University of Reading

The varying rate of phenotypic evolution and natural selection

PhD Thesis

School of Biological Sciences

Joanna Baker

Submitted October 31st 2016

Declaration

I confirm that this is my own work and the use of all material from other sources has been properly and fully acknowledged

Chapter 1 and Chapter 3 are published as (respectively):

Baker J, Meade A, Pagel M & Venditti C (2015). *Adaptive evolution toward larger size in mammals.* Proceedings of the National Academy of Sciences USA, 112 (16): 5093-5098.

Baker J, Meade A, Pagel M & Venditti C (2016). *Positive phenotypic selection inferred from phylogenies.* Biological Journal of the Linnean Society, 118: 95-115.

Author contributions to both publications are as follows:

- Generated hypotheses: **JB**, AM, MP, CV
- Collected data and performed analyses: **JB**
- Writing: **JB** wrote the initial draft of the submitted manuscript; AM, MP and CV contributed to subsequent drafts.

Joanna Baker

Abstract

Changes in the rate at which biological evolution proceeds are widespread and common. Advancements in methodology make it possible to characterize such historical evolution more accurately and reveal complex scenarios where rates vary among organismic groups and even between individual lineages. The work presented here embraces such complexity, seeking to exploit phenotypic rate variation to reconstruct patterns and processes of evolution deep in time with unprecedented accuracy. Chapters 1 and 2 demonstrate for the first time that it is possible to reveal historical, directional trends in morphology that played out over millions and millions of years using only data from living species. These trends arose as a consequence of rapid and repeated instances of directional evolutionary change and the approach employed to detect them may be the only way to study historical adaptive trends in morphology that cannot otherwise be observed in the fossil record. Where evolution is fastest, natural selection has acted more strongly; this idea is developed further in Chapter 3 which presents a novel way to characterize an exceptional subclass of rates of morphological evolution that can be defined as positive phenotypic selection. In both Chapters 3 and 4 it is shown that such intense episodes of natural selection have punctuated the evolution of diverse groups including plants, dinosaurs and hominins. Chapter 5 demonstrates that it is possible to uncover explicit underlying causes of positive phenotypic selection and takes us one step closer to being able to truly understand the drivers of natural diversity. As a complete work, this thesis harnesses and exploits phenotypic rate heterogeneity to inform inferences about patterns and processes of evolution deep in time and to understand how natural selection has acted to sculpt morphology, giving rise to the diversity we observe today both in living species and the fossil record.

Acknowledgements

I first have to thank my supervisor, Chris Venditti for giving me the opportunity to work on so many exciting things. Not only this, but I am eternally grateful for his inspiration and motivation that have driven me to become what I am today – I do not think anyone could wish for a better supervisor.

Secondly, so very much of what I have achieved in this thesis would have been very difficult without my co-supervisor Andrew Meade and absolutely impossible without the countless methods and programs he has implemented to facilitate many of my analyses.

Many of the best ideas are inspired and grown through discussion and debate. For such discussions I thank the highly motivated and productive atmosphere provided by the Evolutionary Biology Group at the University of Reading and all of those people that I have had the joy to share an office with over the last few years. This gratitude also extends to Stuart Humphries who since my days at the University of Hull has been a great source of great advice and inspiration.

This thesis has been a long and – at times – difficult road to travel, and so my thanks go out to all those people who have stood by me, from beginning to end. This includes all of those people above but also my family and friends. A special mention ought to be given to Mathew for putting up with me through all the highs and lows of the last few years without (much) of a complaint.

I must also extend my appreciation for the support and funding I have received from the University of Reading that made all of this possible in the first instance.

I am forever grateful for all of the support I have received over the last four years, and am excited for what the future holds. Thank you.

Contents

Introduction

A paramecium in a pond and a man reading this book, a fish in the sea and a bird in the air, an opossum up a tree and a bat in a cave, or a cat stalking a mouse and a dog chasing a rabbit clearly and significantly represent sharply discontinuous adaptive types with, in the sequence of examples, characteristic differences in the breadth of the discontinuity.

- George Gaylord Simpson, 1953

The diversity of life on Earth is astonishing. For example, even within a single group such as mammals there is variation in body size that spans over 7 orders of magnitude – ranging from tiny bats and shrews right up to enormous elephants and whales (Jones *et al.*, 2009). Most researchers would agree that the ancestral mammal was small in size (Novacek, 1999; Stanley, 1973); some of the earliest mammals known from the fossil record are estimated to be approximately the size of modern shrews (Hu *et al.*, 2010; Ji *et al.*, 2002; Luo *et al.*, 2011). This means that all modern mammals evolved from a single shrew-sized ancestor that existed over 160 million years ago (Bininda-Emonds *et al.*, 2007; Faurby and Svenning, 2015; Fritz *et al.*, 2009; Hedges *et al.*, 2006; Hedges *et al.*, 2015), yet most of the mammals alive today are larger – some by orders of magnitude. This begs the question – how could such diversity arise from a single ancestor during exactly the same evolutionary timeframe?

In order to understand how evolution and adaptation can combine to produce biodiversity, it is first necessary to understand how species are related to one another; species cannot be considered as independent data points as they have shared millions and millions of years of their evolutionary history (Felsenstein, 1985; Harvey and Pagel, 1991) – e.g. chimpanzees and modern humans are similar to each other in many ways owing to their extremely close phylogenetic relationship, only diverging from one another within the last 10 million years or so (Arnason *et al.*, 1998; Dembo *et al.*, 2016; White *et al.*,

2009). It has only been since the advent of modern phylogenetic comparative methods that scientists have been able to detect historical evolutionary relationships and patterns of morphological change that have played out over the branches of a phylogenetic tree (Martins and Hansen, 1997; Pagel, 1997, 1999). The earliest evolutionary models assumed a simple homogenous process, based on an underlying Brownian motion model of change (Felsenstein, 1973; Felsenstein, 1985). The Brownian motion model assumes that at any given instance of time, phenotypic change can occur in any direction regardless of its current or starting value but that such change occurs at a constant rate or tempo across all branches of the phylogenetic tree. Simply put, non-directional changes occur over time, resulting in a constant increase in the variation of the trait (Felsenstein, 1973).

Several modifications to this simple underlying process have also been proposed that allow the evolutionary process to vary in tempo, mode, or both. This includes parameters allowing the tempo of evolution to be time-dependent such that the rate of evolution can change over time e.g. early burst (Harmon *et al.*, 2010) or delta (Pagel, 1997, 1999) models - such approaches have been used to explain or describe adaptive radiations. Other modifications allow trait change to be concentrated at speciation events (e.g. Pagel, 1999) as proposed by the theory of punctuated equilibrium (Eldredge and Gould, 1972) or to model the probability that closely related species will be more similar in trait value than expected by chance (e.g. Pagel, 1999). Yet another model, the Ornstein-Uhlenbeck (OU) process (Butler and King, 2004; Felsenstein, 1985; Hansen, 1997) considers evolution to proceed towards an optimum – species are 'pulled' to this optimum value with some strength. The application of the OU model to biological evolution was originally developed by work at population level (e.g. Lande, 1976) and has been extended to apply to optima across species on the basis of a very simple interpretation of George Gaylord Simpson's classic theory of adaptive zones (Simpson, 1944; Simpson, 1953). The theory of adaptive zones proposes that there are particular 'areas' of morphology in which a species or group of species persists in which they are inherently adapted; where there are shifts in ecology or environment that exert pressures on that morphology to change it can be

considered a shift in adaptive zone – or a change in optimum. Recent modifications to the OU model also allow different optima among groups of species or allow different lineages to experience different 'pulls' towards a single optimum (Butler and King, 2004; Uyeda and Harmon, 2014).

At the heart of all the evolutionary models described thus far is an underlying assumption that simple, homogenous processes can describe the tempo and mode of evolution across groups of species – whether that process be described by a single rate parameter, a change in the rate over time, an amount of trait change expected at any one speciation event, or by one or more optimum trait values or 'pulls' toward any optima. However, even Simpson championed the idea that evolution was unlikely to occur as a constant and uniform process across all species in his classic works (Simpson, 1944; Simpson, 1953). For example, distinctions between adaptive zones are recognized to be blurry depending on the scale of the organisms of interest e.g. two sibling birds within the same nest can be considered to have different adaptive types (Simpson, 1953). This raises questions about how one might define an optimum trait or set of optima that is enforced across multiple species as in the OU model – especially considering distinct discontinuities in adaptive zones across species such as those described in the opening quote of this chapter. The complexities and nuances of biological diversity cast a doubting shadow on just how well evolutionary processes acting across diverse groups of species can be reduced to a simple model describing one or few aspects or patterns of evolution.

There is now abundant evidence that the rate of morphological evolution varies among groups of species and even between individual lineages (Benson *et al.*, 2014; Benson and Choiniere, 2013; Puttick *et al.*, 2014; Rabosky and Adams, 2012; Rabosky *et al.*, 2013; Steeman *et al.*, 2009) using a new generation of phylogenetic comparative methods that automatically identifies instances of rate variation throughout evolutionary history across phylogenetic trees of even thousands of species (Eastman *et al.*, 2011; Kratsch and McHardy, 2014; Landis *et al.*, 2013; Rabosky, 2014; Revell *et al.*, 2012; Thomas and Freckleton, 2012; Venditti *et al.*, 2011). Attempts to characterize the evolutionary process

by using simple models of trait evolution is therefore likely be a naive approach; instead, biological phenotypic diversity arises as a consequence of complex evolutionary processes that cannot be described by simple models.

There are now several different models available for detecting phenotypic rate variation along the branches of a phylogenetic tree (Eastman *et al.*, 2011; Kratsch and McHardy, 2014; Rabosky, 2014; Revell *et al.*, 2012; Thomas and Freckleton, 2012; Venditti *et al.*, 2011). The major difference among each of these lies in how the number of accelerations or decelerations in rate (compared to some background rate of phenotypic change) is defined. Some models work by pre-defining the exact (usually small) number of rate shifts present in a given phylogeny, allowing only a discrete number of rate changes to be modelled as a part of the evolutionary process (Revell *et al.*, 2012; Thomas and Freckleton, 2012). At the other end of the spectrum, there are models that estimate an individual rate along each branch (Bouckaert *et al.*, 2014; Kratsch and McHardy, 2014; Mooers *et al.*, 1999). Both extremes – from either a priori defining the number of different rates to necessarily allowing all rates to vary – suffer as it is impossible to know exactly how many rates may exist in a phylogeny at any one time (Venditti *et al.*, 2011). Several approaches have developed ways of dealing with this by estimating the number of rate shifts within a phylogeny as a part of the process that also estimates the rates themselves (Eastman *et al.*, 2011; Rabosky, 2014; Venditti *et al.*, 2011). *AUTEUR* and *BAMM* are both implemented within a Bayesian Markov chain Monte Carlo (MCMC) framework, placing a prior on the number of rate shifts that can be present in any one phylogeny – this allows individual branches to fall into some undefined rate category, the number of which will be estimated as a part of the model. On the other hand, the variable rates model of Venditti *et al.,* (2011) is also implemented in an MCMC framework but additionally allows for the possibility that there may be nuanced differences in the rate of phenotypic evolution between branches. It is the only one of the three MCMC models developed for simultaneously inferring both the number and magnitude of rate shifts that estimates how many *rates* are present rather than how many *rate categories*. A recent simulation study (Chira and Thomas, 2016) tested

the ability of both the variable rates model and that of *BAMM* – finding that both models tended to accurately recover shifts in evolutionary rate over trees of varying sizes. However, *BAMM* suffered in its ability to detect subtle changes in the rate of evolution along a single lineage. This implies that a model that estimates the number of rate changes present in a phylogeny whilst also allowing for the possibility that individual branches can have their own unique rate of change may provide a more accurate portrayal of the evolutionary process than one assuming a simple homogenous underlying rate of evolution. Therefore, the variable rates model of Venditti *et al.* (2011) is used to estimate phenotypic rate heterogeneity throughout this thesis; the application of the model is described in more detail throughout in addition to its suitability and relevance to the particular evolutionary scenarios of each chapter.

The idea of evolutionary heterogeneity in species phenotypes is in parallel with the picture emerging from similar studies of genetic data which demonstrates that individual lineages can experience intense bursts of change and vary in the rate of genetic evolution (Murrell *et al.*, 2012; Nadeau and Jiggins, 2010; Nikaido *et al.*, 2014; Smith *et al.*, 2015; Yang, 2002, 2006). Evolutionary models which can identify at what time and in which lineages the rate of change was sped up or slowed down in comparison to some background rate of change are therefore gradually becoming the default for comparative analysis – both in a genetic and a phenotypic context.

There is some evidence that including fossil information alongside data from extant species can inform, and even improve, the accuracy of our inferences about the past (Albert *et al.*, 2009; Bokma *et al.*, 2015; Finarelli and Flynn, 2006; Finarelli and Goswami, 2013; Pant *et al.*, 2014; Slater *et al.*, 2012). Despite a growing catalogue of phylogenetic trees that sample both extinct and extant diversity across several groups (Aze *et al.*, 2011; Clarke *et al.*, 2007; Gatesy *et al.*, 2013; Gavryushkina *et al.*, 2016; Geisler *et al.*, 2011; Ksepka *et al.*, 2006; Marx and Fordyce, 2015; Montgomery *et al.*, 2013), there are still many difficulties with constructing and dating phylogenetic trees that include fossil species (Lee *et al.*, 2014; O'Leary *et al.*, 2013; Pyron, 2011; Ronquist *et al.*, 2012; Slater, 2015; Wood *et*

al., 2012) and it continues to be an active field of research. Therefore, if a model that detects and accounts for rate heterogeneity is a good characterization of the true evolutionary processes that gave rise to current diversity, then this may provide an alternative way to more accurately reconstruct historical evolution – without using fossils.

For example, rate heterogeneity may help to reconcile the contrasts between conclusions made by palaeontologists and those made by researchers studying the species alive today. One such example is that of Cope's rule: the enduring idea that there is a general tendency for species to become larger in size through geological time (Stanley, 1973). Evidence taken from the fossil record of mammals seems to provide consistent and robust support for Cope's rule (Alroy, 1998; Bokma *et al.*, 2015; MacFadden, 1986; Raia *et al.*, 2012; Smith *et al.*, 2010; Van Valkenburgh *et al.*, 2004). However, analyses of exclusively extant species find no evidence for any trend in body size during the evolution of modern mammals (Bokma *et al.*, 2015; Monroe and Bokma, 2010) – in stark contrast to results obtained using fossil data. Mammal body size has previously been shown to have evolved with substantial rate heterogeneity (Venditti *et al.*, 2011), but nobody has ever tested how these shifts in rate link to size itself – if increases in body sizes have been favoured throughout mammalian evolution then large increases in size should be associated with increases in the rate of morphological evolution.

Chapter 1[1] uses the variable rates model to scale branches of a phylogeny relative to the rate of evolution through time based on information about living mammal species. This results in a phylogeny where branches are scaled by their rate of evolution – where branches have been made longer, they have evolved at a faster rate, whereas shortened branches have evolved more slowly. If evolution were to be replayed across this scaled tree, there is more opportunity for change in lineages with faster rates of evolution within the same evolutionary timescale. By accounting for the heterogeneous rates of body size

[1] Published as Baker *et al.,* 2015; the rest of this thesis will refer to the publication.

evolution in this way, meaningful variation is introduced into the branch lengths of an extant phylogeny. This offers the ability to test for the presence of trends using the relationship between path length and body size (Pagel, 1997, 1999), where path length in this case is a measure of total changes in the rate of evolution a species has experienced during its entire evolutionary history. A significant positive relationship between body size and rate was found across all mammals and within individual orders; during periods of rapid morphological evolution, mammals preferentially evolved towards larger body sizes. This fits the predictions of Cope's rule and is unlikely to be explained by non-adaptive mechanisms for increasing size (Gould, 1988, 1997; McShea, 1994, 1998; Stanley, 1973; Wagner, 1996). This is the first demonstration that it is possible to recover ancient evolutionary trends using only data from the species that exist today. The results in Chapter 1 imply that incorporation of rate heterogeneity go some way towards improving the accuracy of our historical inferences.

The possibility of revealing adaptive trends in traits is intriguing – but also presents a problem. The strong correlation between body size and many biological characteristics (Bonner, 2011; Peters, 1986; Schmidt-Nielsen, 1984) is likely to complicate and impede our ability to detect adaptive changes in other morphologies – any changes we detect are likely to be confounded by changes in body size. Chapter 2 seeks to determine whether it is possible to detect a trend in a trait whilst simultaneously accounting for its relationship with body size – using a novel extension of the variable rates model described in more detail in Chapter 3[2] – and using vertebrate testes size as a case study. The size of testes in vertebrate species is a measure of reproductive investment – and is tightly linked with body size (Hayward and Gillooly, 2011; MacLeod, 2010, 2014; MacLeod and MacLeod, 2009). However, it is easy to imagine a situation in which a species might benefit from increasing the size of its testes beyond body size. There is a huge body of literature

[2] Published as Baker *et al.,* 2016; the rest of this thesis will refer to the publication.

demonstrating that an increase in the amount of sperm competition is strongly linked to an increase in testes size across many different groups of species (reviewed in Parker *et al.*, 1997). Conversely, it is equally easy to imagine a situation in which a species might benefit from decreasing the size of its testes. Testes are expensive tissues to grow and maintain (Dines *et al.*, 2015; Hayward and Gillooly, 2011; Kenagy and Trombulak, 1986; Pitnick *et al.*, 2006; Simmons and Emlen, 2006; Warren and Iglesias, 2012) – where there has been a reduction in sperm competition or an alternative means of increasing reproductive success e.g. an improvement in ejaculate quality (e.g. Møller, 1988) it would be to the species' advantage to reduce excess investment in reproductive tissues. This is closely linked to the *expensive-tissue hypothesis* (Aiello and Wheeler, 1995) which proposes that an increase in the size of metabolically expensive tissues such as the brain will result in a decrease in the size of other expensive organs – including testes.

The results discussed in Chapter 2 show that the vertebrate tree of life has been punctuated by heritable accelerations in rate – such that a rate increase along a single branch has been inherited by all descendent branches. Ultimately, these give rise to a general tendency for decreases in testes size after accounting for changes in body size. Throughout vertebrate evolutionary history, there has been a strong and significant selective pressure acting to reduce male investment in reproductive tissue.

Chapters 1 and 2 work on the assumption that faster rates of evolution imply adaptive change. This idea has not gone unnoticed in the literature – several authors have linked the rate of morphological evolution to adaptation (Eastman *et al.*, 2011; Kratsch and McHardy, 2014; Kutsukake and Innan, 2013, 2014; Rabosky, 2014; Revell *et al.*, 2012; Thomas and Freckleton, 2012; Venditti *et al.*, 2011). Chapter 3 builds further on this idea, seeking to identify exceptional instances of morphological evolution that might be attributable to historical positive phenotypic selection. Inspiration is taken from analyses of genetic data that detect adaptive changes in genes along individual branches of a phylogenetic tree by looking at the proportion of synonymous neutral substitutions (d_S) to non-synonymous protein-coding substitutions (d_N) (Yang, 2002, 2006). In this context,

a $d_N/d_S > 1$ implies positive selection, where non-synonymous substitutions account for more than half of all substitutions along a branch. The approach proposed in Chapter 3 draws parallels to the genetic d_N/d_S ratio by introducing a branch-specific metric that measures the amount of phenotypic change occurring along a branch that can be attributed to the background (akin to 'neutral'; though see Chapter 3) rate of evolution (Δ_B) and the amount of phenotypic evolution that can be attributed to deviations away from that background rate (Δ_V). The amount of expected phenotypic evolution occurring as a consequence of rate variation necessarily is not independent from the amount expected from the background rate ($\Delta_V = \Delta_B r$, where r = the branch-specific rate of evolution); thus the metric is calculated as Δ_V/Δ_B. This defines the amount of phenotypic change that occurs *beyond* that which is expected given the overall underlying background rate of evolution across the phylogeny as a whole. Along an individual branch, where $\Delta_V/\Delta_B > 2$ it is identified as an exceptional burst of morphological evolution and defined as positive phenotypic selection. This reasoning is applied across five different case studies in Chapter 3, where Δ_V/Δ_B is calculated using a variable rates analysis – and defines positive phenotypic selection in each. Each of the five case studies demonstrates how natural selection has acted to sculpt different morphologies throughout nature: fruit size in plants, limb proportions in dinosaurs, eye shape in mammals, body size in *Anolis* lizards, and tooth area in primates.

The positive phenotypic selection characterizing the evolution of primate teeth found in Chapter 3 occurred primarily during the evolution of hominins – those species that have arisen since the divergence of chimpanzees and modern humans. Positive phenotypic selection has acted to substantially reduce the size of lower molars in the lineage of hominins leading to *Homo erectus, H. neanderthalensis*, and our own species, *H. sapiens*. This reduction in size has previously been linked to reduced feeding times in these hominins (Organ *et al.*, 2011) possibly as a consequence of the uniquely hominin innovation of food processing or cooking with fire (Brace *et al.*, 1987; Organ *et al.*, 2011; Wrangham, 2009). Today, evolutionary biologists unequivocally see humans as just

another branch in the tree of life – yet there lingers in our psyche the notion that the evolutionary route to humanity has been characterised by exceptional periods of strong selection and evolutionary innovation or 'experimentation' (Barton and Venditti, 2014; Hublin, 2015; Hublin *et al.*, 2015; Organ *et al.*, 2011; Pampush, 2015).

Results showing accelerated rates in hominin molars are congruent with the prominent current view that the anatomical features considered to be human – including those associated with behavioural innovations such as tool use or bipedalism – arose as a result of adaptive evolution (Antón *et al.*, 2014; Hublin, 2015; Schroeder *et al.*, 2014). However while this adaptive view is pervasive and fits with the perception that we as humans are somehow 'special' there is counterevidence indicating that at least some of our unique morphology could have been achieved by more passive processes (Barton and Venditti, 2013; Benazzi *et al.*, 2015; Schroeder *et al.*, 2014). Chapter 3 goes some way towards suggesting that, at least in terms of dental evolution, that there has in fact been strong selection pressures driving morphological changes in the lineage leading to our own species.

Chapter 4 demonstrates further the role of natural selection during the evolution of hominins including our own species by focussing on changes in the semicircular canals: organs within the inner ear responsible for balance and motion (Day and Fitzpatrick, 2005) that have been implicated in the origin of human bipedalism (Spoor *et al.*, 1994, 1996). Larger vertically oriented and smaller horizontally oriented canals are thought to have facilitated a transition from early hominins such as *Australopithecus* which simply possessed the ability to walk on two legs to fully committed bipedal species with the ability to run and jump (Harcourt-Smith, 2015; Spoor *et al.*, 1994) including *H. erectus* and *H. sapiens* (Wood, 2002). Chapter 4 uses the variable rates model in combination with the metric for detecting positive phenotypic selection introduced in Chapter 3 to determine where, when, and how strongly natural selection has acted to sculpt the sizes of the semicircular canals during primate evolution. Rather than pinpointing specific adaptive shifts in morphology that can be simply linked with the origins of bipedalism, Chapter 4

finds results implying a complex scenario – where mosaic evolution has played a key role in the evolution of hominins, including modern humans.

Although the results discussed in Chapter 4 find some evidence supporting the idea that there have been directional shifts in the size of the semicircular canals directly attributable to positive phenotypic selection, these are not clearly linked to the origin of bipedalism. Instead, during the evolution of the hominins, there have been multiple shifts towards both larger and smaller semicircular canals along different branches of the phylogenetic tree and at different times. This implies a diverse mosaic of selection pressures acting to shape semicircular canal morphology in hominins. Like the positive phenotypic selection we identify in the five case studies in Chapter 3, it is only possible to speculate about the underlying ecological, environmental or biological drivers of the selective changes without additional analyses.

Chapter 5 demonstrates that it is possible to provide an explanation for observable positive phenotypic selection. Although it is difficult to imagine a scenario in which a rapid burst of morphological change does not come about as a consequence of natural selection, demonstrating the ecological, behavioural or environmental causes underlying positive phenotypic selection can reveal and confirm the relationship between the rate of evolution and the strength of natural selection. In Chapter 3 positive phenotypic selection is found acting to increase variation in eye shapes amongst mammals. Variation in eye morphology has previously been shown to be associated with activity pattern (e.g. Banks *et al.*, 2015; Hall and Ross, 2007; Hall, 2008; Kirk and Kay, 2004; Schmitz and Motani, 2010; Schmitz and Wainwright, 2011) and Chapter 3 discusses the possibility that high rates of evolution and positive phenotypic selection acting to alter eye shape could be linked to activity pattern. For example, in carnivores – where it is found that positive phenotypic selection has acted continuously throughout the history of the group – there have been many more transitions among the three different activity patterns (diurnal, nocturnal or cathemeral) than in any other mammal order (also shown in Chapter 3).

Chapter 5 finds that nearly 90% of all positive phenotypic selection can be attributed directly to the distinct evolutionary trajectory observed in mammals of different activity patterns: activity pattern explains an extraordinary amount of all observed positive phenotypic selection acting on eye shape in mammals. This demonstrates the ability to uncover explicit ecological underlying causes of the exceptional instances of morphological evolution that have been defined as positive phenotypic selection.

This thesis harnesses the variation in the morphological rate of evolution that has been found to be so prevalent throughout nature and exploits it to reveal otherwise impossible patterns of historical evolution that gave rise to the diversity of life on Earth. It demonstrates the ability to uncover historical adaptive trends from living species; providing an approach that researchers can use to detect signatures for adaptive directional change amongst groups of species. It also introduces a novel metric that can be used to identify instances of positive phenotypic selection that have sculpted the morphology of both living and extinct species – and detect such natural selection in multiple different groups, including our own species. Finally, it also shows how to find the underlying causes of this historical positive phenotypic selection – making it possible to understand how abiotic and biotic factors may combine and interact to exert selection pressures leading to adaptation and results in variation among species. Ultimately, this thesis provides a groundwork upon which researchers can build upon to really begin to understand not only that there is rate variation in nature, but how exactly that variation has arisen and what it means for the evolution of diversity.

References

Aiello LC & Wheeler P 1995. The expensive-tissue hypothesis: the brain and the digestive system in human and primate evolution. *Current anthropology,* 36 (2): 199-221.

Albert JS, Johnson DM & Knouft JH 2009. Fossils provide better estimates of ancestral body size than do extant taxa in fishes. *Acta Zoologica,* 90: 357-384.

Alroy J 1998. Cope's rule and the dynamics of body mass evolution in North American fossil mammals. *Science,* 280 (5364): 731-734.

Antón SC, Potts R & Aiello LC 2014. Evolution of early *Homo*: An integrated biological perspective. *Science,* 345 (6192).

Arnason U, Gullberg A & Janke A 1998. Molecular timing of primate divergences as estimated by two nonprimate calibration points. *Journal of Molecular Evolution,* 47 (6): 718-727.

Aze T, Ezard TH, Purvis A, Coxall HK, Stewart DR, Wade BS & Pearson PN 2011. A phylogeny of Cenozoic macroperforate planktonic foraminifera from fossil data. *Biological Reviews of the Cambridge Philosophical Society,* 86 (4): 900-27.

Baker J, Meade A, Pagel M & Venditti C 2015. Adaptive evolution toward larger size in mammals. *Proceedings of the National Academy of Sciences USA,* 112 (16): 5093-5098.

Baker J, Meade A, Pagel M & Venditti C 2016. Positive phenotypic selection inferred from phylogenies. *Biological Journal of the Linnean Society,* 118: 95-115.

Banks MS, Sprague WW, Schmoll J, Parnell JAQ & Love GD 2015. Why do animal eyes have pupils of different shapes? *Science Advances,* 1 (7): e1500391.

Barton RA & Venditti C 2013. Human frontal lobes are not relatively large. *Proceedings of the National Academy of Sciences USA,* 110 (22): 9001-9006.

Barton RA & Venditti C 2014. Rapid evolution of the cerebellum in humans and other great apes. *Current Biology,* 24 (20): 2440-2444.

Benazzi S, Nguyen HN, Kullmer O & Hublin J-J 2015. Exploring the biomechanics of taurodontism. *Journal of Anatomy,* 226 (2): 180-188.

Benson RBJ, Campione NE, Carrano MT, Mannion PD, Sullivan C, Upchurch P & Evans DC 2014. Rates of dinosaur body mass evolution indicate 170 million years of sustained ecological innovation on the avian stem lineage. *PLoS Biology,* 12 (5): e1001853.

Benson RBJ & Choiniere JN 2013. Rates of dinosaur limb evolution provide evidence for exceptional radiation in Mesozoic birds. *Proceedings of the Royal Society B: Biological Sciences,* 280 (1768): e20131780.

Bininda-Emonds OR, Cardillo M, Jones KE, MacPhee RD, Beck RM, Grenyer R, Price SA, Vos RA, Gittleman JL & Purvis A 2007. The delayed rise of present-day mammals. *Nature,* 446 (7135): 507-12.

Bokma F, Godinot M, Maridet O, Ladevèze S, Costeur L, Solé F, Gheerbrant E, Peigné S, Jacques F & Laurin M 2015. Testing for Depéret's rule (body size increase) in mammals using combined extinct and extant data. *Systematic Biology,* 65 (1): 98-108.

Bonner JT 2011. *Why size matters: from bacteria to blue whales,* Princeton, New Jersey, Princeton University Press.

Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu C-H, Xie D, Suchard MA, Rambaut A & Drummond AJ 2014. BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Computational Biology,* 10 (4): e1003537.

Brace CL, Rosenberg KR & Hunt KD 1987. Gradual change in human tooth size in the Late Pleistocene and post-Pleistocene. *Evolution,* 41 (4): 705-720.

Butler MA & King AA 2004. Phylogenetic comparative analysis: a modeling approach for adaptive evolution. *The American Naturalist,* 164 (6): 683-695.

Chira AM & Thomas GH 2016. The impact of rate heterogeneity on inference of phylogenetic models of trait evolution. *Journal of Evolutionary Biology,* [Epub ahead of print].

Clarke JA, Ksepka DT, Stucchi M, Urbina M, Giannini N, Bertelli S, Narváez Y & Boyd CA 2007. Paleogene equatorial penguins challenge the proposed relationship between biogeography, diversity, and Cenozoic climate change. *Proceedings of the National Academy of Sciences USA,* 104 (28): 11545-11550.

Day BL & Fitzpatrick RC 2005. The vestibular system. *Current Biology,* 15 (15): R583-R586.

Dembo M, Radovčić D, Garvin HM, Laird MF, Schroeder L, Scott JE, Brophy J, Ackermann RR, Musiba CM, de Ruiter DJ, Mooers AØ & Collard M 2016. The evolutionary relationships and age of *Homo naledi*: An assessment using dated Bayesian phylogenetic methods. *Journal of Human Evolution,* 97: 17-26.

Dines JP, Mesnick SL, Ralls K, May-Collado L, Agnarsson I & Dean MD 2015. A trade-off between precopulatory and postcopulatory trait investment in male cetaceans. *Evolution,* 69 (6): 1560-1572.

Eastman JM, Alfaro ME, Joyce P, Hipp AL & Harmon LJ 2011. A novel comparative method for identifying shifts in the rate of character evolution on trees. *Evolution,* 65 (12): 3578-89.

Eldredge N & Gould SJ 1972. Punctuated equilibria: An alternative to phyletic gradualism. *In:* Schopf TJM (ed.) *Models in Paleobiology.* San Francisco: Freeman, Cooper and Company.

Faurby S & Svenning J-C 2015. A species-level phylogeny of all extant and late Quaternary extinct mammals using a novel heuristic-hierarchical Bayesian approach. *Molecular Phylogenetics and Evolution,* 84: 14-26.

Felsenstein J 1973. Maximum-likelihood estimation of evolutionary trees from continuous characters. *American journal of human genetics,* 25 (5): 471.

Felsenstein J 1985. Phylogenies and the comparative method. *The American Naturalist,* 125 (1): 1-15.

Finarelli JA & Flynn JJ 2006. Ancestral state reconstruction of body size in the Caniformia (Carnivora, Mammalia): the effects of incorporating data from the fossil record. *Systematic Biology,* 55 (2): 301-313.

Finarelli JA & Goswami A 2013. Potential pitfalls of reconstructing deep time evolutionary history with only extant data, a case study using the Canidae (Mammalia, Carnivora). *Evolution,* 67 (12): 3678-3685.

Fritz SA, Bininda-Emonds ORP & Purvis A 2009. Geographical variation in predictors of mammalian extinction risk: big is bad, but only in the tropics. *Ecology letters,* 12 (6): 538-549.

Gatesy J, Geisler JH, Chang J, Buell C, Berta A, Meredith RW, Springer MS & McGowen MR 2013. A phylogenetic blueprint for a modern whale. *Molecular Phylogenetics and Evolution,* 66 (2): 479-506.

Gavryushkina A, Heath TA, Ksepka DT, Stadler T, Welch D & Drummond AJ 2016. Bayesian total evidence dating reveals the recent crown radiation of penguins. *Systematic Biology,* [Epub ahead of print].

Geisler JH, McGowen MR, Yang G & Gatesy J 2011. A supermatrix analysis of genomic, morphological, and paleontological data from crown Cetacea. *BMC Evolutionary Biology,* 11: 112.

Gould SJ 1988. Trends as changes in variance: A new slant on progress and directionality in evolution. *Journal of Paleontology,* 62 (3): 319-329.

Gould SJ 1997. Cope's rule as psychological artefact. *Nature,* 385: 199-200.

Hall M & Ross C 2007. Eye shape and activity pattern in birds. *Journal of Zoology,* 271 (4): 437-444.

Hall MI 2008. Comparative analysis of the size and shape of the lizard eye. *Zoology,* 111 (1): 62-75.

Hansen TF 1997. Stabilizing selection and the comparative analysis of adaptation. *Evolution,* 51 (5): 1341-1351.

Harcourt-Smith WE 2015. The origins of bipedal locomotion. *In:* Henke W & Tattersall I (eds.), *Handbook of paleoanthropology.* 2nd ed. Berlin: Springer-Verlag.

Harmon LJ, Losos JB, Jonathan Davies T, Gillespie RG, Gittleman JL, Bryan Jennings W, Kozak KH, McPeek MA, Moreno-Roark F, Near TJ, Purvis A, Ricklefs RE, Schluter D, Schulte Ii JA, Seehausen O, Sidlauskas BL, Torres-Carvajal O, Weir JT & Mooers AO 2010. Early bursts of body size and shape evolution are rare in comparative data. *Evolution,* 64 (8): 2385-96.

Harvey PH & Pagel M 1991. *The comparative method in evolutionary biology,* Oxford, Oxford University Press.

Hayward A & Gillooly JF 2011. The cost of sex: Quantifying energetic investment in gamete production by males and females. *PLoS ONE,* 6 (1): e16557.

Hedges SB, Dudley J & Kumar S 2006. TimeTree: a public knowledge-base of divergence times among organisms. *Bioinformatics,* 22 (23): 2971-2972.

Hedges SB, Marin J, Suleski M, Paymer M & Kumar S 2015. Tree of life reveals clock-like speciation and diversification. *Molecular biology and evolution,* 32 (4): 835-845.

Hu Y, Meng J, Li C & Wang Y 2010. New basal eutherian mammal from the Early Cretaceous Jehol biota, Liaoning, China. *Proceedings of the Royal Society B: Biological Sciences,* 277 (1679): 229-236.

Hublin J-J 2015. Paleoanthropology: How old is the oldest human? *Current Biology,* 25 (11): R453-R455.

Hublin J-J, Neubauer S & Gunz P 2015. Brain ontogeny and life history in Pleistocene hominins. *Philosophical Transactions of the Royal Society B: Biological Sciences,* 370 (1663).

Ji Q, Luo Z-X, Yuan C-X, Wible JR, Zhang J-P & Georgi JA 2002. The earliest known eutherian mammal. *Nature,* 416 (6883): 816-822.

Jones KE, Bielby J, Cardillo M, Fritz SA, O'Dell J, Orme CDL, Safi K, Sechrest W, Boakes EH, Carbone C, Connolly C, Cutts MJ, Foster JK, Grenyer R, Habib M, Plaster CA, Price SA, Rigby EA, Rist J, Teacher A, Bininda-Emonds ORP, Gittleman JL, Mace GM, Purvis A & Michener WK 2009. PanTHERIA: a species-level database of life history, ecology, and geography of extant and recently extinct mammals. *Ecology,* 90 (9): 2648.

Kenagy G & Trombulak SC 1986. Size and function of mammalian testes in relation to body size. *Journal of Mammalogy,* 67 (1): 1-22.

Kirk EC & Kay RF 2004. The evolution of high visual acuity in the Anthropoidea. *In:* Ross C & Kay RF (eds.), *Anthropoid Origins.* New York: Kluwer Academic/Plenum Publishers.

Kratsch C & McHardy AC 2014. RidgeRace: ridge regression for continuous ancestral character estimation on phylogenetic trees. *Bioinformatics,* 30 (17): i527-i533.

Ksepka DT, Bertelli S & Giannini NP 2006. The phylogeny of the living and fossil Sphenisciformes (penguins). *Cladistics,* 22 (5): 412-441.

Kutsukake N & Innan H 2013. Simulation-based likelihood approach for evolutionary models of phenotypic traits on phylogeny. *Evolution,* 67 (2): 355-367.

Kutsukake N & Innan H 2014. Detecting phenotypic selection by Approximate Bayesian Computation in phylogenetic comparative methods. *In:* Garamszegi LZ (ed.) *Modern Phylogenetic Comparative Methods and Their Application in Evolutionary Biology.* Berlin: Springer-Verlag.

Lande R 1976. Natural Selection and Random Genetic Drift in Phenotypic Evolution. *Evolution,* 30 (2): 314-334.

Landis MJ, Schraiber JG & Liang M 2013. Phylogenetic analysis using Lévy processes: finding jumps in the evolution of continuous traits. *Systematic Biology,* 62 (2): 193-204.

Lee MS, Cau A, Naish D & Dyke GJ 2014. Morphological clocks in paleontology, and a mid-Cretaceous origin of crown Aves. *Systematic biology,* 63 (3): 442-449.

Luo Z-X, Yuan C-X, Meng Q-J & Ji Q 2011. A Jurassic eutherian mammal and divergence of marsupials and placentals. *Nature,* 476 (7361): 442-445.

MacFadden BJ 1986. Fossil horses from "*Eohippus*" (*Hyracotherium*) to *Equus*: Scaling, Cope's law and the evolution of body size. *Paleobiology,* 12 (4).

MacLeod CD 2010. The relationship between body mass and relative investment in testes mass in cetaceans: Implications for inferring interspecific variations in the extent of sperm competition. *Marine Mammal Science,* 26 (2): 370-380.

MacLeod CD 2014. Exploring and explaining complex allometric relationships: A case study on amniote testes mass allometry. *Systems,* 2 (3): 379-392.

MacLeod CD & MacLeod R 2009. The relationship between body mass and relative investment in testes mass in amniotes and other vertebrates. *Oikos,* 118 (6): 903-916.

Martins EP & Hansen TF 1997. Phylogenies and the comparative method: a general approach to incorporating phylogenetic information into the analysis of interspecific data. *The American Naturalist,* 149 (4): 646-667.

Marx FG & Fordyce RE 2015. Baleen boom and bust: a synthesis of mysticete phylogeny, diversity and disparity. *Royal Society open science,* 2 (4).

McShea DW 1994. Mechanisms of large-scale evolutionary trends. *Evolution,* 48 (6): 1747-1763.

McShea DW 1998. Possible Largest-Scale Trends in Organismal Evolution: Eight "Live Hypotheses". *Annual Review of Ecology and Systematics,* 29: 293-318.

Møller AP 1988. Ejaculate quality, testes size and sperm competition in primates. *Journal of Human Evolution,* 17 (5): 479-488.

Monroe MJ & Bokma F 2010. Little evidence for Cope's rule from Bayesian phylogenetic analysis of extant mammals. *Journal of evolutionary biology,* 23 (9): 2017-21.

Montgomery SH, Geisler JH, McGowen MR, Fox C, Marino L & Gatesy J 2013. The evolutionary history of cetacean brain and body size. *Evolution,* 67 (11): 3339-3353.

Mooers AØ, Vamosi SM & Schluter D 1999. Using phylogenies to test macroevolutionary hypotheses of trait evolution in cranes (Gruinae). *The American Naturalist,* 154 (2): 249-259.

Murrell B, Wertheim JO, Moola S, Weighill T, Scheffler K & Kosakovsky Pond SL 2012. Detecting individual sites subject to episodic diversifying selection. *PLoS Genetics,* 8 (7): e1002764.

Nadeau NJ & Jiggins CD 2010. A golden age for evolutionary genetics? Genomic studies of adaptation in natural populations. *Trends in Genetics,* 26 (11): 484-492.

Nikaido M, Ota T, Hirata T, Suzuki H, Satta Y, Aibara M, Mzighani SI, Sturmbauer C, Hagino-Yamagishi K & Okada N 2014. Multiple episodic evolution events in V1R receptor genes of East-African cichlids. *Genome Biology and Evolution,* 6 (5): 1135-1144.

Novacek MJ 1999. 100 million years of land vertebrate evolution: the Cretaceous-early Tertiary transition. *Annals of the Missouri Botanical Garden,* 86 (2): 230-258.

O'Leary MA, Bloch JI, Flynn JJ, Gaudin TJ, Giallombardo A, Giannini NP, Goldberg SL, Kraatz BP, Luo ZX, Meng J, Ni X, Novacek MJ, Perini FA, Randall ZS, Rougier GW, Sargis EJ, Silcox MT, Simmons NB, Spaulding M, Velazco PM, Weksler M, Wible JR & Cirranello AL 2013. The placental mammal ancestor and the post-K-Pg radiation of placentals. *Science,* 339 (6120): 662-7.

Organ C, Nunn CL, Machanda Z & Wrangham RW 2011. Phylogenetic rate shifts in feeding time during the evolution of *Homo*. *Proceedings of the National Academy of Sciences USA,* 108 (35): 14555-14559.

Pagel M 1997. Inferring evolutionary processes from phylogenies. *Zoologica Scripta,* 26 (4): 331-348.

Pagel M 1999. Inferring the historical patterns of biological evolution. *Nature,* 401: 877-884.

Pampush JD 2015. Selection played a role in the evolution of the human chin. *Journal of Human Evolution,* 82: 127-136.

Pant SR, Goswami A & Finarelli JA 2014. Complex body size trends in the evolution of sloths (Xenarthra: Pilosa). *BMC Evolutionary Biology,* 14 (1): 184.

Parker G, Ball M, Stockley P & Gage M 1997. Sperm competition games: a prospective analysis of risk assessment. *Proceedings of the Royal Society of London B: Biological Sciences,* 264 (1389): 1793-1802.

Peters RH 1986. *The ecological implications of body size,* UK, Cambridge University Press.

Pitnick S, Jones KE & Wilkinson GS 2006. Mating system and brain size in bats. *Proceedings of the Royal Society of London B: Biological Sciences,* 273 (1587): 719-724.

Puttick MN, Thomas GH & Benton MJ 2014. High rates of evolution preceded the origin of birds. *Evolution,* 68 (5): 1497-1510.

Pyron RA 2011. Divergence time estimation using fossils as terminal taxa and the origins of Lissamphibia. *Systematic Biology,* 60 (4): 466-481.

Rabosky DL 2014. Automatic detection of key innovations, rate shifts, and diversity-dependence on phylogenetic trees. *PLoS One,* 9 (2): e89543.

Rabosky DL & Adams DC 2012. Rates of morphological evolution are correlated with species richness in salamanders. *Evolution,* 66 (6): 1807-1818.

Rabosky DL, Santini F, Eastman J, Smith SA, Sidlauskas B, Chang J & Alfaro ME 2013. Rates of speciation and morphological evolution are correlated across the largest vertebrate radiation. *Nature Communications,* 4.

Raia P, Carotenuto F, Passaro F, Fulgione D & Fortelius M 2012. Ecological specialization in fossil mammals explains Cope's rule. *The American Naturalist,* 179 (3): 328-337.

Revell LJ, Mahler DL, Peres-Neto PR & Redelings BD 2012. A new phylogenetic method for identifying exceptional phenotypic diversification. *Evolution,* 66 (1): 135-146.

Ronquist F, Klopfstein S, Vilhelmsen L, Schulmeister S, Murray DL & Rasnitsyn AP 2012. A total-evidence approach to dating with fossils, applied to the early radiation of the Hymenoptera. *Systematic Biology,* 61 (6): 973-999.

Schmidt-Nielsen K 1984. *Scaling: Why is Animal Size so Important?,* Cambridge, Cambridge University Press.

Schmitz L & Motani R 2010. Morphological differences between the eyeballs of nocturnal and diurnal amniotes revisited from optical perspectives of visual environments. *Vision research,* 50 (10): 936-946.

Schmitz L & Wainwright PC 2011. Nocturnality constrains morphological and functional diversity in the eyes of reef fishes. *BMC evolutionary biology,* 11 (1): 1.

Schroeder L, Roseman CC, Cheverud JM & Ackermann RR 2014. Characterizing the evolutionary path(s) to early *Homo*. *PLoS ONE,* 9 (12): e114307.

Simmons LW & Emlen DJ 2006. Evolutionary trade-off between weapons and testes. *Proceedings of the National Academy of Sciences,* 103 (44): 16346-16351.

Simpson GG 1944. *Tempo and mode in evolution,* London, GB, Columbia University Press.

Simpson GG 1953. *The major features of evolution,* London, GB, Columbia University Press.

Slater GJ 2015. Iterative adaptive radiations of fossil canids show no evidence for diversity-dependent trait evolution. *Proceedings of the National Academy of Sciences USA,* 112 (16): 4897-4902.

Slater GJ, Harmon LJ & Alfaro ME 2012. Integrating fossils with molecular phylogenies improves inference of trait evolution. *Evolution,* 66 (12): 3931-3944.

Smith FA, Boyer AG, Brown JH, Costa DP, Dayan T, Ernest SKM, Evants AR, Fortelius M, Gittleman JL, Hamilton MJ, Harding LE, Lintualaakso K, Lyons SK, McCain C, Okie JG, Saarinen JJ, Sibly RM, Stephens PR, Theodor J & Uhen MD 2010. The evolution of maximum body size of terrestrial mammals. *Science,* 330: 1216-1219.

Smith MD, Wertheim JO, Weaver S, Murrell B, Scheffler K & Pond SLK 2015. Less is more: an adaptive branch-site random effects model for efficient detection of episodic diversifying selection. *Molecular biology and evolution,* 32 (5): 1342-1353.

Spoor F, Wood B & Zonneveld F 1994. Implications of early hominid labyrinthine morphology for evolution of human bipedal locomotion. *Nature,* 369 (6482): 645-648.

Spoor F, Wood B & Zonneveld F 1996. Evidence for a link between human semicircular canal size and bipedal behaviour. *Journal of Human Evolution,* 30 (2): 183-187.

Stanley SM 1973. An explanation for Cope's rule. *Evolution,* 27 (1): 1-26.

Steeman ME, Hebsgaard MB, Fordyce RE, Ho SY, Rabosky DL, Nielsen R, Rahbek C, Glenner H, Sørensen MV & Willerslev E 2009. Radiation of extant cetaceans driven by restructuring of the oceans. *Systematic Biology,* 58 (6): 573-585.

Thomas GH & Freckleton RP 2012. MOTMOT: Models of trait macroevolution on trees. *Methods in Ecology and Evolution,* 3 (1): 145-151.

Uyeda JC & Harmon LJ 2014. A Novel Bayesian Method for Inferring and Interpreting the Dynamics of Adaptive Landscapes from Phylogenetic Comparative Data. *Systematic Biology,* 63 (6): 902-918.

Van Valkenburgh B, Wang X & Damuth J 2004. Cope's rule, hypercarnivory, and extinction in North American canids. *Science,* 306 (5693): 101-104.

Venditti C, Meade A & Pagel M 2011. Multiple routes to mammalian diversity. *Nature,* 479 (7373): 393-396.

Wagner PJ 1996. Contrasting the underlying patterns of active trends in morphologic evolution. *Evolution,* 50 (3): 990-1007.

Warren DL & Iglesias TL 2012. No evidence for the 'expensive-tissue hypothesis' from an intraspecific study in a highly variable species. *Journal of Evolutionary Biology,* 25 (6): 1226-1231.

White TD, Asfaw B, Beyene Y, Haile-Selassie Y, Lovejoy CO, Suwa G & WoldeGabriel G 2009. *Ardipithecus ramidus* and the paleobiology of early hominids. *Science,* 326 (5949): 64-86.

Wood B 2002. Palaeoanthropology: Hominid revelations from Chad. *Nature,* 418 (6894): 133-135.

Wood HM, Matzke NJ, Gillespie RG & Griswold CE 2012. Treating fossils as terminal taxa in divergence time estimation reveals ancient vicariance patterns in the palpimanoid spiders. *Systematic biology,* 62 (2): 264-284.

Wrangham R 2009. *Catching fire: How cooking made us human,* New York, Basic Books.

Yang Z 2002. Inference of selection from multiple species alignments. *Current Opinion in Genetics & Development,* 12 (6): 688-694.

Yang Z 2006. *Computational molecular evolution,* Oxford, Great Britain, Oxford University Press.

Chapter 1

Adaptive evolution towards larger size in mammals

(Published as: Baker J, Meade A, Pagel M & Venditti C 2015. *Adaptive evolution toward larger size in mammals.* Proceedings of the National Academy of Sciences USA, 112 (16): 5093-98.)

Abstract

The notion that large body size confers some intrinsic advantage to biological species has been debated for centuries. Using a phylogenetic statistical approach that allows the rate of body size evolution to vary across a phylogeny, we find a long-term directional bias towards increasing size in the mammals. This pattern holds separately in ten of eleven orders for which sufficient data are available and arises from a tendency for accelerated rates of evolution to produce increases, but not decreases, in size. On a branch-by-branch basis, increases in body size have been more than twice as likely as decreases, yielding what amounts to millions and millions of years of rapid and repeated increases in size away from the small ancestral mammal. These results are the first evidence, to our knowledge, from extant species that are compatible with Cope's rule: the pattern of body size increase through time observed in the mammalian fossil record. We show that this pattern is unlikely to be explained by several non-adaptive mechanisms for increasing size, and most likely represents repeated responses to new selective circumstances. By demonstrating that it is possible to uncover ancient evolutionary trends from a combination of a phylogeny and appropriate statistical models, we illustrate how data from extant species can complement paleontological accounts of evolutionary history, opening up new avenues of investigation for both.

Introduction

The idea that large size confers some intrinsic advantage has lingered in the psyche of biologists for centuries. Researchers have proposed that bigger body sizes can increase tolerance to environmental extremes (Peters, 1986), reduce mortality (Brown and Sibly, 2006) and enhance predation success (Hone and Benton, 2005), among other advantages. In support of these conjectures, analyses from a range of different taxonomic groups demonstrate that within populations larger individuals have significantly enhanced survival, fecundity and mating success (Kingsolver and Pfennig, 2004, 2007). If these advantages are general and have played out over long timescales they could explain the existence of Cope's rule (Cope, 1896): a broad trend towards increasing size through time (Hunt *et al.*, 2010; Kingsolver and Pfennig, 2004, 2007).

Mammals evolved from a relatively small common ancestor over 165 Ma (Fritz *et al.*, 2009; Ji *et al.*, 2002; O'Leary *et al.*, 2013) and went on to form one of the largest and most successful vertebrate radiations in Earth's history. Mammals vary greatly in size, spanning almost eight orders of magnitude. This variation implies that some groups have experienced much greater evolutionary change in size from the ancestral form than others. Indeed, the mammalian fossil record provides the clearest evidence in support of Cope's rule over long evolutionary timescales (Alroy, 1998; Cope, 1896; Van Valkenburgh *et al.*, 2004).

Despite the paleontological support, evidence for Cope's rule remains elusive from studies of extant data alone (Knouft and Page, 2003; Moen, 2006; Pianka, 1995), including studies of the mammals (Monroe and Bokma, 2010). A possible reason for the discrepancy between paleontological and extant data might be that conventional comparative methods for studying trends within extant data implicitly assume homogenous evolutionary patterns and processes. When these assumptions are violated, it renders the homogenous modelling approach incomplete at best and at worst, a source of potential bias in the study of historical evolutionary change; for example, reconstructions of

probable ancestral values can be biased towards average or intermediate values (Elliot and Mooers; Pagel, 1999) which would thereby mask long-term evolutionary trends that are apparent from the fossil record.

Previously, we have shown that rates of body size evolution in the mammals routinely violate the assumption of homogeneity (Venditti *et al.*, 2011) but how these rate changes might be related to size itself has not been studied. If changes towards larger size in the mammals have consistently occurred at rates that differ from changes to smaller size then reconstructed ancestral states accounting for these rate differences may track more closely the observed fossil record. Such a pattern would allow the detection of size-related evolutionary trends from extant data (Figure A1.1).

Here we apply a statistical phylogenetic approach for reconstructing mammalian evolutionary history that allows the rate of evolution to vary throughout a phylogenetic tree without prior knowledge or specification of where and when rate-shifts occurred. We use this method to test for size-related biases in rates of morphological change and ask whether accounting for any such bias allows us to predict a generalised pattern of size increase in the mammals in line with the generalised pattern of size increase observed in the fossil record. Finally, we consider whether a size-related bias in the rate of morphological evolution can help to choose among the several macro-evolutionary processes that have been suggested to give rise to Cope's rule.

Results and Discussion

Because the rate of morphological evolution has varied considerably among mammals throughout their history, branch lengths measured in time can overestimate or underestimate the amount of change expected under a homogeneous Brownian motion model (Venditti *et al.*, 2011). We therefore scale time by an amount reflecting the rate of morphological evolution along individual branches of the mammalian phylogeny (Venditti *et al.*, 2011, Appendix 1). Longer rate-scaled branches have experienced more change than

would be expected given their length in time (Methods). If body size increase has been disproportionally favoured, we expect to find that longer rate-scaled branches are linked to larger increases in size throughout the phylogeny. If this pattern has been repeated over many branches, we expect to find them associated with a long-term historical trend towards increasing size (Cope, 1896).

Across all mammals we find a significant positive relationship between *path-wise rates* (the sum of all rate-scaled branches along the evolutionary path leading to individual species, Methods) and body size [likelihood ratio (D) test, compared to a homogenous Brownian motion model: D = 359.85, P < 0.001, *df* = 2, Figure 1A; this relationship holds in all of 500 randomly selected trees from the posterior distribution of rate-scaled phylogenies]. Allowing the slope of the relationship between size and path-wise rate to vary among orders (separate-slopes model, Figure 1B) significantly improves on the model relying on a single common slope (D = 252.24, P < 0.001, *df* = 31; this relationship also holds in all of 500 randomly selected trees from the posterior distribution of rate-scaled phylogenies), and reveals that the positive relationship is maintained separately within 10 of 11 mammalian orders (Figure 1B, Table A1.1): the only exception is the marsupial order Diprotodontia, where the path-wise rate is largest in the evolutionary paths leading to smaller species (Figure 1B).

We visualize the importance of detecting variation in the rate of evolution by simulating body sizes from the separate-slopes regression model (Methods) and from a conventional homogenous Brownian motion model assuming a single uniform rate of change. The separate-slopes model simulates values that symmetrically bracket the observed body size distribution (Figure 2). By comparison, the homogeneous model systematically overestimates small sizes and underestimates large sizes (Figure 2, Inset). This poor fit to the real data arises by virtue of the homogenous model missing the historical bias towards rapid rates leading to larger size.

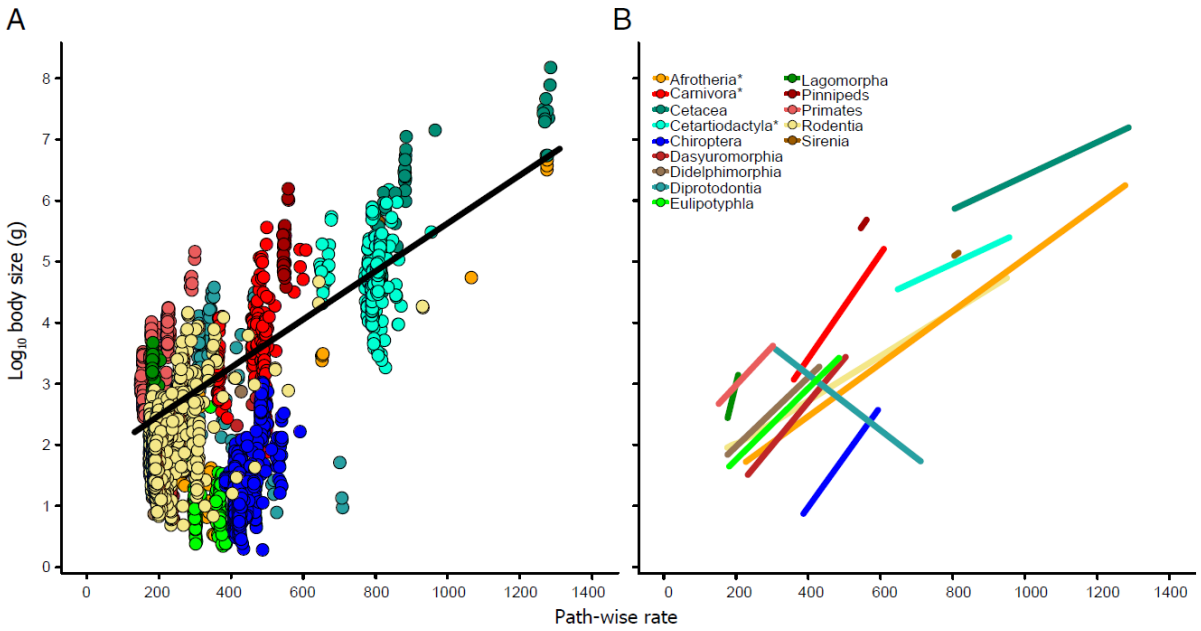


Figure 1: Faster path-wise rates have led to larger body size in mammals. (A) Relationship across all mammals is plotted – data points are coloured by order (n = 3,321). The black line is the fitted phylogenetic slope of the relationship between body size and path-wise rates (Methods) across all mammals. (B) Fitted phylogenetic slopes of the relationship within each of the 11 mammalian orders investigated here. Orders that contain aquatic groups are indicated by an asterisk; for these orders, only the terrestrial members are plotted. Aquatic groups are plotted separately (Cetacea, pinnipeds and Sirenia).

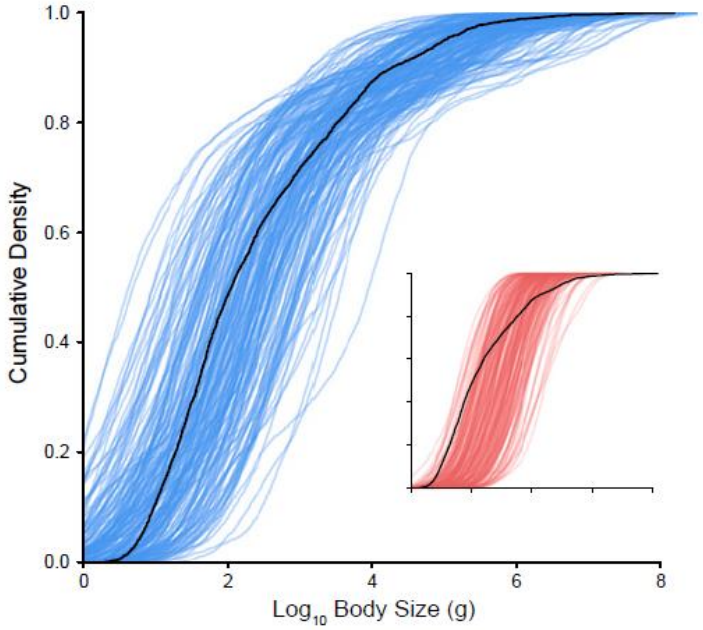


Figure 2: Comparisons between the cumulative distribution of observed mammalian body sizes (n = 3,321, black lines) and simulated data (n = 1,000, coloured lines). The real data are compared with simulations generated from our separate-slopes regression model (blue lines) and a conventional homogenous Brownian motion model (red lines, Inset).

Using our separate-slopes model we infer the ancestral body size at each internal node of the mammalian phylogeny (Methods, Figure 3A*,* Figure A1.2). The tendency for body size increase can be studied quantitatively by finding the difference in body size from the start to the end of each branch of the phylogeny (n = 5,233). We term these differences *phylogenetic ancestor-descendant* (PAD) comparisons to contrast with the paleontological approach where *fossil ancestor-descendant* (FAD) comparisons (Alroy, 1998; Alroy, 2000; Raia *et al.*, 2012) are made between the sizes of taxonomically paired species found in the fossil record.

Our PAD comparisons demonstrate that not only are size increases more common but they also tend to be greater in magnitude and occur at a faster rate compared with body size decreases (Figure 3). Of the 5,233 PAD comparisons, 3,496 or 66.8% showed an increase in size (exact binomial test, P < 0.001). On average, descendant species are 0.10±0.004 $\log_{10}$ units, or 6±0.25% larger than their ancestors (Figure 3B), although this figure varies between 1.4 and16.9% in individual orders (-3.8±0.63% in Diprotodontia*,* Table A1.2). These figures compare favourably to results from paleontological data, where North American Cenozoic mammals are, on average, 9% larger than their ancestors (Alroy, 1998).

We find that on a branch-by-branch basis the largest increases in size are associated with the fastest rates of evolution (β = 0.015, P < 0.001, Figure 3C). One argument for such a pattern is based on the premise that phyletic increases in size arise simply as a consequence of evolutionary divergence away from a small ancestral value, where there is some lower physiological limit on size (McShea, 1994; Stanley, 1973). In this scenario, a taxon's "*maximal potential adaptive zone*" (Stanley, 1973) is always skewed such that larger species will evolve and that those species will be specialized (Cope, 1896; Stanley, 1973).

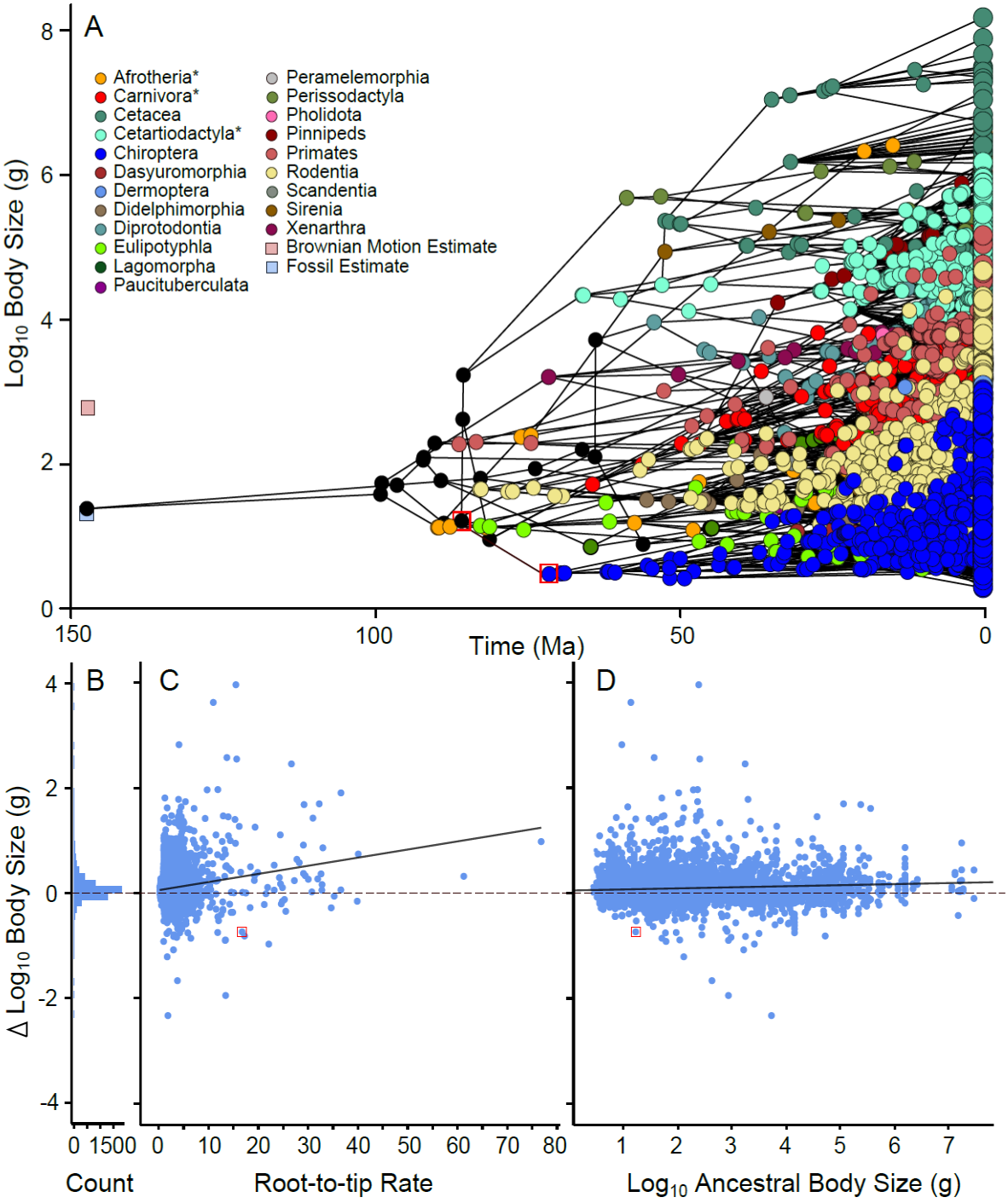


Figure 3: PAD comparisons and reconstructed ancestral sizes. (A) Projection of ancestral state reconstructions into a phylomorphospace (n = 5,234 including all tips and internal nodes). Points are connected by phylogeny and each internal node of the tree has been reconstructed using the parameters of our separate-slopes regression model. Our estimate for the therian root (24.5g) falls within the ranges given by the paleontological data (20-25g, midpoint indicated by the pale blue square). This estimate is in contrast to the estimate made by a conventional homogenous Brownian motion model, which is more than an order of magnitude too large (pale pink square, 610.7g). **(Figure caption continues overleaf.)**

Figure 3 (cont.): Orders which contain aquatic groups are indicated by an asterisk; for these orders, only the terrestrial members are plotted. Aquatic groups are plotted separately (Cetacea, pinnipeds and Sirenia). (B-D) PAD changes (Δlog10 body size) across every branch of the mammalian phylogeny (n = 5,233). The red dashed line indicates no change in size. (B) Frequency (f) distribution of Δlog10 body size across individual branches. There is a significant bias towards body size increase (exact binomial test, P < 0.001). (C) Plot of the inferred rate of evolution along individual branches (Methods) against Δlog10 body size. The regression line is significantly positive (β= 0.015, P < 0.0001). (D) Ancestral body size against body size change across individual branches. The grey bars represent the SD of Δlog10 body size calculated from the variance associated with each data point (Methods, $\sigma^2_{\Delta \log_{10} \text{body size}}$). The regression line and the SDs in D have been corrected for the regression to the mean artifact (Methods, Appendix 1). The slope of the relationship between ancestral size and Δlog10 body size is significantly positive (β = 0.020, P = 0.0006). Highlighted by a red square on each of these plots is the branch leading to modern bats.

We use our PAD comparisons to test for the presence of a lower bound by drawing on ideas developed in the paleontological literature (Alroy, 1998; Alroy, 2000; McShea, 1994; Wagner, 1996) whilst explicitly accounting for shared ancestry. If some lower boundary on size is enforced, we expect most ancestor-descendant size changes to be positive when the ancestral size is near to that limit; it is only possible to get larger. However, as the ancestral state moves away from this limit, we predict that the distribution of body size change will become increasingly centred about zero i.e. size decreases are equally likely as size increases (Wagner, 1996). Taken over all branches of the phylogeny, this pattern predicts a negative relationship between a branch's ancestral size and the average body size change observed along that branch (Alroy, 1998; Alroy, 2000). When ancestral size is small changes will tend to be positive, but when ancestral size is large size can change in either direction.

We do not find the predicted negative relationship (Figure 3D*;* Appendix 1). Instead, we find that size change actually slightly increases in magnitude when ancestral size is larger (β = 0.020, P < 0.001, Figure 3D). This pattern is also found in the paleontological data using FAD comparisons (Alroy, 1998). To retain the idea that some physiological lower limit could produce these PAD changes and results from paleontological data (Alroy, 1998), proponents would have to invoke a new physiological lower limit for each new

species that comes into existence. Why, or according to what processes these mysterious and dynamically shifting constraints arise impose a steep hill for this explanation to climb.

The notion that 'adaptive zones' litter the morphological landscape has often been wielded as a driver for large-scale macro-evolutionary patterns (Hansen, 1997; Hunt, 2012; Simpson, 1953; Slater, 2013; Uyeda *et al.*, 2011). With this view, one might expect fast evolutionary rates to be the result of shifts from one zone to another or in the position of the adaptive peak through time (Arnold, 2014; Estes and Arnold, 2007; Hansen, 1997; Simpson, 1953). If the occupation of new adaptive zones is constantly associated with changes toward large size or there is some sort of continuously moving optima, such that large size is favoured, this view would be consistent with the pattern we observe here, although there is nothing in the patterns we observe that requires the existence of discrete adaptive zones.

It has been suggested that large-bodied species may have an inherently faster rate of evolution owing to the relaxation of some unspecified size-linked constraint (Simpson, 1953; Stanley, 1979) (e.g. genetic, developmental, biomechanical). If such constraints were operating we would expect to observe that larger-bodied species change disproportionately more along the branches of the phylogeny than smaller-bodied ones, leading to the prediction that the variance of body size change should be positively correlated with ancestral size: small-bodied species change less than large-bodied ones. We calculated the variance for all PAD comparisons ($\sigma^2_{\Delta log_{10}\ body\ size}$, n = 5,233), after adjusting for the regression to the mean artifact (Appendix 1, Alroy, 1998; Kelly and Price, 2005). We then regressed log-transformed $\sigma^2_{\Delta \log 10\ body\ size}$ onto log-transformed ancestral size (i.e. size reconstructed at the start of a branch) across all branches of the phylogeny. We do not find the expected positive relationship (β = -0.017, t = -1.47, P = 0.14, see Figure 3D). Therefore, and in agreement with previous work (Cooper and Purvis, 2009; Smith *et al.*, 2004), we see no reason to invoke the release of constraints as a force driving rate variation or size changes in mammals.

A possible difficulty for our model is that it predicts that mammals will become increasingly and indefinitely larger over long periods of time even though there must be some physical limit on the maximum size a terrestrial vertebrate can attain. Usefully, it seems that mammals have not reached those limits: even the largest ever-known terrestrial mammals (Alexander, 1998; Smith *et al.*, 2010) fall well below the proposed maximum masses for terrestrial animals of between 20,000 to 1 million kg (Economos, 1981; Hokkanen, 1986). If extant mammals had reached their maxima, it would be reflected in additional parameters (quadratic effects) in our model that would account for a slowing of the trajectory towards increasing size; however, at least for now, quadratic models are not necessary (Appendix 1).

A second difficulty is that if large body size is continuously favoured one would expect that there must have come a point at which it was advantageous for species to become small, exploiting niches made available by continued size increases in competing taxa. In fact, size reduction was common in the evolutionary history of mammals (1,737 of our PAD comparisons or 33.2%) and often occurred at rapid rates (Evans *et al.*, 2012, Figure 3C). For example, there was rapid evolutionary change towards body size decrease in the branch leading to extant bats, although subsequent evolution within this group returned to a general pattern of body size increase (Figure 3, Figure 1B). In the special case of Diprotodontia, it appears that rapid changes resulting in smaller size dominated, although we do still observe some large body size increases in this group. A possible explanation for this pattern is that these species might have become smaller in response to nutrient-poor environments in Australian habitats (Milewski and Diamond, 2000; Orians and Milewski, 2007).

The consistent signal for directional evolutionary change in size implies a relatively small common ancestor of mammals. Previously, ancestral state reconstruction in the face of such a trend has been problematic; conventional comparative methods make it impossible to detect evolutionary trends using extant data. Incorporating fossils into a phylogeny improves ancestral state estimates (Finarelli and Flynn, 2006; Finarelli and Goswami, 2013;

Oakley and Cunningham, 2000; Slater *et al.*, 2012); however, here we test a long-posited suggestion that it is possible to infer from extant data alone the existence of ancient forms whose size or shape is not intermediate to the range of present diversity (Pagel, 1999). If our characterization of mammalian size evolution is a good description of the historical processes that led to contemporary mammal species, we should be able to infer ancestral states that are closer to those observed in the fossil record than estimates derived from conventional homogenous models without using fossil data. These expectations are borne out (Figure 3A). We estimate that the ancestral size at the root of therian mammals was 24.5g. This value falls within the fossil body size range (20-25g) of *Eomaia scansoria* (Ji *et al.*, 2002), which has been recently suggested to lie close to the root of all placental and marsupial mammals (O'Leary *et al.*, 2013). In contrast, the homogenous Brownian motion model reconstructs the ancestral body size to be greater than 600g, which is more than an order of magnitude too large.

It may be wrong to assume that fossil species are directly ancestral to extant groups (Foote, 1996). Accordingly, we reconstructed body sizes for 65 unique fossil taxa (Table A1.3) that represent the oldest or basal members of several mammal groups (Methods). Homogenous Brownian motion reconstructions of these taxa yield sizes that are systematically larger than paleontological estimates (t = 4.68, P < 0.001, Figure 4A). In contrast, the separate-slopes regression model reconstructs body sizes that do not differ significantly from paleontological estimates (t = 0.76, P = 0.45, Figure 4B).

Taken together, our results demonstrate that mammals have consistently evolved towards larger size, almost certainly reflecting an adaptive response to new selective circumstances such as competition (Benson *et al.*, 2014), climate changes (Hunt and Roy, 2006; Hunt *et al.*, 2010) or dietary specialization (Van Valkenburgh *et al.*, 2004). These results are not compatible with purely passive explanations for trends through time (Gould, 1988; Wagner, 1996). Instead, rapid and repeated instances of evolutionary change toward bigger body size have consistently shaped mammalian diversity, allowing mammal species to attain larger sizes over the millions and millions of years of their evolutionary history.

Our findings represent unique support for an adaptive explanation for Cope's rule, one of the most enduring and iconic notions in evolutionary biology . The ability to detect and characterize trends within extant taxa provides the attractive opportunity to study a broad number of taxonomic groups using the vast amounts of data available for extant species. Such analyses should be viewed as complementary to work based on fossil evidence which benefits from the ability to study morphology directly through time.

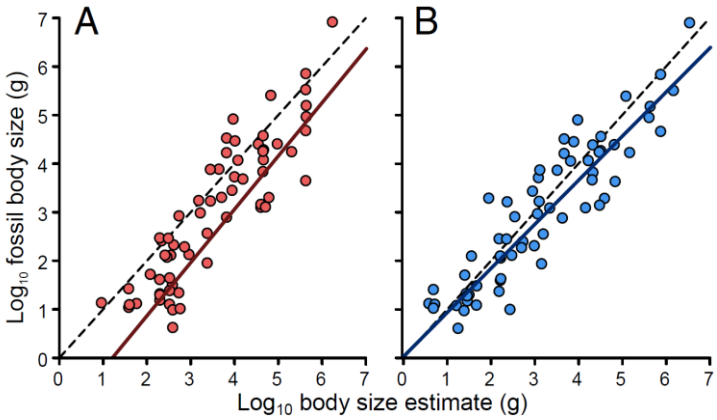


Figure 4: Comparisons of reconstructed body sizes with fossil estimates. The solid coloured lines in both plots are the predicted phylogenetic slopes from a regression model of fossil sizes as given in the paleontological literature against reconstructed values (n = 65). The dashed black lines indicate a one-to-one relationship which is the expected slope if models are predicting body sizes accurately. (A) Predicted body sizes from a homogenous Brownian motion model compared with fossil estimates. (B) Predicted body sizes from our separate-slopes model in comparison to the fossil record.

Methods

Data

We used a comprehensive time-scaled phylogenetic tree (Fritz *et al.*, 2009) of extant mammals (n = 3,321) along with body size data from two major databases (Ernest, 2003; Jones et al., 2009). Body sizes were log-transformed. Our analyses are based on the assumption that the tree of Fritz et al (2009) provides a relatively reliable estimate of mammalian phylogeny and divergence times. We classified the mammals into orders following Bininda-Emonds *et al.*, (2007). To measure the rate of body size evolution in our

mammal dataset, we apply a recently developed phylogenetic statistical approach that detects regions of the tree that have undergone especially fast or slow rates of change (Venditti *et al.*, 2011). Our approach stretches or compresses time-measured branch lengths by an amount reflecting the inferred rate of evolution in that branch (Venditti *et al.*, 2011, Appendix 1). Stretched branches reflect increased rates of change and compressed branches reflect regions where size has changed less than expected under background rates.

Detecting Trends

We use our rate-scaled branch lengths to study long-term trends. We sum all the rate-scaled branches along the evolutionary path of a species, leading from the root to the tip. These summed branches equate to *path-wise rates*, a measure of the total changes in rate a species has experienced during the course of its evolution. If elevated rates have been disproportionately associated with size increases, we expect to find that species with greater path-wise rates will be larger in size. To test this idea, we regressed log body mass onto path-wise rate using phylogenetic generalized least squares models (Freckleton *et al.*, 2002; Pagel, 1997) in a maximum-likelihood framework.

To test for different patterns among mammal orders we allowed the relationship between path-wise rate and body size to vary among those orders where sample size was large enough for analysis ($n \geq 40$, Appendix 1). Owing to the small sample sizes of orders within the monophyletic superorder Afrotheria, we study Afrotheria as a single group (Appendix 1). Because aquatic species may have different patterns and processes of body size evolution (Evans *et al.*, 2012; Schmidt-Nielsen, 1984) we allowed the magnitude of the relationship to vary for these groups (Pinnipeds, Sirenia and Cetacea). We compared nested models using the likelihood-ratio test statistic (D).

Reconstructing historical body sizes

We estimated ancestral body sizes at each node of the mammal phylogeny using a phylogenetic predictive modelling approach that incorporates the parameters of our

separate-slopes regression model (Franks *et al.*, 2012; Jetz and Freckleton, 2015; Organ *et al.*, 2007, Figure A1.2). We then tracked body size change and rates on a branch-by-branch basis across the entire phylogeny. We refer to these branchwise changes as PAD comparisons to contrast with the paleontological method of FAD comparisons (Appendix 1).

Using the same predictive modelling approach as for ancestral state reconstruction, we assessed how well our results could be reconciled with paleontological data by estimating the expected size of 65 unique fossil taxa, given their proposed phylogenetic position (Appendix 1, Figure A1.3). We compared these reconstructions with the paleontological estimates and reconstructions using conventional homogenous Brownian motion methods (Appendix 1).

Quantifying Constraints

We assessed whether our data fit the predictions made by the presence of a size-linked constraint using our PAD comparisons. If some unspecified constraint is acting to restrict evolutionary potential in smaller species, we would expect to see released pressure in larger species allowing for more evolutionary change; as ancestral size increases, we should observe an increase in variance around the observed change in body size ($\Delta \log_{10}$ body size). To test this, we calculated the variance in body size change ($\sigma^2_{\Delta \log 10 \text{ body size}}$) for all PAD comparisons (n = 5,233) across every branch of the phylogeny, after adjusting for the regression to the mean artifact (Appendix 1, Alroy, 1998; Kelly and Price). We assessed whether there was a significant increase in variance with increasing ancestral size by regressing $\sigma^2_{\Delta \log_{10} \text{body size}}$ onto log-transformed ancestral body mass. A visualization of the observed variance in body size change is shown in Figure 3D.

References

Alexander RM 1998. All-time giants: the largest animals and their problems. *Palaeontology,* 41 (6): 1231-1245.

Alroy J 1998. Cope's rule and the dynamics of body mass evolution in North American fossil mammals. *Science,* 280 (5364): 731-734.

Alroy J 2000. Understanding the dynamics of trends within evolving lineages. *Paleobiology,* 26 (3): 319-329.

Arnold SJ 2014. Phenotypic evolution: the ongoing synthesis. *The American Naturalist,* 183 (6): 729-746.

Benson RBJ, Frigot RA, Goswami A, Andres B & Butler RJ 2014. Competition and constraint drove Cope's rule in the evolution of giant flying reptiles. *Nature Communications,* 5 (3567).

Bininda-Emonds OR, Cardillo M, Jones KE, Macphee RD, Beck RM, Grenyer R, Price SA, Vos RA, Gittleman JL & Purvis A 2007. The delayed rise of present-day mammals. *Nature,* 446 (7135): 507-12.

Brown JH & Sibly RM 2006. Life-history evolution under a production constraint. *Proceedings of the National Academy of Sciences USA,* 103 (47): 17595-17599.

Cooper N & Purvis A 2009. What factors shape rates of phenotypic evolution? A comparative study of cranial morphology of four mammalian clades. *Journal of Evolutionary Biology,* 22: 1024-1035.

Cope ED 1896. *The primary factors of organic evolution,* Chicago, Open Court Publishing Company.

Economos AC 1981. The largest land mammal. *Journal of theoretical biology,* 89 (2): 211-214.

Elliot MG & Mooers AO 2014. Inferring ancestral states without assuming neutrality or gradualism using a stable model of continuous character evolution. *BMC Evolutionary Biology,* 14 (226).

Ernest SKM 2003. Life history characteristics of placental nonvolant mammals. *Ecology,* 84 (12): 3402.

Estes S & Arnold SJ 2007. Resolving the paradox of stasis: models with stabilizing selection explain evolutionary divergence on all timescales. *The American Naturalist,* 169 (2): 227-44.

Evans AR, Jones D, Boyer AG, Brown JH, Costa DP, Ernest SKM, Fitzgerald EMG, Fortelius M, Gittleman JL, Hamilton MJ, Harding LE, Lintulaakso K, Lyons SK, Okie JG, Saarinen JJ, Sibly RM, Smith FA, Stephens PR, Theodor JM & Uhen MD 2012. The maximum rate of mammal evolution. *Proceedings of the National Academy of Sciences USA,* 109 (11): 4187-4190.

Finarelli JA & Flynn JJ 2006. Ancestral state reconstruction of body size in the Caniformia (Carnivora, Mammalia): the effects of incorporating data from the fossil record. *Systematic Biology,* 55 (2): 301-313.

Finarelli JA & Goswami A 2013. Potential pitfalls of reconstructing deep time evolutionary history with only extant data, a case study using the Canidae (Mammalia, Carnivora). *Evolution,* 67 (12): 3678-3685.

Foote M 1996. On the probability of ancestors in the fossil record. *Paleobiology,* 22 (2): 141-151.

Franks PJ, Freckleton RP, Beaulieu JM, Leitch IJ & Beerling DJ 2012. Megacycles of atmospheric carbon dioxide concentration correlate with fossil plant genome size. *Philosophical Transactions of the Royal Society B: Biological Sciences,* 367 (1588): 556-64.

Freckleton RP, Harvey PH & Pagel M 2002. Phylogenetic analysis and comparative data: A test and review of evidence. *The American Naturalist,* 160 (6): 712-726.

Fritz SA, Bininda-Emonds ORP & Purvis A 2009. Geographical variation in predictors of mammalian extinction risk: big is bad, but only in the tropics. *Ecology letters,* 12 (6): 538-549.

Gould SJ 1988. Trends as changes in variance: A new slant on progress and directionality in evolution. *Journal of Paleontology,* 62 (3): 319-329.

Hansen TF 1997. Stabilizing selection and the comparative analysis of adaptation. *Evolution,* 51 (5): 1341-1351.

Hokkanen JEI 1986. The size of the largest land animal. *Journal of theoretical biology,* 118 (4): 491-499.

Hone DW & Benton MJ 2005. The evolution of large size: How does Cope's rule work? *Trends in Ecology & Evolution,* 20 (1): 4-6.

Hunt G 2012. Measuring rates of phenotypic evolution and the inseparability of tempo and mode. *Paleobiology,* 38 (3): 351-373.

Hunt G & Roy K 2006. Climate change, body size evolution, and Cope's rule in deep-sea ostracodes. *Proceedings of the National Academy of Sciences USA,* 103 (5): 1347-52.

Hunt G, Wicaksono SA, Brown JE & Macleod KG 2010. Climate-driven body-size trends in the ostracod fauna of the deep Indian Ocean. *Palaeontology,* 53 (6): 1255-1268.

Jetz W & Freckleton RP 2015. Towards a general framework for predicting threat status of data-deficient species from phylogenetic, spatial and environmental information. *Philosophical Transactions of the Royal Society B: Biological Sciences,* 370 (1662).

Ji Q, Luo Z-X, Yuan C-X, Wible JR, Zhang J-P & Georgi JA 2002. The earliest known eutherian mammal. *Nature,* 416 (6883): 816-822.

Jones KE, Bielby J, Cardillo M, Fritz SA, O'dell J, Orme CDL, Safi K, Sechrest W, Boakes EH, Carbone C, Connolly C, Cutts MJ, Foster JK, Grenyer R, Habib M, Plaster CA, Price SA, Rigby EA, Rist J, Teacher A, Bininda-Emonds ORP, Gittleman JL, Mace GM, Purvis A &

Michener WK 2009. PanTHERIA: a species-level database of life history, ecology, and geography of extant and recently extinct mammals. *Ecology,* 90 (9): 2648.

Kelly C & Price TD 2005. Correcting for regression to the mean in behavior and ecology. *The American Naturalist,* 166 (6): 700-707.

Kingsolver JG & Pfennig DW 2004. Individual-level selection as a cause of Cope's rule of phyletic size increase. *Evolution,* 58 (7): 1608-1612.

Kingsolver JG & Pfennig DW 2007. Patterns and power of phenotypic selection in nature. *Bioscience,* 57 (7): 561-572.

Knouft JH & Page LM 2003. The evolution of body sze in extant groups of North American freshwater fishes: Speciation, size distributions, and Cope's rule. *The American Naturalist,* 161 (3): 413-421.

Mcshea DW 1994. Mechanisms of large-scale evolutionary trends. *Evolution,* 48 (6): 1747-1763.

Milewski A & Diamond R 2000. Why are very large herbivores absent from Australia? A new theory of micronutrients. *Journal of biogeography,* 27 (4): 957-978.

Moen DS 2006. Cope's rule in cryptodiran turtles: Do the body sizes of extant species reflect a trend of phyletic size increase? *Journal of Evolutionary Biology,* 19 (4): 1210-1221.

Monroe MJ & Bokma F 2010. Little evidence for Cope's rule from Bayesian phylogenetic analysis of extant mammals. *Journal of evolutionary biology,* 23 (9): 2017-21.

O'leary MA, Bloch JI, Flynn JJ, Gaudin TJ, Giallombardo A, Giannini NP, Goldberg SL, Kraatz BP, Luo ZX, Meng J, Ni X, Novacek MJ, Perini FA, Randall ZS, Rougier GW, Sargis EJ, Silcox MT, Simmons NB, Spaulding M, Velazco PM, Weksler M, Wible JR & Cirranello AL 2013. The placental mammal ancestor and the post-K-Pg radiation of placentals. *Science,* 339 (6120): 662-7.

Oakley TH & Cunningham CW 2000. Independent contrasts succeed where ancestor reconstruction fails in a known bacteriophage phylogeny. *Evolution,* 54 (2): 397-405.

Organ CL, Shedlock AM, Meade A, Pagel M & Edwards SV 2007. Origin of avian genome size and structure in non-avian dinosaurs. *Nature,* 446 (7132): 180-4.

Orians GH & Milewski AV 2007. Ecology of Australia: the effects of nutrient-poor soils and intense fires. *Biological Reviews of the Cambridge Philosophical Society,* 82 (3): 393-423.

Pagel M 1997. Inferring evolutionary processes from phylogenies. *Zoologica Scripta,* 26 (4): 331-348.

Pagel M 1999. Inferring the historical patterns of biological evolution. *Nature,* 401: 877-884.

Peters RH 1986. *The ecological implications of body size,* UK, Cambridge University Press.

Pianka ER 1995. Evolution of body size: Varanid lizards as a model system. *The American Naturalist,* 146 (3): 398-414.

Raia P, Carotenuto F, Passaro F, Fulgione D & Fortelius M 2012. Ecological specialization in fossil mammals explains Cope's rule. *The American Naturalist,* 179 (3): 328-337.

Schmidt-Nielsen K 1984. *Scaling: Why is Animal Size so Important?,* Cambridge, Cambridge University Press.

Simpson GG 1953. *The major features of evolution,* London, GB, Columbia University Press.

Slater GJ 2013. Phylogenetic evidence for a shift in the mode of mammalian body size evolution at the Cretaceous-Palaeogene boundary. *Methods in Ecology and Evolution,* 4 (8): 734-744.

Slater GJ, Harmon LJ & Alfaro ME 2012. Integrating fossils with molecular phylogenies improves inference of trait evolution. *Evolution,* 66 (12): 3931-3944.

Smith FA, Boyer AG, Brown JH, Costa DP, Dayan T, Ernest SKM, Evants AR, Fortelius M, Gittleman JL, Hamilton MJ, Harding LE, Lintualaakso K, Lyons SK, Mccain C, Okie JG, Saarinen JJ, Sibly RM, Stephens PR, Theodor J & Uhen MD 2010. The evolution of maximum body size of terrestrial mammals. *Science,* 330: 1216-1219.

Smith FA, Brown JH, Haskell JP, Lyons SK, Alroy J, Charnov E, Dayan T, Enquist BJ, Morgan E, Hadly E, Jones KE, Kaufman DM, Marquet PA, Maurer BM, Niklas KJ, Porter WP, Tiffney B & Willig MR 2004. Similarity of mammalian body size across the taxonomic hierarchy and across space and time. *The American Naturalist,* 163 (5): 672-691.

Stanley SM 1973. An explanation for Cope's rule. *Evolution,* 27 (1): 1-26.

Stanley SM 1979. *Macroevolution, pattern and process,* San Francisco, USA, W.H. Freeman and Company.

Uyeda JC, Hansen TF, Arnold SJ & Pienaar J 2011. The million-year wait for macroevolutionary bursts. *Proceedings of the National Academy of Sciences USA,* 108 (38): 15908-15913.

Van Valkenburgh B, Wang X & Damuth J 2004. Cope's rule, hypercarnivory, and extinction in North American canids. *Science,* 306 (5693): 101-104.

Venditti C, Meade A & Pagel M 2011. Multiple routes to mammalian diversity. *Nature,* 479 (7373): 393-396.

Wagner PJ 1996. Contrasting the underlying patterns of active trends in morphologic evolution. *Evolution,* 50 (3): 990-1007.

Appendix 1

Data collection protocol

We used a comprehensive phylogenetic tree (Fritz *et al.*, 2009) of extant mammals (n = 3,321) along with body size data from two major databases (Ernest, 2003; Jones *et al.*, 2009). Where body size was available from both sources, we prioritized data from PanTHERIA (Jones *et al.*, 2009). In some instances, species that are placed within a polytomy on the phylogeny were recorded to be identical in size. For these cases we randomly removed all but one - it is unlikely that all members of a group have exactly the same mass (Venditti *et al.*, 2011). However, results do not qualitatively change if we include these in our analysis. Body sizes were log-transformed. We classified the mammals into orders following Bininda-Emonds *et al.* (2007).

Evolutionary rates and morphological change

The variable rates model

We detected variation in the rate of body size evolution using a recently developed phylogenetic statistical method: the variable rates model (Venditti *et al.*, 2011). This model affords the opportunity to automatically detect heterogeneous rates of evolution leading to a given distribution of data at the tips of a time-calibrated phylogenetic tree (e.g. Figure A1.1A). The model detects significant shifts (either acceleration or slowing down) from an underlying homogeneous Brownian motion model of evolution. It can detect shifts within individual branches or entire clades, and without prior knowledge or specification of where and when such rate shifts occurred.

The model works by finding a set of branch-length scalars, r $(0 < r > \infty)$ that optimize the fit of the observed data (e.g., a morphological trait) to a homogeneous Brownian motion model when applied to the original branches of the phylogeny. This procedure is

equivalent to asking how much each branch of the tree should be stretched or compressed in length to make the trait best follow a homogeneous Brownian process. The result is an estimate of the underlying Brownian rate ($\sigma^2_{\text{Brownian}}$) and a set of scaled branch lengths whose scalars reveal by how much the underlying rate must be accelerated or decelerated in any given branch to conform to the Brownian process. Branches which have been stretched have experienced faster rates of change, and compressed branches have evolved at slower rates.

We fit the model using a generalized least squares (GLS) approach (Pagel, 1999) within a Bayesian Markov chain Monte Carlo (MCMC) framework. The method places a gamma prior on each rate parameter with parameter α = 1.1, and parameter β re-scaled such that the median of the distribution is 1; thus ensuring that an even number of rate increases and rate decreases are proposed. Increasing the number of rate scalars increases the number of priors, creating a trade-off between the number of rate parameters and priors (see Venditti *et al.*, 2011 for more details).There is no explicit prior on the number of rate parameters, allowing for 0 to n parameters with equal probability (where n is the number of nodes, including terminal tips, in the phylogeny).

The models we present in the main text are the output of a variable rates model run across the mammal phylogenetic tree and data. The MCMC chain was run for 2 billion iterations, sampling every 500,000 iterations after convergence. Over all iterations, we record the mean branch-specific rate scalar (the factor by which to stretch or compress the branch), generating a posterior sample of rate scalars for each branch in the tree. The posterior distribution of rate parameters differs from the prior, identifying 252 (range 205-301) rate parameters of a possible 5,234. We find an average likelihood increase of 2.56 log units per parameter.

We then use the average scalar for each branch to stretch (or if the scalar is estimated to be less than one, compress) the original time-measured branches to represent the evolution of body size better throughout the mammalian phylogeny. We then use these

newly estimated rate-scaled branch lengths in two ways. One way is to study relationships between changes along individual branches, and the other way is to study long-term trends (see below, Results and Discussion). We repeated the analysis with multiple MCMC chains to ensure convergence was achieved.

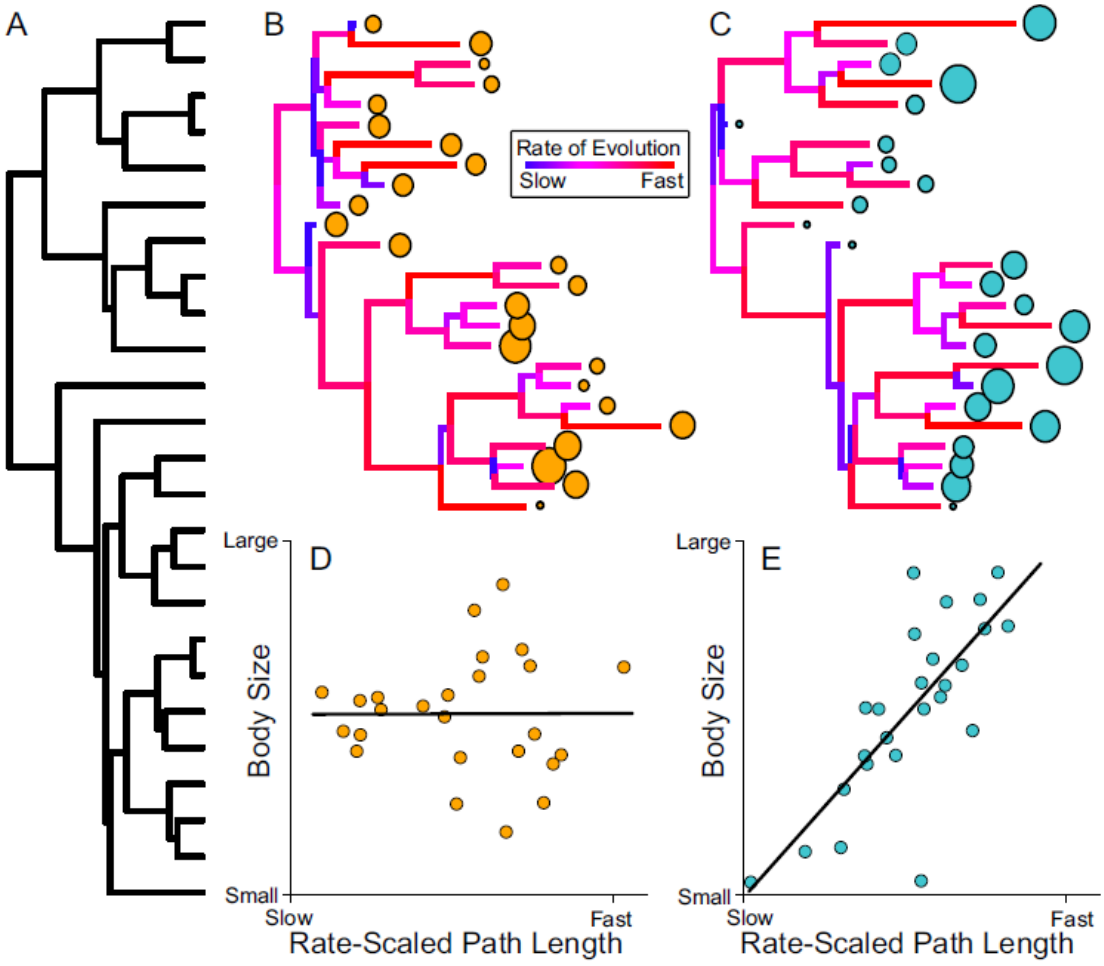


Figure A1.1: Schematic depicting how variation in the rate of morphological evolution can reveal biases in the direction of evolutionary change. (A) Ultrametric, time-scaled phylogeny that can be used with the variable rates model. (B and C) Results of scaling the time-measured branches of the phylogeny in (A) by the rate of body size evolution calculated using the variable rates model. The rate-scaled branches are calculated given the distribution of body size data at the tips represented by the orange and blue circles. (D) If rates have not been biased towards larger or smaller body sizes (as in B), then we would not observe a relationship between sizes at the tips and the path-wise rate of a species after accounting for the shared ancestry as implied by the phylogeny. (E) If there has been a bias towards smaller or larger sizes (as in C), we may expect to see a positive relationship between body size and path-wise rate, such that species that have experienced greater rates of change throughout their evolutionary history tend to have extreme values of body size.

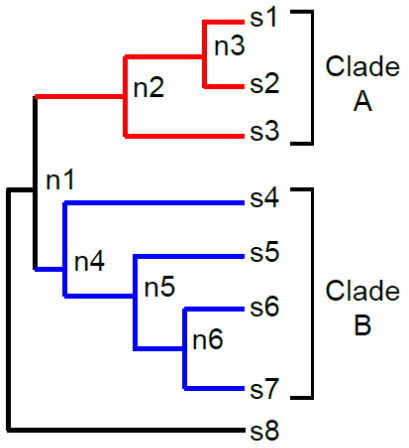
Directionality in the rate of change of body size

If the rate of morphological evolution is idiosyncratic with respect to body size, we expect no association between path-wise rate and the sizes of species at the tips (Figure A1.1B, A1.1D). No relationship would suggest that there had been no directional trends in body size throughout the evolutionary history of the group. However, if during the course of evolution, there was a tendency for rapid rates to lead to larger sizes, we would expect to find a positive relationship between size and path-wise rate (Figure A1.1C, A1.1E). Such a pattern would indicate that where there was a large amount of evolutionary change, as reflected by elevated rates of evolution, species tended to become bigger in size.

To test this idea, we regressed $\log_{10}$ body mass against path-wise rate across all mammals using phylogenetic GLS models (Freckleton *et al.*, 2002; Pagel, 1997) in a maximum-likelihood framework (Franks *et al.*, 2012; Jetz and Freckleton, 2015; Organ *et al.*, 2007). We then used a model which allowed the relationship between path-wise rate and body size to vary among groups (*separate-slopes model*) in order to determine whether there are differences in the patterns of body size evolution among the mammalian orders. We compared nested models using the likelihood-ratio test statistic (D) to determine whether it was significantly better to estimate separate relationships among mammalian orders.

To allow for differential patterns of body size evolution within the mammal orders in our separate-slopes model, all taxa were assigned to an order (Bininda-Emonds *et al.*, 2007) using standard contrast ('dummy') coding. The relationships between body size and path-wise rate were then estimated within each group simultaneously given their taxonomic affiliation (Figure A1.2). To avoid over-parameterization, we did not estimate slopes for orders where sample size was not large enough to ensure ten data points per parameter estimate (Freckleton and Watkinson, 2001). Our model estimates four parameters: the intercept, slope, phylogenetic variance, and a measure of phylogenetic signal λ (Pagel, 1999); we therefore only allow a separate intercept and slope for orders with a sample size of $n \geq 40$. The following orders all have too few species to study individually but fall within

the monophyletic superorder Afrotheria: Macroscelidea (n = 14), Afrosoricida (n = 34), Tubulidentata (n = 1), Hyracoidea (n = 4), Proboscidea (n = 3) and Sirenia (n = 4). For this reason, we study Afrotheria (n = 60) as a single group. It has been suggested that aquatic species may have different patterns and processes of body size evolution (Evans *et al.*, 2012; Schmidt-Nielsen, 1984) . We therefore allowed the magnitude of the relationship to vary within aquatic groups by giving them a separate intercept.



Node / Species	Clade A	Clade B	Clade AB
s1	1	0	1
s2	1	0	1
s3	1	0	1
s4	0	1	1
s5	0	1	1
s6	0	1	1
s7	0	1	1
s8	0	0	0
n1	-	-	1
n2	1	0	-
n3	1	0	-
n4	0	1	-
n5	0	1	-
n6	0	1	-

Figure A1.2: Use of standard contrast coding in the predictive modelling approach is illustrated on a portion of a small schematic phylogeny. Terminal taxa s1 – s7 are assigned to either Clade A or Clade B (s8 is an outgroup) as shown in the table on the right. Internal nodes are assigned to orders depending on their placement in the phylogeny using the same standard coding, as shown in the table. For example, nodes n2 and n3 are treated as a member of Clade A. Hypothetically, the parameters from a regression model across the terminal taxa allowing the relationship between path-wise rate and body size to vary among all of the clades are then used to reconstruct these internal nodes. Nodes which lead to multiple groups (e.g. n1) are allocated to a new group which includes all members of all descendants of that node. For example, n1 contains all members of Clade A and all members of Clade B. An example of the regression model used to reconstruct n1, as an example, is illustrated in the shaded column of the table.

Polytomies

We test whether the number of unresolved nodes that occur along a species path causes any systematic bias in the results by counting the number of polytomies along the evolutionary path leading to each species and assessing the impact of polytomy number in our models.

We adapted the main analysis of the paper by running a multiple regression, where body size was regressed on path-wise rate and polytomy count – allowing for the effect of polytomies on the relationship between body size and path-wise rate. This also did not provide a significant improvement over the simple single-slope model (D = 0.301, P = 0.582, *df* = 1).

We then ran our full separate-slopes model but also allowed the effect of polytomies to vary within each of the 11 orders i.e. each order was given a separate parameter both for path-wise rate and for polytomy count. This did not provide a significant improvement over our separate-slopes model (D = 7.74, P = 0.74, *df* = 11). Based on these results we conclude that the major conclusions of this paper remain unaffected by the presence of unresolved nodes in the phylogeny.

Approaching an upper size boundary

If mammals are approaching some upper size limit, then it will be less advantageous for body size to increase when species are already large or are approaching a size close to that limit. We would therefore expect a decrease in the rate of change towards body size increase at larger sizes (i.e. we should find evidence for concave quadratic curves in the relationship between body size and path-wise rate) (Alroy, 1998). Fitting a single quadratic term to the model across all mammals does not significantly improve the likelihood over a linear model (D = 0.150, P = 0.69, *df* = 1) Allowing the curvilinear relationship to vary among the orders does not reveal a significant curve in the expected direction for any group. The finding that no single order (nor all mammals when analysed as a single group) significantly fits a concave quadratic curve does not lend support to the suggestion that mammals are approaching an upper limit on maximum size.

Reconstructing historical body sizes

Estimating ancestral states

We then used the parameters of our separate-slopes regression model (Table A1.1) to estimate the ancestral body size at each internal node of the mammalian phylogeny (n = 1913) using a phylogenetic predictive modelling approach adapted from standard phylogenetic GLS models (Franks *et al.*, 2012; Jetz and Freckleton, 2015; Organ *et al.*, 2007). This approach uses the parameters of an evolutionary regression model to infer maximum likelihood estimates for unknown data. This approach allows us to account for trends in body size evolution and variation in the rate of morphological evolution observed in the data (Organ *et al.*, 2007; Pagel, 1999). This approach therefore avoids limitations of traditional homogenous Brownian motion methods for ancestral state reconstruction that necessarily constrain inferences to be intermediate to the range of extant data (Pagel, 1999; Webster and Purvis, 2001). We adapted this method to estimate ancestral sizes by placing zero branch-length 'false tips' at each internal node (Figure A1.2). All false tips were assigned to an order in the same way as terminal taxa using standard contrast coding and were reconstructed as a member of that group (Figure A1.2).

It was necessary to modify slightly the predictive modelling approach for deeper nodes, which lead to multiple orders (e.g. n1, Figure A1.2). For each deep node (n = 20) an additional and separate regression model was run where all taxa which descend from the given node were reconstructed as a separate group. Given the parameters from this new regression model, body sizes were reconstructed using the same predictive modelling method described above, where the false tip placed at the deeper node was assigned to the new group using standard contrast coding (Figure A1.2, grey column).

Where a given node fell within a group where sample sizes were too small for analysis ($n < 40$), it was not ignored but, instead, estimated using a homogenous Brownian motion

model. By including all possible species, allowed deeper nodes to be reconstructed whilst still accounting for the existence of smaller groups in the phylogeny (e.g. the node leading to Pholidota [n = 7] and Carnivora [n = 249] was estimated using all 256 species). For comparison, all ancestral sizes were also reconstructed using the conventional homogenous Brownian motion method. We then projected the phylogeny and size data (both reconstructed nodes and tips, n = 5,233) into a phylomorphospace (Sidlauskas, 2008) using R package phytools (Revell, 2012).

Table A1.1: Parameter estimates from the models presented in the main text.

Order	**Slope (β)**	**SE**	***t***	***P***
Afrotheria	0.0043	0.0006	7.181	<0.001
Carnivora	0.0086	0.0012	7.386	<0.001
Cetartiodactyla	0.0028	0.0005	6.074	<0.001
Chiroptera	0.0082	0.001	8.328	<0.001
Dasyuromorphia	0.0071	0.0008	9.031	<0.001
Didelphimorphia	0.0057	0.0016	3.542	<0.001
Diprotodontia	-0.0046	0.0008	-6.081	<0.001
Eulipotyphla	0.0058	0.0013	4.455	<0.001
Lagomorpha	0.0241	0.0118	2.035	0.042
Primates	0.0063	0.0025	2.548	0.011
Rodentia	0.0037	0.0003	10.829	<0.001
All mammals	0.0039	0.0002	19.536	<0.001

Phylogenetic ancestor-descendant comparisons and regression to the mean

We studied the evolution of body size in the face of variable rates of evolution quantitatively by finding the difference between $\log_{10}$ ancestral size and $\log_{10}$ descendant size of each branch of the phylogeny ($\Delta\log_{10}$ body size, n = 5,233) . We refer to these as PAD comparisons.

For each order and across all mammals we calculated the average percentage difference of all PAD size changes; these are shown in Table A1.2.

Table A1.2: Average percentage size increases found using PAD comparisons.

Order	N	% Increase	SE
Afrotheria	99	16.91	4.28
Carnivora	452	6.80	0.52
Cetartiodactyla	444	1.71	0.27
Chiroptera	1062	11.24	0.66
Dasyuromorphia	111	10.61	2
Didelphimorphia	91	8.27	1.65
Diprotodontia	205	-3.84	0.63
Eulipotyphla	245	13.57	1.9
Lagomorpha	116	10.71	2.11
Primates	443	1.37	0.22
Rodentia	1790	3.45	0.31
All Mammals	5233	5.86	0.25

We use these PAD comparisons to test the idea that mammalian evolution has proceeded in the presence of some lower bound (Stanley, 1973). If some lower boundary is enforced preventing mammal species from decreasing in size, we would predict a negative relationship between ancestral size and $\Delta\log_{10}$ body size (more details are provided in the main text). We also test for evidence for some constraint that restricts smaller-bodied species from having higher rates (Simpson, 1953; Stanley, 1979). If a constraint is enforced that is relaxed in larger-bodied species, allowing them to experience greater evolutionary change, we would expect to observe an increase in the variance of body size change as ancestral size increases (more details are provided in the main text).

For both of these tests we are interested in the relationship between log-transformed ancestral size and the inferred $\Delta\log_{10}$ body size along an individual branch. When we

regress $\Delta\log_{10}$ body size, on $\log_{10}$ ancestral size, it is the equivalent of a repeated-measures analysis and it is therefore necessary to correct for the so-called "regression to the mean" effect (Alroy, 1998; Kelly and Price, 2005). This confounding effect arises as a consequence of taking continuous measurements from the same sample at different time points – when we compare two time intervals, individuals with higher values than the average will tend to have lower than average values at the second interval (Kelly and Price, 2005).

Following the method of Kelly and Price (2005), we first tested to see if there is a differential effect between $\log_{10}$ ancestral size and $\log_{10}$ descendant size that pushes values towards the mean using Pitman's test (T) (Pitman, 1939) for the equality of variances in paired samples:

$$T = \frac{\sqrt{n-2}[(s_1/s_2) - (s_2/s_1)]}{2\sqrt{1-r^2}}$$

where n refers to the number of branches; s_1 and s_2 are the sample SDs associated with $\log_{10}$ ancestral size and $\log_{10}$ descendant size respectively; and r is the observed correlation coefficient between $\log_{10}$ ancestral size and $\log_{10}$ descendant size. This test demonstrated that the null hypothesis was rejected; we find significant deviation from the predictions made by equality of variances (T = 165.35, P < 0.001, df = 5231).

Following the method of Kelly and Price (2005) we then adjusted each computed $\Delta\log_{10}$ body size by subtracting the change expected as a consequence of regression to the mean (Δ_{adj}). We modified the equations presented in Kelly and Price (2005) to accommodate the direction of change that is necessary for interpretation in our models. In the below equation we retain the original notation used in Kelly and Price (2005) for ease of comparison and interpretation where X_1 refers to $\log_{10}$ ancestral size and X_2 refers to $\log_{10}$ descendant size:

$$\Delta^* = r(X_1 - \bar{X}_1) - (X_2 - \bar{X}_2),$$

where *r* would be replaced by a correction factor ρ if the null hypothesis of equal variances was accepted [details are provided in Kelly and Price (2005)]. Our adjusted values for PAD size changes were therefore calculated as:

$$\Delta_{adj}\log_{10}\text{body size} = \Delta\log_{10}\text{body size} - \Delta^{*}$$

We then use $\Delta_{adj}\log_{10}$ body size to analyse the relationship between $\log_{10}$ ancestral size and the amount of size change that has occurred along an individual branch of the phylogenetic tree (as discussed in the main text).

Comparisons to the fossil record

It may be misleading to assume that fossil taxa are ancestral to extant groups (Foote, 1996) and therefore it may not be useful to compare estimated nodal values directly with the paleontological data. To determine the accuracy of our ancestral state reconstructions, we reconstructed body sizes of fossil taxa by mapping them onto the extant phylogeny according to the following protocol.

First, we obtained body size data and age ranges for 65 unique fossil taxa that are recorded in the literature as the oldest and closest sister group or basal to 49 different taxonomic groups in the mammalian phylogeny (Table A1.3). These fossil taxa represent 10 of the 11 mammalian orders explored in the main text (the only order with no fossil representatives in this dataset is Rodentia). Some orders and lower level taxonomic groups have multiple fossils which have been proposed to be the oldest or most basal member. Wherever we found multiple representatives of a single group, each fossil was analysed in separate analyses (Table A1.3). Where fossils are suggested by different sources to be a member of more than one group, they have been entered as duplicate rows in the table (Table A1.3). This scenario occurred for three of the 65 fossil taxa: *E. scansoria* was originally proposed to be ancestral to Eutheria (Ji *et al.*, 2002) but in another source, it is suggested to be a therian ancestor (O'Leary *et al.*, 2013); *Eritherium azzouzorum* is placed as a member of

either Tethytheria (O'Leary *et al.*, 2013) or Proboscidea (Sen, 2013); and finally, *Icaronycteris* has been placed as a basal bat (Gunnell and Simmons, 2005) but has also been described as the earliest microchiropteran (Simmons and Geisler, 1998). We therefore considered the placement of 68 representatives.

Where an entry was genus level and information on body sizes existed for multiple species within that genus, body size was taken as an average of all the available data. Where body size ranges are recorded, the midpoint was taken for comparison. For *Indohyus*, a conservative point estimate of 5000g was used (described as <5000g, Table A1.3).

Second, we then inserted each of the fossil representatives into the time-scaled extant phylogeny along the branch leading to the group to which it has been assigned. Because each fossil has an age range, rather than an individual time point, the resolution of the divergence date fell into one of five potential scenarios (Figure A1.3). Branch length was determined as described for each scenario. We then placed the fossil taxa onto the rate-scaled phylogeny in the equivalent position (i.e. if the fossil occurred halfway along a branch on the time tree, it was placed halfway along the corresponding branch on the scaled phylogeny). We then scaled the branch by the rate as described below for each scenario and measured the rate-scaled path length to measure the intensity of evolutionary change experienced by each fossil taxon in the same way as described for extant taxa.

The first scenario (blue, Figure A1.3) was when both the maximum and minimum age of the fossil fall along the branch leading to the root of the group. This branch is henceforth referred to as the *leading branch* (branch nA-nB in Fig. A1.3). Sixteen fossils fell into this scenario. The divergence date for taxa within this scenario was placed at the maximum fossil age along the leading branch. Branch length was then extended to the minimum

age of the fossil, and was scaled by the average rate along the equivalent leading branch in the rate-scaled phylogeny.

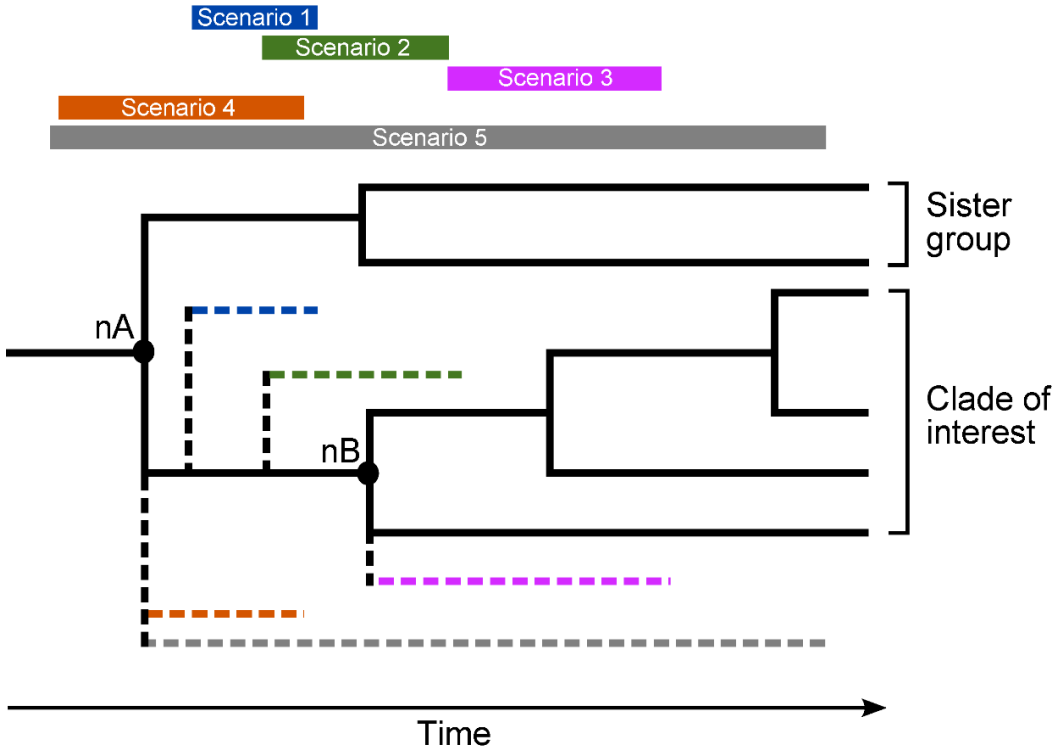


Figure A1.3: This schematic illustrates placement of fossil taxa onto the extant time-scaled phylogeny. Each taxon in our dataset (Table A1.3) has an age range (first appearance date to last appearance date), and thus can fall into five different scenarios depending on its duration in the fossil record. The coloured bars at the top of the figure represent the different placement scenarios. Each fossil has an assumed phylogenetic position that falls somewhere along the leading branch – the branch leading to the clade of interest. This leading branch is represented by the branch between nodes nA and nB in the schematic. For each scenario, taxa were then placed onto the phylogeny (see text for details) at the position indicated by the corresponding coloured dashed branches. Branch length was extended to the minimum fossil age as shown in the figure.

The second scenario (green, Fig. A1.3) was when the maximum age of the fossil was along the leading branch, but the minimum age was younger than the divergence at the root. Six placements fell within this scenario. Placement of these taxa was resolved in an identical manner to the first scenario.

The third scenario (pink, Figure A1.3) occurred when both the maximum and minimum age of the fossil were younger than the root of the group of interest. Because most of the

data available were for the oldest members of (rather than directly ancestral to) a group, this scenario was the most likely scenario, and occurred for 44 of the fossils in our dataset. The divergence date for each of the taxa in this scenario was placed at the base of the group the fossil was assigned to, within a polytomy (Figure A1.3). The branch length was extended to the minimum fossil age, and then, on the rate-scaled phylogeny, was scaled given the average rate of all other branches which stem from the basal node (i.e. all branches that start at node B in Figure A1.3).

The fourth scenario (orange, Figure A1.3) occurred when the maximum age of the fossil was older than the initial divergence between the group of interest and its closest sister group, but the minimum age did fall along the leading branch. This scenario occurred for a single fossil. The placement of this taxon was resolved in a manner identical to the first scenario.

The fifth scenario (grey, Figure A1.3) occurred when the maximum fossil age predated the initial divergence of the group to which the taxa was assigned but the minimum age was also much younger than the base of the group. Just one of the fossils we looked at fell into this scenario. In this case, the placement was resolved by bringing forward the divergence date of the fossil to fall within a polytomy at the basal split between the assigned group and the sister group (e.g. node nA, Figure A1.3). The branch length was then extended to the minimum age of the fossil, and then was scaled by the average rate of the leading branch.

Finally, although an alternative scenario was possible where both the maximum and minimum age of a fossil could be older than the initial divergence between the group of interest and its closest sister group, this occurred for none of the fossil representatives in our dataset.

After each of the fossils had been placed onto both the time-scaled and rate-scaled phylogeny, and given appropriate branch lengths as described, we reconstructed body sizes using the same approach described for internal nodes. These taxa were assigned to the order that the group they represent belongs to using standard contrast coding and their body sizes were estimated using the phylogenetic predictive modelling procedure described earlier, given the observed relationship between body size and rate of evolution within that order. Taxa representing multiple groups e.g. *E. scansoria* (Table A1.3) were assigned to groups in the same way described above for ancestral state reconstruction at deep nodes (Figure A1.2). For comparison, each taxon's mass was also estimated using a homogenous Brownian motion model.

The accuracy of these reconstructions was then compared using phylogenetic GLS regression models with the paleontological body size as the independent variable and estimated body size as the dependent variable. A phylogenetic *t*-test was used to compare mean sizes between methods. To avoid non-independence of data points, the models presented in the main text consider *E. scansoria* as a therian mammal, *E. azzourum* to be ancestral to Tethytheria, and *Icaronycteris* to be an early bat. These taxa are indicated with an asterisk in Table A1.3. When we ran regression models and *t* tests that considered the alternative placements of these fossils, results were unaffected.

References for Appendix 1 Text[1]

Alroy J 1998. Cope's rule and the dynamics of body mass evolution in North American fossil mammals. *Science,* 280 (5364): 731-734.

Bininda-Emonds OR, Cardillo M, Jones KE, Macphee RD, Beck RM, Grenyer R, Price SA, Vos RA, Gittleman JL & Purvis A 2007. The delayed rise of present-day mammals. *Nature,* 446 (7135): 507-12.

Ernest SKM 2003. Life history characteristics of placental nonvolant mammals. *Ecology,* 84 (12): 3402.

Evans AR, Jones D, Boyer AG, Brown JH, Costa DP, Ernest SKM, Fitzgerald EMG, Fortelius M, Gittleman JL, Hamilton MJ, Harding LE, Lintulaakso K, Lyons SK, Okie JG, Saarinen JJ, Sibly RM, Smith FA, Stephens PR, Theodor JM & Uhen MD 2012. The maximum rate of mammal evolution. *Proceedings of the National Academy of Sciences USA,* 109 (11): 4187-4190.

Foote M 1996. On the probability of ancestors in the fossil record. *Paleobiology,* 22 (2): 141-151.

Franks PJ, Freckleton RP, Beaulieu JM, Leitch IJ & Beerling DJ 2012. Megacycles of atmospheric carbon dioxide concentration correlate with fossil plant genome size. *Philosophical Transactions of the Royal Society B: Biological Sciences,* 367 (1588): 556-64.

Freckleton RP, Harvey PH & Pagel M 2002. Phylogenetic analysis and comparative data: A test and review of evidence. *The American Naturalist,* 160 (6): 712-726.

Freckleton RP & Watkinson AR 2001. Nonmanipulative determination of plant community dynamics. *Trends in Ecology & Evolution,* 16 (6): 301-307.

Fritz SA, Bininda-Emonds ORP & Purvis A 2009. Geographical variation in predictors of mammalian extinction risk: big is bad, but only in the tropics. *Ecology letters,* 12 (6): 538-549.

Gunnell GF & Simmons NB 2005. Fossil evidence and the origin of bats. *Journal of Mammalian Evolution,* 12 (1-2): 209-246.

Jetz W & Freckleton RP 2015. Towards a general framework for predicting threat status of data-deficient species from phylogenetic, spatial and environmental information. *Philosophical Transactions of the Royal Society B: Biological Sciences,* 370 (1662).

Ji Q, Luo Z-X, Yuan C-X, Wible JR, Zhang J-P & Georgi JA 2002. The earliest known eutherian mammal. *Nature,* 416 (6883): 816-822.

Jones KE, Bielby J, Cardillo M, Fritz SA, O'dell J, Orme CDL, Safi K, Sechrest W, Boakes EH, Carbone C, Connolly C, Cutts MJ, Foster JK, Grenyer R, Habib M, Plaster CA, Price SA,

[1] References for Table A1.3 are listed separately.

Rigby EA, Rist J, Teacher A, Bininda-Emonds ORP, Gittleman JL, Mace GM, Purvis A & Michener WK 2009. PanTHERIA: a species-level database of life history, ecology, and geography of extant and recently extinct mammals. *Ecology,* 90 (9): 2648.

Kelly C & Price TD 2005. Correcting for regression to the mean in behavior and ecology. *The American Naturalist,* 166 (6): 700-707.

O'leary MA, Bloch JI, Flynn JJ, Gaudin TJ, Giallombardo A, Giannini NP, Goldberg SL, Kraatz BP, Luo ZX, Meng J, Ni X, Novacek MJ, Perini FA, Randall ZS, Rougier GW, Sargis EJ, Silcox MT, Simmons NB, Spaulding M, Velazco PM, Weksler M, Wible JR & Cirranello AL 2013. The placental mammal ancestor and the post-K-Pg radiation of placentals. *Science,* 339 (6120): 662-7.

Organ CL, Shedlock AM, Meade A, Pagel M & Edwards SV 2007. Origin of avian genome size and structure in non-avian dinosaurs. *Nature,* 446 (7132): 180-4.

Pagel M 1997. Inferring evolutionary processes from phylogenies. *Zoologica Scripta,* 26 (4): 331-348.

Pagel M 1999. Inferring the historical patterns of biological evolution. *Nature,* 401: 877-884.

Pitman E 1939. A note on normal correlation. *Biometrika,* 31 (1/2): 9-12.

Revell LJ 2012. phytools: an R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution,* 3: 217-223.

Schmidt-Nielsen K 1984. *Scaling: Why is Animal Size so Important?,* Cambridge, Cambridge University Press.

Sen S 2013. Dispersal of African mammals in Eurasia during the Cenozoic: Ways and whys. *Geobios,* 46 (1–2): 159-172.

Sidlauskas B 2008. Continuous and arrested morphological diversification in sister clades of characiform fishes: a phylomorphospace approach. *Evolution,* 62 (12): 3135-56.

Simmons NB & Geisler JH 1998. Phylogenetic relationships of Icaronycteris, Archaeonycteris, Hassianycteris, and Palaeochiropteryx to extant bat lineages, with comments on the evolution of echolocation and foraging strategies in Microchiroptera. *Bulletin of the American Museum of Natural History,* 235.

Simpson GG 1953. *The major features of evolution,* London, GB, Columbia University Press.

Stanley SM 1973. An explanation for Cope's rule. *Evolution,* 27 (1): 1-26.

Stanley SM 1979. *Macroevolution, pattern and process,* San Francisco, USA, W.H. Freeman and Company.

Venditti C, Meade A & Pagel M 2011. Multiple routes to mammalian diversity. *Nature,* 479 (7373): 393-396.

Webster AJ & Purvis A 2001. Testing the accuracy of methods for reconstructing ancestral states of continuous characters. *Proceedings of the Royal Society B: Biological Sciences,* 269: 143-149.

Table A1.3: Data and sources used for fossil comparisons and the results of body mass reconstructions. References for the placement of each fossil taxa are given. The $\log_{10}$ body mass (g) for each species is given as reported by the literature (BM) and as reconstructed in both our homogenous (BM-H) and separate-slopes (BM-S) models. The age of each fossil and the scenario (Sc) in which it fell for placement on the phylogeny is shown. All references are shown to the right of the relevant column and are listed in detail separately below the table. *Included in the analyses reported in the main text as a member of the group indicated in the table, but is reported to be the oldest member of more than one taxonomic group.

Order	Group	Taxa	Refs	BM	Refs	Min Age	Max Age	Refs	Sc	BM-H	BM-S
Basal / Multiple	Australidelphia	*Djarthia murgonensis*	(1; 2)	1.63	(3)	48.6	55.8	(4)	S3	2.35	2.26
	Placentalia	*Eomaia scansoria*	(5)	1.35	(5)	125	130	(4)	S1	2.85	1.46
	Theria	*Eomaia scansoria**	(6)	1.35	(5)	125	130	(4)	S3	2.79	1.48
	Marsupialia	*Sinodelphys szalayi*	(1; 6)	1.51	(7)	125	130	(4)	S1	2.64	1.72
Afrotheria	Macroscelidea	*Chambius kasserinensis*	(8; 9)	1.11	(9)	40.4	55.8	(4)	S2	2.57	1.71
	Proboscidea	*Eritherium azzouzorum*	(10; 11)	3.74	(10; 12)	55.8	58.7	(4)	S1	4.72	3.78
	Tethytheria	*Eritherium azzouzorum**	(6)	3.74	(10; 12)	55.8	58.7	(4)	S3	4.05	3.12
	Proboscidea	*Mammuthus*	(13)	6.94	(14)	0	7.25	(4)	S3	6.27	6.56
	Proboscidea	*Phosphatherium escuilliei*	(8; 11)	4.10	(15)	48.6	59.2	(4; 15)	S1	4.70	4.26
	Sirenia	*Protosiren smithae*	(8)	5.43	(16)	33.9	40.4	(4)	S3	4.87	5.11
	Hyracoidea	*Saghatherium bowni*	(17)	3.90	(17)	28.4	33.9	(4)	S1	3.70	3.15
Carnivora	Caniformes	*Hesperocyon gregarius*	(6)	3.47	(14)	30.8	37.2	(4)	S3	3.99	2.99
	Felidae	*Homotherium serum*	(18)	4.94	(14)	0.01	0.3	(4)	S3	4.02	4.02
	Ursidae	*Parictis*	(1)	3.12	(14)	33.9	33.3	(4)	S3	4.65	3.38
	Carnivora	*Protictis haydenianus*	(6)	3.32	(14; 19)	56.8	61.7	(4)	S3	3.76	1.99
	Procyonidae	*Pseudobassaris*	(1)	3.25	(20)	29	30	(21)	S1	3.50	2.41
Cetartiodactyla	Tylopoda	*Antiacodon*	(22)	3.32	(14)	37.2	50.3	(4)	S1	4.83	4.62
	Cetartiodactyla	*Cainotherium*	(6)	3.18	(23)	12.8	37.2	(4)	S3	4.65	4.51
	Hippopotamidae	*Cebochoerus*	(22)	3.67	(24)	33.9	40.4	(4)	S1	5.66	4.87
	Cetacea	*Dalanistes ahmedi*	(25)	5.88	(26)	40.4	48.6	(4)	S3	5.67	5.90

Table A1.3 (cont.):

Order	Group	Taxa	Refs	BM	Refs	Min Age	Max Age	Refs	Sc	BM-H	BM-S
Cetartiodactyla	Ruminantia	*Diacodexis*	(22)	3.12	(14)	46.2	50.3	(4)	S2	4.75	4.19
(cont.)	Moschidae	*Dremotherium feignouxi*	(27)	4.42	(28)	20	22.4	(4)	S1	4.59	4.36
	Bovidae	*Eotragus noyei*	(27)	4.30	(29)	18	18.3	(30)	S3	4.71	4.54
	Suina	*Helohyus*	(22)	4.27	(14)	40.4	46.2	(4)	S2	4.70	4.50
	Cetacea	*Indohyus*	(22)	4.70	(31)	40.4	48.6	(4)	S3	5.66	5.90
	Delphinoidea	*Kentriodon pernix*	(25)	4.26	(26)	15.9	20.4	(4)	S3	5.34	5.20
	Ruminantia	*Leptomeryx*	(22)	3.85	(14)	24.8	33.9	(4)	S3	4.69	4.36
	Suina	*Perchoerus*	(22)	4.59	(14)	30.8	37.2	(4)	S3	4.70	4.54
	Tylopoda	*Poebrotherium*	(22)	4.42	(14)	30.8	33.9	(4)	S1	5.02	4.85
	Cetacea	*Saghacetus oskia*	(25)	5.54	(26)	33.9	37.2	(4)	S3	5.66	6.19
	Odontoceti	*Squaloziphius emlongi*	(25)	5.22	(26)	20.4	23	(4)	S3	5.68	5.66
	Odontoceti	*Xenorophus*	(25)	4.99	(26)	23	28.4	(4)	S3	5.67	5.64
Chiroptera	Microchiroptera	*Archaeonycteris*	(32)	1.44	(33)	40.4	48.6	(4)	S3	1.65	0.73
	Hipposideros	*Hipposideros schlosseri*	(34)	1.15	(35)	33.9	37.2	(4)	S3	1.03	0.63
	Chiroptera	*Icaronycteris**	(32)	1.13	(36)	46.2	55.8	(4)	S3	1.83	0.77
	Microchiroptera	*Icaronycteris*	(33)	1.13	(36)	46.2	55.8	(4)	S3	1.65	0.69
	Microchiroptera	*Paleochiropteryx*	(32)	1.05	(33)	40.4	48.6	(4)	S3	1.65	0.73
Dasyuromorphia	Dasyuridae	*Barinya wangala*	(37)	1.11	(38)	16	23	(4)	S3	1.67	1.26
Didelphimorphia	Didelphimorphia	*Nortedelphys*	(39)	1.73	(40)	66	70.6	(4)	S1	2.14	1.45
Diprotodontia	Petauridae	*Djaludjangi*	(37)	2.43	(38)	11.6	23	(41)	S3	2.39	2.77
Eulipotyphla	Erinaceomorpha	*Adunator*	(1)	1.20	(14; 42)	56.8	61.7	(4)	S3	2.36	1.50
	Eulipotyphla	*Batodon*	(6)	0.63	(14)	66	70.6	(4)	S3	2.65	1.29
	Erinaceomorpha	*Diacocherus*	(1)	1.33	(14)	50.3	56.8	(4)	S3	2.36	1.54
	Soricomorpha	*Leptacodon*	(1)	1.00	(14)	50.3	66	(4)	S3	2.65	1.43
	Erinaceidae	*Litolestes*	(1)	1.31	(14)	56.8	61.7	(4)	S3	2.36	1.50

Table A1.3 (cont.):

Order	Group	Taxa	Refs	BM	Refs	Min Age	Max Age	Refs	Sc	BM-H	BM-S
Lagomorpha	Lagomorpha	*Mytonolagus*	(1)	2.13	(14)	37.2	40.4	(4)	S3	2.60	1.60
	Leporidae	*Palaeolagus*	(1; 43)	2.14	(14)	20.4	33.9	(4)	S2	3.02	2.52
Primates	Tarsiiformes	*Archicebus achilles*	(44)	1.40	(44)	54.8	55.8	(44)	S1	2.58	2.23
	Anthropoidea	*Biretia peveteaui*	(6)	2.58	(45)	33.9	37.2	(4)	S3	3.43	3.23
	Platyrrhini	*Branisella boliviana*	(6)	3.00	(46)	21	29	(4)	S2	3.27	3.10
	Catarrhini	*Catopithecus brownii*	(6; 47)	2.91	(48)	28.4	33.9	(4)	S3	3.87	3.67
	Primates	*Dryomomys szalayi*	(49)	1.02	(49)	55	75	(49)	S3	2.82	2.48
	Anthropoidea	*Eosimias sinensis*	(49)	1.97	(46)	37.2	48.6	(4)	S3	3.43	3.19
	Hominoidea	*Kalepithecus songhorensis*	(47)	3.70	(46)	13.7	20.4	(47)	S3	4.24	4.34
	Catarrhini	*Kamoyapithecus hamiltoni*	(47)	4.54	(50)	23	28.1	(51)	S3	3.87	3.71
	Lemuriformes	*Karanisia clarki*	(6)	2.30	(52)	33.9	37.2	(4)	S3	2.92	2.74
	Galagidae	*Komba*	(49)	2.48	(49)	13.7	22.4	(4)	S3	2.35	2.23
	Lorisiformes	*Mioeuoticus*	(53)	2.48	(46)	18	19	(54)	S3	2.55	2.40
	Strepsirrhini	*Plesiopithecus*	(55)	2.94	(48)	28.4	33.9	(4)	S3	2.79	2.58
	Hominoidea	*Proconsul*	(46)	4.49	(46)	11.6	28.4	(4)	S2	4.06	3.93
	Hominoidea	*Rukwapithecus fleaglei*	(47)	4.09	(47)	24	26	(47)	S1	4.13	3.85
	Catarrhini	*Saadanius hijazensis*	(47)	4.24	(56)	23	33.9	(4)	S3	3.87	3.71
	Galagidae	*Saharagalago misrensis*	(1; 6; 52)	2.09	(52)	33.9	37.2	(4)	S1	2.52	2.26
	Alouatta	*Stirtonia*	(46)	3.90	(46)	7.25	23	(4)	S4	3.50	3.55
	Tarsiiformes	*Teilhardina belgica*	(49)	1.66	(49)	48.6	55.8	(4)	S1	2.58	2.27
	Aotus	*Tremacebus harringtoni*	(46)	3.26	(46)	16	28.1	(46)	S5	3.24	3.14
	Galagidae	*Wadilemur*	(1)	2.13	(48)	28.4	33.9	(4)	S1	2.47	2.25

References for fossil data

1 Rose KD 2006. *The Beginning of the Age of Mammals*, Baltimore, MD, John Hopkins University Press.

2 Beck RMD, Godthelp H, Weisbecker V, Archer M & Hand SJ 2008. Australia's oldest marsupial fossils and their biogeographical implications. *PLoS ONE*, 3 (3): e1858.

3 Beck RM 2013. A peculiar faunivorous metatherian from the early Eocene of Australia. *Acta Palaeontologica Polonica,* 60 (1): 123-129

4 Fossilworks. The Paleobiology Database. http://fossilworks.org. [Accessed 02/10/14].

5 Ji Q, Luo Z-X, Yuan C-X, Wible JR, Zhang J-P & Georgi JA 2002. The earliest known eutherian mammal. *Nature*, 416 (6883): 816-822.

6 O'leary MA, Bloch JI, Flynn JJ, Gaudin TJ, Giallombardo A, Giannini NP, Goldberg SL, Kraatz BP, Luo ZX, Meng J, Ni X, Novacek MJ, Perini FA, Randall ZS, Rougier GW, Sargis EJ, Silcox MT, Simmons NB, Spaulding M, Velazco PM, Weksler M, Wible JR & Cirranello AL 2013. The placental mammal ancestor and the post-K-Pg radiation of placentals. *Science*, 339 (6120): 662-7.

7 Luo Z-X, Ji Q, Wible JR & Yuan C-X 2003. An Early Cretaceous tribosphenic mammal and metatherian evolution. *Science*, 302 (5652): 1934-1940.

8 Seiffert ER 2007. A new estimate of Afrotherian phylogeny based on simultaneous analysis of genomic, morphological, and fossil evidence. *BMC Evolutionary Biology*, 7: 224.

9 Tabuce R, Marivaux L, Adaci M, Bensalah M, Hartenberger JL, Mahboubi M, Mebrouk F, Tafforeau P & Jaeger JJ 2007. Early Tertiary mammals from North Africa reinforce the molecular Afrotheria clade. *Proceedings of the Royal Society of London, Series B*, 274 (1614): 1159-66.

10 Sen S 2013. Dispersal of African mammals in Eurasia during the Cenozoic: Ways and whys. *Geobios*, 46 (1–2): 159-172.

11 Gheerbrant E, Amaghzaz M, Bouya B, Goussard F & Letenneur C 2014. Ocepeia (Middle Paleocene of Morocco): The Oldest Skull of an Afrotherian Mammal. *PloS one*, 9 (2): e89739.

12 Gheerbrant E 2009. Paleocene emergence of elephant relatives and the rapid radiation of African ungulates. *Proceedings of the National Academy of Sciences USA*, 106 (26): 10717-10721.

13 Shoshani J 1998. Understanding proboscidean evolution: a formidable task. *Trends in Ecology & Evolution*, 13 (12): 480-487.

14 Alroy J 2000. New methods for quantifying macroevolutionary patterns and processes. *Paleobiology*, 26 (4): 707-733.

15 Gheerbrant E 1998. The oldest known proboscidean and the role of Africa in the radiation of modern orders of placentals. *Bulletin of the Geological Society of Denmark*, 44: 181-185.

16 Sarko DK, Domning DP, Marino L & Reep RL 2010. Estimating body size of fossil sirenians. *Marine Mammal Science*, 26 (4): 937-959.

17 Schwartz GT, Rasmussen DT & Smith RJ 1995. Body-size diversity and community structure of fossil hyracoids. *Journal of Mammalogy*, 76 (4): 1088-1099.

18 Agnarsson I, Kuntner M & May-Collado LJ 2010. Dogs, cats, and kin: A molecular species-level phylogeny of Carnivora. *Molecular Phylogenetics and Evolution*, 54 (3): 726-745.

19 Polly PD 2001. Paleontology and the comparative method: Ancestral node reconstructions versus observed node values. *American Naturalist*, 157 (6): 596-609.

20 Finarelli JA & Flynn JJ 2009. Brain-size evolution and sociality in Carnivora. *Proceedings of the National Academy of Sciences USA*, 106 (23): 9345-9.

21 Wolsan M & Lange-Badré B 1996. An arctomorph carnivoran skull from the Phosphorites du Quercy and the origin of procyonids. *Acta Palaeontologica Polonica*, 41 (3): 277-298.

22 Gatesy J, Geisler JH, Chang J, Buell C, Berta A, Meredith RW, Springer MS & Mcgowen MR 2013. A phylogenetic blueprint for a modern whale. *Molecular Phylogenetics and Evolution*, 66 (2): 479-506.

23 Mennecart B, Scherler L, Hiard F, Becker D & Berger J-P 2012. Large mammals from Rickenbach (Switzerland, reference locality MP29, Late Oligocene): biostratigraphic and palaeoenvironmental implications. *Swiss Journal of Palaeontology*, 131 (1): 161-181.

24 Orliac M & Gilissen E 2012. Virtual endocranial cast of earliest Eocene Diacodexis (Artiodactyla, Mammalia) and morphological diversity of early artiodactyl brains. *Proceedings of the Royal Society of London, Series B*, 22 (27): 3670-3677

25 Montgomery SH, Geisler JH, Mcgowen MR, Fox C, Marino L & Gatesy J 2013. The evolutionary history of cetacean brain and body size. *Evolution*, 67 (11): 3339-3353.

26 Marino L, Mcshea DW & Uhen MD 2004. Origin and evolution of large brains in toothed whales. *Anatomical Record Part A*, 281A (2): 1247-1255.

27 Bibi F 2013. A multi-calibrated mitochondrial phylogeny of extant Bovidae (Artiodactyla, Ruminantia) and the importance of the fossil record to systematics. *BMC Evolutionary Biology*, 13 (166).

28 Costeur L, Maridet O, Peigné S & Heizmann EPJ 2012. Palaeoecology and palaeoenvironment of the Aquitanian locality Ulm-Westtangente (MN2, Lower Freshwater Molasse, Germany). *Swiss Journal of Palaeontology*, 131 (1): 183-199.

29 Tellgren A, Berglund AC, Savolainen P, Janis CM & Liberles DA 2004. Myostatin rapid sequence evolution in ruminants predates domestication. *Molecular Phylogenetics and Evolution*, 33 (3): 782-90.

30 Solounias N, Barry JC, Bernor RL, Lindsay EH & Raza SM 1995. The oldest Bovid from the Siwaliks, Pakistan. *Journal of Vertebrate Paleontology*, 15 (4): 806-814.

31 Thewissen JG, Cooper LN, Clementz MT, Bajpai S & Tiwari BN 2007. Whales originated from aquatic artiodactyls in the Eocene epoch of India. *Nature*, 450 (7173): 1190-4.

32 Gunnell GF & Simmons NB 2005. Fossil evidence and the origin of bats. *Journal of Mammalian Evolution*, 12 (1-2): 209-246.

33 Simmons NB & Geisler JH 1998. Phylogenetic relationships of *Icaronycteris*, *Archaeonycteris*, *Hassianycteris*, and *Palaeochiropteryx* to extant bat lineages, with comments on the evolution of echolocation and foraging strategies in Microchiroptera. *Bulletin of the American Museum of Natural History*, 235.

34 Eiting TP & Gunnell GF 2009. Global Completeness of the Bat Fossil Record. *Journal of Mammalian Evolution*, 16 (3): 151-173.

35 Yao L, Brown JP, Stampanoni M, Marone F, Isler K & Martin RD 2012. Evolutionary change in the brain size of bats. *Brain Behavior and Evolution*, 80 (1): 15-25.

36 Rydell J & Speakman JR 1995. Evolution of nocturnality in bats: Potential competitors and predators during their early history. *Biological Journal of the Linnean Society*, 54: 183-191.

37 Meredith RW, Westerman M, Case JA & Springer MS 2008. A phylogeny and timescale for marsupial evolution based on sequences for five nuclear genes. *Journal of Mammalian Evolution*, 15 (1): 1-36.

38 Travouillon KJ, Legendre S, Archer M & Hand SJ 2009. Palaeoecological analyses of Riversleigh's Oligo-Miocene sites: implications for Oligo-Miocene climate change in Australia. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 276 (1): 24-37.

39 Beck RMD 2008. A dated phylogeny of marsupials using a molecular supermatrix and multiple fossil constraints. *Journal of Mammalogy*, 89 (1): 175-189.

40 Wilson GP 2013. Mammals across the K/Pg boundary in northeastern Montana, USA: dental morphology and body-size patterns reveal extinction selectivity and immigrant-fueled ecospace filling. *Paleobiology*, 39 (3): 429-469.

41 Brammall J 1999. A new petauroid possum from the Oligo-Miocene of Riversleigh, northwestern Queensland. *Alcheringa*, 23 (1): 31-50.

42 Wilf P, Beard KC, Davies-Vollum KS & Norejko JW 1998. Portrait of a late Paleocene (early Clarkforkian) terrestrial ecosystem: Big multi quarry and associated strata, Washaki Basin, Southwestern Wyoming. *PALAIOS*, 13 (6): 514-532.

43 Rose KD, Deleon VB, Missiaen P, Rana RS, Sahni A, Singh L & Smith T 2008. Early Eocene lagomorph (Mammalia) from Western India and the early diversification of Lagomorpha. *Proceedings of the Royal Society of London, Series B*, 275: 1203-1208.

44 Ni X, Gebo DL, Dagosto M, Meng J, Tafforeau P, Flynn JJ & Beard KC 2013. The oldest known primate skeleton and early haplorhine evolution. *Nature*, 498 (7452): 60-64.

45 Jaeger JJ, Beard KC, Chaimanee Y, Salem M, Benammi M, Hlal O, Coster P, Bilal AA, Duringer P, Schuster M, Valentin X, Marandat B, Marivaux L, Metais E, Hammuda O & Brunet M 2010. Late middle Eocene epoch of Libya yields earliest known radiation of African anthropoids. *Nature*, 467 (7319): 1095-8.

46 Fleagle JG 1999. *Primate adaptation and evolution*, State University of New York, Academic Press.

47 Stevens NJ, Seiffert ER, O'Connor PM, Roberts EM, Schmitz MD, Krause C, Gorscak E, Ngasala S, Hieronymus TL & Temu J 2013. Palaeontological evidence for an Oligocene divergence between Old World monkeys and apes. *Nature*, 497 (7451): 611-614.

48 Kirk EC & Simons EL 2001. Diets of fossil primates from the Fayum Depression of Egypt: a quantitative analysis of molar shearing. *Journal of Human Evolution*, 40 (3): 203-29.

49 Boyer DM, Seiffert ER, Gladman JT & Bloch JI 2013. Evolution and Allometry of Calcaneal Elongation in Living and Extinct Primates. *PLoS ONE*, 8 (7): e67792.

50 Begun DR 2013. *The Miocene Hominoid Radiations*. In: Begun DR (ed.) A Companion to Paleoanthropology. Wiley Online Library: Blackwell Publishing Ltd.

51 Hartwig WC 2002. *The primate fossil record*, Cambridge, Cambridge University Press.

52 Seiffert ER, Simons EL & Attia Y 2003. Fossil evidence for an ancient divergence of lorises and galagos. *Nature*, 422 (6930): 421-424.

53 Gingerich PD 1984. Primate Evolution: Evidence from the fossil record, comparative morphology, and molecular biology. *Yearbook of Physical Anthropology*, 27: 57-72.

54 Harrison T 2010. *Later Tertiary Lorisiformes*. In: Werdelin L & Sanders WJ (eds.), Cenozoic mammals of Africa. Berkeley: University of California Press.

55 Tabuce R, Marivaux L, Lebrun R, Adaci M, Bensalah M, Fabre P-H, Fara E, Rodrigues HG, Hautier L & Jaeger J-J 2009. Anthropoid versus strepsirhine status of the African Eocene primates Algeripithecus and Azibius: craniodental evidence. *Proceedings of the Royal Society of London, Series B,* 276 (1679): 4087-4094.

56 Zalmout IS, Sanders WJ, Maclatchy LM, Gunnell GF, Al-Mufarreh YA, Ali MA, Nasser AA, Al-Masari AM, Al-Sobhi SA, Nadhra AO, Matari AH, Wilson JA & Gingerich PD 2010. New Oligocene primate from Saudi Arabia and the divergence of apes and Old World monkeys. *Nature*, 466 (7304): 360-364.

Chapter 2

Multiple evolutionary radiations lead to smaller testes sizes in vertebrates

Abstract

The size of vertebrate testes is one measure of how much a male invests in reproduction. Like many organs, testes size is strongly linked to the body size of an animal yet even after considering the effects of body size there is considerable variation in testes sizes among species. Identifying instances of rapid evolution in testes size that occur beyond changes in body size may reveal how such diversity has arisen. We find that the vertebrate tree of life has been punctuated by many intense radiations acting to increase variation of testes sizes. These radiations occurred in groups of varying sizes ranging from neognath birds and boreoeutherian mammals right down to individual genera. Ultimately, these multiple bursts of rapid variation combined to form an overwhelming tendency for testes sizes to have adaptively decreased in size over more than 400 million years of vertebrate evolutionary history. The negative trend that we detect demonstrates that it is possible to reveal patterns of biological evolution that might otherwise be masked by strong associations with other traits such as body size. The adaptive significance of decreasing investment into reproductive output might reflect increasing investment into other energetically expensive tissues or the evolution of progressively more diverse and complex mating strategies.

Introduction

Testes size varies widely among vertebrates, even after considering its underlying relationship with body size (Figure 1; Parker, 2016). Decades of research have sought to explain this variation among taxa ranging from entire classes such as mammals, birds, reptiles, and fish (Garamszegi *et al.*, 2005; Jetz and Freckleton, 2015; Kahrl *et al.*, 2016; Kenagy and Trombulak, 1986; Møller, 1988b, 1989; Møller and Briskie, 1995; Pitcher *et al.*, 2005; Soulsbury, 2010; Stockley *et al.*, 1997) right down to individual species (e.g. Bidau and Medina, 2013; Ginsberg and Rubenstein, 1990; Harris and Moore, 2005; Heske and Ostfeld, 1990; Jetz and Freckleton, 2015; Newlin, 1976; Ota *et al.*, 2012; Tomkins and Simmons, 2002).

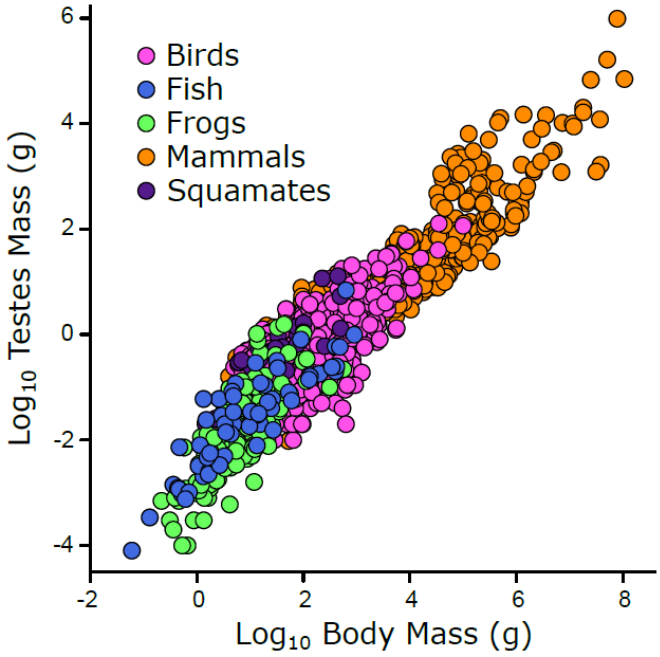


Figure 1: Testes size and body size for 1,872 vertebrate species. Colours indicate major vertebrate clades (see legend). r^2 = 0.21 from a phylogenetic generalized least squares (GLS) regression across this data.

Species with high levels of sperm competition, i.e. those where females mate with multiple males, are expected to invest more heavily in the size and development of their testes than those with lower levels of sperm competition such as monogamous or polygynous species (Birkhead and Møller, 1998; Harcourt *et al.*, 1981; Harcourt *et al.*, 1995; Parker, 1970). This relationship holds in many groups of animals (reviewed in Parker *et al.*, 1997;

Simmons and Fitzpatrick, 2012), ranging from birds (Møller, 1994; Møller and Briskie, 1995) to bats (Hosken, 1997, 1998) to butterflies (Gage, 1994).

However, not all testes size variation among species can be explained by sperm competition. In some cases multi-male mating systems and increased risk of sperm competition have no (or at least no substantial) effect on testes size (e.g. Byrne *et al.*, 2002; Iossa *et al.*, 2008; Kappeler, 1997; Pyron, 2000; Simmons and Fitzpatrick, 2012). As expected then, many other factors have been shown to be important drivers of testes size differences among species (e.g. Dunn *et al.*, 2001; Heske and Ostfeld, 1990; Iossa *et al.*, 2008; Kenagy and Trombulak, 1986; Ota *et al.*, 2012; Pitcher *et al.*, 2005). For example, migration habits and breeding location have been shown to strongly influence testes size (Dunn *et al.*, 2001; Pitcher *et al.*, 2005) after accounting for both body size and sperm competition. In mammals, there is some evidence that seasonal breeding plays a role (Heske and Ostfeld, 1990; Iossa *et al.*, 2008; Kenagy and Trombulak, 1986).

Research seeking to understand testes evolution on a larger scale is sparse (Hayward and Gillooly, 2011; MacLeod, 2014; though see MacLeod and MacLeod, 2009). Theoretically there is no reason to believe that testes size would not conform to a simple linear relationship with body size (Hayward and Gillooly, 2011). However, others have suggested that the relationship between testes size and body size is far more complex than one might expect from the results of simple analyses (MacLeod, 2014; MacLeod and MacLeod, 2009), differing across animals of different sizes. MacLeod (2014) speculates that, at least in amniotes, there are size-related biases in reproductive investment that may be driven by energetic costs required to maintain large testes in combination with the necessary minimum size required for fertilization success. In such a scenario, proportionally larger testes in the largest amniotes do not provide any further reproductive advantages.

Testes – like many other organs (Peters, 1986) – are energetically expensive to develop and maintain (Kenagy and Trombulak, 1986; Schulte-Hostedde *et al.*, 2005; Simmons and Emlen, 2006) yet increasing testes size relative to body mass is one way to provide an

animal with a greater probability of contributing genetic material to the next generation (Harcourt *et al.*, 1981; see Ramm and Schärer, 2014). This indicates that changes in the size of vertebrate testes are necessarily a product of natural selection – especially where such changes have led to size increases. To the extent that this is true, prior suggestions that it is not necessary to study testes size evolution within a phylogenetic context (MacLeod, 2014; MacLeod and MacLeod, 2009) may be unfounded.

Even if changes in testes sizes are strongly associated with other morphology such as body size, we must still be able to study this phylogenetically: species cannot be considered as statistically independent owing to millions of years of shared ancestry (Felsenstein, 1985; Harvey and Pagel, 1991). In fact, we expect there to be strong phylogenetic signal in the residual variance between two evolving biological traits (Felsenstein, 1985; Freckleton, 2002; Revell, 2010). Testes sizes, like all morphologies, are shaped by natural selection (e.g. Simpson, 1953) and we can study such adaptation phylogenetically (Baker *et al.*, 2016; Blomberg and Garland Jr, 2002; Harvey and Pagel, 1991; Pagel 1994; Chapters 4 and 5) .

A new generation of phylogenetic comparative methods mean that it is now not only possible to test if there is phylogenetic signal, but also automatically identify instances of rate variation throughout evolutionary history (Eastman *et al.*, 2011; Kratsch and McHardy, 2014; Kutsukake and Innan, 2013, 2014; Rabosky, 2014; Revell *et al.*, 2012; Thomas and Freckleton, 2012; Venditti *et al.*, 2011). We can reveal where in time and in which lineages that rapid changes in testes sizes may have been linked to instances of evolutionary innovation and the accumulation of biodiversity (Baker *et al.*, 2016, Chapters 4 and 5). Rapid changes in morphology indicate periods of adaptive evolution (Eastman *et al.*, 2011; Kratsch and McHardy, 2014; Kutsukake and Innan, 2013, 2014; Rabosky, 2014; Revell *et al.*, 2012; Thomas and Freckleton, 2012; Venditti *et al.*, 2011); when these adaptive shifts are played out across the branches of a phylogenetic tree it is possible for them to combine and give rise to long-term directional evolutionary change (Baker et al. 2015).

We bring the study of testes size evolution into a phylogenetic context across all vertebrates for the first time, determining whether variation in testes size can be explained using a simple linear relationship with body size or if the pattern is more complex across vertebrate species. We also test the idea that it may be possible to detect adaptive trends in morphology after accounting for the association with body size or any other underlying trait. Although increasing the size of testes beyond body size may be beneficial in some situations, it is also likely that reducing investment in expensive reproductive tissues may have been advantageous during the course of vertebrate evolution, allowing for investment in other energetically expensive morphologies or behaviours. We identify where there have been rapid bursts of testes size evolution over the last 400 million years of vertebrate history and determine whether these have combined to give rise to sustained directional change in relative investment into male reproductive organs.

Methods

Data and phylogenetic tree

We collected testes mass and body mass data across all major vertebrate clades from the literature (Figure 1). We used the unsmoothed version of the phylogeny of the time tree of life limited to vertebrates for all analyses (Hedges *et al.*, 2015); we used this version as it retains polytomies and allowed for a more conservative estimation of rate heterogeneity and any underlying relationship in the data (Venditti *et al.*, 2011).

Species were matched to the time tree of life using major taxonomic references (Appendix 1). We retained only major clades of vertebrates where sample size was large enough for subsequent analyses ($n \geq 30$). The final dataset contained testes mass (g) and male body mass (g) for 1,872 vertebrate species (61 Actinopterygii, 176 Anura, 983 Aves, 621 Mammalia, and 31 Squamata). All variables were $\log_{10}$ transformed before analysis (Figure 1).

Heterogeneity in testes size evolution

The *variable rates model* (Venditti *et al.*, 2011) works within a phylogenetic generalised least squares (GLS) framework to estimate the rate of morphological evolution along the branches of a phylogeny assuming an underlying homogenous Brownian motion process ($\sigma^2_{\text{Brownian}}$). Simultaneously, the model identifies areas of the tree where the phenotypic evolution exceeds (or is less than) that expected given this homogenous variance and adjusts branch lengths to represent such shifts in the rate of evolution. This results in a scaled phylogeny where branches that have been stretched compared to those measured in time indicate a faster rate of morphological evolution, and compressed branches have a slower rate. If we were to replay evolution across this scaled tree, there is more opportunity for change in lineages with faster rates of evolution, i.e. longer branches, within the same evolutionary timeframe.

Here we use a variant of the variable rates model: the *variable rates regression model* (Baker *et al.*, 2016) which detects shifts in the rate of evolution of the phylogenetically structured residual error of a regression model. The model scales the branches of a phylogenetic tree such that they are relative to the rate of evolution of testes size through time, given its underlying relationship with body size.

Positive evidence for rate heterogeneity was determined using a Bayes Factor (BF), calculated as $BF = -2\log_e[m_1/m_0]$, comparing the marginal likelihood of our variable rates model (m_1) to that of a model with no variable rates (m_0) i.e. with a single homogenous underlying Brownian motion rate of evolution. Marginal likelihoods were estimated using a stepping-stone sampler (Xie *et al.*, 2010) as implemented in BayesTraits (Pagel *et al.*, 2004). For each of 200 stones, we ran 1 million iterations drawing values from a beta-distribution (α = 0.4, β = 1) (Xie *et al.*, 2010) and discarded the first 250,000 iterations as burn-in.

We compare the posterior distribution of estimated rates of ancestor-descendant branch pairs. Where a branch differs in rate from its ancestor in ≥ 50% of the posterior distribution

but there is no subsequent rate shift in any descendant branches (Figure 2), it is identified as a *heritable rate shift*. This way, entire clades can be identified as evolving at a different rate of evolution as compared to their parent clade and in the same way, sub-clades within such groups can also be identified. Where we observe heritable rate shifts, this implies an increase in variance throughout the entire clade – the clade has more variation about the regression line than would be expected given the underlying relationship between testes size and body size.

In the context of large datasets spanning large taxonomic ranges and containing many hundreds of species, it is then possible that there may be nested bursts of rates giving rise to differences in the underlying evolutionary relationship (Figure 2) – that they have a fundamentally different *evolutionary trajectory*. In the case of vertebrate testes sizes, we are unlikely to observe a negative relationship with body size although there may be some variation in the magnitude of the positive evolutionary trajectory among groups (e.g. Hayward and Gillooly, 2011; Kenagy and Trombulak, 1986; MacLeod and MacLeod, 2009). Differences in evolutionary trajectories observed among groups of species have the potential to affect our inferences of rate variation (see Chapter 5). Therefore, in order to confirm that the heritable shifts in rate we observe do not arise as a consequence of differences in the relationship between testes size and body size, we ran an additional variable rates regression model that allowed each of the groups large enough to analyse separately ($n \geq 20$) (Freckleton and Watkinson, 2001) that are identified as part of a clade shift (e.g. the red clade in Figure 2) to have an independent evolutionary trajectory (their own intercept and slope). Where groups are nested, the relationship is compared to that observed in its parent clade.

We assessed significance of regression parameters using the proportion of the posterior distribution that crosses zero (P_x). This proportion can be interpreted in a similar way to a traditional P value: where less than 5% of a posterior crosses zero, $P_x < 0.05$ and that variable is estimated to be significantly different from zero. To assess whether two parameters significantly differed from one another, we conducted post-hoc comparisons

between the differences of two parameters at each iteration. Where less than 5% of the posterior distribution of differences crosses zero (P_x < 0.05), two parameters are considered to be distinct. Where a relationship is found to not significantly differ to its parent clade, the shift in rate can therefore be interpreted as a shift in variance about the underlying evolutionary relationship between testes size and body size (Figure 2).

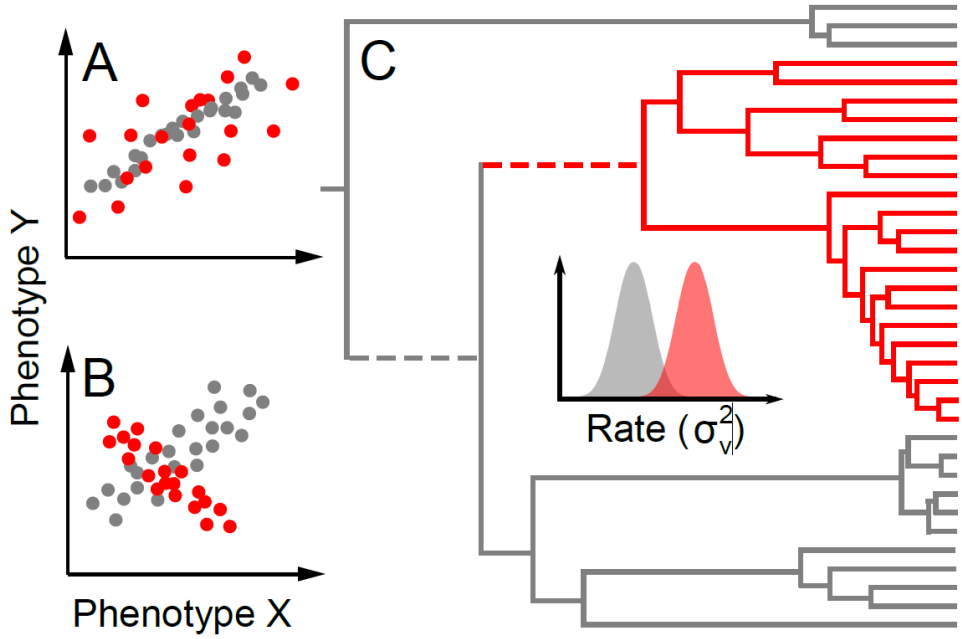


Figure 2: An example of how we identify and interpret rate heterogeneity in the form of clade shifts given two hypothetical datasets (A) and (B) and a phylogenetic tree (C). In (A), a monophyletic group of species (red) has increased residual variance in the relationship between two traits (Phenotypes Y and X). In (B) the same clade has a different relationship in terms of directionality. If we run the variable rates regression model estimating a linear relationship between Phenotypes Y and X across both datasets, we would obtain identical results: it is impossible to distinguish between these scenarios using rate heterogeneity alone. Post-hoc tests comparing the relationship within such groups to their parent clade are therefore necessary in order to explain these clade shifts. More than 50% of the posterior distribution of the estimated rate along the branch leading to the red clade (dashed, red) is greater than that on its immediately ancestral lineage (dashed, grey). This is a heritable rate shift: all descendant branches do not differ in rate to their ancestor – the entire clade inherits a new rate of evolution.

A *single branch modification,* where a rate change is observed along a branch that is not inherited by its daughter clade, indicates that the group of species descending from this branch differ in the intercept of the underlying modelled relationship. Here, descendant

branches would not differ in rate from the branches immediate preceding that upon which the single-branch shift was identified.

Identifying trends in testes size evolution

Our method of detecting rate heterogeneity introduces meaningful variation into the branch lengths of an extant phylogeny, which makes it possible to study adaptive trends in trait evolution (Baker *et al.*, 2015). The root-to-tip distance for each species on the scaled tree (henceforth *path-wise rate*) output by the variable rates regression model gives us a measure of total amount of change in the rate of testes size evolution a species has been subject to during its evolutionary history. We use this measure and the methods of Baker *et al.* (2015) to quantify long-term adaptive trends in testes size evolution across the history of the vertebrate radiation. We test to see if these path-wise rates are linked to testes size: a positive relationship would indicate that on average, adaptive evolution has led to larger testes.

Results and Discussion

We find that testes mass evolution is best described by a variable rates regression model allowing a different slope within each of the 5 major vertebrate clades. There is significant rate heterogeneity in this model (BF = 442.51). Fish, frogs, mammals and birds all have distinct evolutionary trajectories – they differ significantly in their relationship between testes size and body size (Figure 3) and so we prefer this model over one that estimates only a single slope over all species (the slope for squamates is poorly estimated, perhaps owing to small sample size [n = 31] and overlaps all other groups, Figure 3).

On the bases of both theoretical considerations and previous analyses (Hayward and Gillooly, 2011), we would predict testes sizes to have a simple linear relationship with body size much like that observed in other organs (Peters, 1986). However, MacLeod has suggested that the relationship between testes size and body size is much more complex – and that species of different sizes have different relationships (MacLeod, 2010, 2014;

MacLeod and MacLeod, 2009), resulting in an overall non-linear relationship. If the testes sizes of large-, intermediate-, and small-bodied species (however these might be defined) differ in relative investment into testes sizes independently and regardless of phylogenetic ancestry, then we should find a cubic relationship between testes size and body size. To test this, we ran an additional variable rates regression model that estimates a cubic curvature in the relationship between testes size and body size. Although we do find a significant cubic parameter (P_x = 0.01) in a model that estimates a single slope over all vertebrates, when we allow for different slopes among the major clades, the parameter provides only a negligible a mean R^2 improvement (just 0.003 when we compare a model with and without the cubic parameter), and is non-significant (P_x = 0.07). We therefore find limited evidence to support the suggestion that species of different sizes experience different evolutionary trajectories.

MacLeod and MacLeod (2009) found that a single slope could explain most variation in amniote species. In contrast, we find significant differences between each of the major vertebrate clades studied with the exception of squamates (Figure 3). Our results are more in line with a recent analysis by Hayward and Gillooly (2011) which found significant differences among most animal groups.

One of the main reasons that our analyses differ from previous work may arise from the fact that all research looking at testes evolution at large scales to date has been conducted without taking into consideration shared ancestry (Hayward and Gillooly, 2011; MacLeod, 2014; MacLeod and MacLeod, 2009). However, significant differences in the slope of the relationship between testes size and body size between major monophyletic animal clades (Figure 3, Hayward and Gillooly, 2011; MacLeod and MacLeod, 2009) represent fundamental differences in evolutionary trajectory explicitly attributed to phylogenetic relatedness. The nature of our phylogenetic analysis provides another advantage over previous analyses in that it is possible to detect and identify species or groups of species that deviate from the underlying relationships between testes size and body size in the form of the significant evolutionary rate heterogeneity we find in our model.

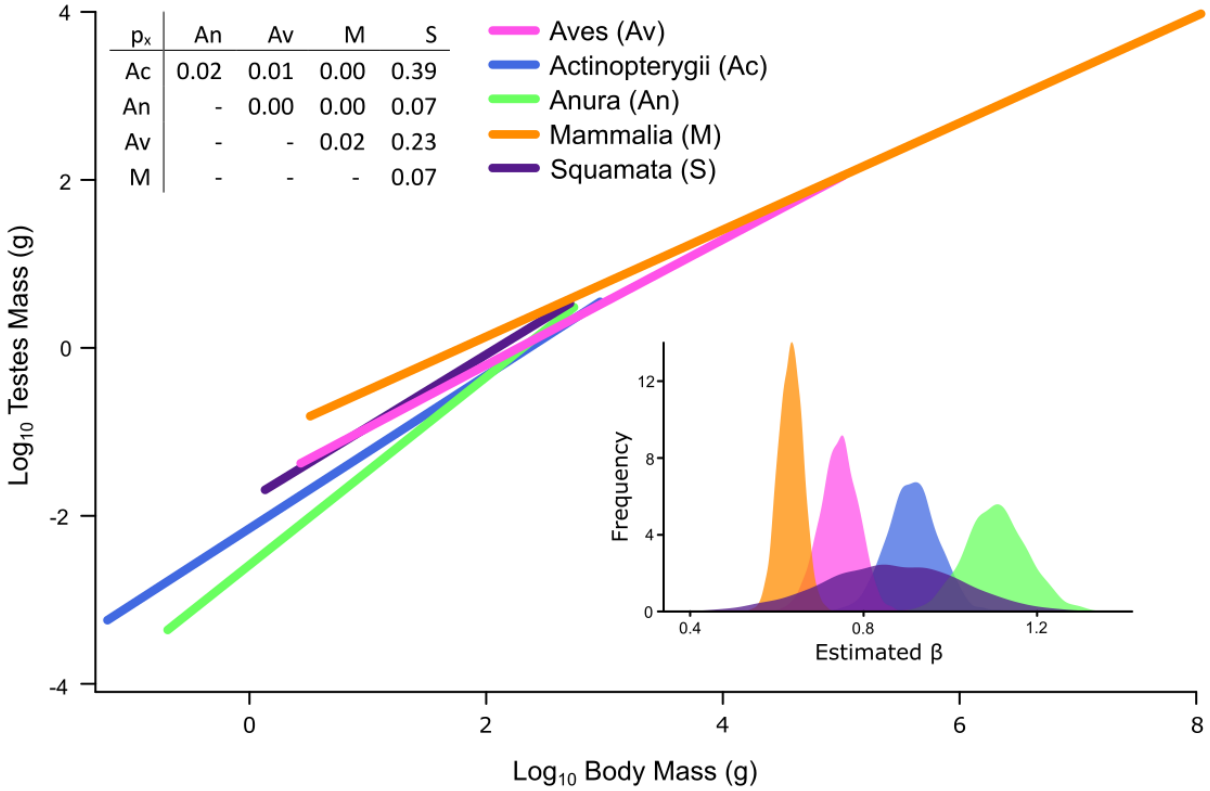


Figure 3: The relationship between testes size and body size across different groups of vertebrates (indicated by different colours) as estimated in the variable rates regression analysis. For each of the major vertebrate clades, we conduct post-hoc pair-wise comparisons assessing the proportion of the posterior distribution of estimated β coefficients that overlap. The table of pair-wise P_x comparisons are shown inset (top left). Where P_x < 0.05, the posterior distribution of the estimated for each of the two groups are considered to be significantly different from one another. All posteriors are also shown inset (bottom right).

The significant rate heterogeneity we find in our model arises in the form of 21 unique *heritable rate shifts* littered throughout the vertebrate phylogeny (Table 1). There are 9 shifts that represent independent episodes of elevated rates of testes size evolution away from the background rate of evolution (Figure 4) – we identify these as *primary shifts*. All other shifts are nested within these (see Table 1) and are termed *nested shifts*. The data and phylogenetic position for all groups where we identify heritable shifts are plotted in Appendix 2.

Table 1: Vertebrate heritable rate shifts. Modal rate increase (r) calculated using kernel density estimation. More details can be found in in Appendix 2.

Group	Species included	n	r
Birds	Neognathae (All birds excluding ratites and tinamous)	979	6.00
	Pipilo chlorurus, *P. erythrophthalmus*, *P. maculatus* and *P. ocai* (Rufous-sided towhee complex)	4	11.94
	Corvus (crows and ravens excluding *C. monedula)*	7	32.43
	Oriolus flavocinctus and *O. szalayi* (Green and Brown oriole)	2	71.57
	Gerygone (peep-warblers excluding *G. cinerea*)	7	41.67
	Alcedines (Kingfishers)	8	6.60
	Charadriidae (Plovers and oystercatchers)	18	7.27
	Oxyura (Stiff-tailed ducks)	2	45.89
Squamates	*Laticauda colubrina*, *Notechis scutatus*, *Pseudonaja textilis* (3 Members of the snake family Elapidae)	3	8.20
Mammals	Boreoeutheria (placental mammals excluding Afrotheria and Xenarthra)	564	5.08
	Cetacea (Whales and dolphins)	58	21.43
	Ovis (sheep)	3	48.30
	Equus grevyi, E. burchellii and *E. zebra* (Zebras)	3	52.94
	7 species of the megabat family Pteropodidae	7	7.29
	Pseudomys (false mice, excluding *P. fumeus*)	10	7.87
Frogs	Bufonoidea (frog superfamily)	124	6.74
	Rana and *Odorana*	10	7.62
	Dicroglossidae and Rhacophoridae (fork-tongued and foam-nesting frogs)	16	6.37
Fish	A small clade within the Syngnathidae family (pipefish)	12	1.46
	Labridae (parrotfish)	10	4.22
	Nocomis and *Campostoma*	5	11.32

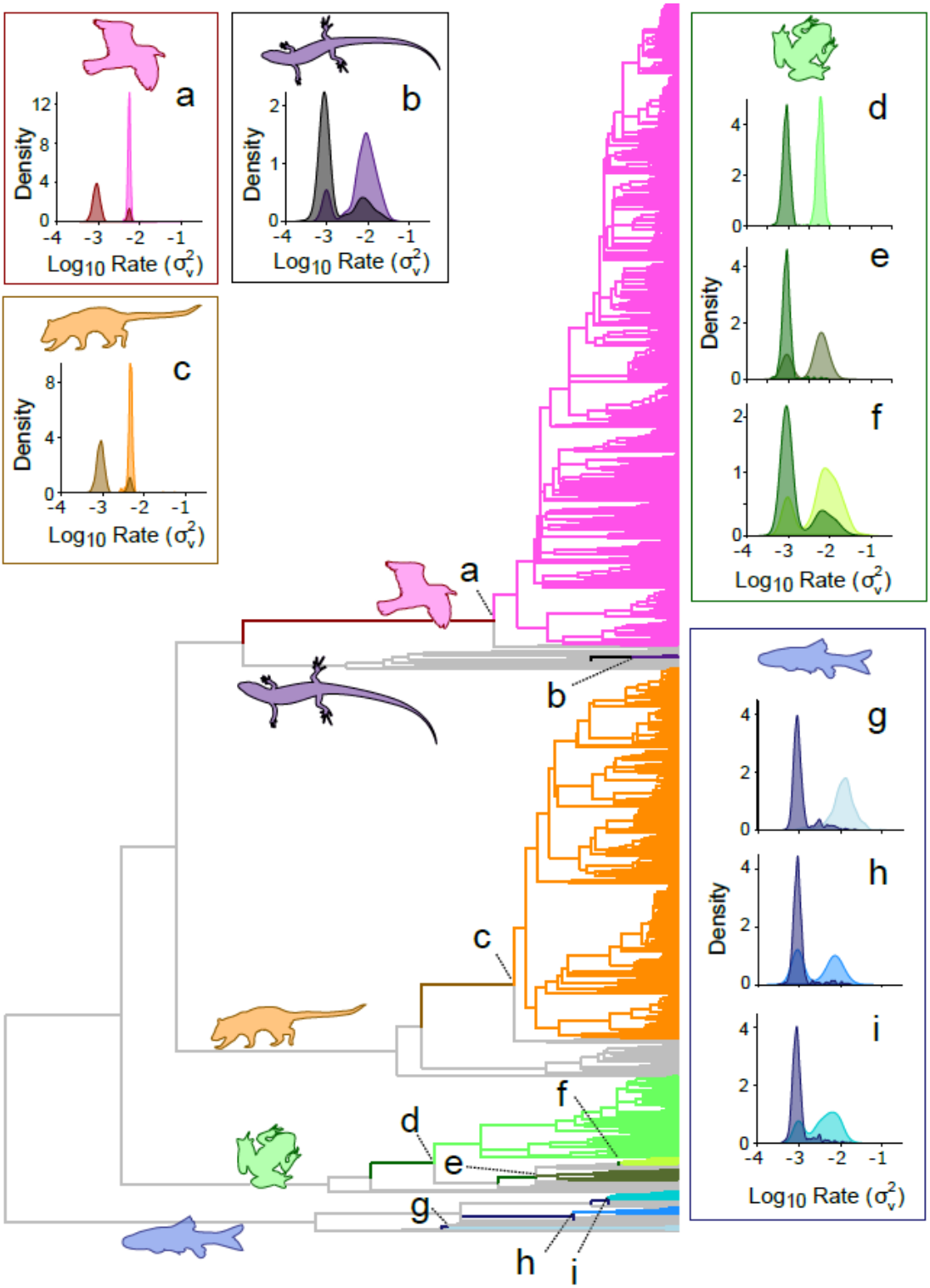


Figure 4: Nine primary heritable rate shifts. The distributions compare the rate along the branch leading to the entire clade (bright colours) to that observed giving rise to the ancestral lineage (dark colours). Each shift is indicated by a different colour and letter. (a) neognath birds; (b) *Laticauda colubrina, Notechis scutatus*, and *Pseudonaja textilis*; (c) boreoeutherian mammals; (d) Bufonoidea; (e) *Rana* and *Odorana*; (f) Dicroglossidae and Rhacophoridae; (g) pipefish; (h) parrotfish; (i) *Nocomis* and *Campostoma*.

In no case do we find any evidence that testes sizes have a different evolutionary trajectory in those groups identified as heritable shifts compared to the group within which they are nested (Appendix 2). All heritable rate shifts can therefore be interpreted as increases in variance (Figure 2). These variance increases are scattered across the vertebrate phylogeny ranging from huge groups like boreoeutherian mammals (all placental mammals excluding Afrotheria and Xenarthra) and neognath birds (all birds excluding ratites and tinamous) all the way to species of a single genus or family such as zebras, crows, parrotfish or foam-nesting frogs (Table 1). Each represent a distinct radiation of testes size.

The largest and most widespread rate increases we observe encompass boreoeutherian mammals and neognath birds (Figure 4); both very diverse groups in terms of their biology and ecology. Specifically, the proportion of socially monogamous species in birds is higher than in any other group of vertebrate species – Pitcher *et al.* (2005) classify 75% of the 1,031 species in their dataset as monogamous although estimates of the proportion of birds that exhibit social monogamy go as high as 90% (Cockburn, 2006; Lack, 1968). This is followed by some considerable margin in mammals where just ~9% of all species (Lukas and Clutton-Brock, 2013) are classified as monogamous and even then is more frequent in some orders than in others (Opie *et al.*, 2013). Monogamy in other vertebrates, though it does exist (e.g. Bull, 2000; Takegaki, 2000; Whiteman and Côté, 2003; Whiteman and Côté, 2004) is much rarer (Bull, 2000; Mock and Fujioka, 1990). An increased variability of mating strategies might provide means of escape from the costs of investing heavily into reproductive tissues. Sperm competition explains much of the variation in testes sizes – particularly within mammals and birds (e.g. Ginsberg and Rubenstein, 1990; Harcourt *et al.*, 1981; Hosken, 1997, 1998; Møller, 1988a, 1989, 1994; Møller and Briskie, 1995, though see Iossa et al. 2008). It therefore follows that the variance increases that we detect throughout vertebrate evolution and specifically within these groups might represent radiations in relative reproductive investment triggered by the evolution of more diverse biological and ecological characteristics allowing for less reliance on high investment in testes.

As an example, one of the heritable rate shifts we observe is within cetaceans (Table 1, Figure 5). Mating systems are notoriously elusive for this group of species (Mann *et al.*, 2000) – and are often inferred simply from measurements of testes size (e.g. Brownell and Ralls, 1986; Dines *et al.*, 2014). It has previously been argued that non-linearity in the relationship between testes size and body size can lead to misleading inferences of mating systems specifically regarding cetaceans (MacLeod, 2010). However, our results demonstrate that the testes sizes of these species are neither larger (also found by Kenagy and Trombulak, 1986) nor subject to a different evolutionary trajectory when compared to the testes of other mammals (Figure 5, Appendix 2). Instead, testes sizes within cetaceans conform to those expected from the relationship observed within other mammals – given the magnitude of the rate increase that we detect (on average, 21.43 times higher than that within boreoeutheria).

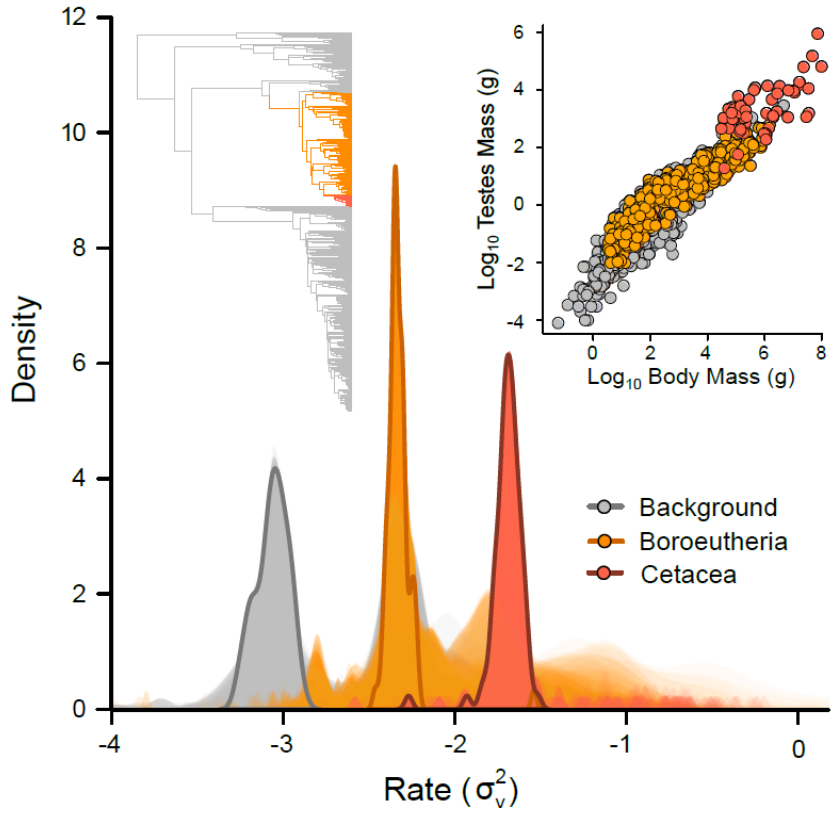


Figure 5: Rate shift in cetaceans. The posterior distribution of rates are shown for 3 subsets of branches: those evolving at the background rate; boreoeutherian mammals; and cetaceans. The basal branch of each group is outlined; the heritable nature of shifts can be seen as lack of variation away from the ancestral distribution. Inset are the phylogeny and data.

However, similarity in the relationship does not preclude a difference in process: cetaceans are evolving at a highly elevated rate which implies some trigger for intense adaptive change in testes sizes. Upon entering the water, cetacean species have experienced a novel set of selective pressures that differs from that observed in any other terrestrial mammal group. This may indicate that cetaceans have considerable diversity in mating systems and other factors that are important underlying factors of testes size variation in terrestrial mammals and other vertebrates (Dines *et al.*, 2015; Kelley *et al.*, 2014) – or even that there may be entirely different factors driving testes size evolution in these species.

Interestingly, we find very limited evidence for mean shifts in testes size throughout the course of vertebrate evolution - instances where the testes size of all species of an entire clade is either increased or decreased in size compared to its relatives. A mean shift implies a change in the intercept of the regression relationship between testes size and body size, and in the variable rates model manifests as a rate increase along an internal branch (Baker *et al.*, 2016). We observe only a single such instance within mammals – along the branch leading to Australian hopping mice (genus *Notomys*). This genera has been noted to have remarkably small testes (Breed and Jason, 2000). *Notomys alexis* has been better studied than other species in this genus (e.g. Breed, 1981; Breed and Washington, 1991; Peirce and Breed, 2001) and is believed to be monogamous (Lukas and Clutton-Brock, 2013) with low levels of sperm competition (Breed and Jason, 2000; Happold, 1976). Breed and Jason (2000) suggest that the tiny testes observed in this genus may be associated with increased efficiency of sperm transport within the female reproductive tract of this species (Breed, 1997; Breed and Washington, 1991).

We detect a number of rate shifts in individual species of birds and mammals that, like a mean shift deeper within the phylogenetic tree, also represent bursts of directional change towards either larger or smaller testes. In these cases, the rapid bursts of evolution could represent shifts in testes sizes linked to unique behavioural or ecological characteristics favouring change in relative reproductive investment – but might also indicate an incipient

heritable rate shift wherein future species arising from that lineage would inherit the same rate of evolution.

Multiple radiations in testes size evolution (including those observed in tip branches) combine to introduce meaningful variation into the path length of extant phylogenies that can be used to study trends (Baker *et al.*, 2015). We find a significant negative relationship (β = -0.00028, P < 0.001) between testes size and path-wise rate using a maximum-likelihood phylogenetic GLS regression that simultaneously accounts for the various relationships observed with body size among major clades (Figure 6). The model does not improve when we allow the slope of the relationship to vary amongst vertebrate groups (likelihood ratio [D] test, compared to a model where the slopes do not vary: D = 6.997, P = 0.13, df = 4): there is a general negative trend across all species. We also find no evidence that species of different body sizes experience differences in rate by testing a model that includes an interaction between body size and path length (D = 2.54, P = 0.11, df = 1); this provides further counterevidence for the suggestion that large and small species have biologically meaningful differences in their testes size evolution as compared to intermediate-sized species (MacLeod, 2014; MacLeod and MacLeod, 2009).

Throughout vertebrate evolutionary history, there has been a general tendency for adaptive evolution to have driven testes sizes smaller and smaller, translating into a reduction in testes size of 0.99g per million years on average. We visualize the amount of adaptive change in testes size by predicting testes sizes given the lowest and highest path-wise rates observed within each group – where path-wise rate is higher, a lineage will have experienced more adaptive change. This way, we show the actual expected amount of change in testes size given both the minimal and maximal amount of adaptive evolution observed (Figure 6).

Adaptive trends can translate into substantial changes in morphology. For example, birds have experienced more adaptive evolutionary change in testes size compared to other vertebrates (Figure 6). Holding all else equal (and factoring out the variance associated

with phylogeny) the predicted testes size for a bird of average size given the minimum amount of adaptive evolution (3.87g) is considerably larger than a bird of the same size allowing for the maximum amount of adaptive change (0.09g) evolving over the same evolutionary timescale (Figure 6).

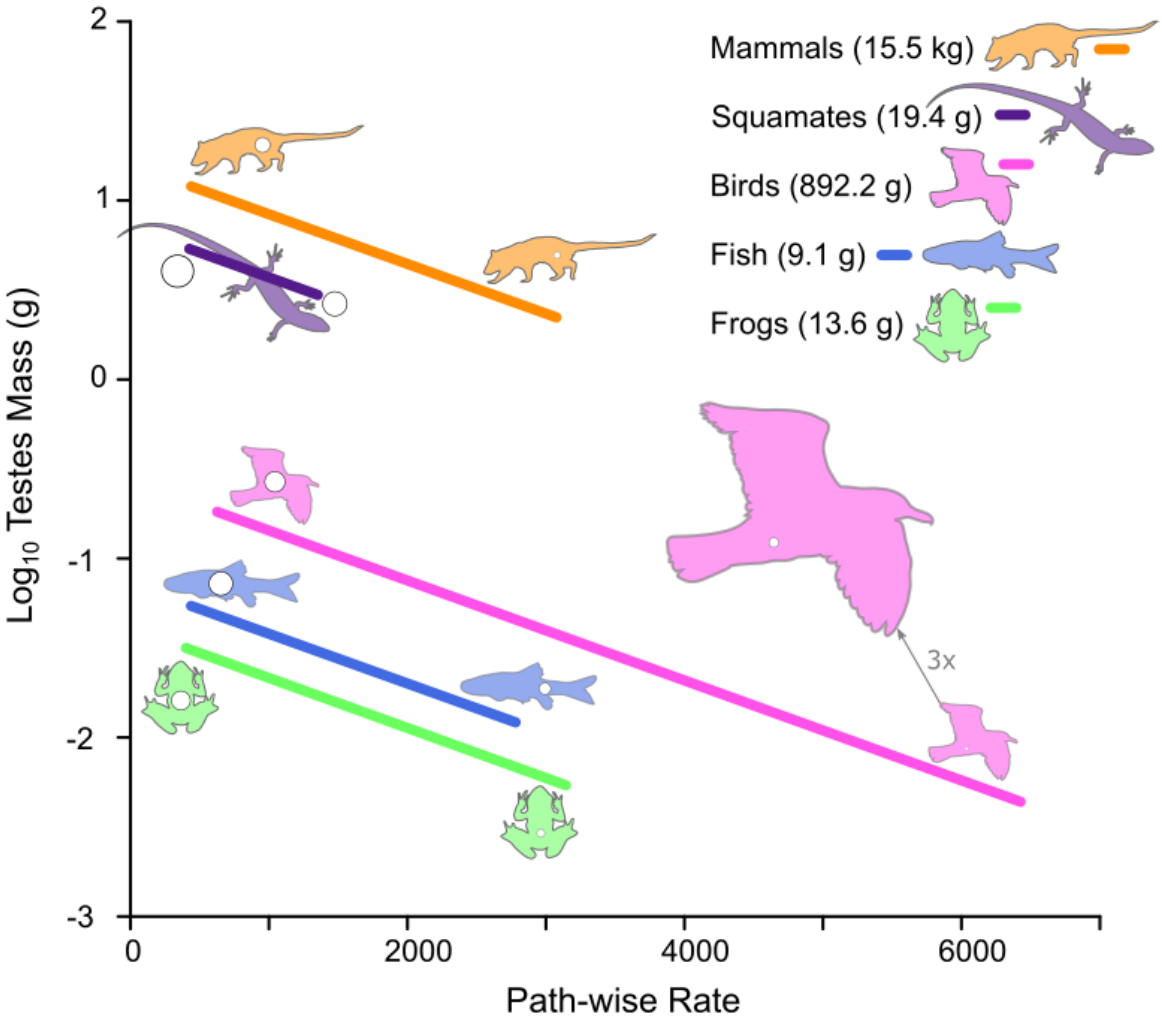


Figure 6: Rapid shifts in the rate of evolution across the vertebrate tree of life ultimately give rise to an overwhelming tendency for testes size reduction. This pattern is universal; there is no difference in the adaptive trend among major vertebrate clades (colours and silhouettes). The fitted slopes are phylogenetic predictions calculated holding the body size of each group at its mean (see legend), calculated using a phylogenetic GLS model (Pagel, 1999). The realized magnitude of testes size change is indicated by the silhouettes. Each animal silhouette is identical in area; the area of the circles representing the predicted testes sizes given the smallest and largest path-wise rate in each group are in proportion assuming that each silhouette represents the average body size observed in that group.

A similar pattern has been demonstrated in the case of mammal body size – where increases in body size over millions of years are the result of adaptive evolutionary change (Baker *et al.*, 2015). Larger body sizes confer a number of adaptive advantages across species (Brown and Sibly, 2006; Hone and Benton, 2005; Peters, 1986) but also to individuals within populations (Kingsolver and Pfennig, 2004, 2007). Over time, there is a general tendency for body size to increase in many animal groups (e.g. Alroy, 1998; Benson *et al.*, 2014; Hone and Benton, 2007; Hunt and Roy, 2006; Raia *et al.*, 2012). This means that although larger testes may be desirable in order to increase reproductive output (Harcourt *et al.*, 1981), evolutionary increases in testes size must also trade-off with any overall advantage conveyed by increasing overall body mass.

This is related to the *expensive-tissue hypothesis* which proposes a trade-off between brain size and that of other metabolically expensive organs: one cannot increase in size without a decrease in another or an increase in metabolic rate (Aiello and Wheeler, 1995; Isler and van Schaik, 2006). Support for the expensive-tissue hypothesis in the form of a negative correlation between brain size and the trait of interest varies depending on the group and the tissue being studied (e.g. Aiello and Wheeler, 1995; Bordes *et al.*, 2011; Jin *et al.*, 2015; Lemaître *et al.*, 2009; Liao *et al.*, 2016; Pitnick *et al.*, 2006; Warren and Iglesias, 2012). Testes sizes should conform to the predictions of the expensive-tissue hypothesis (Lemaître *et al.*, 2009; Pitnick *et al.*, 2006) though support for this is currently lacking. Although Pitnick *et al.* (2006) argue that testes traded off with brains during the evolutionary history of bats, other authors disagree with this (Lemaître *et al.*, 2009) and there is limited evidence in other mammalian clades (Bordes *et al.*, 2011; Kelley *et al.*, 2014; Schillaci, 2006). Evidence within individual species of fish and frogs overwhelmingly demonstrate little evidence for trade-offs between testes size and brain size (Jin *et al.*, 2015; Liu *et al.*, 2014; Warren *et al.*, 2011; Zeng *et al.*, 2016; Zhao *et al.*, 2016).

Besides the brain, there are many other ecological and biological factors which are likely to trade off with testes sizes. In humans at least, testes size has found to compromise with parenting effort – fathers with smaller testes are more likely to be more nurturing (Mascaro

et al., 2013). Generally, when males invest in secondary sexual traits, it tends to come at the expense of reproductive investment (Buzatto *et al.*, 2015; Dines *et al.*, 2015; Dunn *et al.*, 2015; Fitzpatrick *et al.*, 2012) (but see Ferrandiz-Rovira *et al.*, 2014; Simmons and Fitzpatrick, 2016). This suggests that – given the expense of maintaining large testes – animals have found ways of increasing fitness and reproductive success whilst minimizing investment in testes. For example, males of the chorusing frog *Crinia georgiana* compete for mates using their arms as weapons – but have been found to preferentially invest in arm girth over testes size at lower population densities where there is lower risk of sperm competition (Buzatto *et al.*, 2015). This results in a negative correlation between testes size and arm girth across populations and implies a trade-off between sexual characters and reproductive investment (Dines *et al.*, 2015; Fitzpatrick *et al.*, 2012; Lüpold *et al.*, 2014). Direct trade-offs between such sexual characteristics and testes have been found in beetles (Simmons and Emlen, 2006) – where males manipulated to be physically unable to develop horns grew significantly larger testes.

These examples highlight the variation in testes size and degree of sperm competition that can occur at small taxonomic scales (e.g. Ginsberg and Rubenstein, 1990; Harris and Moore, 2005; Heske and Ostfeld, 1990; Newlin, 1976; Ota *et al.*, 2012; Tomkins and Simmons, 2002). Even within a single species, males can take strikingly different morphological forms depending on their mating behaviour (Gross, 1996). Such heterogeneity of data within species can cause problems for comparative analyses (Garamszegi and Møller, 2010; Garamszegi and Møller, 2011) - and are reported to be particularly problematic for testes sizes owing to differences in seasonal breeding testes asymmetry, and geographical variation (Calhim and Birkhead, 2007). However, previous analyses comparing 'reliable' testes size measures (Calhim and Birkhead, 2007) to those obtained as averages from multiple individuals or from measurements of museum specimens (Dunn *et al.*, 2001) have shown that the conclusions of analyses seeking general evolutionary relationships do not significantly differ (MacLeod and MacLeod, 2009).

We demonstrate that variation in the rate of evolution across vertebrate species gives rise to a general tendency for evolution to favour smaller testes sizes throughout their evolutionary history – regardless of their relationship with body size. A combination of novel and numerous mating strategies, investment in alternative methods of increasing reproductive success (e.g. Parker, 2016; Ramm and Schärer, 2014) and trade-offs with other metabolically expensive organs such as brains or secondary sexual traits could have driven testes size decrease in the vertebrate radiation. Our results demonstrate the exciting opportunity to reveal historical trends in traits even after considering the effect of other factors such as body size. Soft tissues of vertebrates, including testes are not often preserved well by the fossil record (Brusatte, 2012) - or at least not well enough and often enough for comparative analysis. The approach we present here therefore allows us to reveal patterns and processes of evolution that occurred deep in time that may otherwise be impossible to detect.

References

Aiello LC & Wheeler P 1995. The expensive-tissue hypothesis: the brain and the digestive system in human and primate evolution. *Current anthropology,* 36 (2): 199-221.

Alroy J 1998. Cope's rule and the dynamics of body mass evolution in North American fossil mammals. *Science,* 280 (5364): 731-734.

Baker J, Meade A, Pagel M & Venditti C 2015. Adaptive evolution toward larger size in mammals. *Proceedings of the National Academy of Sciences USA,* 112 (16): 5093-5098.

Baker J, Meade A, Pagel M & Venditti C 2016. Positive phenotypic selection inferred from phylogenies. *Biological Journal of the Linnean Society,* 118: 95-115.

Benson RBJ, Frigot RA, Goswami A, Andres B & Butler RJ 2014. Competition and constraint drove Cope's rule in the evolution of giant flying reptiles. *Nature Communications,* 5 (3567).

Bidau CJ & Medina AI 2013. Sexual size dimorphism and testis size allometry in tuco-tucos (Rodentia: Ctenomyidae). *Mammalia,* 77 (1): 81-93.

Birkhead TR & Møller AP 1998. *Sperm competition and sexual selection,* San Diego, CA, Academic Press.

Blomberg SP & Garland Jr T 2002. Tempo and mode in evolution: phylogenetic inertia, adaptation and comparative methods. *Journal of Evolutionary Biology,* 15: 899-910.

Bordes F, Morand S & Krasnov BR 2011. Does investment into "expensive" tissue compromise anti-parasitic defence? Testes size, brain size and parasite diversity in rodent hosts. *Oecologia,* 165 (1): 7-16.

Breed W 1981. Unusual anatomy of the male reproductive tract in *Notomys alexis* (Muridae). *Journal of Mammalogy,* 62 (2): 373-375.

Breed W 1997. Interspecific variation of testis size and epididymal sperm numbers in Australasian rodents with special reference to the genus *Notomys*. *Australian Journal of Zoology,* 45 (6): 651-669.

Breed W & Washington J 1991. Mating behaviour and insemination in the hopping mouse (*Notomys alexis*). *Journal of reproduction and fertility,* 93 (1): 187-194.

Breed WG & Jason T 2000. Body mass, testes mass, and sperm size in murine rodents. *Journal of Mammalogy,* 81 (3): 758-768.

Brown JH & Sibly RM 2006. Life-history evolution under a production constraint. *Proceedings of the National Academy of Sciences USA,* 103 (47): 17595-17599.

Brownell R & Ralls K 1986. Potential for sperm competition in baleen whales. *report of the International Whaling Commission,* 8: 110-141.

Brusatte SL 2012. *Dinosaur paleobiology,* UK, John Wiley & Sons, Ltd.

Bull CM 2000. Monogamy in lizards. *Behavioural Processes,* 51 (1–3): 7-20.

Buzatto BA, Roberts JD & Simmons LW 2015. Sperm competition and the evolution of precopulatory weapons: Increasing male density promotes sperm competition and reduces selection on arm strength in a chorusing frog. *Evolution,* 69 (10): 2613-2624.

Byrne PG, Roberts JD & Simmons LW 2002. Sperm competition selects for increased testes mass in Australian frogs. *Journal of Evolutionary Biology,* 15 (3): 347-355.

Calhim S & Birkhead TR 2007. Testes size in birds: quality versus quantity - assumptions, errors, and estimates. *Behavioral Ecology,* 18 (1): 271-275.

Cockburn A 2006. Prevalence of different modes of parental care in birds. *Proceedings of the Royal Society B: Biological Sciences,* 273 (1592): 1375-1383.

Dines JP, Mesnick SL, Ralls K, May-Collado L, Agnarsson I & Dean MD 2015. A trade-off between precopulatory and postcopulatory trait investment in male cetaceans. *Evolution,* 69 (6): 1560-1572.

Dines JP, Otárola-Castillo E, Ralph P, Alas J, Daley T, Smith AD & Dean MD 2014. Sexual selection targets cetacean pelvic bones. *Evolution,* 68 (11): 3296-3306.

Dunn JC, Halenar LB, Davies TG, Cristobal-Azkarate J, Reby D, Sykes D, Dengg S, Fitch WT & Knapp LA 2015. Evolutionary trade-off between vocal tract and testes dimensions in howler monkeys. *Current Biology,* 25 (21): 2839-2844.

Dunn PO, Whittingham LA & Pitcher TE 2001. Mating systems, sperm competition, and the evolution of sexual dimorphism in birds. *Evolution,* 55 (1): 161-175.

Eastman JM, Alfaro ME, Joyce P, Hipp AL & Harmon LJ 2011. A novel comparative method for identifying shifts in the rate of character evolution on trees. *Evolution,* 65 (12): 3578-89.

Felsenstein J 1985. Phylogenies and the comparative method. *The American Naturalist,* 125 (1): 1-15.

Ferrandiz-Rovira M, Lemaître J-F, Lardy S, López BC & Cohas A 2014. Do pre-and post-copulatory sexually selected traits covary in large herbivores? *BMC evolutionary biology,* 14 (1): 1.

Fitzpatrick JL, Almbro M, Gonzalez-Voyer A, Kolm N & Simmons LW 2012. Male contest competition and the coevolution of weaponry and testes in pinnipeds. *Evolution,* 66 (11): 3595-3604.

Freckleton RP 2002. On the misuse of residuals in ecology: regression of residuals vs. multiple regression. *Journal of Animal Ecology,* 71 (3): 542-545.

Freckleton RP & Watkinson AR 2001. Nonmanipulative determination of plant community dynamics. *Trends in Ecology & Evolution,* 16 (6): 301-307.

Gage MJ 1994. Associations between body size, mating pattern, testis size and sperm lengths across butterflies. *Proceedings of the Royal Society of London B: Biological Sciences,* 258 (1353): 247-254.

Garamszegi LZ, Eens M, Hurtrez-Boussès S & Møller AP 2005. Testosterone, testes size, and mating success in birds: a comparative study. *Hormones and Behavior,* 47 (4): 389-409.

Garamszegi LZ & Møller AP 2010. Effects of sample size and intraspecific variation in phylogenetic comparative studies: a meta-analytic review. *Biological Reviews,* 85 (4): 797-805.

Garamszegi LZ & Møller AP 2011. Nonrandom variation in within-species sample size and missing data in phylogenetic comparative studies. *Systematic biology,* 60 (6): 876-880.

Ginsberg JR & Rubenstein DI 1990. Sperm competition and variation in zebra mating behavior. *Behavioral Ecology and Sociobiology,* 26 (6): 427-434.

Gross MR 1996. Alternative reproductive strategies and tactics: diversity within sexes. *Trends in Ecology & Evolution,* 11 (2): 92-98.

Happold M 1976. Social behavior of the conilurine rodents (Muridae) of Australia. *Zeitschrift für Tierpsychologie,* 40 (2): 113-182.

Harcourt AH, Harvey PH, Larson SG & Short R 1981. Testis weight, body weight and breeding system in primates. *Nature,* 293 (5827): 55-57.

Harcourt AH, Purvis A & Liles L 1995. Sperm Competition: Mating system, not breeding season, affects testes size of primates. *Functional Ecology,* 9 (3): 468-476.

Harris W & Moore P 2005. Sperm competition and male ejaculate investment in *Nauphoeta cinerea*: effects of social environment during development. *Journal of evolutionary biology,* 18 (2): 474-480.

Harvey PH & Pagel M 1991. *The comparative method in evolutionary biology,* Oxford, Oxford University Press.

Hayward A & Gillooly JF 2011. The cost of sex: Quantifying energetic investment in gamete production by males and females. *PLoS ONE,* 6 (1): e16557.

Hedges SB, Marin J, Suleski M, Paymer M & Kumar S 2015. Tree of life reveals clock-like speciation and diversification. *Molecular biology and evolution,* 32 (4): 835-845.

Heske EJ & Ostfeld RS 1990. Sexual dimorphism in size, relative size of testes, and mating systems in North American voles. *Journal of Mammalogy,* 71 (4): 510-519.

Hone DW & Benton MJ 2005. The evolution of large size: How does Cope's rule work? *Trends in Ecology & Evolution,* 20 (1): 4-6.

Hone DW & Benton MJ 2007. Cope's rule in the Pterosauria, and differing perceptions of Cope's rule at different taxonomic levels. *Journal of Evolutionary Biology,* 20 (3): 1164-70.

Hosken DJ 1997. Sperm competition in bats. *Proceedings of the Royal Society B: Biological Sciences,* 264 (1380): 385-392.

Hosken DJ 1998. Testes mass in megachiropteran bats varies in accordance with sperm competition theory. *Behavioral Ecology and Sociobiology,* 44 (3): 169-177.

Hunt G & Roy K 2006. Climate change, body size evolution, and Cope's rule in deep-sea ostracodes. *Proceedings of the National Academy of Sciences USA,* 103 (5): 1347-52.

Iossa G, Soulsbury CD, Baker PJ & Harris S 2008. Sperm competition and the evolution of testes size in terrestrial mammalian carnivores. *Functional Ecology,* 22 (4): 655-662.

Isler K & van Schaik C 2006. Costs of encephalization: the energy trade-off hypothesis tested on birds. *Journal of Human Evolution,* 51 (3): 228-243.

Jetz W & Freckleton RP 2015. Towards a general framework for predicting threat status of data-deficient species from phylogenetic, spatial and environmental information. *Philosophical Transactions of the Royal Society B: Biological Sciences,* 370 (1662).

Jin L, Zhao L, Liu WC, Zeng Y & Liao WB 2015. Evidence for the expensive-tissue hypothesis in the Omei Wood Frog (*Rana omeimontis*). *The Herpetological Journal,* 25 (2): 127-130.

Kahrl AF, Cox CL & Cox RM 2016. Correlated evolution between targets of pre- and postcopulatory sexual selection across squamate reptiles. *Ecology and Evolution,* 6 (18): 6452-6459.

Kappeler PM 1997. Intrasexual selection and testis size in strepsirhine primates. *Behavioral Ecology,* 8 (1): 10-19.

Kelley TC, Higdon JW & Ferguson SH 2014. Large testes and brain sizes in odontocetes (order Cetacea, suborder Odontoceti): the influence of mating system on encephalization. *Canadian Journal of Zoology,* 92 (8): 721-726.

Kenagy G & Trombulak SC 1986. Size and function of mammalian testes in relation to body size. *Journal of Mammalogy,* 67 (1): 1-22.

Kingsolver JG & Pfennig DW 2004. Individual-level selection as a cause of Cope's rule of phyletic size increase. *Evolution,* 58 (7): 1608-1612.

Kingsolver JG & Pfennig DW 2007. Patterns and power of phenotypic selection in nature. *Bioscience,* 57 (7): 561-572.

Kratsch C & McHardy AC 2014. RidgeRace: ridge regression for continuous ancestral character estimation on phylogenetic trees. *Bioinformatics,* 30 (17): i527-i533.

Kutsukake N & Innan H 2013. Simulation-based likelihood approach for evolutionary models of phenotypic traits on phylogeny. *Evolution,* 67 (2): 355-367.

Kutsukake N & Innan H 2014. Detecting phenotypic selection by Approximate Bayesian Computation in phylogenetic comparative methods. *In:* Garamszegi LZ (ed.) *Modern Phylogenetic Comparative Methods and Their Application in Evolutionary Biology.* Berlin: Springer-Verlag.

Lack DL 1968. *Ecological adaptations for breeding in birds,* London, Methuen.

Lemaître JF, Ramm SA, Barton RA & Stockley P 2009. Sperm competition and brain size evolution in mammals. *Journal of Evolutionary Biology,* 22 (11): 2215-2221.

Liao WB, Lou SL, Zeng Y, Kotrschal A, Leips J & Bronstein JL 2016. Large brains, small guts: The expensive tissue hypothesis supported within anurans. *The American Naturalist,* 188 (6).

Liu J, Zhou CQ & Liao WB 2014. Evidence for neither the compensation hypothesis nor the expensive-tissue hypothesis in *Carassius auratus*. *Animal Biology,* 64 (2): 177-187.

Lukas D & Clutton-Brock TH 2013. The evolution of social monogamy in mammals. *Science,* 341 (6145): 526-530.

Lüpold S, Tomkins JL, Simmons LW & Fitzpatrick JL 2014. Female monopolization mediates the relationship between pre-and postcopulatory sexual traits. *Nature communications,* 5.

MacLeod CD 2010. The relationship between body mass and relative investment in testes mass in cetaceans: Implications for inferring interspecific variations in the extent of sperm competition. *Marine Mammal Science,* 26 (2): 370-380.

MacLeod CD 2014. Exploring and explaining complex allometric relationships: A case study on amniote testes mass allometry. *Systems,* 2 (3): 379-392.

MacLeod CD & MacLeod R 2009. The relationship between body mass and relative investment in testes mass in amniotes and other vertebrates. *Oikos,* 118 (6): 903-916.

Mann J, Connor RC, Tyack PL & Whitehead H 2000. *Cetacean societies: field studies of dolphins and whales,* Chicago and London, University of Chicago Press.

Mascaro JS, Hackett PD & Rilling JK 2013. Testicular volume is inversely correlated with nurturing-related brain activity in human fathers. *Proceedings of the National Academy of Sciences,* 110 (39): 15746-15751.

Mock DW & Fujioka M 1990. Monogamy and long-term pair bonding in vertebrates. *Trends in Ecology & Evolution,* 5 (2): 39-43.

Møller AP 1988a. Ejaculate quality, testes size and sperm competition in primates. *Journal of Human Evolution,* 17 (5): 479-488.

Møller AP 1988b. Testes size, ejaculate quality and sperm competition in birds. *Biological Journal of the Linnean Society,* 33 (3): 273-283.

Møller AP 1989. Ejaculate quality, testes size and sperm production in mammals. *Functional Ecology,* 3 (1): 91-96.

Møller AP 1994. Directional selection on directional asymmetry: testes size and secondary sexual characters in birds. *Proceedings of the Royal Society of London B: Biological Sciences,* 258 (1352): 147-151.

Møller AP & Briskie JV 1995. Extra-pair paternity, sperm competition and the evolution of testis size in birds. *Behavioral Ecology and Sociobiology,* 36 (5): 357-365.

Newlin ME 1976. Reproduction in the Bunch Grass Lizard, *Sceloporus scalaris*. *Herpetologica,* 32 (2): 171-184.

Opie C, Atkinson QD, Dunbar RIM & Shultz S 2013. Male infanticide leads to social monogamy in primates. *Proceedings of the National Academy of Sciences,* 110 (33): 13328-13332.

Ota K, Hori M & Kohda M 2012. Testes investment along a vertical depth gradient in an herbivorous fish. *Ethology,* 118 (7): 683-693.

Pagel M 1994. The adaptationist wager. *In:* Eggleton P & Vane-Wright RI (eds.), *Phylogenetics and ecology.* London: Academic Press.

Pagel M 1999. Inferring the historical patterns of biological evolution. *Nature,* 401: 877-884.

Pagel M, Meade A & Barker D 2004. Bayesian estimation of ancestral character states on phylogenies. *Systematic Biology,* 53 (5): 673-84.

Parker G, Ball M, Stockley P & Gage M 1997. Sperm competition games: a prospective analysis of risk assessment. *Proceedings of the Royal Society of London B: Biological Sciences,* 264 (1389): 1793-1802.

Parker GA 1970. Sperm competition and its evolutionary consequences in the insects. *Biological Reviews,* 45 (4): 525-567.

Parker GA 2016. The evolution of expenditure on testes. *Journal of Zoology,* 298 (1): 3-19.

Peirce E & Breed W 2001. A comparative study of sperm production in two species of Australian arid zone rodents (*Pseudomys australis*, *Notomys alexis*) with marked differences in testis size. *Reproduction,* 121 (2): 239-247.

Peters RH 1986. *The ecological implications of body size,* UK, Cambridge University Press.

Pitcher T, Dunn P & Whittingham L 2005. Sperm competition and the evolution of testes size in birds. *Journal of evolutionary biology,* 18 (3): 557-567.

Pitnick S, Jones KE & Wilkinson GS 2006. Mating system and brain size in bats. *Proceedings of the Royal Society of London B: Biological Sciences,* 273 (1587): 719-724.

Pyron M 2000. Testes mass and reproductive mode of minnows. *Behavioral Ecology and Sociobiology,* 48 (2): 132-136.

Rabosky DL 2014. Automatic detection of key innovations, rate shifts, and diversity-dependence on phylogenetic trees. *PLoS One,* 9 (2): e89543.

Raia P, Carotenuto F, Passaro F, Fulgione D & Fortelius M 2012. Ecological specialization in fossil mammals explains Cope's rule. *The American Naturalist,* 179 (3): 328-337.

Ramm SA & Schärer L 2014. The evolutionary ecology of testicular function: size isn't everything. *Biological Reviews,* 89 (4): 874-888.

Revell LJ 2010. Phylogenetic signal and linear regression on species data. *Methods in Ecology and Evolution,* 1 (4): 319-329.

Revell LJ, Mahler DL, Peres-Neto PR & Redelings BD 2012. A new phylogenetic method for identifying exceptional phenotypic diversification. *Evolution,* 66 (1): 135-146.

Schillaci MA 2006. Sexual selection and the evolution of brain size in primates. *PLoS ONE,* 1 (1): e62.

Schulte-Hostedde AI, Millar JS & Hickling GJ 2005. Condition dependence of testis size in small mammals. *Evolutionary ecology research,* 7 (1): 143-149.

Simmons LW & Emlen DJ 2006. Evolutionary trade-off between weapons and testes. *Proceedings of the National Academy of Sciences,* 103 (44): 16346-16351.

Simmons LW & Fitzpatrick JL 2012. Sperm wars and the evolution of male fertility. *Reproduction,* 144 (5): 519-534.

Simmons LW & Fitzpatrick JL 2016. Sperm competition and the coevolution of pre- and postcopulatory traits: Weapons evolve faster than testes among onthophagine dung beetles. *Evolution,* 70 (5): 998-1008.

Simpson GG 1953. *The major features of evolution,* London, GB, Columbia University Press.

Soulsbury CD 2010. Genetic patterns of paternity and testes size in mammals. *PLoS ONE,* 5 (3): e9581.

Stockley P, Gage M, Parker G & Møller A 1997. Sperm competition in fishes: the evolution of testis size and ejaculate characteristics. *The American Naturalist,* 149 (5): 933-954.

Takegaki T 2000. Monogamous mating system and spawning cycle in the gobiid fish, *Amblygobius phalaena* (Gobiidae). *Environmental Biology of Fishes,* 59 (1): 61-67.

Thomas GH & Freckleton RP 2012. MOTMOT: Models of trait macroevolution on trees. *Methods in Ecology and Evolution,* 3 (1): 145-151.

Tomkins JL & Simmons LW 2002. Measuring relative investment: a case study of testes investment in species with alternative male reproductive tactics. *Animal Behaviour,* 63 (5): 1009-1016.

Venditti C, Meade A & Pagel M 2011. Multiple routes to mammalian diversity. *Nature,* 479 (7373): 393-396.

Warren DL & Iglesias TL 2012. No evidence for the 'expensive-tissue hypothesis' from an intraspecific study in a highly variable species. *Journal of Evolutionary Biology,* 25 (6): 1226-1231.

Warren J, Topping CJ & James P 2011. An evolutionary modelling approach to understanding the factors behind plant invasiveness and community susceptibility to invasion. *Journal of Evolutionary Biology,* 24 (10): 2099-109.

Whiteman EA & Côté IM 2003. Social monogamy in the cleaning goby *Elacatinus evelynae*: ecological constraints or net benefit? *Animal Behaviour,* 66 (2): 281-291.

Whiteman EA & Côté IM 2004. Monogamy in marine fishes. *Biological Reviews,* 79 (2): 351-375.

Xie W, Lewis PO, Fan Y, Kuo L & Chen M-H 2010. Improving marginal likelihood estimation for Bayesian phylogenetic model selection. *Systematic Biology,* 60 (2): 150-160.

Zeng Y, Lou SL, Liao WB, Jehle R & Kotrschal A 2016. Sexual selection impacts brain anatomy in frogs and toads. *Ecology and Evolution,* 6 (19): 7070-7079.

Zhao L, Mao M & Liao WB 2016. No evidence for the 'expensive-tissue hypothesis' in the dark-spotted frog, *Pelophylax nigromaculatus*. *Acta Herpetologica,* 11 (1): 69-73.

Appendix 1

Testes mass and body mass data

Testes size and body size for vertebrate species were collected from the literature. We prioritized collation of data from sources which contained data for multiple species, and supplemented these with data for individual species.

To avoid issues with differences among datasets, we enforced a standardized protocol to ensure a single measurement for each species: Firstly, we preferred as a priority any sources where testes mass and male body mass were explicitly measured from individual specimens (e.g. Jennions and Passmore, 1993). Where a source was itself a compilation from the literature, we preferred those that at least attempted to obtain body sizes measured in the same individuals as testes size – or failing that, in individuals from the same geographic regions (see Kenagy and Trombulak, 1986). Where single values (for either testes mass or body mass) are provided to represent multiple individuals or populations, we preferred mean values over maximums. Data extracted directly were preferred over values extracted using data extraction software from images (e.g. Fitzpatrick *et al.* 2012). Finally, where we have multiple sources that fit the above criteria, we preferred the most recently published values. Ambiguous data sources were automatically placed at the end of the priority list. Where datasets have the same priority and the same date, we arbitrarily prioritized the dataset with the largest sample size.

Any sources that only contained testes sizes were added to the dataset at the end, and supplemented with body size data from additional sources. We manually modified the values for the following species: *Peromyscus californicus* is incorrectly reported in Soulsbury (2010) and so we corrected it to reflect the value in the original cited source (Nelson *et al.,* 1995); the body masses of two bird species were identified by Peter Dunn (personal communication) to be incorrect in the data set of Calhim and Birkhead (2006)

and so these were adjusted according to his suggestions (Dunn *et al.*, 2001; Pitcher *et al.* 2005).

Matching species names

We matched species to the time tree of life (Hedges *et al.*, 2015). After adjusting for errors and alternative spellings, we checked for species synonyms using several major taxonomic resources: (AmphibiaWeb, 2005; BirdLife Taxonomic Working Group, 2015; Froese and Pauly, 2012; Lepage, 2009; Roskov *et al.*, 2016; Species Survival Commision, 2001).

To maximize the amount of available data, we incorporated species that did not match after checking spellings and synonymy by checking if other members of the genus were included in the phylogeny. Where a species is the sole member of its genus found in the data (but not in the tree), we substituted the data for that species with another member of the same genus that is found in the phylogeny. We only did this where the genus comprised a monophyletic clade, and where inclusion of an additional member did not alter relationships among species already found in the data and phylogeny (Table A1.1).

Table A1.1: Species incorporated using genus matching.

Group	Species (Hedges *et al.* 2015)	Species from Data
Actinopterygii	*Doryrhamphus excisus*	*D. negrosensis*
	Gila conspersa	*G. atraria*
	Hippichthys penicillus	*H. heptagonus*
	Hippocampus kelloggi	*H. angustus*
	Lythrurus roseipinnis	*L. bellus*
	Melanotaenia sp. AC-2011	*M. eachamensis*
	Solegnathus hardwickii	*S. spinosissimus*
Anura	*Physalaemus cuvieri*	*P. crombiei*
	Kaloula conjuncta	*K. verrucosa*
	Brachytarsophrys feae	*B. chuannanensis*
Aves	*Ailuroedus melanotis*	*A. buccoides*
	Anhinga anhinga	*A. melanogaster*
	Calandrella acutirostris	*C. cinerea*
	Cinclus cinclus	*C. leucocephalus*
	Cotinga cayana	*C. amabilis*
	Cyphorhinus arada	*C. phaeocephalus*
	Ducula aenea	*D. spilorrhoa*
	Grallaria andicolus	*G. quitensis*
	Hylopezus berlepschi	*H. perspicillatus*
	Laniocera hypopyrra	*L. rufescens*
	Oncostoma cinereigulare	*O. olivaceum*
	Sphecotheres vieilloti	*S. viridis*
Mammalia	*Acerodon jubatus*	*A. mackloti*
	Axis axis	*A. calamianensis*
	Bunomys andrewsi	*B. fratrorum*
	Epomophorus crypturus	*E. labiatus*
	Epomops franqueti	*E. buettikoferi*
	Saccolaimus flaviventris	*S. peli*

References for testes and body size data

Anderson MJ, Nyholt J & Dixson AF 2004. Sperm competition affects the structure of the mammalian *vas deferens*. *Journal of Zoology,* 264 (1): 97-103.

Birkhead TR & Møller AP 1998. *Sperm competition and sexual selection,* San Diego, CA, Academic Press.

Breed WG & Jason T 2000. Body mass, testes mass, and sperm size in murine rodents. *Journal of Mammalogy,* 81 (3): 758-768.

Byrne PG, Roberts JD & Simmons LW 2002. Sperm competition selects for increased testes mass in Australian frogs. *Journal of Evolutionary Biology,* 15 (3): 347-355.

Calhim S & Birkhead TR 2007. Testes size in birds: quality versus quantity - assumptions, errors, and estimates. *Behavioral Ecology,* 18 (1): 271-275.

Dines JP, Mesnick SL, Ralls K, May-Collado L, Agnarsson I & Dean MD 2015. A trade-off between precopulatory and postcopulatory trait investment in male cetaceans. *Evolution,* 69 (6): 1560-1572.

Eyo VO, Ekanem AP & Jimmy U-IU 2014. A comparative study of the gonado-somatic index (GSI) and gonad gross morphology of African catfish (*Clarias gariepinus*) fed unical aqua feed and coppens commercial feed. *Croatian Journal of Fisheries,* 72 (2): 63-69.

Fitzpatrick JL, Almbro M, Gonzalez-Voyer A, Kolm N & Simmons LW 2012. Male contest competition and the coevolution of weaponry and testes in pinnipeds. *Evolution,* 66 (11): 3595-3604.

Ginsberg JR & Rubenstein DI 1990. Sperm competition and variation in zebra mating behavior. *Behavioral Ecology and Sociobiology,* 26 (6): 427-434.

Harcourt AH, Purvis A & Liles L 1995. Sperm Competition: Mating system, not breeding season, affects testes size of primates. *Functional Ecology,* 9 (3): 468-476.

Hayward A & Gillooly JF 2011. The cost of sex: Quantifying energetic investment in gamete production by males and females. *PLoS ONE,* 6 (1): e16557.

Hosken D & Ward P 2001. Experimental evidence for testis size evolution via sperm competition. *Ecology Letters,* 4 (1): 10-13.

Iossa G, Soulsbury CD, Baker PJ & Harris S 2008. Sperm competition and the evolution of testes size in terrestrial mammalian carnivores. *Functional Ecology,* 22 (4): 655-662.

Jennions MD & Passmore NI 1993. Sperm competition in frogs: testis size and a 'sterile male'experiment on *Chiromantis xerampelina* (Rhacophoridae). *Biological Journal of the Linnean Society,* 50 (3): 211-220.

Kusano T, Mitsuhiko T & Fukuyama K 1991. Testes size and breeding systems in Japanese annrans with special reference to large testes in the treefrog, *Rhacophorus arboreus* (Amphibia: Rhacophoridae). *Behavioral Ecology and Sociobiology,* 29 (1): 27-31.

Kvarnemo C & Simmons LW 2004. Testes investment and spawning mode in pipefishes and seahorses (Syngnathidae). *Biological Journal of the Linnean Society,* 83 (3): 369-376.

Liao WB, Mi ZP, Zhou CQ, Jin L, Lou SL, Han X & Ma J 2011. Relative testis size and mating systems in anurans: large testis in multiple-male mating in foam-nesting frogs. *Animal Biology,* 61 (2): 225-238.

Macleod CD 2010. The relationship between body mass and relative investment in testes mass in cetaceans: Implications for inferring interspecific variations in the extent of sperm competition. *Marine Mammal Science,* 26 (2): 370-380.

Macleod CD & Macleod R 2009. The relationship between body mass and relative investment in testes mass in amniotes and other vertebrates. *Oikos,* 118 (6): 903-916.

Møller AP 1988. Ejaculate quality, testes size and sperm competition in primates. *Journal of Human Evolution,* 17 (5): 479-488.

Nelson RJ, Gubernick DJ & Blom JM 1995. Influence of photoperiod, green food, and water availability on reproduction in male California mice (*Peromyscus californicus*). *Physiology & behavior,* 57 (6): 1175-1180.

Olsson M, Madsen T & Møller A 1998. Sexual selection and sperm competition in reptiles. *In:* Birkhead TR & Møller A (eds.), *Sperm Competition and Sexual Selection.* San Diego: Academic Press.

Pitcher T, Dunn P & Whittingham L 2005. Sperm competition and the evolution of testes size in birds. *Journal of evolutionary biology,* 18 (3): 557-567.

Pitnick S, Jones KE & Wilkinson GS 2006. Mating system and brain size in bats. *Proceedings of the Royal Society of London B: Biological Sciences,* 273 (1587): 719-724.

Prado CPA & Haddad CF 2003. Testes size in leptodactylid frogs and occurrence of multimale spawning in the genus *Leptodactylus* in Brazil. *Journal of Herpetology,* 37 (2): 354-362.

Pyron M, Pitcher T & Jacquemin S 2013. Evolution of mating systems and sexual size dimorphism in North American cyprinids. *Behavioral Ecology and Sociobiology,* 67 (5): 747-756.

Ramm SA 2007. Sexual selection and genital evolution in mammals: a phylogenetic analysis of baculum length. *The American Naturalist,* 169 (3): 360-369.

Ramm SA, Parker GA & Stockley P 2005. Sperm competition and the evolution of male reproductive anatomy in rodents. *Proceedings of the Royal Society B: Biological Sciences,* 272 (1566): 949-955.

Reynolds J, Rommel SA & Pitchford ME 2004. The likelihood of sperm competition in manatees - explaining an apparent paradox. *Marine Mammal Science,* 20 (3): 464-476.

Rose RW, Nevison CM & Dixson AF 1997. Testes weight, body weight and mating systems in marsupials and monotremes. *Journal of Zoology,* 243 (3): 523-531.

Soulsbury CD 2010. Genetic patterns of paternity and testes size in mammals. *PLoS ONE,* 5 (3): e9581.

Todd AC 2008. Using testis size to predict the mating systems of New Zealand geckos. *New Zealand Journal of Zoology,* 35 (2): 103-114.

White CR, Phillips NF & Seymour RS 2006. The scaling and temperature dependence of vertebrate metabolism. *Biology Letters,* 2 (1): 125-127.

Zeng Y, Lou SL, Liao WB & Jehle R 2014. Evolution of sperm morphology in anurans: insights into the roles of mating system and spawning location. *BMC evolutionary biology,* 14 (1): 1.

Taxonomic resources

Amphibiaweb. 2005. *AmphibiaWeb: Information on amphibian biology and conservation* [Online]. Available: http://amphibiaweb.org/ [Accessed March 2016].

Birdlife Taxonomic Working Group. 2015. *BirdLife Taxonomic Checklist v8.0* [Online]. Available: http://www.birdlife.org/datazone/info/taxonomy [Accessed March 2016].

Froese R & Pauly D. 2012. *FishBase* [Online]. Available: http://www.fishbase.org [Accessed March 2016].

Hedges SB, Marin J, Suleski M, Paymer M & Kumar S 2015. Tree of life reveals clock-like speciation and diversification. *Molecular biology and evolution,* 32 (4): 835-845.

Lepage D. 2009. *Avibase–the world bird database* [Online]. Available: http://avibase.bsc-eoc.org [Accessed March 2016].

Roskov Y, Abucay L, Orrell T, Nicolson D, Flann C, Bailly N, Kirk P, Bourgoin T, Dewalt R, Decock W & De Wever A 2016. Species 2000 & ITIS Catalogue of Life, 2016 Annual Checklist. Leiden, the Netherlands: Naturalis.

Species Survival Commision 2001. IUCN red list categories and criteria: version 3.1. *Prepared by the IUCN Species Survival Commission*.

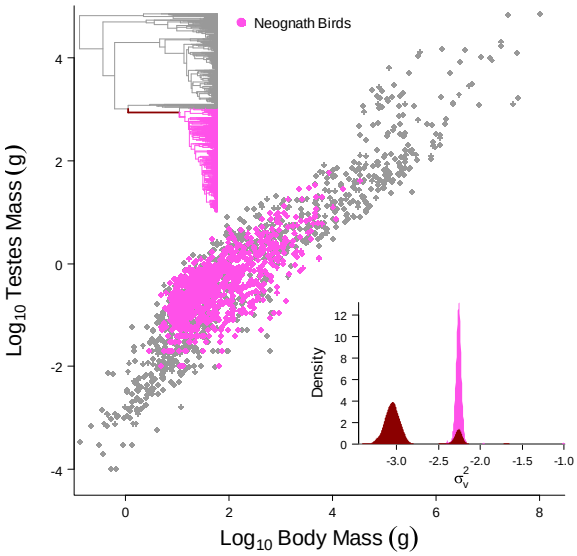
Appendix 2

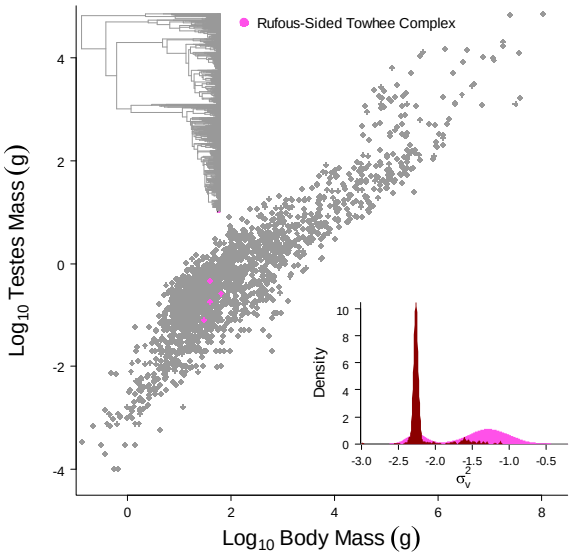
Rate heterogeneity in vertebrate testes size evolution

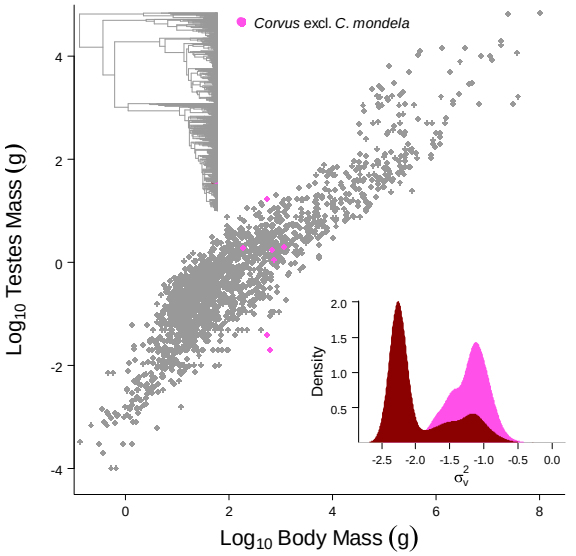
We find 21 radiations in testes sizes scattered throughout the vertebrate phylogeny. Nine of these represent primary heritable rate shifts where entire clades have experienced a rate acceleration away from the background rate of morphological evolution across all other vertebrates. We plot the posterior distribution of rates estimated along ancestor-descendant lineage pairs for each of these groups in the main text (Figure 4). The remaining eleven of these shifts are nested shifts found within birds and mammals and represent a secondary rate acceleration away from the primary heritable rate shift in either neognath birds (all birds excluding ratites and tinamous) or boreoeutherian mammals (all mammals excluding Afrotheria and Xenarthra).

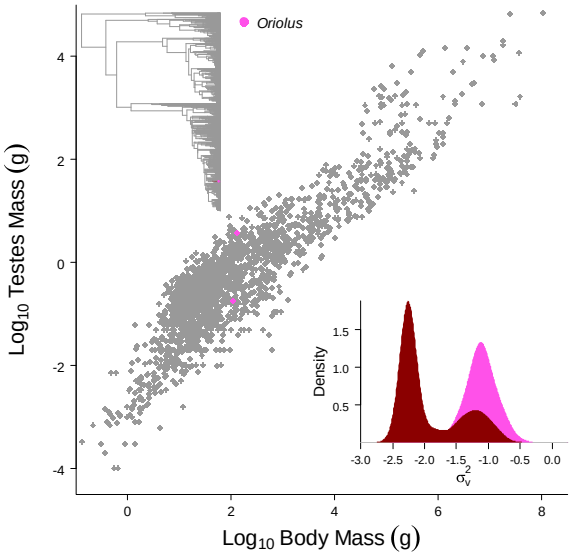
4 of the heritable rate shifts we identify occurred within large groups of species ($n \geq 20$). We repeated the variable rates analysis as described in the main text but allowed a separate slope and intercept for each of these groups: Boreoeutherian mammals, cetaceans, neognath birds, and the anuran superfamily Bufonoidea. None of these groups have a significant difference in intercept or slope compared to the group within which they are nested ($P_x > 0.05$ in all cases).

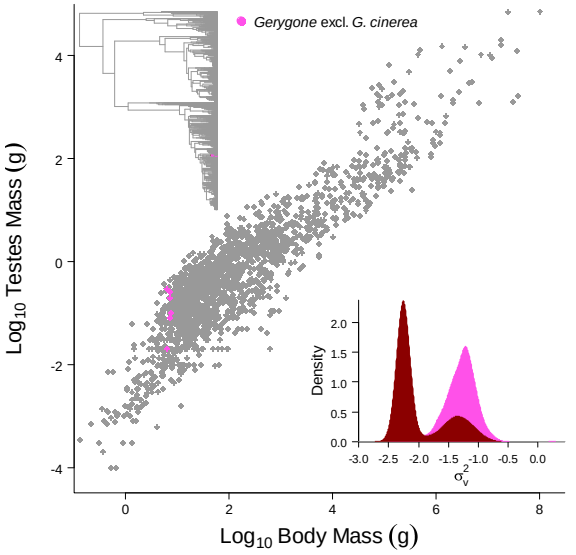
Here, we show the phylogenetic position and ancestor-descendant lineage rate comparisons presented in the context of the data for all heritable rate shifts and the mean shift observed within *Notomys* species (Figure A2.1).

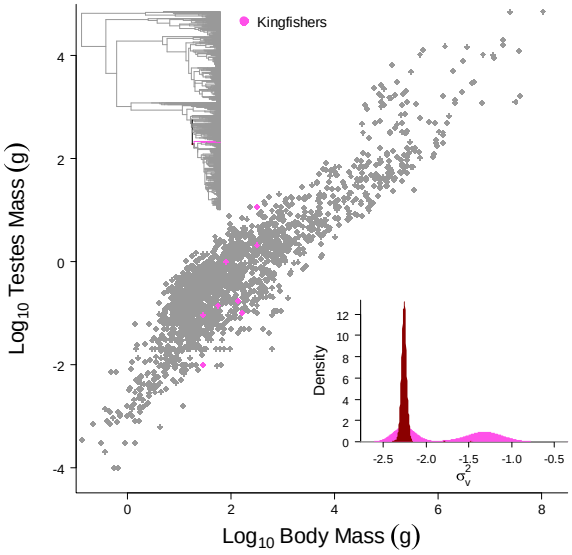

Neognath Birds
Log10 Testes Mass (g)
Log10 Body Mass (g)
Density
σ2v

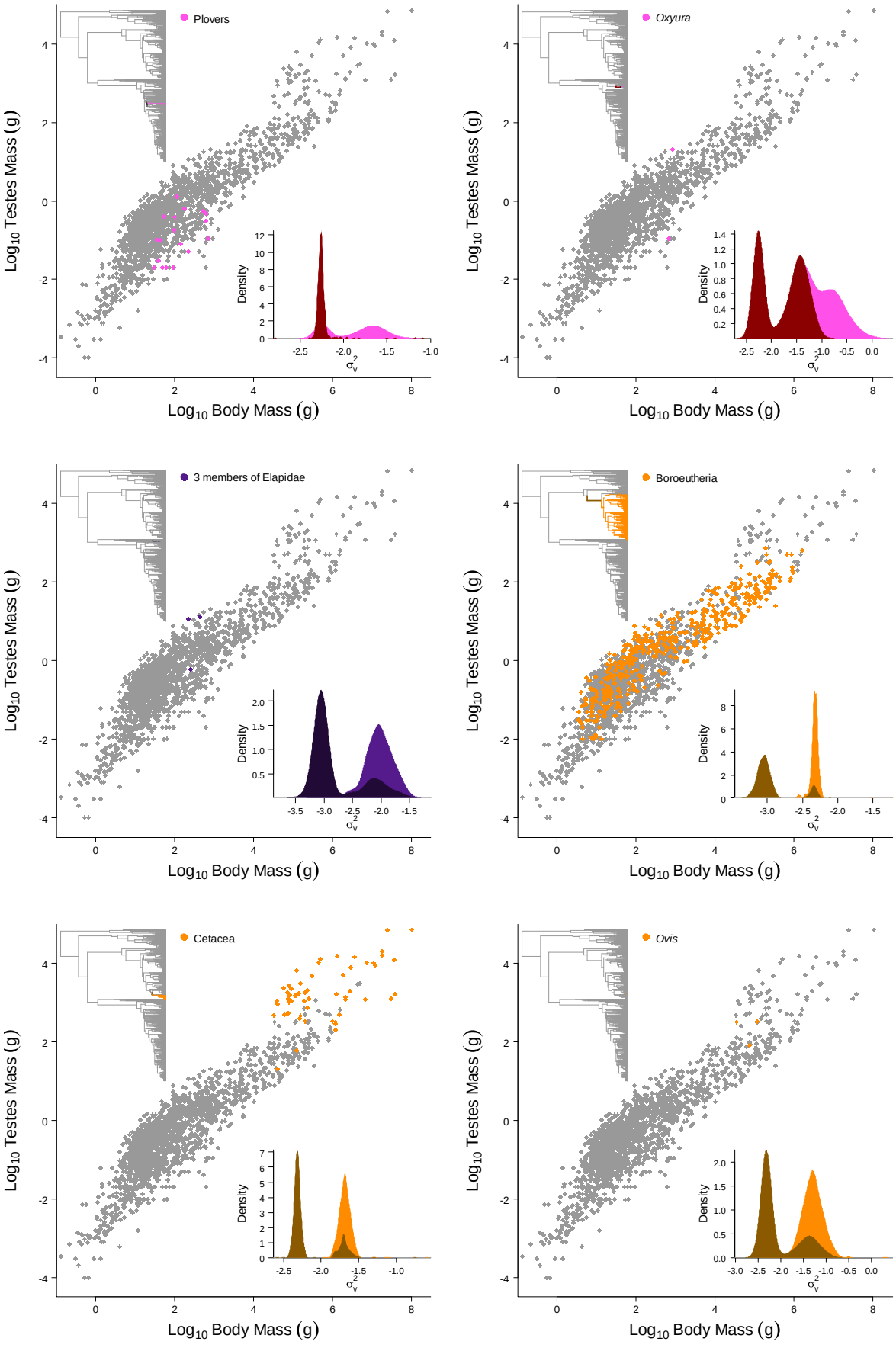

Rufous-Sided Towhee Complex
Log10 Testes Mass (g)
Log10 Body Mass (g)
Density
σ2v

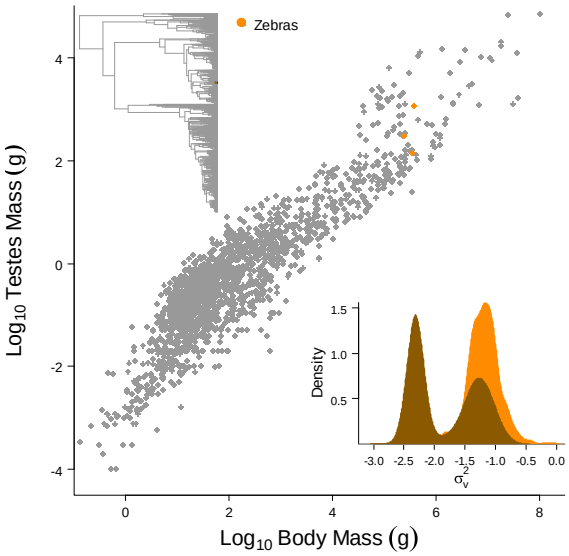

Corvus excl. C. mondela
Log10 Testes Mass (g)
Log10 Body Mass (g)
Density
σ2v

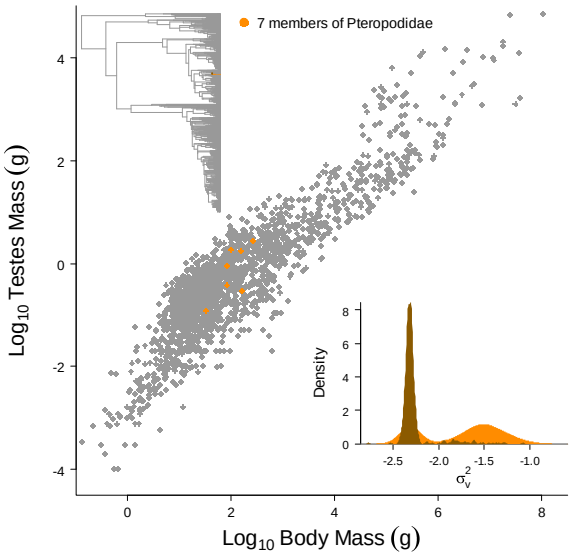

Oriolus
Log10 Testes Mass (g)
Log10 Body Mass (g)
Density
σ2v

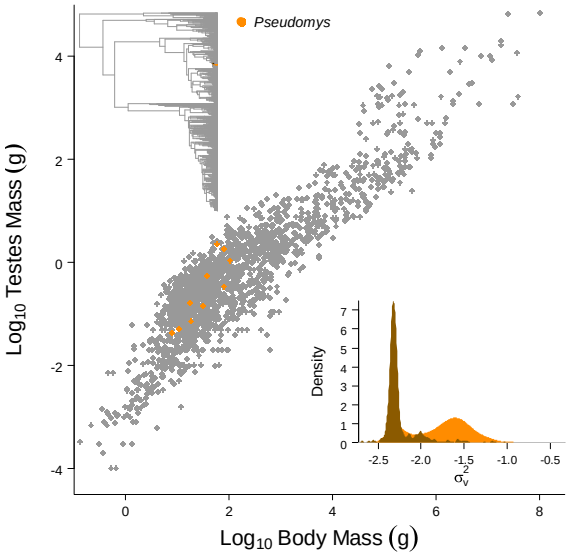

Gerygone excl. G. cinerea
Log10 Testes Mass (g)
Log10 Body Mass (g)
Density
σ2v

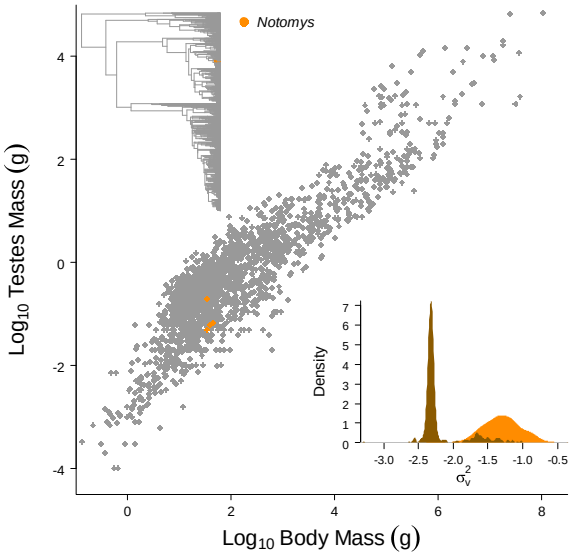

Kingfishers
Log10 Testes Mass (g)
Log10 Body Mass (g)
Density
σ2v

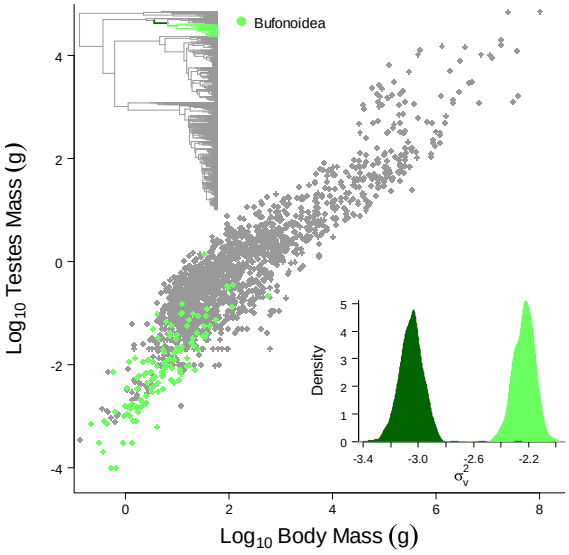

Plovers
Oxyura
3 members of Elapidae
Boroeutheria
Cetacea
Ovis
Log10 Testes Mass (g)
Log10 Body Mass (g)
Density
σ²v

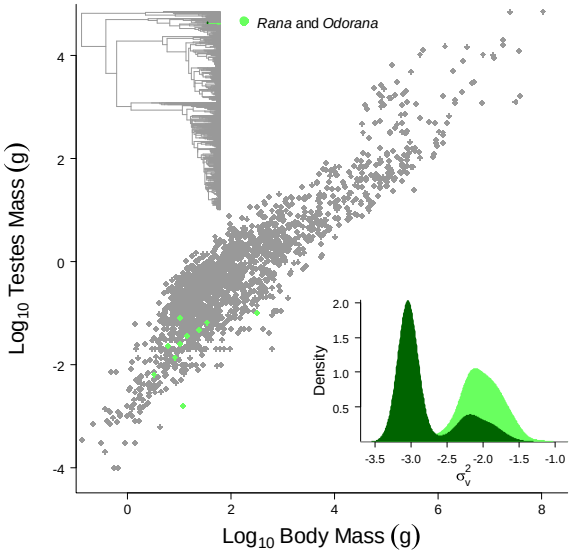

Zebras
Log10 Testes Mass (g)
Log10 Body Mass (g)
Density
σ²v


7 members of Pteropodidae
Log10 Testes Mass (g)
Log10 Body Mass (g)
Density
σ²v


Pseudomys
Log10 Testes Mass (g)
Log10 Body Mass (g)
Density
σ²v


Notomys
Log10 Testes Mass (g)
Log10 Body Mass (g)
Density
σ²v


Bufonoidea
Log10 Testes Mass (g)
Log10 Body Mass (g)
Density
σ²v


Rana and Odorana
Log10 Testes Mass (g)
Log10 Body Mass (g)
Density
σ²v

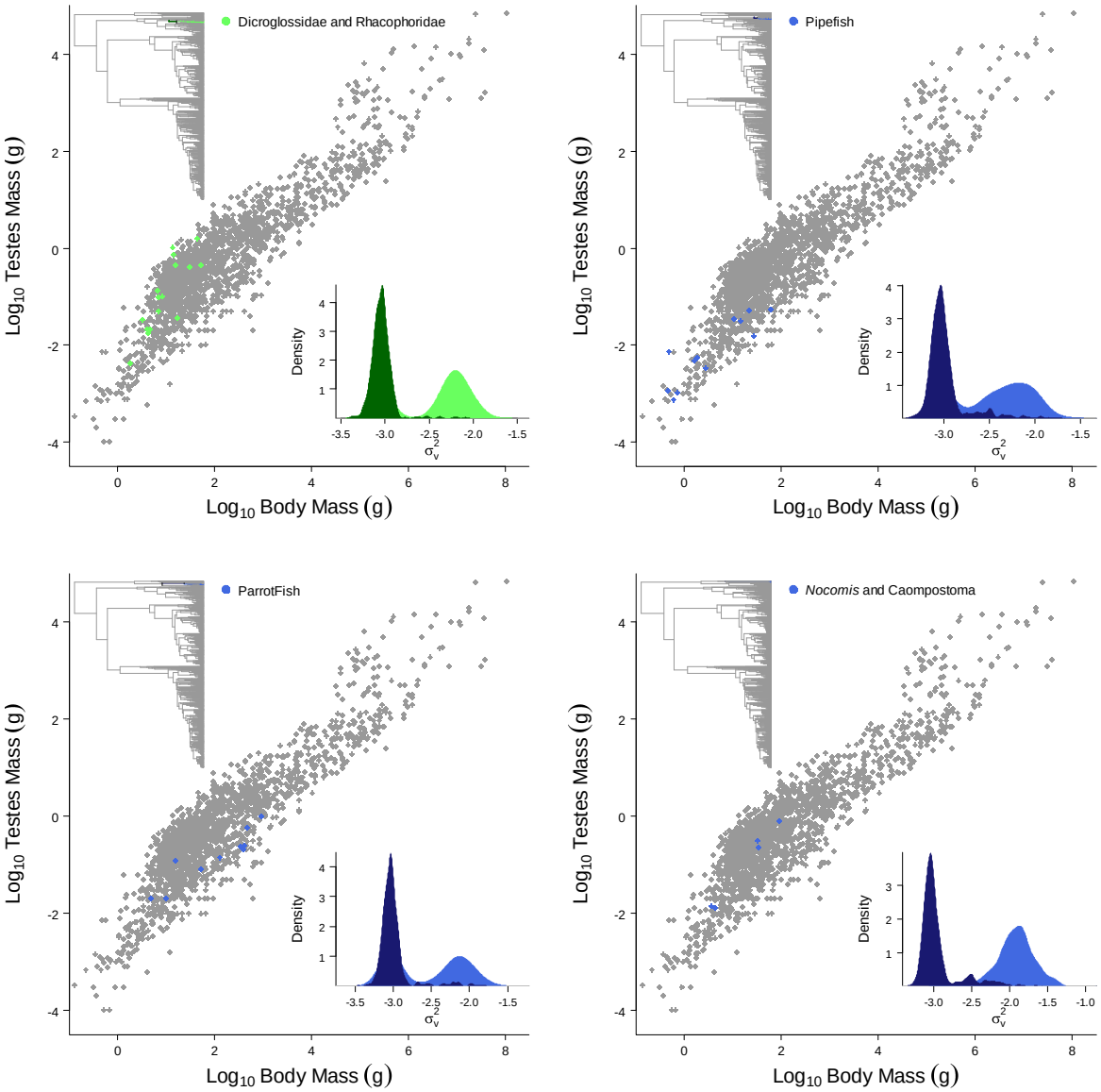


Figure A2.1: The 21 heritable rate shifts observed across the vertebrate tree of life that represent substantial bursts of testes size variance. For each radiation we plot the data, highlighting the clade that has been identified to have an increase in variation and also show where it falls on the phylogeny. We also plot the posterior distribution of rates that are estimated along the branch leading to the clade experiencing the rate acceleration (bright colours) in comparison to the posterior distribution of rates along the directly ancestral branch (darker colours). Also shown is the mean shift observed in *Notomys* which instead of representing a burst of variance represents a substantial testes size reduction. Colours indicate the major vertebrate clade in which the heritable rate shift is found: pink, Aves; purple, Squamata; orange, Mammalia; green, Anura; blue, Actinopterygii.

Chapter 3

Positive phenotypic selection inferred from phylogenies

(Published as: Baker J, Meade A, Pagel M & Venditti C 2016. *Positive phenotypic selection inferred from phylogenies.* Biological Journal of the Linnean Society, 118: 95-115.)

Abstract

Rates of phenotypic evolution vary widely in nature and these rates may often reflect the intensity of natural selection. Here we outline an approach for detecting exceptional shifts in the rate of phenotypic evolution across phylogenies. We introduce a simple new branch-specific metric Δ_V/Δ_B that divides observed phenotypic change along a branch into two components: (1) that attributable to the background rate (Δ_B), and (2) that attributable to departures from the background rate (Δ_V). Where the amount of expected change derived from variation in the rate of morphological evolution doubles that explained by to the background rate ($\Delta_V/\Delta_B > 2$), we identify this as positive phenotypic selection. We apply our approach to six datasets, finding multiple instances of positive selection in each. Our results support the growing appreciation that the traditional gradual view of phenotypic evolution is rarely upheld, with a more episodic view taking its place. This moves focus away from viewing phenotypic evolution as a simple homogeneous process and facilitates reconciliation with macroevolutionary interpretations from a genetic perspective, paving the way to novel insights into the link between genotype and phenotype. The ability to detect positive selection where genetic data are unavailable or unobtainable represents an attractive prospect for extant species, but when applied to fossil data it can reveal patterns of natural selection in deep time that would otherwise be impossible.

Background

George Gaylord Simpson once remarked that: "while individuals live and lineages continue they are ipso facto adapted, and to this extent, at least, adaptation is clearly universal" (Simpson, 1953, p. 161). However, the intensity of natural selection leading to these adaptations can vary and this might be reflected in the rate of evolution. Early burst models of adaptive radiations (Blomberg *et al.*, 2003; Harmon *et al.*, 2010; Pagel, 1999, 2002), for example, are built around this premise.

The literature is now littered with evidence that phenotypic evolution is far from homogeneous (e.g. Eastman *et al.*, 2011; Kratsch and McHardy, 2014; Landis *et al.*, 2013; Rabosky, 2014; Revell *et al.*, 2012; Thomas and Freckleton, 2012; Venditti *et al.*, 2011) – to the extent that rates in excess of 500 times the background rate of evolution have been reported (Puttick *et al.*, 2014). Given this rate heterogeneity and some underlying background rate of evolution, one should be able to identify a class of exceptional shifts in rate that could be indicative of positive phenotypic selection at a macroevolutionary scale.

Here we detect bursts of evolutionary change that can be considered instances of positive phenotypic selection using a novel application of our *variable rates model* for continuously varying data (Venditti *et al.*, 2011). Our approach seeks to automatically identify exceptional shifts in the rate of phenotypic evolution, defined as instances in which accelerated evolutionary change contributes more than half the total amount of expected phenotypic change along a branch. This follows our own work and that of others who have linked the rate of morphological evolution to adaptation (Eastman *et al.*, 2011; Kratsch and McHardy, 2014; Kutsukake and Innan, 2013, 2014; Rabosky, 2014; Revell *et al.*, 2012; Thomas and Freckleton, 2012; Venditti *et al.*, 2011).

We apply our approach across a diverse array of taxonomic groups sampling both extant and extinct species.

Phylogenetic statistical approach

The variable rates model

The variable rates model (Venditti *et al.*, 2011) and others like it (e.g. Eastman *et al.*, 2011; Rabosky, 2014) were originally implemented to detect rate heterogeneity in a single trait. However, many biological traits are closely associated with body size or other morphological traits and so it may be desirable in such cases to study variation in rates of change after accounting for a relationship with another trait or traits. For instance, we might be interested rate changes that have occurred during the evolution of the cerebellum after controlling for its relationship to the neocortex (Barton and Venditti, 2014).

The *variable rates regression model* is an extension of our original variable rates model, designed to detect heterogeneity in the rates of phylogenetically structured residual errors. The method allows for simultaneous estimation of both an overall relationship between the trait of interest and other characters but also any shifts in rate that apply to the residuals. Clades and branches that can be viewed as outliers to the regression line (i.e. with different intercepts or variance) are returned as rate shifts (Figure 1). Note that the original variable rates model is just a special case of the variable rates regression, where one can think of the residual error as from a regression of the trait value about its mean.

A conventional phylogenetic generalised least squares (GLS) model of continuously varying trait evolution (e.g. Pagel, 1999), estimates the instantaneous phenotypic variance of changes (σ^2) that occur along a branch assuming that changes are drawn from an underlying homogenous Brownian process; this parameter is also sometimes referred to as the rate of change (Schluter *et al.*, 1997). The variable rates model works within the same GLS framework but adjusts the lengths of branches in the phylogeny to accommodate areas of the tree where the inferred variance of trait evolution is in excess of (or is less than) that expected from a homogeneous model: increasing the length of a

branch is statistically identical to increasing the variance of changes, and vice versa. The optimal set of stretched and compressed branch lengths yield a tree that conforms to the assumed underlying Brownian motion model of evolution.

The variable rates model partitions the phenotypic variance into two components: a variance parameter that describes the background rate of phenotypic evolutionary change (σ_b^2), and a second set of parameters that identify branch-specific rate shifts. This idea of partitioning the variance is conceptually similar to Bokma's approach to quantifying the homogenous punctuational contribution to phylogenetic variance (Bokma, 2002; Bokma, 2008).

There is no a priori way to know the number of rate shifts hidden within any particular combination of data and phylogeny. The fit to the underlying background rate (σ_b^2) is therefore optimized by estimating a set of branch-specific scalars r ($0 < r > \infty$) that seek to stretch or compress branch lengths relative to the inferred phenotypic variance given σ_b^2 – resulting in our second component: an *optimized variance* for each branch, $\sigma_v^2 = \sigma_b^2 r$. If a branch has experienced greater phenotypic change than would be predicted from the background rate, the rate along that branch must be accelerated and so $r > 1$ ($r < 1$ for branches experiencing less than the expected amount of change). There is no limit or prior expectation imposed on the number of rate scalars, allowing from zero to n parameters with the same probability where n is the number of branches in the phylogeny, making our implementation uniquely placed to study episodes of positive selection.

We simultaneously estimate both the underlying σ_b^2 and all rate scalars in a Bayesian Markov chain Monte Carlo (MCMC) reversible-jump framework (Green, 1995; Pagel and Meade, 2006), generating posterior distributions of r for each branch of the phylogeny. Our method places a gamma prior on each r with an α = 1.1 and a β parameter rescaled such that the median of the distribution is equal to one, i.e. the number of rate increases and decreases proposed is balanced (Venditti *et al.*, 2011).

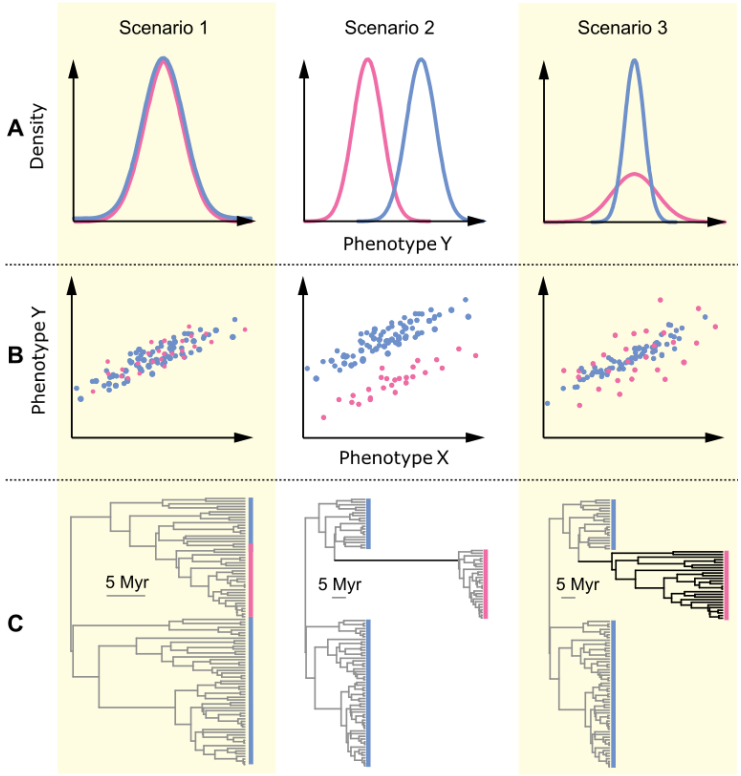


Figure 1: Expectations of how the variable rates model detects rate heterogeneity in a focal clade (pink) under three different evolutionary scenarios (columns). In row (A) we show the distribution of trait data for the focal clade in comparison to the rest of the species in the tree. In row (B) we show the relationship between phenotype Y and phenotype X for the focal clade as compared to the rest of the species in tree. Finally, in row (C) we show how the two distinct forms of rate scalars – single branch modifications and clade modifications – act to stretch (or compress) the branches of a phylogenetic tree given the distribution of data in each scenario. Scenario 1 demonstrates a situation in which there is no rate variation in the data. The focal clade (pink) conforms to the same underlying homogenous Brownian motion model of evolution as all other species in the tree – resulting in a uniform distribution of Phenotype Y at the tips (Scenario 1A). In the form of a regression on phenotype X, the phylogenetically structured residual error is the same for the focal group as it is for the other species in the tree (Scenario 1B). In either of these cases, the variable rates model would detect no rate shifts, and thus no branch scaling would be enforced (Scenario 1C). Scenario 2 demonstrates a situation in which the focal clade (pink) has experienced some jump or shift in the mean value of Phenotype Y (Scenario 2A) – or a shift in the intercept after accounting for its relationship with Phenotype X (Scenario 2B). In both cases, the variable rates model detects a single branch modification that scales the branch leading to the focal clade such that its rate of evolution is increased relative to the rest of the tree (Scenario 2C). Scenario 3 demonstrates a situation in which the focal clade has experienced greater variation in Phenotype Y than would be expected given the variance of the other species in the data (Scenario 3A). In the context of a regression, this can be observed as greater variance in the phylogenetic residuals of a regression on Phenotype X (Scenario 3B). In this case, the variable rates model detects a clade modification, acting to increase the rate of evolution along all branches within the clade of interest (Scenario 3C).

Rate scalars (r) are proposed in two distinct forms: *single branch modifications* and *clade modifications,* given the underlying σ^2_b of the phenotype of interest (e.g. Phenotype Y in Figure 1). If phenotype Y evolved according to a homogenous rate σ^2_b, either independently (Figure 1) or according to some relationship with another phenotypic character(s) (e.g. phenotype X, Figure 1) we expect $r = 1$ for all branches (Figure 1) – the observed tip data can be explained by σ^2_b, and so we expect $\sigma^2_v = \sigma^2_b$.

If an entire clade has experienced a mean shift or jump in trait value (Figure 1), or a shift after accounting for some relationship with phenotype X (Figure 1), then σ^2_b alone is not sufficient to explain the expected phenotypic change that is inferred to have occurred along the branch leading to that clade. The variable rates model in this case would apply a single branch modification ($r > 1, \sigma^2_v > \sigma^2_b$) along the ancestral branch to accommodate the shift in phenotype (Figure 1). One might expect this type of evolutionary change to represent transitions to novel ecological niches or environments (Baker *et al.* 2015, Eastman *et al.* 2013, Venditti *et al.* 2011). For example, in mammals, the branch leading to modern bats has been demonstrated to have a rapid rate of evolution (Venditti *et al.*, 2011) leading to smaller body size (Baker *et al.*, 2015). This reduction in size is associated with the acquisition of a novel aerial lifestyle in this group of animals and interestingly also coincides with strong genetic selection for increased energetic efficiency (Shen *et al.*, 2010; Zhang *et al.*, 2013).

Clade modification scalars scale all branches in a single monophyletic group. If one clade has greater variation than the other (Figure 1) or greater variance in the relationship between phenotype Y and phenotype X (Figure 1) then the expected change predicted by σ^2_b will be inadequate to explain the diversity of observed phenotypes. To explain this, the amount of expected change along each of the branches in the clade must be increased ($r > 1, \sigma^2_v > \sigma^2_b$). The interpretation of each scaled branch in an accelerated clade is similar to those scaled by single branch modifications. If some biological or physiological mechanism acted in a group of species to reduce phenotypic variation e.g.

some form of stabilizing selection, stasis or constraint, we would expect a clade deceleration ($r < 1, \sigma_v^2 < \sigma_b^2$). In our experience such clade decelerations are rare.

In a Bayesian setting 'positive' evidence for rate variation arises from a Bayes factor (BF) taking a value greater than 2 (Table 1, taken from Raftery, 1996). The Bayes Factor is defined as $BF = -2\log_e[m_1/m_0]$, where here m_1 is the marginal likelihood of a variable rates model and m_0 is that of a simple homogeneous Brownian motion model (estimating a single underlying rate). We estimate m_1 and m_0 using stepping-stone sampling (Xie *et al.*, 2010) implemented in BayesTraits (Pagel *et al.*, 2004).

Table 1: Interpretation of Bayes Factors (Raftery, 1996)

Bayes factor $-2\log_e[m_1/m_0]$	**Support**
<0	Negative
0-2	Barely worth mentioning
2-5	Positive
5-10	Strong
>10	Very strong

Phylogenies, dating and uncertainty

The implementation of the variable rates model we use here, like all other similar models (Eastman *et al.*, 2011; Elliot and Mooers, 2014; Kratsch and McHardy, 2014; Landis *et al.*, 2013; Rabosky, 2014; Revell *et al.*, 2012; Thomas and Freckleton, 2012), is confined to using a single phylogenetic tree with meaningful branch lengths. Although mixing may be hard to achieve (Huelsenbeck *et al.*, 2002), there is no theoretical reason that these models could not be adapted to integrate across a posterior sample of phylogenetic trees. This would be an attractive feature as it would allow one to account for phylogenetic uncertainty in terms of topology and divergence times – this will be the subject of future work.

Linked to this is the problem of accurate dating of divergence times themselves (dos Reis *et al.*, 2014; dos Reis *et al.*, 2012; Drummond *et al.*, 2006; Gavryushkina *et al.*, 2014; Heath, 2012; Heath *et al.*, 2014; Meredith *et al.*, 2011; Mooers *et al.*, 2012; Shapiro *et al.*, 2011; Stadler and Yang, 2013), which has further complications if fossil (or other morphological) data are to be included in the tree inference (Lee *et al.*, 2014a; O'Leary *et al.*, 2013; Pyron, 2011; Ronquist *et al.*, 2012; Slater, 2015; Wood *et al.*, 2012). While a detailed discussion of this is beyond the scope of this paper we believe that the recent and intense focus of modern methods for dating trees using and including fossil data are currently revolutionising this field. We look forward to the time when we can incorporate these multiple sources of uncertainty in to our study of phenotypic rates – but for now, at least, we assume that the topology and divergence times of a single phylogenetic tree are reliable.

Detecting positive phenotypic selection

The expected phenotypic variance occurring along a single branch given a gradual background rate is defined by $\Delta_B = \sigma_b^2 t$, where t is the branch length in time. The expected variance along each branch occurring as a consequence of rate variation can be defined as $\Delta_V = \sigma_v^2 t$. Given this, we compare adaptive phenotypic change arising from accelerations or decelerations in rate (Δ_V) to that attributed to the background rate (Δ_B) by calculating the ratio: $\Delta_V/\Delta_B = \sigma_v^2 t/\sigma_b^2 t$.

Although this can be conveniently reduced to r (where $r = \Delta_V/\Delta_B$), we use Δ_V/Δ_B here to highlight the distinction between the expected phenotypic change which occurs as a consequence of rate increases and decreases compared to that associated with the background rate. For each branch we calculate a posterior distribution of Δ_V and Δ_B.

Δ_V/Δ_B, though superficially similar to, and indeed inspired by the well known d_N/d_S ratio widely used for detecting positive genetic selection across phylogenetic trees (Yang, 2002, 2006), has two fundamental differences in interpretation: (1) Δ_B, unlike d_S does not represent neutral evolution – it is the background rate of adaptive evolutionary change

(Blomberg and Garland Jr, 2002; Harvey and Purvis, 1991; Pagel, 1994a); (2) Δ_V and Δ_B, unlike d_S and d_N do not measure two independent rates – instead, Δ_V is not independent of Δ_B ($\Delta_V = \Delta_B r$). Therefore, whereas a $d_N/d_S > 1$ indicates positive selection, where non-synonymous substitutions contribute more than half of the genetic change occurring along a branch, a $\Delta_V/\Delta_B > 1$ simply indicates that there is some non-zero contribution of rapid evolution along that branch (i.e. $\sigma_v^2 > \sigma_b^2$).

With this in mind, we identify an exceptional subclass of branches by drawing parallels to the genetic d_N/d_S ratio – where the phenotypic variance attributed to accelerations in the rate of evolution contributes more than half the total expected phenotypic variance occurring along that branch ($\sigma_v^2 > 2\sigma_b^2$). This brings our interpretation of positive phenotypic selection using $\Delta_V/\Delta_B > 2$ more in line with that of positive genetic selection using $d_N/d_S > 1$. We formally define *positive phenotypic selection* on the basis of two criteria: (1) magnitude; Δ_V/Δ_B must be greater than two and (2) certainty; $\Delta_V/\Delta_B > 2$ must be observed in more than 95% of the posterior distribution of rate scalars for the branch in question.

Instances of significant rate variation ($BF > 2$), even if not identified as exceptional according to our two criteria, provide a unique window into the history of macroevolutionary patterns and processes (Baker *et al.*, 2015; Barton and Venditti, 2013; Barton and Venditti, 2014; Benson and Choiniere, 2013; Eastman *et al.*, 2013; Mahler *et al.*, 2010; Pampush, 2015; Puttick *et al.*, 2014; Rabosky and Adams, 2012; Rabosky *et al.*, 2014; Rabosky *et al.*, 2013; Revell *et al.*, 2012; Venditti *et al.*, 2011). However, this is not the major focus of this work: our goal is to identify exceptional occurrences of rapid evolutionary change ($\Delta_V/\Delta_B > 2$) i.e. positive phenotypic selection.

One could posit three alternative scenarios in which rapid bursts of phenotypic change cannot, in themselves, be considered positive selection: (1) the relaxation of some fundamental constraint restricting phenotypic change could cause a release in the rate of evolution, (2) a founder event or bottleneck causing a severe reduction in phenotypic

diversity could result in a shift in mean phenotype despite a lack of rapid directional evolution (Smith *et al.*, 2015; Yang, 2006), or (3) one may believe that a faster pace of life could result in high rates of morphological evolutionary change. However, we know of no studies that convincingly demonstrate that high rates of change are associated with release from constraint or reduced diversity owing to bottlenecks, and some have suggested little evidence for the former (Baker *et al.*, 2015; Cooper and Purvis, 2009). Intuitively, a faster pace of life is expected to be linked with higher rates of neutral evolution and there is some evidence from molecular studies supporting this idea (e.g. Bromham *et al.*, 1996; Lartillot and Poujol, 2011; Li *et al.*, 1996; Nabholz *et al.*, 2008; Welch *et al.*, 2008). However, there is scant support for a negative association between the speed of life history and either morphological rates of evolution or positive selection in a genetic context (Cooper and Purvis, 2009). As such we suggest that positive selection can be assumed in the absence of contradictory external evidence.

Case studies: trees, datasets and expectations

We studied six datasets for evidence of rate variation (Table 2). We ran Markov chains for at least 100 million iterations in each dataset, sampling every once every 100,000 iterations from the converged chain. The marginal likelihoods m_0 and m_1 are calculated using a stepping stone sampler, sampling 100,000 iterations for each of 1,000 total stones drawn from a beta-distribution ($\alpha = 0.4$ and $\beta = 1$) (Xie *et al.*, 2010).

To ensure our characterisation of positive phenotypic selection is robust, we run a set of simple phylogenetic simulations using R (R Core Team, 2015) and the *phytools* package (Revell, 2012). Using the phylogenetic mean of the focal trait (Phenotype Y in Figure 1), and the background rate σ_b^2 from a standard maximum-likelihood GLS model (Pagel, 1999), we simulate data according to a homogenous Brownian motion model along each time-calibrated phylogeny 100 times.

For traits where we use variable rates regression, we simulate a new dependent variable using the same procedure but use the original independent variable from the real data as the predictor in our analysis to avoid introducing multiple sources of variation. We apply the variable rates (or variable rates regression) model to each of these datasets, treating the simulated data as the focal trait. From these, we can calculate the incidence of false positives and ensure confidence in our results. We expect to find evidence for rate heterogeneity ($BF > 2$, Table 1) in no greater than 5% of each set of simulations.

For five of the six datasets we include here we have identified areas of the phylogeny that may have undergone rapid shifts in morphological evolution (Table 2). We identify whether (1) we detect positive evidence for shifts in the rate of evolution, (2) whether any of those rate shifts correspond to instances of positive phenotypic selection and (3) whether we can identify any instances of positive selection where we have no a priori expectations.

We first discuss our results in relation to the expectations associated with each dataset (Table 2), and then lead into a discussion of how these case studies and our approach fit into the wider context of macroevolutionary biology.

Table 2: Datasets used to test for the presence of positive phenotypic selection using our Δ_V/Δ_B metric ($\Delta_V/\Delta_B > 2$ in 95% of the posterior).

Group	Trait of interest	N	Expectations based on previous work	Expectation Sources	Data Sources	Tree Sources
Anolis lizards	Female snout-vent length, SVL (mm)	100	Multiple rate shifts have been reported in this group but it is unclear whether any of these may represent positive selection.	(Eastman *et al.*, 2013; Revell *et al.*, 2012; Thomas and Freckleton, 2012)	(Thomas and Freckleton, 2012)	(Thomas and Freckleton, 2012)
Angiosperms	Fruit diameter (mm)	250	Fleshy fruited angiosperms exhibit remarkable morphological diversity. They therefore represent an interesting opportunity to explore how natural diversity has accumulated although we have no a priori expectations regarding positive phenotypic selection.	(Jordano, 1995)	(Jordano, 1995)	(Zanne *et al.*, 2014)
Mammals	Semi-circular ear canal radius (mm) *accounting for body size (g)*	234	The ear canal architecture of modern cetaceans is thought to have evolved early and rapidly in their evolutionary history. We may observe positive phenotypic selection leading to this clade.	(Spoor *et al.*, 2002)	(Spoor *et al.*, 2002; Spoor *et al.*, 2007)	(Spoor *et al.*, 2007; Steeman *et al.*, 2009)

Table 2 (cont.):

Group	Trait of interest	N	Expectations based on previous work	Expectation Sources	Data Sources	Tree Sources
Primates	Second lower molar area (mm^2) *accounting for body size (kg)*	74	Previous work has suggested that in order to account for the observed reduction in molar size in the hominin lineage beginning with *Homo erectus*, the rate of evolution would have to be increased by ~50 times. The robust australopithecines belonging to the genus *Paranthropus* are also identified as phylogenetic outliers. We might therefore expect to observe positive selection along each of the branches leading to these species.	(Organ *et al.*, 2011)	(Organ *et al.*, 2011)	(Organ *et al.*, 2011)
Dinosaurs	Humerus length (mm) *in relation to femur length (mm)*	239	There is evidence for rapid morphological evolution in Paraves in particular relating to the relative elongation of the forelimb preceding the emergence of flight. We therefore may observe positive phenotypic selection on relative limb proportions in this group.	(Benson *et al.*, 2014; Dececchi and Larsson, 2013) (Puttick *et al.*, 2014; Turner *et al.*, 2007)	(Benson *et al.*, 2014)	(Benson *et al.*, 2014)

Table 2 (cont.):

Group	Trait of interest	N	Expectations based on previous work	Expectation Sources	Data Sources	Tree Sources
Mammals	Corneal diameter (mm) *in relation to axial eye length (mm)*	242	Anthropoid primates substantially deviate from the pattern of eye shape observed in other mammals, thought to be associated with a shift to diurnal predatory behaviour. We may observe positive selection along this lineage.	(Hall *et al.*, 2012; Kirk, 2006)	(Hall *et al.*, 2012)	(Bininda-Emonds *et al.*, 2007)

All trees used are time-calibrated in the original sources with the exception of the dinosaur tree (Benson *et al.*, 2014). We use the midpoint of age ranges given in Benson et al (2014) to time-scale this tree using branch-sharing methods implemented in the *paleotree* R package (Bapst, 2012), enforcing a minimum branch length of 2 million years. For the mammal ear canal analysis, we grafted the recent well-resolved cetacean phylogeny from Steeman *et al.* (2009) onto the mammal phylogeny used in Spoor *et al.* (2007) using dates from Hedges *et al.* (2006).

Positive phenotypic selection in six case studies

For all datasets we analyse here, we find 'very strong' evidence (Table 1) for heterogeneity in phenotypic evolutionary rate ($BF > 10$, Table 3) in addition to multiple episodes of positive phenotypic selection in all datasets. Our simulations show that the error rate is negligible: considerably less than 5% of each set showed any significant rate variation ($BF > 2$, Table 3).

Table 3: Results of the variable rates model and the application of our metric for positive selection over six case studies. BF = Bayes Factor calculated as $-2\log_e[m_1/m_0]$, (phylogenetic statistical approach); N_{total} = total number of branches in the phylogeny; $N_{positive}$ = number of branches identified as undergoing positive selection; Prop. $BF_{Sim} > 2$ = the proportion of simulated BFs exceeding 2.

Group (trait)	BF	N_{total}	$N_{positive}$	Expectations met?	Prop. $BF_{Sim} > 2$
Anolis (SVL)	12.4	194	10	Partially	0.03
Angiosperms (Fruit diameter)	153.8	498	114	n/a	0.01
Mammals (Ear canal radius)	44.4	466	37	Yes	0.02
Primates (Molar area)	34.1	146	14	Partially	0.00
Dinosaurs (Humerus length)	19.2	454	69	Yes	0.00
Mammals (Corneal size)	104.6	457	101	Yes	0.01

Snout-vent length in *Anolis* lizards

Anolis lizards have been the subject of decades of intense research (Losos, 2009), including studies reporting variation in the rates of phenotypic evolution (Eastman *et al.*, 2013; Revell *et al.*, 2012; Thomas and Freckleton, 2012) linked to ecological transitions to new habitats and environments (Eastman *et al.*, 2013) such as the colonization of new islands (Revell *et al.*, 2012).

We identify ten branches that provide evidence of positive selection on female snout-vent-length (SVL) in the *Anolis* radiation (Figure 2A). Selection along all nine branches in

a group of five species (*A. baleatus, A. barahonae, A. ricordii, A. eugenegrahami, A. christophei*, green, Figure 2B) yields a modal (calculated using kernel density estimation of the posterior distribution) Δ_V/Δ_B ranging from 5.94 - 6.40 (95.5 - 98.6% of the posterior >2). Previous work on this subclade has reported rate shifts along just a couple of these branches (Eastman *et al.*, 2013) whereas another identified these branches as belonging to a part of a larger clade radiation (Thomas and Freckleton, 2012) involving several more species and branches. We also detect positive selection along the branch leading to a group of another five species: *A. equestris, A. luteogularis, A.baracoae, A. noblei* and *A. smallwoodi* (yellow, Figure 2C). This branch has a modal $\Delta_V/\Delta_B = 11.70$ (95.5% of the posterior >2).

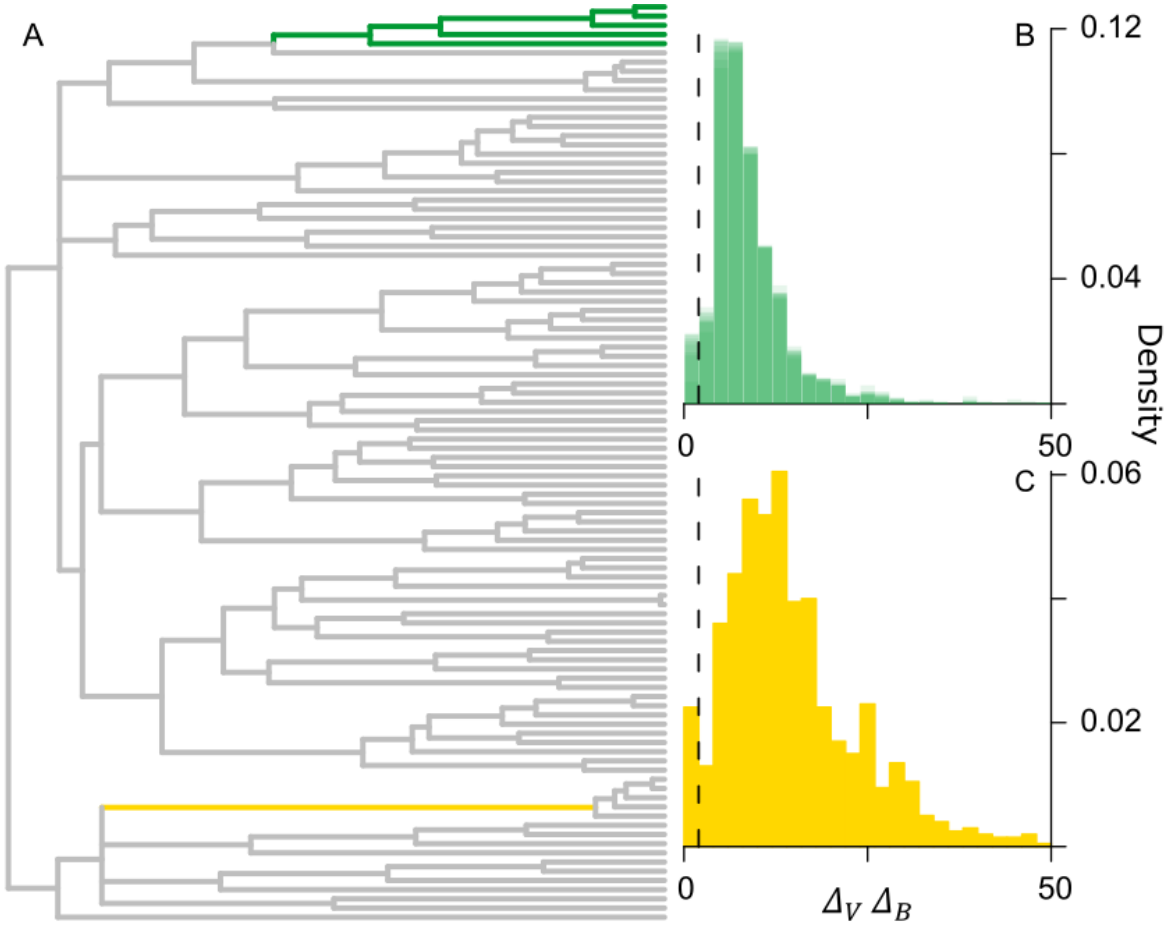


Figure 2: Positive phenotypic selection on female snout-vent-length in *Anolis* lizards. (A) The branches are highlighted in green and yellow. (B) The green branches lead to a clade of five species (*A. ricordii, A. baleatus, A. barahonae, A. eugenegrahami, A. christophei*) and have an average modal Δ_V/Δ_B = 6.07. (C) The yellow branch leads to another five species (*A. equestris, A. luteogularis, A. baracoae, A. noblei, A. smallwoodi*) and has a modal Δ_V/Δ_B = 11.70.

Six of the ten branches on which we do detect positive selection (Figure 2) are within (or leading to) small clades of crown-giant anoles, one of the defining features of which is larger body size than most other species in the tree (Losos, 2009). Our results demonstrate that the large body sizes characterizing this ecomorph tend to arise as the result of strong positive selection. Intriguingly, the only two other crown-giant species not included in those clades within which we detect positive selection (*A. garmani* and *A. cuvieri*) also experience rapid rates of evolutionary change (modal $\Delta_V/\Delta_B = 5.49$ and 5.99 respectively). However, our characterization of positive phenotypic selection is based on two criteria – magnitude and certainty – and these branches fit just the first of the two. These species, among other rapidly evolving branches in the phylogeny not undergoing positive selection according to both of our criteria illustrates one of our major points rather nicely: it is possible to detect exceptional rapid bursts of evolution as a consequence of positive phenotypic selection above and beyond the presence of other interesting rate heterogeneity.

Fleshy-fruited angiosperms and fruit diameter

Despite remarkable variation in morphology (Jordano, 1995), to our knowledge the evolution of fleshy fruits has never been explored using a variable rates model. However, analysing such a dataset has the potential to reveal regimes of selection with no a priori expectations. Applying the variable rates model to the evolution of fruit size indicates significant rate variation littered throughout the phylogeny.

We find that 23% of all branches (114 of 498) have experienced exceptional rapid shifts in morphology indicative of positive phenotypic selection: $\Delta_V/\Delta_B > 2$ in more than 95% of their posterior distribution (Table 4, Figure 3). At least one episode of positive phenotypic selection is found in seven of the 31 orders represented in the tree (Jordano, 1995). These episodes are highlighted on the full phylogeny in Figure 3A where all branches are scaled by the amount of phenotypic evolution that can be attributed solely to exceptional rapid bursts of morphological change (Δ_V- Δ_B). These can be compared to branches

representing phenotypic change that can be attributed to the background rate (Δ_B, Figure 3B).

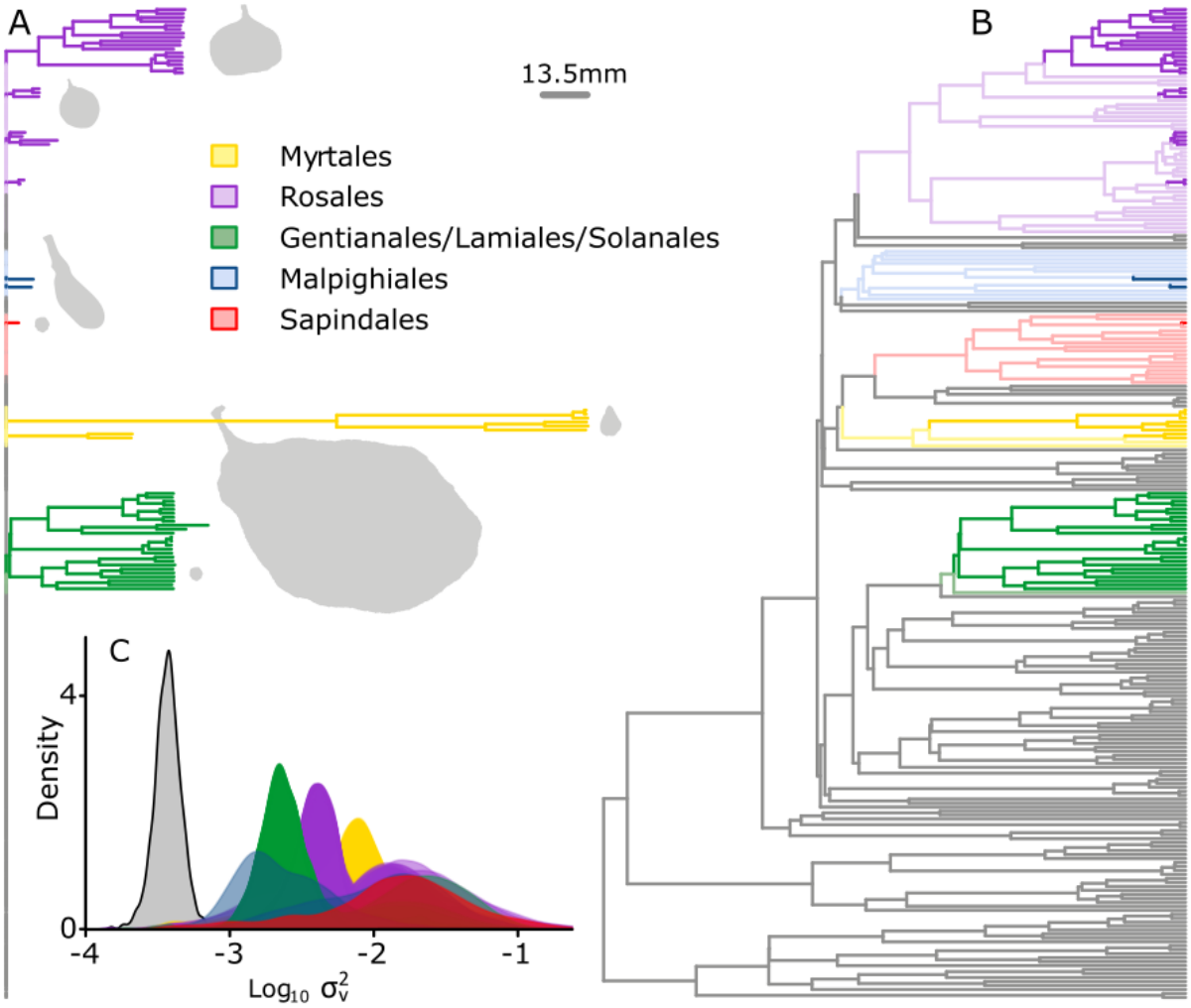


Figure 3: Positive phenotypic selection in fleshy fruit diameter. 23% of all branches in the phylogeny of fleshy fruited angiosperms are identified as having undergone positive phenotypic selection on fruit diameter. Colours indicate orders where branches undergoing positive selection are found: darker branches indicate instances of positive phenotypic selection. (A) Branches of the fruit phylogeny in which we detect positive phenotypic selection are scaled by the amount of phenotypic evolution that can be attributed solely to exceptional rapid bursts of morphological change ($\Delta_V - \Delta_B$). (B) The phenotypic change that can be attributed to the background rate (Δ_B) across the whole phylogeny. (C) The posterior distribution of $\log_{10} \sigma^2_{v,}$ for each of the positively selected branches as compared to the background rate of change ($\log_{10} \sigma^2_{b,}$ grey). Silhouettes represent approximate relative fruit size at the tips and are drawn by the authors and Ciara O'Donovan.

Table 4: Orders undergoing positive phenotypic selection in fleshy-fruited angiosperms. Δ_V/Δ_B reported for only those branches undergoing positive selection. For each order, the average and range of modal Δ_V/Δ_B are presented as well as the range of percentages of the posterior distribution that crosses zero for each branch. ; N_{total} = total number of branches in the phylogeny belonging to the order of interest; $N_{positive}$ = number of branches identified as undergoing positive selection.

Order	N_{total}	$N_{positive}$	Average	Range	% Crossing
Gentianales / Lamiales / Solanales	51	49	5.74	5.31-12.31	98.4-100
Malpighiales	25	2	8.39	3.73-13.05	97.7-99.8
Myrtales	19	14	17.00	14.62-19.42	95.9-100
Rosales	113	48	11.09	6.79-35.14	96.0-100
Sapindales	35	1	20.86	-	97.7

Variation in fruit size across species is often linked to the diversity of vertebrate herbivores that disperse fruit seeds (Burns and Lake, 2009; Jordano, 1995, 2000; Lomáscolo *et al.*, 2010). These, and other pressures (Alcántara and Rey, 2003), can modify fruit size on relatively short timescales (Lord, 2004). The fig genus *Ficus* (Rosales, Table 4) has been described as one of the most important food sources for tropical frugivores (Janzen, 1979; Shanahan *et al.*, 2001), being consumed by over 1200 invertebrate (e.g. ants and crabs) and vertebrate (mammals, fish and birds) species (Shanahan *et al.*, 2001). We identify instances of positive selection acting on fruit size in all 33 branches in the *Ficus* genus, lending support to the proposed key role of dispersers in driving the diversity of fruited angiosperms. Linking phenotypic rates of evolution and fruit morphology as we have here potentially represents a novel approach to the investigation of disperser diversity.

Our results highlight an interesting application of the variable rates model in a context where we did not have any a priori expectations – although we must concede that plant biologists might have been better qualified to make specific predictions. The identification of a number of clades that have experienced episodes of positive phenotypic selection (Table 4) and substantial variation in the strength of these episodes (Figure 3C) highlights potential areas of interest for future research into the evolution of fruit diversity.

Mammalian semi-circular ear canal radius

Cetaceans represent a major transition in mode of life from terrestrial quadrupeds to obligate swimmers (Montgomery *et al.*, 2013; Rabosky, 2014; Steeman *et al.*, 2009; Thewissen *et al.*, 2009; Thewissen and Bajpai, 2001). A rapid and early reduction in the size of the vestibular system of cetaceans (Ketten, 1992a; Spoor *et al.*, 2002) made a return to terrestrial locomotion challenging, potentially paving the way for this extraordinary evolutionary transition (Spoor *et al.*, 2002; Thewissen *et al.*, 2009). Cetaceans then evolved to become obligate aquatic animals with entirely unique inner ear proportions i.e. the semi-circular ear canal of modern cetaceans is nearly three times smaller in radius relative to body size than other mammals (Ketten, 1992a; Spoor *et al.*, 2002; Spoor *et al.*, 2007). For example, although the Indian gazelle (*Gazella bennetti*) and the pygmy right whale (*Caperea marginata*) have very similar canal radii (2.6 mm and 2.5 mm respectively), relative to body size the whale's is drastically smaller (silhouettes, Figure 4).

The otherwise strong positive association between body size and the semi-circular ear canal radius (Figure 4B), exemplifies why it is necessary to control for size in our analyses. If we were to run the variable rates model with the ear canal data in isolation, i.e. as a single trait, any rates recovered would be confounded by size and thus very difficult to interpret. Applying the variable rates regression model, we find that strong positive selection on the branch leading to modern cetaceans (modal Δ_V/Δ_B = 11.23) is associated with a reduction in the relative size of the semicircular canal. This lends support to the idea that the reduction represents a rapid functional modification (Spoor *et al.*, 2002) rather than a vestigial morphology (Ketten, 1992b; Yamato *et al.*, 2012). The rapid, size-independent, rate of evolution we observe along this branch can only be identified by accounting for the shift in the relationship between body size and canal radius in this group of animals (Figure 4).

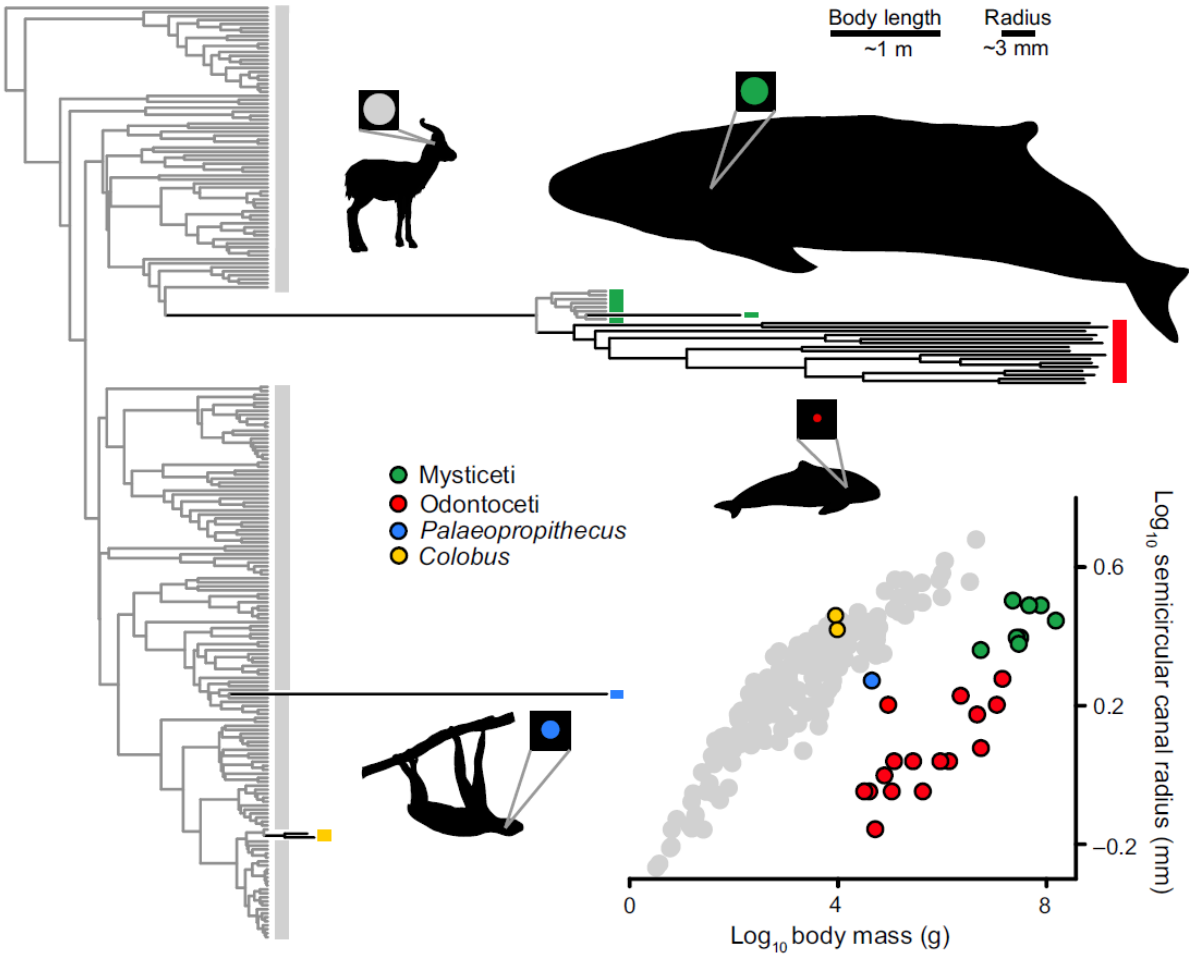


Figure 4: The evolution of the semicircular ear canal in mammals. Each of the branches of the mammal phylogeny undergoing positive selection are scaled by the modal σ_v^2 (black) and all other branches by the modal background rate (σ_b^2, grey). Coloured bars at the tips of the phylogeny indicate clades where rate scalars have acted, and link to the data (inset). The dramatic reduction in the radius of the semi-circular ear canal in cetaceans relative to body size (inset, also see silhouettes) corresponds to a modal Δ_V/Δ_B of 11.23 in the branch leading to all cetaceans. After this initial burst of morphological change, baleen whales (green) experience no further positive selection (with the exception of the acrobatic humpback whale, Megaptera novaeangliae, modal Δ_V/Δ_B = 7.62) but toothed whales (red) experience a continued acceleration of morphological diversification (Δ_V/Δ_B = 6.53, range 7.08-9.55) indicating greater variation in relative ear canal radius for this group (inset). We also observe positive selection in the branch leading to the extinct sloth lemur *Palaeopropithecus* (blue, modal Δ_V/Δ_B = 9.75) and the branches leading to colobus monkeys (yellow, average modal Δ_V/Δ_B = 12.61, range 10.75-15.31) – these species have smaller and larger relative ear canal radiuses as compared to their closest phylogenetic relatives respectively (inset). Silhouettes are taken from phylopic.org or drawn by the authors.

After the initial burst of rapid phenotypic evolution, the radius of the ear canal in baleen whales ceases to experience positive selection (green, Figure 4), with the exception of the acrobatic humpback whale (*Megaptera novaeangliae*, modal Δ_V/Δ_B = 7.62). However, positive selection continued to act on the toothed whales throughout the evolution of the entire clade (red, Figure 4, average modal Δ_V/Δ_B = 6.53, ranging between 7.08- 9.55).

Odontocetes vary dramatically in both habitat and lifestyle, ranging from the deepest-diving extant cetacean, the sperm whale (*Physeter catodon*) to freshwater river-dwelling beasts like the Amazon river dolphin (*Inia geoffrensis*). The vestibular system has been implicated in auditory capacity for low-frequency sounds (Ketten, 1992a, 1992b) and so the development of echolocation behaviour in conjunction with this ecological variability may be further modifying the ear canals in this group of animals. Evidence from other animals supports this idea: echolocating bats tend to exhibit greater inter-specific variation in ear canal size (Davies *et al.*, 2013). It is possible, therefore, that the evolution of multiple novel modes of life within the aquatic environment has placed an extraordinary range of selective regimes on this group of animals.

Unexpectedly, we find that the branch leading to an extinct subfossil sloth lemur *Palaeopropithecus* has experienced intense positive selection (Δ_V/Δ_B = 9.75) resulting in an unusually small ear canal radius for its body size (blue, Figure 4). Ironically, when this species was first discovered, it was thought to be aquatic (Standing, 1903); unfortunately for our narrative this was later shown to be on the basis of misidentified skeletal remains (Godfrey and Jungers, 2003; Jenkins, 1974; Lamberton, 1957). Based on both post-cranial anatomy and the dimensions of the vestibular ear canals, *Palaeopropithecus* is predicted to have had an extraordinarily slow locomotory style, slower even than other closely related sloth lemurs (Godfrey and Jungers, 2003; Jungers *et al.*, 1997; Spoor *et al.*, 2007).

If, for some reason, maintenance of a large semi-circular ear canal system was tremendously costly to an animal, the exploitation of such a novel sluggish lifestyle may have placed selective pressure on this structure to reduce in size. However, this scenario

is very unlikely as *Palaeopropithecus*, along with other sloth lemurs to a slightly lesser extent, possesses remarkable convergent anatomy in many ways with modern extant sloths (Godfrey and Jungers, 2003; Walker *et al.*, 2008), perhaps unsurprisingly given their name. Sloths are well known for their less than lively locomotion, and also possess reduced semi-circular ear canals (Spoor *et al.* 2007; Walker *et al.* 2008), yet we see no evidence for positive selection in this group (Figure 4A). If we can take the *Palaeopropithecus* ear canal data at face value there must therefore be some other adaptive advantage that conferred selection towards smaller relative ear canal radius in *Palaeopropithecus* that is yet to be identified.

Primates, hominins and molar area

Numerous factors have been posited to explain the human lineage's unique trajectory (Maslin *et al.*, 2014; McHenry and Coffing, 2000; Potts, 1998; Wrangham, 2009). For example, the drastic reduction of relative molar area in the evolutionary lineage leading to our own species (Figure 5) is believed to be associated with the innovation of food processing such as cooking (Organ *et al.*, 2011; Wrangham, 2009).

It has been suggested that the marked reduction in molar size observed in *Homo erectus, H. sapiens* and *H. neanderthalensis*, might have required a 50-fold greater rate of evolution than that observed for molar size evolution in all other primates (Organ *et al.*, 2011). These three species, along with two species of *Paranthropus* (*P. boisei* and *P. robustus*) have been identified as phylogenetic outliers in terms of their molar area (Organ *et al.*, 2011). Unlike the three species of *Homo*, which are outliers on the basis of their relatively tiny molars, the two robust australopithecines (*Paranthropus*) instead have the largest molars of any known hominins (Grine, 2007). Therefore, we might expect to observe two independent areas of positive phenotypic selection: along the branches leading to the two species of *Paranthropus,* and along the branches leading to the three *Homo* species.

Although we do find that all of the expected branches experience positive phenotypic selection, we actually find this as a part of a larger clade radiation: all branches within the monophyletic group encompassing these species (the five expected species plus *H. rudolfensis* and *H. habilis*) are undergoing selection (average modal Δ_V/Δ_B = 9.48, range 8.64-10.92 Figure 5).

All *Homo* species have much smaller molar areas than *Paranthropus,* but the positive selection observed in the hominin branches does not indicate a clear trend towards smaller sizes within the genus as a whole (Figure 5). However, in line with results from Organ *et al.* (2011) and the timing of novel cooking behaviour, we observe a pronounced size reduction in the branches associated with *H. erectus, H. sapiens* and *H. neanderthalensis* (Figure 5), although the rate increases we detect (average modal Δ_V/Δ_B = 10.21, range 9.65-10.92) are lower than the predicted 50-fold in Organ *et al.* (2011). This may have something to do with the fact that all rate shifts we observe are estimated simultaneously.

Our approach provides the opportunity to explore the relative contribution of rapid rates of evolution to the realised amount of phenotypic change and to characterize the nature of positive selection acting on individual branches. For example, we identify that the exceptional bursts of rapid phenotypic evolution we observe contribute >90% of phenotypic change along each branch. For the hominin clade, we visualize the effects of these shifts by simulating the amount of expected change in molar area predicted by σ_b^2 and comparing to simulations incorporating the observed positive selection (σ_v^2). Assuming a starting change of zero, we sum the expected change $\sigma^2 t$ along each lineage. We find that a model in which species are allowed to evolve only at the rate of change expected by the background predicts a smaller modal amount of molar size change. For example, along the branch leading to *H. sapiens,* the background rate predicts a decrease of 1.28 mm (±95th percentiles 1.10, 1.56). However, when we allow the species to experience the greater amount of evolutionary change that our model predicts, the trait decreases by a value of 2.04 mm (±95th percentiles 1.40, 4.14). We plot these simulations

for all branches in the hominin lineage beginning with the branch leading to the clade coloured pink in Figure 5B.

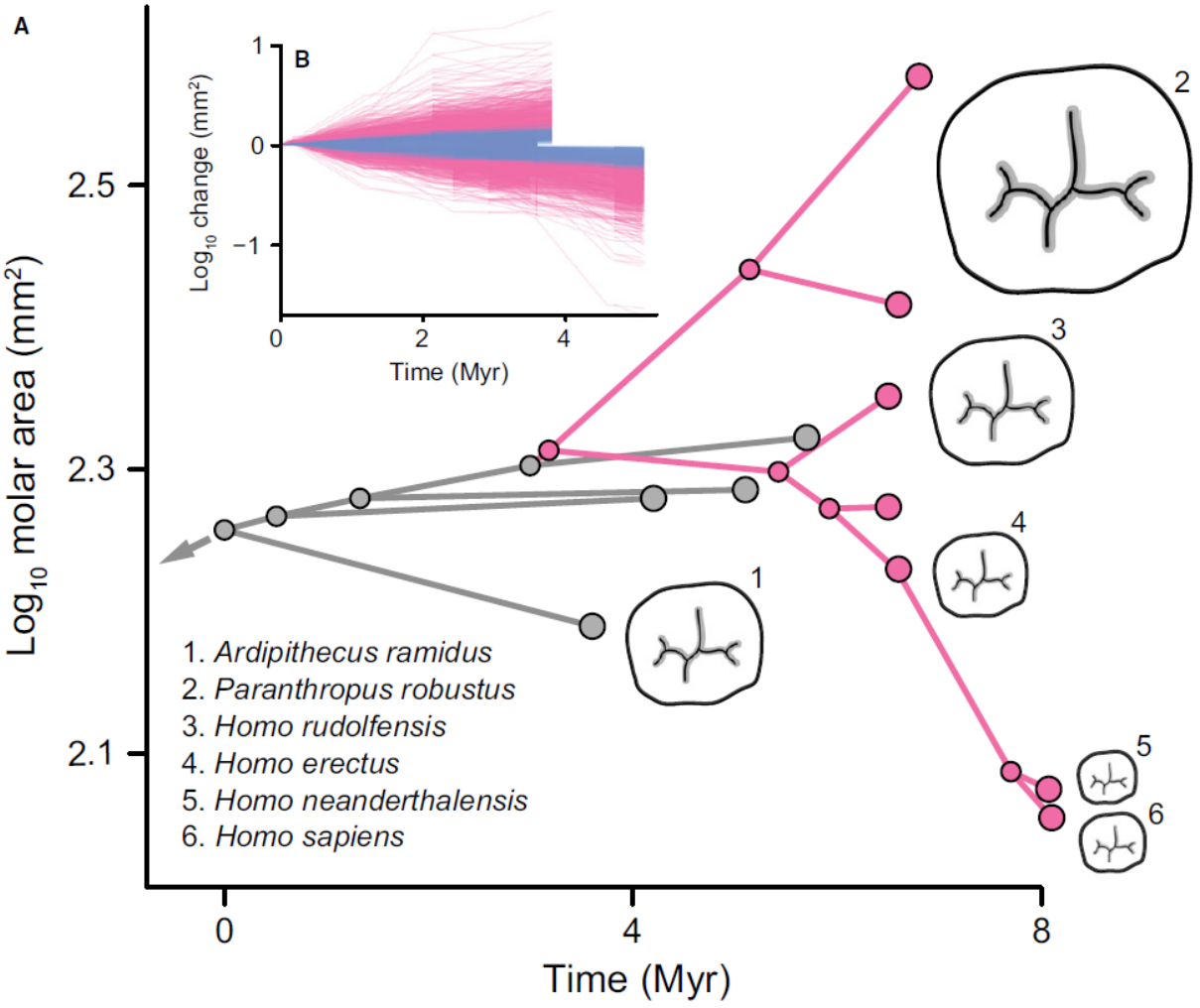


Figure 5: Positive selection for decreased molar size in the human lineage and visualization of expected phenotypic change. (A) The hominin portion of the primate phylogeny is plotted in a phylomorphospace documenting the evolution of molar area through time. Ancestral nodes (small circles) are reconstructed using restricted maximum likelihood in the R package 'ape' (Paradis *et al.*, 2004) on the phylogeny where each of the branches experiencing positive selection has been scaled by its modal Δ_V/Δ_B. We detect exceptional positive selection along 14 branches within the hominin clade (pink) associated with shifts in molar area relative to body size (tooth outlines 1-6). (B) We simulate the amount of expected change along this lineage according to the elevated σ^2_v (pink) and compare it to simulations using the gradual background rate (σ^2_b, blue). We assume that all change in these simulations is in the direction corresponding to that predicted by the ancestral state reconstructions. The grey arrow leads to the rest of primates, where we find no other instances of significant positive selection with the exception of *Papio cynocephalus* (modal Δ_V/Δ_B= 15.31). Tooth outlines are scaled according to approximate relative molar area and were drawn by the authors.

Relative limb proportions in Dinosauria

The origin of modern birds from their dinosaurian ancestors is one of the most celebrated and scrutinized evolutionary transitions (Benson *et al.*, 2014; Benson and Choiniere, 2013; Brusatte *et al.*, 2014; Chiappe, 2009; Dececchi and Larsson, 2013; Puttick *et al.*, 2014; Turner *et al.*, 2007; Zhou, 2004). Particularly in recent years, the rate of morphological change leading to the explosion of avian diversity we see today has been a hot-bed of activity and controversy (Benson *et al.*, 2014; Benson and Choiniere, 2013; Brusatte *et al.*, 2014; Lee *et al.*, 2014b; Puttick *et al.*, 2014). The consensus view is that the traits typically associated with the rise of modern avian diversity actually began to evolve far earlier (Brusatte et al. 2014; Lee *et al.* 2014b; Puttick *et al.* 2014; Turner *et al.* 2007).

In avian theropods the forelimb is elongated relative to body size (Dececchi and Larsson, 2013; Hutchinson and Allen, 2009; Ostrom, 1974; Padian and Chiappe, 1998). There is some evidence that forelimb segment proportions were more constrained within Maniraptora (including birds and their paravian predecessors) when compared to other theropods (Benson and Choiniere, 2013), while hindlimb segments evolved at a faster rate within Avialae (Benson and Choiniere, 2013) and an overall relative elongation of the forelimb occurred in paravians (Dececchi and Larsson, 2013; Puttick *et al.*, 2014). We take a wide taxonomic perspective across all Dinosauria (including 59 sauropods, 93 theropods and 87 ornithischians) to determine if the observed change in relative forelimb length arose as the consequence of positive selection for morphological change in avian or paravian dinosaurs (using humerus and femur length to characterize forelimb and hind limb respectively). Putting the clade of interest in a broad phylogenetic context allows us to reliably estimate a background rate of limb evolution that may have otherwise been difficult in the smaller theropod clade.

Within Paraves we find positive selection not only along the branch leading to all paravians (modal Δ_V/Δ_B = 3.42), but also further episodes of directional selection along 84% of this clade's branches (the modal Δ_V/Δ_B for each branch ranges between 2.85 and 3.60, Figure

6: the average modal Δ_V/Δ_B = 2.95) – the exceptions to this are the Enantiornithines and *Epidipteryx* (grey in Figure 6). We find no evidence that avian theropods experience any rate increase over that experienced over all paravians (Figure 6). Outside of Paraves, our variable rates approach finds no positive phenotypic selection in the rest of dinosaurs, with the exception of the theropod *Caudipteryx* (modal Δ_V/Δ_B = 3.00).

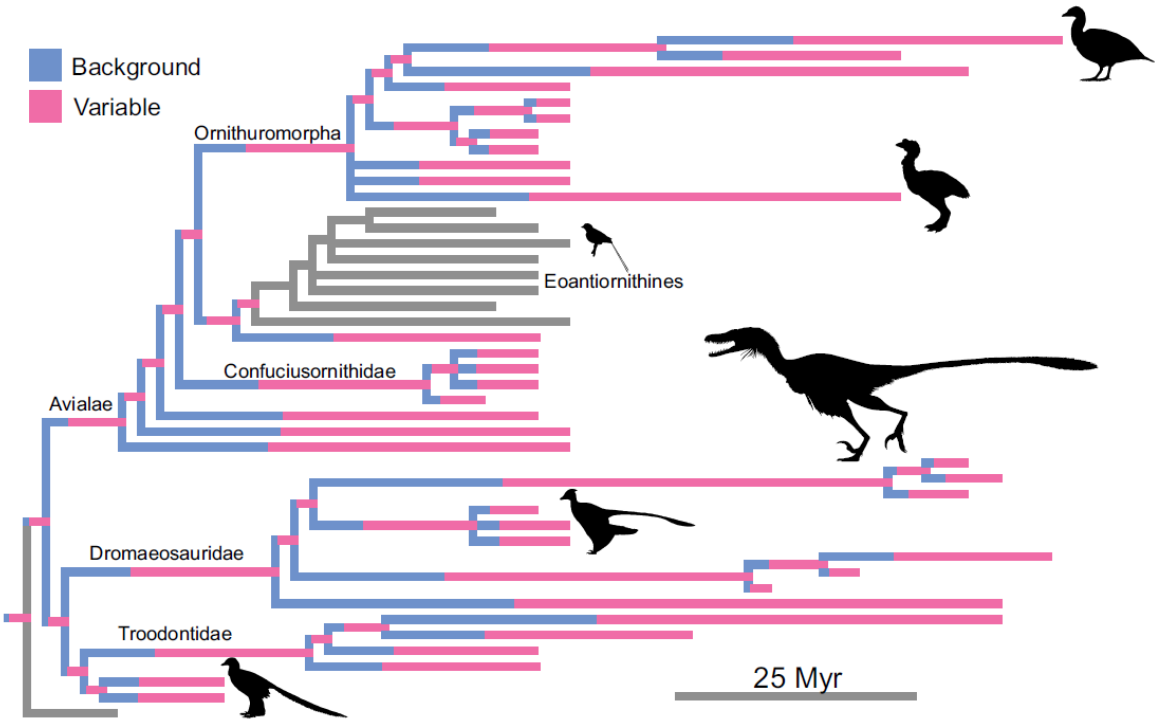


Figure 6: The evolution of relative humeral length in paravian theropod dinosaurs can be explained mostly by rapid shifts in morphology. We plot here only the paravian portion of the dinosaur phylogeny – the grey arrow leads to the rest of theropods and Dinosauria where there is no evidence for positive selection with the exception of *Caudipteryx* (modal Δ_V/Δ_B = 3.00). Branches are coloured in proportion to the amount of phenotypic change occurring in that lineage attributable to the gradual background rate (Δ_B, blue) compared to that explained by variation in the rate of phenotypic evolution (Δ_V, pink). Only those branches experiencing positive selection are coloured – all other branches are grey. Silhouettes of species at the tips are roughly proportional in size and are obtained from phylopic.org or drawn by the authors.

Overall, the amount of phenotypic change attributed to positive selection in paravians is far greater than that which can be explained by the background rate of dinosaurian evolution: rapid bursts of change contribute substantially to morphological diversity in this group (Figure 6). This is in line with recent results from theropod dinosaurs which suggest that morphological changes leading to the evolution of flight preceded the origin

of avian diversity (Lee *et al.*, 2014b; Puttick *et al.*, 2014) although at a much more biologically realistic rate compared to other studies (Puttick *et al.*, 2014).

Mammalian eye shape and visual acuity

Unlike the eyes of all other vertebrates, the shape of a mammalian eye does not reliably predict a species' activity pattern (e.g. nocturnality, diurnality) (Hall and Ross, 2007; Hall, 2008; Hall *et al.*, 2012; Kirk, 2006). The exception to this rule is anthropoid primates (haplorrhines to the exclusion of tarsiers) – this diurnal group of species all have relatively larger eyes and smaller corneas for their body size and are notably different from other mammals (Ross *et al.*, 2007; Ross and Kirk, 2007, Figure 7). The nocturnal niche of early mammals would have promoted the evolution of relatively large eye sizes to improve visual sensitivity in a light-poor context (Hall *et al.*, 2012; Heesy and Hall, 2010). A shift to a novel diurnal lifestyle reliant on visual predation at the origin of anthropoid primates would have favoured enhanced visual acuity and improved daylight hunting success (Kirk and Kay, 2004; Ross, 1996; Ross and Kirk, 2007).

In line with this expectation, we identify positive phenotypic selection towards smaller corneas relative to axial eye length in the branch leading to anthropoid primates (modal Δ_V/Δ_B = 4.47, Figure 7). After this initial shift in morphology, positive selection is observed along all branches within new world monkeys (to the exclusion of Atelidae, modal Δ_V/Δ_B = 4.21, range 3.87-5.22) and also mandrills and mangabeys (average modal Δ_V/Δ_B = 4.29, range 3.82-4.74).

Unexpectedly, we also observe positive phenotypic selection along all branches within Carnivora (Figure 7), where the modal Δ_V/Δ_B ranges from 4.34-43.23 (average modal Δ_V/Δ_B = 6.41). The range of these effects manifests as greater variation among Carnivores in relative corneal size (around their regression line with axial length) compared to other mammals (Figure 7 and compare with Figure 1): phylogenetic sum of square errors ($SSE_{carnivores}$ = 0.00806 compared to $SSE_{mammals}$ = 0.00528 – after removing primates).

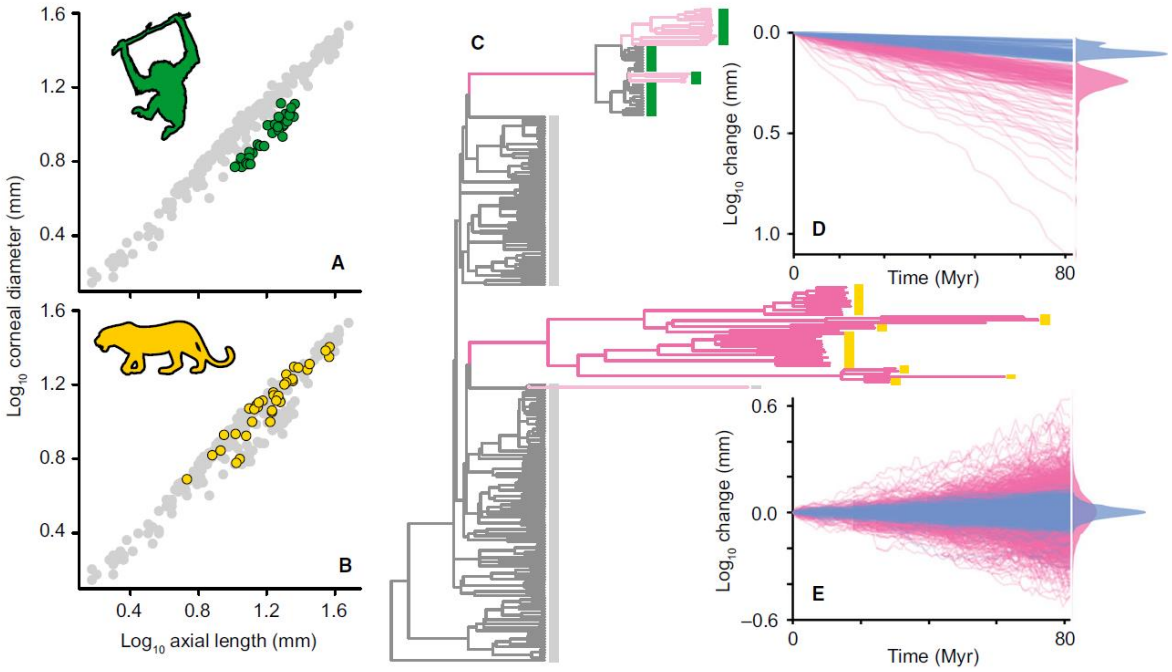


Figure 7: Positive selection on relative corneal diameter in extant mammals. The relationship between axial eye length and corneal diameter is shown for all mammals (A, B), emphasizing points for anthropoid primates in (A) and carnivores in (B). According to this relationship, we find significant positive phenotypic selection in 102 of 457 branches. (C) We scale each of these 102 branches by modal σ^2_v (pink) and all other branches by the modal background rate (σ^2_b, grey). Bars at the tips of the trees correspond to coloured points in (A) and (B). We simulate the expected change occurring along those branches shaded in darker pink and plot the results in (D) for the branch leading to primates and (E) for the carnivoran branches. Anthropoid primates have a much smaller corneal diameter than expected for their eye length (A). This results in a rate increase of more than 4 times the background along the branch leading to this monophyletic clade (dark pink, C). We simulate the amount of expected phenotypic change along this lineage according to the elevated modal σ^2_v (pink) and compare it to simulations using the gradual background rate (σ^2_b, blue) assuming negative directional change. The resulting distribution of data at the end of the branch demonstrates a shift, or jump, in the amount of expected change in relative corneal diameter (C). After this initial shift in morphology, subsequent positive selection is observed in new world monkeys (with the exclusion of Atelidae) and also mandrills and mangabeys – these branches are shaded in light pink in (C). Carnivores exhibit greater variation in relative corneal size compared to all other mammal orders (B, see text). Consequently, rapid phenotypic diversification and positive selection is detected along all of the branches in this clade – coloured dark pink in (C). As with the primates, we simulate the amount of expected change using both the variable and background rate, but this time assume non-directional change (E). These simulations highlight that the amount of expected change in this clade has a much greater variance than that expected by the background model. Note that we also observe a rate increase in the black-bellied pangolin *Manis tetradactyla* (pale pink, B). Silhouettes are taken from phylopic.org.

Research on the relationship between corneal size and axial lengths is highly primate centric (e.g. Kirk and Kay, 2004; Ross *et al.*, 2007; Ross and Kirk, 2007). However, the unexpected pattern we reveal may emerge as a consequence of the high diversity in activity pattern observed in carnivores. In the data we use for our analyses (Hall *et al.*, 2012), the number of cathemeral, nocturnal and diurnal species observed within carnivores are approximately equal – this is not observed in any other order except rodents. A phylogenetic analysis of the transition rates between activity patterns in each of the orders of mammals (Pagel, 1994b) where sample sizes allowed (n >10, six orders) suggests that carnivores have undergone far higher transition rates: at least 3 times higher than any other group. This variation in activity pattern and the potential interaction with other visual adaptations that are likely to be important in this order (Heesy, 2005, 2008; Heesy and Hall, 2010; Stevens, 2006) might explain why we observe such high rates of change and highlight this group as a potential area for future research.

The nature of positive selection we observe is qualitatively identical even after controlling for the non-significant effect of generation time on corneal diameter (generation length data from Pacifici *et al.* (2013) as an additional covariate in the model): we see positive phenotypic selection acting along exactly the same branches as before. This provides further evidence that the pace of life history does not necessarily predict the rate of morphological change (Cooper and Purvis, 2009). The ability to control for other factors that may confound our pursuit of rate heterogeneity and instances of positive phenotypic selection highlights one of the major advantages of the variable rates regression model that we describe here.

To demonstrate the effect of positive phenotypic selection on mammalian corneal diameter, we simulate phenotypic evolution along the branch leading to anthropoid primates (single branch modification) and along all branches in carnivores (clade modification). We simulate according to the elevated σ_v^2 and compare the amount of change to that from simulations using the gradual background rate (σ_b^2). For the branch leading to anthropoid primates, we can infer that there has been rapid directional change

towards smaller relative corneal size – as all anthropoids are shifted in this morphology (Figure 7). Given σ^2_v, we observe a modal directional phenotypic change in corneal size of -1.70 mm (± 95th percentiles -1.47,-4.27, Figure 7). This is greater than that observed when we simulate using the background rate σ^2_b (modal change = -1.29 mm ± 95th percentiles -1.09,-1.38, Fig 7) along the same branch, even if we simulate all change as directional. This is in line with the shift that we observe in the real data (Figure 7). Because we cannot infer a single direction of change along all branches in the Carnivora, we simulate non-directional phenotypic evolution and plot the resulting distribution of trait data at the tips. Data simulated in this way using σ^2_v have a much wider variation (σ = 0.121, Figure 7) than those simulated using the background σ^2_b (σ = 0.038, Figure 7).

Discussion

Positive phenotypic selection as defined by our criteria is present in all datasets we analyse and we suspect – given the preponderance of rate heterogeneity in many biological groups and traits – that it is widespread and common in nature (e.g. Benson and Choiniere, 2013; Eastman *et al.*, 2011; Eastman *et al.*, 2013; Rabosky, 2014; Rabosky *et al.*, 2014; Rabosky *et al.*, 2013; Steeman *et al.*, 2009; Venditti *et al.*, 2011). To the extent this is true, it has the potential to change the way many biologists think about phenotypic evolution. In the wake of the case-studies we report here and the realisation of the limitations of simplistic models (see also Hunt *et al.*, 2015; Uyeda *et al.*, 2011), an attractive new episodic view of evolution emerges where selection acts in a far more idiosyncratic manner (see for example Venditti *et al.* 2011). This provides a macroevolutionary perspective for phenotypic evolution that is consistent with the macroevolutionary picture emerging from genetic data (Murrell *et al.*, 2015; Murrell *et al.*, 2012; Nadeau and Jiggins, 2010; Nikaido *et al.*, 2014; Smith *et al.*, 2015; Yang, 2002, 2006). As such, models which allow for non-homogeneity of the nature we demonstrate should become the default option in the analysis of comparative datasets, replacing analyses comparing simplistic homogenous

models (e.g. Benson *et al.*, 2014; Burbrink *et al.*, 2012; Cardillo, 2015; Cooper and Purvis, 2010; Harmon *et al.*, 2010; Huttenlocker, 2014; Slater, 2013, 2015; Sookias *et al.*, 2012), to name just a few). The magnitude and certainty of the exceptional rates we uncover here in addition to other interesting rate heterogeneity is certain to mislead these simple models (Venditti *et al.* 2011) and could affect conclusions about historical evolutionary processes.

The existence of wide variation in rates of evolution, and of methods that can detect it, opens an interesting avenue for research seeking to link genotype and phenotype. Currently, studies investigating this link examine the relationship between the magnitude of some phenotypic trait and the rate of genetic evolution, typically as measured by d_N/d_S (Lartillot and Poujol, 2011; Nadeau *et al.*, 2007). This can work where natural selection has acted to alter the phenotype directionally, e.g. increased plumage pigmentation in birds (Nadeau *et al.*, 2007), or expansion of the brain during primate evolution (Montgomery *et al.*, 2011; Montgomery and Mundy, 2012), but will be less useful where phenotypic change has not been directional. By comparison, linking the phenotypic rate of change (Δ_V) to the genetic rate of non-synonymous changes (d_N) on a branch-by-branch basis might provide the key to unlocking hidden associations between genotype and phenotype in the absence of sustained directionality across multiple branches.

Whilst studying positive selection at a genetic level is perhaps the most direct way to study adaptation, in some cases molecular data can never be brought to bear, such as will be true for most fossils. In these cases our approach may be the only way to study the historical signals of selection.

References

Alcántara JM & Rey PJ 2003. Conflicting selection pressures on seed size: evolutionary ecology of fruit size in a bird-dispersed tree, *Olea europaea*. *Journal of Evolutionary Biology,* 16 (6): 1168-1176.

Baker J, Meade A, Pagel M & Venditti C 2015. Adaptive evolution toward larger size in mammals. *Proceedings of the National Academy of Sciences USA,* 112 (16): 5093-5098.

Bapst DW 2012. paleotree: an R package for paleontological and phylogenetic analyses of evolution. *Methods in Ecology and Evolution,* 3 (5): 803-807.

Barton RA & Venditti C 2013. Human frontal lobes are not relatively large. *Proceedings of the National Academy of Sciences USA,* 110 (22): 9001-9006.

Barton RA & Venditti C 2014. Rapid evolution of the cerebellum in humans and other great apes. *Current Biology,* 24 (20): 2440-2444.

Benson RBJ, Campione NE, Carrano MT, Mannion PD, Sullivan C, Upchurch P & Evans DC 2014. Rates of dinosaur body mass evolution indicate 170 million years of sustained ecological innovation on the avian stem lineage. *PLoS Biology,* 12 (5): e1001853.

Benson RBJ & Choiniere JN 2013. Rates of dinosaur limb evolution provide evidence for exceptional radiation in Mesozoic birds. *Proceedings of the Royal Society B: Biological Sciences,* 280 (1768): e20131780.

Bininda-Emonds OR, Cardillo M, Jones KE, Macphee RD, Beck RM, Grenyer R, Price SA, Vos RA, Gittleman JL & Purvis A 2007. The delayed rise of present-day mammals. *Nature,* 446 (7135): 507-12.

Blomberg SP & Garland Jr T 2002. Tempo and mode in evolution: phylogenetic inertia, adaptation and comparative methods. *Journal of Evolutionary Biology,* 15: 899-910.

Blomberg SP, Garland Jr T & Ives AR 2003. Testing For phylogenetic signal in comparative data: Behavioural traits are more labile. *Evolution,* 57: 717-745.

Bokma F 2002. Detection of punctuated equilibrium from molecular phylogenies. *Journal of Evolutionary Biology,* 15 (6): 1048-1056.

Bokma F 2008. Detection of "punctuated equilibrium" by Bayesian estimation of speciation and extinction rates, ancestral character states, and rates of anagenetic and cladogenetic evolution on a molecular phylogeny. *Evolution,* 62 (11): 2718-2726.

Bromham L, Rambaut A & Harvey PH 1996. Determinants of rate variation in mammalian DNA sequence evolution. *Journal of Molecular Evolution,* 43 (6): 610-621.

Brusatte SL, Lloyd GT, Wang SC & Norell MA 2014. Gradual assembly of avian body plan culminated in rapid Rates of evolution across the dinosaur-bird transition. *Current Biology,* 24 (20): 2386-2392.

Burbrink FT, Chen X, Myers EA, Brandley MC & Pyron RA 2012. Evidence for determinism in species diversification and contingency in phenotypic evolution during adaptive radiation. *Proceedings of the Royal Society B: Biological Sciences,* 279 (1748): 4817-4826.

Burns KC & Lake B 2009. Fruit–frugivore interactions in two southern hemisphere forests: allometry, phylogeny and body size. *Oikos,* 118 (12): 1901-1907.

Cardillo M 2015. Geographic range shifts do not erase the historic signal of speciation in mammals. *The American Naturalist,* 185 (3): 343-353.

Chiappe LM 2009. Downsized dinosaurs: the evolutionary transition to modern birds. *Evolution: Education and Outreach,* 2 (2): 248-256.

Cooper N & Purvis A 2009. What factors shape rates of phenotypic evolution? A comparative study of cranial morphology of four mammalian clades. *Journal of Evolutionary Biology,* 22: 1024-1035.

Cooper N & Purvis A 2010. Body size evolution in mammals: Complexity in tempo and mode. *The American Naturalist,* 175 (6): 727-738.

Davies KTJ, Bates PJJ, Maryanto I, Cotton JA & Rossiter SJ 2013. The evolution of bat vestibular systems in the face of potential antagonistic selection pressures for flight and echolocation. *PLoS ONE,* 8 (4): e61998.

Dececchi TA & Larsson HCE 2013. Body and limb size dissociation at the origin of birds: uncoupling allometric constraints across a macroevolutionary transition. *Evolution,* 67 (9): 2741-2752.

Dos Reis M, Donoghue PCJ & Yang Z 2014. Neither phylogenomic nor palaeontological data support a Palaeogene origin of placental mammals. *Biology Letters,* 10 (1): e20131003.

Dos Reis M, Inoue J, Hasegawa M, Asher RJ, Donoghue PCJ & Yang Z 2012. Phylogenomic datasets provide both precision and accuracy in estimating the timescale of placental mammal phylogeny. *Proceedings of the Royal Society B: Biological Sciences,* 279 (1742): 3491-3500.

Drummond AJ, Ho SYW, Phillips MJ & Rambaut A 2006. Relaxed phylogenetics and dating with confidence. *PLoS Biology,* 4 (5): e88.

Eastman JM, Alfaro ME, Joyce P, Hipp AL & Harmon LJ 2011. A novel comparative method for identifying shifts in the rate of character evolution on trees. *Evolution,* 65 (12): 3578-89.

Eastman JM, Wegmann D, Leuenberger C & Harmon LJ 2013. Simpsonian 'evolution by jumps' in an adaptive radiation of *Anolis* lizards. arXiv preprint.

Elliot MG & Mooers AO 2014. Inferring ancestral states without assuming neutrality or gradualism using a stable model of continuous character evolution. *BMC Evolutionary Biology,* 14 (226).

Gavryushkina A, Welch D, Stadler T & Drummond AJ 2014. Bayesian inference of sampled ancestor trees for epidemiology and fossil calibration. *PLoS Computational Biology,* 10 (12).

Godfrey LR & Jungers WL 2003. The extinct sloth lemurs of Madagascar. *Evolutionary Anthropology,* 12 (6): 252-263.

Green PJ 1995. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika,* 82 (4): 711-732.

Grine FE 2007. *Evolutionary history of the "robust" Australopithecines,* New Jersey, Transaction Publishers.

Hall M & Ross C 2007. Eye shape and activity pattern in birds. *Journal of Zoology,* 271 (4): 437-444.

Hall MI 2008. Comparative analysis of the size and shape of the lizard eye. *Zoology,* 111 (1): 62-75.

Hall MI, Kamilar JM & Kirk EC 2012. Eye shape and the nocturnal bottleneck of mammals. *Proceedings of the Royal Society B: Biological Sciences,* 279 (1749): 4692-4968.

Harmon LJ, Losos JB, Jonathan Davies T, Gillespie RG, Gittleman JL, Bryan Jennings W, Kozak KH, Mcpeek MA, Moreno-Roark F, Near TJ, Purvis A, Ricklefs RE, Schluter D, Schulte Ii JA, Seehausen O, Sidlauskas BL, Torres-Carvajal O, Weir JT & Mooers AO 2010. Early bursts of body size and shape evolution are rare in comparative data. *Evolution,* 64 (8): 2385-96.

Harvey PH & Purvis A 1991. Comparative methods for explaining adaptations. *Nature,* 351: 619-624.

Heath TA 2012. A hierarchical Bayesian model for calibrating estimates of species divergence times. *Systematic Biology,* 61 (5): 793-809.

Heath TA, Huelsenbeck JP & Stadler T 2014. The fossilized birth–death process for coherent calibration of divergence-time estimates. *Proceedings of the National Academy of Sciences USA,* 111 (29): 2957-2966.

Hedges SB, Dudley J & Kumar S 2006. TimeTree: a public knowledge-base of divergence times among organisms. *Bioinformatics,* 22 (23): 2971-2972.

Heesy CP 2005. Function of the mammalian postorbital bar. *Journal of morphology,* 264 (3): 363-380.

Heesy CP 2008. Ecomorphology of orbit orientation and the adaptive significance of binocular vision in primates and other mammals. *Brain, behavior and evolution,* 71 (1): 54.

Heesy CP & Hall MI 2010. The nocturnal bottleneck and the evolution of mammalian vision. *Brain, behavior and evolution,* 75 (3): 195.

Huelsenbeck JP, Larget B, Miller RE & Ronquist F 2002. Potential applications and pitfalls of Bayesian inference of phylogeny. *Systematic Biology,* 51 (5): 673-688.

Hunt G, Hopkins MJ & Lidgard S 2015. Simple versus complex models of trait evolution and stasis as a response to environmental change. *Proceedings of the National Academy of Sciences USA,* 112 (16): 4885-4890.

Hutchinson JR & Allen V 2009. The evolutionary continuum of limb function from early theropods to birds. *Naturwissenschaften,* 96 (4): 423-448.

Huttenlocker AK 2014. Body size reductions in nonmammalian eutheriodont therapsids (Synapsida) during the end-Permian mass extinction. *PLoS ONE,* 9 (2): e87553.

Janzen DH 1979. How to be a fig. *Annual Review of Ecology and Systematics,* 10: 13-51.

Jenkins FA 1974. *Primate locomotion,* London, Academic Press.

Jordano P 1995. Angiosperm fleshy fruits and seed dispersers: a comparative analysis of adaptation and constraints in plant-animal interactions. *The American Naturalist,* 145 (2): 163-191.

Jordano P 2000. Fruits and frugivory. *In:* Fenner M (ed.) *Seeds: the ecology of regeneration in plant communities.* Wallingford, UK: CABI Publishing.

Jungers WL, Godfrey LR, Simons EL & Chatrath PS 1997. Phalangeal curvature and positional behavior in extinct sloth lemurs (Primates, Palaeopropithecidae). *Proceedings of the National Academy of Sciences USA,* 94 (22): 11998-12001.

Ketten DR 1992a. The cetacean ear: Form, frequency, and evolution. *In:* Thomas J, Kastelein R & Supin A (eds.), *Marine mammal sensory systems.* US: Springer.

Ketten DR 1992b. The marine mammal ear: specializations for aquatic audition and echolocation. *In:* Webster AJ (ed.) *The evolutionary biology of hearing.* New York: Springer.

Kirk EC 2006. Effects of activity pattern on eye size and orbital aperture size in primates. *Journal of Human Evolution,* 51 (2): 159-170.

Kirk EC & Kay RF 2004. The evolution of high visual acuity in the Anthropoidea. *In:* Ross C & Kay RF (eds.), *Anthropoid Origins.* New York: Kluwer Academic/Plenum Publishers.

Kratsch C & Mchardy AC 2014. RidgeRace: ridge regression for continuous ancestral character estimation on phylogenetic trees. *Bioinformatics,* 30 (17): i527-i533.

Kutsukake N & Innan H 2013. Simulation-based likelihood approach for evolutionary models of phenotypic traits on phylogeny. *Evolution,* 67 (2): 355-367.

Kutsukake N & Innan H 2014. Detecting phenotypic selection by Approximate Bayesian Computation in phylogenetic comparative methods. *In:* Garamszegi LZ (ed.) *Modern Phylogenetic Comparative Methods and Their Application in Evolutionary Biology.* Berlin: Springer-Verlag.

Lamberton C 1957. Examen de quelques hypotheses de Sera concernant les lémuriens fossilles et actuels. *Bull Acad Malgache,* 34: 51-65.

Landis MJ, Schraiber JG & Liang M 2013. Phylogenetic analysis using Lévy processes: finding jumps in the evolution of continuous traits. *Systematic Biology,* 62 (2): 193-204.

Lartillot N & Poujol R 2011. A phylogenetic model for investigating correlated evolution of substitution rates and continuous phenotypic characters. *Molecular Biology and Evolution,* 28 (1): 729-744.

Lee MS, Cau A, Naish D & Dyke GJ 2014a. Morphological clocks in paleontology, and a mid-Cretaceous origin of crown Aves. *Systematic biology,* 63 (3): 442-449.

Lee MSY, Cau A, Naish D & Dyke GJ 2014b. Sustained miniaturization and anatomical innovation in the dinosaurian ancestors of birds. *Science,* 345 (6196): 562-566.

Li WH, Ellsworth DL, Krushkal J, Chang BHJ & Hewettemmett D 1996. Rates of nucleotide substitution in primates and rodents and the generation time effect hypothesis. *Molecular Phylogenetics and Evolution,* 5 (1): 182-187.

Lomáscolo SB, Levey DJ, Kimball RT, Bolker BM & Alborn HT 2010. Dispersers shape fruit diversity in *Ficus* (Moraceae). *Proceedings of the National Academy of Sciences USA,* 107 (33): 14668-14672.

Lord JM 2004. Frugivore gape size and the evolution of fruit size and shape in southern hemisphere floras. *Austral Ecology,* 29 (4): 430-436.

Losos JB 2009. *Lizards in an evolutionary tree: ecology and adaptive radiation of anoles,* London, England, University of California Press.

Mahler DL, Revell LJ, Glor RE & Losos JB 2010. Ecological opportunity and the rate of morphological evolution in the diversification of greater Antillean anoles. *Evolution,* 64 (9): 2731-2745.

Maslin MA, Brierley CM, Milner AM, Shultz S, Trauth MH & Wilson KE 2014. East African climate pulses and early human evolution. *Quaternary Science Reviews,* 101: 1-17.

Mchenry HM & Coffing K 2000. *Australopithecus* to *Homo*: transformations in body and mind. *Annual Review of Anthropology,* 29 (6): 125-146.

Meredith RW, Janečka JE, Gatesy J, Ryder OA, Fisher CA, Teeling EC, Goodbla A, Eizirik E, Simão TLL, Stadler T, Rabosky DL, Honeycutt RL, Flynn JJ, Ingram CM, Steiner C, Williams TL, Robinson TJ, Burk-Herrick A, Westerman M, Ayoub NA, Springer MS & Murphy WJ 2011. Impacts of the Cretaceous Terrestrial Revolution and K-Pg extinction on mammal diversification. *Science,* 334 (6055): 521-524.

Montgomery SH, Capellini I, Venditti C, Barton RA & Mundy NI 2011. Adaptive evolution of four microcephaly genes and the evolution of brain size in anthropoid primates. *Molecular Biology and Evolution,* 28 (1): 625-638.

Montgomery SH, Geisler JH, Mcgowen MR, Fox C, Marino L & Gatesy J 2013. The evolutionary history of cetacean brain and body size. *Evolution,* 67 (11): 3339-3353.

Montgomery SH & Mundy NI 2012. Evolution of ASPM is associated with both increases and decreases in brain size in primates. *Evolution,* 66 (3): 927-932.

Mooers A, Gascuel O, Stadler T, Li H & Steel M 2012. Branch lengths on birth–death trees and the expected loss of phylogenetic diversity. *Systematic Biology,* 61 (2): 195-203.

Murrell B, Weaver S, Smith MD, Wertheim JO, Murrell S, Aylward A, Eren K, Pollner T, Martin DP & Smith DM 2015. Gene-wide identification of episodic selection. *Molecular biology and evolution,* 32 (5): 1365-1371.

Murrell B, Wertheim JO, Moola S, Weighill T, Scheffler K & Kosakovsky Pond SL 2012. Detecting individual sites subject to episodic diversifying selection. *PLoS Genetics,* 8 (7): e1002764.

Nabholz B, Glemin S & Galtier N 2008. Strong variations of mitochondrial mutation rate across mammals - the longevity hypothesis. *Molecular Biology and Evolution,* 25 (1): 120-130.

Nadeau NJ, Burke T & Mundy NI 2007. Evolution of an avian pigmentation gene correlates with a measure of sexual selection. *Proceedings of the Royal Society B: Biological Sciences,* 274 (1620): 1807-1813.

Nadeau NJ & Jiggins CD 2010. A golden age for evolutionary genetics? Genomic studies of adaptation in natural populations. *Trends in Genetics,* 26 (11): 484-492.

Nikaido M, Ota T, Hirata T, Suzuki H, Satta Y, Aibara M, Mzighani SI, Sturmbauer C, Hagino-Yamagishi K & Okada N 2014. Multiple episodic evolution events in V1R receptor genes of East-African cichlids. *Genome Biology and Evolution,* 6 (5): 1135-1144.

O'leary MA, Bloch JI, Flynn JJ, Gaudin TJ, Giallombardo A, Giannini NP, Goldberg SL, Kraatz BP, Luo ZX, Meng J, Ni X, Novacek MJ, Perini FA, Randall ZS, Rougier GW, Sargis EJ, Silcox MT, Simmons NB, Spaulding M, Velazco PM, Weksler M, Wible JR & Cirranello AL 2013. The placental mammal ancestor and the post-K-Pg radiation of placentals. *Science,* 339 (6120): 662-7.

Organ C, Nunn CL, Machanda Z & Wrangham RW 2011. Phylogenetic rate shifts in feeding time during the evolution of *Homo*. *Proceedings of the National Academy of Sciences USA,* 108 (35): 14555-14559.

Ostrom JH 1974. *Archaeopteryx* and the origin of flight. *Quarterly Review of Biology,* 49 (1): 27-47.

Pacifici M, Santini L, Di Marco M, Baisero D, Francucci L, Marasini GG, Visconti P & Rondinini C 2013. Generation length for mammals. *Nature Conservation,* 5: 89-94.

Padian K & Chiappe LM 1998. The origin of birds and their flight. *Scientific American,* 278 (2): 28-37.

Pagel M 1994a. The adaptationist wager. *In:* Eggleton P & Vane-Wright RI (eds.), *Phylogenetics and ecology.* London: Academic Press.

Pagel M 1994b. Detecting correlated evolution on phylogenies: a general method for the comparative analysis of discrete characters. *Proceedings of the Royal Society B: Biological Sciences,* 255 (1342): 37-45.

Pagel M 1999. Inferring the historical patterns of biological evolution. *Nature,* 401: 877-884.

Pagel M 2002. Modelling the evolution of continuously varying characters on phylogenetic trees. *In:* Macleod N & Forey PL (eds.), *Morphology, shape and phylogeny.* USA: CRC Press.

Pagel M & Meade A 2006. Bayesian analysis of correlated evolution of discrete characters by reversible-jump Markov chain Monte Carlo. *The American Naturalist,* 167 (6): 808-825.

Pagel M, Meade A & Barker D 2004. Bayesian estimation of ancestral character states on phylogenies. *Systematic Biology,* 53 (5): 673-84.

Pampush JD 2015. Selection played a role in the evolution of the human chin. *Journal of Human Evolution,* 82: 127-136.

Paradis E, Claude J & Strimmer K 2004. APE: Analyses of Phylogenetics and Evolution in R language. *Bioinformatics,* 20 (2): 289-290.

Potts R 1998. Environmental hypotheses of hominin evolution. *American Journal of Physical Anthropology,* 107 (s27): 93-136.

Puttick MN, Thomas GH & Benton MJ 2014. High rates of evolution preceded the origin of birds. *Evolution,* 68 (5): 1497-1510.

Pyron RA 2011. Divergence time estimation using fossils as terminal taxa and the origins of Lissamphibia. *Systematic Biology,* 60 (4): 466-481.

R Core Team 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing.

Rabosky DL 2014. Automatic detection of key innovations, rate shifts, and diversity-dependence on phylogenetic trees. *PLoS One,* 9 (2): e89543.

Rabosky DL & Adams DC 2012. Rates of morphological evolution are correlated with species richness in salamanders. *Evolution,* 66 (6): 1807-1818.

Rabosky DL, Donnellan SC, Grundler M & Lovette IJ 2014. Analysis and visualization of complex macroevolutionary dynamics: an example from Australian scincid lizards. *Systematic biology,* 63 (4): 610-627.

Rabosky DL, Santini F, Eastman J, Smith SA, Sidlauskas B, Chang J & Alfaro ME 2013. Rates of speciation and morphological evolution are correlated across the largest vertebrate radiation. *Nature Communications,* 4.

Raftery AE 1996. Hypothesis testing and model selection. *In:* Gilks WR, Richardson S & Spiegelhalter DJ (eds.), *Markov Chain Monte Carlo in Practice.* London, Great Britain: Chapman & Hall.

Revell LJ 2012. phytools: an R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution,* 3: 217-223.

Revell LJ, Mahler DL, Peres-Neto PR & Redelings BD 2012. A new phylogenetic method for identifying exceptional phenotypic diversification. *Evolution,* 66 (1): 135-146.

Ronquist F, Klopfstein S, Vilhelmsen L, Schulmeister S, Murray DL & Rasnitsyn AP 2012. A total-evidence approach to dating with fossils, applied to the early radiation of the Hymenoptera. *Systematic Biology,* 61 (6): 973-999.

Ross C 1996. Adaptive explanation for the origins of the Anthropoidea (primates). *American Journal of Primatology,* 40 (3): 205-230.

Ross CF, Hall MI & Heesy CP 2007. Were basal primates nocturnal? Evidence from eye and orbit shape. *In:* Ravosa MJ & Dagosto M (eds.), *Primate origins: adaptations and evolution.* New York: Springer.

Ross CF & Kirk EC 2007. Evolution of eye size and shape in primates. *Journal of Human Evolution,* 52 (3): 294-313.

Schluter D, Price T, Mooers AØ & Ludwig D 1997. Likelihood of ancestor states in adaptive radiation. *Evolution,* 51 (6): 1699-1711.

Shanahan M, So S, Compton SG & Corlett R 2001. Fig-eating by vertebrate frugivores: a global review. *Biological Reviews of the Cambridge Philosophical Society,* 76 (4): 529-572.

Shapiro B, Ho SY, Drummond AJ, Suchard MA, Pybus OG & Rambaut A 2011. A Bayesian phylogenetic method to estimate unknown sequence ages. *Molecular Biology and Evolution,* 28 (2): 879-87.

Shen Y-Y, Liang L, Zhu Z-H, Zhou W-P, Irwin DM & Zhang Y-P 2010. Adaptive evolution of energy metabolism genes and the origin of flight in bats. *Proceedings of the National Academy of Sciences USA,* 107 (19): 8666-8671.

Simpson GG 1953. *The major features of evolution,* London, GB, Columbia University Press.

Slater GJ 2013. Phylogenetic evidence for a shift in the mode of mammalian body size evolution at the Cretaceous-Palaeogene boundary. *Methods in Ecology and Evolution,* 4 (8): 734-744.

Slater GJ 2015. Iterative adaptive radiations of fossil canids show no evidence for diversity-dependent trait evolution. *Proceedings of the National Academy of Sciences USA,* 112 (16): 4897-4902.

Smith MD, Wertheim JO, Weaver S, Murrell B, Scheffler K & Pond SLK 2015. Less is more: an adaptive branch-site random effects model for efficient detection of episodic diversifying selection. *Molecular biology and evolution,* 32 (5): 1342-1353.

Sookias RB, Butler RJ & Benson RB 2012. Rise of dinosaurs reveals major body-size transitions are driven by passive processes of trait evolution. *Proceedings of the Royal Society B: Biological Sciences,* 279 (1736): 2180-7.

Spoor F, Bajpai S, Hussain ST, Kumar K & Thewissen JG 2002. Vestibular evidence for the evolution of aquatic behaviour in early cetaceans. *Nature,* 417 (6885): 163-166.

Spoor F, Garland T, Krovitz G, Ryan TM, Silcox MT & Walker A 2007. The primate semicircular canal system and locomotion. *Proceedings of the National Academy of Sciences USA,* 104 (26): 10808-10812.

Stadler T & Yang Z 2013. Dating phylogenies with sequentially sampled tips. *Systematic Biology,* 62 (5): 674-688.

Standing H 1903. Rapport sur des ossements sub-fossiles provenant d'Ampasambazimba. *Bull Acad Malgache,* 2: 227-235.

Steeman ME, Hebsgaard MB, Fordyce RE, Ho SY, Rabosky DL, Nielsen R, Rahbek C, Glenner H, Sørensen MV & Willerslev E 2009. Radiation of extant cetaceans driven by restructuring of the oceans. *Systematic Biology,* 58 (6): 573-585.

Stevens KA 2006. Binocular vision in theropod dinosaurs. *Journal of Vertebrate Paleontology,* 26 (2): 321-330.

Thewissen J, Cooper LN, George JC & Bajpai S 2009. From land to water: the origin of whales, dolphins, and porpoises. *Evolution: Education and Outreach,* 2 (2): 272-288.

Thewissen JGM & Bajpai S 2001. Whale origins as a poster child for macroevolution: Fossils collected in the last decade document the ways in which Cetacea (whales, dolphins, and porpoises) became aquatic, a transition that is one of the best documented examples of macroevolution in mammals. *BioScience,* 51 (12): 1037-1049.

Thomas GH & Freckleton RP 2012. MOTMOT: Models of trait macroevolution on trees. *Methods in Ecology and Evolution,* 3 (1): 145-151.

Turner AH, Pol D, Clarke JA, Erickson GM & Norell MA 2007. A basal Dromaeosaurid and size evolution preceding Avian flight. *Science,* 317 (5843): 1378-1381.

Uyeda JC, Hansen TF, Arnold SJ & Pienaar J 2011. The million-year wait for macroevolutionary bursts. *Proceedings of the National Academy of Sciences USA,* 108 (38): 15908-15913.

Venditti C, Meade A & Pagel M 2011. Multiple routes to mammalian diversity. *Nature,* 479 (7373): 393-396.

Walker A, Ryan TM, Silcox MT, Simons EL & Spoor F 2008. The semicircular canal system and locomotion: The case of extinct lemuroids and lorisoids. *Evolutionary Anthropology: Issues, News, and Reviews,* 17 (3): 135-145.

Welch JJ, Bininda-Emonds OR & Bromham L 2008. Correlates of substitution rate variation in mammalian protein-coding sequences. *BMC evolutionary biology,* 8 (1): 53.

Wood HM, Matzke NJ, Gillespie RG & Griswold CE 2012. Treating fossils as terminal taxa in divergence time estimation reveals ancient vicariance patterns in the palpimanoid spiders. *Systematic biology,* 62 (2): 264-284.

Wrangham R 2009. *Catching fire: How cooking made us human,* New York, Basic Books.

Xie W, Lewis PO, Fan Y, Kuo L & Chen M-H 2010. Improving marginal likelihood estimation for Bayesian phylogenetic model selection. *Systematic Biology,* 60 (2): 150-160.

Yamato M, Ketten DR, Arruda J, Cramer S & Moore K 2012. The auditory anatomy of the minke whale (*Balaenoptera acutorostrata*): A potential fatty sound reception pathway in a baleen whale. *The Anatomical Record,* 295 (6): 991-998.

Yang Z 2002. Inference of selection from multiple species alignments. *Current Opinion in Genetics & Development,* 12 (6): 688-694.

Yang Z 2006. *Computational molecular evolution,* Oxford, Great Britain, Oxford University Press.

Zanne AE, Tank DC, Cornwell WK, Eastman JM, Smith SA, Fitzjohn RG, Mcglinn DJ, O/'Meara BC, Moles AT, Reich PB, Royer DL, Soltis DE, Stevens PF, Westoby M, Wright IJ, Aarssen L, Bertin RI, Calaminus A, Govaerts R, Hemmings F, Leishman MR, Oleksyn J, Soltis PS, Swenson NG, Warman L & Beaulieu JM 2014. Three keys to the radiation of angiosperms into freezing environments. *Nature,* 506 (7486): 89-92.

Zhang G, Cowled C, Shi Z, Huang Z, Bishop-Lilly KA, Fang X, Wynne JW, Xiong Z, Baker ML, Zhao W, Tachedjian M, Zhu Y, Zhou P, Jiang X, Ng J, Yang L, Wu L, Xiao J, Feng Y, Chen Y, Sun X, Zhang Y, Marsh GA, Crameri G, Broder CC, Frey KG, Wang LF & Wang J 2013. Comparative analysis of bat genomes provides insight into the evolution of flight and immunity. *Science,* 339 (6118): 456-60.

Zhou Z 2004. The origin and early evolution of birds: discoveries, disputes, and perspectives from fossil evidence. *Naturwissenschaften,* 91 (10): 455-471.

Chapter 4

Mosaic evolution of hominin semicircular canals and the rise of bipedalism

Abstract

Anatomical features associated with the origin of bipedalism in hominins – species that have emerged since the common ancestor of modern humans and chimpanzees – have received almost endless attention in the literature. Increases in the size of the vertical anterior and posterior semicircular canals coupled with a reduction in the size of the horizontally oriented lateral canal are thought to be associated with the evolution of obligate bipedal locomotion in hominins yet it remains to be seen whether these changes arose as a consequence of strong natural selection. If bipedalism imposed strong selection pressures on semicircular canal size, it should be possible to detect a signal for this in the form of rapid bursts of evolutionary change. Here, we present the first comprehensive phylogenetic analysis of semicircular canals that includes all available data for extinct hominins and explicitly studies phenotypic change in the context of natural selection. In species considered to be unequivocally obligate bipeds, we find that the anterior canal adaptively increases in size relative not only to body size but also to the other canals. For the posterior canal, no positive selection within hominins has acted to increase the size. A decrease in the size of the lateral canal occurred earlier than expected, within *Australopithecus*, and strong positive selection has acted almost constantly to change its size throughout hominin evolution. Our results indicate that the semicircular canals have the capacity to evolve independently under natural selection – although the functional significance of this remains unclear. Positive phenotypic selection has acted to sculpt the morphology of the semicircular canals independently – demonstrating that the semicircular canals evolve in a mosaic fashion. The pattern of evolution we observe in the semicircular canals highlights the growing appreciation that complex morphologies may have evolved in more intricate ways than previously thought.

Introduction

Organs found within the inner ear are vitally important for balance and locomotion (Day and Fitzpatrick, 2005; Ekdale, 2016). One component of the vestibular apparatus is the semicircular canal system which is essential for detecting and compensating for the direction of rotational head movements (Day and Fitzpatrick, 2005). It comprises three semicircular canals – bony structures containing three fluid-filled ducts (anterior, posterior, and lateral) that are oriented such that they approximately encompass three planes of directional motion (Day and Fitzpatrick, 2005; Purves *et al.*, 2001) and therefore allow detection of motion within a three-dimensional space (Day and Fitzpatrick, 2005; Jones and Spells, 1963; Spoor, 2003). As the head is tilted or moved in one way or another, endolymph fluid within the canals is shifted along the plane of movement, triggering both visual and cognitive reflexes for efficient spatial orientation in combination with responses from other vestibular organs (Sipla, 2007; Spoor, 2003).

All jawed vertebrates have three more-or-less orthogonally oriented canals (Janvier, 2001); the evolution of the horizontal canal in the earliest jawed vertebrates (Janvier, 2001; Sipla, 2007 cites Retzius 1881, 1884 as the first recorded evidence) occurred over 300 million years ago in the early Paleozoic (Janvier, 1996). However, differences in the size, shape and orientation of the semicircular canals among species may represent functional specializations and adaptations to novel environments or behaviours (e.g. Billet *et al.*, 2012; Ekdale, 2013; Georgi, 2008; Lebrun *et al.*, 2010; Malinzak, 2010; Malinzak *et al.*, 2012; Pfaff *et al.*, 2015; Ryan *et al.*, 2012; Spoor *et al.*, 2002; Spoor *et al.*, 2007; Spoor *et al.*, 1994, 1996). Differences in morphology may reflect adaptive changes in response to changes in environmental circumstances (e.g. Chapter 5 demonstrates that differences in the eye shape of mammals arise as a consequence of shifts in activity pattern) though this has yet to be ascertained in terms of the semicircular canals. In cetaceans there has been a drastic reduction in the overall size of the semicircular canal system relative to body size (Spoor *et al.*, 2002) that has explicitly been shown to be the result of strong historical positive

selection (Baker *et al.* 2016). Although the specific set of adaptive pressures acting to shape vestibular morphology within cetaceans remains controversial (Ekdale, 2013, 2016; Kandel and Hullar, 2010; Ketten, 1992; Spoor *et al.*, 2002) on the whole aquatic mammals tend to have smaller semicircular canal systems for their body size (Ekdale, 2013).

Specifically, the size and orientation of the semicircular ducts are implicated in sensitivity to external stimuli (Hullar, 2006; Malinzak, 2010; Malinzak *et al.*, 2012; Muller, 1994; Ten Kate *et al.*, 1970; Yang and Hullar, 2007). Orthogonally oriented canals allow for detection of movements evenly in all planes (Malinzak *et al.*, 2012) while larger canals and thus larger ducts (Spoor *et al.*, 1994) allow more fine-tuned responses to movement (Ten Kate *et al.*, 1970). As early as 1908 it was suggested that changes in the sensitivity of the semicircular canals may have implications for behaviour and that the slow, sluggish movement of sloths may be linked to their small canal sizes (Gray, 1908). Within individual species, variation in canal size varies in its ability to predict the sensitivity of the vestibular responses (Hullar, 2006) and orientation has been reported to be a more important predictor of sensitivity than size in a small group (n = 11) of strepsirrhine primates (Malinzak *et al.*, 2012). However in larger comparative analyses of primates and other mammals, canal size has been demonstrated to be linked to greater sensitivity and agility (Cox and Jeffery, 2010; Spoor *et al.*, 2007).

Great apes possess uniquely small semicircular canals when compared to other primates (Ryan *et al.*, 2012; Spoor *et al.*, 2007; Spoor *et al.*, 1994; 1996, Figure 1). A decrease in the overall size and thus sensitivity of the semicircular canals in great apes is therefore secondarily derived (Ryan *et al.*, 2012), indicative that slower locomotion may have characterized the great apes. Overall, the radiation of the great apes involved the evolution of larger species with slower, more deliberate locomotion compared to other anthropoid primates in which leaping and swinging are more commonplace (e.g. Ashton and Oxnard, 1964; Cartmill and Milton, 1977; Gebo, 1996; Henke and Tattersall, 2015; Ryan *et al.*, 2012; Temerin and John, 1983). Even some of the earliest hominoid species possess similar semicircular canal morphology to that of modern great apes (Rook *et al.*, 2004).

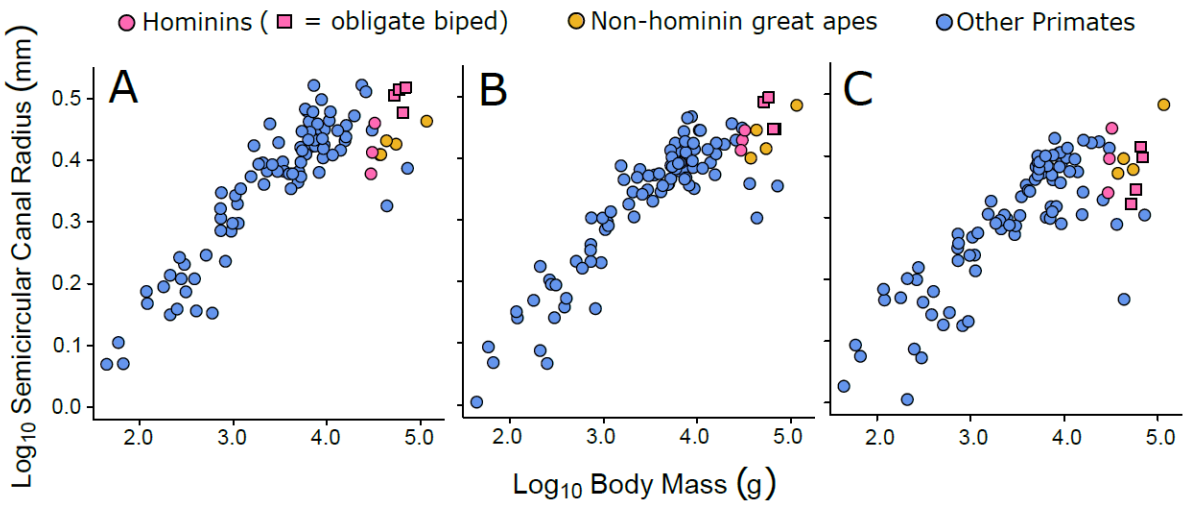


Figure 1: The radius of the semicircular canal plotted against body size for the (A) anterior (B) posterior and (C) lateral canals. Squares represent unequivocally obligate bipeds according to recent consensus (Harcourt-Smith, 2015; Wood, 2002).

For example, on the bases of semicircular canals and other anatomy the early great ape *Oreopithecus* is thought to have been a slow climber (Gebo, 1996; Rook *et al.*, 2004; Ryan *et al.*, 2012; Sarmiento, 1983; Szalay and Langdon, 1986). There is even some evidence that *Oreopithecus* possessed some postcranial adaptations indicative of bipedalism (Köhler and Moyà-Solà, 1997; Moyà-Solà *et al.*, 1999; Rook *et al.*, 1999) although this idea remains controversial (Harcourt-Smith, 2015; Harrison, 1991; Russo and Shapiro, 2013).

Moving through the evolutionary family tree to the lineage comprising our own bipedal species, Spoor *et al.* (1994) presented novel semicircular canal data from 4 extinct fossil hominin species (*Homo erectus, H. habilis, Australopithecus africanus,* and *Paranthropus robustus*) and formulated an original hypothesis linking hominin bipedalism and the relative dimensions of the semicircular ear canals. In particular, Spoor *et al.* (1994) argued that, after a reduction in size within great apes, some hominins subsequently evolved relatively enlarged vertically oriented anterior and posterior semicircular canals and a reduced horizontally positioned lateral canal. These evolutionary changes were observed within *H. erectus* and modern humans (*H.* sapiens) but not earlier hominins (Spoor *et al.*, 1994).

Postcranial evidence generally suggests that although early hominins, including *A. africanus* and *P. robustus,* possessed the ability to walk on two legs they did not do so all the time as they still possessed arboreal adaptations and were unlikely to be able to run or jump (Harcourt-Smith, 2015; Harcourt-Smith and Aiello, 2004; Kappelman *et al.*, 2016; Lovejoy *et al.*, 2009; Spoor *et al.*, 1994; Spoor and Zonneveld, 1998; Stern Jr and Susman, 1983; Susman *et al.*, 1984; Ward, 2002; White *et al.*, 2014). The observed changes in semicircular canal morphology in *H. sapiens* and *H. erectus* were interpreted as evidence supporting the idea that the earliest unequivocal obligate biped – i.e. the earliest species to be fully adapted and committed to walking on two legs – was *H. erectus* (Spoor *et al.*, 1994, 1996). An increase in the size of the anterior and posterior semicircular canals has also been posited to represent adaptation to efficient long-distance endurance running, rather than obligate bipedalism *per se* (Bramble and Lieberman, 2004; Lieberman and Bramble, 2007).

Although Spoor *et al.* (1996) explicitly state that their hypothesis linking semicircular canal size and hominin bipedalism "*refers to the canal sizes standardized against body mass, and not … relative to each other*", there is some ambiguity regarding the importance of the size of an individual semicircular canal size with reference to the others. Within mammals the lateral semicircular canal tends to be the smallest in size and the anterior is generally largest (e.g. Ekdale and Racicot, 2015; Hullar and Williams, 2006; Ramprashad *et al.*, 1984; Spoor *et al.*, 2002; Spoor *et al.*, 2007; Spoor *et al.*, 1994). However, this is not true for all species even within primates – e.g. in the capuchin monkey (*Cebus* sp.), the posterior canal is the smallest of the three (Ramprashad *et al.*, 1984) and the anterior canal is the smallest in the western gorilla (*Gorilla gorilla*) (Spoor *et al.*, 1994). The original comparisons in hominins imply differential shifts in size in each of the canals in association with bipedal locomotion (Spoor *et al.*, 1994) and that each of the semicircular canals may have the potential to evolve independently of one another. However, it is unclear whether these independent changes represent adaptive evolution: Perhaps increasing the size of one of the canals more or less than the other has some functional and selective importance

but this has not so far been considered. It therefore remains to be seen whether changes in the relative dimensions of the semicircular canals represent meaningful differences that might also be linked to the evolution of bipedalism in hominins.

Graf and Vidal (1996) reject the evidence for any link between semicircular canal morphology and hominin bipedalism, arguing that neither modern humans nor birds (which are also bipedal) fall outside the range of expected variation for non-bipedal animals and that great apes do not differ from other primates or even mammals on the whole. However, to make any such inferences meaningful, it is necessary to account for the shared ancestry implied by phylogeny (Felsenstein, 1985; Harvey and Pagel, 1991). Changes in the semicircular canal of a primate cannot be considered comparable to changes observed in the semicircular canals in any other distantly or even closely related group, including birds – without explicitly accounting for phylogeny. Hitherto, all analyses seeking to elucidate the role of the vestibular system in hominin bipedalism and locomotion have been conducted in a non-phylogenetic context (Graf and Vidal, 1996; Spoor *et al.*, 1994, 1996); what such analyses can tell us in terms of how changes in semicircular canals gave rise to the unique anatomy we see today in modern humans and our extinct close relatives is therefore limited.

For example, although living and great apes may look smaller upon visual inspection – particularly in the case of the anterior canal (Figure 1) – this may not necessarily represent anything special in terms of their evolution (for example, see whales in Chapter 2). We can use phylogenetic comparative methods to explicitly test whether differences in the great apes have arisen simply as a product of evolution flowing throughout the branches of the primate phylogenetic tree, or whether there have been shifts in morphology that represent unexpected and substantial bursts of morphological change.

All morphology is shaped by natural selection (Simpson, 1953). The strength of this natural selection is reflected in the rate of morphological evolution (Baker *et al.*, 2016; Eastman *et al.*, 2011; Kratsch and McHardy, 2014; Kutsukake and Innan, 2013, 2014; Rabosky, 2014;

Revell *et al.*, 2012; Thomas and Freckleton, 2012; Venditti *et al.*, 2011) such that where there have been rapid evolutionary change it is an indication that there has been strong selection pressure acting on species morphology. Where there have been major shifts in species ecology or lifestyle, it is possible to identify intense rapid bursts of evolutionary change that represent instances of historical positive phenotypic selection (Baker *et al.*, 2016). We can use these methods – whilst simultaneously testing for differences among great apes – to measure the rate of morphological adaptive change in the size of the semicircular canals in primates. If the evolutionary circumstances leading to the origin of obligate bipedalism exerted strong selection pressure for the sizes of the semicircular ear canals to change – as implied by (Spoor *et al.*, 1994), then it should be possible to detect signal for this among the branches of the phylogenetic tree in the form of intense bursts of historical *positive phenotypic selection*.

Although other analyses have considered average canal sizes (Baker *et al.*, 2016; Graf and Vidal, 1996; Spoor *et al.*, 2002), Spoor (1996) argues against this in the context of bipedalism. We analyse each of the canals separately in order to determine whether we can detect nuanced changes in the canals that would be otherwise impossible to detect were they to be studied as a single unit. We seek to detect increases in the size of the vertically-oriented anterior and posterior semicircular canals and a decrease in size of the lateral canal arising as a consequence of strong historical positive phenotypic selection in hominins. We also consider the importance of the relative sizes of each of the semicircular canals in reference to the others. Any changes in size that we detect in these relative sizes can be attributed to independent shifts in the individual canal rather than changes to the system as a whole or to concerted changes among the canals.

Cutting-edge phylogenetic statistical methodology in combination with semicircular canal data for other extinct hominins including *H. neanderthalensis* and *H. heidelbergensis* (Quam *et al.*, 2016; Spoor *et al.*, 2003) which are also obligate bipeds (Harcourt-Smith, 2015) will allow us to reveal the role of the semicircular canal system during hominin

evolution and identify how natural selection has acted in response to the evolution of obligate bipedalism in hominins and ultimately the lineage leading to our own species.

Methods

Data and tree

We obtained body sizes and measurements of the anterior, posterior and lateral semicircular canals (the arc radius of curvature, see Spoor *et al.,* 1994) for living and recently extinct primates including *H. sapiens* (n = 91) from the dataset of Spoor *et al.* (2007). We then added data for 7 additional extinct hominins from the literature: *H. neanderthalensis, H. heidelbergensis, H. erectus, H. habilis, Australopithecus africanus,* and *Paranthropus robustus* (Holloway *et al.*, 2004; Jungers *et al.*, 2016; Quam *et al.*, 2016; Spoor *et al.*, 1994). All variables were $\log_{10}$-transformed prior to analysis.

We used the primate phylogeny from Spoor *et al.* (2007) which includes all non-hominin primates with semicircular canal measurements – and *H. sapiens*. For the extinct hominins, we grafted the maximum clade credibility phylogeny from the most recent comprehensive Bayesian phylogenetic analysis (Dembo *et al.*, 2015) onto the primate phylogeny. We did not use a more recently published hominin phylogeny (Dembo *et al.*, 2016) as the tree of Dembo *et al.* (2015) was based on a more comprehensively sampled morphological dataset; thus we considered it to be more robust. Because the dates of hominin species were reconstructed using the first appearances in the fossil record in the analysis of Dembo *et al.* (2015), we extended all tip branches to the last known occurrence as reported by (Wood and Lonergan, 2008).

Phylogenetic analyses

We used the *variable rates regression model* (Baker *et al.*, 2016) to detect variation in the rates of evolution within the phylogenetically structured residual variance of a regression. This allowed us to detect shifts in the rate of morphological evolution in one trait (in this

case, semicircular canal radius) whilst simultaneously estimating and accounting for its relationship with one or more other covariates. The variable rates model estimates two components: firstly, a background rate of morphological evolutionary change (σ^2_b) and secondly a set of rate scalars (r) that define how much the rate of evolution along each branch deviates from that explained by the background rate – where $r > 1$, a branch has evolved faster than the background rate and where $r < 1$ a branch has evolved slower. From these two components it is possible to calculate the *optimized rate* of evolution along each branch, ($\sigma^2_v = \sigma^2_b r$).

These two rates of evolution (σ^2_b and σ^2_v) make it possible to define the amount of evolutionary change that is actually occurring along a branch ($\Delta_V = \sigma^2_v t$, where t = the length of the branch in time) vs. that which would be expected given the background rate of evolution ($\Delta_B = \sigma^2_b t$). From this, we can compare the amount of evolutionary change given the background rate of evolution to that arising from shifts in the rate of evolutionary change by calculating the ratio Δ_V/Δ_B.

We identified positive phenotypic selection along branches where the phenotypic variance attributed to increases in the rate of morphological evolution was more than double that attributed to the background variance. The amount of phenotypic evolution attributed to shifts in the rate of evolution is calculated as $\Delta_V = \sigma^2_v t$. Positive phenotypic selection was identified where $\Delta_V/\Delta_B > 2$ in more than 95% of the posterior distribution (Baker *et al.*, 2016). These branches represent exceptional increases in the underlying rate of morphological evolution – where there is significant unexplained residual variance away from an underlying evolutionary relationship.

All models were run for 250 million iterations, removing the first 50 million as burn-in and using a sampling rate of 100,000. This resulted in a total number of 2,000 samples. Convergence was assessed visually and all chains were replicated multiple times to ensure model stability. Positive evidence for rate heterogeneity is accepted where we find a Bayes Factor (BF) greater than 2 (Raftery, 1996) calculated as $BF = -2\log_e[m_1/m_0]$, where m_0 is

the marginal likelihood of a simple Brownian motion model that estimates a single underlying rate of evolution and m_1 is the marginal likelihood of a model that allows the rate of evolution to vary among branches. All marginal likelihoods were estimated using stepping-stone sampling (Xie *et al.*, 2010) implemented in BayesTraits (Pagel *et al.*, 2004). We ran the stepping-stone sampler for 200 stones with a total number of 750,000 iterations per stone after burn-in - drawing values from a beta-distribution (α = 0.4, β = 1) (Xie *et al.*, 2010).

We tested to see if great apes differ in the dimensions of their semicircular canals by estimating a separate intercept for great apes in all models, allowing the magnitude of their semicircular canal radius to differ from other primates. For each of the three canals, we tested two models: one which examined the evolution of the semicircular canal relative to body size whilst allowing for a shift in size within the great apes (*relative canal size* model), and another more complex model that also accounted for the sizes of the other two semicircular canals to detect independent evolution in each of the canals (*independent canal size* model). We assessed the significance of each parameter by identifying the proportion of the distribution that crosses zero (P_x). Where P_x < 0.05 a parameter was considered to be substantially shifted from zero and was therefore significant. Where a parameter was non-significant we removed it from the model; in this way, our *supported model* for each of the three canals retained only significant parameters. We present only the supported models from each analysis.

Given the supported model for each of the three canals, we identified along which, if any, hominin branches there had been positive phenotypic selection. Owing to small sample sizes and short branch lengths, it can be difficult to detect rate shifts along hominin branches that correspond to directional change – we therefore tested to see if these episodes of positive phenotypic selection represented directional changes in the sizes of the semicircular canals by running additional models that allowed a further shift in intercept in these species.

Results

We present the results for each of the three semicircular canals separately. For each canal we present results firstly for the *relative canal size* model that estimates a relationship between semicircular canal radius and body size and secondly the *independent canal size* models that account for any relationships between the sizes of the semicircular canals themselves. All parameter estimates are tabulated in Appendix 2.

Anterior Semicircular Canal

In the supported model for the relative canal size (Table A.2.1) we find that great apes have a significantly reduced anterior semicircular canal radius for their body size (mean β = -0.153, P_x= 0.001, Figure 2A). There is significant rate heterogeneity (BF = 18.16) and within hominins there is positive phenotypic selection (Δ_V/Δ_B > 2 in more than 95% of the posterior distribution) identified along all branches within and leading to the genus *Homo* (modal Δ_V/Δ_B ranges between 8.56 and 9.34). When we estimate a different intercept for *Homo*, we no longer observe any rates that can be defined as positive phenotypic selection within any hominin branches. The anterior semicircular canal radius for *Homo* is significantly larger for its body size compared to great apes (mean difference = 0.070, P_x = 0.000): hominins exhibit a significant directional shift in the relative size of the anterior canal in the opposite direction (Figure 2A).

We also identify several episodes of positive phenotypic selection outside of hominins. In the supported model (one that did not estimate a separate intercept for *Homo*), we identify the branch leading to the blue monkey (*Cercopithecus mitis,* modal Δ_V/Δ_B = 6.62), the extinct sloth lemur *Palaeopropithecus ingens* (modal Δ_V/Δ_B = 9.50) and branches leading to and within the *Colobus* genus (n = 2 species, modal Δ_V/Δ_B ranges between 8.74 and 13.23). After estimating a separate intercept for *Homo*, there is still significant rate heterogeneity (BF = 16.08) and we identify the same branches as positive phenotypic selection in non-hominins.

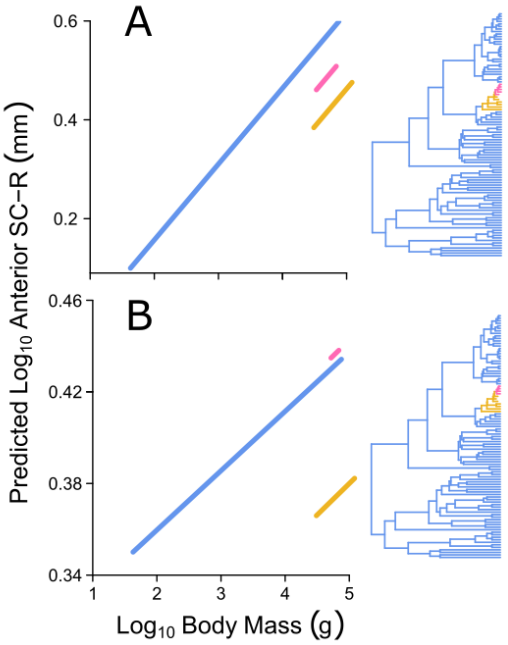


Figure 2: Results for the anterior semicircular canal radius (SC-R) showing the predicted phylogenetic regression lines for (A) the relative canal size model and (B) the independent canal size model. Colours indicate clades with significant differences in intercept: blue, non-hominoids; pink, hominin species where there has been a directional shift; yellow, all remaining great apes. A phylogeny is shown that indicates which species are included in each group. For the fitted line in (B) we hold the size of the posterior and lateral canal at their mean values.

In the independent canal size model, there is a significant reduction in the size of the anterior semicircular canal in great apes (mean β = -0.061, P_x = 0.014, Figure 2B, Table A.2.2). Amongst the significant rate heterogeneity (BF = 33.08), we identify positive phenotypic selection along all branches within the clade comprising *H. erectus, H. sapiens, H. neanderthalensis,* and *H. heidelbergensis* (*erectus-sapiens* clade, modal Δ_V/Δ_B ranges between 8.15 and 9.48). When we estimate a different intercept for the *erectus-sapiens* clade, (Figure 2B) we find that it is significantly different from great apes (mean difference = 0.06, P_x = 0.022). Although we find that variable rates still improves over a single rate model (BF = 28.44), we no longer observe positive phenotypic selection in any hominin branches.

In the supported independent canal size model, the only positive phenotypic selection identified outside of hominins is along the branch leading to *C. mitis* (modal Δ_V/Δ_B = 16.65); this is still the case after allowing for the directional shift in *erectus-sapiens*.

Posterior Semicircular Canal

We find a significant reduction in relative canal size for the posterior canal in great apes (mean β = -0.073, P_x = 0.015, Figure 3A, Table A.2.3). Although we find positive evidence for rate heterogeneity (BF = 14.99), we find no positive phenotypic selection among any hominin branches. The only branches that are scaled to have a Δ_V/Δ_B > 2 in more than 95% of the posterior distribution are those within the *Colobus* genus (modal Δ_V/Δ_B ranges from 10.78 to 11.73).

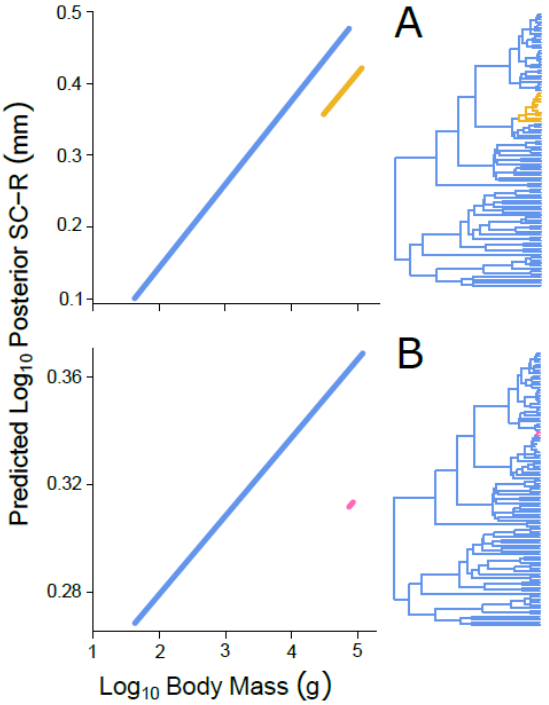


Figure 3: Results for the posterior semicircular canal radius (SC-R) showing the predicted phylogenetic regression lines for (A) the relative canal size model and (B) the independent canal size model. Colours indicate clades with significant intercept differences: blue branches and lines indicate non-hominoid primates; pink branches and line indicate *H. neanderthalensis* and *H. heidelbergensis* (and the branch leading to them in the tree); yellow branches and line represent all great apes. For (B) we predict while holding the anterior and lateral semicircular canal radii at their means.

Our supported independent canal size model (Table A.2.4) estimated only a single intercept across all primates (Figure 3B). Amongst significant rate heterogeneity (BF = 29.04), three branches within hominins were undergoing positive phenotypic selection – those leading to and including the species *H. neanderthalensis* and *H. heidelbergensis* (henceforth the *nean-heid* clade, modal Δ_V/Δ_B ranges between 10.53 and 11.85).

When we allow the intercept of the relationship to vary in the *nean-heid* clade, a variable rates model still improves over a single rate model (BF = 21.37) but we no longer observe positive phenotypic selection in any hominin branches. The intercept of the two hominin species is significantly smaller than that of other primates (Table A.2.4, mean β = -0.051, P_x = 0.009) – they have a significant independent reduction in their posterior semicircular canal radius (Figure 3B). All other hominins and great apes fall within the expected range of variation of the primate clade as a whole (Figure 3B).

Outside of the hominins we observe selection along all branches within the genus *Macaca* (n = 14 branches leading to n = 8 species; modal Δ_V/Δ_B ranges between 5.68 and 6.52) and *C.mitis* (modal Δ_V/Δ_B = 6.23). In the model where we allow a difference in intercept for the *nean-heid* clade these results are qualitatively identical; the same branches are identified as positive phenotypic selection.

Lateral Semicircular Canal

The lateral semicircular canal shows a significant reduction in size within great apes in the relative canal size model (Figure 4, Table A.2.5). There is significant rate variation (BF = 22.78), and we identify that all hominin branches to the exclusion of *Paranthropus* (the clade comprising *Australopithecus* and *Homo;* henceforth *Australopithecus-Homo*) have an estimated $\Delta_V/\Delta_B > 2$ in more than 95% of their posterior distribution (modal Δ_V/Δ_B ranges between 10.16 in *A. africanus* up to 11.84 in *H. heidelbergensis*). The model with a different intercept for *Australopithecus-Homo* still has significant rate variation (BF = 27.85) and all the same hominin branches are identified as positive selection (Figure 4, modal Δ_V/Δ_B ranges between 10.17 and 11.84). Although *Australopithecus-Homo* has a significantly smaller lateral semicircular canal radius for their body size compared to other primates (Table A.2.5, mean difference = -0.110, P_x = 0.019), they do not significantly differ from great apes (mean difference = -0.038, P_x = 0.141).

Outside of hominins, we identify positive phenotypic selection in the branch leading to the gelada baboon *Theropithecus gelada* (modal Δ_V/Δ_B = 4.863) and in the extinct sloth

lemur *Palaeopropithecus ingens* (modal Δ_V/Δ_B = 8.667); these results remain unchanged when we allow a different intercept for the *Australopithecus-Homo* clade.

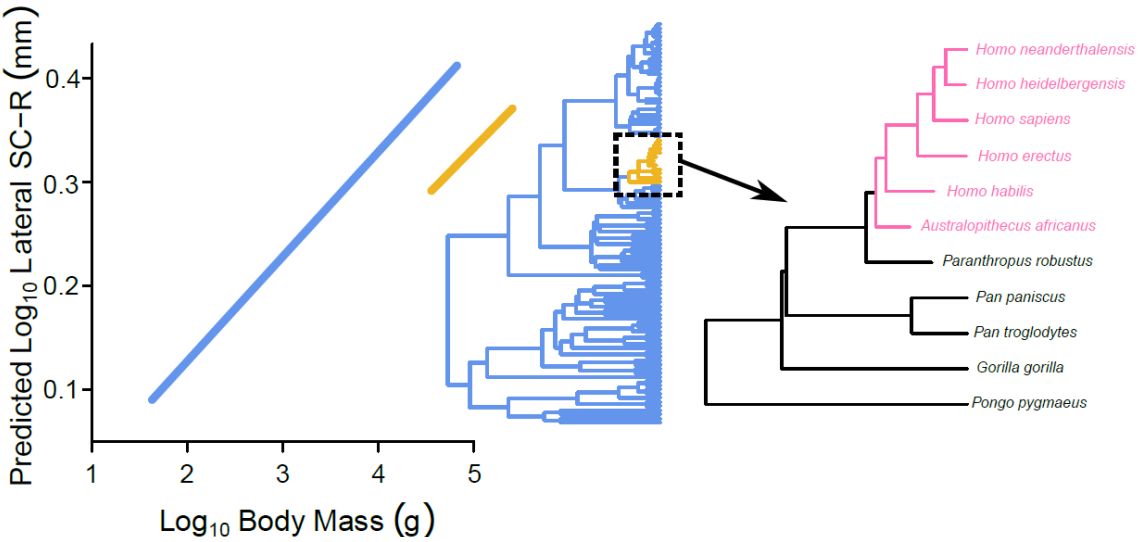


Figure 4: The results of the relative canal size model for the lateral semicircular canal radius (SC-R). There is a significant difference in the size of the lateral SC-R between great apes (yellow) and other primates (blue) but no differences observed within in hominins according to our criteria. However, we observe positive selection along the branches highlighted pink in the zoomed-in great ape portion of the tree.

As with the posterior semicircular canal, there is no significant difference in the lateral semicircular canal size for great ape species compared to other primates when we consider the sizes of the other canals in our independent canal size model (Table A.2.6). There is also no significant link with body size (Table A.2.6). Here, a variable rates model is still preferred over a model estimating just a single rate (BF = 50.51). We identify positive phenotypic selection in all branches within *Australopithecus-Homo* (modal Δ_V/Δ_B ranges between 17.93 and 22.99).

When we allow the intercept of the relationship to vary in *Australopithecus-Homo,* a variable rates model still improves over a single rate model (BF = 48.660). We still observe positive phenotypic selection in all branches within *Homo* (modal Δ_V/Δ_B ranges between 19.00 and 22.87). The branch leading to *A. africanus* and *Homo* is no longer identified; it has a modal Δ_V/Δ_B = 2.68 and is scaled to be >2 in just 67% of the posterior. Similarly, we observe a modal Δ_V/Δ_B = 3.36 and 67% >2 in the *A. africanus* branch. Most importantly,

we find that there is no significant difference in the size of primate species and those hominins that have evolved the sizes of their semicircular canals as a result of positive phenotypic selection (Table A2.6); there is no simple directional shift in the size of the lateral canal within *Australopithecus-Homo*. Instead, we observe an increase in variance – most branches within this clade remain scaled.

Outside of the hominins we observe selection in only two branches – those leading to the species *Macaca nemestrina* (modal Δ_V/Δ_B = 7.76) and *Theropithecus gelada* (modal Δ_V/Δ_B = 5.34). Positive selection in these branches remains qualitatively unchanged when we estimate a separate intercept for *Australopithecus-Homo*.

Discussion

Our results represent the first comprehensive phylogenetic analysis of ear canals in the context of hominin evolution and bipedalism. Previous analyses attempting to link vestibular morphology to hominin locomotion have all previously been non-phylogenetic (Graf and Vidal, 1996; Spoor *et al.*, 1994, 1996) and have not explicitly considered the effect of independent changes across the semicircular canals despite discussing the importance of relative sizes (Spoor *et al.*, 1994, 1996).

During primate evolution, we reveal that intense selection pressure has acted to change the sensitivity of the semicircular canals by increasing and decreasing their overall radius relative to body size. It is perhaps not surprising then that we also observe intense changes in the size of each canal compared to the others (and after accounting for body size). Natural selection has acted not only to sculpt the overall size of the semicircular canal but also to independently change individual semicircular canals relative to one another. Although increasing the size of the semicircular canal relative to body size gives a generalized increase in sensitivity (Ten Kate *et al.*, 1970) independent changes within individual canals implies differential affinity to movement within particular planes (Ekdale,

2016; Graf and Vidal, 1996) – although the exact functional significance of this remains poorly understood (Ekdale, 2016).

In great apes, we find that the relative size of all three semicircular canals are smaller than in other primates (Figures 2A, 3A, 4A). However, we find that only the anterior semicircular canal independently differs within great apes (Table A.2.2, Figure 2A). In both the posterior and lateral semicircular canals, great apes fall within the expected range of morphological diversity implied given the rate of evolution in primates as a whole (Figure 2B, 2C). This means that the anterior semicircular canal has decreased in size in great apes by such an amount that it goes beyond the amount of change observed in the other canals. This is reflected by the larger intercept difference for great apes in the supported relative anterior canal size model (Table A.2.1, Figure 2A) which equates to roughly an 29.7% reduction in size (calculated as the difference between the predicted anterior semicircular canal radius for an ape of average body size [50.12 kg] vs. a non- great ape primate of the same size). On the other hand, changes in the relative size for both the posterior canal (Figure 3A) and the lateral canal (Figure 4A) are smaller in great apes compared to primates (there is a 15.6% decrease in both; Table A.2.3, Table A.2.5).

The original reports that a marked increase in the size of the two vertically-oriented canals (Spoor *et al.*, 1994, 1996) are cited frequently as evidence for morphological evolutionary change linked to the development of bipedal locomotion and adaptation for endurance running within hominins (e.g. Bramble and Lieberman, 2004; Lieberman *et al.*, 2009; Raichlen and Polk, 2013). However, we find only limited evidence to support this hypothesis. Instead, we observe an overall increase in sensitivity in the anterior semicircular canal within all species of the genus *Homo* (Figure 2A). Unlike earlier comparisons among hominin species (Spoor *et al.*, 1994, 1996), we find that this increase is also observed in *H. habilis* – a species that remains contentious with regards to its primary mode of locomotion (Harcourt-Smith, 2015; Harcourt-Smith and Aiello, 2004; Ruff, 2009, see later). However, we find that it is only in species belonging to the lineage including *H. erectus*, *H. heidelbergensis*, *H. neanderthalensis* and *H. sapiens* (the *erectus*-

sapiens clade) that there has been an increase in the anterior semicircular canal independent of all other changes in the posterior and lateral canals (Figure 2B) – to the exclusion of *H. habilis*. We no longer see a significant increase in the size of the anterior canal in *H. habilis* when we account for the sizes of the other canals owing to an increase in the lateral canal along the same lineage (Figure 5, see later).

It is only when we consider both changes in canal size as compared to body size *and* the relative contribution of independent evolution in each of the canals individually that we observe a directional change in hominin species that coincides with the presumed origin for obligate bipedalism (Harcourt-Smith, 2015; Harcourt-Smith and Aiello, 2004; Wood, 2002) and endurance running (Bramble and Lieberman, 2004). Our results therefore highlight that individual shifts in each of the semicircular canals that alter the relative sizes within the entire system (i.e. which canal is largest, and by how much) is a previously unconsidered potential marker for bipedal locomotion in hominins. Increasing or decreasing the size of one semicircular canal beyond that of the others provides a viable substrate upon which natural selection can act; understanding exactly how and why this can happen is worth future investigation.

Unlike the anterior semicircular canal, we observe no differences in the relative size of the posterior canal between hominins and other great apes after accounting for body size (Figure 3A). This therefore puts further doubt on the claim that specific increases in the size of both this and the other vertically-oriented anterior canal are linked with the onset of obligate bipedalism in hominins (Spoor *et al.*, 1994, 1996; Spoor and Zonneveld, 1998). What we do find, however, is that there has been natural selection in *H. neanderthalensis* and *H. heidelbergensis* (Figure 3B, Table A.2.4) to evolve a smaller posterior semicircular canal in comparison to both the anterior and the lateral canals. This actually came about as a consequence of not directional change towards a smaller posterior canal in these species but instead via natural selection driving both the anterior and the lateral canal to increase in size (Figure 2A and Figure 5), ultimately resulting in a significantly smaller posterior canal in comparison to the others.

Neanderthals have previously been recognized to have very distinct labyrinthine morphology to modern humans including a relatively smaller posterior canal (Hublin *et al.*, 1996; Spoor *et al.*, 2003). On the bases of this and other postcranial features, Neanderthals have been interpreted to have been inefficient endurance runners (Raichlen *et al.*, 2011; Spoor *et al.*, 2003; Steudel-Numbers and Tilkens, 2004; Weaver and Steudel-Numbers, 2005), and that in general their locomotion was energetically expensive (Steudel-Numbers and Tilkens, 2004; Weaver and Steudel-Numbers, 2005, but see Higgins and Ruff, 2011) . Much less has been said about the locomotor ability of *H. heidelbergensis* although this species may have similar locomotion to Neanderthals given their close phylogenetic affinity (Buck and Stringer, 2014; Dembo *et al.*, 2015; Dembo *et al.*, 2016; Mounier and Caparrós, 2015; Pablos *et al.*, 2014; Stringer, 2012) – some authors even suggest that they are not distinct species (see Wood, 2011). However, we do find that natural selection has driven differences between these two species in terms of their lateral canal (see later, Table A.3.1). In any case, our results show that both species evolved their unique posterior canal size as a consequence of strong historical positive phenotypic selection rather than simply any lack of adaptation e.g. to endurance running. Changes in adaptive circumstances within these species therefore exerted intense selective pressures on the semicircular canals, allowing a new and unique form to arise from an ancestor more like modern humans or *H. erectus* in terms of their ear canal morphology.

Unlike the anterior and posterior canals, we identify positive phenotypic selection in the size of the lateral semicircular canal that cannot be ascribed to any simple directional change in size (Figure 4, Table A.2.5, Table A.2.6). Instead, natural selection has acted in *Australopithecus-Homo* to increase variation (Table A.2.5, Figure 4). An increase in morphological variation arising as a consequence of strong historical positive phenotypic selection implies ecological or environmental variability. For example – we demonstrate in Chapter 5 that natural selection acting to increase eye shape variation in mammalian carnivores can be almost entirely explained by activity pattern.

This therefore begs the question as to why exactly we observe such a variance increase in the size of the lateral canal in the first place, especially considering that it has arisen as a consequence of strong historical natural selection – and whether we can identify any evolutionary differences in the directionality of the change. We conducted an additional exploratory post-hoc analysis where we re-ran the supported relative canal size model that models the evolution of the lateral semicircular canal while accounting for its relationship with body size whilst also allowing the intercept to differ in great apes. However, in this model we additionally estimated a separate intercept for each of the six species within *Australopithecus-Homo* (Appendix 3). Assessing the posterior distribution of each of the species-intercepts can tell us the direction of change in each species. In some ways this is similar to the phylogenetic outlier test proposed by (Organ *et al.*, 2011) which seeks to identify whether an individual species falls outside the range of variation expected given some underlying model and a phylogenetic tree. Here, instead of identifying individual species as outliers, we can use these species-intercept differences to identify significant differences between species: where these occur within a single lineage, it will shed light on the nuances of how the variance increase in the size of the lateral canal in hominins came about - and will inform us of any potential directionality of the adaptive change. We compare the differences between each species in Table A.3.1. and broadly characterize the timing of directional shifts in the size of the lateral semicircular canal within hominins by tracing the evolutionary path along the hominin portion of the phylogeny (Figure 5).

We find that *A. africanus* was the first hominin species in which there was a substantial reduction in the size of the lateral semicircular canal that arose as a consequence of strong historical positive phenotypic selection. This reduction occurred sometime after the divergence of this species and other great apes (Figure 5). Moving along the branches of the phylogeny we come across the unusually large *H. habilis* (in terms of relative lateral semicircular canal size) in which natural selection drove an evolutionary increase in size (Figure 5). It is worth noting here that *H. habilis* can be considered as the equivalent of a

phylogenetic outlier in this analysis as it has a significantly larger intercept (and thus a larger lateral semicircular canal size) than all other species, including other great apes (Table A.3.1). Our comparisons imply that a further reduction in size occurred along the branch leading to the *erectus-sapiens* clade (Figure 5), finally followed by an adaptive shift towards a secondarily larger semicircular canal in the lineage leading to *H. neanderthalensis* and *H. heidelbergensis* (Figure 5).

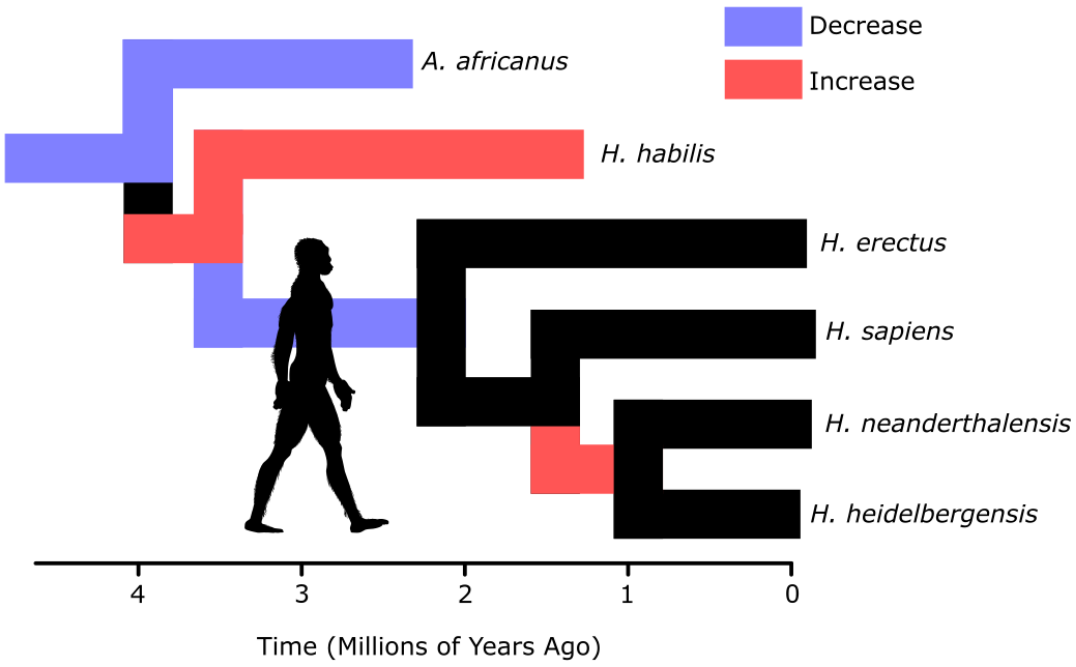


Figure 5: A graphical representation of how natural selection sculpted the size of the lateral canal during hominin evolution. Blue represents a decrease in relative size in response to natural selection; red represents an increase. Black branches show where there has been no significant shifts arising from positive phenotypic selection. The silhouette is taken from phylopic.org and shows African *H. erectus* - the earliest unequivocal obligate bipedal hominin. The justification for shifts are as follows using values from Table A.3.1: *A. africanus* is significantly smaller than other great apes; an initial decrease must therefore have occurred along this branch or along the branch leading to the more inclusive clade of *Australopithecus-Homo*. *H. habilis* is significantly larger than *A. africanus* and so we infer an evolutionary shift towards larger size somewhere along the lineage leading to this species. *H. erectus* is significantly smaller than both *A. africanus* and *H. habilis* so we know that there must have been a further decrease in the lineage leading to this species; because *H. sapiens* does not significantly differ from *H. erectus* in size the most likely timing of this is in the lineage leading to the erectus-sapiens clade. Finally, an adaptive shift towards a secondarily larger semicircular canal occurred in the lineage leading to *H. neanderthalensis* and *H. heidelbergensis* – these species both differ from *H. sapiens*, but not from one another.

When this process was repeated for the independent canal size model, comparing hominins to all other primates as a whole (owing to the lack of difference between great apes and other primates in this model, Table A.2.6) there is no qualitative difference in our conclusion – with one exception: *Homo neanderthalensis* has a significantly larger lateral semicircular canal compared to *H. heidelbergensis* when the sizes of the other canals are taken into consideration (Table A.3.1). This implies a further burst of positive phenotypic selection among the branches within that lineage acting to independently alter the lateral canal size although it is impossible to determine precisely along which branches this occurred (i.e. whether *H. neanderthalensis* increased in size or *H. heidelbergensis* decreased – or both).

In support of Spoor's original suggestion that there may have been a decrease in size of the lateral semicircular canal in association with the evolution of obligate bipedalism in hominins (Spoor *et al.*, 1994, 1996; Spoor and Zonneveld, 1998), we identify a significant adaptive shift towards smaller size that most likely occurred in the lineage leading to *H. erectus* onwards (Figure 5). However, there was also an earlier decrease in size before the proposed origin of obligate bipedalism. Of the three canals, the lateral is most closely linked with agility (Walker *et al.*, 2008). Multiple adaptive shifts in the size of the lateral canals implies that hominins might have been experimenting with different forms of locomotion throughout their evolution involving different forms of agile movement. For example, the unique changes observed as a consequence of natural selection in *H. habilis* may reflect the stronger dependence on arboreality argued to have been present in this species compared to other members of *Homo* (Haeusler and McHenry, 2004; Harcourt-Smith, 2015; Harcourt-Smith and Aiello, 2004; Ruff, 2009). However, the importance of the independent increase in the size of the lateral canal in this scenario is yet to be elucidated: nobody would question that *H. neanderthalensis* and *H. heidelbergensis* were both obligate bipeds (Harcourt-Smith, 2015; Harcourt-Smith and Aiello, 2004; Wood,

2002) – but both of these species also have significantly larger lateral semicircular canals compared to *H. erectus* and *H. sapiens* (Table A.3.1).

Taken together, the selective pressures imposed on *H. habilis* drove this species to evolve semicircular canal morphology unlike any other hominin or primate. We observe positive phenotypic selection to drive both the anterior and lateral canals larger in this species (for the anterior canal, this increase occurs along with other members of the genus *Homo*). Compared to the rest of great apes, this change equates to an approximately 18% increase in the size of the anterior canal. Further natural selection independently acted to drive the lateral canal to increase even further (Figure 5) – assuming that the ancestral form of *H. habilis* was more like *Australopithecus* than *Paranthropus* and other great apes (Figure 5, Appendix 3), then this difference equates to an approximately 26% increase. The particular specimen where the semicircular canal data has been measured for *H. habilis* – Stw 53g (Spoor *et al.*, 1994) – has previously been identified to be unusual in its relative canal dimensions (Spoor *et al.*, 1994) and although current convention ascribes this specimen to *H. habilis* (Wood, 2011) it has previously been the subject of taxonomic debate (Curnoe and Tobias, 2006; Grine, 1993; Grine *et al.*, 1996; Kuman and Clarke, 2000; Wood, 2011). There are many similarities between *H. habilis* and other australopithecines in both locomotor adaptations and other morphological features (Haeusler and McHenry, 2004; Harcourt-Smith, 2015; Harcourt-Smith and Aiello, 2004; Kuman and Clarke, 2000; Ruff, 2009; Wood and Collard, 1999), even prompting some authors to even argue that the species should be considered as a member of the genus *Australopithecus* (Kuman and Clarke, 2000; Wood and Collard, 1999). Our results show that the morphology of *H. habilis* has been sculpted by historical natural selection in distinct ways not only from *Australopithecus* but also all other *Homo* – at least in terms of its semicircular canal morphology.

Though evolutionary circumstances leading to the origin of obligate bipedalism and endurance running in hominins certainly imposed significant selection pressures on hominins, the signal for this in the sizes of the semicircular canals is not as clear as previous

analyses might suggest. We find no evidence to support the suggestion that the posterior canal increases in size during the evolution of obligate bipedalism in hominins (Spoor *et al.*, 1994, 1996). Instead, positive selection has acted to increase the size of the anterior canal beyond both the lateral and posterior canal in the lineage leading to obligate bipeds; this indicates an as of yet undetermined functional role of independent canal changes during the evolution of bipedalism in hominins. Graf and Vidal (1996) also noted that it is only the anterior canal that shows any apparent difference in the radius of curvature among quadrupeds and bipeds although did not explicitly test this. In other groups, this canal has shown to be linked to bipedalism: bipedal dinosaur species tend to possess relatively larger anterior canals compared to quadrupeds (Georgi *et al.*, 2013; Sipla, 2007). Differential decreases in lateral canal size in all hominins excluding *Paranthropus* and *H. habilis* calls into question any simple link between the horizontal canal dimensions and obligate bipedalism but does have important implications for locomotion in these species.

Throughout hominin evolution, we find little evidence for coordinated or concerted changes across the system as a whole. The use of average semicircular canal sizes as an indicator of overall semicircular canal system sensitivity or agility e.g. (Baker *et al.*, 2016; Graf and Vidal, 1996; Spoor *et al.*, 2002; Spoor *et al.*, 2007) may therefore not be appropriate. Although semicircular canal size on average may be linked with agility e.g. (Spoor *et al.*, 2007), the relationship between agility and canal sizes differs within each of the individual canals (Spoor *et al.*, 2007). Previously, we identified positive phenotypic selection acting to sculpt average semicircular canal radi across mammals including primates (Baker, Meade et al. 2015). In this earlier analysis we identified that the extinct sloth lemur *Palaeopropithecus* experienced intense selection pressure to evolve a smaller average semicircular canal size (Baker *et al.*, 2016). The present results demonstrate that only the anterior and lateral canals in this species have experienced a reduction in size as a result of strong historical positive selection – and only the anterior canal is smaller in comparison to the others. Such differences among the canals almost certainly result as a consequence of natural selection independently sculpting the sizes of individual canals

depending on different ecologies and environments. These independent changes in semicircular canal sizes – i.e. differences among species regarding which is the largest or smallest of the canals demonstrate that studying each of the canals separately whilst allowing for associations with the others may be more appropriate than using any single measure of the overall size of the semicircular canal system.

The evolution of the hominin semicircular canal has been characterized by a mosaic of intense bursts of historical positive phenotypic selection. Adaptive changes have occurred in different directions and along different lineages in each of the three canals (Figures 2-4). Changes among different morphologies and behaviours have often been suggested to have occurred in a mosaic fashion during hominin and hominoid evolution (e.g. Ackermann and Smith, 2007; Alba *et al.*, 2012; Delezene, 2015; Holloway and Post, 1982; Kivell *et al.*, 2011; Manthi *et al.*, 2012; Mitteroecker and Bookstein, 2008; Rae, 1999). This has since been demonstrated to characterize the evolution of the mammalian and avian brain in general where different structures and regions have the capacity to change independently of one another (Barton and Harvey, 2000; Iwaniuk and Hurd, 2005).

Although there is some evidence that some of the morphology we consider to be unique to modern humans could have been achieved by passive processes (Barton and Venditti, 2013; Benazzi *et al.*, 2015; Schroeder *et al.*, 2014), the prominent current view holds that the anatomical features and innovations we consider to be human – including bipedalism and associated morphologies – arose as a result of adaptive evolution and strong natural selection (Antón *et al.*, 2014; Baker *et al.*, 2016; Barton and Venditti, 2014; Hublin, 2015; Organ *et al.*, 2011; Pagel, 2002; Pampush, 2015; Schroeder *et al.*, 2014). Our results fit with this adaptive view, demonstrating that hominin evolution was a time of evolutionary innovation and experimentation, where intense natural selection sculpted the diversity observed within extinct hominins, ultimately giving rise to our own species.

References

Ackermann RR & Smith RJ 2007. The macroevolution of our ancient lineage: what we know (or think we know) about early hominin diversity. *Evolutionary Biology,* 34 (1): 72-85.

Alba DM, Almecija S, Casanovas-Vilar I, Mendez JM & Moya-Sola S 2012. A partial skeleton of the fossil great ape *Hispanopithecus laietanus* from Can Feu and the mosaic evolution of crown-Hominoid positional behaviours. *PLoS ONE,* 7 (6): e39617.

Antón SC, Potts R & Aiello LC 2014. Evolution of early *Homo*: An integrated biological perspective. *Science,* 345 (6192).

Ashton E & Oxnard CE 1964. Locomotor patterns in primates. *Journal of Zoology,* 142 (1): 1-28.

Baker J, Meade A, Pagel M & Venditti C 2016. Positive phenotypic selection inferred from phylogenies. *Biological Journal of the Linnean Society,* 118: 95-115.

Barton RA & Harvey PH 2000. Mosaic evolution of brain structure in mammals. *Nature,* 405 (6790): 1055-1058.

Barton RA & Venditti C 2013. Human frontal lobes are not relatively large. *Proceedings of the National Academy of Sciences USA,* 110 (22): 9001-9006.

Barton RA & Venditti C 2014. Rapid evolution of the cerebellum in humans and other great apes. *Current Biology,* 24 (20): 2440-2444.

Benazzi S, Nguyen HN, Kullmer O & Hublin J-J 2015. Exploring the biomechanics of taurodontism. *Journal of Anatomy,* 226 (2): 180-188.

Billet G, Hautier L, Asher RJ, Schwarz C, Crumpton N, Martin T & Ruf I 2012. High morphological variation of vestibular system accompanies slow and infrequent locomotion in three-toed sloths. *Proceedings of the Royal Society B: Biological Sciences,* 279 (1744): 3932-3939.

Bramble DM & Lieberman DE 2004. Endurance running and the evolution of *Homo*. *Nature,* 432 (7015): 345-352.

Buck LT & Stringer CB 2014. *Homo heidelbergensis*. *Current Biology,* 24 (6): R214-R215.

Cartmill M & Milton K 1977. The lorisiform wrist joint and the evolution of "brachiating" adaptations in the Hominoidea. *American Journal of Physical Anthropology,* 47 (2): 249-272.

Cox PG & Jeffery N 2010. Semicircular canals and agility: the influence of size and shape measures. *Journal of Anatomy,* 216 (1): 37-47.

Curnoe D & Tobias PV 2006. Description, new reconstruction, comparative anatomy, and classification of the Sterkfontein Stw 53 cranium, with discussions about the taxonomy of other southern African early *Homo* remains. *Journal of Human Evolution,* 50 (1): 36-77.

Day BL & Fitzpatrick RC 2005. The vestibular system. *Current Biology,* 15 (15): R583-R586.

Delezene LK 2015. Modularity of the anthropoid dentition: Implications for the evolution of the hominin canine honing complex. *Journal of Human Evolution,* 86: 1-12.

Dembo M, Matzke NJ, Mooers AØ & Collard M 2015. Bayesian analysis of a morphological supermatrix sheds light on controversial fossil hominin relationships. *Proceedings of the Royal Society B: Biological Sciences,* 282 (1812).

Dembo M, Radovčić D, Garvin HM, Laird MF, Schroeder L, Scott JE, Brophy J, Ackermann RR, Musiba CM, De Ruiter DJ, Mooers AØ & Collard M 2016. The evolutionary relationships and age of *Homo naledi*: An assessment using dated Bayesian phylogenetic methods. *Journal of Human Evolution,* 97: 17-26.

Eastman JM, Alfaro ME, Joyce P, Hipp AL & Harmon LJ 2011. A novel comparative method for identifying shifts in the rate of character evolution on trees. *Evolution,* 65 (12): 3578-89.

Ekdale EG 2013. Comparative anatomy of the bony labyrinth (inner ear) of placental mammals. *PLoS ONE,* 8 (6): e66624.

Ekdale EG 2016. Form and function of the mammalian inner ear. *Journal of Anatomy,* 228 (2): 324-337.

Ekdale EG & Racicot RA 2015. Anatomical evidence for low frequency sensitivity in an archaeocete whale: comparison of the inner ear of *Zygorhiza kochii* with that of crown Mysticeti. *Journal of anatomy,* 226 (1): 22-39.

Felsenstein J 1985. Phylogenies and the comparative method. *The American Naturalist,* 125 (1): 1-15.

Gebo DL 1996. Climbing, brachiation, and terrestrial quadrupedalism: historical precursors of hominid bipedalism. *American Journal of Physical Anthropology,* 101 (1): 55-92.

Georgi JA. 2008. *Semicircular canal morphology as evidence of locomotor environment in amniotes.* PhD Thesis, State University of New York at Stony Brook.

Georgi JA, Sipla JS & Forster CA 2013. Turning semicircular canal function on its head: dinosaurs and a novel vestibular analysis. *PloS one,* 8 (3): e58517.

Graf W & Vidal P-P 1996. Semicircular canal size and upright stance are not interrelated. *Journal of Human Evolution,* 30 (2): 175-181.

Gray AA 1908. *The labyrinth of animals: including mammals, birds, reptiles and amphibians,* London, J. & A. Churchill.

Grine FE 1993. Description and preliminary analysis of new hominid craniodental fossils from the Swartkrans formation. *In:* Brain C (ed.) *Swartkrans: A Cave's Chronicle of Early Man.* . Pretoria: Transvaal Museum.

Grine FE, Jungers W & Schultz J 1996. Phenetic affinities among early *Homo* crania from East and South Africa. *Journal of Human Evolution,* 30 (3): 189-225.

Haeusler M & Mchenry HM 2004. Body proportions of *Homo habilis* reviewed. *Journal of Human Evolution,* 46 (4): 433-65.

Harcourt-Smith WE 2015. The origins of bipedal locomotion. *In:* Henke W & Tattersall I (eds.), *Handbook of paleoanthropology.* 2nd ed. Berlin: Springer-Verlag.

Harcourt-Smith WEH & Aiello LC 2004. Fossils, feet and the evolution of human bipedal locomotion. *Journal of Anatomy,* 204 (5): 403-416.

Harrison T 1991. The implications of *Oreopithecus bambolii* for the origins of bipedalism. *In:* Coppens Y & Senut B (eds.), *Origine (s) de la bipédie chez les hominidés.* Paris: Editions du Centre National de la Recherche Scientifique.

Harvey PH & Pagel M 1991. *The comparative method in evolutionary biology,* Oxford, Oxford University Press.

Henke W & Tattersall I 2015. *Handbook of paleoanthropology,* Berlin, Springer-Verlag.

Higgins RW & Ruff CB 2011. The effects of distal limb segment shortening on locomotor efficiency in sloped terrain: implications for Neandertal locomotor behavior. *American journal of physical anthropology,* 146 (3): 336-345.

Holloway RL, Broadfield DC & Yuan MS-T 2004. *Brain endocasts - The paleoneurological Evidence,* New Jersey, Wiley-Liss.

Holloway RL & Post DG 1982. The relativity of relative brain measures and hominid mosaic evolution. *In: Primate brain evolution.* US: Springer.

Hublin J-J 2015. Paleoanthropology: How old is the oldest human? *Current Biology,* 25 (11): R453-R455.

Hublin J-J, Spoor F, Braun M, Zonneveld F & Condemi S 1996. A late Neanderthal associated with Upper Palaeolithic artefacts. *Nature,* 381 (6579): 224-226.

Hullar TE 2006. Semicircular canal geometry, afferent sensitivity, and animal behavior. *The Anatomical Record Part A: Discoveries in Molecular, Cellular, and Evolutionary Biology,* 288A (4): 466-472.

Hullar TE & Williams CD 2006. Geometry of the semicircular canals of the chinchilla (*Chinchilla lanigen*). *Hearing Research,* 213 (1–2): 17-24.

Iwaniuk AN & Hurd PL 2005. The evolution of cerebrotypes in birds. *Brain, Behavior and Evolution,* 65 (4): 215-230.

Janvier P 1996. *Early vertebrates,* New York, Oxford University Press.

Janvier P 2001. Ostracoderms and the shaping of the gnathostome characters. *Major Events in Early Vertebrate Evolution*: 172-186.

Jones GM & Spells KE 1963. A theoretical and comparative study of the functional dependence of the semicricular canal upon its physical dimensions. *Proceedings of the Royal Society of London. Series B. Biological Sciences,* 157 (968): 403-419.

Jungers WL, Grabowski M, Hatala KG & Richmond BG 2016. The evolution of body size and shape in the human career. *Philosophical Transactions of the Royal Society B: Biological Sciences,* 371 (1698): 20150247.

Kandel BM & Hullar TE 2010. The relationship of head movements to semicircular canal size in cetaceans. *Journal of Experimental Biology,* 213 (7): 1175-1181.

Kappelman J, Ketcham RA, Pearce S, Todd L, Akins W, Colbert MW, Feseha M, Maisano JA & Witzel A 2016. Perimortem fractures in Lucy suggest mortality from fall out of tall tree. *Nature,* 537 (7621): 503-507.

Ketten DR 1992. The marine mammal ear: specializations for aquatic audition and echolocation. *In:* Webster AJ (ed.) *The evolutionary biology of hearing.* New York: Springer.

Kivell TL, Kibii JM, Churchill SE, Schmid P & Berger LR 2011. *Australopithecus sediba* hand demonstrates mosaic evolution of locomotor and manipulative abilities. *Science,* 333 (6048): 1411-1417.

Köhler M & Moyà-Solà S 1997. Ape-like or hominid-like? The positional behavior of *Oreopithecus bambolii* reconsidered. *Proceedings of the National Academy of Sciences,* 94 (21): 11747-11750.

Kratsch C & Mchardy AC 2014. RidgeRace: ridge regression for continuous ancestral character estimation on phylogenetic trees. *Bioinformatics,* 30 (17): i527-i533.

Kuman K & Clarke RJ 2000. Stratigraphy, artefact industries and hominid associations for Sterkfontein, Member 5. *Journal of Human Evolution,* 38 (6): 827-847.

Kutsukake N & Innan H 2013. Simulation-based likelihood approach for evolutionary models of phenotypic traits on phylogeny. *Evolution,* 67 (2): 355-367.

Kutsukake N & Innan H 2014. Detecting phenotypic selection by Approximate Bayesian Computation in phylogenetic comparative methods. *In:* Garamszegi LZ (ed.) *Modern Phylogenetic Comparative Methods and Their Application in Evolutionary Biology.* Berlin: Springer-Verlag.

Lebrun R, De León MP, Tafforeau P & Zollikofer C 2010. Deep evolutionary roots of strepsirrhine primate labyrinthine morphology. *Journal of Anatomy,* 216 (3): 368-380.

Lieberman DE & Bramble DM 2007. The evolution of marathon running. *Sports medicine,* 37 (4-5): 288-290.

Lieberman DE, Bramble DM, Raichlen DA & Shea JJ 2009. Brains, brawn, and the evolution of human endurance running capabilities. *In: The First Humans–Origin and Early Evolution of the Genus Homo.* Springer.

Lovejoy CO, Suwa G, Spurlock L, Asfaw B & White TD 2009. The pelvis and femur of *Ardipithecus ramidus*: The emergence of upright walking. *Science,* 326 (5949): 71-76.

Malinzak MD. 2010. *Experimental analyses of the relationship between semicircular canal morphology and locomotor head rotations in primates.* PhD Thesis, Duke University.

Malinzak MD, Kay RF & Hullar TE 2012. Locomotor head movements and semicircular canal morphology in primates. *Proceedings of the National Academy of Sciences,* 109 (44): 17914-17919.

Manthi FK, Plavcan JM & Ward CV 2012. New hominin fossils from Kanapoi, Kenya, and the mosaic evolution of canine teeth in early hominins. *South African Journal of Science,* 108 (3-4).

Mitteroecker P & Bookstein F 2008. The evolutionary role of modularity and integration in the hominoid cranium. *Evolution,* 62 (4): 943-958.

Mounier A & Caparrós M 2015. The phylogenetic status of *Homo heidelbergensis* - a cladistic study of Middle Pleistocene hominins. *Bulletins et mémoires de la Société d'anthropologie de Paris,* 27 (3-4): 110-134.

Moyà-Solà S, Köhler M & Rook L 1999. Evidence of hominid-like precision grip capability in the hand of the Miocene ape *Oreopithecus*. *Proceedings of the National Academy of Sciences,* 96 (1): 313-317.

Muller M 1994. Semicircular duct dimensions and sensitivity of the vertebrate vestibular system. *Journal of Theoretical Biology,* 167 (3): 239-256.

Organ C, Nunn CL, Machanda Z & Wrangham RW 2011. Phylogenetic rate shifts in feeding time during the evolution of *Homo*. *Proceedings of the National Academy of Sciences USA,* 108 (35): 14555-14559.

Pablos A, Martínez I, Lorenzo C, Sala N, Gracia-Téllez A & Arsuaga JL 2014. Human calcanei from the Middle Pleistocene site of Sima de los Huesos (Sierra de Atapuerca, Burgos, Spain). *Journal of Human Evolution,* 76: 63-76.

Pagel M 2002. Modelling the evolution of continuously varying characters on phylogenetic trees. *In:* Macleod N & Forey PL (eds.), *Morphology, shape and phylogeny.* USA: CRC Press.

Pagel M, Meade A & Barker D 2004. Bayesian estimation of ancestral character states on phylogenies. *Systematic Biology,* 53 (5): 673-84.

Pampush JD 2015. Selection played a role in the evolution of the human chin. *Journal of Human Evolution,* 82: 127-136.

Pfaff C, Martin T & Ruf I 2015. Bony labyrinth morphometry indicates locomotor adaptations in the squirrel-related clade (Rodentia, Mammalia). *Proceedings of the Royal Society B: Biological Sciences,* 282 (1809).

Purves D, Augustine GJ, Fitzpatrick D, Katz LC, Lamantia A-S, Mcnamara JO & Williams SM 2001. The vestibular system. *In: Neuroscience.* 2 ed. Sunderland (MA): Sinauer Associates.

Quam R, Lorenzo C, Martínez I, Gracia-Téllez A & Arsuaga JL 2016. The bony labyrinth of the middle Pleistocene Sima de los Huesos hominins (Sierra de Atapuerca, Spain). *Journal of Human Evolution,* 90: 1-15.

Rabosky DL 2014. Automatic detection of key innovations, rate shifts, and diversity-dependence on phylogenetic trees. *PLoS One,* 9 (2): e89543.

Rae TC 1999. Mosaic Evolution in the Origin of the Hominoidea. *Folia Primatologica,* 70 (3): 125-135.

Raftery AE 1996. Hypothesis testing and model selection. *In:* Gilks WR, Richardson S & Spiegelhalter DJ (eds.), *Markov Chain Monte Carlo in Practice.* London, Great Britain: Chapman & Hall.

Raichlen DA, Armstrong H & Lieberman DE 2011. Calcaneus length determines running economy: Implications for endurance running performance in modern humans and Neandertals. *Journal of Human Evolution,* 60 (3): 299-308.

Raichlen DA & Polk JD 2013. Linking brains and brawn: exercise and the evolution of human neurobiology. *Proceedings of the Royal Society B: Biological Sciences,* 280 (1750).

Ramprashad F, Landolt JP, Money KE & Laufer J 1984. Dimensional analysis and dynamic response characterization of mammalian peripheral vestibular structures. *American Journal of Anatomy,* 169 (3): 295-313.

Retzius G 1881. *Das Gehörorgan der Wirbelthiere: Das Gehörorgan der Fische und Amphibien*, Samson & Wallin.

Retzius G 1884. *Das Gehörorgan der Reptilien, der Vögel und der Säugethiere*, Gedruckt in der Centraldruckerei in Commission bei Samson & Wallin.

Revell LJ, Mahler DL, Peres-Neto PR & Redelings BD 2012. A new phylogenetic method for identifying exceptional phenotypic diversification. *Evolution,* 66 (1): 135-146.

Rook L, Bondioli L, Casali F, Rossi M, Köhler M, Solá SM & Macchiarelli R 2004. The bony labyrinth of *Oreopithecus bambolii*. *Journal of human evolution,* 46 (3): 347-354.

Rook L, Bondioli L, Köhler M, Moyà-Solà S & Macchiarelli R 1999. *Oreopithecus* was a bipedal ape after all: Evidence from the iliac cancellous architecture. *Proceedings of the National Academy of Sciences,* 96 (15): 8795-8799.

Ruff C 2009. Relative limb strength and locomotion in *Homo habilis*. *American Journal of Physical Anthropology,* 138 (1): 90-100.

Russo GA & Shapiro LJ 2013. Reevaluation of the lumbosacral region of *Oreopithecus bambolii*. *Journal of Human Evolution,* 65 (3): 253-265.

Ryan TM, Silcox MT, Walker A, Mao X, Begun DR, Benefit BR, Gingerich PD, Köhler M, Kordos L, Mccrossin ML, Moyà-Solà S, Sanders WJ, Seiffert ER, Simons E, Zalmout IS & Spoor F 2012. Evolution of locomotion in Anthropoidea: the semicircular canal evidence. *Proceedings of the Royal Society B: Biological Sciences,* 279 (1742): 3467-3475.

Sarmiento EE 1983. The significance of the heel process in anthropoids. *International Journal of Primatology,* 4 (2): 127-152.

Schroeder L, Roseman CC, Cheverud JM & Ackermann RR 2014. Characterizing the evolutionary path(s) to early *Homo*. *PloS one,* 9 (12): e114307.

Simpson GG 1953. *The major features of evolution,* London, GB, Columbia University Press.

Sipla JS. 2007. *The semicircular canals of birds and non-avian theropod dinosaurs.* PhD Thesis, State University of New York at Stony Brook.

Spoor F 2003. The semicircular canal system and locomotor behaviour, with special reference to hominin evolution. *Courier-Forschungsinstitut Senckenberg*: 93-104.

Spoor F, Bajpai S, Hussain ST, Kumar K & Thewissen JG 2002. Vestibular evidence for the evolution of aquatic behaviour in early cetaceans. *Nature,* 417 (6885): 163-166.

Spoor F, Garland T, Krovitz G, Ryan TM, Silcox MT & Walker A 2007. The primate semicircular canal system and locomotion. *Proceedings of the National Academy of Sciences USA,* 104 (26): 10808-10812.

Spoor F, Hublin J-J, Braun M & Zonneveld F 2003. The bony labyrinth of Neanderthals. *Journal of Human Evolution,* 44 (2): 141-165.

Spoor F, Wood B & Zonneveld F 1994. Implications of early hominid labyrinthine morphology for evolution of human bipedal locomotion. *Nature,* 369 (6482): 645-648.

Spoor F, Wood B & Zonneveld F 1996. Evidence for a link between human semicircular canal size and bipedal behaviour. *Journal of Human Evolution,* 30 (2): 183-187.

Spoor F & Zonneveld F 1998. Comparative review of the human bony labyrinth. *American Journal of Physical Anthropology,* 107 (S27): 211-251.

Stern Jr JT & Susman RL 1983. The locomotor anatomy of *Australopithecus afarensis*. *American Journal of Physical Anthropology,* 60 (3): 279-317.

Steudel-Numbers KL & Tilkens MJ 2004. The effect of lower limb length on the energetic cost of locomotion: implications for fossil hominins. *Journal of Human Evolution,* 47 (1–2): 95-109.

Stringer C 2012. The status of *Homo heidelbergensis* (Schoetensack 1908). *Evolutionary Anthropology,* 21 (3): 101-107.

Susman RL, Stern Jr JT & Jungers WL 1984. Arboreality and bipedality in the Hadar hominids. *Folia primatologica,* 43 (2-3): 113-156.

Szalay FS & Langdon JH 1986. The foot of *Oreopithecus*: an evolutionary assessment. *Journal of Human Evolution,* 15 (7): 585-621.

Temerin LA & John GHC 1983. The evolutionary divergence of old world monkeys and apes. *The American Naturalist,* 122 (3): 335-351.

Ten Kate J, Van Barneveld H & Kuiper J 1970. The dimensions and sensitivities of semicircular canals. *J Exp Biol,* 53 (2): 501-514.

Thomas GH & Freckleton RP 2012. MOTMOT: Models of trait macroevolution on trees. *Methods in Ecology and Evolution,* 3 (1): 145-151.

Venditti C, Meade A & Pagel M 2011. Multiple routes to mammalian diversity. *Nature,* 479 (7373): 393-396.

Walker A, Ryan TM, Silcox MT, Simons EL & Spoor F 2008. The semicircular canal system and locomotion: The case of extinct lemuroids and lorisoids. *Evolutionary Anthropology: Issues, News, and Reviews,* 17 (3): 135-145.

Ward CV 2002. Interpreting the posture and locomotion of *Australopithecus afarensis*: Where do we stand? *American journal of physical anthropology,* 119 (S35): 185-215.

Weaver TD & Steudel-Numbers K 2005. Does climate or mobility explain the differences in body proportions between Neandertals and their Upper Paleolithic successors? *Evolutionary Anthropology: Issues, News, and Reviews,* 14 (6): 218-223.

White TD, Lovejoy CO & Suwa G 2014. Ignoring *Ardipithecus* in an origins scenario for bipedality is… lame. *Antiquity,* 88 (341): 919-921.

Wood B 2002. Palaeoanthropology: Hominid revelations from Chad. *Nature,* 418 (6894): 133-135.

Wood B 2011. *Wiley-Blackwell Encyclopedia of human evolution,* New Jersey, John Wiley & Sons.

Wood B & Collard M 1999. The human genus. *Science,* 284 (5411): 65-71.

Wood B & Lonergan N 2008. The hominin fossil record: taxa, grades and clades. *Journal of Anatomy,* 212 (4): 354-376.

Xie W, Lewis PO, Fan Y, Kuo L & Chen M-H 2010. Improving marginal likelihood estimation for Bayesian phylogenetic model selection. *Systematic Biology,* 60 (2): 150-160.

Yang A & Hullar TE 2007. Relationship of semicircular canal size to vestibular-nerve afferent sensitivity in mammals. *Journal of Neurophysiology,* 98 (6): 3197-3205.

Appendix 1

Data for extinct hominins

Table A1.1: The arc radius of curvature (mm) for the three semicircular canals (A, anterior; P, posterior; L, lateral) and body mass (BM, g) for extinct hominins. * specimen Stw 53g.

Species	***A***	***P***	***L***	***BM***	***Refs***
Paranthropus robustus	2.60	2.70	2.50	31700	(1; 2)
Australopithecus africanus	2.40	2.60	2.20	30700	(1; 2)
*Homo habilis**	2.90	2.80	2.80	33700	(1; 2)
Homo erectus	3.20	3.10	2.10	51900	(1; 2)
Homo heidelbergensis	3.30	2.80	2.50	68700	(3; 4)
Homo neanderthalensis	3.00	2.80	2.60	64900	(3; 4)

Sources used in Table A1.1

1 Spoor F, Wood B & Zonneveld F 1994. Implications of early hominid labyrinthine morphology for evolution of human bipedal locomotion. *Nature,* 369 (6482): 645-648.

2 Jungers WL, Grabowski M, Hatala KG & Richmond BG 2016. The evolution of body size and shape in the human career. *Philosophical Transactions of the Royal Society B: Biological Sciences,* 371 (1698): 20150247.

3 Quam R, Lorenzo C, Martínez I, Gracia-Téllez A & Arsuaga JL 2016. The bony labyrinth of the middle Pleistocene Sima de los Huesos hominins (Sierra de Atapuerca, Spain). *Journal of Human Evolution,* 90: 1-15.

4 Holloway RL, Broadfield DC & Yuan MS-T 2004. *Brain endocasts - The paleoneurological Evidence,* New Jersey, Wiley-Liss.

Appendix 2

Parameter estimates from regression analyses

Table A2.1: Parameter estimates from the relative canal size models for the anterior semicircular canal. *The beta estimates for parameters estimating the intercept differences for each group are reported with reference to primates; see text for comparison between the great apes and *Homo* intercept.

Parameter		Supported Model	Supported Model + ***Homo***
	mean	-0.154	-0.150
α (Primates)	95% CI	[-0.235,-0.054]	[-0.230,-0.045]
	P_X	0.003	0.003
	mean	0.155	0.154
β (Body Mass)	95% CI	[0.121,0.183]	[0.118,0.181]
	P_X	0.000	0.000
	mean	-0.153	-0.155
β (Great Apes)*	95% CI	[-0.213,-0.084]	[-0.219,-0.084]
	P_{X^X}	0.001	0.000
	mean		-0.084
β (*Homo*)*	95% CI	n/a	[-0.155,-0.005]
	P_X		0.018

Table A2.2: Parameter estimates from the canal-proportions models for the anterior semicircular canal. The clade comprising all *Homo* with the exception of H. habilis is denoted erectus-sapiens* The beta estimates for parameters estimating the intercept differences for each group are reported with reference to primates; see text for comparison between the great apes and erectus-sapiens intercept.

	Parameter	**Supported Model**	**Supported Model + *erectus-sapiens***
	mean	-0.017	-0.016
α (Primates)	95% CI	[-0.070, 0.035]	[-0.070, 0.034]
	P_X	0.257	0.277
	mean	0.691	0.655
β (Posterior)	95% CI	[0.520, 0.858]	[0.473, 0.843]
	P_X	0.000	0.000
	mean	0.269	0.318
β (Lateral)	95% CI	[0.107, 0.437]	[0.150, 0.486]
	P_X	0.000	0.000
	mean	0.028	0.027
β (Body Mass)	95% CI	[0.005, 0.052]	[0.004, 0.050]
	P_X	0.028	0.012
	mean	-0.061	-0.060
β (Great Apes)*	95% CI	[-0.111, -0.009]	[-0.110, -0.011]
	P_X	0.014	0.015
	mean		0.002
β (*erectus-sapiens*)*	95% CI	n/a	[-0.076, 0.077]
	P_X		0.464

Table A2.3: Parameter estimates from relative canal size models for the posterior semicircular canal

	Parameter	Supported Model
α (Primates)	mean	-0.086
	95% CI	[-0.161,-0.012]
	P_X	0.010
β (Body Mass)	mean	0.115
	95% CI	[0.091,0.140]
	P_X	0.000
β (Great Apes)	mean	-0.073
	95% CI	[-0.141,-0.007]
	P_X	0.015

Table A2.4: Parameter estimates for the relative-canals model for the posterior canal

	Parameter	Supported Model	Supported Model + ***nean-heid***
α (Primates)	mean	-0.028	-0.032
	95% CI	[-0.071,0.017]	[-0.074,0.014]
	P_X	0.110	0.076
β (Anterior)	mean	0.587	0.559
	95% CI	[0.430,0.737]	[0.421,0.708]
	P_X	0.000	0.000
β (Lateral)	mean	0.137	0.145
	95% CI	[-0.025,0.300]	[-0.003,0.303]
	P_X	0.048	0.029
β (Body Mass)	mean	0.025	0.029
	95% CI	[0.004,0.047]	[0.007,0.051]
	P_X	0.009	0.006
β (*nean-heid*)	mean		-0.051
	95% CI	n/a	[-0.087,-0.016]
	P_X		0.009

Table A2.5: Parameter estimates for body-size models in the lateral semicircular canal

Parameter		Supported Model	Supported Model + ***Austra-Homo***
α (Primates)	mean	-0.072	-0.072
	95% CI	[-0.147,0.004]	[-0.142,0.005]
	P_X	0.031	0.031
β (Body Mass)	mean	0.100	0.100
	95% CI	[0.074,0.124]	[0.074,0.123]
	P_X	0.000	0.000
β (Great Apes)	mean	-0.074	-0.072
	95% CI	[-0.145,0.001]	[-0.145,0.000]
	P_X	0.027	0.026
β (*Austra-Homo*)	mean		-0.110
	95% CI	n/a	[-0.205,-0.008]
	P_X		0.019

Table A2.6: Parameter estimates for relative-canals models in the lateral canal

Parameter		Supported Model	Supported Model + ***Austra-Homo***
α (Primates)	mean	0.024	0.023
	95% CI	[-0.005,0.054]	[-0.007,0.051]
	P_X	0.053	0.053
β (Anterior)	mean	0.347	0.345
	95% CI	[0.176,0.520]	[0.179,0.515]
	P_X	0.000	0.000
β (Posterior)	mean	0.367	0.374
	95% CI	[0.167,0.564]	[0.173,0.569]
	P_X	0.001	0.001
β (*Austra-Homo*)	mean		-0.038
	95% CI	n/a	[-0.118, 0.569]
	P_X		0.106

Appendix 3

Pairwise differences for the post-hoc lateral semicircular canal analyses

Table A.3.1: Pairwise differences between great apes and hominins in the (A) supported relative canal size model and (B) supported independent canal size model that both allow a separate intercept for each of the hominin species undergoing positive phenotypic selection. Below the diagonal are mean differences; where this is negative it indicates that the species (or group of species) named in the row is smaller than that in the column. Above the diagonal are the $\boldsymbol{P_x}$ values indicating the proportion of the distribution of the pairwise differences that cross zero; where $\boldsymbol{P_x}$ < 0.05 we consider this to be a significant difference (grey). Names are abbreviated as follows: apes = great apes including *Paranthropus* but excluding other hominins; prim = all primates including great apes and *Paranthropus* but excluding other hominins, aus = *A. africanus*; hab = *H. habilis*; ere = *H. erectus*; sap = *H. sapiens*; nean = *H. neanderthalensis*; heid = *H. heidelbergensis*. Other parameters in both models are qualitatively identical to those presented in (A) Table A.2.5 and (B) Table A.2.6 respectively.

A	**apes**	**aus**	**hab**	**ere**	**sap**	**nean**	**heid**
apes	-	0.01	0.02	0.00	0.02	0.31	0.11
aus	-0.05	-	0.00	<0.05	0.14	0.07	0.20
hab	0.05	0.10	-	0.00	0.01	0.02	0.01
ere	-0.10	-0.04	-0.14	-	0.18	0.01	0.02
sap	-0.08	-0.02	-0.13	0.02	-	0.01	0.04
nean	-0.01	0.04	-0.06	0.08	0.06	-	0.09
heid	-0.03	0.02	-0.08	0.06	0.04	-0.02	-

B	**prim**	**aus**	**hab**	**ere**	**sap**	**nean**	**heid**
prim	-	0.01	0.07	0.00	0.00	0.04	0.30
aus	-0.04	-	0.01	0.01	0.01	0.43	0.09
hab	-0.02	0.06	-	0.00	0.00	0.01	<0.05
erect	-0.13	-0.09	-0.16	-	0.15	0.01	0.00
sap	-0.11	-0.07	-0.14	0.02	-	0.01	0.00
nean	-0.04	0.00	-0.07	0.09	0.07	-	0.02
heid	-0.01	0.03	-0.04	0.12	0.10	0.03	-

Chapter 5

Changes in activity pattern explain exceptional bursts of mammal eye shape evolution

Abstract

The evolution of mammalian eye shape has been punctuated by exceptional and intense bursts of rapid morphological change that can be defined as historical positive phenotypic selection. Among vertebrates, there is a general and predictable association between eye morphology and diel activity pattern – nocturnal species are expected to increase the relative size of their pupil to maximize light availability whereas diurnal species favour a decrease in pupil size facilitating high visual acuity. Activity pattern therefore has the potential to exert strong selection pressures on eye morphology and thus may link to some of the unexplained variance in eye shape that manifests as positive phenotypic selection. Our results suggest that shifts in activity pattern during the course of mammalian evolutionary history explain over 86% of all positive phenotypic selection acting to alter the eye shape of mammal species. Strong selection has been exerted as a result of changes from one activity pattern to another but also selection pressures are distinct among species possessing different activity patterns: nocturnal, cathemeral and diurnal species each experience different evolutionary trajectories. Our approach demonstrates a way to identify and understand the underlying causes of exceptional bursts of morphological evolutionary change attributable to historical positive phenotypic selection.

Introduction

Ecological shifts such as changes in activity pattern – the time of day at which a species is active – have the potential to exert strong selection pressures. Birds with a nocturnal lifestyle tend to have larger olfactory bulbs compared to their diurnal relatives (Healy and Guilford, 1990), nocturnal species of gecko are energetically efficient at lower temperatures than diurnal species (Autumn *et al.*, 1999; Autumn *et al.*, 1994) and social groups tend to be larger and more common in diurnal primates (Shultz *et al.*, 2011; Van Schaik, 1983). In these ways and many others (e.g. De Cock and Matthysen, 2005; Ebensperger and Blumstein, 2006; Lyytinen *et al.*, 2004; MacLean *et al.*, 2009; Santini *et al.*, 2015; Wang *et al.*, 2010), activity pattern is likely a key factor underlying biological diversity.

Across vertebrate species, changes in activity pattern are perhaps reflected best in the shape and structure of the eye (e.g. Banks *et al.*, 2015; Hall and Ross, 2007; Hall, 2008a, 2008b; Kay and Kirk, 2000; Kirk and Kay, 2004; Ross *et al.*, 2007; Schmitz and Motani, 2010, 2011; Schmitz and Wainwright, 2011). Nocturnal species are expected to have eye shapes that maximize the amount of light that enters the retina; this can be achieved by increasing the relative size of the pupil (Figure 1). Conversely, in a diurnal setting a greater focal distance improves image clarity (Figure 1, Heesy and Hall, 2010; Ross *et al.*, 2007) and heightens visual acuity (Veilleux and Kirk, 2014). This pattern generally holds among different groups of vertebrates (Hall and Ross, 2007; Hall, 2008b; Kirk, 2004; Motani and Schmitz, 2011; Schmitz and Motani, 2010).

Within mammals, it has been demonstrated that eye shape is overall a relatively poor predictor of activity pattern (Hall *et al.*, 2012; Ross and Kirk, 2007, but see Schmitz and Motani 2010). This is attributed to the idea that most mammals have a typically 'nocturnal eye' (Heesy and Hall, 2010; Lovegrove, 2016) as a result of a *nocturnal bottleneck* early in their evolutionary history (Crompton *et al.*, 1978; Hall *et al.*, 2012; Heesy and Hall, 2010; Walls, 1942) – a long period of nocturnality during which the mammalian eye evolved

many adaptations associated with improved vision in light-poor environments. Some authors have suggested that after initial and intense adaptation to a nocturnal lifestyle, selection pressures on eye shape may not have been sufficient to observe substantial changes in morphology, even if there was a return to a predominantly diurnal lifestyle (Hall *et al.*, 2012; Hall, 2006; Heesy and Hall, 2010). This would imply some sort of constraint acting to restrict change in eye shape, even when it may be otherwise inherently adaptive to do so.

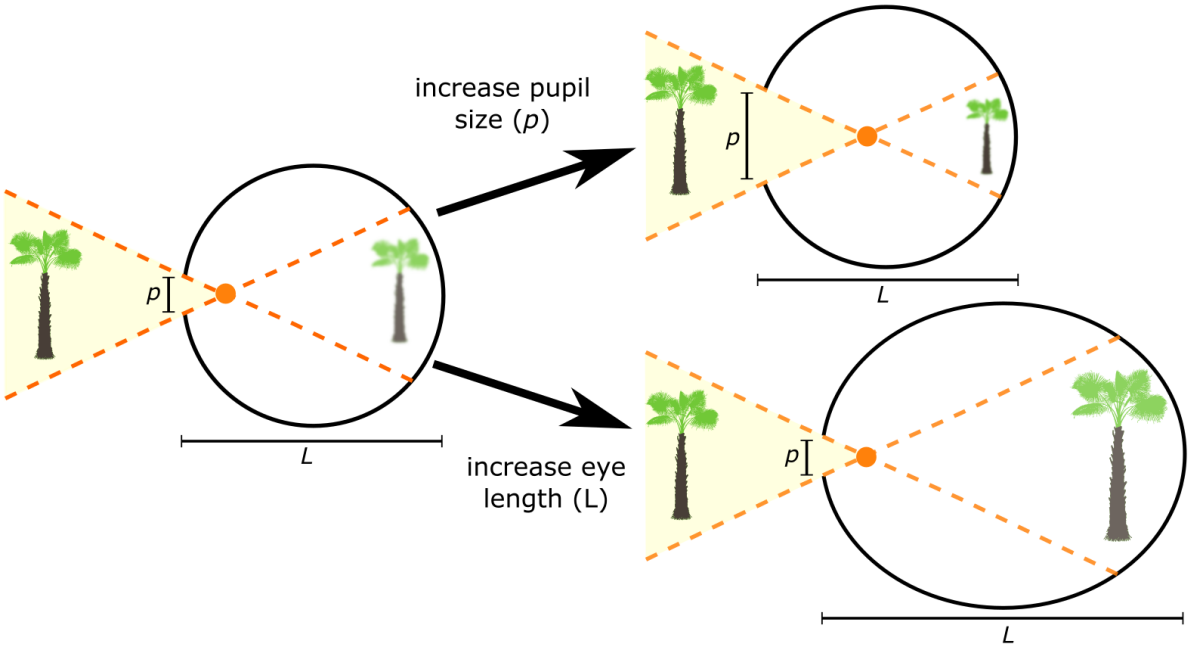


Figure 1: A simplified diagram of the eye that shows how changes in eye shape can influence vision given the same environmental conditions. Increasing pupil size while maintaining the same eye length allows a greater amount of light to enter the eye and results in a brighter image but no increase in clarity. Holding pupil size constant while lengthening focal distance either as a result of increasing eye length (as shown) or enlarging the entire eye will increase image clarity by enlarging the observed image though brightness will remain the same. The silhouette of the palm species *Sabal umbraculifera* is taken from phylopic.org. Image is inspired by and adapted from Figure 3 in Ross *et al.* (2007).

Alone among mammals, diurnal anthropoid primates seem to have escaped the clutches of any such constraint; the corneal size of these animals is substantially smaller than that of other mammal species (e.g. Hall *et al.*, 2012; Ravosa and Savakova, 2004; Ross, 2000; Ross and Kirk, 2007, Figure 2). The unique eye shape of this group is argued to be associated with their shift to a diurnal lifestyle – favouring visual acuity and thus a relatively

reduced corneal diameter compared to eye length (Kirk and Kay, 2004; Ross, 2000; Ross and Kirk, 2007; Veilleux and Kirk, 2014). However, this does not explain why selection could act to change anthropoid eyes but not those of other mammals (Ross and Kirk, 2007) especially considering multiple other transitions to a diurnal lifestyle across the mammalian phylogeny (Gerkema *et al.*, 2013; Roll *et al.*, 2006; Santini *et al.*, 2015).

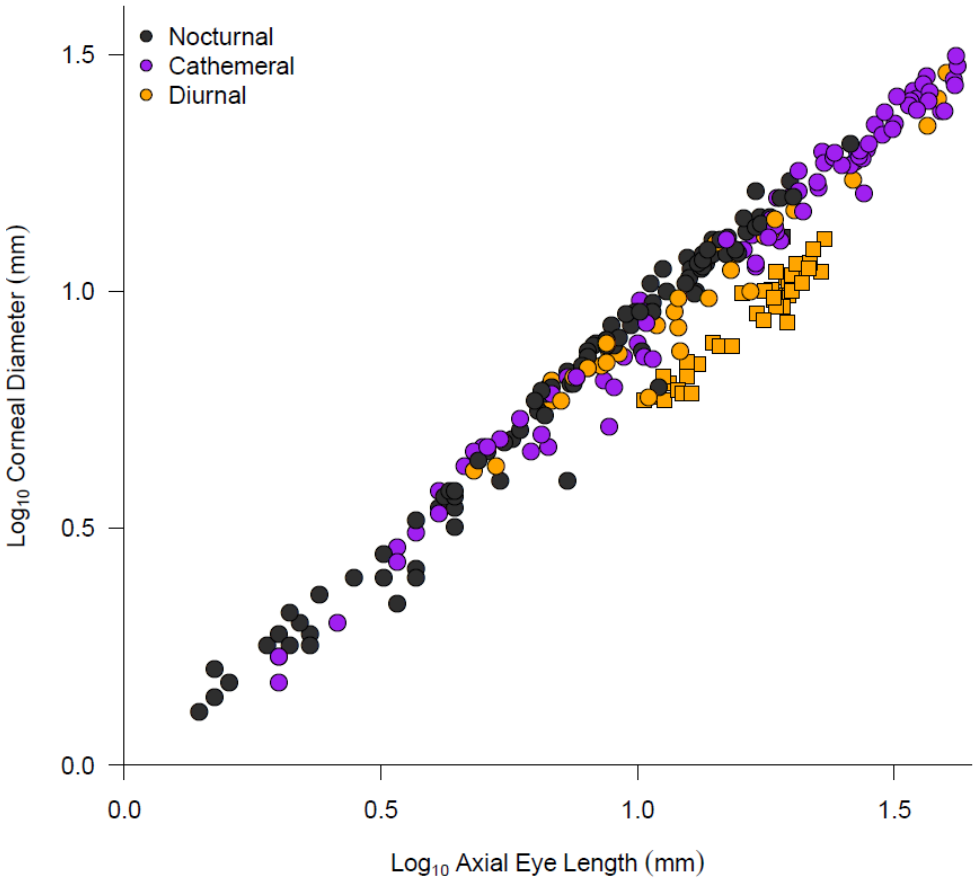


Figure 2: The eye shape of mammal species plotted as the (log_{10}) relationship between corneal diameter and axial eye length. Points are coloured according to activity pattern (see legend). Anthropoid primates are indicated by the squares, all other groups are indicated by circles.

The role of activity pattern in explaining eye shape variation in mammals remains to be fully elucidated. However, weak associations or poor predictive power do not necessarily preclude the fact that activity pattern may have played a role in the evolution of the modern mammalian eye; there may be an alternative way to explain these patterns, or lack thereof.

Initially, such a statement might seem paradoxical, but consider a hypothetical situation in which intense selection pressures are imposed on eye morphology as a consequence of changes in activity pattern. Within one particular clade in the phylogenetic tree, activity pattern may have evolved rapidly along the branches with many transitions between states occurring within a short period of time. We might in this situation expect a reduction in the phylogenetic signal of activity pattern owing to the fast rates of change in one clade of the tree. As eye morphology has been strongly affected by activity pattern, this would also result in rapid rates of evolution in eye shape – which could reduce the strength of its association with activity pattern when we simply look across the tips of the tree. This hypothetical situation highlights the possibility that weak associations might actually result from strong selection pressures: it may therefore be possible to observe historical natural selection in eye shape that is explained by activity patterns in mammals.

Researchers are now beginning to appreciate that there is variation in the rate of morphological evolution that almost certainly reflects adaptive responses to new selective circumstances (e.g. Baker *et al.* 2015; 2016, Eastman *et al.* 2013, Kratsch and McHardy 2014, Kutsukake and Innan 2013; 2014, Rabosky 2014, Venditti *et al.* 2011). Amongst such rate heterogeneity, it is possible to detect exceptional and intense bursts of rapid evolutionary change that can be attributed to positive phenotypic selection (Baker *et al.*, 2016, Chapter 4).

Previously, we have revealed that the evolution of eye shape in mammals has been punctuated by episodes of historical positive phenotypic selection in primates, pangolins and carnivores (Baker *et al.*, 2016), representing substantial and significant unexplained residual variance away from the underlying evolutionary relationship between corneal diameter and axial eye length. If we can identify ecological factors that can explain this unexplained residual variance, such as activity pattern, then we can provide an explicit empirical link between ecology and positive phenotypic selection.

Carnivore species have transitioned between activity patterns up to three times faster than any other order of mammals (Baker *et al.*, 2016). As in the hypothetical situation described above, these rapid transition rates could have exerted strong selection pressures on eye morphology and thus could explain the strong positive phenotypic selection we observe among the branches within this clade – this is despite the poor power of eye shape for predicting activity pattern in carnivores (Hall *et al.*, 2012).

Across all mammals, if activity pattern is the underlying cause of observed positive phenotypic selection acting to sculpt eye shape we would expect a reduction in the number of branches identified as exceptional bursts of evolutionary change when we account for differences in activity patterns. Moreover, we might also expect that each activity pattern might itself impose differential selection pressures on eye shape beyond a simple shift in relative size – the relationship between corneal diameter and axial eye length has been demonstrated to differ in magnitude and even gradient between activity patterns or environmental light intensity among birds and lizards (Hall and Ross, 2007; Hall, 2008b) such that diurnal, nocturnal and cathemeral mammals may each have a distinct *evolutionary trajectory* (Chapter 2).

Here, in the hope of elucidating a direct relationship between ecology and natural selection, we test if activity pattern – a potential important driver of eye shape – can explain episodes of phenotypic selection across the mammal radiation.

Methods

Data and phylogenetic tree

Changes in the shape of the mammal eye were described using the relationship between corneal diameter (as a proxy for pupil size) and axial eye length (as a proxy for focal distance (Hall *et al.*, 2012; Kirk, 2004; Ross *et al.*, 2007). All data was obtained from Hall *et al.* (2012) and included axial eye length (mm), corneal diameter (mm), and activity pattern (nocturnal, diurnal, cathemeral) for n = 266 species spanning 26 orders of mammals.

We used the comprehensive time-scaled phylogeny of mammals from Bininda-Emonds *et al.* (2007). All variables were $\log_{10}$ transformed before analysis. A plot of the data can be seen in Figure 2.

Phylogenetic analyses of eye shape

To detect variation and positive phenotypic selection in the rate of evolution of eye shape in mammals we used the *variable rates regression model* (Baker *et al.*, 2016). This is a recently introduced extension of the *variable rates model* (Venditti *et al.*, 2011), which seeks to identify changes or shifts in the rate of morphological evolution along the branches of a phylogenetic tree. The variable rates regression model is explicitly designed to detect variation in the rates of evolution in phylogenetically structured residual variance of a regression. This allows us to detect instances of fast or slow morphological evolution in one trait (in this case, corneal diameter) after accounting for its relationship with one or more other covariates (axial eye length and activity pattern).

The model simultaneously estimates the relationship between traits whilst also identifying areas of the phylogeny where there has been shifts in the rate of evolution. It works by partitioning the underlying Brownian phylogenetic variance (σ^2) of a continuously varying GLS model of trait evolution (e.g. Pagel, 1999) acting across the branches of the phylogenetic tree into two components: (1) a background rate of morphological evolution (σ^2_{b}) and (2) a set of rate scalars r that define branch-specific shifts away from the background rate. These rate scalars estimate an *optimized variance* for each branch ($\sigma^2_{\mathrm{v}} = \sigma^2_{\mathrm{b}} r$), identifying where a branch has evolved faster ($r > 1$) or slower ($0 \leq r < 1$) as compared to the background rate (where $r = 1$, $\sigma^2_{\mathrm{v}} = \sigma^2_{\mathrm{b}}$ and a branch has evolved at the background rate).

Rate heterogeneity was identified using Bayes factors (BF), calculated as $BF = -2\log_e[m_1/m_0]$, where m_0 is the marginal likelihood of a model that estimates a single Brownian motion rate of evolution across the entire phylogeny and m_1 is the marginal likelihood of a model that allows each branch to have a different rate of evolution (the

variable rates regression model). Where $BF > 2$ it was considered positive support for rate variation (Raftery, 1996).

Positive phenotypic selection was defined by exceptional bursts in the rate of morphological evolution – representing substantial and significant unexplained residual variance away from the underlying evolutionary relationship, in this case, between corneal diameter and axial eye length. We calculated the amount of phenotypic variance along an individual branch as expected from the background rate of change ($\Delta_B = \sigma_b^2 t$) where t is the branch length in time. For each branch, we compared this *background variance* to that arising from shifts in the rate of morphological evolution ($\Delta_V = \sigma_v^2 t$) by calculating the ratio $\Delta_V/\Delta_B = \sigma_v^2 t/\sigma_b^2 t$. Where $\Delta_V/\Delta_B > 2$, the phenotypic variance attributed to increases in the rate of morphological evolution was more than double that attributed to the background variance. The variable rates regression model is implemented within a Bayesian Markov chain Monte Carlo (MCMC) reversible-jump framework thus we calculated a posterior distribution of Δ_V/Δ_B for each branch. We defined an exceptional subclass of branches as episodes of positive phenotypic selection where $\Delta_V/\Delta_B > 2$ in more than 95% of the posterior distribution (Baker *et al.*, 2016).

We first identified positive phenotypic selection using a bivariate regression between corneal diameter and axial eye length (*bivariate linear model*). We then compared the subset of branches identified in this model to those identified as undergoing positive phenotypic selection in two further models: one allowing for different intercepts in the relationship for each of the three activity patterns (*separate-intercepts model)* and another allowing for different slopes in the relationship *(separate-slopes model*).

Branches undergoing positive phenotypic selection in the initial model are those in which there is significant unexplained variance about the underlying relationship between corneal diameter and axial eye length. If this unexplained variance is explained simply by shifts in eye shape corresponding to changes in activity pattern, then we would expect to see a reduction in the number of branches undergoing positive phenotypic selection in

the separate-intercepts model – where we allow different activity patterns to differ in the size of their relative corneal diameter. If positive phenotypic selection can be explained by differential evolutionary trajectories faced by the different activity patterns – i.e. different relationships between corneal diameter and axial length between the different activity patterns e.g. in birds (Hall and Ross, 2007) – we would observe a further reduction in the number of identified branches in our separate-slopes model.

All chains were run for a total of 200 million iterations, sampling every 100,000 iterations after convergence to ensure independence of successive samples. We repeated the analysis with multiple MCMC chains to ensure convergence was achieved.

Significance of the regression parameters was assessed by the proportion of the posterior distribution that crosses zero (P_x). Where less than 5% of a posterior crosses zero, $P_x < 0.05$ and that variable is estimated to be significantly different from zero. To compare parameters amongst different groups i.e. activity pattern, we conducted post-hoc pairwise comparisons between the differences of two parameters at each iteration. Where less than 5% of the posterior distribution of differences crosses zero ($P_x < 0.05$), two parameters are considered to be distinct.

Phylogenetic analyses of activity pattern

We tested the idea that selection pressures on eye shape have not been sufficient to observe substantial changes in eye morphology associated with shifts away from nocturnality (Hall *et al.*, 2012; Hall, 2006; Heesy and Hall, 2010) by linking transitions between activity patterns with results obtained from the variable rates regression analyses.

In order to understand how activity pattern has evolved across mammals, we used a Continuous-time Markov transition model to estimate the relative transition rates between all states across all branches of the phylogenetic tree (Pagel and Meade, 2006). These models estimate the rate at which one state shifts to another given the observed states across the tips and a model of evolution. All rates were estimated within a Bayesian MCMC reversible jump framework (Pagel and Meade, 2006). A gamma distribution was placed as

the prior on transition rates; both the shape and scale parameters of this distribution were seeded from a uniform hyper-prior ranging between 0 and 2 (Currie *et al.*, 2010).

The reversible-jump framework allows for the possibility that rates can be zeroed i.e. there are no shifts observed between those states. It also means that a transition rate can be estimated as identical to one or more other rates i.e. there can be changes in the dimensionality of the model wherein the number of estimated rate parameters can be reduced or increased dynamically. Activity patterns were classified as nocturnal (0), cathemeral (1), or diurnal (2), and we estimated transition rates between each of these states whilst simultaneously estimating trait values at each internal node of the phylogeny. The results we present did not qualitatively differ with alternative prior distributions (Appendix 1).

Results and Discussion

The relationship between corneal diameter and axial eye length estimated in the *bivariate linear model* is depicted in Figure 3A (predicted lines are calculated using 500 random values from the posterior distributions of regression parameters). The posterior distribution of the regression coefficient for the slope is significantly shifted from zero (mean β = 0.92, P_x < 0.0001). There is significant rate heterogeneity in the model (BF = 112.54), and 150 branches are identified as instances of positive phenotypic selection (Figure 3B). In line with our previous analysis (Baker *et al.*, 2016) these branches fall exclusively within anthropoid primates (n = 74), carnivores (n = 75), and the black-bellied pangolin *Manis tetradactyla* (n = 1).

Within anthropoid primates we observe selection along all branches to the exclusion of Atelidae (modal Δ_V/Δ_B ranges between 3.63 and 6.22 along these branches, mode calculated using kernel density estimation across the posterior distribution).

All carnivoran branches are identified as positively selected (modal Δ_V/Δ_B = 3.63-47.11), implying that these mammals have experienced substantial change in eye shape

throughout their evolutionary history (Figure 3B). The rate of morphological evolution, indicated by the length of the stretched branches in Figure 3B, is generally much higher in carnivores than in primates.

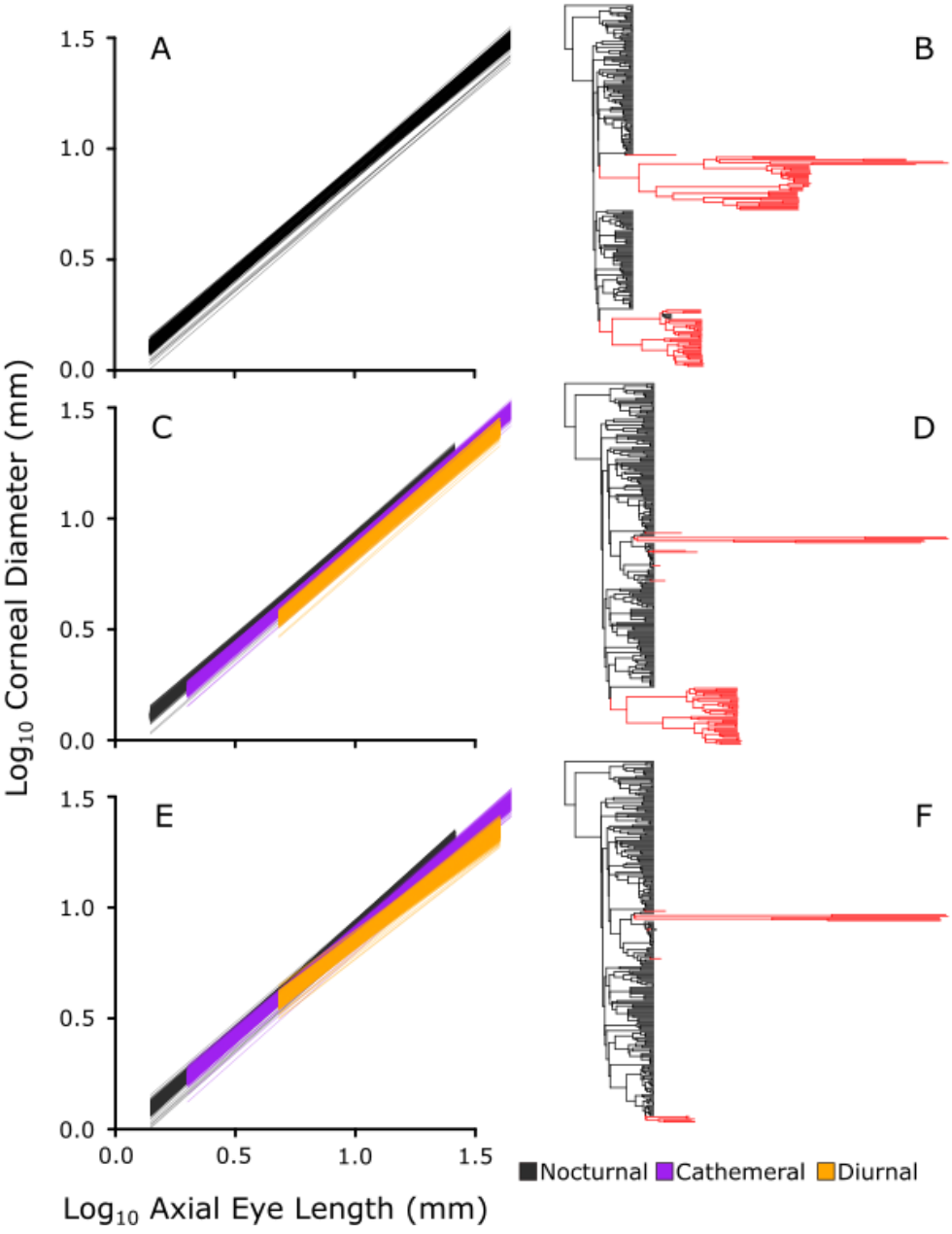


Figure 3: Results from variable rates regression analyses. We show the posterior distribution of predicted phylogenetic regression slopes along with the mammal tree where branches undergoing positive phenotypic selection are coloured red and are stretched according how much they deviate from the background rate of morphological evolution (new branch lengths = rt, where r is the modal rate scalar acting along that branch and t is the original branch length measured in time, modes calculated using kernel density estimation). (A) and (B) show results for the bivariate linear model; (C) and (D) the separate-intercepts model; and (E) and (F) the separate-slopes model.

When we allow each of the different activity patterns to have differences in the intercept of the relationship whilst estimating only a single slope (the *separate-intercepts model*), we find that each of the three groups are significantly different from each other (Figure 3C). Nocturnal species have the largest intercept, indicating that they have significantly larger corneal diameters than both diurnal and cathemeral species (P_x = 0.003 and P_x = 0.009 respectively for the posterior distribution of the differences between intercepts). Cathemeral species have intermediate relative cornea sizes and are significantly larger than diurnal species (P_x = 0.035). This follows the expected pattern of variation (Hall and Ross, 2007; Hall, 2008b; Hall *et al.*, 2012). Although there has been some difficulty in statistically distinguishing cathemeral from both nocturnal and diurnal mammals on the bases of eye shape (Hall and Ross, 2007; Hall *et al.*, 2012; Kirk, 2006; Ross and Kirk, 2007, though see Kirk 2006 and Schmitz and Motani 2010), if we account for heterogeneity in the rate of morphological evolution then it is possible to identify significant differences among the eye shapes of mammal species with different activity patterns.

Accounting for the differences in relative corneal sizes attributed to activity pattern reduces the number of branches that are identified as undergoing positive phenotypic selection by 36% (Figure 3D). Despite this, there is still significant rate heterogeneity (BF = 70.41) and 96 branches are still identified as episodes of historical positive phenotypic selection. Positive phenotypic selection within anthropoid primates, some carnivores, and the pangolin is not sufficiently explained by the grade shifts in relative corneal size associated with transitions between activity patterns (Figure 3D). The unique morphology observed in anthropoid primates therefore cannot be explained simply by a difference in corneal size in diurnal species (Kirk and Kay, 2004; Ross, 2000; Ross and Kirk, 2007). Instead, the variance that is still identified as positive phenotypic selection both in this group and the carnivores may be attributable to differential evolutionary trajectories of each of the activity pattern (as described in Chapter 2).

Within carnivores, we observe many *activity pattern mediated episodes of positive phenotypic selection* (i.e. positive phenotypic selection that disappears after accounting

for activity pattern, Figure 3D). The underlying causes of adaptive eye shape variation observed in most modern carnivores can be attributed to differences in the relative corneal diameter between different activity patterns. This reconciles the apparent paradox presented by the hypothetical situation posed in the introduction; despite the poor predictive power of eye shape (Hall *et al.*, 2012), we observe positive phenotypic selection in carnivores that can largely be explained by shifts in activity pattern. This lends power to the idea that intense and rapid bursts of phenotypic evolution can truly be considered as exceptional instances of historical natural selection as they have an explicit underlying ecological cause.

The *separate-slopes model* detects a significantly increasing slope linked to the amount of daylight within which a species is active: cathemeral, nocturnal, and diurnal species all experience different evolutionary trajectories (Figure 3E). The relationship between corneal diameter and axial length is sharpest in nocturnal mammals (mean β = 0.96) and is significantly different to the relationship observed in both cathemeral (P_x = 0.001) and diurnal (P_x = 0.036) species. Nocturnal species with larger eyes have a larger cornea compared with cathemeral or diurnal species with eyes of a similar size. Conversely, diurnal species have a much shallower trajectory (mean β = 0.84) that is significantly different to that observed for cathemeral species (mean β = 0.92, P_x = 0.012). Although anthropoid primates make up 56% of all diurnal species in the dataset, these relationships do not qualitatively differ when they are excluded from the analysis (Appendix 2).

When we account for the different relationships that exist among activity pattern groups in mammals, there is a further 78% reduction in the number of branches experiencing positive phenotypic selection though there is still significant rate heterogeneity (BF = 66.66).

In primates, 88% of all positive phenotypic selection acting on eye shape can be explained by differences in slope between species of different activity patterns. Owing to the high proportion of diurnality within anthropoid primates, most of the adaptive changes in eye

morphology within this group can be explained by the unique evolutionary trajectory of diurnal species (Figure 3E). Therefore, bursts of evolution in the eye shape of anthropoid primates can be explained not only by their diurnal niche but also by subsequent strong and persistent selection pressure for substantially reduced corneas relative to eye lengths that is experienced by all diurnal species and across all mammals. In this context, anthropoid primates are not 'special'; even when they are removed from the analyses we find the same qualitative pattern in other mammals (Appendix 2).

In the separate slopes model, we observe a reduction not only in the magnitude of the rate of evolution (represented by modal Δ_V/Δ_B) but also in the frequency at which the branches have been scaled in the posterior distribution (Figure 4) – these two components together define positive phenotypic selection (see methods). Across all mammals, 129 branches are undergoing activity pattern mediated episodes of positive phenotypic selection that can only be explained by the different evolutionary trajectory we find in the separate slopes model.

In 62 of the 64 episodes of activity pattern mediated selection identified within carnivores (Figure 3) the modal Δ_V/Δ_B in the separate slopes model is reduced by more than 50% compared to the bivariate linear model (52.7-91.6%, Figure 4A, inset) and is coupled with a reduction in frequency ranging between 12-61% (Figure 4B, inset). Two additional episodes of activity pattern mediated selection occur along branches within the genus *Leopardus* without a huge deceleration in the rate of morphological evolution: the magnitude of Δ_V/Δ_B decreases by 16% in *L. wiedii* and just 0.3% in *L. tigrinus*. However, both of these branches are scaled just 88% of the time when we account for activity pattern (compared to 100% in the bivariate linear model). Most positive phenotypic selection in carnivores is therefore easily explained by activity pattern which acts to sculpt eye shape in a consistent and repeatable manner within this group.

Within anthropoid primates, there are 65 activity pattern mediated episodes of positive phenotypic selection (Figure 3). One subset of these branches is estimated to have a

greater than 60% reduction in the rate of evolution – this can be seen in the division of the distribution of modal Δ_V/Δ_B and reduction within primates shown in Figure 3C. These branches are evolving at approximately the same as the background rate (modal $r = 1$), are scaled ~30% less frequently in the posterior (29.9-34.4%, Figure 4D), and encompass all old world monkeys and apes (excluding macaques), and the new world monkey family Pithecidae.

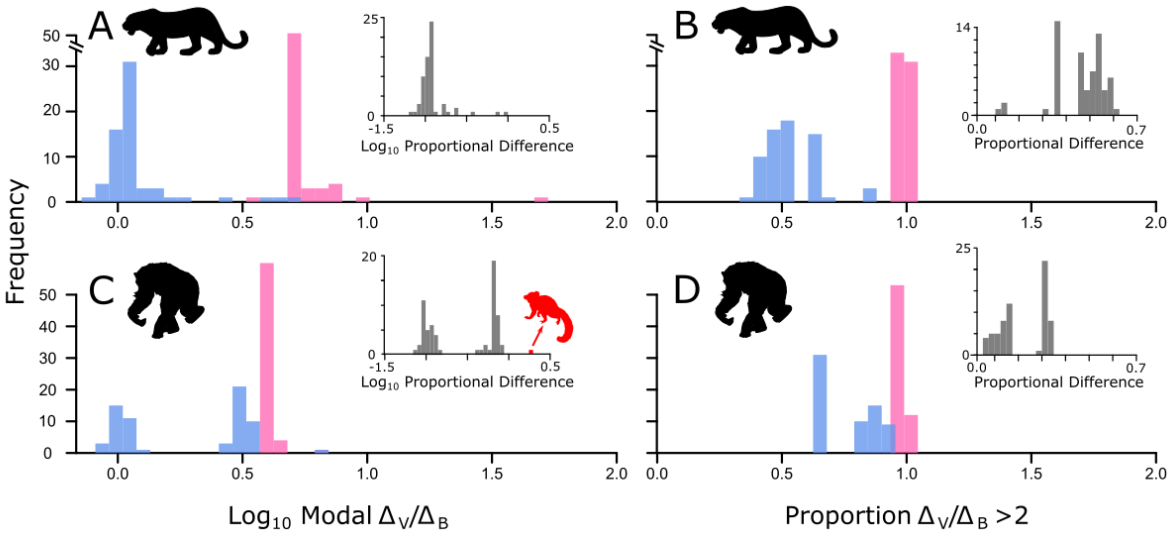


Figure 4: Activity pattern mediated episodes of positive phenotypic selection acting on eye shape. (A) The modal Δ_V/Δ_B across carnivore branches as estimated by the bivariate linear model (pink) compared to that estimated from the separate slopes model (blue). Inset is the distribution of absolute proportional changes in $\log_{10}$ modal Δ_V/Δ_B between the two models. (B) The frequency at which carnivore branches have been scaled in the bivariate linear model (pink) compared to the separate slopes model (blue). Inset is the distribution of proportional changes in frequency between the two models. Negative values indicate that the branch is scaled less frequently in the separate slopes model. (C) and (D) are as in (A) and (B) but for episodes of activity pattern mediate positive phenotypic selection in anthropoid primates. In (C, inset) the common marmoset *Callithrix jacchus* is highlighted in red; this species has a faster rate of evolution after accounting for activity pattern.

The remaining primate branches that reflect activity pattern mediated episodes of selection include the rest of new world monkeys; these branches are scaled up to 20% less often in the separate slopes model compared to the bivariate linear model (Figure 4D) corresponding with only a ~20% drop in modal Δ_V/Δ_B. Note that the branch leading to

the common marmoset (*Callithrix jacchus*) actually has a faster rate of evolution after accounting for activity pattern (highlighted in Figure 4C, inset) but is not identified as undergoing positive phenotypic selection owing to a small reduction in the frequency at which it has been scaled in the posterior distribution (6%). The high rate along the branch leading to this species, like many other branches throughout the phylogeny represents adaptive change in eye shape that cannot be explained by differential trajectories of evolution among activity patterns. However, this variance is not so great such that it can be defined as positive phenotypic selection.

Even after accounting for the different evolutionary trajectories of species of different activity patterns, there are 21 branches identified as experiencing positive phenotypic selection (Figure 3F). These fall predominantly within two groups: in carnivores all branches within and leading to Herpestidae and Eupleridae (mongooses and Malagasy carnivores) have experienced intense and rapid bursts of change in eye shape (modal Δ_V/Δ_B ranges between 11.96-13.94) and within primates, we observe positive phenotypic selection in the monophyletic group comprising mandrills and baboons (modal Δ_V/Δ_B ranges between 10.13-34.72). The variation in eye shape observed in both of these groups of species is sufficiently high such that it cannot be explained simply by activity pattern. Other factors may impose different or even more important selective pressures on eye morphology among these species e.g. brain size (Barton, 2004; Garamszegi *et al.*, 2002) and running speed (Heard-Booth and Kirk, 2012) both have implications for eye size and diet has been shown to be strongly linked to visual acuity (Veilleux and Kirk, 2014).

In addition, positive phenotypic selection is detected along three lineages leading to individual species (Figure 3F). One of these is the branch leading to the black-bellied pangolin *Manis tetradactyla* (modal Δ_V/Δ_B = 24.47). The other two branches fall within carnivores: one leads to the two members of the genus *Leopardus* (modal Δ_V/Δ_B = 6.37) and the other leads to the hairy-nosed otter (*Lutra sumatrana,* modal Δ_V/Δ_B = 4.56). We can identify the direction of change in these individual species by comparing the actual eye morphology to that inferred using phylogenetic imputation (Franks *et al.*, 2012; Jetz

and Freckleton, 2015; Organ *et al.*, 2007). For each of these 4 species (there are two members of the genus *Leopardus*), we predicted corneal diameter from axial eye length using our separate slopes regression model across the stretched tree. Both members of *Leopardus* and also the pangolin have much smaller corneal diameters than would be predicted from the separate slopes model whereas the otter has a much larger cornea than would be expected from our separate slopes model.

These few branches aside, 86% of all positive phenotypic selection acting on eye shape in mammals can be explained entirely by activity pattern (Figure 3). Our results imply that strong selection pressures are exerted on mammal eye morphology as a result of transitions from one activity pattern to another but also that distinctive selection pressures are faced by species possessing different activity patterns. In order to obtain a clear picture of how exactly activity pattern has acted to sculpt natural selection on eye morphology in mammals, we reconstructed transition rates in activity pattern across the mammal phylogeny. We estimated two models independently: one for carnivores and another for the rest of mammals, as we expect carnivores to have higher transition rates among activity patterns (Baker *et al.*, 2016).

Across all mammals (excluding carnivores), we find all transitions are approximately equally likely with one exception – we never observe evolution from a nocturnal ancestor to a diurnal descendant (Figure 5). This implies that in order for diurnal lifestyles to evolve in non-carnivoran mammals, they had to first pass through an intermediate cathemeral phase. On the other hand, diurnal lineages frequently became both nocturnal and cathemeral over the course of mammalian evolutionary history. There are two major differences in the transition rates among carnivoran species compared to other mammals: firstly, nocturnal carnivores frequently evolved diurnality; and secondly, transition rates between cathemerality and diurnality are frequently estimated to be zero in both directions (Figure 5). These results were obtained independently of eye shape, and demonstrate that there has been many shifts in activity pattern across the mammal tree of life (Figure 6, Gerkema *et al.*, 2013; Roll *et al.*, 2006; Santini *et al.*, 2015).

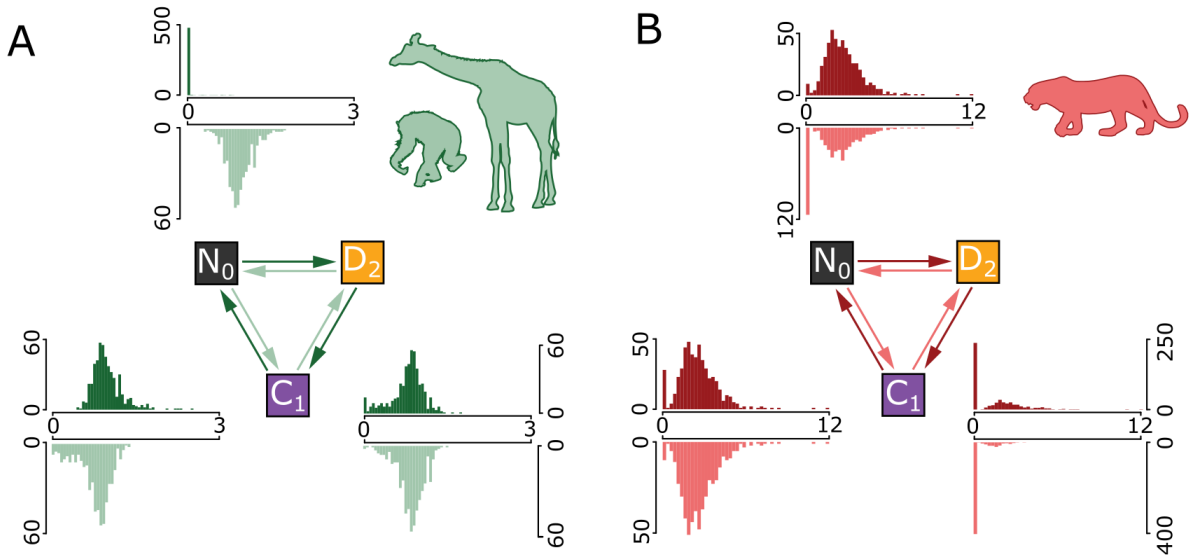


Figure 5: Transition rates between activity patterns in mammals. A) Rates from a model across all mammals excluding carnivores and B) from a model only including carnivores. The transition between each pair of traits is indicated by an arrow – shaded to match the corresponding distribution of estimated rates.

The previously posited release of selection pressures invoked to explain the 'nocturnal' mammalian eye shape (Hall *et al.*, 2012; Hall, 2006; Heesy and Hall, 2010) is associated explicitly with evolution away from nocturnality rather than just changes in activity pattern *per se*. However, when we use our estimated transition rates (Figure 5) to estimate ancestral activity patterns across carnivores we observe multiple instances of activity pattern mediated phenotypic selection along branches that are inferred to have evolved from a nocturnal ancestor (Figure 6). For example, we infer that the ancestral canid was almost certainly nocturnal, but gave rise to both cathemeral and diurnal descendants – all of which experienced positive phenotypic selection acting to change their eye shape (Figure 6). This explicitly demonstrates that selection pressures associated with a shift to day-active lifestyles are substantial and can drive changes in eye morphology.

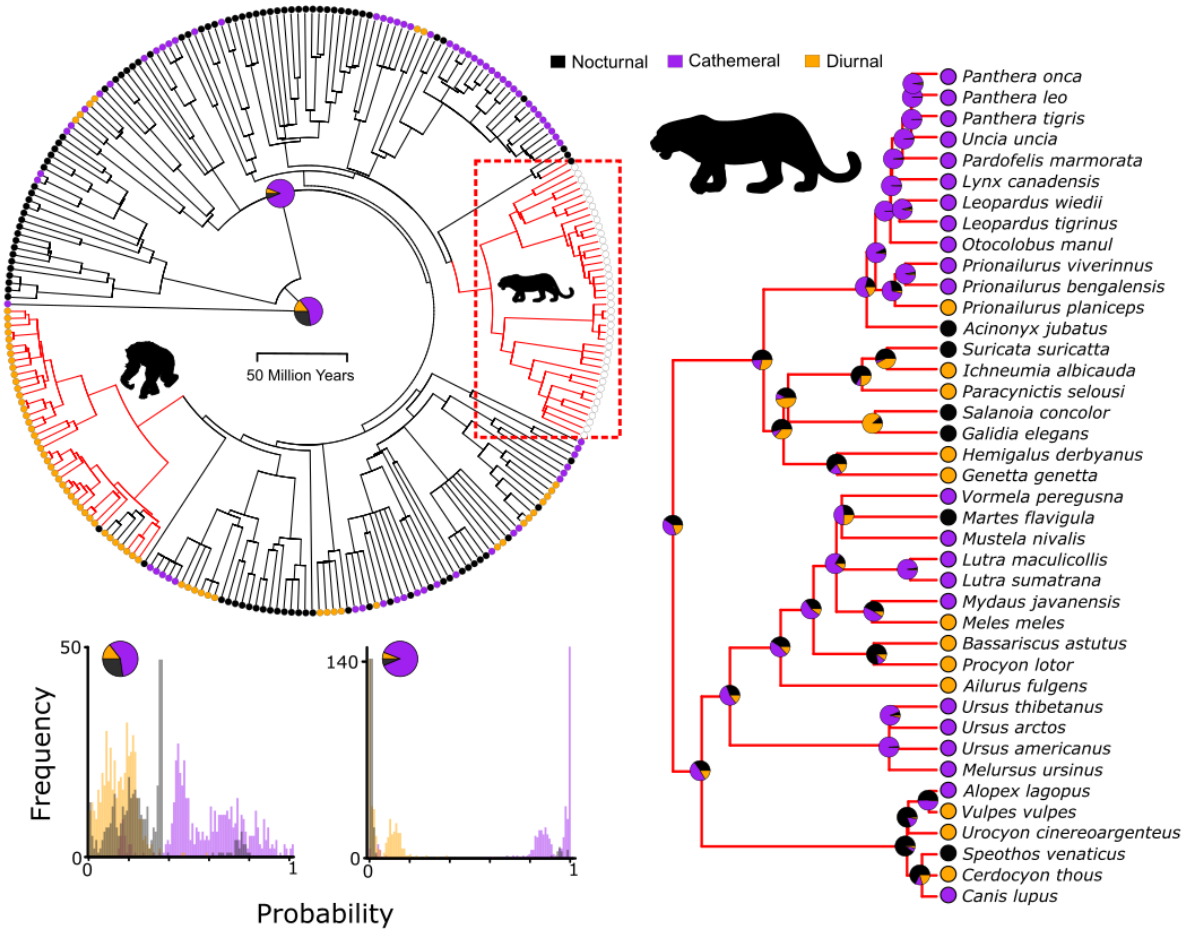


Figure 6: The tree with branches measured in millions of years is shown with activity patterns of each species indicated by coloured points at the tips. Branches coloured red indicate positive phenotypic selection (carnivores are indicated by the grey dashed branches on the full tree). Pie charts indicate reconstructed ancestral states at each of two nodes reconstructed using the model across all mammals excluding carnivores: the root of all mammals and the root of eutherian mammals. The posterior distribution of the estimated probability that the node falls in each state is shown below the tree. The carnivore clade is highlighted and expanded, depicting ancestral state reconstructions at all nodes based on the model of evolution reconstructed across these species only.

Moreover, when we use our model to reconstruct the activity pattern at the root of all mammals, we find that the most likely state was cathemeral (mean probability = 52%, Figure 6). The same is true for the root of eutherians (78% cathemeral, figure 6). This is in contrast to the long-standing idea that the ancestral mammal was nocturnal (Crompton *et al.*, 1978; Walls, 1942). Many morphological and genetic features of mammals are suggested to reflect nocturnality (e.g. Bickelmann *et al.*, 2015; Gerkema *et al.*, 2013; Hall

et al., 2012; Heesy and Hall, 2010), though these features, including eye morphology, are often made with reference to other vertebrate groups However, the degree to which such comparisons might be useful is limited owing to the shared ancestry among species implied by their phylogeny (Felsenstein, 1985; Harvey and Pagel, 1991, Chapter 2, Chapter 4). Although the eye shape of most mammals may on appearances be similar to that of nocturnal birds or reptiles (Hall *et al.*, 2012; Heesy and Hall, 2010), it has achieved this via an entirely different evolutionary route – in this context, the eye shape of mammals is not so much a 'nocturnal' eye as simply a 'mammalian' eye. This might be the case for other so-called nocturnal adaptations; simply possessing these morphologies does not necessarily preclude a non-nocturnal ancestor (Angielczyk and Schmitz, 2014).

Our reconstruction agrees with recent syntheses of photopigment evolution in modern mammals which suggest that the earliest mammals may have lived in low-light but not nocturnal conditions (Davies *et al.*, 2012; Davies *et al.*, 2015; Gerkema *et al.*, 2013). Davies *et al.* (2012) have even proposed a '*twilight-bottleneck*' during which time cellular and molecular features associated with dim-light environments that are still present in modern species could have evolved. An ancestral cathemeral mammal fits with the idea that the earliest members of a group tend to be generalists that radiate into more specialized forms, evolving adaptations to particular lifestyles and niches (Cope, 1896; Mayr, 1942; Simpson, 1953) – in this case into diurnal or nocturnal niches.

These results are the first demonstration that analyses seeking to identify heterogeneity in the rate of morphological change and instances of positive phenotypic selection can be placed within an explicitly ecological context. The framework we develop in this chapter offers researchers a way to analyse direct links between ecology and morphology even in the absence of directional change; providing an opportunity to obtain a deeper understanding of what factors truly drive the evolution of diversity.

References

Angielczyk K & Schmitz L 2014. Nocturnality in synapsids predates the origin of mammals by over 100 million years. *Proceedings of the Royal Society of London B: Biological Sciences,* 281 (1793).

Autumn K, Jindrich D, DeNardo D & Mueller R 1999. Locomotor performance at low temperature and the evolution of nocturnality in geckos. *Evolution,* 53 (2): 580-599.

Autumn K, Weinstein RB & Full RJ 1994. Low cost of locomotion increases performance at low temperature in a nocturnal lizard. *Physiological Zoology,* 67 (1): 238-262.

Baker J, Meade A, Pagel M & Venditti C 2015. Adaptive evolution toward larger size in mammals. *Proceedings of the National Academy of Sciences USA,* 112 (16): 5093-5098.

Baker J, Meade A, Pagel M & Venditti C 2016. Positive phenotypic selection inferred from phylogenies. *Biological Journal of the Linnean Society,* 118: 95-115.

Banks MS, Sprague WW, Schmoll J, Parnell JAQ & Love GD 2015. Why do animal eyes have pupils of different shapes? *Science Advances,* 1 (7): e1500391.

Barton RA 2004. Binocularity and brain evolution in primates. *Proceedings of the National Academy of Sciences of the United States of America,* 101 (27): 10113-10115.

Bickelmann C, Morrow JM, Du J, Schott RK, van Hazel I, Lim S, Müller J & Chang BSW 2015. The molecular origin and evolution of dim-light vision in mammals. *Evolution,* 69 (11): 2995-3003.

Bininda-Emonds OR, Cardillo M, Jones KE, MacPhee RD, Beck RM, Grenyer R, Price SA, Vos RA, Gittleman JL & Purvis A 2007. The delayed rise of present-day mammals. *Nature,* 446 (7135): 507-12.

Cope ED 1896. *The primary factors of organic evolution,* Chicago, Open Court Publishing Company.

Crompton A, Taylor CR & Jagger JA 1978. Evolution of homeothermy in mammals. *Nature,* 272 (5651): 333-336.

Currie TE, Greenhill SJ, Gray RD, Hasegawa T & Mace R 2010. Rise and fall of political complexity in island South-East Asia and the Pacific. *Nature,* 467 (7317): 801-804.

Davies WIL, Collin SP & Hunt DM 2012. Molecular ecology and adaptation of visual photopigments in craniates. *Molecular Ecology,* 21 (13): 3121-3158.

Davies WIL, Tamai TK, Zheng L, Fu JK, Rihel J, Foster RG, Whitmore D & Hankins MW 2015. An extended family of novel vertebrate photopigments is widely expressed and displays a diversity of function. *Genome Research,* 25 (11): 1666-1679.

De Cock R & Matthysen E 2005. Sexual communication by pheromones in a firefly, *Phosphaenus hemipterus* (Coleoptera: Lampyridae). *Animal Behaviour,* 70 (4): 807-818.

Eastman JM, Wegmann D, Leuenberger C & Harmon LJ 2013. Simpsonian 'evolution by jumps' in an adaptive radiation of *Anolis* lizards. arXiv preprint.

Ebensperger LA & Blumstein DT 2006. Sociality in New World hystricognath rodents is linked to predators and burrow digging. *Behavioral Ecology,* 17 (3): 410-418.

Felsenstein J 1985. Phylogenies and the comparative method. *The American Naturalist,* 125 (1): 1-15.

Franks PJ, Freckleton RP, Beaulieu JM, Leitch IJ & Beerling DJ 2012. Megacycles of atmospheric carbon dioxide concentration correlate with fossil plant genome size. *Philosophical Transactions of the Royal Society B: Biological Sciences,* 367 (1588): 556-64.

Garamszegi LZ, Møller AP & Erritzøe J 2002. Coevolving avian eye size and brain size in relation to prey capture and nocturnality. *Proceedings of the Royal Society of London B: Biological Sciences,* 269 (1494): 961-967.

Gerkema MP, Davies WI, Foster RG, Menaker M & Hut RA 2013. The nocturnal bottleneck and the evolution of activity patterns in mammals. *Proceedings of the Royal Society of London B: Biological Sciences,* 280 (1765): 20130508.

Hall M & Ross C 2007. Eye shape and activity pattern in birds. *Journal of Zoology,* 271 (4): 437-444.

Hall MI 2008a. The anatomical relationships between the avian eye, orbit and sclerotic ring: implications for inferring activity patterns in extinct birds. *Journal of Anatomy,* 212 (6): 781-794.

Hall MI 2008b. Comparative analysis of the size and shape of the lizard eye. *Zoology,* 111 (1): 62-75.

Hall MI, Kamilar JM & Kirk EC 2012. Eye shape and the nocturnal bottleneck of mammals. *Proceedings of the Royal Society B: Biological Sciences,* 279 (1749): 4692-4968.

Hall MIM. 2006. *The roles of function and phylogeny in the morphology of the diapsid visual system.* PhD Thesis, State University of New York at Stony Brook.

Harvey PH & Pagel M 1991. *The comparative method in evolutionary biology,* Oxford, Oxford University Press.

Healy S & Guilford T 1990. Olfactory-bulb size and nocturnality in birds. *Evolution,* 44 (2): 339-346.

Heard-Booth AN & Kirk EC 2012. The influence of maximum running speed on eye size: a test of Leuckart's Law in mammals. *The Anatomical Record,* 295 (6): 1053-1062.

Heesy CP & Hall MI 2010. The nocturnal bottleneck and the evolution of mammalian vision. *Brain, behavior and evolution,* 75 (3): 195-203.

Jetz W & Freckleton RP 2015. Towards a general framework for predicting threat status of data-deficient species from phylogenetic, spatial and environmental information. *Philosophical Transactions of the Royal Society B: Biological Sciences,* 370 (1662).

Kay RF & Kirk EC 2000. Osteological evidence for the evolution of activity pattern and visual acuity in primates. *American Journal of Physical Anthropology,* 113 (2): 235-262.

Kirk EC 2004. Comparative morphology of the eye in primates. *The Anatomical Record Part A: Discoveries in Molecular, Cellular, and Evolutionary Biology,* 281 (1): 1095-1103.

Kirk EC 2006. Eye morphology in cathemeral lemurids and other mammals. *Folia Primatologica,* 77 (1-2): 27-49.

Kirk EC & Kay RF 2004. The evolution of high visual acuity in the Anthropoidea. *In:* Ross C & Kay RF (eds.), *Anthropoid Origins.* New York: Kluwer Academic/Plenum Publishers.

Kratsch C & McHardy AC 2014. RidgeRace: ridge regression for continuous ancestral character estimation on phylogenetic trees. *Bioinformatics,* 30 (17): i527-i533.

Kutsukake N & Innan H 2013. Simulation-based likelihood approach for evolutionary models of phenotypic traits on phylogeny. *Evolution,* 67 (2): 355-367.

Kutsukake N & Innan H 2014. Detecting phenotypic selection by Approximate Bayesian Computation in phylogenetic comparative methods. *In:* Garamszegi LZ (ed.) *Modern Phylogenetic Comparative Methods and Their Application in Evolutionary Biology.* Berlin: Springer-Verlag.

Lovegrove BG 2016. A phenology of the evolution of endothermy in birds and mammals. *Biological Reviews,* Early View.

Lyytinen A, Lindström L & Mappes J 2004. Ultraviolet reflection and predation risk in diurnal and nocturnal Lepidoptera. *Behavioral Ecology,* 15 (6): 982-987.

MacLean EL, Barrickman NL, Johnson EM & Wall CE 2009. Sociality, ecology, and relative brain size in lemurs. *Journal of Human Evolution,* 56 (5): 471-478.

Mayr E 1942. *Systematics and the origin of species, from the viewpoint of a zoologist,* Cambridge (MA), Harvard University Press.

Motani R & Schmitz L 2011. Phylogenetic versus functional signals in the evolution of form–function relationships in terrestrial vision. *Evolution,* 65 (8): 2245-2257.

Organ CL, Shedlock AM, Meade A, Pagel M & Edwards SV 2007. Origin of avian genome size and structure in non-avian dinosaurs. *Nature,* 446 (7132): 180-4.

Pagel M 1999. Inferring the historical patterns of biological evolution. *Nature,* 401: 877-884.

Pagel M & Meade A 2006. Bayesian analysis of correlated evolution of discrete characters by reversible-jump Markov chain Monte Carlo. *The American Naturalist,* 167 (6): 808-825.

Rabosky DL 2014. Automatic detection of key innovations, rate shifts, and diversity-dependence on phylogenetic trees. *PLoS One,* 9 (2): e89543.

Raftery AE 1996. Hypothesis testing and model selection. *In:* Gilks WR, Richardson S & Spiegelhalter DJ (eds.), *Markov Chain Monte Carlo in Practice.* London, Great Britain: Chapman & Hall.

Ravosa MJ & Savakova DG 2004. Euprimate origins: the eyes have it. *Journal of Human Evolution,* 46 (3): 355-362.

Roll U, Dayan T & Kronfeld-Schor N 2006. On the role of phylogeny in determining activity patterns of rodents. *Evolutionary Ecology,* 20 (5): 479-490.

Ross CF 2000. Into the light: the origin of Anthropoidea. *Annual Review of Anthropology,* 29: 147-194.

Ross CF, Hall MI & Heesy CP 2007. Were basal primates nocturnal? Evidence from eye and orbit shape. *In:* Ravosa MJ & Dagosto M (eds.), *Primate origins: adaptations and evolution.* New York: Springer.

Ross CF & Kirk EC 2007. Evolution of eye size and shape in primates. *Journal of Human Evolution,* 52 (3): 294-313.

Santini L, Rojas D & Donati G 2015. Evolving through day and night: origin and diversification of activity pattern in modern primates. *Behavioral Ecology*: arv012.

Schmitz L & Motani R 2010. Morphological differences between the eyeballs of nocturnal and diurnal amniotes revisited from optical perspectives of visual environments. *Vision research,* 50 (10): 936-946.

Schmitz L & Motani R 2011. Nocturnality in dinosaurs inferred from scleral ring and orbit morphology. *Science,* 332 (6030): 705-708.

Schmitz L & Wainwright PC 2011. Nocturnality constrains morphological and functional diversity in the eyes of reef fishes. *BMC evolutionary biology,* 11 (1): 1.

Shultz S, Opie C & Atkinson QD 2011. Stepwise evolution of stable sociality in primates. *Nature,* 479 (7372): 219-222.

Simpson GG 1953. *The major features of evolution,* London, GB, Columbia University Press.

Van Schaik CP 1983. Why are diurnal primates living in groups? *Behaviour,* 87 (1): 120-144.

Veilleux CC & Kirk EC 2014. Visual acuity in mammals: effects of eye size and ecology. *Brain, behavior and evolution,* 83 (1): 43-53.

Venditti C, Meade A & Pagel M 2011. Multiple routes to mammalian diversity. *Nature,* 479 (7373): 393-396.

Walls GL 1942. *The vertebrate eye and its adaptive radiation,* Bloomfield Hills (MI), Cranbrook Instititue of Science, Hafner Publishing Company.

Wang G, Zhu Z, Shi P & Zhang Y 2010. Comparative genomic analysis reveals more functional nasal chemoreceptors in nocturnal mammals than in diurnal mammals. *Chinese Science Bulletin,* 55 (34): 3901-3910.

Appendix 1

Phylogenetic analyses of activity pattern

We reconstructed activity patterns across the mammal phylogeny using a Continuous-time Markov transition model allowing all rates to vary (Pagel and Meade, 2006). We used a hyper-prior approach to reduce inherent uncertainty and biases in prior choice (Organ *et al.*, 2009; Pagel and Meade, 2006). We placed a gamma distribution as the prior on transition rates; seeding both shape and scale parameters from a uniform prior ranging between 0 and 2 (Currie *et al.*, 2010). We additionally tested an exponential distribution seeding the λ parameter from a uniform prior ranging between 0 and 2 (Currie *et al.*, 2010; Organ *et al.*, 2009; Shultz *et al.*, 2011). We also repeated both analyses using a covarion model (Tuffley and Steel, 1998) which allows rates to be zeroed; this automatically identifies where transition rates are identical. All analyses using alternative prior distributions produced qualitatively identical results to those presented in the main analysis for models in carnivores and the rest of mammals independently.

References

Currie TE, Greenhill SJ, Gray RD, Hasegawa T & Mace R 2010. Rise and fall of political complexity in island South-East Asia and the Pacific. *Nature,* 467 (7317): 801-804.

Organ CL, Janes DE, Meade A & Pagel M 2009. Genotypic sex determination enabled adaptive radiations of extinct marine reptiles. *Nature,* 461 (7262): 389-92.

Pagel M & Meade A 2006. Bayesian analysis of correlated evolution of discrete characters by reversible-jump Markov chain Monte Carlo. *The American Naturalist,* 167 (6): 808-825.

Shultz S, Opie C & Atkinson QD 2011. Stepwise evolution of stable sociality in primates. *Nature,* 479 (7372): 219-222.

Tuffley C & Steel M 1998. Modeling the covarion hypothesis of nucleotide substitution. *Mathematical biosciences,* 147 (1): 63-91.

Appendix 2

Analyses excluding anthropoid primates

We find positive phenotypic selection in only two groups of mammals: carnivores and anthropoid primates. No other mammal branches, despite multiple shifts in activity pattern across the phylogeny (Figure 5) have any discernible natural selection acting to sculpt eye shape over the millions of years of their evolutionary history (Figure 2) - even in groups with high proportions of day-active species (e.g. 94% of artiodactyl species in this dataset are diurnal). By themselves, anthropoid primates make up 56% of all diurnal species in the dataset. To ensure that the differential evolutionary trajectories that we observe in the main analysis are not being driven by the large proportion of diurnal species in this group, we repeat the bivariate-linear, separate-intercepts, and separate-slopes models as in the main text, but exclude data for anthropoid primates.

In the bivariate linear model excluding anthropoid primates, there is significant rate heterogeneity (BF = 121.75) and 76 branches are identified as instances of positive phenotypic selection. As expected, these fall only within carnivores (and that leading to *M. tetradactyla*). In this model, we find an overall significant and positive relationship between corneal diameter and axial length (mean β = 0.92, P_x = 0.000).

In the separate-intercepts model, we still find distinct grade shifts among the eye shapes of species with different activity patterns. Nocturnal species have the largest intercept, and have significantly larger corneal diameters than both diurnal and cathemeral species (P_x = 0.000 and P_x = 0.000 respectively for the posterior distribution of the differences between intercepts). Cathemeral species have intermediate relative cornea sizes and are significantly larger than diurnal species (P_x = 0.003). There is a 79% reduction in the number of branches identified as positive phenotypic selection compared to the bivariate linear model – amongst the significant rate heterogeneity (BF = 79.46), 16 branches are still identified. 15 of those fall within carnivores – and are identical to those identified in the model including anthropoids (see Figure 2B in the main text). We also observe positive

selection within a single rodent species, *Marmota monax* (the groundhog); in the analysis including anthropoid primates presented in the main text, this branch had a $\Delta_V/\Delta_B > 2$ in 93% of its posterior distribution.

As with the main analysis, we still find significant differences in evolutionary trajectory of eye shape for mammals of different activity patterns even when the predominantly diurnal anthropoid primates are removed. The relationship between corneal diameter and axial length is sharpest in nocturnal mammals (average $\beta = 0.95$) and is significantly different to the relationship observed in both cathemeral ($P_x = 0.04$) and diurnal (P_x 0.001) species. Conversely, diurnal species have a much shallower trajectory (mean $\beta = 0.87$) and is significantly different ($P_x = 0.04$) to that observed for cathemeral species (mean $\beta = 0.92$). There is still significant rate heterogeneity ($BF = 74.52$), but there is just a 25% reduction in the number of episodes of positive phenotypic selection compared to the separate-intercepts model.

The 12 branches that remain as instances of intense bursts of positive phenotypic selection even after accounting for activity pattern are identical to those that remain in the main analysis (excluding those identified with anthropoid primates). Overall, the results presented here demonstrate that excluding the anthropoid primates does not qualitatively affect the conclusions of our main analysis.

Discussion

Researchers now recognize that the evolutionary history of most biological groups has been characterized by variation in the rate of morphological change among and between lineages. However, less is understood about what this means for natural selection. Throughout this thesis it is demonstrated that embracing and exploiting phenotypic rate heterogeneity in evolutionary analyses can reveal patterns and processes of adaptive change deep in time that are otherwise impossible to uncover. The ability to detect complex evolutionary patterns and processes highlights the growing appreciation that tempo and mode are not mutually exclusive components of biological evolution (e.g. Kaliontzopoulou and Adams, 2016). Instead, variation in the rate of phenotypic evolution has the potential to reveal nuanced evolutionary change among lineages and within single groups yet can still give rise to broad-scale patterns of evolution.

For example, Chapter 1 finds that rapid and repeated instances of evolutionary change have driven mammals to larger body sizes over the millions and millions of years of their history. This suggests that Cope's rule – the iconic idea that species have a tendency to increase in size through time (Cope, 1896) – can be explained in mammals (e.g. Alroy, 1998) by adaptive changes in body size playing out over the branches of the phylogenetic tree. With the exception of a recent analysis in a small group of rodents (Avaria-Llautureo *et al.*, 2012), previous attempts to explain or detect historical trends in body size using only data from extant species have been unsuccessful (Bokma *et al.*, 2015; Knouft and Page, 2003; Moen, 2006; Monroe and Bokma, 2010; Pianka, 1995). However, most of these studies assume or test the idea that changes in body size are concentrated in speciation events (Avaria-Llautureo *et al.*, 2012; Bokma *et al.*, 2015; Knouft and Page, 2003; Moen, 2006; Monroe and Bokma, 2010), inspired by the theory of punctuated equilibrium developed by Eldredge and Gould (1972) . While there is some contention regarding a link between speciation and rates of morphological evolution (e.g. Adams *et al.*, 2009; Rabosky

and Adams, 2012; Rabosky and Matute, 2013; Venditti *et al.*, 2011), Cope's rule is defined as the tendency for species within a group to evolve to a larger body size through time (Alroy, 1998; Hone and Benton, 2005) and makes no specific predictions regarding speciation events. The methods outlined in Chapter 1 may therefore be the only current approach available for detecting and understanding patterns of directional change using data from living species.

Using evidence from extinct species, palaeontologists have found Cope's rule in some groups, but not others (e.g. Alroy, 1998; Arnold *et al.*, 1995; Hone and Benton, 2007; Hunt and Roy, 2006; Jablonski, 1997; Raia *et al.*, 2012; Sookias *et al.*, 2012; Van Valkenburgh *et al.*, 2004; Zanno and Makovicky, 2013). Approaches that allow us to understand the patterns of evolution in deep time using living species will help us to determine just how universal Cope's rule might be. With the advent of modern tree-building methods (Bouckaert *et al.*, 2014; Drummond *et al.*, 2006; Drummond *et al.*, 2012; Ronquist *et al.*, 2012) and an increasing availability of genomic and morphological data, phylogenetic trees of extant species are now available that contain thousands of species and span entire vertebrate classes such as mammals (Bininda-Emonds *et al.*, 2007; Faurby and Svenning, 2015; Fritz *et al.*, 2009; Hedges *et al.*, 2006; Hedges *et al.*, 2015), birds (Davis and Page, 2008; Jetz *et al.*, 2012; Prum *et al.*, 2015), amphibians (Pyron and Wiens, 2013), and bony fish (Rabosky *et al.*, 2013) even up to the recently published tree of life which includes more than 50,000 species from all walks of life including bacteria (Hedges *et al.*, 2015). Such trees can help us recover patterns of adaptive evolution in other groups, affording the opportunity to understand not only the prevalence of Cope's rule and how this matches with evidence from the fossil record, but also how other trends have acted during historical evolution.

The ability to study trends in morphology that occur deep in time using only data from living species provides a unique and novel way to study trends in biological characteristics that are rare or otherwise impossible to recover in the fossil record such as soft tissues, diet and behaviour, or pigmentation. This idea is developed further in Chapter 2 in which

adaptive trends towards smaller testes sizes are found throughout vertebrate evolution. Testes (or gonads), like other soft tissues, are occasionally fossilized (e.g. Sutton *et al.*, 2006) but are not found frequently enough in the fossil record to make meaningful inferences by any other means.

Phenotypic rate heterogeneity has the potential to reveal adaptive changes in many traits and across many groups by looking at broad scale patterns across whole phylogenetic trees. However, understanding fine scale differences in rates between individual species may reveal how the strength of natural selection has acted to sculpt the morphology of individual species and this varies among lineages. If the rate of evolution represents the intensity of natural selection, then it should be possible to identify an exceptional subclass of branches that experience bursts of morphological change that can be attributed to positive selection. On this basis, Chapters 3 and 4 introduce a novel metric for detecting positive phenotypic selection based entirely on variation in the rate of morphological evolution. Both chapters demonstrate that episodic instances of positive phenotypic selection have been an important factor during the evolution of all groups studied. It is shown that selective pressures vary in strength and direction, but also that positive phenotypic selection differs in how often it punctuates the evolution of different groups. In Chapter 3 it is found that over 20% of all fleshy-fruited plants have experienced intense selection pressures acting to sculpt fruit sizes whereas only 5% of branches in *Anolis* represent bursts of change in body size attributable to positive phenotypic selection.

Much like the approaches described in Chapters 1 and 2, using phenotypic rate heterogeneity to identify episodes of intense historical natural selection provides the only opportunity to study positive selection deep in time – an ability previously only accessible to geneticists (Murrell *et al.*, 2013; Nadeau and Jiggins, 2010; Yang, 2002, 2006). However, the link between genotype and phenotype is often not clear - in fact, there is a huge body of literature seeking to find genes where changes in sequence results in direct phenotypic consequences (e.g. Lartillot and Poujol, 2011; Nadeau *et al.*, 2007). It is possible to detect positive selection in a genetic context where changes in sequence result in predictable

changes in morphology e.g. (Montgomery *et al.*, 2011; Montgomery and Mundy, 2012; Nadeau *et al.*, 2007). However, where such links are less clear for a given phenotype – (where there are a combination of genes acting or it is not known which genes are important) – the approaches described in Chapters 3, 4 and 5 may be able to help. That is, where we detect positive *phenotypic* selection acting to change a given morphology in a group of species or even in individual species, it may be possible to find corresponding positive *genetic* selection in the same lineages. This may reveal links between genotype and phenotype in the absence of a general directional effect. The approaches presented throughout this thesis can tell us about directionality of adaptive evolution and where natural selection has acted to drive morphological change in a particular way – for example Chapters 1 and 2 but also Chapters 3 and 4 where we find directional change acting on both molar size and the dimensions of the semicircular canals within the inner ear that can be linked to key evolutionary transitions during hominin evolution.

Looking at all of these things – directionality, trends, and positive phenotypic selection – as a part of a bigger picture may help us to understand how evolution acted over important ecological transitions and during times of innovation – including the many that are believed to have characterized the human lineage (e.g. Antón *et al.*, 2014). Finding positive phenotypic selection in morphologies that can be linked to traversal of novel adaptive zones (Simpson, 1944; Simpson, 1953) will enhance our understanding of how natural selection has acted to give rise to modern diversity. However, the ultimate goal of any study of adaptation is to truly understand the drivers of diversity. In Chapter 5 it is demonstrated that it is possible identify the underlying causes of positive phenotypic selection. This is presented in the context of natural selection acting on eye shape in mammals; an extraordinary 86% of all positive phenotypic selection can be explained simply by changes in activity pattern (the time of day at which a species is active). The same logic can also be extended to any study of adaptive change using the approaches described throughout this thesis.

In Chapter 2 we discuss the idea that adaptive decreases in testes size across vertebrate evolutionary history may be linked to the expensive- tissue hypothesis (Aiello and Wheeler, 1995) which suggests that the size of tissues that are energetically expensive to grow and maintain trade-off with the size of other expensive organs – one cannot increase without a decrease in another. In the context of testes sizes, evidence for this hypothesis is conflicting: There is some evidence that there has been a trade-off between brain and testes size during the evolution of bats (Pitnick *et al.*, 2006), though other authors disagree with this (Lemaître *et al.*, 2009) and there has been little evidence found in cetaceans (Kelley *et al.*, 2014), primates (Schillaci, 2006), or rodents (Bordes *et al.*, 2011). Generally, the expensive-tissue hypothesis is investigated by testing correlations between two morphologies: where this is negative, this is taken as support for a trade-off (e.g. Aiello and Wheeler, 1995; Liao *et al.*, 2016; Warren and Iglesias, 2012). This thesis provides a novel angle of attack from which to address this idea. Perhaps metabolic rate explains some of the rate heterogeneity in testes size – even after accounting for body sizes. It should even be possible to determine whether there have been directional shifts to smaller testes arising as a consequence of positive phenotypic selection that correspond with any similar directional shifts in the opposite direction in other tissues and organs.

It is certain that the growing interest in phenotypic heterogeneity will hold its place in the minds of evolutionary biologists and palaeontologists for many years to come. The approaches outlined in this thesis provide a framework within which it is possible to describe and understand the processes of phenotypic evolution at a far more nuanced level than was previously conceivable. The promise and potential power of this framework comes by virtue of the ability to detect episodes of adaptive evolution – that in some cases can translate to historical trends over millions and millions of years – combined with a unique approach to understanding the underlying drivers of such adaptive changes. Looking to the future, researchers should be seeking to identify gaps in knowledge that can be both explained by and revealed by heterogeneity in rates.

References

Adams DC, Berns CM, Kozak KH & Wiens JJ 2009. Are rates of species diversification correlated with rates of morphological evolution? *Proceedings of the Royal Society B: Biological Sciences,* 276 (1668): 2729-2738.

Aiello LC & Wheeler P 1995. The expensive-tissue hypothesis: the brain and the digestive system in human and primate evolution. *Current anthropology,* 36 (2): 199-221.

Alroy J 1998. Cope's rule and the dynamics of body mass evolution in North American fossil mammals. *Science,* 280 (5364): 731-734.

Antón SC, Potts R & Aiello LC 2014. Evolution of early *Homo*: An integrated biological perspective. *Science,* 345 (6192).

Arnold AJ, Kelly DC & Parker WC 1995. Causality and Cope's rule: Evidence from the planktonic foraminifera. *Journal of Paleontology,* 69 (2): 203-210.

Avaria-Llautureo J, Hernandez CE, Boric-Bargetto D, Canales-Aguirre CB, Morales-Pallero B & Rodriguez-Serrano E 2012. Body size evolution in extant Oryzomyini rodents: Cope's rule or miniaturization? *PLoS One,* 7 (4): e34654.

Bininda-Emonds OR, Cardillo M, Jones KE, MacPhee RD, Beck RM, Grenyer R, Price SA, Vos RA, Gittleman JL & Purvis A 2007. The delayed rise of present-day mammals. *Nature,* 446 (7135): 507-12.

Bokma F, Godinot M, Maridet O, Ladevèze S, Costeur L, Solé F, Gheerbrant E, Peigné S, Jacques F & Laurin M 2015. Testing for Depéret's rule (body size increase) in mammals using combined extinct and extant data. *Systematic Biology,* 65 (1): 98-108.

Bordes F, Morand S & Krasnov BR 2011. Does investment into "expensive" tissue compromise anti-parasitic defence? Testes size, brain size and parasite diversity in rodent hosts. *Oecologia,* 165 (1): 7-16.

Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu C-H, Xie D, Suchard MA, Rambaut A & Drummond AJ 2014. BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Computational Biology,* 10 (4): e1003537.

Cope ED 1896. *The primary factors of organic evolution,* Chicago, Open Court Publishing Company.

Davis KE & Page RD 2008. Reweaving the tapestry: a supertree of birds. *PLoS currents,* 6.

Drummond AJ, Ho SYW, Phillips MJ & Rambaut A 2006. Relaxed phylogenetics and dating with confidence. *PLoS Biology,* 4 (5): e88.

Drummond AJ, Suchard MA, Xie D & Rambaut A 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular biology and evolution,* 29 (8): 1969-1973.

Eldredge N & Gould SJ 1972. Punctuated equilibria: An alternative to phyletic gradualism. *In:* Schopf TJM (ed.) *Models in Paleobiology.* San Francisco: Freeman, Cooper and Company.

Faurby S & Svenning J-C 2015. A species-level phylogeny of all extant and late Quaternary extinct mammals using a novel heuristic-hierarchical Bayesian approach. *Molecular Phylogenetics and Evolution,* 84: 14-26.

Fritz SA, Bininda-Emonds ORP & Purvis A 2009. Geographical variation in predictors of mammalian extinction risk: big is bad, but only in the tropics. *Ecology letters,* 12 (6): 538-549.

Hedges SB, Dudley J & Kumar S 2006. TimeTree: a public knowledge-base of divergence times among organisms. *Bioinformatics,* 22 (23): 2971-2972.

Hedges SB, Marin J, Suleski M, Paymer M & Kumar S 2015. Tree of life reveals clock-like speciation and diversification. *Molecular biology and evolution,* 32 (4): 835-845.

Hone DW & Benton MJ 2005. The evolution of large size: How does Cope's rule work? *Trends in Ecology & Evolution,* 20 (1): 4-6.

Hone DW & Benton MJ 2007. Cope's rule in the Pterosauria, and differing perceptions of Cope's rule at different taxonomic levels. *Journal of Evolutionary Biology,* 20 (3): 1164-70.

Hunt G & Roy K 2006. Climate change, body size evolution, and Cope's rule in deep-sea ostracodes. *Proceedings of the National Academy of Sciences USA,* 103 (5): 1347-52.

Jablonski D 1997. Body-size evolution in Cretaceous molluscs and the status of Cope's rule. *Nature,* 385: 250-252.

Jetz W, Thomas G, Joy J, Hartmann K & Mooers A 2012. The global diversity of birds in space and time. *Nature,* 491 (7424): 444-448.

Kaliontzopoulou A & Adams DC 2016. Phylogenies, the Comparative Method, and the Conflation of Tempo and Mode. *Systematic Biology,* 65 (1): 1-15.

Kelley TC, Higdon JW & Ferguson SH 2014. Large testes and brain sizes in odontocetes (order Cetacea, suborder Odontoceti): the influence of mating system on encephalization. *Canadian Journal of Zoology,* 92 (8): 721-726.

Knouft JH & Page LM 2003. The evolution of body sze in extant groups of North American freshwater fishes: Speciation, size distributions, and Cope's rule. *The American Naturalist,* 161 (3): 413-421.

Lartillot N & Poujol R 2011. A phylogenetic model for investigating correlated evolution of substitution rates and continuous phenotypic characters. *Molecular Biology and Evolution,* 28 (1): 729-744.

Lemaître JF, Ramm SA, Barton RA & Stockley P 2009. Sperm competition and brain size evolution in mammals. *Journal of Evolutionary Biology,* 22 (11): 2215-2221.

Liao WB, Lou SL, Zeng Y, Kotrschal A, Leips J & Bronstein JL 2016. Large brains, small guts: The expensive tissue hypothesis supported within anurans. *The American Naturalist,* 188 (6).

Moen DS 2006. Cope's rule in cryptodiran turtles: Do the body sizes of extant species reflect a trend of phyletic size increase? *Journal of Evolutionary Biology,* 19 (4): 1210-1221.

Monroe MJ & Bokma F 2010. Little evidence for Cope's rule from Bayesian phylogenetic analysis of extant mammals. *Journal of evolutionary biology,* 23 (9): 2017-21.

Montgomery SH, Capellini I, Venditti C, Barton RA & Mundy NI 2011. Adaptive evolution of four microcephaly genes and the evolution of brain size in anthropoid primates. *Molecular Biology and Evolution,* 28 (1): 625-638.

Montgomery SH & Mundy NI 2012. Evolution of ASPM is associated with both increases and decreases in brain size in primates. *Evolution,* 66 (3): 927-932.

Murrell B, Moola S, Mabona A, Weighill T, Sheward D, Pond SLK & Scheffler K 2013. FUBAR : A Fast, Unconstrained Bayesian AppRoximation for inferring selection. *Molecular biology and evolution,* 30 (5): 1196-1205.

Nadeau NJ, Burke T & Mundy NI 2007. Evolution of an avian pigmentation gene correlates with a measure of sexual selection. *Proceedings of the Royal Society B: Biological Sciences,* 274 (1620): 1807-1813.

Nadeau NJ & Jiggins CD 2010. A golden age for evolutionary genetics? Genomic studies of adaptation in natural populations. *Trends in Genetics,* 26 (11): 484-492.

Pianka ER 1995. Evolution of body size: Varanid lizards as a model system. *The American Naturalist,* 146 (3): 398-414.

Pitnick S, Jones KE & Wilkinson GS 2006. Mating system and brain size in bats. *Proceedings of the Royal Society of London B: Biological Sciences,* 273 (1587): 719-724.

Prum RO, Berv JS, Dornburg A, Field DJ, Townsend JP, Lemmon EM & Lemmon AR 2015. A comprehensive phylogeny of birds (Aves) using targeted next-generation DNA sequencing. *Nature,* 526 (7574): 569-573.

Pyron RA & Wiens JJ 2013. Large-scale phylogenetic analyses reveal the causes of high tropical amphibian diversity. *Proceedings of the Royal Society of London B: Biological Sciences,* 280 (1770): 20131622.

Rabosky DL & Adams DC 2012. Rates of morphological evolution are correlated with species richness in salamanders. *Evolution,* 66 (6): 1807-1818.

Rabosky DL & Matute DR 2013. Macroevolutionary speciation rates are decoupled from the evolution of intrinsic reproductive isolation in *Drosophila* and birds. *Proceedings of the National Academy of Sciences USA,* 110 (38): 15354-15359.

Rabosky DL, Santini F, Eastman J, Smith SA, Sidlauskas B, Chang J & Alfaro ME 2013. Rates of speciation and morphological evolution are correlated across the largest vertebrate radiation. *Nature Communications,* 4.

Raia P, Carotenuto F, Passaro F, Fulgione D & Fortelius M 2012. Ecological specialization in fossil mammals explains Cope's rule. *The American Naturalist,* 179 (3): 328-337.

Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA & Huelsenbeck JP 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic biology,* 61 (3): 539-542.

Schillaci MA 2006. Sexual selection and the evolution of brain size in primates. *PLoS ONE,* 1 (1): e62.

Simpson GG 1944. *Tempo and mode in evolution,* London, GB, Columbia University Press.

Simpson GG 1953. *The major features of evolution,* London, GB, Columbia University Press.

Sookias RB, Butler RJ & Benson RB 2012. Rise of dinosaurs reveals major body-size transitions are driven by passive processes of trait evolution. *Proceedings of the Royal Society B: Biological Sciences,* 279 (1736): 2180-7.

Sutton MD, Briggs DEG, Siveter DJ & Siveter DJ 2006. Fossilized soft tissues in a Silurian platyceratid gastropod. *Proceedings of the Royal Society B: Biological Sciences,* 273 (1590): 1039-1044.

Van Valkenburgh B, Wang X & Damuth J 2004. Cope's rule, hypercarnivory, and extinction in North American canids. *Science,* 306 (5693): 101-104.

Venditti C, Meade A & Pagel M 2011. Multiple routes to mammalian diversity. *Nature,* 479 (7373): 393-396.

Warren DL & Iglesias TL 2012. No evidence for the 'expensive-tissue hypothesis' from an intraspecific study in a highly variable species. *Journal of Evolutionary Biology,* 25 (6): 1226-1231.

Yang Z 2002. Inference of selection from multiple species alignments. *Current Opinion in Genetics & Development,* 12 (6): 688-694.

Yang Z 2006. *Computational molecular evolution,* Oxford, Great Britain, Oxford University Press.

Zanno LE & Makovicky PJ 2013. No evidence for directional evolution of body mass in herbivorous theropod dinosaurs. *Proceedings of the Royal Society B: Biological Sciences,* 280 (1751).