

Rapid molecular evolution across amniotes of the IIS/TOR network

Suzanne E. McGaugh^{a,1}, Anne M. Bronikowski^{b,1}, Chih-Horng Kuo^{c,1}, Dawn M. Reding^{b,2}, Elizabeth A. Addis^{b,3}, Lex E. Flagel^{d,e}, Fredric J. Janzen^b, and Tonia S. Schwartz^{f,1}

^aDepartment of Ecology, Evolution, and Behavior, University of Minnesota, Saint Paul, MN 55108; ^bDepartment of Ecology, Evolution, and Organismal Biology, Iowa State University, Ames, IA 50011; ^cInstitute of Plant and Microbial Biology, Academia Sinica, Taipei 11529, Taiwan; ^dDepartment of Plant Biology, University of Minnesota, Saint Paul, MN 55108; ^eMonsanto Company, Chesterfield, MO 63017; and ^fOffice of Energetics, School of Public Health, University of Alabama, Birmingham, AL 35294

Edited by David M. Hillis, The University of Texas at Austin, Austin, TX, and approved April 16, 2015 (received for review October 20, 2014)

The insulin/insulin-like signaling and target of rapamycin (IIS/TOR) network regulates lifespan and reproduction, as well as metabolic diseases, cancer, and aging. Despite its vital role in health, comparative analyses of IIS/TOR have been limited to invertebrates and mammals. We conducted an extensive evolutionary analysis of the IIS/TOR network across 66 amniotes with 18 newly generated transcriptomes from nonavian reptiles and additional available genomes/transcriptomes. We uncovered rapid and extensive molecular evolution between reptiles (including birds) and mammals: (i) the IIS/TOR network, including the critical nodes insulin receptor substrate (IRS) and phosphatidylinositol 3-kinase (PI3K), exhibit divergent evolutionary rates between reptiles and mammals; (ii) compared with a proxy for the rest of the genome, genes of the IIS/TOR extracellular network exhibit exceptionally fast evolutionary rates; and (iii) signatures of positive selection and coevolution of the extracellular network suggest reptile- and mammal-specific interactions between members of the network. In reptiles, positively selected sites cluster on the binding surfaces of insulin-like growth factor 1 (IGF1), IGF1 receptor (IGF1R), and insulin receptor (INSR); whereas in mammals, positively selected sites clustered on the IGF2 binding surface, suggesting that these hormone-receptor binding affinities are targets of positive selection. Further, contrary to reports that IGF2R binds IGF2 only in marsupial and placental mammals, we found positively selected sites clustered on the hormone binding surface of reptile IGF2R that suggest that IGF2R binds to IGF hormones in diverse taxa and may have evolved in reptiles. These data suggest that key IIS/TOR paralogs have sub- or neofunctionalized between mammals and reptiles and that this network may underlie fundamental life history and physiological differences between these amniote sister clades.

insulin signaling | insulin growth factor | molecular evolution | rapamycin

The last 20 y has provided overwhelming support that the insulin- and insulin-like signaling/target of rapamycin (IIS/TOR) molecular network responds to stress and nutrients and underlies a wide range of physiological functions (1); cancer, metabolic syndrome, and diabetes (2); and the timing of life events (e.g., growth, maturation, reproduction, and aging) (3). The vertebrate IIS/TOR network consists of peptide hormones, binding proteins that regulate hormone bioavailability, and cell membrane receptors (hereafter, extracellular proteins of the IIS/TOR network) that induce an intracellular signaling cascade (hereafter, intracellular proteins of the IIS/TOR network) to stimulate cell proliferation, survival, and metabolism (Fig. S1). The core intracellular signal transduction genes in this network are largely conserved across deep phylogenetic time (4, 5). In contrast, genes encoding the IIS/TOR extracellular network have diverged in the vertebrate lineage (6–8) and may have variable roles among taxa (9, 10). Despite its central role in health, comparative analyses of IIS/TOR have been limited to model invertebrates and mammals. Here we conduct evolutionary analyses of IIS/TOR across amniotes: i.e., mammals and their reptile sister clade, which includes birds (Fig. S2). Many life

history and metabolic traits differ substantially between mammals and reptiles (11, 12), and the IIS/TOR network influences these traits (13–15).

Within vertebrates, many IIS/TOR extracellular genes have evolved through gene duplication and thus are paralogs. Duplications of an insulin-like progenitor gene resulted in genes encoding insulin (INS) and insulin-like growth factors 1 and 2 (IGF1 and IGF2) (6). These paralogous hormones bind the similarly paralogous receptors, insulin receptor (INSR) and insulin-like growth factor 1 receptor (IGF1R), and this binding initiates the intracellular signaling cascade through insulin receptor substrate (IRS) and through the phosphatidylinositol 3-kinase (PI3K) and serine/threonine protein kinase intracellular nodes (Fig. S1) (5, 7). Repeated duplication of the gene encoding the ancestral IGF-binding proteins (IGFBP) resulted in six binding proteins that regulate bioavailability of the hormones (8). Generally, these receptors, hormones, and binding proteins maintain the ability for cross-talk, but binding affinities differ (16). An additional receptor, IGF2R, is a co-opted mannose-6 phosphate receptor that regulates IGF2 bioavailability for activating IIS/TOR

Significance

Comparative analyses of central molecular networks uncover variation that can be targeted by biomedical research to develop insights and interventions into disease. The insulin/insulin-like signaling and target of rapamycin (IIS/TOR) molecular network regulates metabolism, growth, and aging. With the development of new molecular resources for reptiles, we show that genes in IIS/TOR are rapidly evolving within amniotes (mammals and reptiles, including birds). Additionally, we find evidence of natural selection that diversified the hormone-receptor binding relationships that initiate IIS/TOR signaling. Our results uncover substantial variation in the IIS/TOR network within and among amniotes and provide a critical step to unlocking information on vertebrate patterns of genetic regulation of metabolism, modes of reproduction, and rates of aging.

Author contributions: S.E.M., A.M.B., D.M.R., E.A.A., F.J.J., and T.S.S. designed and performed research; S.E.M., A.M.B., C.-H.K., L.E.F., and T.S.S. analyzed data; and S.E.M., A.M.B., C.-H.K., D.M.R., E.A.A., L.E.F., F.J.J., and T.S.S. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The sequences reported in this paper have been deposited in the National Center for Biotechnology Information (NCBI) Sequence Read Archive, www.ncbi.nlm.nih.gov/sra (accession nos. SRA062458 and SRP017466). Transcriptome assemblies, annotation summaries, and alignments for protein coevolution analyses are available through Dryad ([10.5061/dryad.vn872](https://doi.org/10.5061/dryad.vn872)).

¹To whom correspondence may be addressed. Email: smcgaugh@umn.edu (assemblies, alignments, and molecular evolution analyses), abroniko@iastate.edu (the study itself, transcriptomes, and accessing data), chk@gate.sinica.edu.tw (OrthoMCL analyses), or tschwartz@uab.edu (protein predictions and coevolutionary and network interactions).

²Present address: Department of Biology, Luther College, Decorah, IA 52101.

³Present Address: Department of Biology, Gonzaga University, Spokane, WA 99258.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1419659112/-DCSupplemental.

by bringing IGF2 into the cell for degradation (17). It is widely hypothesized that IGF2-IGF2R binding is unique to therian mammals (marsupials and placentals) due to sexual conflict in regulating paternal IGF2 during placental and embryo development (10, 18–21).

Previous evolutionary analyses of IIS/TOR in vertebrate and invertebrate lineages suggest that extracellular genes (Table S1 and Fig. S1) often experienced positive selection, whereas intracellular genes often experienced purifying selection (22–28), especially farther downstream in the intracellular network. These previous findings have supported the prediction that upstream extracellular factors (e.g., the initial components that interact with environmental stimuli in signal transduction pathways) may have larger impacts on signaling through the network than downstream components. However, comparative studies of (co)evolution among extracellular components of the network have not been possible with studies from invertebrates due to the near absence of these paralogs. In addition, mammalian studies of this network have not included the other half of the amniote group (avian and nonavian reptiles), except chickens. Thus, a general understanding of the evolution of this network and the co-evolutionary relationships among the proteins of this network has not been possible.

We analyze coding sequence data from 32 species of mammal—an order of magnitude higher than previous comparative studies of mammals—and 34 species of reptile (10 species of birds and 24 nonavian reptiles; Fig. S2). Analyses of this improved sampling revealed that members of the IIS/TOR network, particularly extracellular and critical intracellular genes, exhibit exceptionally fast evolutionary rates between mammals and reptiles relative to the rest of the genome. Additionally, strong positive selection occurs at amino acid sites important for hormone-receptor protein interactions, and this selection likely shapes binding affinities in reptile- and mammal-specific ways.

Results

We identified an average of 31,060 unique ORFs per species (range, 15,893–102,156) and used OrthoMCL (29) and quality control methods to produce alignments of putative orthologs across 66 species (Table S2) (see data deposition footnote and SI Text).

We focused on 61 genes from the IIS/TOR network that were identified through KEGG pathways (Kyoto encyclopedia of genes and genomes, ref. 30) and/or previous publications (Fig. S1 and Table S1) (8, 24). Alignments of these focal genes contained 19–66 species (median = 62, mode = 66; mean = 58.4; Table S1). To provide a proxy for evolution of the noninsulin signaling genes in the genome, we used 1,417 putative orthologs that contained all 66 species and referred to these as control genes. We also analyzed (i) 48 of the 61 focal genes that had greater representation of species within the alignments (56–66 species, median = 63.5, mode = 66, mean = 62.6), and (ii) a control and IIS/TOR focal gene set that contained phylogenetically matched species (43 control genes and 43 focal genes with identical species). Both of these analyses are presented in the SI Materials and Methods and are consistent with the analysis reported below that extracellular genes of the network are highly divergent outliers.

IIS/TOR Network Contains Fast-Evolving Outliers. Twenty-six genes (of 61 analyzed) from the IIS/TOR network exhibited divergent evolutionary rates between reptiles and mammals (i.e., significant likelihood ratio test between the null and alternative models) using the Clade model in PAML (31) and a P value estimated following refs. 32 and 33 and corrected for multiple tests by sequential Bonferroni (Table S3, CMC reptiles). For 20 of these 26 divergent genes, the ω [nonsynonymous substitutions per nonsynonymous site (K_a)/synonymous substitutions per synonymous site (K_s)] for reptiles was significantly greater than the ω for the rest of the tree (e.g., mammals, $\chi^2 = 7.54$, $P = 0.006$). We obtained similar results for a paired Wilcoxon test ($P = 0.056$), and this

difference in divergence between clades was also seen among control genes (SI Results).

Extracellular genes of the IIS/TOR network exhibited greater divergence between mammals and reptiles than 1,417 control genes and intracellular genes when measured by the median of all pairwise mammal-reptile K_a/K_s measures. Extracellular genes had equivalent K_s compared with control genes (Wilcoxon rank sum test, $W = 5476$, $P = 0.2154$), but had notably greater median ω ($W = 2998$, $P = 0.002$) and K_a ($W = 2333.5$, $P < 0.001$; Fig. 1). Compared with intracellular genes, extracellular genes also had significantly higher ω and K_a (ω : $W = 129.5$, $P = 0.02$; K_a : $W = 88$, $P = 0.001$), but K_s did not differ ($W = 209$, $P = 0.375$).

Collectively, the intracellular IIS/TOR genes did not have elevated median K_a , K_s , or ω compared with control genes ($P > 0.467$ in all cases; Fig. 1). In all cases, the median for intracellular and control genes was identical to the hundredths place. The median K_s for extracellular genes was 1.59, and the median K_s for intracellular and control genes was 1.60. The median K_a for extracellular genes was 0.17, and the median K_a for intracellular and control genes was 0.09. The median ω for extracellular genes was 0.11 and for intracellular and control genes ω was 0.06.

When comparing the distribution of ω values for each group of IIS/TOR genes to the distribution of the ω values for the control genes, the extracellular genes were 8.4 times more likely than control genes to reside in the highest 5% of ω values (OR, 8.37; 95% CI, 2.12, 33.08). In comparison, the intracellular genes were not significantly more likely than controls to be in top 5% (OR, 2.21; 95% CI, 0.82, 5.51). These odds ratios imply that the extracellular group contains the fastest evolving components of the IIS/TOR network.

Evidence for Positive Selection. To understand how positive selection may have shaped the genes within the IIS/TOR network, we used the branch-site test in PAML, which tests for molecular evolution at the nucleotide level with functional impacts at the protein level. In the first analysis, the branch leading to reptiles was tested for evidence of positive selection (i.e., was placed in the “foreground,” which functions to test the predicted ancestral reptile against all mammals). Eighteen genes showed significant signatures of positive selection along this branch leading to reptiles, six of which remained significant after sequential Bonferroni correction: IGF2R and five intracellular genes [protein kinase C gamma (*PRKCG*), inositol polyphosphate phosphatase-like 1, phosphatidylinositol 3-kinase regulatory subunit (*PIK3R*), *IRS1*, and *IRS2*] (Table S3). In the second analysis, with the branch leading to mammals designated as the foreground branch, testing this predicted ancestral mammalian branch against all reptiles, 23 genes showed significant signatures of positive selection, 9 of which remained significant after sequential Bonferroni correction. This group included most of the genes that were significant along the

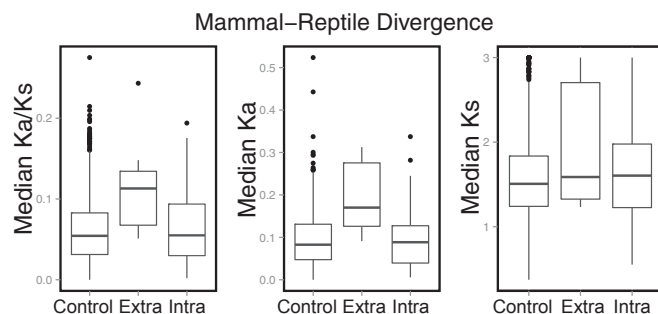


Fig. 1. Medians of pairwise measures between all reptiles and mammals per gene for K_a , K_s , and K_a/K_s calculated in PAML. Control = 1,417 genes not in the IIS/TOR network, with 66 taxa represented in the alignments. Intracellular = 51 genes. Extracellular = 10 genes (hormones, receptors, IGF binding proteins). Extracellular genes exhibit significantly greater median K_a/K_s and median K_a than control or intracellular genes.

reptile branch (*IGF2R*, *IRS1*, *IRS2*, *PRKCG*, and *PIK3R*), as well as others (Table S3).

We also performed the branch-site test with specific lineages within reptiles, because previous research indicated that genes of the IIS/TOR network may be under strong positive selection in Squamata (lizards and snakes) (34) (Table S3). Overall, the branch leading to Squamata had more genes under positive selection (number of genes = 7 of 61) than on the branches leading to crocodylians ($n = 6$), birds ($n = 1$), and turtles ($n = 5$) (when separate tests were conducted for each), i.e., minor differences (Table S3). In additional tests using the clade model, we found that snakes had larger ω relative to the rest of the tree (paired Wilcoxon-sign rank test, $V = 24$, $P = 0.04$) across the 15 IIS/TOR network genes that were significant after multiple test correction.

Positive Selection in the IIS/TOR Hormones and Receptors Show Clade-Specific Patterns. Positively selected sites [i.e., those indicated by PAML branch-site models with Bayes Empirical Bayes (BEB) score of 0.9 or greater] were analyzed in the context of the protein structure and predicted protein-protein interactions between insulin/IGF hormones and their receptors. We found reptile- and mammal-specific patterns of positive selection in the hormone and receptor domains that are important for binding affinity. First, the mature INS hormone (containing protein domains A and B) was conserved in reptiles and mammals, whereas the C-peptide that is cleaved from the mature insulin protein (35) contained four positively selected sites in mammals. Second, 5 of the 12 amino acids of the C-domain in IGF1 in reptiles, but none in mammals, were positively selected. Third, for IGF2, 3 of the 16 amino acids of the C-domain in mammals, but none in reptiles, were positively selected (Fig. 2A).

IGF hormones bind to the receptors IGF1R and INSR by interacting with specific domains on each receptor (L1, CR, and L2) (36, 37). Variation in the C-domain of IGF1 and IGF2 can regulate binding specificity to IGF1R (38) and to INSR (39) through the interactions of the C-domain of the hormones with the CR-domain in the binding pocket of IGF1R and INSR (36) (Fig. 2B). Specifically, previous mutagenesis studies revealed that altering one of the positively selected sites (IGF1 C-domain R37, human numbering used throughout) disrupts IGF1-IGF1R binding (16, 40). In reptiles, positively selected sites were clustered on the hormone-binding surface of the IGF1R CR domain and in the binding pocket of INSR. They include IGF1R site F251,

which affects IGF1-IGF1R binding in humans through its interaction with the IGF1 C-domain (36) (Table S4).

In therian mammals, IGF2R binds IGF2 with relatively high affinity, but studies of this interaction in reptiles (mainly chickens) have yielded conflicting results (18, 21, 41, 42). We found that IGF2R has been shaped by putatively strong positive selection within reptiles and positively selected sites clustered on the IGF2R protein surface in domain 11, which is intimately involved in binding IGF1 and IGF2. Several of the positively selected sites on the protein surface of IGF2R in reptiles are essential for binding IGF2 based on mutagenesis studies and the crystal structure of the IGF2R-IGF2 complex (e.g., Y1542) (43–45) (Fig. 2C and Table S4). Although some variants in IGF2R would predict decreased binding to IGF2, such as in chicken, many variants in snakes and lizards predict increased binding to IGF2 and/or IGF1 because they exhibit similar biochemical properties as the human amino acids (e.g., Y1542F in snakes and Y1542L/M in lizards; Table S4). Utilizing Coevolutionary Analysis Using Protein Sequences (CAPS) (46), we identified that amino acid site P4 of IGF2 is coevolving with the positively selected site on the binding surface of IGF2R (site R1623, $\rho = 0.4$, $P < 0.01$) in reptiles (Fig. 2C). Among reptiles, MatrixMatchMaker version II (MMMvII) (47) identified sunbeam and viper boa snakes as having the tightest coevolutionary signal between IGF2 and IGF2R ($\rho = 1$), and identified brown anole, green anole, and gecko lizards as having the tightest coevolutionary signal between IGF1 and IGF2R ($\rho = 0.33$). Thus, among reptiles, IGF2R binding of IGFs is most likely to be found in the Squamates.

Binding Proteins Exhibit Putative Truncations of Important Functional Domains. IGF binding proteins bind to IGF1 and IGF2 in the bloodstream to regulate their bioavailability (48, 49). These IGFbps are characterized by N- and C-terminal domains that cooperate to bind IGFs; protease cleavage separating these domains decreases affinity to IGFs. In both reptiles and mammals (except primates), many of our assembled IGFbp transcripts were either completely missing the N-terminal domain or it was truncated (Table S5 and Fig. S3). As assembled, these transcripts would produce truncated proteins with diminished binding affinity to IGF1 and IGF2. We summarize putative losses and truncations in Table S4 to serve as a hypothesis-generating resource for future validation work. Most evident is IGFbp6, which was neither found in any archosaurs (birds and crocodylians) nor in platypus (8). The 5' end of IGFbp6 was truncated in nearly all other reptiles including genome-derived ENSEMBL sequences of the *Anolis* lizard and the *Pelodiscus*

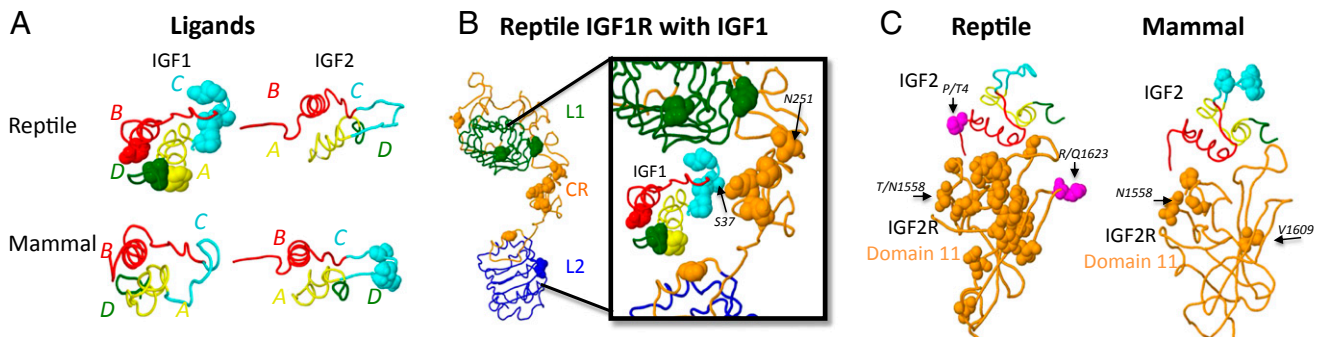


Fig. 2. Protein structures for reptile and mammal IGF hormones and receptors. Reptile protein structures predicted from snake sequence homology modeled onto human protein structures from the Protein Data Bank. Enlarged positions indicate the amino acid sites predicted to be under positive selection (Table S3). (A) Reptile and mammal IGF hormones with their protein domains color coded. Positively selected sites cluster on the C-domain of reptile IGF1 but are not present in the C-domain of the reptilian IGF2. In contrast, positively selected sites cluster on the C-domain of mammal IGF2. (B) The α chain of reptile IGF1R homodimer with hormone binding domains L1, CR, and L2 labeled. The square is an enlargement with IGF1 orientated in the IGF1R binding pocket to demonstrate the clustering of positively selected sites on the interacting IGF1-IGF1R binding surfaces (36). Labeled sites (IGF1 S37 and N251; human numbering) are known to affect IGF hormone and receptor binding (Table S4). (C) Domain 11 of reptile and mammal IGF2R with IGF2 oriented toward the binding pocket to demonstrate the clustering of positively selected sites on the reptile IGF2R binding surfaces (43, 44). The magenta sites on reptile IGF2 and IGF2R were identified as coevolving amino acids using CAPS (46). Labeled sites IGF2R (1558 and 1609; human numbering) are predicted to regulate IGF2-IGF2R binding (Table S4). Like mammals, some lizards have IGF2R N1558.

turtle. Furthermore, in examining the three N-terminal amino acids that are conserved across all binding proteins in humans (49), only one of these amino acids was conserved in only two snake species in IGFBP6, although all three sites were conserved across the reptile IGFBP2-5. For those amino acids important for binding IGFs and specific to IGFBP6 (49), only 7 of 12 are conserved in reptiles. Two of these seven conserved amino acids have additional functions beyond IGF binding, which requires conservation (Fig. S3). These multiple lines of evidence suggest that IGBP6 does not function as an IGF binding protein across the reptile clade.

Discussion

We conducted extensive evolutionary analyses of the IIS/TOR network in amniotes (i.e., mammals and reptiles, including birds) and uncovered fundamental differences between reptiles and mammals in the evolution of this centrally important network. Our analyses revealed that members of the IIS/TOR network have exceptionally fast evolutionary rates between reptiles and mammals compared with a proxy for the rest of the genome. More specifically, the extracellular network is a target of positive selection, and the location of the selected sites suggests changes in the hormone-receptor binding relationships in reptile- and mammal-specific patterns.

Members of IIS/TOR Network Are Outliers in Evolutionary Rate.

Members of the IIS/TOR network, especially the extracellular hormones, receptors, and binding proteins, exhibit remarkably high reptile-mammal divergence compared with control genes. Our results complement those of ref. 24, who found that the IIS/TOR network across human populations is enriched for genes evolving under positive selection relative to a sample representing the genomic background. Across the amniote scale that we examined, many evolutionary innovations have arisen (e.g., feathers/hair, leglessness, endothermy), and each was likely accompanied by substantial molecular evolution. However, within the 1,478 total genes that we analyzed (61 IIS/TOR network genes plus 1,417 non-IIS/TOR genes), the evolution of the IIS/TOR extracellular network is a prominent outlier in reptile-mammal divergence. Our results provide additional evidence that the phenotypes governed by this pathway, including metabolism and life histories, are key differences between reptiles and mammals.

Our data show that multiple IIS/TOR genes are under positive selection in one or more lineages of amniotes. Importantly, these include genes that encode proteins in critical nodes of the IIS/TOR network that mediate the intracellular signal (e.g., IRS and PI3K) (5) and extracellular nodes that regulate the initiation of the cascade (IGF1R, INSR, IGFBP4, IGFBP5, and IGF2R). Although these genes are implicated in aging and disease phenotypes (3, 50), here we find they are also under positive selection among amniote species. Because vertebrate IIS/TOR connects with many other networks, we cannot directly compare our results to studies in the more simplified invertebrate network (23, 26, 27). However, our findings of elevated ω s agree with those of ref. 28 and supply further support that extracellular components are among the fastest evolving genes in the IIS/TOR network (22), as is likely true in other networks. Overall, our data are in agreement with reports that receptors and other extracellular components of signal transduction pathways appear to be under less purifying selection than intracellular components (51–53). Indeed, our data indicate that one potential driver of differences in evolutionary rates among genes in the network may be the number of interactions that a gene has with other genes or proteins (i.e., connectivity; *SI Results*) similar to what has been seen in other systems (54–56, but see refs. 27, 57, and 58).

Evolving Interactions in the Extracellular Network. Our data strongly support the conclusion that many of the IIS/TOR extracellular proteins have undergone positive selection. Detailed evaluation of the protein structure of the hormones, receptors, and binding

proteins of IIS/TOR suggests that these binding relationships are targets of clade-specific selection between mammals and reptiles. In reptiles, structural evaluations indicated that residues on the interacting binding surfaces of IGF1 C-domain and the IGF1R CR-domain are under positive selection. Specifically, positively selected amino acid sites identified by our models have previously been shown to modulate binding when altered in humans (16). More broadly, mutations predicted to affect this binding relationship are associated with longevity in humans (59) and model organisms (3). In contrast, positive selection on the IGF2 C-domain in mammals suggests that IGF2 binding affinities with both INSR and IGF1R may be targets of positive selection in mammals. The positive selection putatively affecting hormone-receptor binding relationships across amniote species equates to selection at the cell surface start of the IIS/TOR signaling cascade. Functional studies are necessary to further our understanding of the regulatory effects of these changes and the stability of the physiological roles of extracellular IIS/TOR proteins across amniotes.

Juvenile and adult *IGF2* gene expression is observed in humans (60, 61) and fish (62), but not in adult mice and rats (the typical vertebrate models for studying the IIS/TOR network) (63). We found *IGF1* and *IGF2* gene expression in each of our reptile transcriptomes, regardless of whether the source liver was from a juvenile or an adult (Table S2). This observation underscores the importance of broad taxonomic sampling for understanding the function and evolution of pathways important to human health—for which rodent models may not always be the most appropriate. Together, our molecular evolution and expression data suggest that the IGF2 protein may have a more stable role in IIS/TOR signaling across reptiles, in contrast to the more variable and specialized roles of IGF2 across mammals (64).

IGF2R-IGF2 binding is believed to have evolved in therian mammals for maternal regulation of paternally imprinted IGF2 (10). This hypothesis of mammalian-specific function was bolstered by early studies showing that IGF2R does not effectively bind IGF2 in chicken (18, 21, 41), *Xenopus* (18), or monotremes (10), and IGF2R has lower affinity for IGF2 in marsupials compared with placental mammals (19, 20). However, more sensitive assays have indicated that IGF2R-IGF2 binding occurs in chicken, trout, and garden lizards (42, 65, 66), which counters the claim that measurable IGF2-IGF2R binding is confined to mammals. Our data provide support for the hypothesis that positive selection drove the high-affinity binding between IGF2R and IGF2 in placental mammals relative to monotremes and marsupials. Our data also call into question the assumption that IGF2R does not bind IGF hormones in reptiles. IGF2R in chicken contains a substitution thought to inhibit IGF2 binding ability (isoleucine to leucine at 1572, I1572L) (67, 68). However, our work shows that this amino acid is a conserved isoleucine in many reptile species, even within other birds (66). Additionally, many of the sites that are important for binding of IGF2 to IGF2R in mammals are conserved across most reptiles in our study. Because chickens have typically been used as the sole representative of the reptile clade, we suggest that this narrow sampling promoted the premature conclusion that IGF2 binds IGF2R only in mammals. Further, in reptiles, we found a signal of co-evolution between IGF2 and IGF2R in our CAPs and MMMvII analyses. Additionally, we found three sites under positive selection on the surface of the IGF1 that would likely promote binding with IGF2R (34, 67). Thus, by extending the comparative genomic landscape, we suggest that IGF-IGF2R binding may not be unique to therian mammals but also may occur in some reptile species.

IGF binding proteins regulate the ability of hormones to activate receptors through steric hindrance, thereby limiting the bioavailability of IGF1 and IGF2 to initiate the IIS/TOR signaling cascade (49). Intriguingly, we found that many reptile species appear to have truncated or missing N-terminal domains across the IGFbps that would decrease IGF binding affinity. Confirming results from ref. 8, IGFBP6 was not recovered from opossum, platypus, or any bird or crocodile. When identified in our other reptile transcriptomes and ENSEMBL-derived

genomic data, the N terminus of the protein was truncated. These data suggest that across reptiles, IGF1-IGF1R binding, IGF2-IGF2R binding, IGF2-IGFBP6 binding in mammals functions to regulate IGF2 levels during embryo development in placental mammals (69). The putative loss of this regulatory mechanism in both reptiles and some nonplacental mammals is particularly interesting given that placentation has evolved not only in mammals but also in various snake and lizard species (70). Thus, our data suggest that in many reptiles (i) IGF1-IGF1R binding has been lost, (ii) IGF2R binds IGF hormones, and (iii) novel positive selection characterizes IGF1-IGF1R binding. Therefore, future functional assays should address the role of IIS/TOR extracellular signaling in the evolution of viviparity and placentation in Squamates, relative to that in placental mammals (10) and placental fish (9).

Comparative Genomics Approach. The insights our study provide into the evolution of the IIS/TOR network were previously unattainable without adequate molecular resources in reptiles. Our work adds to the recent discoveries of rapid evolution of genes involved in development and metabolism in the branch leading to modern snakes (71) and of regulatory innovation in IGF1-IGF1R and IGF2-IGF2R in the branch leading to modern birds (72). Although de novo transcriptome assemblies may not fully reveal all biologically important signals in data (such as species-specific isoforms and very recent paralogs) (73), when combined with available genomes, ours revealed insights into the evolution of the IIS/TOR network. Although the core of the IIS/TOR network is conserved in animals (4, 5), we found high divergence and selection on genes in this network between mammals and their sister clade reptiles (including birds). The extracellular genes of this network had exceptionally fast divergence between reptiles and mammals relative to genomic background, and many genes have been shaped by positive selection. Hormones, receptors, and binding proteins that are essential for producing a physiological response to environmental stimuli have undergone taxon-specific patterns of positive selection. Our results suggest that key paralogs have subfunctionalized or neofunctionalized between reptiles and mammals and that this network may underlie fundamental life history and physiological differences between these clades.

In a larger context, the strength of comparative biology in understanding human health and disease lies in its power to distinguish conserved vs. flexible mechanisms of normal and disease states and thereby suggest worthy targets of biomedical research into future interventions (74, 75). For example, lifespan extension is observed with mutant IGF1, IGF1R, and IRS across diverse model species (3, 76–79)—where a shared effect on IIS/TOR signaling is to either decrease rates of signaling by disrupting protein-protein interactions or to decrease normal levels of hormone or receptor. In addition, the IIS/TOR network has

been associated with longevity in humans (59, 78, 80–82). Likewise, our comparative genomic analyses show that many IIS/TOR genes are variable across amniotes and that the binding affinities of IGF1, IGF1R and INSR, and thereby the initiation of IIS/TOR signaling, is likely impacted. Future comparative analyses of the IIS/TOR network across amniotes and within reptiles may provide unique insights into the regulation of body size, reproductive investment (e.g., placentation), and rates of aging (83).

Materials and Methods

We used transcriptomic and genomic data across amniotes to evaluate molecular evolution of the IIS/TOR pathway between reptiles and mammals. All animal protocols were approved by the Iowa State University Institutional Animal Care and Use Committee (log 3-2-5125J). De novo liver transcriptome assembly was performed in Trinity (Table S2), and some gene sets were obtained through past studies (Table S2). The longest ORF from each assembled transcript was used for defining homologs through OrthoMCL (29). Sequences within each putative ortholog were further clustered so that a single transcript represented each ortholog from each species. Transcripts were translated, and amino acid sequences were aligned with MSAPROBS (84). Alignments were back-translated to the original nucleic acids with RevTrans (85) and trimmed of poorly aligned regions using Gblocks (86).

These cleaned nucleotide alignments were analyzed for molecular evolutionary parameters and models of sequence evolution in PAML (31). Positively selected sites for extracellular genes were predicted for reptiles and mammals using the branch-site model in PAML. Sites with signatures of positive selection were evaluated for putative functional significance on human protein structures from the Protein Data Bank (PDB) or predicted reptile structures from homology modeling of snake sequences onto human structures. Hormone and IGF2R amino acid alignments were used for co-evolution analyses with CAPS (46) [significance of permutations ($P < 0.01$) detailed in *SI Materials and Methods*] and MMMVII (47) (tolerance level: 0.2).

We describe each of these steps in detail in *SI Materials and Methods*.

ACKNOWLEDGMENTS. We thank D. Warner, R. Telemeco, A. Cordero, N. Ford, K. Wray, T. Owerkowicz, and C. Watson for contributing specimens; E. Tillier for advice on MMMVII; and A. Brown, J. P. de Magalhães, and an anonymous reviewer for useful comments. We thank the Baylor College of Medicine and The Genome Institute at Washington University in St. Louis for use of the unpublished genomic sequence. We thank the Broad Institute Genomics Platform, Vertebrate Genome Biology group, J. Alföldi, and K. Lindblad-Toh for making the *Mustela putorius* and *Microtus ochrogaster* data available. We are grateful for resources from the University of Minnesota Supercomputing Institute, University of Alabama at Birmingham Office of Energetics, and the Iowa State University High Performance Computing facility. This research was supported by National Science Foundation (NSF) Grants IOS-0922528 and IOS-1253896 (to A.M.B.) and DEB-DDIG-1011350 (to A.M.B. and T.S.S.) and grants from the Iowa State University Center for Integrated Animal Genomics (to A.M.B. and F.J.J.). We acknowledge additional support from the NSF (Graduate Research Fellowship to S.E.M.), the James S. McDonnell Foundation (postdoctoral fellowship to T.S.S.), the Howard Hughes Medical Institute (postdoctoral support to E.A.A.), and Academia Sinica (C.H.K.).

- Wullschlegel S, Loewith R, Hall MN (2006) TOR signaling in growth and metabolism. *Cell* 124(3):471–484.
- Zoncu R, Efeyan A, Sabatini DM (2011) mTOR: From growth signal integration to cancer, diabetes and ageing. *Nat Rev Mol Cell Biol* 12(1):21–35.
- Kenyon CJ (2010) The genetics of ageing. *Nature* 464(7288):504–512.
- Oldham S (2011) Obesity and nutrient sensing TOR pathway in flies and vertebrates: Functional conservation of genetic mechanisms. *Trends Endocrinol Metab* 22(2):45–52.
- Taniguchi CM, Emanuelli B, Kahn CR (2006) Critical nodes in signalling pathways: Insights into insulin action. *Nat Rev Mol Cell Biol* 7(2):85–96.
- Olinski RP, Lundin L-G, Hallböök F (2006) Conserved synteny between the *Ciona* genome and human paralogs identifies large duplication events in the molecular evolution of the insulin-relaxin gene family. *Mol Biol Evol* 23(1):10–22.
- Hernández-Sánchez C, Mansilla A, de Pablo F, Zardoya R (2008) Evolution of the insulin receptor family and receptor isoform expression in vertebrates. *Mol Biol Evol* 25(6):1043–1053.
- Daza DO, Sundström G, Bergqvist CA, Duan C, Larhammar D (2011) Evolution of the insulin-like growth factor binding protein (IGFBP) family. *Endocrinology* 152(6):2278–2289.
- O'Neill MJ, et al. (2007) Ancient and continuing Darwinian selection on insulin-like growth factor II in placental fishes. *Proc Natl Acad Sci USA* 104(30):12404–12409.
- Killian JK, et al. (2000) M6P/IGF2R imprinting evolution in mammals. *Mol Cell* 5(4):707–716.
- Schwartz TS, Bronikowski AM (2011) Molecular stress pathways and the evolution of life histories in reptiles. *Molecular Mechanisms of Life History Evolution*, ed Heyland F (Oxford Univ Press, Oxford, UK).
- de Magalhães JP, Toussaint O (2002) The evolution of mammalian aging. *Exp Gerontol* 37(6):769–775.
- Swanson EM, Dantzer B (2014) Insulin-like growth factor-1 is associated with life-history variation across Mammalia. *Proc Royal Soc B Biol Sci* 281(1782):20132458.
- Sparkman AM, Vleck CM, Bronikowski AM (2009) Evolutionary ecology of endocrine-mediated life-history variation in the garter snake *Thamnophis elegans*. *Ecology* 90(3):720–728.
- Sparkman AM, Byars D, Ford NB, Bronikowski AM (2010) The role of insulin-like growth factor-1 (IGF-1) in growth and reproduction in female brown house snakes (*Lamprophis fuliginosus*). *Gen Comp Endocrinol* 168(3):408–414.
- Denley A, Cosgrove LJ, Booker GW, Wallace JC, Forbes BE (2005) Molecular interactions of the IGF system. *Cytokine Growth Factor Rev* 16(4-5):421–439.
- Ghosh P, Dahms NM, Kornfeld S (2003) Mannose 6-phosphate receptors: New twists in the tale. *Nat Rev Mol Cell Biol* 4(3):202–212.
- Clairmont KB, Czech MP (1989) Chicken and *Xenopus* mannose 6-phosphate receptors fail to bind insulin-like growth factor II. *J Biol Chem* 264(28):16390–16392.
- Dahms NM, Brzycki-Wessell MA, Ramanujam KS, Seetharam B (1993) Characterization of mannose 6-phosphate receptors (MPRs) from opossum liver: Opossum cation-independent MPR binds insulin-like growth factor-II. *Endocrinology* 133(2):440–446.

20. Yandell CA, Dunbar AJ, Wheldrake JF, Upton Z (1999) The kangaroo cation-independent mannose 6-phosphate receptor binds insulin-like growth factor II with low affinity. *J Biol Chem* 274(38):27076–27082.
21. Canfield WM, Kornfeld S (1989) The chicken liver cation-independent mannose 6-phosphate receptor lacks the high affinity binding site for insulin-like growth factor II. *J Biol Chem* 264(13):7100–7103.
22. Alvarez-Ponce D, Aguadé M, Rozas J (2013) comment on “The Molecular evolutionary patterns of the Insulin/FOXO signaling pathway”. *Evol Bioinform Online* 9:229–234.
23. Alvarez-Ponce D, et al. (2012) Molecular population genetics of the insulin/TOR signal transduction pathway: A network-level analysis in *Drosophila melanogaster*. *Mol Biol Evol* 29(1):123–132.
24. Luisi P, et al. (2012) Network-level and population genetics analysis of the insulin/TOR signal transduction pathway across human populations. *Mol Biol Evol* 29(5):1379–1392.
25. Alvarez-Ponce D, Aguadé M, Rozas J (2011) Comparative genomics of the vertebrate insulin/TOR signal transduction pathway: A network-level analysis of selective pressures. *Genome Biol Evol* 3:87–101.
26. Alvarez-Ponce D, Aguadé M, Rozas J (2009) Network-level molecular evolutionary analysis of the insulin/TOR signal transduction pathway across 12 *Drosophila* genomes. *Genome Res* 19(2):234–242.
27. Jovelín R, Phillips PC (2011) Expression level drives the pattern of selective constraints along the insulin/Tor signal transduction pathway in *Caenorhabditis*. *Genome Biol Evol* 3:715–722.
28. Wang M, et al. (2013) The molecular evolutionary patterns of the Insulin/FOXO signaling pathway. *Evol Bioinform Online* 9:1–16.
29. Li L, Stoeckert CJ, Jr, Roos DS (2003) OrthoMCL: Identification of ortholog groups for eukaryotic genomes. *Genome Res* 13(9):2178–2189.
30. Kanehisa M, Goto S (2000) KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 28(1):27–30.
31. Yang Z (2007) PAML 4: Phylogenetic analysis by maximum likelihood. *Mol Biol Evol* 24(8):1586–1591.
32. Self SG, Liang K-L (1987) Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. *J Am Stat Assoc* 82(398):605–610.
33. Goldman N, Whelan S (2000) Statistical tests of gamma-distributed rate heterogeneity in models of sequence evolution in phylogenetics. *Mol Biol Evol* 17(6):975–978.
34. Sparkman AM, et al. (2012) Rates of molecular evolution vary in vertebrates for insulin-like growth factor-1 (IGF-1), a pleiotropic locus that regulates life history traits. *Gen Comp Endocrinol* 178(1):164–173.
35. Wahren J (2004) C-peptide: New findings and therapeutic implications in diabetes. *Clin Physiol Funct Imaging* 24(4):180–189.
36. Keyhanfar M, Booker GW, Whittaker J, Wallace JC, Forbes BE (2007) Precise mapping of an IGF-I-binding site on the IGF-1R. *Biochem J* 401(1):269–277.
37. Epa VC, Ward CW (2006) Model for the complex between the insulin-like growth factor I and its receptor: Towards designing antagonists for the IGF-1 receptor. *Protein Eng Des Sel* 19(8):377–384.
38. Bayne ML, et al. (1989) The C region of human insulin-like growth factor (IGF) I is required for high affinity binding to the type 1 IGF receptor. *J Biol Chem* 264(19):11004–11008.
39. Denley A, et al. (2004) Structural determinants for high-affinity binding of insulin-like growth factor II to insulin receptor (IR)-A, the exon 11 minus isoform of the IR. *Mol Endocrinol* 18(10):2502–2512.
40. Zhang W, Gustafson TA, Rutter WJ, Johnson JD (1994) Positively charged side chains in the insulin-like growth factor-1 C- and D-regions determine receptor binding specificity. *J Biol Chem* 269(14):10609–10613.
41. Yang YW, Robbins AR, Nissley SP, Rechler MM (1991) The chick embryo fibroblast cation-independent mannose 6-phosphate receptor is functional and immunologically related to the mammalian insulin-like growth factor-II (IGF-II)/man 6-P receptor but does not bind IGF-II. *Endocrinology* 128(2):1177–1189.
42. Koduru S, Yadavalli S, Nadimpalli SK (2006) Mannose 6-phosphate receptor (MPR 300) proteins from goat and chicken bind human IGF-II. *Biosci Rep* 26(2):101–112.
43. Brown J, Jones EY, Forbes BE (2009) Interactions of IGF-II with the IGF2R/cation-independent mannose-6-phosphate receptor mechanism and biological outcomes. *Vitam Horm* 80:699–719.
44. Williams C, et al. (2012) An exon splice enhancer primes IGF2:IGF2R binding site structure and function evolution. *Science* 338(6111):1209–1213.
45. Brown J, et al. (2008) Structure and functional analysis of the IGF-1I/IGF2R interaction. *EMBO J* 27(1):265–276.
46. Fares MA, McNally D (2006) CAPS: Coevolution analysis using protein sequences. *Bioinformatics* 22(22):2821–2822.
47. Rodionov A, Bezginov A, Rose J, Tillier ER (2011) A new, fast algorithm for detecting protein coevolution using maximum compatible cliques. *Algorithms Mol Biol* 6(1):17.
48. Duan C, Xu Q (2005) Roles of insulin-like growth factor (IGF) binding proteins in regulating IGF actions. *Gen Comp Endocrinol* 142(1–2):44–52.
49. Forbes BE, McCarthy P, Norton RS (2012) Insulin-like growth factor binding proteins: A structural perspective. *Front Endocrinol (Lausanne)* 3:38.
50. Moloney AM, et al. (2010) Defects in IGF-1 receptor, insulin receptor and IRS-1/2 in Alzheimer's disease indicate possible resistance to IGF-1 and insulin signalling. *Neurobiol Aging* 31(2):224–243.
51. Han M, et al. (2013) Evolutionary rate patterns of genes involved in the *Drosophila* Toll and Imd signaling pathway. *BMC Evol Biol* 13(1):245.
52. Song X, Jin P, Qin S, Chen L, Ma F (2012) The evolution and origin of animal Toll-like receptor signaling pathway revealed by network-level molecular evolutionary analyses. *PLoS ONE* 7(12):e51657.
53. Cui Q, Purisima EO, Wang E (2009) Protein evolution on a human signaling network. *BMC Syst Biol* 3(1):21.
54. Montanucci L, Laayouni H, Dall'Olio GM, Bertranpetit J (2011) Molecular evolution and network-level analysis of the N-glycosylation metabolic pathway across primates. *Mol Biol Evol* 28(1):813–823.
55. Fraser HB, Hirsh AE, Steinmetz LM, Scharfe C, Feldman MW (2002) Evolutionary rate in the protein interaction network. *Science* 296(5568):750–752.
56. Fraser HB, Wall DP, Hirsh AE (2003) A simple dependence between protein evolution rate and the number of protein-protein interactions. *BMC Evol Biol* 3(1):11.
57. Bloom JD, Adami C (2004) Evolutionary rate depends on number of protein-protein interactions independently of gene expression level: Response. *BMC Evol Biol* 4(1):14.
58. Larracuente AM, et al. (2008) Evolution of protein-coding genes in *Drosophila*. *Trends Genet* 24(3):114–123.
59. Suh Y, et al. (2008) Functionally significant insulin-like growth factor I receptor mutations in centenarians. *Proc Natl Acad Sci USA* 105(9):3438–3442.
60. Hawkes C, Kar S (2004) The insulin-like growth factor-II/mannose-6-phosphate receptor: Structure, distribution and function in the central nervous system. *Brain Res Brain Res Rev* 44(2–3):117–140.
61. Russo VC, Gluckman PD, Feldman EL, Werther GA (2005) The insulin-like growth factor system and its pleiotropic functions in brain. *Endocr Rev* 26(7):916–943.
62. Yuan X-N, Jiang X-Y, Pu J-W, Li Z-R, Zou S-M (2011) Functional conservation and divergence of duplicated insulin-like growth factor 2 genes in grass carp (*Ctenopharyngodon idellus*). *Gene* 470(1–2):46–52.
63. Brown AL, et al. (1986) Developmental regulation of insulin-like growth factor II mRNA in different rat tissues. *J Biol Chem* 261(28):13144–13150.
64. Killian JK, et al. (2001) Monotreme IGF2 expression and ancestral origin of genomic imprinting. *J Exp Zool* 291(2):205–212.
65. Méndez E, Planas JV, Castillo J, Navarro I, Gutiérrez J (2001) Identification of a type II insulin-like growth factor receptor in fish embryos. *Endocrinology* 142(3):1090–1097.
66. Sivaramakrishna Y, Amancha PK, Siva Kumar N (2009) Reptilian MPR 300 is also the IGF-1IR: Cloning, sequencing and functional characterization of the IGF-II binding domain. *Int J Biol Macromol* 44(5):435–440.
67. Zhou M, Ma Z, Sly WS (1995) Cloning and expression of the cDNA of chicken cation-independent mannose-6-phosphate receptor. *Proc Natl Acad Sci USA* 92(21):9762–9766.
68. Garmroudi F, Devi G, Slentz DH, Schaffer BS, MacDonald RG (1996) Truncated forms of the insulin-like growth factor II (IGF-II)/mannose 6-phosphate receptor encompassing the IGF-II binding site: Characterization of a point mutation that abolishes IGF-II binding. *Mol Endocrinol* 10(6):642–651.
69. Gadd TS, Osgerby JC, Wathes DC (2002) Regulation of insulin-like growth factor binding protein-6 expression in the reproductive tract throughout the estrous cycle and during the development of the placenta in the ewe. *Biol Reprod* 67(6):1756–1762.
70. Murphy BF, Thompson MB (2011) A review of the evolution of viviparity in squamate reptiles: The past, present and future role of molecular biology and genomics. *J Comp Physiol B* 181(5):575–594.
71. Castoe TA, et al. (2013) The Burmese python genome reveals the molecular basis for extreme adaptation in snakes. *Proc Natl Acad Sci USA* 110(51):20645–20650.
72. Lowe CB, Clarke JA, Baker AJ, Haussler D, Edwards SV (2015) Feather development genes and associated regulatory innovation predate the origin of *Dinosauria*. *Mol Biol Evol* 32(1):23–28.
73. Vijay N, Poelstra JW, Künstner A, Wolf JBW (2013) Challenges and strategies in transcriptome assembly and differential gene expression quantification. A comprehensive *in silico* assessment of RNA-seq experiments. *Mol Ecol* 22(3):620–634.
74. Austad SN (2010) Cats, “rats,” and bats: The comparative biology of aging in the 21st century. *Integr Comp Biol* 50(5):783–792.
75. Alberts SC, et al. (2013) Reproductive aging patterns in primates reveal that humans are distinct. *Proc Natl Acad Sci USA* 110(33):13440–13445.
76. Yamamoto R, Tatar M (2011) Insulin receptor substrate chico acts with the transcription factor FOXO to extend *Drosophila* lifespan. *Aging Cell* 10(4):729–732.
77. Bartke A (2008) Impact of reduced insulin-like growth factor-1/insulin signaling on aging in mammals: Novel findings. *Aging Cell* 7(3):285–290.
78. Tacutu R, et al. (2013) Human Ageing Genomic Resources: Integrated databases and tools for the biology and genetics of ageing. *Nucleic Acids Res* 41(Database issue, D1):D1027–D1033.
79. Li Y, de Magalhães JP (2013) Accelerated protein evolution analysis reveals genes and pathways associated with the evolution of mammalian longevity. *Age (Dordr)* 35(2):301–314.
80. Soerensen M, et al. (2012) Human longevity and variation in GH/IGF-1/insulin signaling, DNA damage signaling and repair and pro/antioxidant pathway genes: Cross sectional and longitudinal studies. *Exp Gerontol* 47(5):379–387.
81. Ziv E, Hu D (2011) Genetic variation in insulin/IGF-1 signaling pathways and longevity. *Ageing Res Rev* 10(2):201–204.
82. de Magalhães JP (2014) Why genes extending lifespan in model organisms have not been consistently associated with human longevity and what it means to translation research. *Cell Cycle* 13(17):2671–2673.
83. Miller DAW, Janzen FJ, Fellers GM, Kleeman PM, Bronikowski A (2014) Biodemography of ectothermic tetrapods provides insights into the evolution and plasticity of mortality trajectories. *Sociality, Hierarchy, Health: Comparative Demography Advances in Biodemography*, eds Weinstein M, Lane MA (The National Academies Press, Washington, DC).
84. Liu Y, Schmidt B, Maskell DL (2010) MSAProbs: Multiple sequence alignment based on pair hidden Markov models and partition function posterior probabilities. *Bioinformatics* 26(16):1958–1964.
85. Wernersson R, Pedersen AG (2003) RevTrans: Multiple alignment of coding DNA from aligned amino acid sequences. *Nucleic Acids Res* 31(13):3537–3539.
86. Castresana J (2000) Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol* 17(4):540–552.