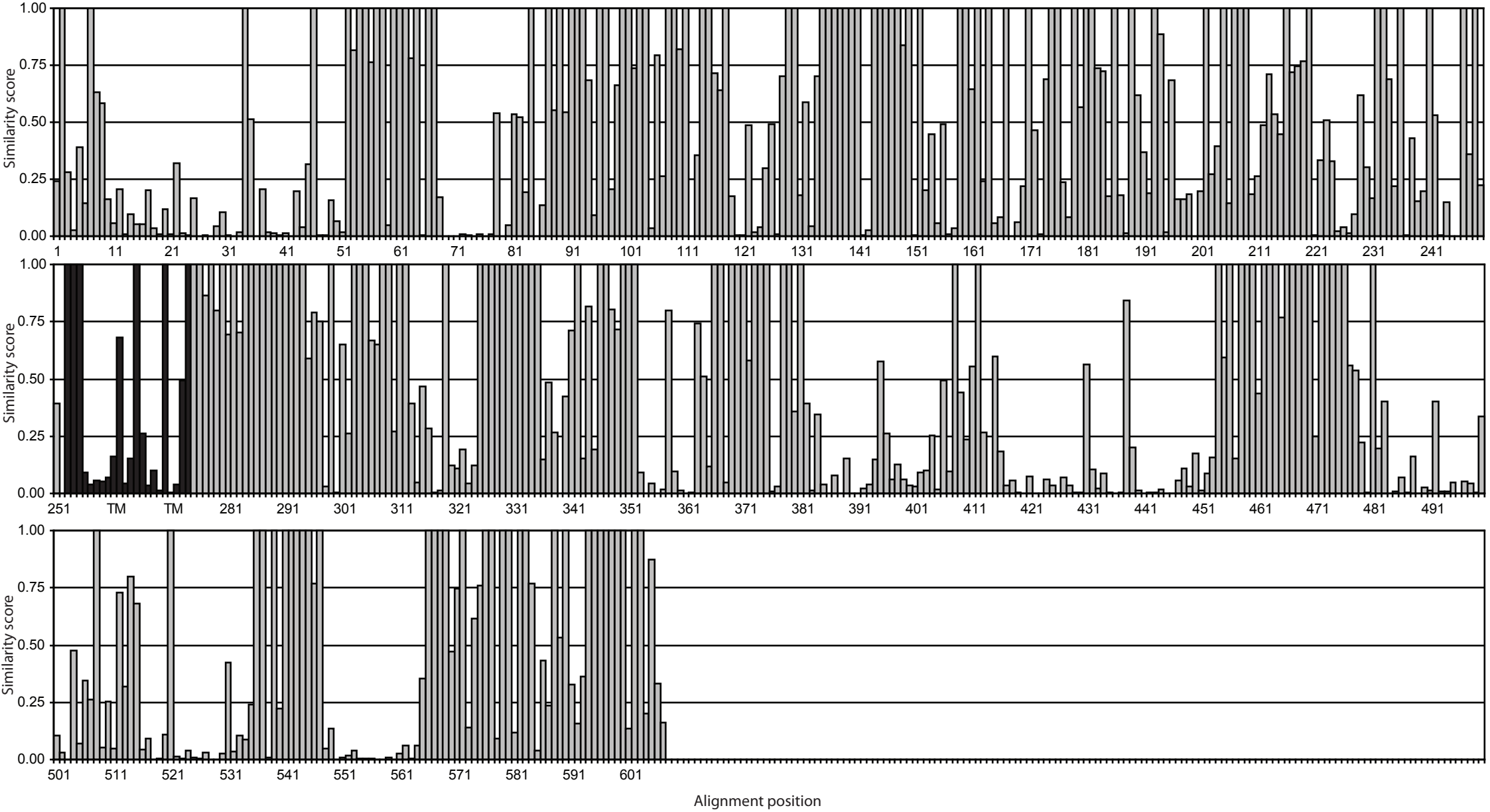


Appendix D Figure S2

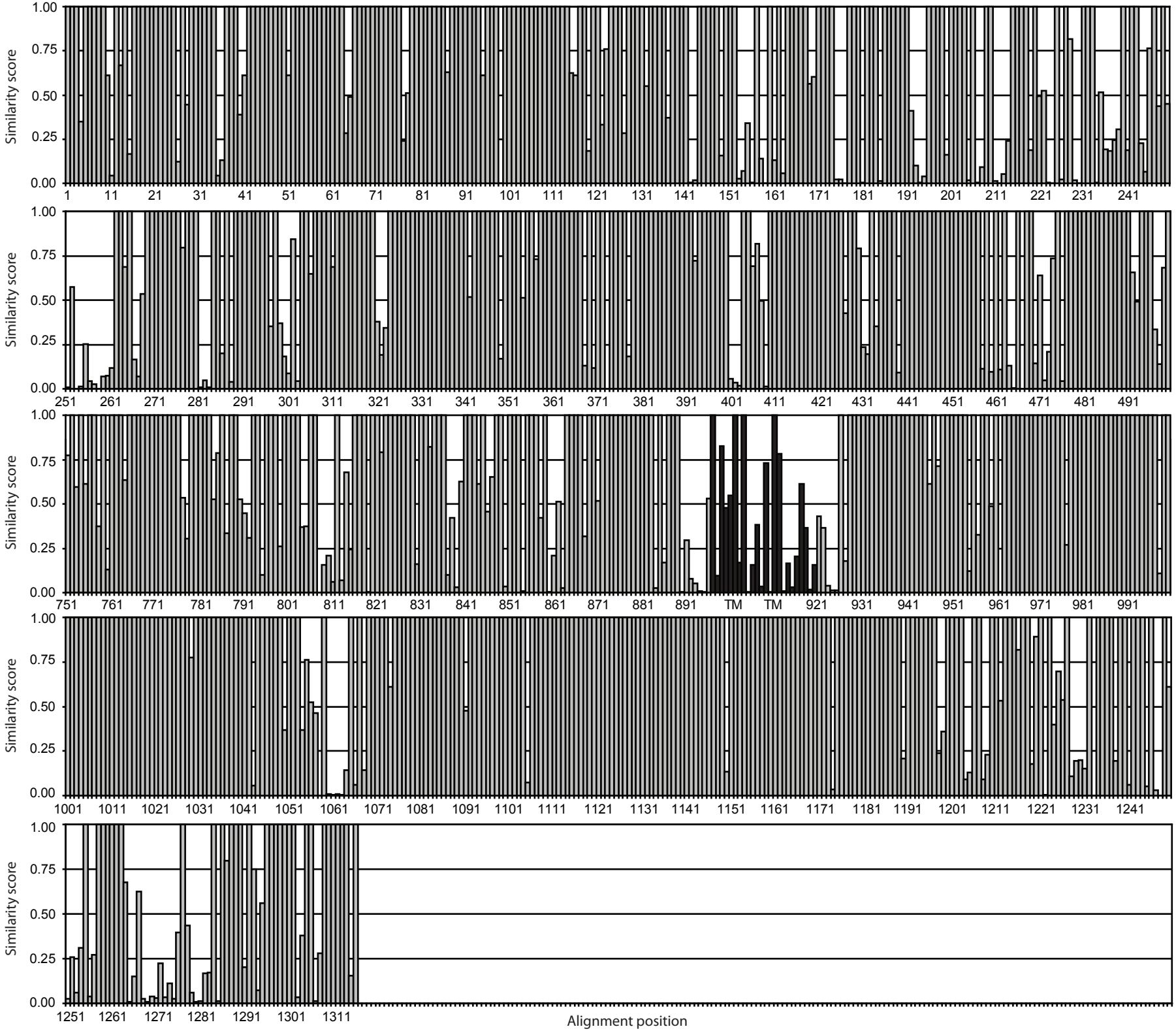


Appendix E Figure 3A

(A)



(B)



Appendix E
Figure S3B