Dental characters used in phylogenetic analyses of mammals show higher rates of evolution, but not reduced independence (#41216)

First submission

Guidance from your Editor

Please submit by **21 Oct 2019** for the benefit of the authors (and your $200 publishing discount) .

Structure and Criteria
Please read the 'Structure and Criteria' page for general guidance.

Raw data check
Review the raw data. Download from the materials page.

Image check
Check that figures and images have not been inappropriately manipulated.

Privacy reminder: If uploading an annotated PDF, remove identifiable information to remain anonymous.

Files
Download and review all files from the materials page.

4 Figure file(s)
5 Table file(s)

Structure and Criteria

Structure your review

The review form is divided into 5 sections. Please consider these when composing your review:

1. **BASIC REPORTING**
2. **EXPERIMENTAL DESIGN**
3. **VALIDITY OF THE FINDINGS**
4. General comments
5. Confidential notes to the editor

You can also annotate this PDF and upload it as part of your review

When ready submit online.

Editorial Criteria

Use these criteria points to structure your review. The full detailed editorial criteria is on your guidance page.

BASIC REPORTING

- Clear, unambiguous, professional English language used throughout.
- Intro & background to show context. Literature well referenced & relevant.
- Structure conforms to PeerJ standards, discipline norm, or improved for clarity.
- Figures are relevant, high quality, well labelled & described.
- Raw data supplied (see PeerJ policy).

EXPERIMENTAL DESIGN

- Original primary research within Scope of the journal.
- Research question well defined, relevant & meaningful. It is stated how the research fills an identified knowledge gap.
- Rigorous investigation performed to a high technical & ethical standard.
- Methods described with sufficient detail & information to replicate.

VALIDITY OF THE FINDINGS

- Impact and novelty not assessed. Negative/inconclusive results accepted. *Meaningful* replication encouraged where rationale & benefit to literature is clearly stated.
- All underlying data have been provided; they are robust, statistically sound, & controlled.
- Speculation is welcome, but should be identified as such.
- Conclusions are well stated, linked to original research question & limited to supporting results.

Standout reviewing tips

The best reviewers use these techniques

Tip	*Example*
Support criticisms with evidence from the text or from other sources	*Smith et al (J of Methodology, 2005, V3, pp 123) have shown that the analysis you use in Lines 241-250 is not the most appropriate for this situation. Please explain why you used this method.*
Give specific suggestions on how to improve the manuscript	*Your introduction needs more detail. I suggest that you improve the description at lines 57- 86 to provide more justification for your study (specifically, you should expand upon the knowledge gap being filled).*
Comment on language and grammar issues	*The English language should be improved to ensure that an international audience can clearly understand your text. Some examples where the language could be improved include lines 23, 77, 121, 128 – the current phrasing makes comprehension difficult.*
Organize by importance of the issues, and number your points	*1. Your most important issue* *2. The next most important item* *3. …* *4. The least important points*
Please provide constructive criticism, and avoid personal opinions	*I thank you for providing the raw data, however your supplemental files need more descriptive metadata identifiers to be useful to future readers. Although your results are compelling, the data analysis should be improved in the following ways: AA, BB, CC*
Comment on strengths (as well as weaknesses) of the manuscript	*I commend the authors for their extensive data set, compiled over many years of detailed fieldwork. In addition, the manuscript is clearly written in professional, unambiguous language. If there is a weakness, it is in the statistical analysis (as I have noted above) which should be improved upon before Acceptance.*

Dental characters used in phylogenetic analyses of mammals show higher rates of evolution, but not reduced independence

Neil Brocklehurst [Corresp., 1], **Gemma L Benevento** [2]

[1] Department of Earth Sciences, University of Oxford, Oxford, United Kingdom

[2] School of Geography, Earth and Environmental Sciences, University of Birmingham, Birmingham, United Kingdom

Corresponding Author: Neil Brocklehurst
Email address: neil.brocklehurst@earth.ox.ac.uk

Accurate reconstructions of phylogeny are essential for studying the evolution of a clade. Morphological characters are necessarily used for the reconstruction of fossil organism relationships, however variation in their evolutionary modes not accounted for in analyses may be leading to unreliable phylogenies. A recent study suggested that phylogenetic analyses of mammals may be suffering from a dominance of dental characters, which were shown to have lower phylogenetic signal than osteological characters and produced phylogenies less congruent with molecular benchmarks. Here we build on this previous work by testing seven additional morphological partitions for phylogenetic signal and examining what aspects of dental and other character evolution may be affecting this, by fitting models of discrete character evolution to phylogenies inferred and time calibrated using molecular data. Results indicate that the phylogenetic signal of discrete characters correlate most strongly with rates of evolution, with increased rates driving increased homoplasy. Dental characters have higher rates of evolution than other partitions. They do not, however, fit a model of independent character evolution any worse than other regions. Marsupials show different patterns to other mammal clades, with dental characters evolving at slower rates and being more heavily integrated (less independent). While the dominance of dental characters in analysis of mammals could be leading to inaccurate phylogenies, the issue is not unique to dental characters and the results are not consistent across datasets. Molecular benchmarks (being entirely independent of the character data) provide a framework for examining each dataset individually to assess the evolution of the characters used.

Dental Characters Used in Phylogenetic Analyses of Mammals Show Higher Rates of Evolution, but Not Reduced Independence

Neil Brocklehurst[1], Gemma Louise Benevento[2]

[1] *Department of Earth Sciences, University of Oxford, South Parks Road, Oxford, UK, OX1 3QR*

[2] *School of Geography, Earth and Environmental Sciences, University of Birmingham, Birmingham, UK, B15 2TT,*

ABSTRACT

Accurate reconstructions of phylogeny are essential for studying the evolution of a clade. Morphological characters are necessarily used for the reconstruction of fossil organism relationships, however variation in their evolutionary modes not accounted for in analyses may be leading to unreliable phylogenies. A recent study suggested that phylogenetic analyses of mammals may be suffering from a dominance of dental characters, which were shown to have lower phylogenetic signal than osteological characters and produced phylogenies less congruent with molecular benchmarks. Here we build on this previous work by testing seven additional morphological partitions for phylogenetic signal and examining what aspects of dental and other character evolution may be affecting this, by fitting models of discrete character evolution to phylogenies inferred and time calibrated using molecular data. Results indicate that the phylogenetic signal of discrete characters correlate most strongly with rates of evolution, with increased rates driving increased homoplasy. Dental characters have higher rates of evolution than other partitions. They do not, however, fit a model of independent character evolution any worse than other regions. Marsupials show different patterns to other mammal clades, with dental characters evolving at slower rates and being more heavily integrated (less independent). While the dominance of dental characters in analysis of mammals could be leading to inaccurate phylogenies, the issue is not unique to dental characters and the results are not consistent across datasets. Molecular benchmarks (being entirely independent of the character data) provide a framework for examining each dataset individually to assess the evolution of the characters used.

Keywords: Evolutionary Rates; Homoplasy; Independence; Phylogeny; Mammalia

47

48

49 INTRODUCTION

50 Accurate reconstructions of phylogenetic relationships are essential for studying the
51 evolutionary history of a clade, with hypotheses being based on molecular or morphological
52 data, or both. While it is comparatively straightforward to observe patterns of evolution in
53 molecular sequence data and therefore develop models more closely representing the
54 evolutionary processes, this is more difficult in the case of morphological characteristics due to a
55 poorer understanding of how novel morphology is evolved from ancestral traits. Nonetheless,
56 morphological data is our only means of reconstructing the phylogenetic relationships of fossil
57 organisms that are too old to preserve DNA. It is therefore imperative that we strive to better
58 understand the evolutionary modes of morphological traits. In recent years many studies have
59 examined how variation in their evolutionary patterns accounted for by current analyses may be
60 affecting phylogenetic inferences (e.g. O'Keefe & Wagner, 2001; Scotland, 2003; DeGusta,
61 2004, Sansom et al. 2017; Billet & Bardin, 2018).

62 The high percentage of dental characters used in the reconstruction of fossil mammal
63 phylogenies has become a particular cause for concern. Numerous studies have highlighted
64 issues such as the non-independent evolution of dental characters (Kangas et al., 2004; Kavanagh
65 et al., 2007; Harjunma et al., 2014; Dávalos et al., 2014; Billet & Bardin, 2018) and increased
66 convergence due to ecological selective pressures (Evans et al., 2007; Kavanagh et al., 2007). In
67 a recent meta-analysis, Sansom et al. (2017) examined the phylogenetic signal of tooth and
68 osteological character partitions, using phylogenies derived from molecular data as a benchmark.
69 This study found that osteological characters were more consistent with the molecular

phylogenies and contained greater phylogenetic signal than dental characters, while parsimony analyses with only dental characters produced results less similar to the molecular phylogenies than analyses where the same number of characters were selected at random from both partitions.

This paper builds on the work of Sansom et al. (2017) in two principle ways. Sansom et al. (2017) employed two partitions, dental and osteological, to assess whether dental characters performed more poorly than osteological characters in phylogenetic analyses. As such, while dental characters have been demonstrated to potentially be problematic, an understanding of whether this problem was limited to them, or whether it is seen in other partitions, is lacking. We therefore examine phylogenetic signal in eight morphological partitions in mammals in order to establish whether any other skeletal regions may be a poor indicator of phylogeny.

Secondly, we also aim to understand why dental characters may be producing phylogenies less congruent with molecular benchmarks. It is becoming more well established that morphological characters frequently violate at least some of the principle assumptions of parsimony: between-character rate homogeneity (all characters being just as likely to transition), within-character rate homogeneity (all character states within the same character being similarly likely to transition than others), and character independence (see below). We test each morphological character partition for variation in rates of state transition within characters, variation in rates of evolution between characters, and character independence.

In most published phylogenetic analyses performed using parsimony, the characters are weighted equally (Källersjö et al., 1999; Kluge, 2005; Goloboff et al., 2008). Under such a scheme, a change in any character is given equal emphasis in determining tree length. However, such a scheme only produces reliable results when the characters are all equally likely to change. If, however, there is variation in the rates of character evolution, certain characters will change

more frequently and are more likely to show homoplasy (Felsenstein, 1981; Goloboff, 1993). While parsimony analysis does not incorporate an explicit evolutionary model, an equal weights analysis does rely on equal between-character rates for its accuracy.

Furthermore, in most published phylogenetic analyses, transitions between different combinations of character states are given equal weight (i.e. a transition from state 0 to state 1 is just as likely as a transition from state 1 to state 0; an assumption of within-character rate homogeneity). This assumption may be relaxed by incorporating step matrices which give greater weight to particular transitions (Sankoff & Kedergren, 1983), or by ordering (Fitch, 1971), an extreme modification of step matrices, setting the possibility of most transitions to 0. However, such modifications are rarely employed (see Marjanović & Laurin, 2019 for summary of their history) and most analysis assume equality of within-character rates.

Finally, all phylogenetic analyses (parsimony, Bayesian, and likelihood methods), treat all characters as independent of one another (i.e. an assumption that a change in one character will have no effect on another character transition probability). Extensive study has shown this assumption of independence to be frequently violated (e.g. Kangas et al., 2004; Kavanagh et al., 2007; Harjunma et al., 2014; Dávalos et al., 2014; Billet and Bardin, 2018), with many traits or regions forming integrated modules that change as a unit (Goswami, 2006, 2007; Goswami & Polly, 2010).

By analysing phylogenetic signal, between- and within-character rates, and character independence across eight morphological partitions and four mammalian clades, we aim to better understand how the use of mammalian morphological characters can be optimised during phylogenetic analyses. The results should provide future studies that intend to reconstruct the

relationships of fossil mammals with a framework to enable more evidence-based decisions about which characters are more reliable for use in phylogenetic analyses.

MATERIALS AND METHODS

Data

This study builds on the protocol established by Sansom et al. (2017), where molecular phylogenies are used as the framework over which morphological evolution may be analysed. This allows the evolutionary patterns of the characters to be examined over a phylogeny produced from data entirely independent of those characters. For mammals the time-scaled molecular phylogeny was taken from Meredith et al. (2011), and the morphological data from Bi et al. (2014), both recent and comprehensive datasets. The morphological characters were divided between eight partitions: dental, cranial, axial, pectoral girdle, pelvic girdle, forelimb, hindlimb, and soft tissue. Taxa not present in both the morphological matrix and molecular tree were dropped. If, after doing so, a character showed no variation in score among the remaining taxa, that character was also dropped from subsequent analyses.

As well as the global analysis of mammals, three subclades were subjected to the same analyses to test for variation in the macroevolutionary patterns within Mammalia. The clades chosen were: Cetartiodactyla (Molecular tree from Hassanin et al. [2012], Morphological matrix from O'Leary & Gatesy [2008]), Primates (Molecular tree from Perelman et al. [2011], Morphological matrix from Pattinson et al. [2015]) and Marsupialia (Molecular tree from Mitchell et al. [2014], Morphological matrix from Beck [2017]). These clades were chosen for the following reasons: 1) they have been analysed using morphological character matrices containing characters from all eight of the morphological partitions; 2) there exist time calibrated

138 molecular phylogenies with substantial taxonomic overlap with the morphological matrixes; 3)
139 the character list, data matrix and time calibrated phylogeny were available in usable formats;
140 and 4) they are morphologically and ecologically diverse lineages, and therefore the
141 morphological characters have the potential to be heavily influenced by functional and ecological
142 constraints.
143
144 *Phylogenetic Signal*
145 Levels of homoplasy relative to the molecular phylogeny were used as an estimate of the
146 phylogenetic signal of the characters, measured using Pagel's lambda (Pagel, 1999). This
147 statistic produces a value between 0 and 1, where 0 indicates that character states are distributed
148 independent of phylogeny (no phylogenetic signal). Other methods of calculating phylogenetic
149 signal in discrete characters, for example Moran's I (Gittleman & Kot, 1990) or Fritz & Purvis's
150 D (Fritz & Purvis, 2010), were not used as they are only suitable for binary characters and would
151 require a large proportion of characters to be dropped. For each character, taxa scored as
152 unknown were dropped from the tree. If more than a quarter of the taxa were scored as unknown,
153 the character was not considered in this or subsequent analyses. Pagel's lambda was calculated in
154 R version 3.3.2 (R core team, 2016) using the *fitDiscrete* function in the package Geiger
155 (Harmon et al., 2007).
156
157 *Testing the Assumptions of Phylogenetic Analysis*
158 Within-character rate homogeneity was tested by fitting models of discrete character
159 evolution to the observed phylogeny and trait values using the function *fitDiscrete* in the R
160 package Geiger. Two models were compared: an equal rates (ER) model, where every possible

161 character state transformation has the same rate, and an all-rates-different (ARD) model, where
162 every possible character state transformation is allowed a different rate. The models are
163 compared using the Akaike information criterion, which penalises the parameter-rich ARD
164 model. The Akaike weights of the ER model are used as a metric to assess how well a character
165 obeys the assumption of within-character rate homogeneity.

166 The *fitDiscrete* function also allows testing of between-character rate homogeneity. As
167 well as identifying the model of discrete character evolution that best fits the trait and phylogeny,
168 it also identifies the rates of character-state transformation that best fits the observed data. A
169 higher rate of change means a character is more likely to change multiple times by convergence.
170 If a character was found to best fit the ER model in the above analysis, then the single rate of
171 change was assigned to the character. If the ARD model was found to fit best, the rate assigned
172 to that character was the mean of all rates assigned to each possible transformation, weighted by
173 the number of times each transformation occurred over the phylogeny. The number of transitions
174 was inferred by stochastically mapping the character over the phylogeny 1000 times using the
175 *make.simmap* function in in the R package phytools (Revell, 2012), and calculating the mean
176 frequency of each possible transition.

177 To test character independence, the method of Pagel (1994) was applied to pairwise
178 comparisons of characters. This is again a model-fitting approach, where dependant and
179 independent models of character evolution are fit to pairs of traits and the observed phylogeny.
180 Under the dependant model, the rate of character change in trait 1 will depend on which
181 character state is observed in trait 2, and vice versa. Under the independent model, both
182 characters change state independently of each other. Again, the two models may be compared via
183 the Akaike information criterion, and the Akaike weights of the independent model may be used

as a metric for how well a pair of characters obeys the assumption of independent evolution. Unfortunately, this method is only applicable to binary characters, so non-binary characters were not considered in this section of the analyses. The analysis was implemented using the function *fitPagel* in phytools.

Statistical Comparisons

Pagel's lambda values for each character partition were compared using generalised least squares (GLS), using the R package nlme. For each partition, a null model where all the phylogenetic signal of all partitions comes from the same distribution, was compared to a model where the partition of interest had a different phylogenetic signal to the others (H1). The Akaike weights was used to infer which best fit the data. Partitions that better fit the H1 model were deemed to have significantly different phylogenetic signals than the other partitions, with the GLS coefficient used to identify whether higher or lower. The same method was also applied to the rate values, the support for the ER model, and support for the independent model of evolution.

The rate of character change for each character, and the Akaike weight for the ER model for each character, were both compared to Pagel's lambda using the Kendall's tau correlation coefficient. This latter test could not be applied to the Akaike weights values of the independent model of evolution because these represent pairwise comparisons of characters rather than individual characters.

RESULTS

Results from the Total Mammalian Dataset

The median phylogenetic signal calculated from the Bi et al. (2014) character matrix (the total Mammalia dataset) was 1 for all partitions (Fig. 1A). This indicates that at least half of the characters in each partition are synapomorphies for a single clade. The dental characters do show a larger range of lambda values than most of the other partitions. However, the range of values observed for cranial characters is even wider, indicating that for Mammalia the cranium possesses the largest number of characters with reduced phylogenetic signal. In the GLS analysis, cranial characters are the only partition to not fit the null model best; instead they are found to have significantly lower phylogenetic signal than other partitions (Table 1).

Dental characters show no evidence of increased within-character rate heterogeneity than do the other partitions (Fig. 1B). In fact, the Akaike weights of the equal rates (ER) model are the highest of all the partitions, and in the GLS analysis the dental partition is the only one have significantly better support for the ER model than other partitions (Table 2). Dental characters also show no evidence of increased non-independence (Fig. 1C). Only the pectoral girdle partition was found to have significantly worse support for the independent model of evolution than other partitions (Table 3). The forelimb was found to have significantly better support for the independent model.

However, dental characters have the highest median rates of evolution compared to all other partitions (Fig. 1D), and the increase in rates is significant according to the GLS analysis (Table 4). The pectoral girdle was found to have reduced rates of evolution relative to other partitions, albeit only a marginally significant reduction.

Results from Mammalian Subclade Datasets

The Cetartiodactyla datasets produced similar results to those of mammals overall, albeit with considerably more variation in phylogenetic signal from the vertebral, forelimb and soft tissue characters (Fig. 2). The dental characters are the only partition where the GLS analysis found phylogenetic signal to be significantly reduced (Table S1). Rates of dental evolution are again significantly higher than for other partitions (Fig. 2, Table S4). There is no significant difference found between the Akaike weights support for the ER model of evolution in teeth (Table S2), nor the support for the independent model of character evolution, compared to other partitions (Table S3).

The primate dataset showed less variation in the performance of the various character partitions compared to the cetartiodatyl dataset (Fig. 3). The dental characters again show significantly lower phylogenetic signal than other partitions (Table S5). The range of Pagel's lambda values obtained for the dental characters was wider than other partitions, as was that of forelimb characters (Fig. 3A). However, there is no significant difference in their support for an ER model of evolution compared to other partitions, and their fit to the independent model of evolution is actually significantly better than other partitions. (Tables S6-S7). Rates of evolution in primate dental characters are faster than most other partitions, but the difference is not significant. The only partition to show significantly high rates of character evolution is the pectoral girdle (Table S8)

The marsupial dataset produced results conflicting with the other subclades (Fig. 4). While many of the character partitions, including dentition, show a wide range of Pagel's lambda values, the lambda values of the tooth characters are more concentrated towards higher values than other partitions. The tooth characters show no significant difference in their phylogenetic signal than other partitions (Table S9). The dental characters showed no significant difference

252 from any other partitions in support for the ER model of evolution (Table S10), and no
253 significant difference in rates (Fig. 4D). In contrast to the other datasets, however, the marsupial
254 dataset does support increased character non-independence of dental characters relative to other
255 partitions, with median Akaike weights support for the independent model of evolution lower
256 than all other partitions except the pelvic girdle (Fig. 4C; Table S11).
257
258 *Correlation Tests*
259 In all four datasets, there is a significant negative correlation between lambda and rate of
260 character evolution (Table 5). The correlation between the lambda values and Akaike weights of
261 the ER model is weaker in all four, but in some is still significant.
262
263 DISCUSSION
264 Mammalian tooth characters have been a source of much discussion over the last two
265 decades, due in part to their dominance of the character lists used in morphological phylogenetic
266 analyses of mammals, itself largely a product of their dominance in the mammalian fossil record.
267 Teeth have been shown to suffer from issues such as large amounts of homoplasy (Evans et al.,
268 2007; Davalos et al., 2014) and non-independence (Kangas et al., 2004; Harjunmaa et al., 2014).
269 While these issues clearly do impact on the utility of dental characters in phylogenetic analysis,
270 what has received less attention is whether dental characters are in fact worse affected than other
271 body partitions in these regards. The majority of studies cited above focus solely on teeth, but
272 issues of homoplasy due to ecological and functional constraints might be expected to affect
273 other character partitions (e.g. limb characters being functionally linked to locomotion). Indeed,
274 ecological constraint and developmental linkage has been demonstrated in cranial and limb

275 characters across various tetrapod groups, including mammals (Ruvinsky & Gibson-Brown,
276 2000; Young & Hallgrimson, 2005; Sadleir & Mackovicky, 2008). The same argument could be
277 made for the issue of character non-independence: while this has been demonstrated to be a
278 problem with mammal dentition, recent work on modularity and integration highlights that this
279 issue might just as strongly impact on non-dental characters (Goswami 2006, 2007; Goswami &
280 Polly 2010).

281 Our analyses suggest that increased homoplasy driven by increased rates of evolution
282 may affect dental characters to a greater extent than other partitions. Dental characters from the
283 total mammalian dataset and the cetartiodactyl dataset are found to evolve at faster rates than the
284 other character partitions, and so are more likely to transition multiple times. Moreover, the
285 strong inverse correlation between phylogenetic signal and rates of evolution in all tested
286 datasets indicates that rate variation is likely to be the main driving force behind loss of
287 phylogenetic signal, more so than within-character rate heterogeneity. However, this signal is not
288 consistent across all the tested clades. In the marsupial dataset, for example, dental characters
289 have lower rates (and higher phylogenetic signal) than most other partitions.

290 Moreover, while the results obtained here seem to suggest that dental characters have
291 lower phylogenetic signal than some other characters when optimised over a molecular-based
292 phylogeny, they are not alone in this respect. The mammal dataset indicates that cranial
293 characters also produce low phylogenetic signal. In primates, the forelimb characters have a
294 similar range of Pagel's lambda values to the dental characters (Fig 3A), and in cetartiodactyls
295 the same may be said for hindlimb characters (Fig 2A). One might take this as an indication that,
296 while it is not unreasonable to expect dental characters to contain a strong ecological signal, such
297 a signal is likely to be found in other regions. The limbs of cetartiodactyls, for example, will be

heavily constrained by locomotor type and, in particular, the restrictions placed on the hindlimb by a cursorial lifestyle may be responsible for the reduced phylogenetic signal of hindlimb characters. The hindlimbs in cursorial artiodactyls, as well as in cursorial perissodactyls, have been shown to be responsible for providing the majority of the driving force for such locomotion (Merkens et al., 1993; Dutto et al., 2004; Vaughn et al., 2011). The architecture of the limbs in both clades independently reflects this, with more limited ranges of stance and planes of movement (Liem et al., 2001) and increased muscle mass relative to length (Crook et al., 2008). However, as a counter-point to the suggestion that the constraints of cursoriality are responsible for the reduced phylogenetic signal in cetartiodactyl hindlimbs, one might ask why it is only the hindlimbs that are affected in this way. The forelimbs, for example, while not as important in driving locomotion, should be constrained by the need to "catch" the weight of the animal as it lands (McGuigan & Wilson, 2003; Witte et al., 2004; Vaughn et al., 2011), and so their architecture is constrained by the need to support greater forces. A potential area of future study is to examine whether forelimbs or hindlimbs in cursorial mammals show greater ranges of morphological variability or convergence.

The results observed in cetartiodactyls raise a possibility that might warrant future study: the increase in rates of dental evolution observed might be due to the dominance of herbivores in this dataset. Herbivory has been suggested to be a driver of dental disparity in mammals (Jernvell et al., 1996, 2000) as they morphology tracks a constantly changing resource (plants). Since the functional requirements of eating meat has not changed over time, carnivorous mammals show reduced dental disparity and less evolutionary change (Van Valkenburgh, 1988; Wesley-Hunt et al., 2005). In an analysis of diversification patterns across all mammals, herbivores showed significantly higher diversification rates than carnivores or omnivores (Price & Hopkins, 2015).

While this analysis focussed on lineage diversification, the authors cited increased specialisation and niche-subdivision as a potential driving force behind diversification patterns, and morphological diversification patterns should respond to these drivers in the same way.

It is finally worth noting that in the total-mammal dataset and the two placental subclades tested, there is little evidence that tooth characters are affected by non-independence to any greater extent than the other morphological partitions. The marsupial dataset is the exception, with dental characters showing a weaker fit to the independent model than almost all other partitions, with the exception of pectoral girdle. That pectoral characters are strongly affected by character non-independence in marsupials is unsurprising due to the developmental constraints placed on this girdle and the forelimb; the need for neonatal marsupials, born extremely early in their development, to crawl to the pouch requires these structures to develop precocially, and therefore potentially from a more integrated module (Sears, 2004; Cooper & Steppan, 2010). The integration of the dental characters and their low rates of evolution is likely due to similar constraints; the need to attach to the teat leads to precocial development of the jaw and facial region in marsupials (Smith, 1996, 2006) and they do show reduced dental disparity relative to placentals (Werdlin, 1987)

The concept pioneered by Sansom et al. (2017), of testing morphological discrete characters over a molecular benchmark, is a powerful tool, and it would be highly recommended that researchers studying clades where molecular phylogenies exist examine the performance of their characters in this manner. But given the extremely wide variation in results found by this study, where different partitions produced different relative phylogenetic signals (with the marsupials in particular producing results conflicting strongly with the other datasets studied), one should perhaps be cautious of basing assumption of character quality on the results of large

meta-analyses. While the latter are useful for identifying broad-scale patterns, it is necessary that each dataset be examined individually, and decisions made based on the macroevolutionary patterns observed in that clade. While dental characters have been shown to suffer from issues of homology and non-independence (Kangas et al., 2004; Evans et al., 2007; Harjunmaa et al., 2014), the comparison of the dental characters to finer partitions of data presented here demonstrates that these issues are not unique to teeth. In fact, in some cases other regions perform even worse, and that the nature of these issues varies from clade to clade.

ACKNOWLEDGMENTS

We would like to thank Roger Benson and Robert Sansom for helpful comments and discussion. NB's research is funded by Deutsche Forschungsgemeinschaft grant number BR 5724/1-1. GLB's research is funded by a NERC studentship from the Oxford DTP in Environmental Research (NE/L0021612/1), alongside additional support provided by the European Research Council (grant agreement 637483).

REFERENCES

Beck R. M. 2017. The skull of *Eepidolops ameghinoi* from the early Eocene Itaboraí fauna, southeastern Brazil, and the affinities of the extinct marsupialiform order Polydolopimorphia. *Journal of Mammalian Evolution, 24*, 373–414

Bi S., Wang Y., Guan J., & Meng J. 2014.Three new Jurassic euharamiyidan species reinforce early divergence of mammals. *Nature, 514*, 579–584.

365 Billet, G. & Bardin, J. 2018. Serial homology and correlated characters in morphological
366 phylogenetics: modelling the evolution of dental crests in placentals. *Systematic Biology,*
367 *In Press*

368 Dávalos L.M., Velazco P.M., Warsi O.M., Smits P.D., & Simmons N.B. 2014. Integrating
369 incomplete fossils by isolating conflicting signal in saturated and non-independent
370 morphological characters. *Systematic Biology, 63*, 582–600.

371 DeGusta D. 2004. A method for estimating the relative importance of characters in cladistic
372 analyses. *Systematic Biology, 53*, 529–532

373 Evans A.R., Wilson G.P., Forelius M., & Jernvall J. 2007. High-level similarity of dentitions in
374 carnivorans and rodents. *Nature, 445,* 78–81.

375 Felsenstein, J. 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach.
376 *Journal of Molecular Evolution, 17*, 368-376

377 Fitch W.M. 1971. Toward defining the course of evolution: minimum change for a specific tree
378 topology. *Systematic Biology, 20*, 406-416

379 Fritz S. A., & Purvis A. 2010. Selectivity in mammalian extinction risk and threat types: a new
380 measure of phylogenetic signal strength in binary traits. *Conservation Biology, 24*, 1042-
381 1051.

382 Gittleman J. L. & Jot, M. 1990. Adaptaion-statistics and a null model for estimating phylogenetic
383 effects. *Systematic Zoology, 39*, 227-241

384 Goloboff, P.A. 1993. Estimating character weights during tree search. *Cladistics, 9*, 83-91

385 Goloboff, P.A, Carpenter, J.M., Arias, J.S. & Esquivel, D.R.M. 2008. Weighting against
386 homoplasy improves phylogenetic analysis of morphological datasets. *Cladistics, 24*,
387 758-773.

388 Goswami A. 2006. Cranial modularity shifts during mammalian evolution. *American Naturalist,*
389 *168*, 270–280.

390 Goswami, A. 2007. Cranial modularity and sequence heterochrony in mammals. *Evolution and*
391 *Development, 9*, 290–298.

392 Goswami A. & Polly P. D. 2010. The influence of modularity on cranial morphological disparity
393 in Carnivora and Primates (Mammalia). *PLoS One, 5*, e9517.

394 Harjunmaa E. Seidel K., Häkinen T., Renvoisé, Corfe I.J. Kallonen A., Zhang, Z.Q., Evans A.R.,
395 Mikkola M.L., Klein O.D., & Jernvall, J. 2014. Replaying evolutionary transitions from
396 the dental fossil record. *Nature, 512*, 44–48.

397 Harmon L. J., Weir J. T., Brock C. D., Glor R. E., & Challenger W. 2007. GEIGER:
398 investigating evolutionary radiations. *Bioinformatics, 24*, 129–131

399 Hassanin A., Delsuc F., Ropiquet A., Hammer C., van Vuuren B. J., Matthee C., Ruiz-Garcia M.,
400 Catzeflis F., Areskoug V., Nguyen T. T., & Couloux A. 2012. Pattern and timing of
401 diversification of Cetartiodactyla (Mammalia, Laurasiatheria), as revealed by a
402 comprehensive analysis of mitochondrial genomes. *Comptes Rendus Biologie, 335,* 32–
403 50.

404 Jernvall J., Hunter J. P., & Fortelius M. 1996. Molar tooth diversity, disparity, and ecology in
405 Cenozoic ungulate radiations. *Science, 274*, 1489-1492.

406 Jernvall J., Hunter J. P., & Fortelius M. 2000. Trends in the evolution of molar crown types in
407 ungulate mammals: evidence from the Northern Hemisphere. In Teaford M. F., Smith M.
408 M., & Ferguson M. W. J. (eds) *Development, Function and Evolution of Teeth* (p.p. 269-
409 281). Cambridge University Press, Cambridge.

410 Kälersjö, M., Albert, V.A., & Farris, J.S. 1999. Homoplasy increases phylogenetic structure.
411 *Cladistics, 15*, 91-93

412 Kavanagh, K.D., Evans, A.R. & Jernvall, J., 2007. Predicting evolutionary patterns of
413 mammalian teeth from development. *Nature, 449*, 427–432.

414 Kangas A.T., Evans A.R., Thesleff I., & Jernvall J. 2004. Nonindependence of mammalian
415 dental characters. *Nature, 432*, 211–214.

416 Kluge A.G. 2005. What is the rationale for 'Ockham's razor' (aka parsimony) in phylogenetic
417 inference? In Albert, V.A. (ed) *Parsimony, Phylogeny and Genomics* (pp. 15-42). Oxford
418 University Press, Oxford

419 Marjanović, D., & Laurin, M. 2019. Phylogeny of Palaeozoic limbed vertebrates reassessed
420 through revision and expansion of the largest published relevant data matrix. *PeerJ, 6*,
421 e5565

422 Meredith R.W., Janecka J.E., Gatesy J., Ryder O.A., Fisher C.A., Teeling E.C., Goodbla A.,
423 Eizirik E., Simão T.L., Stadler T., & Rabosky D.L. 2011. Impacts of the Cretaceous
424 Terrestrial Revolution and KPg extinction on mammal diversification. *Science, 334,* 521-
425 524

426 Mitchell K. J., Pratt R. C., Watson L. N., Gibb G. C., Llamas B., Kasper M., Edson J., Hopwood
427 B., Male D., Armstrong K. N., Meyer M., Hofreiter M., Austin J., Donnellan S. C., Lee
428 M. S. Y., Philips M. J., & Cooper A. 2014. Molecular phylogeny, biogeography, and
429 habitat preference evolution of marsupials. *Molecular Biology and Evolution, 31*, 2322–
430 2330.

431 O'Keefe F.R., & Wagner P.K. 2001. Inferring and testing hypotheses of cladistic character
432 dependence by using character compatibility. *Systematic Biology,* 50, 657–675.

433 O'Leary M.A., & Gatesy, J. 2008. Impact of increased character sampling on the phylogeny of
434 Cetartiodactyla (Mammalia): combined analysis including fossils. *Cladistics, 24*, 397–
435 442

436 Pagel M. 1994. Detecting correlated evolution on phylogenies: a general method for the
437 comparative analysis of discrete characters. *Proceedings of the Royal Society B, 255*, 37–
438 45

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

440 Pattinson D. J., Thompson R. S., Piotrowski A. K., & Asher R. J. 2015. Phylogeny,
441 paleontology, and primates: do incomplete fossils bias the Tree of Life? *Systematic*
442 *Biology, 64*,169–186

443 Perelman P., Johnson W. E., Roos, C., Seuánez H. N., Horvath J. E., Moreira M. A. M., Kessing
444 B., Pontius J., Roelke M., Rumpler Y., Schneider M. P. C., Siva A., O'Brien S. J., &
445 Pecon-Slattery J. 2011. A molecular phylogeny of living primates. *PLoS Genetics, 7*,
446 e1001342

447 Price S. A., Hopkins S. S. B., Smith K. K. & Roth, V. L. 2012. Tempo of trophic evolution and
448 its impact on mammalian diversification. *Proceedings of the National Academy of*
449 *Science, 109*, 7008-7012.

450 R Core Team. 2016. *R: a language and environment for statistical computing*. Vienna, Austria.

451 Revell L. J. 2012. phytools: an R package for phylogenetic comparative biology (and other
452 things. *Methods in Ecology and Evolution, 3*, 217–223

453 Ruvinksy I., & Gibson-Brown J. J. 2000. Genetic and developmental bases of serial homology in
454 vertebrate limb evolution. *Development, 127*, 5233–5244.

455 Sadleir R.W., & Makovicky P.J. 2008. Cranial shape and correlated characters in crocodilian
456 evolution. *Journal of Evolutionary Biology, 21*, 1578–1596.

457 Sankoff D., & Cedergren R.J. 1983. Simultaneous comparison of three or more sequences related
458 by a tree. In Sankoff, D., Kruskal, J.B. (eds). *Time Warps, String Edits, and*
459 *Macromolecules: The Theory and Practice of Sequence Comparison* (pp. 253-263).
460 Addison-Wesley, Reading, Massachusetts

461 Sansom, R. S., Wills, M. A., & Williams, T. 2017. Dental data perform relatively poorly in
462 reconstructing mammal phylogenies: morphological partitions evaluated with molecular
463 benchmarks. *Systematic Biology, 66*, 813–822.

464 Scotland R.W., Olmstead R.G., & Bennett J.R. 2003. Phylogeny reconstruction, the role of
465 morphology. *Systematic Biology, 52,* 529–548.

466 Van Valkenburgh B. 1988. Trophic diversity in past and present guilds of large predatory
467 mammals. *Paleobiology, 14*, 155-173

468 Wesley-Hunt G.D. 2005. The morphological diversification of carnivores in North America.
469 *Paleobiology, 31*, 35-55

470 Young N., & Hallgrímsson B. 2005. Serial homology and the evolution of mammalian limb
471 covariation structure. *Evolution, 59*, 2691–2704.

472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487

488 FIGURE CAPTIONS

489 Figure 1: Results from the Bi et al. (2014) character matrix (total Mammalia). A) Pagel's lambda
490 values (phylogenetic signal) of each character. A value of 0 indicates no phylogenetic signal,

491 while a value of 1 indicates high phylogenetic signal. B) Akaike weights support for the ER
492 model of evolution of each character. Characters with an Akaike weights score of 1 have equal
493 rates of within-character evolution between each state, while characters with a score of 0 display
494 unequal rates of within-character state evolution. C) Akaike weights support for the independent
495 model of evolution of all pairwise comparisons of characters in each partition. Pairwise
496 comparisons that have an Akaike weights score of 1 evolve independently of one another, while
497 pairwise comparisons with a score of 0 display character non-independence. D) Rates of
498 character evolution of each character (log scale).
499
500 Figure 2: Results from O'Leary & Gatesy (2008) matrix (Cetartiodactyla). A) Pagel's lambda
501 values (phylogenetic signal) of each character. B) Akaike weights support for the ER model of
502 evolution of each character. C) Akaike weights support for the independent model of evolution
503 of all pairwise comparisons of characters in each partition. D) Rates of character evolution of
504 each character (log scale).
505
506 Figure 3: Results from the Pattinson et al. (2005) matrix (Primata). A) Pagel's lambda values
507 (phylogenetic signal) of each character. B) Akaike weights support for the ER model of
508 evolution of each character. C) Akaike weights support for the independent model of evolution
509 of all pairwise comparisons of characters in each partition. D) Rates of character evolution of
510 each character (log scale).
511
512 Figure 4: Results from the Beck (2017) matrix (Marsupialia). A) Pagel's lambda values
513 (phylogenetic signal) of each character. B) Akaike weights support for the ER model of

514 evolution of each character. C) Akaike weights support for the independent model of evolution
515 of all pairwise comparisons of characters in each partition. D) Rates of character evolution of
516 each character (log scale).

517

518 TABLES

519 Table 1. Results of GLS analyses of Pagel's λ (phylogenetic signal of character partitions) in
520 mammals. Rows coloured are those where the partition best fits the H1 model (partition has a
521 different lambda value to all others); blue indicates lower phylogenetic signal.

Partition	Median λ	GLS Coefficient	lnL (null)	lnL (H1)	AIC (null)	AIC (H1)	Akaike weights (null)	Akaike weights (H1)
Teeth	1	-0.05	-62.01	-63.54	128.03	133.08	0.93	0.07
Skull	1	-0.11	-62.01	-60.24	128.03	126.47	0.31	0.69
Vertebrae	1	0.13	-62.01	-62.65	128.03	131.30	0.84	0.16
Pectoral girdle	1	0.13	-62.01	-62.01	128.03	130.02	0.73	0.27
Pelvic girdle	1	0.12	-62.01	-62.70	128.03	131.40	0.84	0.16
Forelimb	1	0.13	-62.01	-62.28	128.03	130.55	0.78	0.22
Hindlimb	1	0.04	-62.01	-63.69	128.03	133.39	0.94	0.06
Soft tissue	1	0.12	-62.01	-62.58	128.03	131.17	0.83	0.17

522

523

524 Table 2. Results of GLS analyses of Akaike weight support for the equal rates (ER) model of
525 character evolution in mammals. Rows coloured are those where the partition best fits the H1
526 model (partition has a different rate value to all others); red indicates higher support for equal
527 rates.

Partition	Median weight	GLS Coefficient	lnL (null)	lnL (H1)	AIC (null)	AIC (H1)	Akaike weights (null)	Akaike weights (H1)
Teeth	0.99	0.11	54.95	56.91	-105.9	-107.8	0.28	0.72

Skull	0.67	-0.02	54.95	52.53	-105.9	-99.06	0.97	0.04
Vertebrae	0.64	-0.12	54.95	54.74	-105.9	-103.5	0.77	0.23
Pectoral girdle	0.64	-0.04	54.95	53.08	-105.9	-100.2	0.95	0.05
Pelvic girdle	0.70	0.02	54.95	53.24	-105.9	-100.5	0.94	0.06
Forelimb	0.67	-0.08	54.95	54.02	-105.9	-102.0	0.87	0.13
Hindlimb	0.72	0.02	54.95	52.63	-105.9	-99.25	0.97	0.03
Soft tissue	0.66	0.03	54.95	53.72	-105.9	-100.4	0.90	0.10

528

529 Table 3. Results of GLS analyses of Akaike weight support for the independent model of
530 character evolution in mammals. Rows coloured are those where the partition best fits the H1
531 model (partition has a different rate value to all others); blue indicates lower Akaike weights, red
532 indicates higher.

Partition	Median weight	GLS Coefficient	lnL (H0)	lnL (H1)	AIC (H0)	AIC (H1)	Akaike weights (H0)	Akaike weights (H1)
Teeth	0.75	0.04	1028.8	1027.6	-2054	-2049	0.90	0.10
Skull	0.73	-0.02	1028.8	1028.5	-2054	-2051	0.78	0.22
Vertebrae	0.59	-0.02	1028.8	1026.5	-2054	-2047	0.96	0.04
Pectoral girdle	0.63	-0.06	1028.8	1030.2	-2054	-2054	0.39	0.61
Pelvic girdle	0.80	0.04	1028.8	1026.9	-2054	-2048	0.94	0.06
Forelimb	0.76	0.04	1028.8	1035.7	-2054	-2065	0.01	0.99
Hindlimb	0.60	0.02	1028.8	1026.5	-2054	-2047	0.96	0.04
Soft tissue	0.75	0.08	1028.8	1028.1	-2054	-2050	0.84	0.16

533

534

535 Table 4. Results of GLS analyses of rates of character evolution in mammals. Rows coloured are
536 those where the partition best fits the H1 model (partition has a different rate value to all others);
537 blue indicates lower rate, red indicates higher rate.

Partition	Median rate	GLS Coefficient	lnL (null)	lnL (H1)	AIC (null)	AIC (H1)	Akaike weights	Akaike weights

							(null)	(H1)
Teeth	0.0016	0.21	-98.17	-94.88	200.34	195.76	0.09	0.91
Skull	0.0010	0.04	-98.17	-99.99	200.34	205.98	0.94	0.06
Vertebrae	0.0007	-0.22	-98.17	-97.78	200.34	201.56	0.64	0.35
Pectoral girdle	0.0007	-0.19	-98.17	-96.98	200.34	199.96	0.45	0.55
Pelvic girdle	0.0009	-0.06	-98.17	-99.20	200.34	204.40	0.88	0.12
Forelimb	0.0010	-0.07	-98.17	-99.38	200.34	204.77	0.90	0.10
Hindlimb	0.0008	-0.03	-98.17	-100.0	200.34	206.05	0.95	0.05
Soft tissue	0.0006	-0.21	-98.17	-98.33	200.34	202.65	0.76	0.24

538

539

540 Table 5. Results of Kendal's tau correlation tests

	Pagel's lambda vs Rates of character evolution	Pagel's lambda vs Akaike weight support for ER model of character evolution
Total mammal dataset	-0.27 ($p = 8.91 \times 10^{-9}$)	-0.062 (p = 0.1906)
Cetartiodactyl dataset	-0.25 ($p=3.75 \times 10^{-11}$)	0.15 ($p=6.37 \times 10^{-5}$)
Primate Dataset	-0.23 ($p=8.26 \times 10^{-8}$)	-0.033 (p=0.44)
Marsupial dataset	-0.37 ($p=6.46 \times 10^{-12}$)	0.18 (p=0.0063)

541

542

543

Figure 1

Figure 1: Results of analyses from the Bi et al. (2014) character matrix (total Mammalia)

A) Pagel's lambda values (phylogenetic signal) of each character. A value of 0 indicates no phylogenetic signal, while a value of 1 indicates high phylogenetic signal. B) Akaike weights support for the ER model of evolution of each character. Characters with an Akaike weights score of 1 have equal rates of within-character evolution between each state, while characters with a score of 0 display unequal rates of within-character state evolution. C) Akaike weights support for the independent model of evolution of all pairwise comparisons of characters in each partition. Pairwise comparisons that have an Akaike weights score of 1 evolve independently of one another, while pairwise comparisons with a score of 0 display character non-independence. D) Rates of character evolution of each character (log scale).

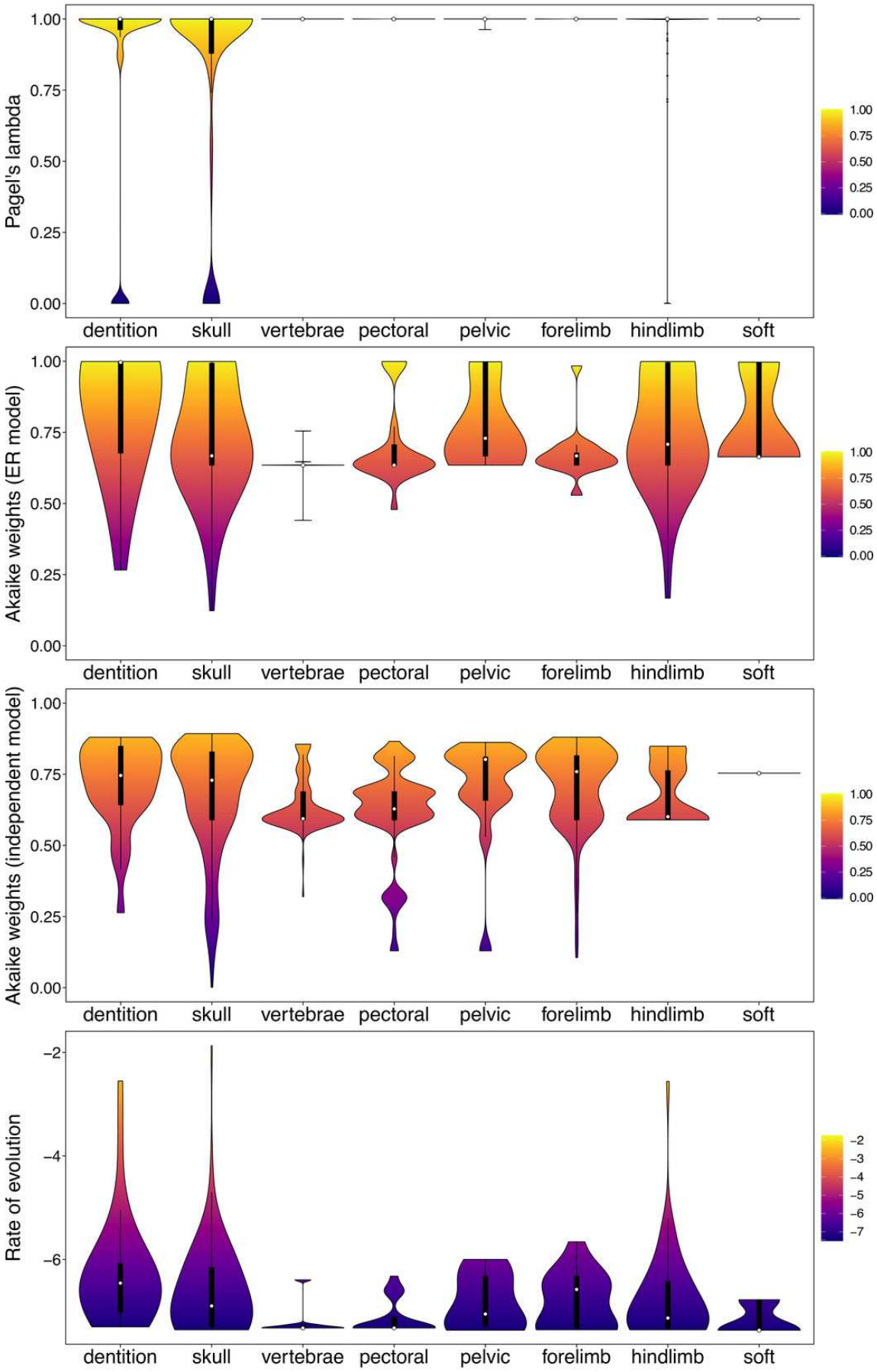

Pagel's lambda
1.00
0.75
0.50
0.25
0.00
dentition
skull
vertebrae
pectoral
pelvic
forelimb
hindlimb
soft
Akaike weights (ER model)
Akaike weights (independent model)
Rate of evolution
−2
−4
−6
−3
−5
−7

Figure 2

Figure 2: Results from O'Leary & Gatesy (2008) matrix (Cetartiodactyla)

A) Pagel's lambda values (phylogenetic signal) of each character. B) Akaike weights support for the ER model of evolution of each character. C) Akaike weights support for the independent model of evolution of all pairwise comparisons of characters in each partition. D) Rates of character evolution of each character (log scale).

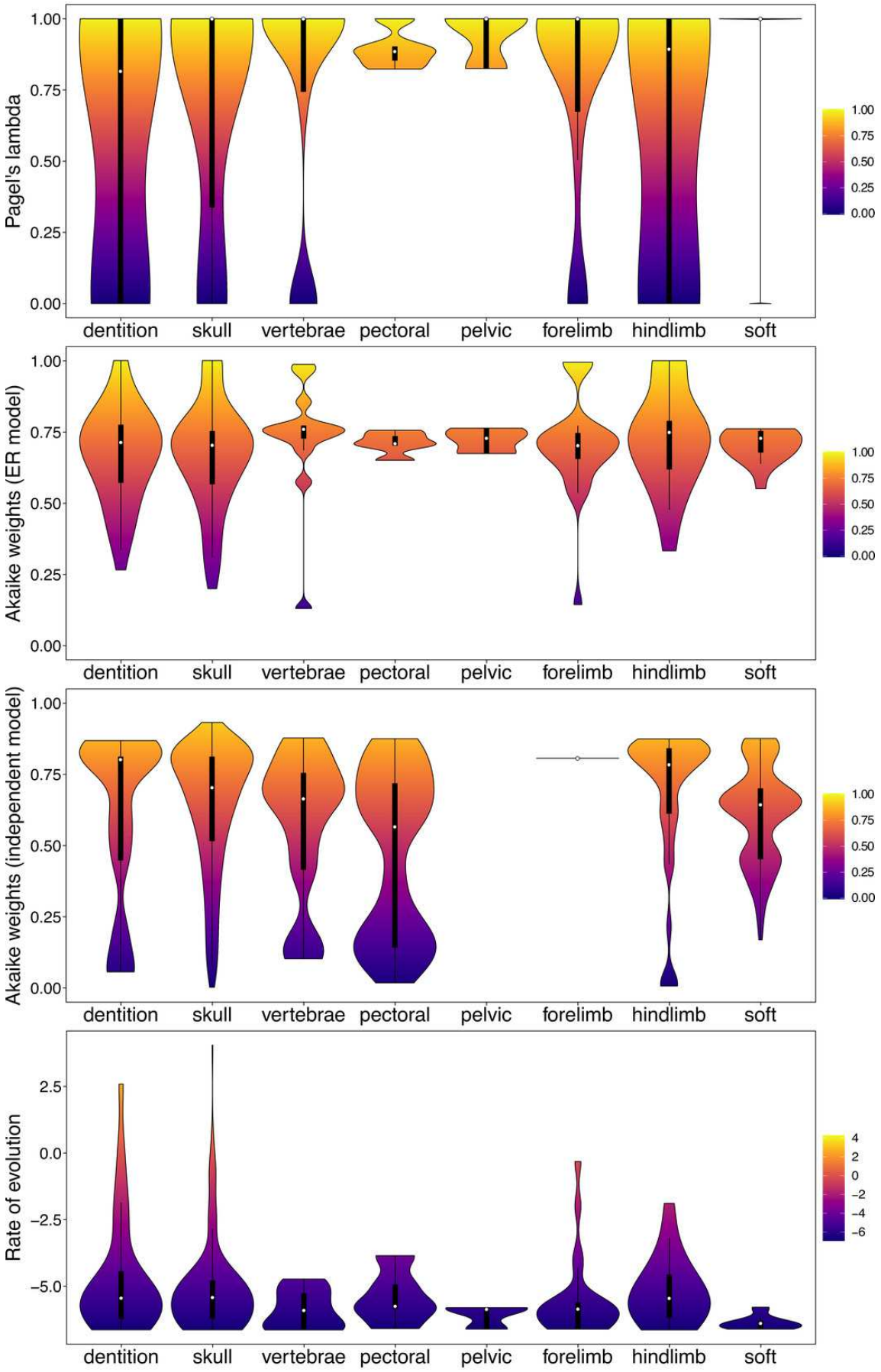

Pagel's lambda
1.00
0.75
0.50
0.25
0.00
dentition
skull
vertebrae
pectoral
pelvic
forelimb
hindlimb
soft
Akaike weights (ER model)
Akaike weights (independent model)
Rate of evolution
2.5
0.0
−2.5
−5.0
4
2
0
−2
−4
−6

Figure 3

Figure 3: Results from the Pattinson et al. (2005) matrix (Primates)

A) Pagel's lambda values (phylogenetic signal) of each character. B) Akaike weights support for the ER model of evolution of each character. C) Akaike weights support for the independent model of evolution of all pairwise comparisons of characters in each partition. D) Rates of character evolution of each character (log scale).

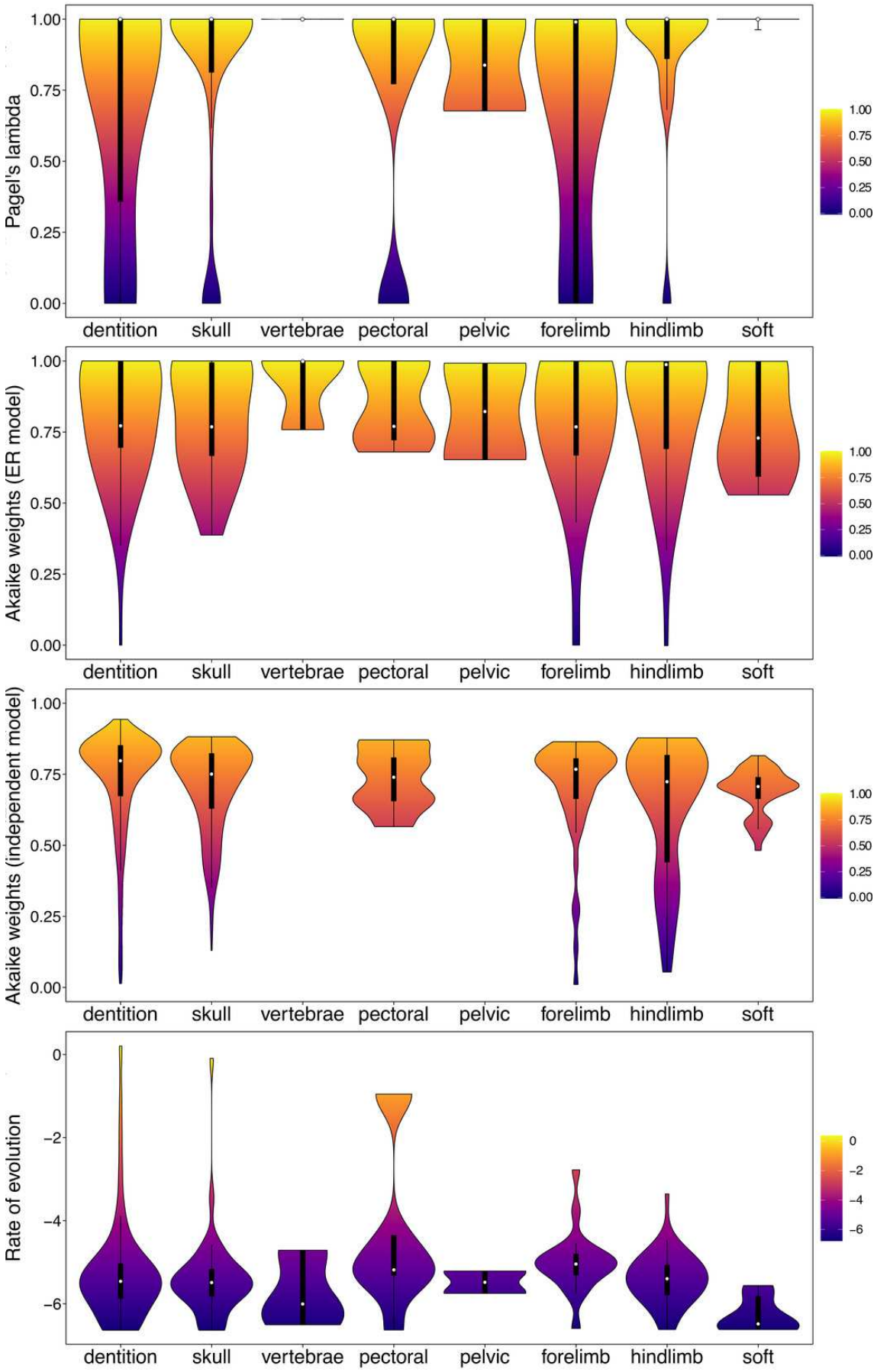
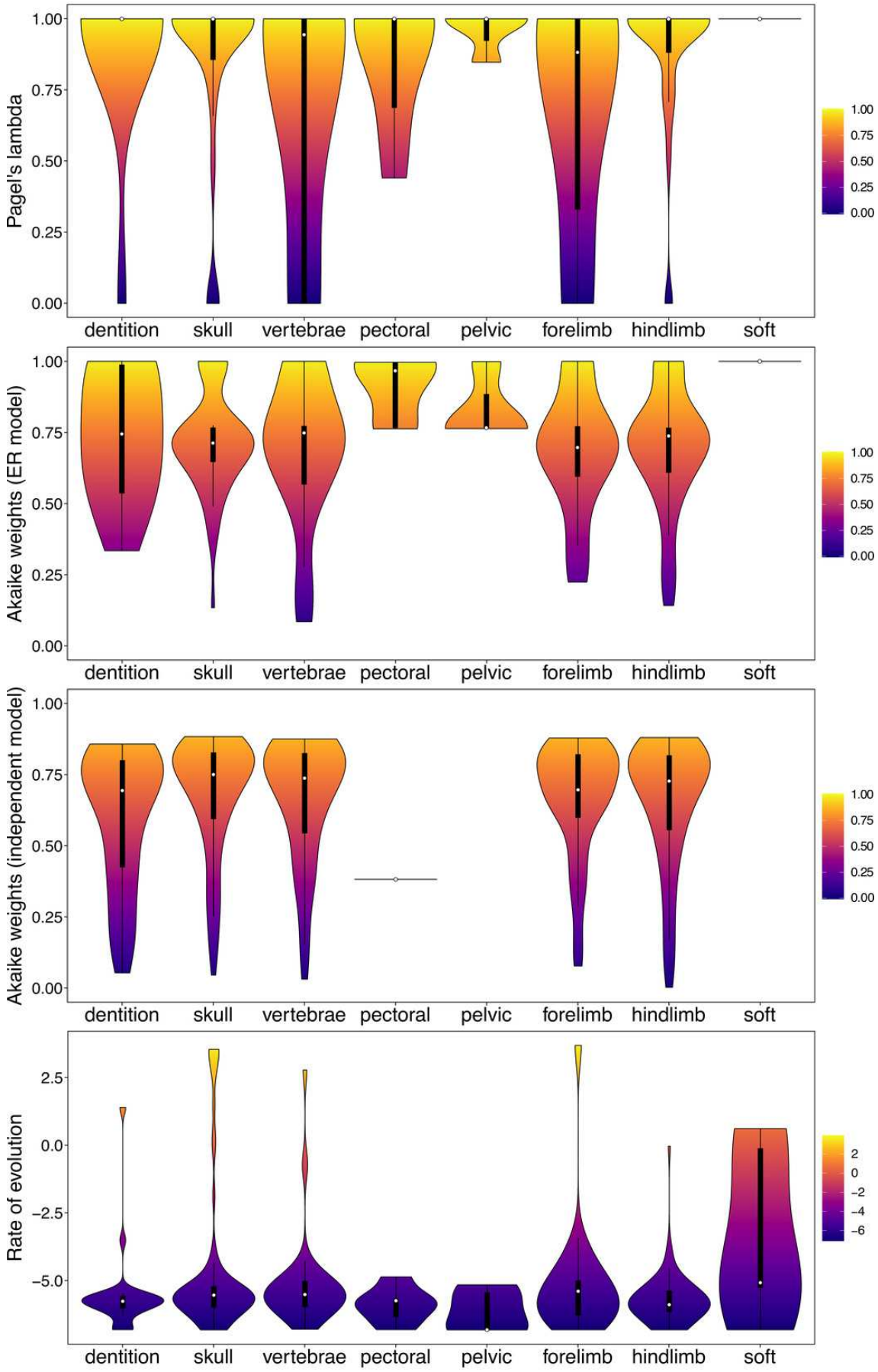

Pagel's lambda
1.00
0.75
0.50
0.25
0.00
dention
skull
vertebrae
pectoral
pelvic
forelimb
hindlimb
soft
Akaike weights (ER model)
Akaike weights (independent model)
Rate of evolution
0
−2
−4
−6

Figure 4

Figure 4: Results from the Beck (2017) matrix (Marsupialia)

A) Pagel's lambda values (phylogenetic signal) of each character. B) Akaike weights support for the ER model of evolution of each character. C) Akaike weights support for the independent model of evolution of all pairwise comparisons of characters in each partition. D) Rates of character evolution of each character (log scale).


Pagel's lambda
1.00
0.75
0.50
0.25
0.00
dentition
skull
vertebrae
pectoral
pelvic
forelimb
hindlimb
soft
Akaike weights (ER model)
Akaike weights (independent model)
Rate of evolution
2.5
0.0
−2.5
−5.0
2
0
−2
−4
−6