Testing Correlations in Morphological and Molecular Evolution: a Meta-analysis Approach

A thesis submitted to the University of Manchester for the degree of
Master of Philosophy
in the Faculty of Science and Engineering

2019

Joel A Heath

School of Natural Sciences
Department of Earth and Environmental Sciences

List of Contents

Word count: 20219

1 Abstract:

Understanding the relationship between genomic and phenotypic evolution, and the factors that facilitate interactions between these processes, is of central importance in evolutionary biology. Mutations giving rise to beneficial phenotypes provide the foundation for descent with modification, and play a major role in evolutionary radiations. However, most genome-wide substitutions occur stochastically, whilst the tempo of phenotypic evolution is driven directly by selection, and it is unclear how these processes are linked. Deciphering the relationship between genomic and phenotypic rates of evolution will yield crucial insight into molecular processes driving adaptive evolution, and allow us to better understand the underlying forces structuring biodiversity. Here, correlations between molecular and morphological branch lengths were tested in a wide range of animal phylogenies reconstructed using Bayesian inference with datasets from 12 recently published total evidence studies. Correlations between rates of molecular and morphological evolution along branches were also tested in time calibrated phylogenies reconstructed using relaxed clock methods. Only very weak correlations were detected between molecular and morphological evolution in terms of both branch lengths and their rates, with the majority of groups studied showing no correlation at all. Analysing rates through time suggests that this apparent disconnect may stem from molecular evolution remaining unaffected by the intensity of selection that often drives accelerated morphological evolution during periods of heightened ecological opportunity. These results show that morphological characters often contain signal that is undetected in the molecular datasets most often used in taxonomic and phylogenetic research. This is likely because morphological variation in animals is predominantly driven by mutations occurring in non-coding regulatory sequences, rather than the molecular markers typically used to infer evolutionary relationships between taxa. Whilst the signal contained within molecular and morphological characters is conflicting, it may also be complimentary, and these dynamics should be considered in all studies aiming to test evolutionary hypotheses and reconstruct the tree of life.

2 Declaration

No portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

3 Copyright statement

i. The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the "Copyright") and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

ii. Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made only in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.

iii. The ownership of certain Copyright, patents, designs, trademarks and other intellectual property (the "Intellectual Property") and any reproductions of copyright works in the thesis, for example graphs and tables ("Reproductions"), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

iv. Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see http://documents.manchester.ac.uk/DocuInfo.aspx?DocID=24420), in any relevant Thesis restriction declarations deposited in the University Library, The UniversityLibrary's regulations (see http://www.library.manchester.ac.uk/about/ regulations/) and in The University's policy on Presentation of Theses.

4 Acknowledgements

I would like to thank my supervisors Dr. Robert Sansom and Dr. Russell Garwood for the continued support throughout the year, and for guiding me throughout all stages of the project. It has been a privilege working under their supervision. I would also like to thank Dr. Joseph Keating who had planned the initial stages of the project and gathered all total evidence datasets used, and who trained and assisted me in R programming and Bayesian phylogenetics.

5 List of tables

6 List of figures

7 Introduction:

Understanding the relationship between genomic and phenotypic evolution is a major goal for evolutionary biologists. The rate in which phenotypic features evolve can vary widely amongst lineages, often reflecting the intensity of natural selection occurring in different ecological conditions (Gavrilets and Losos, 2009; Rundell and Price, 2009; Glor, 2010). Heritable mutations reaching fixation in the genome provide the foundation for adaptive evolution, and are thus central components of evolutionary radiations. However, because the vast majority of mutations occurring in the genome are selectively neutral, some have argued that genome-wide rates of change have no effect on the tempo of phenotypic evolution (Kimura, 1968, 1983; Bromham *et al*., 2002; Davies and Savolainen, 2006). Studying correlations between rates of genomic and phenotypic evolution across the tree of life can help decipher this relationship, which will ultimately yield crucial insight into the molecular mechanisms that structure biodiversity, and allow us to better understand the macroevolutionary patterns characteristic of evolutionary radiations.

7.1 Measuring the timing and rates of evolution

Traditionally, measuring rates of evolution and estimating the timing of evolutionary transitions was largely restricted to analysing macroevolutionary patterns in the fossil record. This was until Zuckerkandl and Pauling's (1962, 1965) ground-breaking discovery that genomic variation between lineages may be used as a molecular clock. After studying mutations rates in haemoglobin genes, they were able to demonstrate that the number of amino acid differences accumulating between lineages is roughly proportional to the time since their divergence from a common ancestor. Later work would provide evidence to suggest that the molecular clock amongst different organisms may operate at widely different rates, which ultimately led to the development of relaxed molecular clocks that allow the rate of evolution within and between lineages to vary (Laird *et al*., 1969; Welch and Bromham, 2005). George Gaylord Simpson had previously highlighted the importance of rate variation in his seminal work, *Tempo and Mode in Evolution*, where he suggested that the rate of evolution between lineages is primarily driven by determinants such as population size, generation time, and natural selection (Simpson, 1944). Using a synthesis of palaeontology and Mendelian genetics, Simpson proposed that evolutionary transitions occurring deep in time were not only the result of

a fragmented fossil record, but also due to sudden bursts of evolution following shifts into novel adaptive zones (Simpson, 1944). Eldredge and Gould (1972) would later expand upon this theory to challenge Darwin's concept of phyletic gradualism, instead suggesting that lineages evolve in rare and abrupt bursts of speciation separated by bouts of evolutionary stasis. These concepts were used to explain the rapid emergence of major Metazoan phyla in the fossil record during the Cambrian period, which did not fit well with Darwin's theory of descent with slow and gradual modification. It has now been shown empirically that rates of evolution during this period were accelerated relative to the subsequent Phanerozoic periods due to a series of unique ecological, environmental, and developmental factors (Marshall, 2006; Lee *et al*., 2013).

Bursts of morphological evolution occurring during events such as these likely occur as a result of intense selection following the sudden emergence of ecological opportunity, but the neutral theory of molecular evolution predicts that genomic rates will not follow the same pattern (Gavrilets and Losos, 2009; Rundell and Price, 2009; Glor, 2010). Kimura's neutral theory proposes that because only a small fraction of the genome contains protein-coding sequences, the majority of genes are selectively neutral, and that mutations at these regions will arise predominantly due to random drift, largely independent of phenotypic evolution (Kimura, 1968, 1983). If this is true, then it would seem reasonable to expect a disconnect between rates of molecular and phenotypic evolution through time. This was demonstrated empirically in a number of studies performed throughout the latter half of the 20th century studying correlations between molecular and phenotypic evolution using genetic distances (Wallace *et al*., 1971; King and Wilson, 1975; Wilson *et al*., 1977; Cherry *et al*., 1978).

7.2 Comparing rates of molecular and morphological evolution

Omland (1997) questioned the validity of results from previous studies, suggesting that in the majority of cases, taxa that had been chosen for study were expected to exhibit highly contrasting rates of evolution *a priori*. For example, Cherry *et al.* (1978) compare differences in continuous morphological traits between frog species, with those between humans and chimpanzees. They argue that compared to frog species, the extensive morphological differentiation observed between humans and chimpanzees suggests that in these lineages, the pace of phenotypic evolution is accelerated relative to genetic rates. Omland (1997) was adamant that such studies could not be used to support the consensus that molecular and phenotypic rates of evolution are decoupled, instead suggesting that

investigations focusing on a wider range of taxa were needed. Omland (1997) decided to readdress the question by looking for correlations between molecular and morphological evolution using phylogenetic tree branch lengths proportional to character change. Omland (1997) obtained trees from the literature constructed using maximum parsimony on molecular sequence data and discrete morphological characters for a variety of animal and plant groups, and measured similarities between resultant branch lengths using Pearson's product moment and Spearman's rank correlation tests. Omland (1997) also measured root to tip path lengths for all trees in order to quantify the rates of character change, as the length of time evolution has been occurring is constant from the base of the tree to all terminal branches. He found that in nearly every case, molecular and morphological branch lengths were correlated, and that the same was found when analysing rates of evolution, going on to suggest that selective sweeps may induce the accumulation of neutral or slightly deleterious mutations as morphological characters evolve (Gillespie, 1991)

These results were later challenged by Bromham *et al*. (2002), who did not believe that such selective sweeps would be sufficient to influence genome-wide rates of evolution, especially as only a small proportion of nucleotide sites are known to be associated with morphological development (Kimura, 1968, 1983; Barrier *et al*., 2001; Carroll, 2008; Frankel *et al*., 2011; Gaunt and Paul, 2012; Wittkopp and Kalay, 2012). Instead, they suggest that rates of genetic and phenotypic evolution could only correlate indirectly through confounding variables such as population size, and that in most cases, should not be correlated at all, suspecting that Omland's (1997) results were likely caused by a failure to correct for phylogenetic non-independence. This occurs where closely related taxa are more likely to exhibit similar branch lengths due to shared ancestry, leading to clustering of data points which cannot be considered statistically independent when performing correlation tests (Felsenstein, 1985). They also suggest that Omland (1997) had failed to consider artefacts introduced by the node density effect, where root to tip pathways broken up by a larger number of nodes permit a larger number of character state changes to be inferred, whilst multiple hits along long unbroken branches remain undetected. In a parsimony framework, this will artificially reduce rates of evolution in clades possessing fewer taxa, whilst the more diverse groups will show a tendency to exhibit higher rates of evolution for both molecular and morphological characters. Finally, Bromham *et al.* (2002) were not convinced that Omland's (1997) methods had accounted ascertainment bias in phenotypic datasets. Because the majority of phenotypic datasets have been put

together in the framework of maximum parsimony, invariant and derived characters in extant taxa are often ignored, as possible tree topologies remain unaffected regardless of whether these characters are included in the analysis or not (Yeates, 1992; Bryant, 1995). This means that terminal branches in morphological trees are likely to be underestimated, resulting in artificially lower rates of evolution relative to internal branches.

Bromham *et al*. (2002) devised their own methodology to account for such artefacts. Reconstructing molecular trees with maximum likelihood, and morphological trees with maximum parsimony, the authors tested for correlation between molecular and morphological branch rates using three different statistical approaches: PARE, APPLE, and GRAPE. In the PARE analysis, Sanderson's rate smoothing was used to estimate divergence times from input trees with branch lengths in units of molecular character change. Rates were then calculated by dividing branch lengths representing morphological and molecular change by those representing time (Sanderson, 1997). The APPLE analysis was designed to account for the node density effect by comparing average branch lengths between phylogenetically independent sister clades from each phylogeny (normally two), after randomly removing lineages to produce equal numbers of taxa for comparison. Because sister taxa by definition originate from the same point in evolutionary time, the amount of character change along each branch can be considered proportional to the rate of evolution. A positive relationship between rates of molecular and morphological evolution was considered to be present in the event that one clade exhibited higher rates for both molecular and morphological branch lengths. The same principle was applied to the GRAPE analysis, where sister clades were pruned to include just one taxon for each comparison, making sure to select those of the greatest distance in the tree. The authors postulated that any correlation between molecular and morphological evolution would be easier to detect when comparing taxa separated by the greatest evolutionary distance. Using these methods, Bromham *et al.* (2002) found that across 13 vertebrate phylogenies and 3 of the original datasets from Omland's (1997) analysis, no correlation was detected between molecular and morphological evolutionary rates, concluding that Omland's (1997) findings were likely due to phylogenetic bias.

However, Davies and Savolainen (2006) have since claimed that the manner in which various taxa were removed from the APPLE and GRAPE analyses may have reduced the statistical power of correlation tests, ultimately preventing any significant association from being detected. Moreover, they point out that Sanderson's (1997) nonparametric rate

smoothing algorithm estimates divergence times based on the assumption that rates along consecutive lineages correlate, which is likely to be violated in some datasets. Davies and Savolainen (2006) reanalysed the relationship between phenotypic and genomic evolutionary rates using phylogenetic independent contrasts on 13 phylogenies reconstructed using Bayesian inference and parsimony for genetic and morphological data respectively (Felsenstein, 1985; Huelsenbeck *et al.*, 2001). Phylogenetic independent contrasts is a phylogenetic comparative method developed by Felsenstein (1985) for the specific purpose of correcting for non-independence when studying trait evolution in phylogenies. Assuming a Brownian motion model of trait evolution, weighted differences between terminal sister taxa are used to estimate relative rates of character evolution. The algorithm transforms these values into statistically independent data points by taking into account shared ancestry between taxa, which may then be used in regression analyses (Felsenstein, 1985). Using this method, they found that when combining unweighted raw plant contrasts, there was weak but significant correlation between molecular and morphological rates, whilst those standardised using phylogenetically independent contrasts were also significant, albeit slightly less so. Overall, the results revealed that when combining all datasets for plants, around 2-11% of morphological rate variation can be explained by molecular rates of change, whilst no significant correlation was detected when analysing animal groups. Davies and Savolainen (2006) speculate that this may have been due to the lower statistical power of correlation tests used to analyse the much smaller animal datasets, as for any given correlation there is, by definition, a simple monotonic relationship between P value and sample size. This is supported by the fact that when analysing rates within individual plant datasets, no significant correlation was found.

The conflicting results produced by these studies highlight the sensitivity that such analyses have to analytical methodologies and to phylogenetic artefacts produced by factors such as the node density effect and ascertainment bias. The extent that these factors may effect conclusions reached in rate correlations tests was investigated by Seligmann (2010), where datasets from Bromham *et al.* (2002) and Davies and Savolainen (2006) were statistically reanalysed. Seligmann (2010) found evidence for significant correlation between molecular and morphological rates in datasets used by Bromham *et al.* (2002), showing that the decision to remove tips in consideration of ascertainment bias had inadvertently lowered the statistical power of correlation tests such that any positive association would not be detectable. Interestingly, Seligmann also

found that the number of characters used to estimate rates of morphological evolution was directly proportional to the statistical power of correlation tests performed. A similar pattern was found to occur when analysing molecular data from both studies, where p values were found to decrease with increasing sequence length up until around 1300 characters, before rising again as number of nucleotide sites extend beyond this point.

Based on the results of these studies, it remains unclear exactly how molecular and morphological rates of evolution are associated. Each of these previous attempts to test this relationship have also relied on the use of maximum parsimony to reconstruct morphological character evolution, the accuracy of which has been questioned in some studies (Felsenstein, 1978; Wright and Hillis, 2014; O'Reilly *et al*., 2016). Parsimony estimates character evolution along a branch based on the least possible number of changes being the most likely to have occurred. This method fails to take into account multiple changes along the same branch, meaning that long unbroken branches are likely to be underestimated, whilst multiple state changes will be detected along pathways separated by a greater number of intervening nodes (Wright and Hillis, 2014). Issues such as these are known to introduce error from artefacts such as long-branch attraction, where extensive molecular or morphological character change occurring amongst distantly related lineages increases the probability of the same character developing at a given site at the same time, which can incorrectly infer them to be closely related (Felsenstein, 1978). In contrast, Bayesian methods may be better able to account for these issues, and allow parallelisms and reversals to take place as characters evolve along a branch (Wright and Hillis, 2014).

7.3 Aims and objectives

This study aims to make a revised attempt at deciphering the relationship between rates of molecular and morphological evolution. Bayesian relaxed clock methods that are able to make use of fossil age information to calibrate trees offer a powerful alternative to traditional comparative methods used for estimating rates of evolution. In the context of the present study, relaxed clock methods are preferred for a number of reasons. Firstly, phylogenetic comparative methods often rely on assumptions that may be violated in some datasets, such as in independent contrasts, where phenotypic evolution must be modelled in a Brownian motion process (Freckleton and Harvey, 2006). Moreover, the majority of these tests have been designed specifically for studies focusing on extant taxa, and rarely take into consideration evolutionary rates occurring along internal branches

constructed using discrete morphological character data. Finally, such methods are only able to provide relative rates of evolution, and offer no way of analysing rate variation through time. By studying rate patterns through time, it may be possible to identify potential geological, evolutionary, or ecological factors that could have influenced the rate of molecular and morphological evolution in a manner that may have led to their coupling or decoupling. Relaxed clock methods have recently been used to show that rates of genomic and phenotypic evolution were elevated in arthropods during the Cambrian explosion, before decelerating and plateauing throughout the subsequent Phanerozoic periods (Lee *et al.*, 2013). These results provide an explanation for the sudden diversification of metazoan morphological complexity in the fossil record during the Cambrian period, suggesting that the intensity of selection driven by ecological opportunity drove the accelerated evolution of animal morphology and housekeeping genes. Applying these methods to other interesting radiations has the potential to yield insight into the interplay between molecular and morphological evolution during periods of ecophenotypical diversification, and compare any patterns observed to those displayed in Arthropods during the Cambrian or in other groups undergoing adaptive radiations. Whilst morphological evolution will likely accelerate during these periods of intense selection, the neutral theory of molecular evolution predicts that the majority of the genome will remain unaffected, and that only those mutations providing adaptive advantages will reach fixation at a higher rate (Kimura, 1968, 1983). Thus, studying rates through time may reveal interesting evolutionary dynamics occurring during radiations that could help to explain the results of correlation tests performed.

These concepts are also of fundamental importance in the context of modern phylogenetic and taxonomic studies. Traditionally, research in these areas has been based on assessing morphological differences exhibited between taxa. However, the plethora of genomic data available in the modern era has led to a surge in the number of studies utilising nucleotide sequence information to reconstruct evolutionary relationships and test phylogenetic hypotheses (Lee and Palci, 2015). Unfortunately, genetic information is not directly observable deep in time, whilst morphological features from fossilised organisms can play a critical role in identifying evolutionary relationships that cannot be inferred from molecular data (Lee and Palci, 2015). Moreover, the majority of sequences analysed in taxonomic and phylogenetic studies are not thought to be those responsible for driving morphological variation, meaning that they may be evolving independently (Carroll, 2008; Frankel *et al*., 2011; Gaunt and Paul, 2012; Wittkopp and Kalay, 2012; Patwardhan

et al., 2014). If genomic and phenotypic evolution operate at independent rates, it would imply that there is some signal contained in morphological characters that is unobservable when focusing on molecular data alone (and *vice versa*) (Lee and Palci, 2015). This would mean that studies choosing to incorporate just one character type into their methodology may fail to uncover underlying evolutionary dynamics that would otherwise be detected when synthesising both character types into the same analysis.

This study aims to test correlations between molecular and morphological evolution and their rates under a Bayesian framework. Bayesian inference was conducted using MrBayes to reconstruct phylogenies with branch lengths proportional to the amount of evolution (average number of state changes per character), whilst BEAST2 was used to reconstruct rates of molecular and morphological evolution along time-calibrated branches. A meta-analysis approach is implemented here with the aim to ensure that the maximum possible number of animal groups are represented in the analysis. Molecular and morphological datasets were obtained from 34 recently published total-evidence studies, consisting of a wide range of taxonomic groups, and at a variety of different taxonomic levels. Rates through time were also analysed in order to observe how rates fluctuate temporally throughout the evolutionary history of animal groups. Phylogenetic artefacts introduced by ascertainment bias and rate autocorrelation are considered and accounted for in all analyses, as well as the effect of low sample sizes on correlation p values. A flow chart outlining the complete analytical procedure is given at the end of the methods in figure 1.

8 Methods

8.1 Data compilation

Molecular and discrete morphological character matrices were obtained for animal groups from 100 previously published total-evidence studies. Any dataset containing less than 10 parsimony informative characters for the molecular partition was discarded due to the lack of phylogenetic signal that would be contained in these alignments. Datasets consisting of less than three sampled genes were also excluded to account for biases operating on single genes such as incomplete lineage sorting and convergence. Datasets absent of any molecular gene partitions were also discarded based on the fact that

specifying models under Bayesian analysis requires partitions to be present. All taxa without either molecular or morphological character data were immediately removed from the analysis, which included all fossil taxa from tip-dating studies. Following taxon removal, any dataset containing less than 10 taxa was discarded, whilst datasets containing morphological partitions with less than 1.5 parsimony informative characters per taxon were also removed. In the event that matrices from two different datasets overlapped by more than 50%, the dataset published at an earlier date was excluded from the analysis. Finally, any datasets containing component partitions that would not converge when analysed independently were discarded. Based on this exclusion criteria, a total of 34 datasets were put forward for analysis (table 1). The majority of molecular datasets consist of DNA sequence data, with the exception of Mollusca (Vinter *et al*., 2017), which contains amino acid sequences. Each dataset was analysed using non-clock and time calibrated relaxed clock approaches with MrBayes and BEAST2 respectively. Morphological characters were given their own partition and treated as unordered.

8.2 Bayesian analyses: model selection

In order to accurately reconstruct character evolution, it is essential to select evolutionary models and sequence partitioning schemes that most accurately fit the data. This is most often assessed using statistical model selection software such as PartitionFinder *(*Lanfear *et al*., 2017) and JmodelTest2 (Darriba *et al*., 2012). However, as this process had already been carried out in each of the original total evidence studies, it was determined most appropriate to use the models and partitioning schemes that had been selected by the previous authors. See supplementary information (SI1 and SI2) for the nexus and XML files containing the models used.

8.3 MrBayes non-clock analysis

All non-clock analyses were performed in MrBayes v3.2.6. In the context of the current study, it is important that the molecular and morphological trees of each dataset possess identical topologies. Initial attempts at testing correlations between molecular and morphological branch lengths involved attaining previously published trees from the literature. Because these molecular and morphological trees did not always possess the same topology, it was impossible to directly compare branch lengths. To ensure topologies are identical, molecular and morphological character matrices were concatenated into nexus files and analysed using MrBayes to jointly infer tree topology

whilst using the unlink brlens=() command to unlink molecular and morphological branch lengths. This produces an output of two trees with identical topology, one with branch lengths proportional to molecular character change, and one with branch lengths proportional to morphological character change (average number of state changes per character).

8.4 Phylogenetic inference

All combined datasets were analysed using MrBayes v3.2.6, each using 2 independent runs of 1 cold and 3 heated chains, starting out at 50,000 generations with the first 25% discarded as burnin. Sampling frequency was kept at 0.0001% of the number of generations throughout all analyses, whilst temperature was set to 0.1 to promote chain mixing. Convergence was assessed using the standard deviation of split frequencies, with values below 0.01 being considered acceptable, whilst effective sample size (ESS) for parameters were assessed in Tracer v1.7.1, accepting only those with values greater than 200. Final trees were generated as 50% majority rule consensus trees (the default MrBayes output). See table 1 for all datasets analysed with the final number of generations used in each run.

Table 1: Final datasets used in MrBayes non-clock anlayses

Taxonomic group (common names)	Taxonomic Rank	Authors	Number of Terminals	Sites	Molecular sequence length	Morphological characters	Number of generations
Abrotrichini (South American rodents)	Tribe	Teta *et al*., 2017	24	cytb, DMP1, IRBP, Adh, bfbg, Ins	4441	96	25,000,000
Actinopterygii (ray-finned fishes)	Class	Giles *et al*., 2017	10	12 nuclear loci	10674	265	50,000
Apinae (bees)	Subfamily	Cameron and Mardulyn, 2001	10	Cytb, 16s, 28, opsin	2442	95	10,000,000
Araneomorphae (spiders)	Superfamily	Wood *et al*., 2012b	32	28S, 18S, H3, COI	5185	117	10,000,000
Arthropods	Phylum	Lee *et al*., 2013	53	62 nuclear loci	22131	395	50,000,000
Caviidae (South American rodents)	Family	Pérez and Pol, 2012	11	Ghr, Tth, 12s, cyb	4014	96	25,000,000
Cephalotes (turtle ants)	Genus	Price *et al*., 2016	59	LR, EF1aF2, WG, COI, COII,Cytb	3875	98	50,000,000
Cetaceans (whales, dolphins)	Infraorder	Geisler *et al*., 2011	29	16 mitochondial, 69 nuclear loci	59811	304	25,000,000

Taxonomic group (common names)	Taxonomic Rank	Authors	Number of Terminals	Sites	Molecular sequence length	Morphological characters	Number of generations
Chiroptera (bats)	Order	Dávalos *et al*., 2014	85	7 nuclear, 3 mitochondrial loci	9584	278	50,000,000
Coccomorpha (scale insects)	Infraorder	Vea and Grimaldi, 2016	48	18S, 28S, EF1a1, EF1a2, EF1a3	3731	174	50,000,000
Cynipidae (gall wasps)	Family	Ronquist *et al*., 2015	97	COI, EF1aF1, 28S, EF1aF2, LWRh,	4975	239	20,000,000
Fabriciidae (polychaete worms)	Family	Huang *et al*., 2011	26	18S, D1-28S, H3	3107	47	25,000,000
Hamadryas (cracker butterflies)	Genus	Garzón-Orduña *et al*., 2013	43	COI, RpS5, CAD, WG, EF1a	4576	92	25,000,000
Hexactinellida (sponges)	Class	Dohrmann *et al*., 2017	56	16S, 18S, 28S, COI	4806	107	20,000,000
Hydrophilinae (water scavenger beetles)	Subfamily	Short *et al*., 2017	19	COI , Ark, 18S, 28S	4294	42	20,000,000
Hydroptilidae (caddisflies)	Family	Santos *et al*., 2016	62	CAD, COI, 28S, EF-1α, H3	2451	114	50,000,000
Hymenoptera (ants, sawflies, wasps ect.)	Order	Ronquist *et al*., 2012	59	12S, 16S, 18S, 28S, CO1, EF1a-F1, EF1a-F2	4449	353	50,000,000
Lemuroidea (lemurs)	Superfamily	Herrera and Dávalos, 2016	57	cytb, PAST, ADORA, CNR, RAG1, RAG2	5767	369	25,000,000
Malacostraca (crabs, lobsters, shrimp etc.)	Class	Jenner *et al*., 2009	15	18S, 28S, COI, 16S	8316	177	10,000,000
Mammalia (mammals)	Class	Lee, 2016	46	27 nuclear loci	36 860	4541	50,0000
Mollusca (molluscs)	Phylum	Vinther *et al*., 2017	37	EEF1a1, ALDOA, CAT, MAT1A, ATP5F1B, TPI1, PFKM	2030	102	50,0000
Mysticeti (baleen whales)	Parvorder	Marx and Fordyce, 2015	12	45 nuclear, 12 mitochondrial loci	37285	272	25,000,000
Nephilidae (spiders)	Subfamily	Kuntner *et al*., 2013	43	12S, 16S, COX1, COX13, 28S, H3, 18S	4197	231	25,000,000
Opiliones (harvestmen)	Order	Garwood *et al*., 2014	40	18S, 28S, 16S, COI, H3	3885	123	50,000,000
Ostariophysi (fish)	Superorder	Giles *et al*., 2017	10	9 nuclear loci	7173	128	25,000,000

Taxonomic group (common names)	Taxonomic Rank	Authors	Number of Terminals	Sites	Molecular sequence length	Morphological characters	Number of generations
Osteoglossiformes (fish)	Order	Lavoué, 2016	12	11 nuclear loci, 1 mitogenome	20935	84	50,000,000
Serpentes (snakes)	Suborder	Harrington and Reeder, 2017	37	44 nuclear loci	33717	670	30,000,000
Sphenisciformes (penguins)	Order	Gavryushkina *et al*., 2017	19	RAG1, cytb, COI, 12s, 16s	8145	252	150,000,000
Squamata (lizards, snakes ect.)	Order	Wiens *et al.*, 2010	45	22 nuclear loci	15794	363	50,000,000
Stygnopsidae (harvestmen)	Family	Cruz-López and Francke, 2017	44	28S, 16S, COI	4836	72	50,000,000
Tetraodontiformes (fish)	Order	Arcila *et al*., 2015	20	15 nuclear, 1 mitochondrial loci	14648	210	20,000,000
Theridiidae (cobweb spiders)	Family	Agnarsson, 2006	36	H3, 18S, 28S, 16S, COI	2490	215	50,000,000
Tribelocephalinae (assassin bugs)	Subfamily	Forthman and Weirauch, 2017	56	162, 182, 28s, COI, wg	4055	199	20,000,000

8.5 BEAST2 relaxed clock analysis

BEAST2 v2.5.1 was used for the Bayesian relaxed clock analyses to estimate rates of molecular and morphological evolution, using the same 34 total evidence datasets used in the non-clock analysis. BEAST2 is advantageous for this because divergence times, and rates of evolution, can be simultaneously estimated for both molecular and morphological characters when incorporating fossil calibration information into the analysis. By attaining branch lengths as both units of time and as units of evolutionary rate, it is not only possible to investigate correlations between rates of molecular and morphological evolution, but also to study patterns in rate variation through time. This was done by plotting the midpoint of branch lengths as units of time against branch lengths proportional to evolutionary rate, a method that was recently used to visualise rates of phenotypic and genomic evolution in arthropods during the Cambrian explosion and subsequent Phanerozoic periods (see section 8.13) (Lee *et al*., 2013).

Third codon positions were removed from all sequences before estimating rates. Initial investigations that included third codon positions gave rate estimates several orders of magnitude higher than those attained by Lee *et al.* (2013) for arthropods. It was suspected that saturation at third codon positions was generating noise in the data in such a way that analysing any patterns in rate fluctuations through time would be difficult. Because third

codon positions are usually the least functionally constrained, they are often free to evolve independently from selection, displaying accelerated substitutions rates relative to those of first and second codon positions (Bofkin and Goldman, 2007). Removing third codon positions therefore, would allow correlations to be tested using primarily non-synonymous sites directly linked to phenotypic evolution, whilst reducing the amount of noise generated in the data. This procedure was also carried out by Lee *et al*. (2013) when analysing arthropod rates, as well as Halliday *et al*. (2019) when analysing mammalian rates, allowing results here to compared to those of previous studies. Polymorphic morphological characters were handled as ambiguous states rather than ignored (the BEAST2 default). However, a major limitation of BEAST2 is that without character state labels (a list of morphological character descriptions with the number of states for each character), invariant characters cannot be included in the analysis. Character state labels were not available for the majority of datasets, meaning that invariant characters needed to be removed from all analyses to keep methods consistent across all datasets. The number of independent runs used for each dataset was determined based on the approach from each of the original total-evidence studies, whilst molecular and morphological characters were each assigned an independent clock rate, with a birth-death tree prior. The Yule (Yule, 1925) or birth-death tree priors (Kendall, 1948; Nee *et al*., 1994; Stadler, 2013) in BEAST2 are most suitable for use when studying relationships between species, rather than coalescent priors, which are most often used when analysing intraspecific relationships within populations (Kingman, 1982; Drummond *et al*., 2005). In order to enable direct comparison of results here to those of Lee *et al*. (2013), the birth death prior was selected, which assigns species with a constant rate of extinction and speciation. Similarly, the uncorrelated lognormal relaxed clock was used in each analysis to allow comparison to previous results, but also because implementing any of the autocorrelated models provided by BEAST2 may result in non-independent branch rates, which would be unsuitable for correlation testing. A single chain was used for each run, with random starting trees and random number seeds generated prior. Assessment of log files was conducted in Tracer v1.7.1, accepting combined runs with ESS scores above 200 for all parameters. Tree files were combined using LogCombiner v2.5.1.0, before being summarised as common ancestor (CA) trees using TreeAnnotator v.2.5.1.0. Initially, maximum clade credibility (MCC) trees with mean node heights were used, but this resulted in some trees possessing negative branch lengths that would be unsuitable for further analysis, as the true number of character state changes occurring in nature cannot be negative (even if a character reverts to its ancestral state, the number of changes

occurring will always be positive). MCC trees are often generated by taking the average node height across all trees in the sample containing a given clade. If a particular clade occurs at low frequency, it may not occur in the trees containing the ancestral clade. This can cause mean or median node values to be higher than mean or median values of the parent node, ultimately resulting in negative branch lengths. The CA tree avoids this issue, by assigning every clade an age using the mean of all posterior trees (Heled and Bouckaert, 2013).

8.6 Fossil calibration

Fossil calibration information was taken from the same total-evidence analyses from which the molecular and morphological datasets were attained. Of the 34 datasets available, only 18 of these were taken from studies estimating species divergence times. Of these 18 studies, 6 use both tip dating and node dating approaches, 6 incorporate only node calibration information into the analysis, whilst 6 focus solely on a tip dating approach. Each approach has advantages and disadvantages, but it was determined that a node calibration approach would be most suitable for the relaxed clock analyses for a number of reasons, discussed below (Arcila *et al.*, 2015; O'Reilly *et al*., 2015; Donoghue and Yang, 2016). To keep methods consistent across all datasets analysed, a decision to use either node dating or tip dating was required. Initial attempts were conducted using BEAST2 and MrBayes to repeat the tip dating analyses as performed in the original studies, which frequently resulted in poor ESS scores for the majority of parameters when assessing log marginal likelihoods in Tracer, despite rigorous attempts at tuning MCMC parameters to optimal settings. Additionally, the estimated time for completion of the majority of tip dating analyses was frequently unfeasible given the time available to the present study (some datasets exceeding 3000 hours when attempted in MrBayes), whilst node-dating analyses finished with more manageable convergence issues and in comparatively shorter periods of time. Finally, as well as being less computationally demanding, a node dating approach would allow the direct comparison of patterns of rate variation through time to those exhibited by arthropods during the Cambrian explosion and subsequent Phanerozoic periods as in Lee *et al.* (2013). To date, this is the most significant study to have measured evolutionary rates deep in time, and so the protocol used by the authors is implemented here. Thus, the relaxed clock analysis focused solely on the 12 datasets for which adequate node calibration information was available. Unfortunately, 2 of the datasets did not possess sufficient taxa for correlation testing (see

section 8.12) (Caviidae n = 11 and Osteoglossiformes n = 12), and so it was decided that they would not be used in the final relaxed clock analysis.

8.7 Minimum and maximum age constraints

The approach for assigning appropriate probability densities and age constraints to calibration points was based on the rationale used in Lee *et al*. (2013). It is often not possible to empirically assign parameters for calibration densities such as the offset lognormal or exponential distributions, meaning values must be selected arbitrarily. Studies have shown that divergence time estimates using these complex calibration densities can be highly sensitive to parameter choice, resulting in conflicting values differing by hundreds of millions of years (Ho and Phillips, 2009; Warnock *et al*., 2012). Using such calibration densities may also run the risk of over estimating divergence dates due to the age of the peak prior probability deviating beyond the minimum bound set by the fossil (Warnock *et al*., 2012). Lee *et al*. (2013) reason that these methods are unsuitable when estimating rates deep in time, as any increases in divergence time estimates will likely compress more internal branches in the tree and artificially elevate rates of evolution deeper in time. Thus, all calibration points used here are assigned uniform distributions with simple minima only. This allows clades to be calibrated to times more closely resembling the age of the first fossil, providing lineages with a sufficient amount of time to evolve differences (Lee *et al*., 2013).

8.8 Stem and crown calibrations

Another advantage with using BEAST2 for divergence dating analyses is that node calibrations have the option of being treated as either stem or crown. Crown calibrations incorporate fossil ages into the last common ancestor of the calibrated clade, whilst stem calibrations apply this information into the stem lineage leading up to the calibrated clade. Because the fossil record only represents an infinitesimally small fraction of the number of species to have ever lived, it is difficult to be sure that a fossil displays features representative of the crown group of the clade rather than stem lineage leading up to it. If there is uncertainty as to whether a fossil should be allocated to the basal node of a clade, then there is a high chance that a more accurate placement is at the stem. Thus, the conservative approach is to use stem calibrations over crown calibrations where there is any uncertainty over the placement of the fossil (Forest, 2009; Wang and Mao, 2016). Moreover, in contrast to calibrating at the crown node, stem calibrations do not require

constraining any particular sister group to the calibrated clade, which instead can be estimated as part of the analysis (Lee *et al*., 2013). For these reasons, all calibrations are treated as stem, unless explicitly stated otherwise in the original source.

8.9 Prior parameters

As the original total-evidence studies rarely included information on prior selection in BEAST or BEAST2 analyses, and the fact that some of these node-dating studies were performed in MrBayes, and thus, had used priors currently unavailable in BEAST2, MCMC parameters were selected based on those used in Lee *et al.* (2013), the majority of which were simply left as default values. Where any prior information was available from the original authors, this was implemented in the analysis. For each dataset, the ucld mean and stdev parameters for molecular and morphological partitions were assigned lognormal and exponential distributions respectively, which do not allow negative values, and weight values closer to 0. Naturally, default prior settings were often suboptimal when applied to the majority of datasets, which resulted in convergence issues. ESS values were assessed in Tracer v1.7.1, and alternate prior distributions were chosen for parameters where mixing was poor. To assess the joint effect of priors on the posterior, each analysis was repeated using "empty" alignments (no molecular or morphological data) to sample from the prior only. In each dataset analysed, the sample from the posterior consistently overlapped with the sample from the prior, but was much more concentrated, suggesting that priors were reasonable and that the data were informing the posterior (Nascimento *et al*., 2017). Moreover, the topologies of timetrees produced by the empty data appeared random and highly conflicting with topologies generated by the data, with very low posterior support for all nodes and entirely different branch rate patterns, indicating that the data were generating the results. For detailed information on the final priors chosen for each analysis, see supplementary information (SI2) for XML files used, as well as table 2 for information on the number of generations used for each run.

Table 2: Final datasets used in BEAST2 relaxed clock analyses.

Taxonomic group (common names)	Taxonomic Rank	Authors	Number of Terminals	Sites	Molecular sequence length	Morphological characters	Generations	Burnin
Araneomorphae (spiders)	Infraorder	Wood *et al*., 2012b	32	28S, 18S, H3, COI	5185	117	6x30m	10%
Cephalotes (turtle ants)	Genus	Price *et al*., 2016	59	LR, EF1aF2, WG, COI, COII, Cytb	3875	131	4x50m	10%
Coccomorpha (scale insects)	Infraorder	Vea and Grimaldi, 2016	48	18S, 28S, EF1a1, EF1a2, EF1a3	3731	174	4x200m	30%
Hydroptilidae (caddisflies)	Family	Santos *et al*., 2016	62	CAD, COI, 28S, EF1, HEX	2451	112	4x50m	10%
Hymenoptera (ants, sawflies, wasps ect.)	Order	Ronquist *et al*., 2012	59	12S, 16S, 18S, 28S, CO1,EF1aF1, EF1aF2	4449	253	4x60m	10%
Opiliones (harvestmen)	Order	Garwood *et al*., 2014	40	18S, 28S, 16S, COI, H3	3885	149	4x30m	10%
Mammalia (mammals)	Class	Lee, 2016	46	27 nuclear loci	36 860	4541	4x50m	10%
Cetaceans (whales, dolphins)	Infraorder	Geisler *et al*., 2011	29	16 mitochondial, 69 nuclear loci	59811	304	4x50m	10%
Tetraodontiformes (fish)	Order	Arcila *et al*., 2015	20	1 mitochnodrial and 15 nuclear loci	14648	210	4x50m	10%
Serpentes (snakes)	Suborder	Harrington and Reeder, 2017	37	44 nuclear loci	33717	670	4x50m	10%

8.10 Ascertainment bias

Invariant and derived characters in extant taxa (autapomorphies) are often underrepresented in phenotypic datasets, which have traditionally been constructed in the framework of maximum parsimony. Such characters have no effect on tree topologies when estimated using parsimony, and are thus, typically excluded from these analyses. However, autapomorphies are important when estimating branch lengths and rates. Missing character data in extant taxa may result in an underestimation of terminal branch lengths, artificially lowering rates of evolution relative to internal branches. This ascertainment bias is unlikely to affect rates along internal branches, as character state changes occurring amongst extinct taxa are informative when using parsimony, and are thus, sampled to a much greater degree of accuracy.

Specialised implementations of the Lewis (2001) Mk model have been adapted in consideration for this bias, with the Mki model correcting for only including parsimony informative characters, and the Mkv model for only variable characters. The Mkv or Mki model were used as appropriate in each non-clock analysis, based on the model choice made by authors in the original total evidence studies. See supplementary information (SI1) for MrBayes nexus files containing models used. Unfortunately, the Mki model is unavailable in BEAST2, and so the Mkv model was used for all relaxed clock analyses. It must be noted however, that correcting for ascertainment bias does not necessarily restore character state changes at tips (Lewis, 2001). Thus, as a precautionary measure, all terminal branches were excluded when performing correlation tests, and terminal branches for morphology were not considered when analysing rates through time. It has been suggested that the lack of correlation found between molecular and morphological rates in some previous studies is largely due to the effect of excluding the tips, which has been often justified using the same rationale implemented. This is thought to have resulted in a loss of statistical power through the reduction of sample size in such a way that any significant association becomes difficult to detect (Seligmann, 2010). To investigate this as a factor, tips were also included in supplementary correlation tests for comparison to the focal analysis.

8.11 Phylogenetic autocorrelation

Another consideration that needed to be made was that branch lengths and rates may autocorrelate in some datasets. Autocorrelation occurs in phylogenetic trees where branches at a closer proximity to one another display more similar values than those located at greater distances, as character differences accumulate between lineages in proportion to evolutionary time (Felsenstein, 1985). Data points from any autocorrelated trees are unreliable for correlation testing, as they may cluster in a manner that could spuriously suggest that molecular and morphological rates or branch lengths are correlated (statistical non-independence) (Felsenstein, 1985).

Although there is also a possibility that autocorrelated trees could also spuriously suggest non-correlation, it is more likely that these trees will enhance correlation between two traits. This was the main point of concern of Felsenstein (1985) who was among the first to highlight this issue, leading to the development of a plethora of phylogenetic comparative methods designed to account for statistical non-independence when comparing biological trait data across taxa in phylogenies. Examples of such methods

include phylogenetically independent contrasts and phylogenetic generalised least squares, which have often been implemented in previous studies testing rate correlations in molecular and morphological evolution (Bromham *et al.*, 2002; Davies and Savolainen, 2006).

However, the extent that rates of molecular and morphological evolution are inherited is likely to be highly variable amongst different lineages. This is because in some groups, such as mammals, heritable life history traits such as body size and generation time are known to correlate negatively with DNA sequence evolution rate, whilst some studies on invertebrates have found no such link (Thomas *et al*., 2006; Lanfear *et al*., 2007; Bromham, 2011). Moreover, rate autocorrelation will not necessarily be present in trees constructed using the uncorrelated lognormal clock in BEAST2, where branch rates are drawn independently from a lognormal distribution (Drummond *et al*., 2006). Therefore, it is more appropriate to test for autocorrelation in each phylogeny rather than assume that it is present or absent. By isolating trees showing autocorrelation, it is then possible to identify where this may be causing type 1 error (false positives) in any correlation tests.

Unfortunately, the majority of tests for phylogenetic non-independence apply to continuous traits observed in the tips rather than internal branches. This required alternate ways of identifying autocorrelation to be developed. Here, two different but complimentary approaches are used. For the first, the pairwise differences between branch lengths were compared with the number of intervening nodes between each branch in the tree. Lower values for branch length differences would imply that they are more similar (a value of 0 would mean they are the same length), whilst branches with very different lengths would display much higher values. If branch lengths were autocorrelated, then branches separated by more nodes would show more variation, which could then be assessed graphically. A Mantel test was also used to compare the pairwise difference matrices with node distance matrices in order to assess correlation between phylogenetic distance and branch length similarity. To test the effectiveness of this method, trees were drawn in Mesquite 3.6 (Maddison and Maddison, 2018), with branch lengths deliberately designed to autocorrelate throughout the tree (for example, branch lengths increasing twofold ascending from branch to branch). Unfortunately, the Mantel test on the autocorrelated data did not detect any significant relationship between branch difference and node distance, implying that the method was not a statistically sound approach for detecting autocorrelation. The second method for detecting autocorrelation was an adaptation of the method used by Ronquist *et al*., (2012). This test compares the

distribution of parent-daughter branch ratios to the branch ratios of all other branches in the tree using the Kolmogorov-Smirnov test for equality of distributions. Any significant difference between the parent-daughter and random branch ratios in the tree would imply that there is some autocorrelation present, and thus, this could not be ruled out as a factor when detecting correlations between molecular and morphological branch lengths and rates. This method showed that parent-daughter branch ratios were significantly different from all other branch ratios in the trees designed for autocorrelation testing, and was thus employed here. Terminal branches were not included in the autocorrelation tests, using the same rationale implemented in the molecular and morphological branch correlation testing.

8.12 Correlation testing: branch lengths and rates

For trees estimated from both non-clock and relaxed clock analyses, the distribution of branch lengths was tested using a one sample Kolmogorov-Smirnov test, and by graphical assessment. The data distribution were non-normal, and so molecular and morphological branch lengths were compared using the Spearman's correlation test for non-parametric data. Branches for all non-clock analyses were first standardised by dividing each branch by the tree height. This was not necessary for relaxed clock trees, as character change was already proportional to time. Correlation tests were performed using only internal branches for both molecular and morphological data, whilst datasets containing fewer than 20 taxa were not analysed in order to discount results being due to low sample sizes reducing statistical power. Outgroup taxa were excluded from all correlation analyses for a number of reasons. Firstly, outgroups in phylogenies are almost always long branches due to incomplete taxon sampling relative to the ingroup, and are thus not an accurate representation of the true evolutionary rate of the outgroup taxa. Because including outgroup taxa would sometimes mean the difference between a significant and a non-significant result, it was decided that excluding outgroup taxa would avoid altering the outcome of correlation tests using inaccurate data. Moreover, this procedure allows the methodology to more closely follow the protocol used by Lee *et al.,* (2013), who opted to exclusively analyse ingroup branches to account for rooting artefacts, whilst removing outgroups will also allow the present study to focus on only those taxa explicitly contained within the clades of interest. In addition to testing correlations within datasets, a combined analysis was performed to test for overall correlation between molecular and morphological branch lengths and their rates in animals using linear mixed effects

models. Although the Fisher's combined probability test used by Seligmann (2010) can offer an effective way of summarising p values from multiple independent test, in the context of the current study, mixed-effect models offer a more statistically robust alternative due to the manner in which variable variances can be accounted for between datasets analysed. The relationship between molecular and morphological branch lengths was tested with each taxonomic group representing a random effect, as well as branch number as a nested random effect, performed in R v3.5.1 (R Core Team 2018) using the package nlme (Pinheiro *et al.*, 2018). A likelihood ratio test was used to compare the quality of the fully parametrised mixed effects models with several less parameterised models.

8.13 Rates through time

A separate analysis focusing on temporal variation in branch rates was also performed. This was done using the same data generated from the time calibrated relaxed clock analyses, by plotting the rate of each molecular and morphological branch against the age of the same branch. To calculate branch age, the midpoint of each branch in units of time (millions of years (My)) was subtracted from the tree height (My). The BEAST2 tree output assigns the root node an age of 0 My, and the tips the same age as the tree height, so to calculate the branch ages in units of time from the present, they must be subtracted from the tree height. For example, if the age of tips are 100 My, subtracting this from the tree height (100 My) would give 0 My from present, whilst subtracting the root node age (0 My) from the tree height would give 100 My from present.

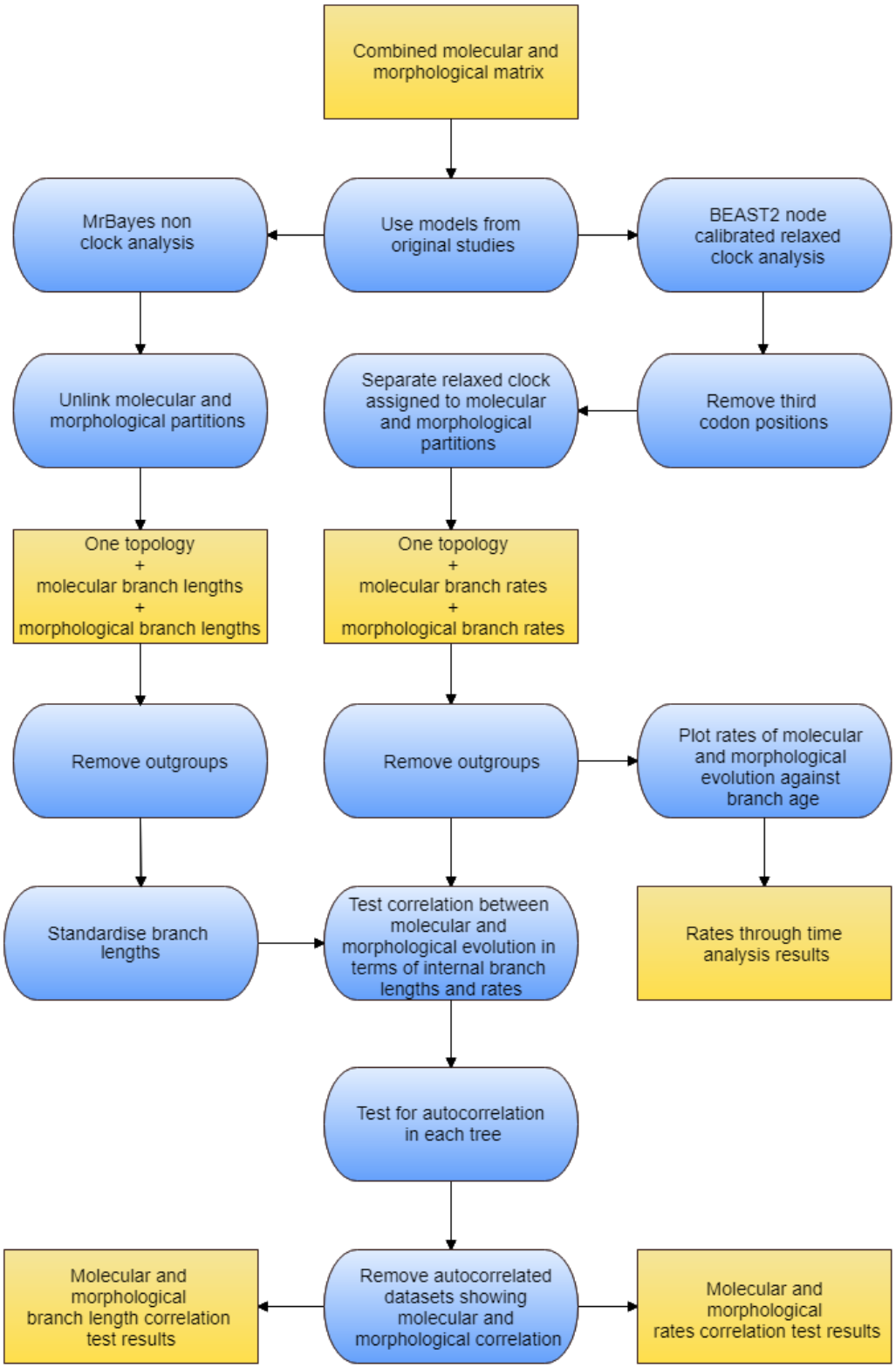

Figure 1: Flow chart summarising the steps taken in each analysis. See the methods section under the relevant headings for detailed information and justification for each step taken.

9 Results

9.1 MrBayes non-clock analyses

The final data sample comprised 10 datasets with both molecular and morphological branch lengths for at least 20 taxa. The majority of datasets analysed in the MrBayes non-clock analysis did not converge using standard settings. In many cases, convergence was achieved by increasing the number of generations of each run. Despite these measures, many datasets would not converge even after increasing the number of generations to 100 million, with extremely poor ESS scores (in many cases remaining below 20). Increasing generation time above 100 million was not feasible with many datasets, which were estimated to complete in excess of 1000 hours, whilst changing MCMC temperature from 0.1 to 0.2 failed to facilitate chain mixing. In some cases, even when starting out runs at 50,000 generations, the estimated time for completion would exceed 3000 hours. Whilst many of these molecular and morphological character matrices have been used without issue for phylogenetic inference in previous studies, combining matrices to infer tree topology whilst unlinking branch lengths for molecular and morphological partitions may be too computationally demanding when analysing some datasets. It is possible that in some cases the signal contained in the molecular and morphological data conflicts to such an extent that achieving convergence using this method would be near impossible, regardless of the computational resources available. See supplementary information (SI3) for final trees used in the analysis.

9.2 BEAST2 relaxed-clock analyses

The final data sample for molecular and morphological rates analyses comprised 6 datasets of at least 20 taxa. Some datasets did not converge after initial runs, in which case convergence was generally achieved by increasing the number of generations for each run, increasing the sampling frequency, or by combining more independent runs. In some cases, convergence was only achieved after tuning operator parameters in accordance to the performance suggestions printed following completion of initial runs. At the end of each BEAST2 run, a list of performance suggestions are given for the operators of parameters that are not mixing well. Adjusting operator values according to these suggestions allow those operators to explore state space more efficiently, which may facilitate mixing. However, for some datasets, achieving convergence was especially

difficult. The larger datasets containing molecular sequence data in excess of 14,000 nucleotides consistently failed to reach stationarity in log likelihoods, and any further efforts did not seem practical given the frame of time available to the present study. See supplementary information (SI4) for final trees used in the analysis.

9.3 Autocorrelation tests

Of the ten non-clock molecular and morphological sets of trees, significant rate autocorrelation was detected in the branch lengths of the Hymenoptera molecular tree, and the Araneomorph molecular tree (table 3). Significant autocorrelation was more prevalent in the relaxed clock trees, being detected in rates of molecular evolution in the Hydroptilidae and *Cephalotes* trees and the rates of morphological evolution in the branches of the Araneomorph tree (table 4).

Table 3: MrBayes non-clock analysis Kolmogorov-Smirnov test for equality of distributions between parent-daughter log branch length ratios and all other log branch length ratios in the tree. Significant scores highlighted in bold.

Taxonomic group	Molecular parent-daughter ratio vs random ratio		Morphological parent-daughter ratio vs random ratio	
	D statistic	P value	D statistic	P value
Abrotrichini	0.230	0.456	0.314	0.139
Cephalotes	0.135	0.297	0.093	0.756
Fabriciidae	0.269	0.145	0.246	0.223
Hamadryas	0.150	0.450	0.188	0.192
Hexactinellidae	0.191	0.087	0.112	0.651
Hymenoptera	0.224	**0.016**	0.188	0.068
Mollusca	0.122	0.888	0.197	0.362
Opiliones	0.222	0.112	0.092	0.965
Araneomorphae	0.274	**0.046**	0.198	0.279
Theridiidae	0.172	0.356	0.163	0.419

Table 4: BEAST2 relaxed clock analysis Kolmogorov-Smirnov test for equality of distributions between parent-daughter log rate ratios and all other log rate ratios in the tree. Significant scores highlighted in bold.

Taxonomic group	Molecular parent-daughter ratio vs random ratio		Morphological parent-daughter ratio vs random ratio	
	D statistic	P value	D statistic	P value
Araneomorphae	0.097	0.903	0.258	**0.021**
Cephalotes	0.245	**0.004**	0.144	0.231
Coccomorpha	0.157	306	0.131	0.53
Hydroptilidae	0.201	**0.02**	0.118	0.404
Hymenoptera	0.163	0.155	0.201	**0.041**
Opiliones	0.166	0.396	0.121	0.784

9.4 Non-clock branch length correlation tests

For the focal analysis excluding tips, there was no evidence for significant correlation between molecular and morphological branch lengths in 7 of the 10 sets of trees generated in the MrBayes non-clock analysis (table 5). Of the datasets showing correlation between molecular and morphological branch lengths, Hymenoptera also exhibit significant rate autocorrelation in the molecular tree analysed. Because of this, correlation between molecular and morphological branch lengths in this dataset cannot be discounted as being an artefact of rate autocorrelation. Spearman's Rho statistic scores for the *Hamadryas* and the *Cephalotes* indicate moderate positive correlation with no evidence for rate autocorrelation in either molecular or morphological trees. This leaves a total of 7 of 9 datasets showing no evidence for correlation between molecular and morphological branch lengths, suggesting that in the majority of animal groups studied, molecular and morphological evolution are not correlated (figure 2).

When including tips, 5 of the 10 datasets show significant correlation between molecular and morphological branch lengths (SI5 supplementary table 2). However, three of these datasets were also found to autocorrelate when including tips, meaning that this cannot be discounted as a causative factor (SI5 supplementary table 1). Excluding the 3 autocorrelated datasets, 5 of 7 groups show no evidence for correlation between molecular and morphological branch lengths. Thus, in the majority of cases, molecular and morphological evolution show no significant association regardless of whether tips are included or excluded.

The most parametrised linear mixed effects model (lme4) was found to provide the best fit to the data (table 6) and found no significant association between molecular and morphological branch lengths across all datasets analysed (ANOVA p = 0.9817). Hymenoptera were not included in mixed effects models given their autocorrelation in molecular branches.

Table 5: Taxonomic groups analysed in MrBayes non-clock analysis with tips excluded to account for ascertainment bias. Spearman's Rho statistics and P values from molecular vs morphological rates correlation test for each group shown, as well as results from the autocorrelation tests. Significant correlation highlighted in bold. Autocorrelation was considered present when detected in either a molecular or a morphological tree from a given dataset.

Taxonomic group	Autocorrelation of either molecular or morphological branch lengths (p value)	Correlation between molecular and morphological branch lengths (Spearman's Rho)	P value
Abrotrichini	Absent	0.014	0.960
Cephalotes	Absent	0.251	0.061
Fabriciidae	Absent	**0.505**	**0.024**
Hamadryas	Absent	**0.553**	**0.000**
Hexactinellida	Absent	0.208	0.164
Hymenoptera	Present (molecular p = 0.016)	**0.375**	**0.006**
Molluscs	Absent	-0.104	0.626
Opiliones	Absent	0.245	0.167
Araneomorphae	Present (molecular p = 0.046)	-0.108	0.572
Theridiidae	Absent	0.190	0.295

Table 6: Mr Bayes non-clock analysis likelihood ratio test comparing the relative quality of four linear mixed effects models measuring the relationship between internal molecular and morphological branch lengths in 9 animal phylogenies. With lme1 taking into account taxonomic group as a random effect (lme(BL~Type, random=~1|Group)), lme2 taking into account branch number as a random effect (lme(BL~Type,random=~1|Branch)), lme3 taking into account taxonomic group as a random effect with branch number as a nested random effect (lme(BL~Type, random=~1|Group/Branch)), and lme4 taking into account the same random effects as lme3 but whilst allowing variable variances between datasets (lme(BL~Type, random=~1|Group/Branch, weights=varIdent(form=~1|Group)). Where BL = branch length, type = molecular and morphological partitions, Group = taxonomic group, and Branch = branch pairing.

Model	df	AIC	Log likelihood	Test	Likelihood Ratio	p-value
lme1	4	-1071.956	539.9779			
lme2	4	-1065.603	536.8015			
lme3	5	-1074.494	542.247	2 vs 3	10.89091	0.001
lme4	13	-1089.168	557.5841	3 vs 4	30.67427	< 0.000

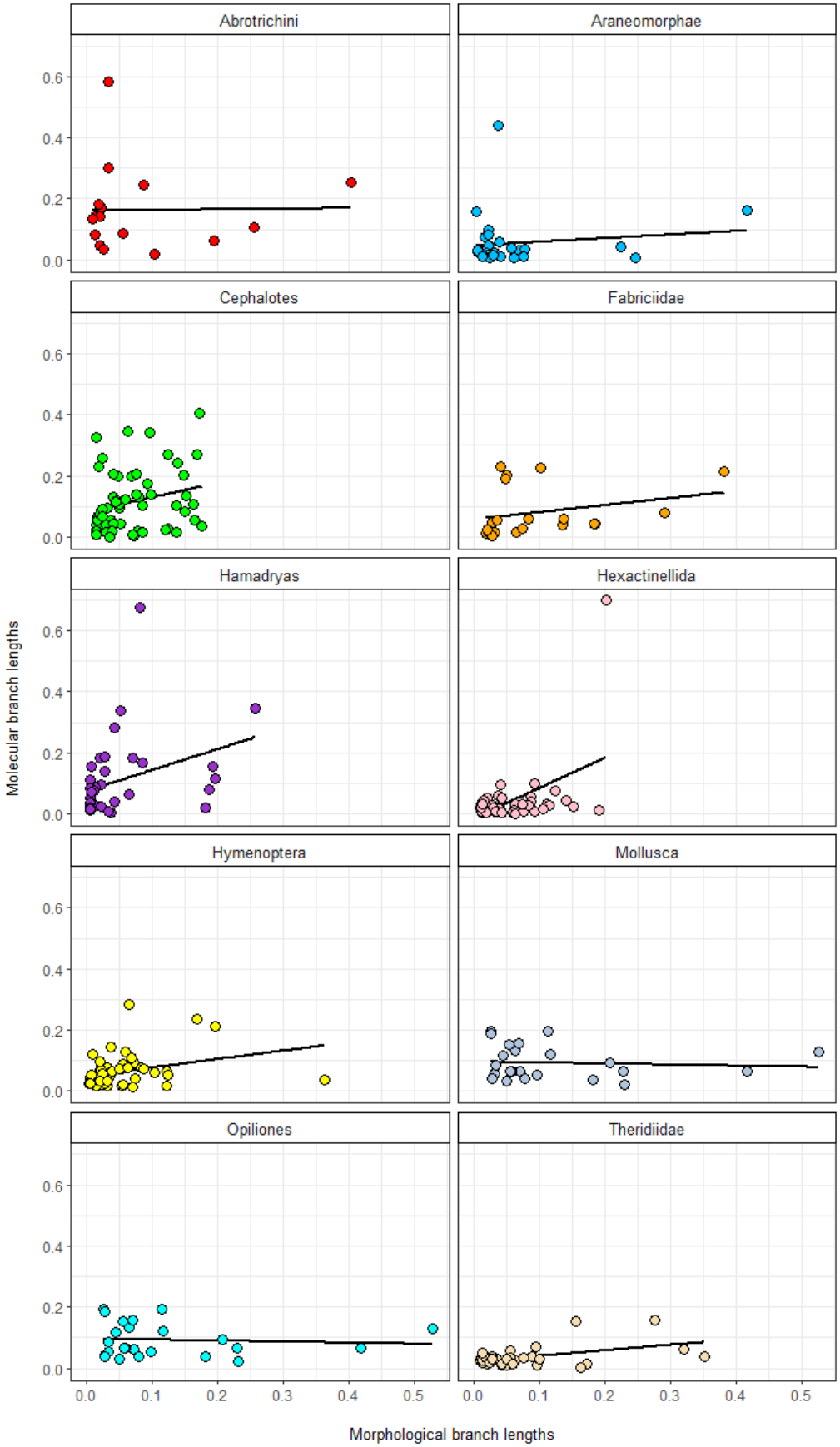

Figure 2: Non-clock molecular vs morphological branch lengths in 10 animal groups. Branch lengths estimated using MrBayes. Branch lengths represent average number of state changes per character (number of character state changes divided by sequence length). Trend lines (linear correlation) are for visualisation only.

9.5 Relaxed-clock rates correlation tests

Of the six datasets analysed in the relaxed clock analysis, only the *Cephalotes* show significant correlation between molecular and morphological rates of evolution (figure 3 and table 7). However, significant rate autocorrelation was also detected in the *Cephalotes* dataset, meaning type 1 error cannot be ruled out. As such, 5 out of 5 datasets without autocorrelation show no evidence for significant correlation between molecular and morphological rates of evolution. Although there is no significant correlation, when assessing the data graphically the pattern across each group appears to be that there is an inverse relationship between molecular and morphological rates of evolution. Increases in rates of morphological evolution are often met with decreases in molecular evolution, and there does not appear to be linear relationship in any of the datasets (figure 3). The most parametrised linear mixed effects model (lme8) was found to provide the best fit to the data (table 8) and found a significant association between molecular and morphological branch lengths across all datasets analysed (ANOVA $p = 0.0397$), whilst all other models were non-significant. The relationship is inverse, with a correlation of fixed effects score of -0.133. This means that as molecular rates increase, morphological rates are expected to decrease. *Cephalotes* were not included in mixed effects models given their autocorrelation in molecular rates. When including tips, results remain similar, the only difference being that Araneomorphae show significant correlation between molecular and morphological rates. Despite autocorrelation being present in 2 datasets, 4 of 6 datasets show no evidence for significant correlation between molecular and morphological rates of evolution when tips are included (SI5 supplementary tables 3 and 4).

Table 7: Taxonomic groups analysed in BEAST2 relaxed-clock analyses excluding tips to account for ascertainment bias. Spearman's Rho statistics and P values from molecular vs morphological rates correlation test for each group shown, as well as results from the autocorrelation tests. Significant correlation highlighted in bold. Autocorrelation was considered present when detected in either a molecular or a morphological tree from a given dataset.

Taxonomic group	Autocorrelation of either molecular or morphological branch lengths (p value)	Correlation between molecular and morphological branch lengths (Spearman's Rho)	P value
Araneomorphae	Present (morphology p =0.021)	-0.208	0.276
Cephalotes	Present (molecular p = 0.004)	**-0.361**	**0.007**
Coccomorpha	Absent	-0.184	0.243
Hydroptilidae	Present (molecular p = 0.020)	-0.251	0.053
Hymenoptera	Present (morphology p = 0.041)	-0.163	0.248
Opiliones	Absent	-0.206	0.249

Table 8: BEAST2 relaxed clock analysis likelihood ratio test comparing the relative quality of four linear mixed effects models measuring the relationship between internal molecular and morphological rates in 5 animal phylogenies. With lme1 taking into account taxonomic group as a random effect (lme(Rate~Type, random=~1|Group)), lme2 taking into account branch number as a random effect (lme(Rate~Type,random=~1|Branch)), lme3 taking into account taxonomic group as a random effect with branch number as a nested random effect (lme(Rate~Type, random=~1|Group/Branch)), and lme4 taking into account the same random effects as lme3 but whilst allowing variable variances between datasets (lme(Rate~Type ,random=~1|Group/Branch, weights=varIdent(form=~1|Group)). Where Rate = rate of evolution, type = molecular and morphological partitions, Group = taxonomic group, and Branch = branch pairing.

Model	df	AIC	Log likelihood	Test	Likelihood Ratio	p-value
lme5	4	-3381.024	-3364.677	1694.512		
lme6	4	-3364.539	-3348.192	1686.269		
lme7	5	-3379.024	-3358.59	1694.512	2vs3	16.4852
lme8	9	-3725.742	-3688.961	1871.871	3vs4	354.7179

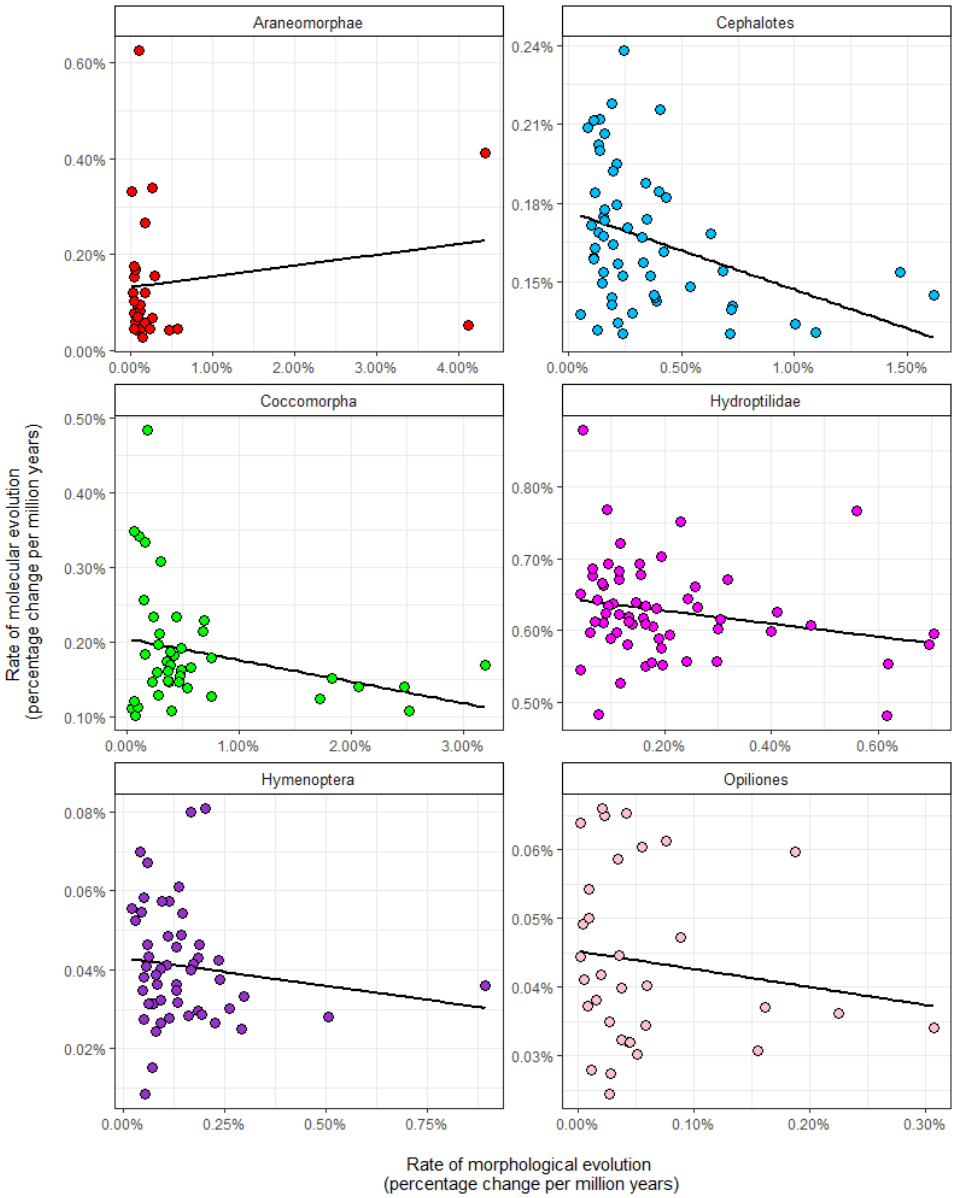

Figure 3: Rates of molecular vs morphological evolution in 6 animal groups. Rates estimated using BEAST2. Rate of evolution as percentage of character state changes occurring per million years (rate of 1.00 = 100% character state changes per million years). Trend lines (linear correlation) are for visualisation only.

9.6 Rates through time analysis

The rates through time analysis revealed a variety of different results. In the majority of cases, rates of molecular and morphological evolution appear to accelerate towards the present, with some interesting exceptions. Individual results for each dataset are as follows.

9.6.1 Araneomorphae:

Rates of molecular evolution in Araneomorphae remain static for most lineages, with a number of internal branches showing elevated rates throughout the Jurassic and Cretaceous as well as in several extant taxa (figure 4). These instances appear to represent rapid molecular evolution in the Palpimanoidea clades Archaeidae and Stenochilidae following their divergence from Mecysmaucheniidae and Huttoniidae, before splitting and undergoing continued rapid genetic differentiation, suggesting that rates of molecular evolution are particularly fast in these lineages. Morphological rates of evolution are relatively constant throughout Araneomorph evolution, remaining below 1% character change change per million years (figure 5). There are two exceptional increases in rates of morphological evolution occurring as Palpimanoidea branch off from their sister clade, the Entelegynae, and in Archaeids after diverging from Stenochilids, where rates of morphological character change reach over 4% per million years.

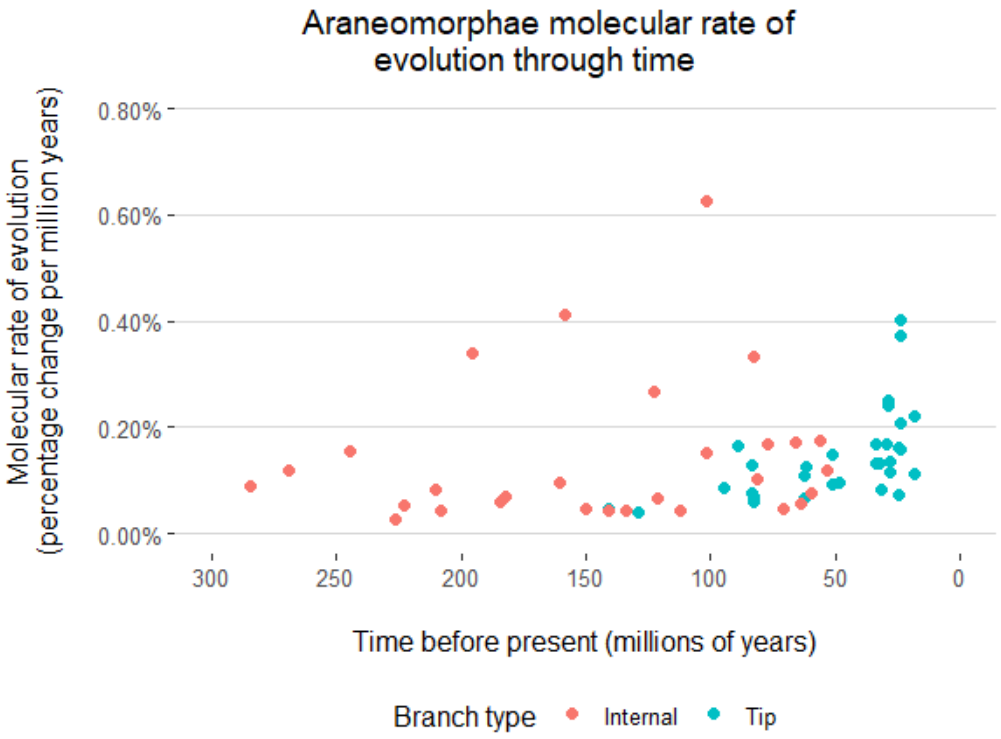

Figure 4: Rates of molecular evolution in Araneomorphae are mostly static, with several internal branches showing accelerated rates during the Jurassic and Cretaceous. Rate of evolution as percentage of character state changes occurring per million years (rate of 1.00 = 100% character state changes per million years) against time before present (midpoint of branch age subtracted from tree height). Terminal and internal branches both considered in molecular rates analyses.

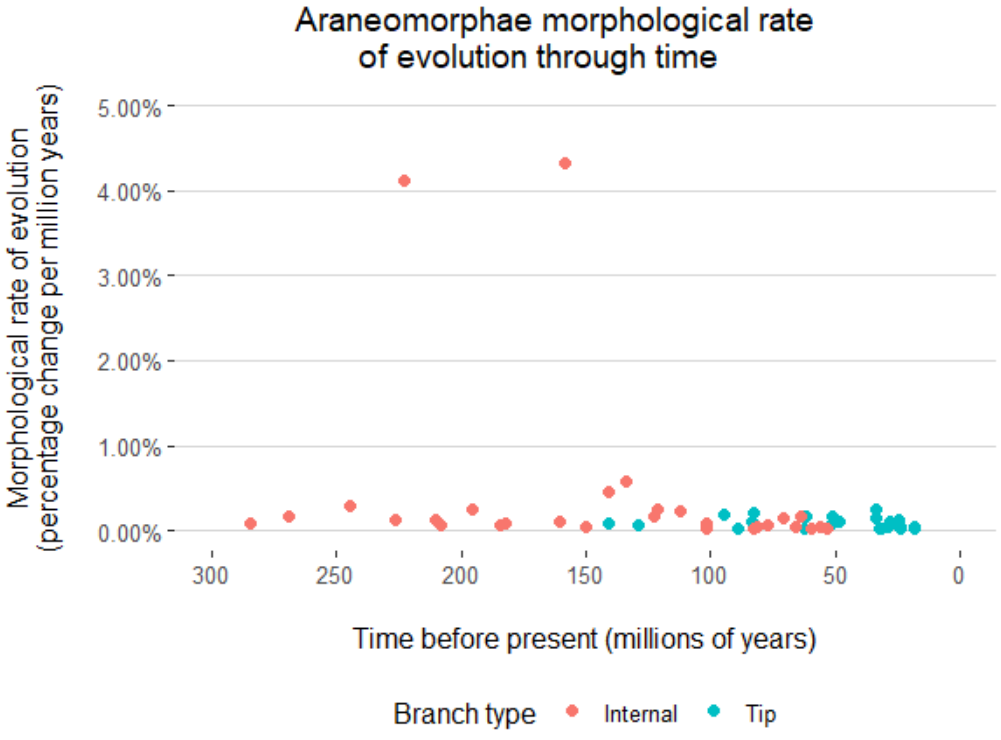

Figure 5 Rates of morphological evolution in Araneomorphae are static, with the exception of two internal branches that show exceptionally high rates during the late Triassic and the late Jurassic. Rate of evolution as percentage of character state changes occurring per million years (rate of 1.00 = 100% character state changes per million years) against time before present (midpoint of branch age subtracted from tree height). Only internal branches considered in morphological rates analyses in consideration of ascertainment bias.

9.6.2 Cephalotes:

Molecular evolution throughout the evolutionary history of the *Cephalotes* appears static, with rates slightly accelerating and decelerating in some lineages towards the present (figure 6). This is likely attributable to the heteroscedasticity of the data with the number of data points increasing from the base of the tree to terminal branches. Rates of morphological evolution show a different pattern, appearing to accelerate prior to several nested radiations (figure 7). There is an exceptionally high rate of morphological evolution in the branch leading to the clade comprising most Central and South American taxa after diverging from their Antillean sister group.

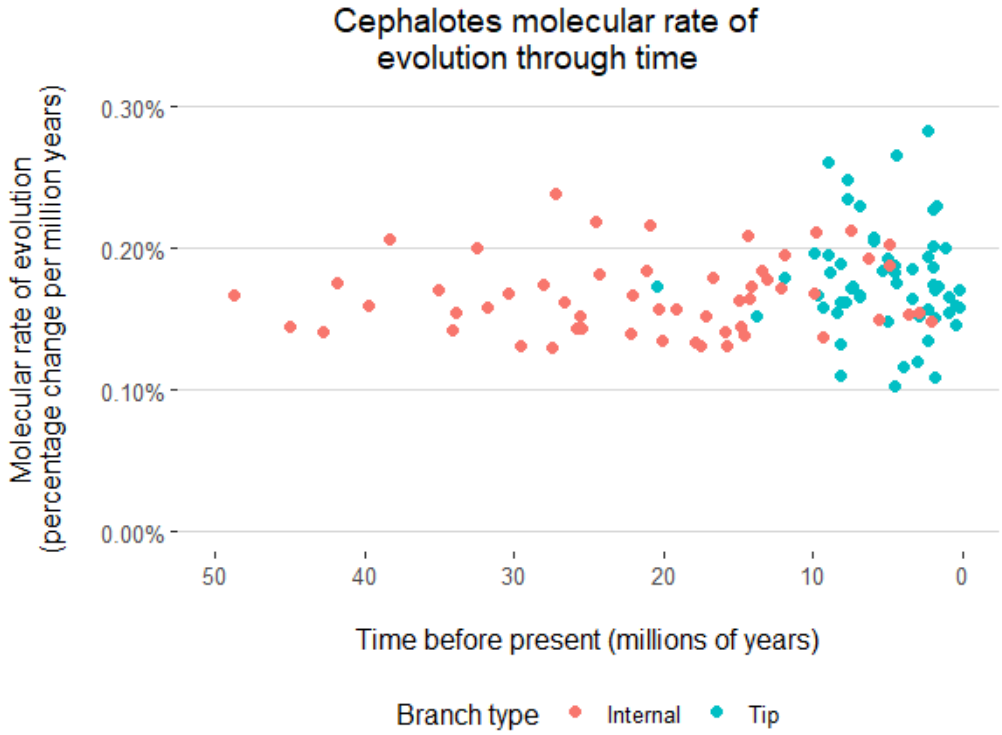

Figure 6 Rates of molecular evolution in Cephalotes are static with very little change through time. Rate of evolution as percentage of character state changes occurring per million years (rate of 1.00 = 100% character state changes per million years) against time before present (midpoint of branch age subtracted from tree height). Terminal and internal branches both considered in molecular rates analyses.

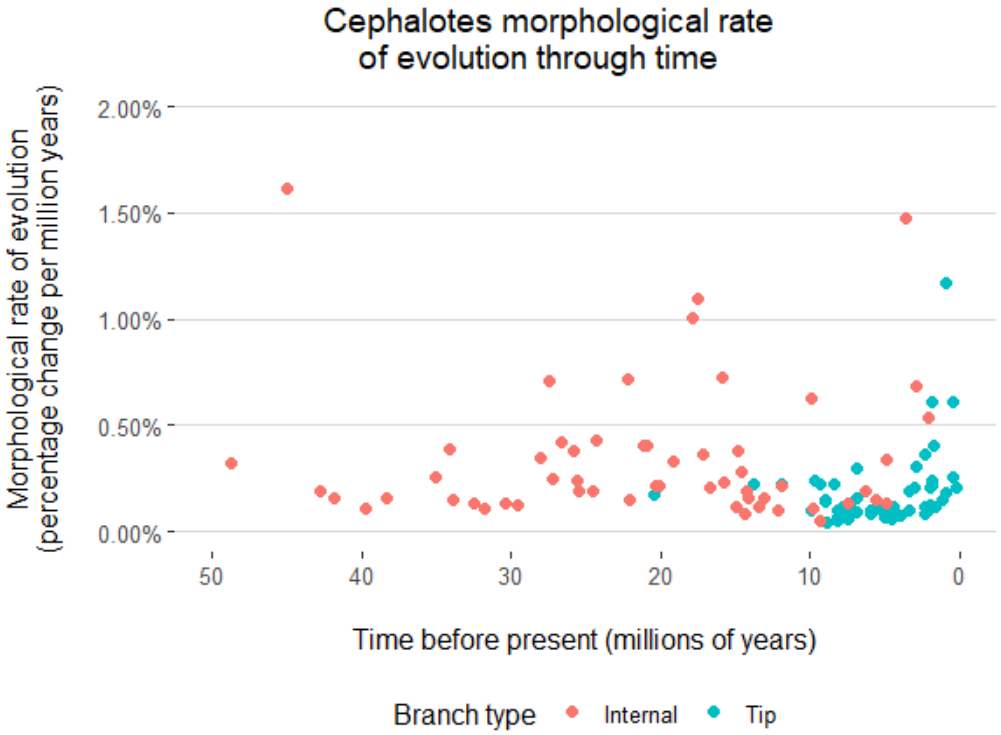

Figure 7: Rates of morphological evolution in the Cephalotes appear to increase through time. The accelerated rates occurring early in the evolutionary history of this group occur at the stem of the clade comprising all South American taxa following their divergence from the Antillean sister group. Rate of evolution as percentage of character state changes occurring per million years (rate of 1.00 = 100% character state changes per million years) against time before present (midpoint of branch age subtracted from tree height). Only internal branches considered in morphological rates analyses in consideration of ascertainment bias.

9.6.3 Coccomorpha:

Initial rates of molecular evolution were static throughout the Jurassic and early Cretaceous, remaining at less than 0.20% character change occurring every million years, before accelerating throughout the Cenozoic to up to over 0.60% character change every million years (figure 8). Rates of morphological evolution are largely static throughout Coccomorphan evolution, remaining at below 1% character change per million years for most lineages. There are several branches with accelerated morphological rates during the late Cretaceous and early Paleogene, whilst rates are also elevated in two branches at the root of the tree, consistent with the divergence of the clade comprising Putoidae, Phenacoleachiidae and Steingeliidae from their sister group, the Neococcoids and *Pityococcus* (figure 9).

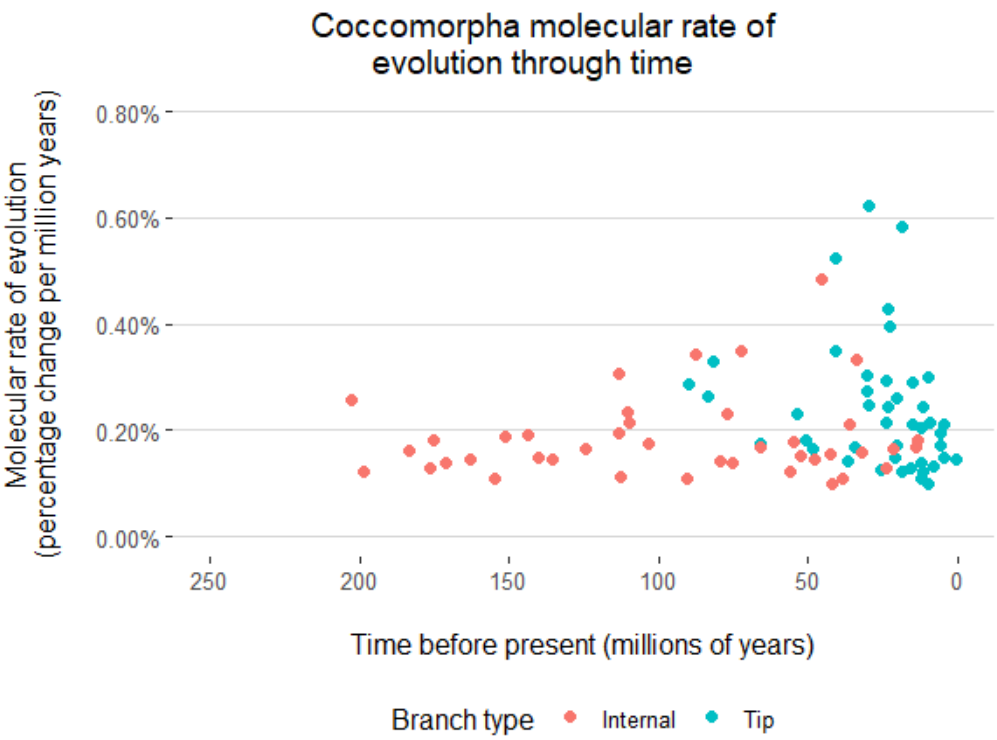

Figure 8: Rates of molecular evolution in Coccomorpha are static before increasing progressively from the mid cretaceous. Rate of evolution as percentage of character state changes occurring per million years (rate of 1.00 = 100% character state changes per million years) against time before present (midpoint of branch age subtracted from tree height). Terminal and internal branches both considered in molecular rates analyses.

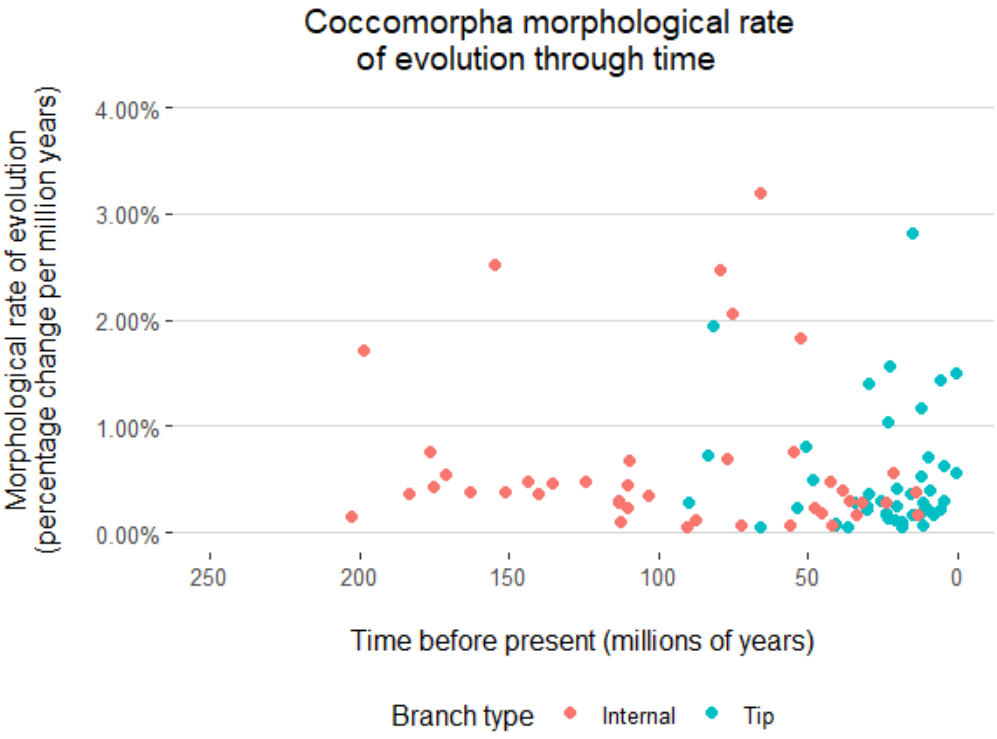

Figure 9: Rates of morphological evolution in Coccomorpha are static, with two branches showing elevated rates during the Jurassic, before rates accelerate again throughout the late Cretaceous. Rate of evolution as percentage of character state changes occurring per million years (rate of 1.00 = 100% character state changes per million years) against time before present (midpoint of branch age subtracted from tree height). Only internal branches considered in morphological rates analyses in consideration of ascertainment bias.

9.6.4 Hydroptilidae:

Rates of molecular evolution in Hydroptilidae follow a similar pattern to the one seen in *Cephalotes*, where lineages start out around 0.60% character change per million years, before rates amongst some lineages very gradually accelerate and decelerate through time, which may correspond to the heteroscedasticity of the data (figure 10). The rate of molecular evolution appears to be high relative to other groups, with the majority of lineages undergoing 0.60-0.80% character change per million years. Rates of morphological evolution are also mostly static, with some branches accelerating towards the present (figure 11).

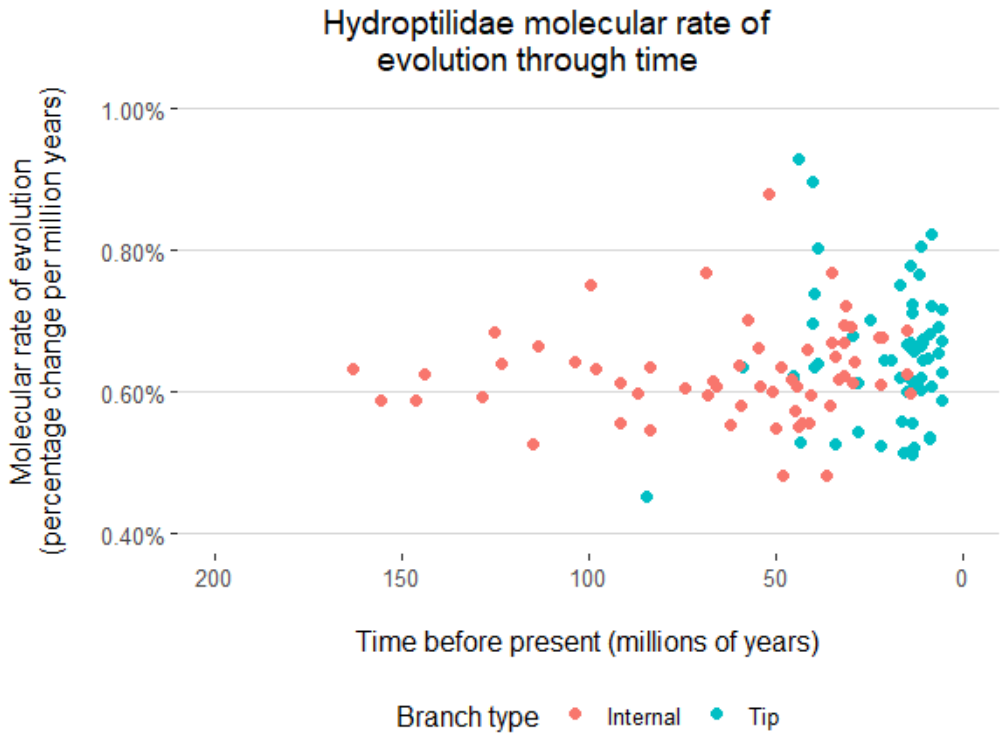

Figure 10: Rates of molecular evolution in Hydroptilidae are static with some lineages showing slightly accelerated and decelerated rates towards the present. Rate of evolution as percentage of character state changes occurring per million years (rate of 1.00 = 100% character state changes per million years) against time before present (midpoint of branch age subtracted from tree height). Terminal and internal branches both considered in molecular rates analyses.

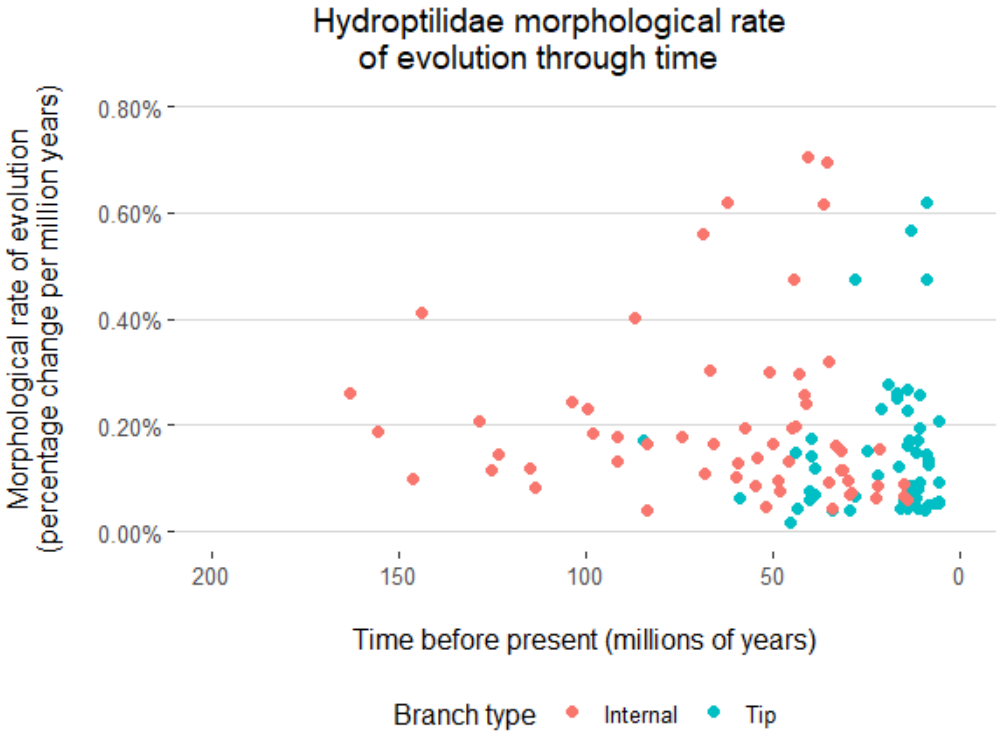

Figure 11: Rates of morphological evolution in Hydroptilidae accelerate after the Cretaceous Rate of evolution as percentage of character state changes occurring per million years (rate of 1.00 = 100% character state changes per million years) against time before present (midpoint of branch age subtracted from tree height). Only internal branches considered in morphological rates analyses in consideration of ascertainment bias

9.6.5 Hymenoptera:

Rates of molecular evolution in the Hymenoptera are relatively low compared to other groups (< 0.15% change per million years) with the highest rates seen in extant taxa (> 0.10% change per million years) (figure 12). Rates of morphological evolution are static, with notable increases occurring after the divergence of Vespina from Xiphidrioidea and after both groups divergence from their sister clade, the Siricoidea, implying rapid morphological differentiation (figure 13).

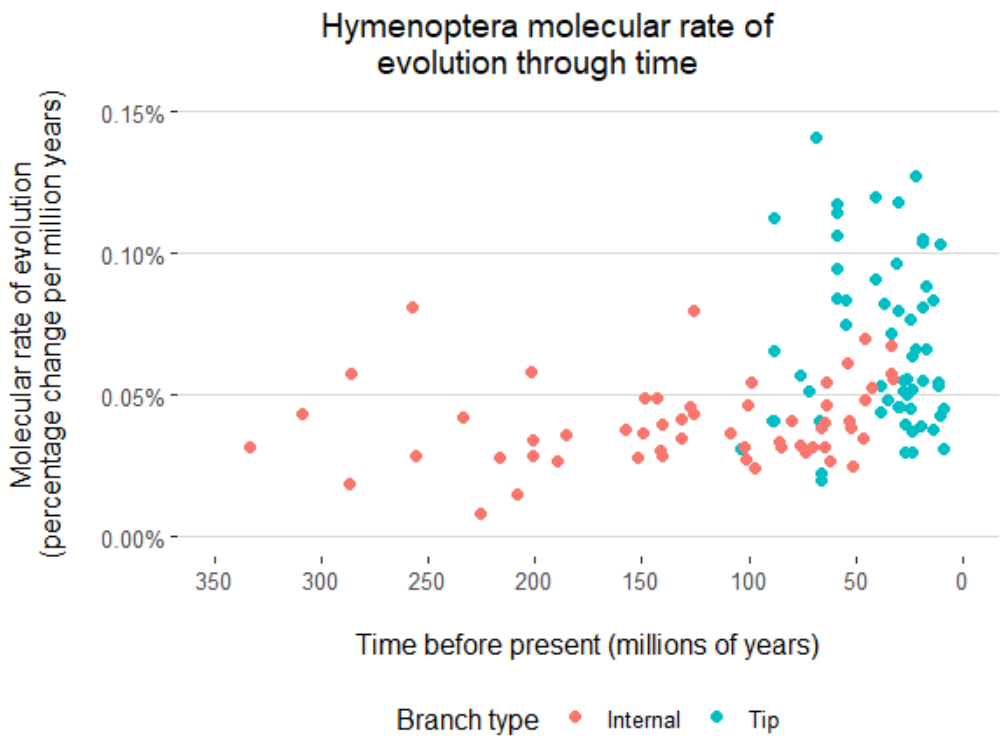

Figure 12: Rates of molecular evolution in Hymenoptera are elevated in terminal taxa, but internal branch rates show no pattern. Rate of evolution as percentage of character state changes occurring per million years (rate of 1.00 = 100% character state changes per million years) against time before present (midpoint of branch age subtracted from tree height). Terminal and internal branches both considered in molecular rates analyses.

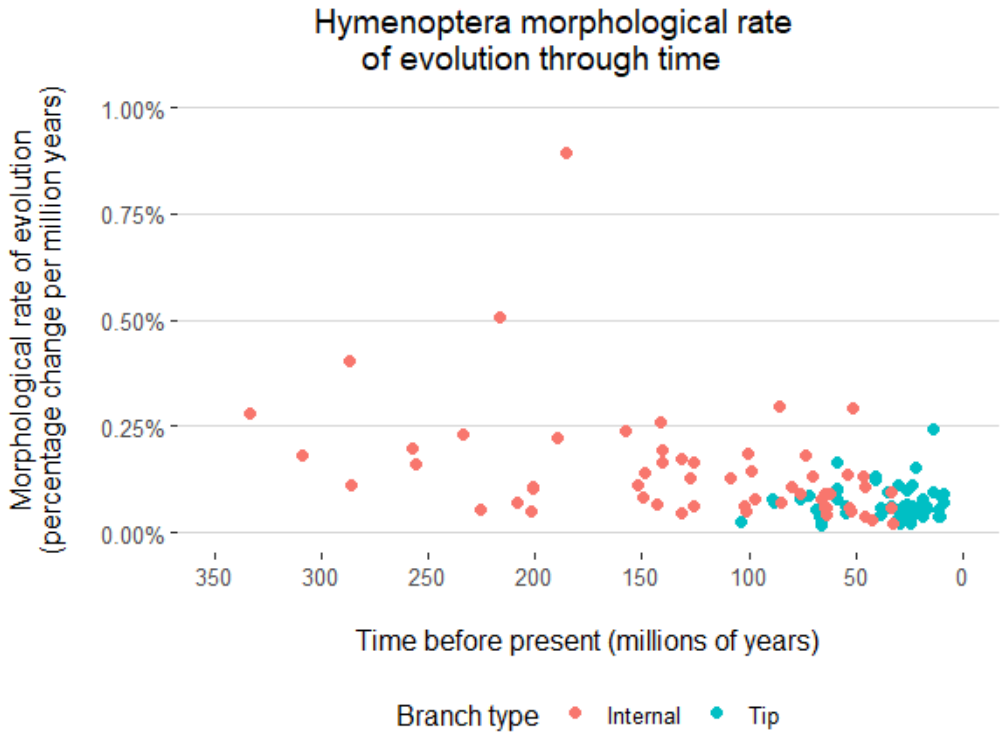

Figure 13: Rates of morphological evolution in Hymenoptera are mostly static, with rates accelerating at the stem of Vespina around the late Triassic early Jurassic period. Rate of evolution as percentage of character state changes occurring per million years (rate of 1.00 = 100% character state changes per million years) against time before present (midpoint of branch age subtracted from tree height). Only internal branches considered in morphological rates analyses in consideration of ascertainment bias

9.6.6 Opiliones:

Rates of molecular evolution in the Opiliones do not reach more than 0.10% change per million years, and remain relatively unchanged through time (figure 14). Morphological rates of evolution are elevated from the Devonian and throughout the Carboniferous and Permian, before decreasing and plateauing throughout the Mesozoic (figure 15). There are two branches with very high rates of morphological evolution within the Dyspnoi clade after the divergence of *T. nepaeformis, D. soerenseni, C. dubium, N. bimaculatum,* and *Ortholasma spp.*, from *N. abei*, and following the divergence of *S. cavicolens, Taracus spp.*, and *H. modestum* from *I. luteipes* and *C. tricantha* suggesting this clade has evolved to be particularly disparate.

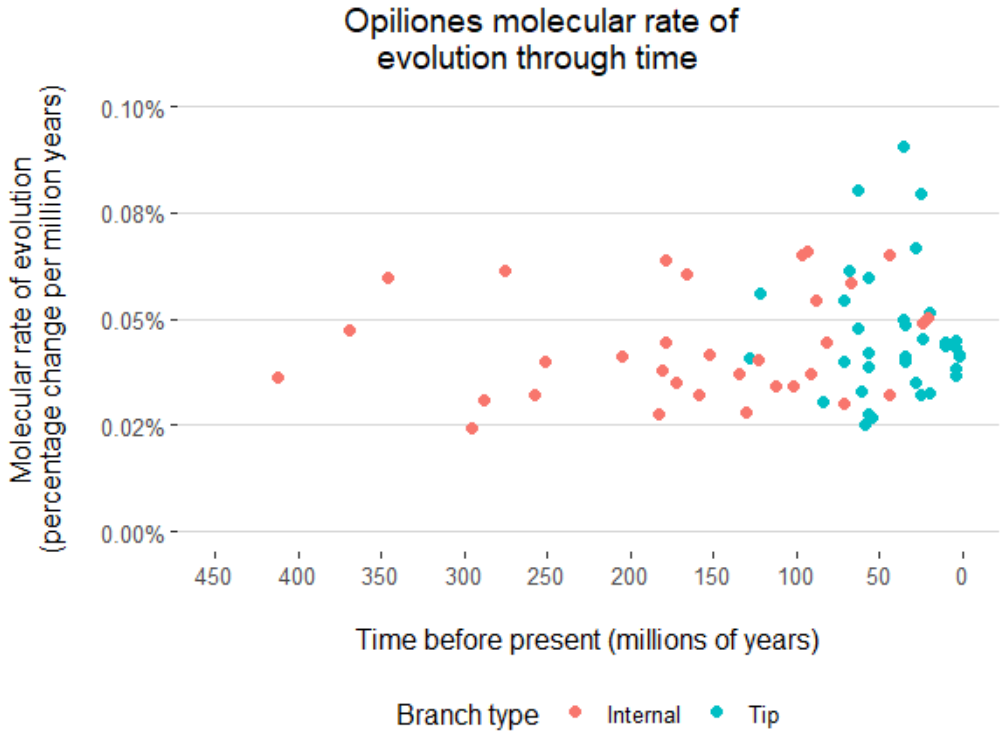

Figure 14: Rates of molecular evolution in Opiliones are static through time. Rate of evolution as percentage of character state changes occurring per million years (rate of 1.00 = 100% character state changes per million years) against time before present (midpoint of branch age subtracted from tree height). Terminal and internal branches both considered in molecular rates analyses

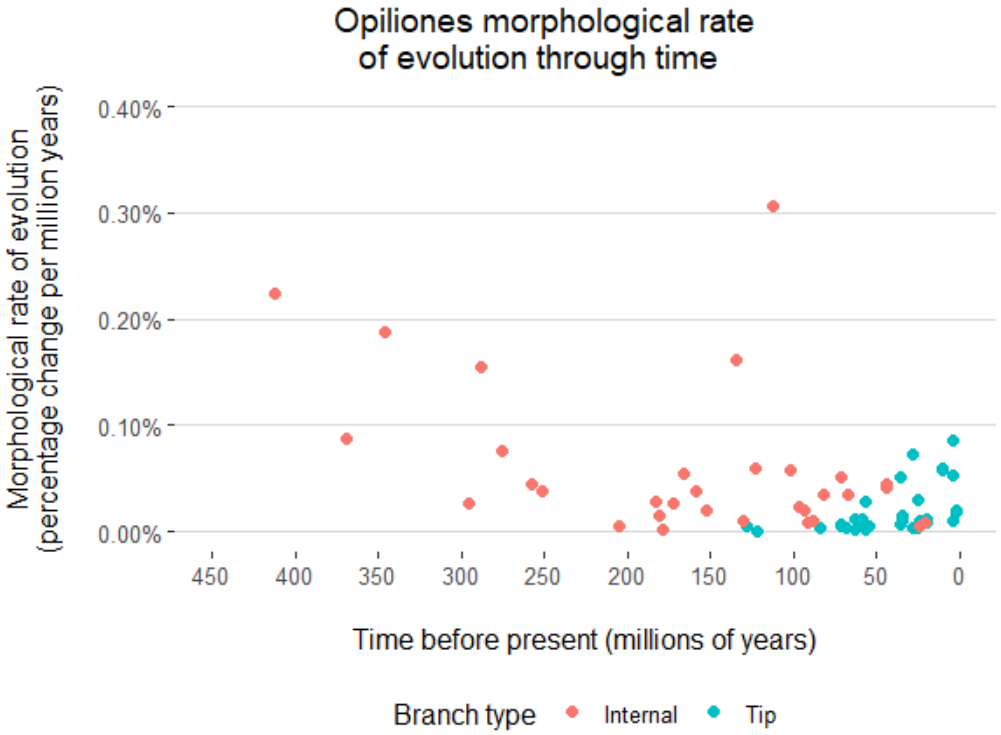

Figure 15: Rates of morphological evolution in Opiliones are elevated from the early Devonian and Carboniferous, before progressively declining towards the Mesozoic. Rates are static throughout the Mesozoic and Cenozoic, with two branches showing accelerated rates within the Dyspnoi during the Cretaceous. Rate of evolution as percentage of character state changes occurring per million years (rate of 1.00 = 100% character state changes per million years) against time before present (midpoint of branch age subtracted from tree height). Only internal branches considered in morphological rates analyses in consideration of ascertainment bias.

10 Discussion

10.1 Correlation tests

Overall, the results of these analyses show that there is no correlation between molecular and morphological branch lengths in animal groups, whilst rates of molecular and morphological evolution along branches are also uncorrelated, supporting the findings of Bromham *et al.* (2002) and Davies and Savolainen (2006). Seligmann (2010) has suggested that results from these previous studies stem from a loss of statistical power following the decision to exclude tips from analyses when testing correlations. It has been shown here that whilst correlation was detected more often when including tips, many of those datasets also show significant autocorrelation. Excluding autocorrelated trees, the pattern that molecular and morphological evolution do not correlate remains the same whether tips are included or excluded, suggesting that that the lack of correlation detected when removing tips is biologically meaningful, and not due to a loss of statistical power

from reduced sample size. This is supported by the fact that excluding tips from the rates correlation tests results in decreased Rho statistic scores and increased p values for Hymenoptera, Opiliones, Coccomorpha, and Hydroptilidae, whilst the two largest datasets (Coccomorpha n = 60, and Hydroptilidae n = 60) do not show any evidence for significant correlation when including or excluding extant taxa (Tables 5 and 7, SI5 supplementary tables 2 and 4).

Phylogenetic autocorrelation has been identified as a potential underlying factor causing correlation in several datasets. Although there are autocorrelated trees present in some datasets that do not show significant correlation in branch lengths and rates, there is greater risk that correlation between two traits would be enhanced in these trees rather than reduced. Regardless, the decision to exclude autocorrelated datasets was taken as a precautionary measure, and the finding that molecular and morphological branch lengths and rates are uncorrelated still stands whether autocorrelated datasets are included or not. Issues such as the node density effect can also be discounted as a factor. The node density effect is an artefact of phylogenetic reconstruction where branch lengths located in areas of the tree containing more terminals tend to be better estimated than those represented by fewer taxa. This is because branches occupying areas of the tree containing more taxa are separated by a greater number of intervening nodes, and this allows multiple character state changes to occur at a given site. Conversely, branch lengths located in areas of the tree where taxon sampling is poor will be underestimated due to a lower number of intervening nodes preventing multiple character state changes from being detected. This ultimately has the effect of artificially lowering root to tip path lengths amongst the least diverse clades in a tree. Although this is primarily an issue when using maximum parsimony, there is some evidence to suggest that maximum likelihood and Bayesian methods may also suffer from this effect (Hugall and Lee, 2007). However, bias introduced by the node density effect would apply to both character types analysed. This would only produce artificial correlation between molecular and morphological evolution, whilst the majority of datasets analysed here show no evidence for correlation at all. Moreover, the node density effect was primarily an issue in previous studies because they had chosen specifically to use root to tip pathways in correlation tests, an approach that has not been implemented here.

Unfortunately, there are some unavoidable issues present when estimating branch lengths and rates of evolution using the methods implemented here. Discrete morphological character data have proven to be an exceptionally useful tool in studies aiming to test

phylogenetic hypotheses and reconstruct evolutionary relationships within the tree of life. However, a potential issue with the information contained within discrete morphological character sets can stem from inconsistencies in the criteria used for character state coding, ordering, and weighting used by researchers. Due to the subjective nature of many qualitative characters, there can often be a great deal of variation in the coding procedure chosen between researchers, which will inevitably have an effect on the reconstructed rate of morphological evolution across different datasets (Pleijel, 1995; Wilkinson, 1995). However, this is precisely the reason that a meta-analysis approach has been implemented here. When analysing just one dataset, it is difficult to rule out measurement error as a factor influencing the results of the analysis. When combining a variety of different clades from datasets constructed by various different authors, there can be more confidence that these results represent a biologically meaningful pattern, as measurement error is highly unlikely to have been the consistent across all datasets analysed.

Another area of uncertainty may lie in the choice of fossil age information when time calibrating phylogenies, as well as in the probability densities chosen to estimate the exact date that the fossil represents. The present study has attempted to avoid seriously biasing divergence date analyses by avoiding the implementation of complex calibration parameters such as the commonly used offset lognormal distribution, which has been shown have a major effect on estimates when used incorrectly (Ho and Phillips, 2009; Warnock *et al*., 2012). Taking the conservative approach by applying a uniform distribution to fossil calibrations with only minimum bounds also allows dating analyses to be consistent across datasets. For simplicity, fossil calibrations used here have been made using the same temporal information incorporated in each original total evidence study. The dates used to calibrate these fossils were based on the judgement of previous authors, and have been well justified by experts specialising in each taxonomic group analysed. Unfortunately, calibration information was only available for a handful of groups, meaning that the relaxed clock rates analysis was largely restricted to those datasets that had been used previously in node dating analyses. Online resources such as the Fossil Calibration Database have recently begun to provide peer reviewed catalogues containing the most recent and well supported fossil calibrations (Ksepka *et al*., 2015). However, sufficient calibration data relevant to the majority of taxa analysed in the present study was unavailable from such a resource. Continued cataloguing of well-justified fossil calibrations in these databases will provide a consistent and reliable means of calibrating these phylogenies, which should be of the benefit to future studies aiming

to measure molecular and morphological rates and estimate the timing of evolutionary transitions.

Whilst the general rule in animals appears to be that molecular and morphological evolution in terms of branch lengths and rates are not correlated, significant correlation was detected in a small minority of datasets. Without conducting any further analyses, it is difficult to establish exactly what factors are causative in this relationship, but it seems likely that underlying mechanisms are at work rather than direct interactions between genotype and phenotype. There is mounting evidence that the primary mechanisms driving morphological variation in animals are mutations occurring in non-coding regions affecting gene regulatory networks (GRNs) responsible for the expression of genes associated with morphogenesis (Carroll, 2008; Stern and Orgogozo, 2008; Frankel *et al.*, 2011; Gaunt and Paul, 2012; Wittkopp and Kalay, 2012). Cis-regulatory elements (CREs) are thought to be central components of GRNs, where mutations arising in these sequences are thought to alter the manner in which transcription factors bind to them. Tighter or looser binding of transcription factors to CREs then has the effect of up or down regulating the transcription of a particular gene associated with the development of a morphological feature (Carroll 2008; Frankel *et al*. 2011; Gaunt and Paul 2012; Wittkopp and Kalay 2012). When analysing rates of molecular evolution in Hawaiian silverswords, Barrier *et al*. (2001) found that nonsynonymous relative to synonymous substitution rates were elevated in regulatory gene sequences, but the same was not observed when analysing rates of evolution in structural genes. The authors found no such acceleration in regulatory sequence evolution rate when analysing closely related and less morphologically diverse North American taxa, concluding that rapid morphological evolution occurring in adaptive radiations may be accompanied by accelerated rates of regulatory gene evolution.

It seems unlikely then, that the genetic markers most often used in phylogenetic analyses are the areas of the genome most likely to drive morphological evolution. Thus, in the rare event that rates correlate, there must be an underlying process indirectly driving the association. (Seligmann, 2010) suggests that rates of evolution at mitochondrial sites frequently used in phylogenetic studies may have an indirect effect on morphological rates of evolution by increasing the efficiency of tissue metabolism, growth, and differentiation. Omland (1997) instead posits that molecular and morphological rates will accelerate and consequently correlate as large populations go through bottlenecks and pre-existing deleterious or nearly neutral alleles arise to fixation more rapidly. When

conducting future investigations into correlations between molecular and morphological evolution, it would be interesting to see how results differ when analysing those sequences most often associated with morphological differentiation, rather than the commonly used markers employed in analyses here.

Even when significant correlation is detected, there is no evidence for any linear relationship. Visual assessment of the non-clock data suggests that longer molecular branch lengths are not accompanied by morphological branch lengths of equivalent lengths, and *vice versa*. The relaxed clock rate correlation tests revealed an inverse relationship in all groups analysed, meaning that as morphological evolution accelerates, molecular evolution decelerates. It is difficult to establish exactly what factors may have driven this effect without conducting further investigation, but it is possible that life history traits such as body size, longevity, and generation time may have an effect (Bromham, 2011). It has been suggested that larger bodied, longer lived organisms accumulate more cells over a lifetime, and thus, must replicate DNA more frequently, meaning that selection will favour more efficient DNA repair machinery to counter the increased likelihood of deleterious mutations arising in the genome (Bromham *et al*., 1996). Moreover, the frequency of meiotic division occurring throughout the life of larger bodied species will be lower than in smaller taxa with shorter generation times. This results in fewer circumstances where DNA copy errors may occur, and ultimately has the effect of slowing genome-wide substitution rate (Thomas *et al*., 2010; Bromham, 2011). The rate of substitution is also thought to scale with metabolic rate (which itself scales inversely with body size) due to a greater oxidative output generating elevated levels of mutagenic by-products (Bromham 2011). The effect that these cumulative effects may have through time is that as organisms evolve larger body sizes, rates of genomic evolution may slow down. Because large body sizes are associated with greater morphological complexity, these lower rates of molecular evolution may be met with accelerated rates of morphological evolution (Bell and Mooers, 1997). When studying mammalian groups, Cooper and Purvis (2009) found that rates of morphological evolution were correlated with body size, but that rodents with intermediate sizes were found to have the slowest rates. The authors also found that rates of morphological evolution increased with elevated environmental temperatures in phyllostomid bats, suggesting that there may be a variety of factors influencing morphological rates of evolution in different lineages. Based on the results here and the findings of previous studies, it would appear that in general, molecular and morphological evolution in animals

take place at independent rates, but that rates may rarely correlate due to confounding variables such as environmental factors and life history traits.

10.2 Rates through time

When analysing rates through time, there does not appear to be a universal pattern across animal groups. In contrast to the findings of Lee *et al*. (2013), molecular evolutionary rates amongst taxa analysed here appear to be either static, or accelerate towards the present. There are a number of clades exhibiting several interesting nested radiations accompanied by rapid morphological evolution, such as in the *Cephalotes*, which likely corresponds to the geographic expansion of this group across South America, resulting in several clades diversifying into isolated adaptive zones (Price *et al*., 2014). These results are in line with the findings of Price *et al.* (2016), who found that rates of cladogenesis and phenotypic diversification in *Cephalotes* were accelerated during periods of ecological transition. Conversely, rates of molecular evolution in this group have remained relatively unchanged through time. Interestingly, the same pattern is observed in the Hydroptilidae, which show a gradual acceleration of morphological rates through time, whilst the molecular rate remains static. This suggests that whilst these particular groups are still ecophenotypically diversifying, the molecular sequences under study have been relatively unaffected by the selective pressures driving morphological evolution.

Other groups show increases in molecular evolutionary rates through time, such as in the Coccomorpha, where rates are static up until around the mid Cretaceous, before accelerating throughout the remainder of the Cretaceous and the subsequent Cenozoic periods. This may be linked to the rapid diversification of many Coccomorphan lineages following a dietary shift to angiosperms, which are thought to have radiated during this period (Vea and Grimaldi, 2016). High substitution rate has been frequently linked with elevated diversification rate amongst plant and animal groups but it is often difficult to determine what is causative (Webster *et al*., 2003; Eo and DeWoody, 2010; Lanfear *et al.*, 2010). It has been suggested that fast substitution rates in some lineages may accelerate genetic differentiation, which would increase the likelihood of traits linked to reproduction and speciation diverging in sub-populations (Venditti and Pagel, 2010). Genetic drift in these small isolated populations would be of greater intensity than in the larger parent population, leading to an increased frequency of nearly neutral and slightly deleterious mutations arising, further accelerating rates throughout the genome (Webster *et al*., 2003; Venditti and Pagel, 2010). Alternatively, rates of molecular evolution

accelerating through time may occur as a result of the evolution of progressively smaller body sizes. This theory fits well with insect groups, which are thought to have evolved increasingly smaller body sizes as atmospheric oxygen levels began to drop following the Permian (Harrison *et al*., 2010).

Another interesting pattern observed in the data is that morphological evolution in several groups takes place at a relatively constant rate through time, with rates abruptly increasing at infrequent intervals. A particularly interesting pattern is seen in the Coccomorpha, where rapid morphological evolution takes place along the branch preceding the Neococcoidea. This basal diversification of morphological traits appears to loosely resemble an early burst model, where lineages phenotypically diversify early in a clades history as ecological space is filled (Gavrilets and Losos, 2009; Rundell and Price, 2009; Czekanski-Moir and Rundell, 2019). However, rates do not plateau as expected. Rather, several bursts of morphological evolution take place amongst Neococcoids throughout the Cretaceous, which may correspond to the rapid Neococcoid diversification associated with ecophenotypical expansion during the angiosperm radiation (Vea and Grimaldi, 2016). Thus, it would seem that amongst this group, genetic and morphological rates of differentiation may have been elevated during this period due to a confounding variable. Rates of molecular and morphological evolution are however, uncorrelated in the Coccomorpha, suggesting that the relationships is perhaps more complex than this.

Accelerated rates of morphological evolution are also noticeable in the Hymenoptera, occurring at the stems of Vespina and Xiphydrioidea. This may represent the switch from phytophagy to parasitoidism in wasps, which facilitated an explosive radiation accompanied by the evolution of the wasp-waist, the venomous stinger, eusociality and a switch to pollen-feeding (Whitfield, 2003; Heraty *et al*., 2011; Peters *et al*., 2017). This abrupt acceleration of evolutionary rate does not seem to take place in the molecular sequence data analysed, suggesting that these sites were not under the same selective pressure during this evolutionary transition. Similarly, the rate of morphological evolution in Araneomorph spiders has remained relatively constant through time, with the exception of two branches showing high rates as Palpimanoidea branch off from Entelegynes, and as Archaeids branch off from Stenochilids. Palpimanoidea, also known as assassin spiders, are specially adapted to preying on other spiders, and have evolved an array of extreme morphological adaptations to suit their unique lifestyle (Pekár *et al*., 2011; Wood *et al*., 2012; Wheeler *et al*., 2017). Amongst this group, the Archaeids have

evolved a particularly elongated cephalic region and set of chelicera to accommodate their specialised method of seizing other spider species from their webs (Wood *et al*., 2012a). Elevated rates of evolution at the base of these clades may correspond to rapid anatomical innovation as they began to expand into these unique niches. Unlike in the Hymenoptera, DNA sequence evolution during these periods appears to have also occurred at an elevated rate, whilst there are several other instances of rate acceleration occurring during the Cretaceous that are not matched by the morphological data. This perhaps suggests that rapid molecular evolution continued to take place after important adaptive features had evolved early in the history of the Palpimanoidea.

The only group that seem to have evolved under an early burst model of evolution similar to that of arthropods during the Cambrian is the Opiliones, where morphological rates are elevated during the Devonian and Carboniferous before decelerating and plateauing throughout the Mesozoic. This may correspond to the early diversification of extant suborders in the Euramerican coal forests of the Carboniferous, where modern body forms were established and remained relatively unchanged throughout subsequent Phanerozoic periods (Garwood *et al.*, 2011, 2014). It is interesting that molecular evolutionary rates in this group do not follow this pattern and remain static throughout all time periods, suggesting that whatever conditions accelerated morphological rates during the Devonian and Carboniferous had no effect on the molecular sequences analysed here.

These results appear to suggest that episodes of rapid morphological evolution occurring during evolutionary transitions in these animal groups are often not accompanied by equivalent patterns in molecular evolutionary rates, which more often seem to remain static or increase gradually as lineages diversify through time. It would seem that molecular and morphological evolution in most cases are operating independently, which may stem from selection operating differently on genetic and phenotypic features. Because selection acts directly on the phenotype and only indirectly on a small fraction of protein coding sites, genomic evolution may not necessarily scale with the accelerated morphological rates of evolution that often occur during evolutionary transitions. For example, a recent study analysing 15,306 nuclear and 12 mitochondrial genes has shown that molecular and morphological rates of evolution throughout the evolutionary history of placental mammals have been largely uncoupled. Notably, they found that mammals had undergone a rapid acceleration of morphological evolution as lineages filled vacant niches following the Cretaceous-Paleogene (K-Pg) extinction event, whilst the same

pattern was not observed when analysing molecular rates (Halliday *et al.*, 2019). Conversely, during long periods of morphological stasis, the vast majority of genomic sites may continue to evolve through genetic drift. For example, a decoupling of rates of molecular and morphological evolution has been observed in meiobenthic copepods, where several cryptic species were found to have undergone extensive genetic differentiation with long-term morphological stasis persisting prior to speciation (Rocha-Olivares *et al.*, 2001). Rates of molecular and morphological evolution have also been found to be disassociated in 'living fossils' such as the tuatara (*Sphenodon punctatus*), as well as within the Platyhelminthes genus *Lamellodiscus*, where rapid genetic differentiation was not accompanied by equivalent rates of morphological diversification (Hay *et al*., 2008; Poisot *et al*., 2011).

One potential reason why the majority of groups analysed here do not follow the early burst model of evolution most often associated with adaptive radiations is that the molecular sequences under study are typically of shorter length (< 6000 nucleotides) than the arthropod sequence analysed by Lee *et al.* (2013), and it may be the case that covering more extensive regions of the genome will provide different estimates. Branch lengths and rates constructed from these sequences only represent the rate of evolution at phylogenetically informative molecular markers, and do not necessarily represent genome-wide rates. Sampling greater areas of the genome, particularly those that are expected to be under direct selection during radiations, would likely produce entirely different patterns of evolution (Bromham and Woolfit, 2004). Lee *et al.* (2013) also made the decision to remove synonymous fast evolving sites form their sequence data. Such sites present in any datasets analysed here may mask interesting rate patterns from being detected in the molecular data. However, systematically assessing and removing fast evolving sites from each dataset would be unfeasible given the time-frame available to the present study. Thus, there are likely to be some differences between the results attained by Lee *et al*. (2013) and the results attained here. However, the fact that rates of morphological evolution also rarely follow an early burst model suggests that there may be other reasons for why rates do not resemble the patterns found in arthropod during the Cambrian explosion. Lee *et al*. (2013) made the decision to study arthropod rates during the Cambrian with the *a priori* expectation that rates would be elevated relative to the subsequent Phanerozoic periods, based on the nature in which modern metazoan phyla abruptly appear in the fossil record during this period. Whilst the ecological principles of this radiation may also apply to other periods of intense diversification throughout the

history of life, the Cambrian explosion was a particularly unique event, and it is unlikely that the groups analysed here will have been exposed to similarly profound ecological and environmental conditions (Rundell and Price, 2009; Czekanski-Moir and Rundell, 2019). A notable exception would be the Opiliones, which are thought to have been one of several arthropod groups undergoing adaptive radiations during the establishment of terrestrial ecosystems throughout the Devonian and Carboniferous, and so it is interesting that morphological rates in this group are elevated during these periods. Morphological rates in mammals have recently been found to have been static during the initial diversification of early lineages throughout the Cretaceous, before accelerating relative to molecular rates following the K-Pg extinction event (Halliday *et al*., 2019). In this sense, mammals did not undergo an early burst of evolution when they appeared during the Cretaceous, and it was not until the K-Pg mass extinction event that the sudden availability of underutilised ecological resources drove morphological rates to accelerate to such a degree (Halliday *et al*., 2016, 2019). The same concept can be applied to several groups analysed here, such as the Coccomorpha, where rates remain static before accelerating during the angiosperm radiation, and the Hymenoptera, where a significant increase in rate of morphological evolution is detected at the stem of Vespina following the acquisition of a key innovation (parasitoidism) (Vea and Grimaldi, 2016; Peters *et al*., 2017).

An alternative explanation for the rapid rates of morphological evolution occurring along the stem of some clades is that there may have been a bias in character selection in some datasets. That is, previous authors may have explicitly sampled synapomorphies that unite specific clades of interest, whilst ignoring those uniting clades considered less important to the investigation. For example the Araneomorphae dataset used by Wood *et al*. (2012b) was used to investigate and compare Archaeid biogeographical patterns to other Palpimanoid groups, with no focus on the remainder of Araneomorph clades within the phylogeny. This dataset is predominantly made up of Archaeid traits, which may allow a larger number of character state changes to be inferred along branches leading to these clades. This may have the overall effect of artificially accelerating rates in the Palpimanoidea, particularly leading up to Archaeid taxa. However, the molecular data also show accelerated rates amongst these groups, which could suggest that some of the signal contained within the morphological characters is biologically meaningful. For other groups, character sampling seems to have been a bit more consistent, such as in the Hymenoptera, where characters uniting all basal clades have been explicitly sampled.

However, this highlights a potential issue with the morphological datasets that are often used in phylogenetic analysis. If there is an inconsistency in character sampling within morphological datasets, this will likely have an effect on the accuracy of studies relying on this information for divergence dating and estimation of evolutionary rates. In order to reconstruct the tree of life and calculate the timing of evolutionary transitions, phylogenies must be dated using information from the fossil record. Because this relies on the accurate phylogenetic placement of fossils, for which molecular data are impossible to attain, this must be performed via the assessment and comparison of morphological characters exhibited between extinct and extant taxa. Incomplete morphological datasets will likely hinder this process, which may lead to incorrect conclusions when testing evolutionary hypotheses. Thus, it is essential that character states relevant to all levels of the phylogeny are sampled in morphological datasets in order to maximise their phylogenetic utility. However, it does not seem that the lack of correlation detected between molecular and morphological evolution is driven solely by incomplete character sampling. The elevated rates that may be attributable to character sampling bias are only found in a minority of branches within each dataset, and some datasets possess similarly elevated rates in the molecular data (Coccomorpha, Araneomorphae). This implies that regardless of sampling bias, morphological characters hold signal that is not found in the genetic markers typically used in phylogenetic analysis. This conflicting signal may stem from the fact that the molecular sequences analysed here do not seem to be being influenced by the same selective pressures as the morphological characters under study, meaning that both character types provide separate, yet equally important information about the evolutionary history of these groups. This has important implications for the results of past and future taxonomic and phylogenetic studies, as analyses focussing on just one character type when testing evolutionary hypotheses may be failing to detect fundamental evolutionary dynamics that would otherwise be picked up in approaches incorporating both character types in a combined analysis.

11 Conclusions

Using different methods, the results of this study support those of Bromham *et al*. (2002) and Davies and Savolainen (2006), that in general, molecular and morphological evolution in animals take place at independent rates. Non-clock branch lengths constructed using molecular and morphological characters are rarely correlated, suggesting that the same principle applies even without incorporating temporal information into analyses. Individually inspecting branch rates as they change through time reveals that the lack of correlation detected may partially stem from molecular sequences remaining unaffected by the intensity of selection occurring as animal groups undergo rapid morphological diversification during evolutionary transitions. This is likely because morphological evolution is primarily driven by changes occurring in regulatory gene sequences, and not by the genetic markers most often used taxonomic and phylogenetic research. Maximising the phylogenetic utility of morphological datasets will require that characters are sampled at all levels in animal datasets, as including only synapomorphies uniting specific clades of interest may inflate inferred evolutionary rates along the stem of these groups. Regardless, there is clearly signal contained within morphological characters that is not picked up in genetic markers, meaning that approaches incorporating just one character type into an analysis may fail to detect fundamental evolutionary dynamics central to the history of the group under study. Thus, it is recommended that all future studies consider this when conducting phylogenetic and taxonomic investigations.

12 References

Agnarsson, I., (2006). Phylogenetic placement of Echinotheridion (Araneae : Theridiidae) – do male sexual organ removal, emasculation, and sexual cannibalism in Echinotheridion and Tidarren represent evolutionary replicas? *Invertebrate Systematics*, 20, 415–429.

Arcila, D., Alexander Pyron, R., Tyler, J.C., Ortí, G., Betancur-R., R., (2015). An evaluation of fossil tip-dating versus node-age calibrations in tetraodontiform fishes (Teleostei: Percomorphaceae). *Molecular Phylogenetics and Evolution*, 82, 131–145.

Barrier, M., Robichaux, R.H., Purugganan, M.D., (2001). Accelerated regulatory gene evolution in an adaptive radiation. *Proceedings of the National Academy of Sciences*, 98, 10208–10213.

Bell, G., Mooers, A.O., (1997). Size and complexity among multicellular organisms. *Biological Journal of the Linnean Society*, 60, 345–363.

Bofkin, L., Goldman, N., (2007). Variation in Evolutionary Processes at Different Codon Positions. *Molecular Biology and Evolution*, 24, 513–521.

Bromham, L., (2011). The genome as a life-history character: why rate of molecular evolution varies between mammal species. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366, 2503–2513.

Bromham, L., Rambaut, A., Harvey, P.H., (1996). Determinants of rate variation in mammalian DNA sequence evolution. *Journal of Molecular Evolution*, 43, 610–621.

Bromham, L., Woolfit, M., (2004). Explosive Radiations and the Reliability of Molecular Clocks: Island Endemic Radiations as a Test Case. *Systematic Biolgy*, 53, 758–766.

Bromham, L., Woolfit, M., Lee, M.S.Y., Rambaut, A., (2002). Testing the relationship between morphological and molecular rates of change along phylogenies. *Evolution*, 56: 1921–1930.

Bryant, H.N., (1995). Why autapomorphies should be removed: a reply to yeates. *Cladistics*, 11: 381– 4.

Cameron, S.A., Mardulyn, P., (2001). Multiple molecular data sets suggest independent origins of highly eusocial behavior in bees (Hymenoptera:Apinae). *Systematic Biology*, 50, 194–214.

Carroll, S.B., (2008). Evo-Devo and an Expanding Evolutionary Synthesis: A Genetic Theory of Morphological Evolution. *Cell*, 134, 25–36.

Cherry, L.M., Case, S.M., Wilson, A.C., (1978). Frog perspective on the morphological difference between humans and chimpanzees. *Science*, 200, 209–211.

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.

Cruz-López, J.A., Francke, O.F., (2017). Total evidence phylogeny of the North American harvestman family Stygnopsidae (Opiliones : Laniatores : Grassatores) reveals hidden diversity. *Invertebrate Systematics*, 31, 317–360.

Czekanski-Moir, J.E., Rundell, R.J., (2019). The Ecology of Nonecological Speciation and Nonadaptive Radiations. *Trends in Ecology and Evolution*, 34, 400–415.

Darriba, D., Taboada, G.L., Doallo, R., Posada, D., (2012). jModelTest 2: more models, new heuristics and high-performance computing. *Nature Methods*, 9, 772.

Dávalos, L.M., Velazco, P.M., Warsi, O.M., Smits, P.D., Simmons, N.B., (2014). Integrating Incomplete Fossils by Isolating Conflicting Signal in Saturated and Non-Independent Morphological Characters. *Systematic Biolgy*, 63, 582–600.

Davies, T.J., Savolainen, V., (2006). Neutral Theory, Phylogenies, and the Relationship Between Phenotypic Change and Evolutionary Rates. *Evolution*, 60, 476–483.

Dohrmann, M., Kelley, C., Kelly, M., Pisera, A., Hooper, J.N.A., Reiswig, H.M., (2017). An integrative systematic framework helps to reconstruct skeletal evolution of glass sponges (Porifera, Hexactinellida). *Frontiers in Zoology*, 14, 18.

Donoghue, P.C.J., Yang, Z., (2016). The evolution of methods for establishing evolutionary timescales. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371, 20160020.

Drummond, A.J., Ho, S.Y.W., Phillips, M.J., Rambaut, A., (2006). Relaxed Phylogenetics and Dating with Confidence. *PLOS Biology*, 4, e88.

Drummond, A.J., Rambaut, A., Shapiro, B., Pybus, O.G., (2005). Bayesian Coalescent Inference of Past Population Dynamics from Molecular Sequences. *Molecular Biology and Evolution*, 22, 1185–1192.

Eldredge, N., Gould S. J., (1972). Punctuated equilibria: an alternative to phyletic gradualism. In: Schopf, T. J. M., ed. *Models in Paleobiology*. Freeman, Cooper and Company San Francisco, pp. 82-115

Eo, S.H., DeWoody, A.J., (2010). Evolutionary rates of mitochondrial genomes correspond to diversification rates and to contemporary species richness in birds and reptiles. *Proceedings of the Royal Society B: Biological Sciences*, Sci. 277, 3587–3592.

Felsenstein, J., (1978). Cases in which Parsimony or Compatibility Methods will be Positively Misleading. *Systematic Biology*, 27, 401–410.

Felsenstein, J., (1985). Phylogenies and the Comparative Method. *The American Naturalist*, 125, 1–15.

Forest, F., (2009). Calibrating the Tree of Life: fossils, molecules and evolutionary timescales. *Annals of Botany*, 104, 789–794.

Forthman, M., Weirauch, C., (2017). Millipede assassins and allies (Heteroptera: Reduviidae: Ectrichodiinae, Tribelocephalinae): total evidence phylogeny, revised classification and evolution of sexual dimorphism. *Systematic Entomology*, 42, 575–595.

Frankel, N., Erezyilmaz, D.F., McGregor, A.P., Wang, S., Payre, F., Stern, D.L., (2011). Morphological evolution caused by many subtle-effect substitutions in regulatory DNA. *Nature*, 474, 598–603.

Freckleton, R.P., Harvey, P.H., (2006). Detecting Non-Brownian Trait Evolution in Adaptive Radiations. *PLOS Biology*, 4.

Garwood, R.J., Dunlop, J.A., Giribet, G., Sutton, M.D., (2011). Anatomically modern Carboniferous harvestmen demonstrate early cladogenesis and stasis in Opiliones. *Nature Communications*, 2, 1–7.

Garwood, R.J., Sharma, P.P., Dunlop, J.A., Giribet, G., (2014). A Paleozoic Stem Group to Mite Harvestmen Revealed through Integration of Phylogenetics and Development. *Current Biology*, 24, 1017–1023.

Garzón-Orduña, I.J., Marini-Filho, O., Johnson, S.G., Penz, C.M., (2013). Phylogenetic relationships of Hamadryas (Nymphalidae: Biblidinae) based on the combined analysis of morphological and molecular data. *Cladistics*, 29, 629–642.

Gaunt, S.J., Paul, Y.-L., (2012). Changes in Cis-regulatory Elements during Morphological Evolution. *Biology*, 1, 557–574.

Gavrilets, S., Losos, J.B., (2009). Adaptive Radiation: Contrasting Theory with Data. *Science*, 323, 732–737.

Gavryushkina, A., Heath, T.A., Ksepka, D.T., Stadler, T., Welch, D., Drummond, A.J., (2017). Bayesian Total-Evidence Dating Reveals the Recent Crown Radiation of Penguins. *Systematic Biology*, 66, 57–73.

Geisler, J.H., McGowen, M.R., Yang, G., Gatesy, J., (2011). A supermatrix analysis of genomic, morphological, and paleontological data from crown Cetacea. *BMC Evolutionary Biology*, 11, 112.

Giles, S., Xu, G.-H., Near, T.J., Friedman, M., (2017). Early members of 'living fossil' lineage imply later origin of modern ray-finned fishes. *Nature*, 549, 265–268.

Gillespie J. H., (1991). *The Causes of Molecular Evolution*. Oxford University Press, Oxford

Glor, R.E., (2010). Phylogenetic Insights on Adaptive Radiation. *Annuals Review of Ecology, Evolution, and Systematics*, 41, 251–270.

Halliday, T.J.D., dos Reis, M., Tamuri, A.U., Ferguson-Gow, H., Yang, Z., Goswami, A., (2019). Rapid morphological evolution in placental mammals post-dates the origin of the crown group. *Proceedings of the Royal Society B: Biological Sciences*, 286, 20182418.

Halliday, T.J.D., Upchurch, P., Anjali, G., (2016). Eutherians experienced elevated evolutionary rates in the immediate aftermath of the Cretaceous–Palaeogene mass extinction. *Proceedings of the Royal Society B: Biological Sciences*, 283, 20153026.

Harrington, S.M., Reeder, T.W., (2017). Phylogenetic inference and divergence dating of snakes using molecules, morphology and fossils: new insights into convergent evolution of feeding morphology and limb reduction. *Biological Journal of the Linnean Society*, 121, 379–394.

Harrison, J.F., Kaiser, A., VandenBrooks, J.M., (2010). Atmospheric oxygen level and the evolution of insect body size. *Proceedings of the Royal Society B: Biological Sciences*, 277, 1937–1946.

Hay, J.M., Subramanian, S., Millar, C.D., Mohandesan, E., Lambert, D.M., (2008). Rapid molecular evolution in a living fossil. *Trends in Genetics*, 24, 106–109.

Heled, J., Bouckaert, R.R., (2013). Looking for trees in the forest: summary tree from posterior samples. *BMC Evolutionary Biology*, 13, 221.

Heraty, J., Ronquist, F., Carpenter, J.M., Hawks, D., Schulmeister, S., Dowling, A.P., Murray, D., Munro, J., Wheeler, W.C., Schiff, N., Sharkey, M., (2011). Evolution of the hymenopteran megaradiation. *Molecular Phylogenetics and Evolutionm* 60, 73–88.

Herrera, J.P., Dávalos, L.M., (2016). Phylogeny and Divergence Times of Lemurs Inferred with Recent and Ancient Fossils in the Tree. *Systematic Biology*, 65, 772–791.

Ho, S.Y.W., Phillips, M.J., (2009). Accounting for Calibration Uncertainty in Phylogenetic Estimation of Evolutionary Divergence Times. *Systematic Biology*, 58, 367–380.

Huang, D., Fitzhugh, K., Rouse, G.W., (2011). Inference of phylogenetic relationships within Fabriciidae (Sabellida, Annelida) using molecular and morphological data. *Cladistics*, 27, 356–379.

Huelsenbeck, J.P., Ronquist, F., Nielsen, R., Bollback, J.P., (2001). Bayesian Inference of Phylogeny and Its Impact on Evolutionary Biology. *Science*, 294, 2310–2314.

Hugall, A.F., Lee, M.S.Y., (2007). The Likelihood Node Density Effect and Consequences for Evolutionary Studies of Molecular Rates. *Evolution*, 61, 2293–2307.

Jenner, R.A., Dhubhghaill, C.N., Ferla, M.P., Wills, M.A., (2009). Eumalacostracan phylogeny and total evidence: limitations of the usual suspects. *BMC Evolutionary Biology*, 9, 21.

Kendall, D.G., (1948). On the Generalized "Birth-and-Death" Process. The Annals of Mathematical Statistics, 19, 1–15.
Kimura, M., (1968). Evolutionary rate at the molecular level. *Nature*, 217, 624–626.
Kimura, M., (1983). *The Neutral Theory of Molecular Evolution*. Cambridge University Press, Cambridge.
King, M.C., Wilson, A.C., (1975). Evolution at two levels in humans and chimpanzees. *Science*, 188, 107–116.
Kingman, J.F.C., (1982). The coalescent. *Stochastic Processes and their Applications*, 13, 235–248.
Ksepka, D.T., Parham, J.F., Allman, J.F., Benton, M.J., Carrano, M.T., Cranston, K.A., Donoghue, P.C.J., Head, J.J., Hermsen, E.J., Irmis, R.B., Joyce, W.G., Kohli, M., Lamm, K.D., Leehr, D., Patané, J.L., Polly, P.D., Phillips, M.J., Smith, N.A., Smith, N.D., Van Tuinen, M., Ware, J.L., Warnock, R.C.M., (2015). The Fossil Calibration Database—A New Resource for Divergence Dating. *Systematic Biolgy*, 64, 853–859.
Kuntner, M., Arnedo, M.A., Trontelj, P., Lokovšek, T., Agnarsson, I., (2013). A molecular phylogeny of nephilid spiders: Evolutionary history of a model lineage. *Molecular Phylogenetics and Evolution*, 69, 961–979.
Laird, C.D., Mcconaughy, B.L., Mccarthy, B.J., (1969). Rate of Fixation of Nucleotide Substitutions in Evolution. *Nature*, 224, 149–154.
Lanfear, R., Frandsen, P.B., Wright, A.M., Senfeld, T., Calcott, B., 2017. PartitionFinder 2: New Methods for Selecting Partitioned Models of Evolution for Molecular and Morphological Phylogenetic Analyses. Mol. Biol. Evol. 34, 772–773.
Lanfear, R., Ho, S.Y.W., Love, D., Bromham, L., (2010). Mutation rate is linked to diversification in birds. *Proceedings of the National Academy of Sciences*, 107, 20423–20428.
Lanfear, R., Thomas, J.A., Welch, J.J., Brey, T., Bromham, L., (2007). Metabolic rate does not calibrate the molecular clock. *Proceedings of the National Academy of Sciences*, 15388–15393.
Lavoué, S., (2016). Was Gondwanan breakup the cause of the intercontinental distribution of Osteoglossiformes? A time-calibrated phylogenetic test combining molecular, morphological, and paleontological evidence. *Molecular Phylogenetics and Evolution*, 99, 34–43.
Lee, M.S.Y., (2016). Multiple morphological clocks and total-evidence tip-dating in mammals. *Biology Letters*, 12, 20160033.
Lee, M.S.Y., Palci, A., (2015). Morphological Phylogenetics in the Genomic Age. *Current Biology*, 25, R922–R929.
Lee, M.S.Y., Soubrier, J., Edgecombe, G.D., (2013). Rates of Phenotypic and Genomic Evolution during the Cambrian Explosion. *Current Biology*, 23, 1889–1895.
Lewis, P.O., (2001). A Likelihood Approach to Estimating Phylogeny from Discrete Morphological Character Data. *Systematic Biology*, 50, 913–925.
Maddison, W. P., Maddison, D. R., (2018). *Mesquite: a modular system for evolutionary analysis*. Version 3.51, http://www.mesquiteproject.org
Marshall, C.R., 2006. Explaining the Cambrian "Explosion" of Animals. Annual Review of Earth and Planetary Sciences 34, 355–384. https://doi.org/10.1146/annurev.earth.33.031504.103001
Marx, F.G., Fordyce, R.E., (2015). Baleen boom and bust: a synthesis of mysticete phylogeny, diversity and disparity. *Royal Society Open Science*, 2, 140434.
Nascimento, F.F., dos Reis, M., Yang, Z., (2017). A biologist's guide to Bayesian phylogenetic analysis. *Nature Ecology & Evolution*, 1, 1446–1454.

Nee, S., May, R.M., Harvey, P.H., (1994). The reconstructed evolutionary process. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 344, 305–311.
Omland, K.E., (1997). Correlated Rates of Molecular and Morphological Evolution. *Evolution*, 51, 1381–1393.
O'Reilly, J.E., dos Reis, M., Donoghue, P.C.J., (2015). Dating Tips for Divergence-Time Estimation. *Trends in Genetics*, 31, 637–650.
O'Reilly, J.E., Puttick, M.N., Parry, L., Tanner, A.R., Tarver, J.E., Fleming, J., Pisani, D., Donoghue, P.C.J., (2016). Bayesian methods outperform parsimony but at the expense of precision in the estimation of phylogeny from discrete morphological data. *Biology Letters*, 12, 20160081.
Patwardhan, A., Ray, S., Roy, A., (2014). Molecular Markers in Phylogenetic Studies-A Review. *Journal of Phylogenetics & Evolutionary Biology* 2, 1–9.
Pekár, S., Šobotník, J., Lubin, Y., (2011). Armoured spiderman: morphological and behavioural adaptations of a specialised araneophagous predator (Araneae: Palpimanidae). *Naturwissenschaften*, 98, 593–603.
Pérez, M.E., Pol, D., (2012). Major Radiations in the Evolution of Caviid Rodents: Reconciling Fossils, Ghost Lineages, and Relaxed Molecular Clocks. *PLOS ONE*, 7, e48380.
Peters, R.S., Krogmann, L., Mayer, C., Donath, A., Gunkel, S., Meusemann, K., Kozlov, A., Podsiadlowski, L., Petersen, M., Lanfear, R., Diez, P.A., Heraty, J., Kjer, K.M., Klopfstein, S., Meier, R., Polidori, C., Schmitt, T., Liu, S., Zhou, X., Wappler, T., Rust, J., Misof, B., Niehuis, O., (2017). Evolutionary History of the Hymenoptera. *Current Biology*, 27, 1013–1018.
Pinheiro J., Bates D., DebRoy S., Sarkar D., R Core Team (2018). *nlme: Linear and No nlinear Mixed Effects Models*. R package version 3.1-137, https://CRAN.R-proje ct.org/package=nlme
Pleijel, F., (1995). On Character Coding for Phylogeny Reconstruction. *Cladistics*, 11, 309–315.
Poisot, T., Verneau, O., Desdevises, Y., (2011). Morphological and Molecular Evolution Are Not Linked in Lamellodiscus (Plathyhelminthes, Monogenea). *PLOS ONE*, 6, e26252.
Price, S.L., Etienne, R.S., Powell, S., (2016). Tightly congruent bursts of lineage and phenotypic diversification identified in a continental ant radiation. *Evolution*, 70, 903–912.
Price, S.L., Powell, S., Kronauer, D.J.C., Tran, L. a. P., Pierce, N.E., Wayne, R.K., (2014). Renewed diversification is associated with new ecological opportunity in the Neotropical turtle ants. *Journal of Evolutionary Biology*, 27, 242–258.
R Core Team (2018). *R: A language and environment for statistical computing. R Found ation for Statistical Computing*. Vienna, Austria. https://www.R-project.org/
Rocha-Olivares, A., Fleeger, J.W., Foltz, D.W., (2001). Decoupling of Molecular and Morphological Evolution in Deep Lineages of a Meiobenthic Harpacticoid Copepod. *Molecular Biology and Evolution*, 18, 1088–1102.
Ronquist, F., Klopfstein, S., Vilhelmsen, L., Schulmeister, S., Murray, D.L., Rasnitsyn, A.P., (2012). A Total-Evidence Approach to Dating with Fossils, Applied to the Early Radiation of the Hymenoptera. *Systematic Biology*, 61, 973–999.
Ronquist, F., Nieves-Aldrey, J.-L., Buffington, M.L., Liu, Z., Liljeblad, J., Nylander, J.A.A., (2015). Phylogeny, Evolution and Classification of Gall Wasps: The Plot Thickens. *PLOS ONE*, 10, e0123301.
Rundell, R.J., Price, T.D., (2009). Adaptive radiation, nonadaptive radiation, ecological speciation and nonecological speciation. *Trends in Ecology & Evolution*, 24, 394–399.

Sanderson, M.J., (1997). A Nonparametric Approach to Estimating Divergence Times in the Absence of Rate Constancy. *Molecular Biology and Evolution*, 14, 1218–1231.

Santos, A.P.M., Nessimian, J.L., Takiya, D.M., (2016). Revised classification and evolution of leucotrichiine microcaddisflies (Trichoptera: Hydroptilidae) based on morphological and molecular data. *Systematic Entomology*, 41, 458–480.

Seligmann, H., (2010). Positive correlations between molecular and morphological rates of evolution. *Journal of Theoretical Biology*, 264, 799–807.

Short, A.E.Z., Cole, J., Toussaint, E.F.A., (2017). Phylogeny, classification and evolution of the water scavenger beetle tribe Hydrobiusini inferred from morphology and molecules (Coleoptera: Hydrophilidae: Hydrophilinae). *Systematic Entomology*, 42, 677–691.

Simpson, G., (1944). *Tempo and mode in evolution*. Columbia University Press, New York

Stadler, T., (2013). How Can We Improve Accuracy of Macroevolutionary Rate Estimates? *Systematic Biology*, 62, 321–329.

Stern, D.L., Orgogozo, V., (2008). The loci of evolution: how predictable is genetic evolution? *Evolution*, 62, 2155–2177.

Teta, P., Cañón, C., Patterson, B.D., Pardiñas, U.F.J., (2017). Phylogeny of the tribe Abrotrichini (Cricetidae, Sigmodontinae): integrating morphological and molecular evidence into a new classification. *Cladistics*, 33, 153–182.

Thomas, J.A., Welch, J.J., Lanfear, R., Bromham, L., (2010). A Generation Time Effect on the Rate of Molecular Evolution in Invertebrates. *Molecular Biology and Evolution*, 27, 1173–1180.

Thomas, J.A., Welch, J.J., Woolfit, M., Bromham, L., (2006). There is no universal molecular clock for invertebrates, but rate variation does not scale with body size. *Proceedings of the National Academy of Sciences*, 103, 7366–7371.

Vea, I.M., Grimaldi, D.A., (2016). Putting scales into evolutionary time: the divergence of major scale insect lineages (Hemiptera) predates the radiation of modern angiosperm hosts. *Scientific Reports*, 6, 23487.

Venditti, C., Pagel, M., (2010). Speciation as an active force in promoting genetic evolution. *Trends in Ecology and Evolution*, 25, 14–20.

Vinther, J., Parry, L., Briggs, D.E.G., Van Roy, P., (2017). Ancestral morphology of crown-group molluscs revealed by a new Ordovician stem aculiferan. Nature 542, 471–474.

Wallace, D.G., Maxson, L.R., Wilson, A.C., (1971). Albumin Evolution in Frogs: A Test of the Evolutionary Clock Hypothesis. *Proceedings of the National Academy of Sciences of the United States of America*, 68, 3127–3129.

Wang, Q., Mao, K.-S., (2016). Puzzling rocks and complicated clocks: how to optimize molecular dating approaches in historical phytogeography. *New Phytologist*, 209, 1353–1358.

Warnock, Yang Ziheng, Donoghue Philip C. J., (2012). Exploring uncertainty in the calibration of the molecular clock. *Biology Letters*, 8, 156–159.

Webster, A.J., Payne, R.J.H., Pagel, M., (2003). Molecular Phylogenies Link Rates of Evolution and Speciation. *Science*, 301, 478–478.

Welch, J.J., Bromham, L., (2005). Molecular dating when rates vary. *Trends in Ecology and Evolution*, 20, 320–327.

Wheeler, W.C., Coddington, J.A., Crowley, L.M., Dimitrov, D., Goloboff, P.A., Griswold, C.E., Hormiga, G., Prendini, L., Ramírez, M.J., Sierwald, P., Almeida-Silva, L., Alvarez-Padilla, F., Arnedo, M.A., Silva, L.R.B., Benjamin, S.P., Bond, J.E., Grismado, C.J., Hasan, E., Hedin, M., Izquierdo, M.A., Labarque, F.M., Ledford, J., Lopardo, L., Maddison, W.P., Miller, J.A.,

Piacentini, L.N., Platnick, N.I., Polotow, D., Silva-Dávila, D., Scharff, N., Szűts, T., Ubick, D., Vink, C.J., Wood, H.M., Zhang, J., (2017). The spider tree of life: phylogeny of Araneae based on target-gene analyses from an extensive taxon sampling. *Cladistics*, 33, 574–616.

Whitfield, J.B., (2003). Phylogenetic Insights into the Evolution of Parasitism in Hymenoptera, *Advances in Parasitology*, 54, 69–100.

Wiens, J.J., Kuczynski, C.A., Townsend, T., Reeder, T.W., Mulcahy, D.G., Sites, J.W., (2010). Combining Phylogenomics and Fossils in Higher-Level Squamate Reptile Phylogeny: Molecular Data Change the Placement of Fossil Taxa. *Systematic Biology*, 59, 674–688.

Wilkinson, M., (1995). A Comparison of Two Methods of Character Construction. *Cladistics*, 11, 297–308.

Wilson, A.C., Carlson, S.S., White, T.J., (1977). Biochemical Evolution. *Annual Review of Biochemistry*, 46, 573–639.

Wittkopp, P.J., Kalay, G., (2012). Cis-regulatory elements: molecular mechanisms and evolutionary processes underlying divergence. *Nature Reviews Genetics*, 13, 59–69.

Wood, H.M., Griswold, C.E., Gillespie, R.G., (2012a). Phylogenetic placement of pelican spiders (Archaeidae, Araneae), with insight into evolution of the "neck" and predatory behaviours of the superfamily Palpimanoidea. *Cladistics*, 28, 598–626.

Wood, H.M., Matzke, N.J., Gillespie, R.G., Griswold, C.E., (2012b). Treating Fossils as Terminal Taxa in Divergence Time Estimation Reveals Ancient Vicariance Patterns in the Palpimanoid Spiders. *Systematic Biology*, 62, 264–284.

Wright, A.M., Hillis, D.M., (2014). Bayesian Analysis Using a Simple Likelihood Model Outperforms Parsimony for Estimation of Phylogeny from Discrete Morphological Data. *PLOS ONE*, 9.

Yeates, D., (1992). Why remove autapomorphies? *Cladistics*, 8(4), 387–389

Yule, G.U., (1925). A mathematical theory of evolution, based on the conclusions of Dr. J. C. Willis, F. R. S. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 213, 21–87.

Zuckerkandl, E., Pauling, L., (1962). Molecular disease, evolution, and genetic heterogeneity. In: Kasha, M., Pullman, B., eds. *Horizons in biochemistry*. Academic Press, New York Pp. 189-225

Zuckerkandl, E., Pauling, L., (1965). Evolutionary divergence and convergence in proteins. In: Bryson, V., and Vogel, H. J., eds. *Evolving genes and proteins*. Academic Press, New York. Pp. 97-166

13 List of supplementary information

Supplementary information included with this thesis:

SI1: Molecular and morphological matrices for MrBayes non-clock analyses

SI2: XML files used in all BEAST2 relaxed clock analyses

SI3: MrBayes non-clock trees

SI4: BEAST2 relaxed clock trees

SI5: Supplementary tables S1-4