

RH: Tempo and mode in mammals evolution

# **Extinction effects on niche occupancy in mammals after the K-Pg boundary**

THOMAS GUILLERME<sup>1,2\*</sup>, AND NATALIE COOPER<sup>1,2</sup>

<sup>1</sup>*School of Natural Sciences, Trinity College Dublin, Dublin 2, Ireland.*

<sup>2</sup>*Trinity Centre for Biodiversity Research, Trinity College Dublin, Dublin 2, Ireland.*

**\*Corresponding author.** *Zoology Building, Trinity College Dublin, Dublin 2, Ireland; E-mail: guillert@tcd.ie; Fax: +353 1 6778094; Tel: +353 1 896 2571.*

## Abstract

Massive global extinctions have a turn-over effect on biodiversity. When some large group of taxa suffers from a high rate of extinction, it is expected that niches becomes available for potentially unrelated clades that can undergo an adaptive radiation to fill these vacant niches. The causes and effects of such events are well understood for marine organisms with a good fossil record (e.g. Ammonoidea and Foraminifera) but the effects remains unclear on some iconic vertebrate groups.

Typically, placental mammals (eutherians) are shown by some studies to be undergoing an adaptive radiation after the Cretaceous-Palaeogene mass extinction event (K-Pg) by originating shortly before the K-Pg event and displaying high morphological evolutionary rates leading to high diversification during the Palaeogene. However, some other studies have demonstrated that eutherians originates during the Cretaceous and don't display significantly high diversification after the K-Pg event.

Here we propose a new approach to test if eutherians undergo an adaptive radiation after the K-Pg event. We use trees containing both living and fossil taxa based on all the available data (Total Evidence) and the state-of-the-art method in dating (tip dating) along side with a better proxy for niche occupancy (morphological diversity as opposed to taxonomic diversity) and finer grain analysis through time (introducing a time slicing method).

Our results shows that eutherians don't display significantly higher changes in morphological disparity that expected under a Brownian motion after the K-Pg boundary. We therefore propose that eutherian mammals don't undergo an

<sup>24</sup> adaptive radiation during the Palaeogene.

<sup>25</sup> (Keywords: disparity, diversity, punctuate equilibrium, constant evolution, time slicing)

## INTRODUCTION

Abiotic and biotic variations can have a cascade effect on biodiversity and lead to drastic changes in species richness and composition as well as changes in species ecological roles and dominant clades (Brusatte et al. 2008). Mass extinctions events can lead to a whole gradient of effects ranging from complete extinction of entire clades to an explosion in diversity and a diversification of ecological role in other ones (Erwin 1998). These changes in biodiversity can sometimes be linked to ecological niche replacement through time where species can either occupy the same niche through time or shift towards different ones (Pearman et al. 2008). In the case when niches remains stable through time, it is expected that extinctions can leave these niches vacant allowing other clade to speciate and occupy these vacant niches (Pearman et al. 2008). Mass extinction events can lead to biodiversity turn-overs where a clade that suffers a high level of extinctions leaves vacant ecological niches that can be filled by another unrelated clade that undergoes an adaptive radiation to fill these vacant ecological niches. For example Brachiopoda diversity declines at end-Permian mass extinction event and are rapidly replaced by Bivalvia as the dominant filter-feeding shelled organism displaying a pattern of adaptive radiation where the decline in diversity in one clade (Brachiopoda) leaves vacant ecological niches that can be occupied by an unrelated clade (Bivalvia) (e.g. Sepkoski 1981, Clapham et al. 2006 but see Payne et al. 2014).

However, both the timing of the extinction event and the radiation, as well as the

concept of vacant ecological niche are crucial to determine whether a clade truly displays an adaptive radiation pattern by filling vacant niches following the extinction event or if the observed radiation is due to a competitive radiation process where a clade replaces another through direct competition and not as a result of a mass extinction event (e.g. Brusatte et al. 2008; Benton et al. 2014). It is therefore crucial to understand the real tempo and mode of a clade's evolution through time to be able to determine if it radiates because of niche availability let vacant by previous extinctions.

The most recent mass extinction at the Cretaceous-Paleogene (K-Pg) boundary has been well studied with particular focus on the timing of events (asteroid impact at 66 Mya; Renne et al. 2013), the causes (e.g. Renne et al. 2013; Brusatte et al. 2015) and the effect on species extinction (e.g. Erwin 1998; Brusatte et al. 2015) as well as species diversification (e.g. Stadler 2011; Meredith et al. 2011; O'Leary et al. 2013). This mass extinction event is therefore well suited for testing whether groups that goes extinct leaves vacant niches allowing the surviving groups to undergo adaptive radiation. For example, the extinction of 90% of the Cretaceous Foraminifera species at the K-Pg boundary is followed by the radiation of surviving groups in the Palaeogene (D'Hondt et al. 1996; Coxall et al. 2006). However, because of the scarcity of the fossil record, the patterns are less clear in vertebrates groups, especially mammals (e.g. Meredith et al. 2011; O'Leary et al. 2013). Within Eutheria (crown and stem placental mammals), two hypothesis about the processes leading to diversification during the Tertiary are still debated. The first one (hereafter the "adaptive radiation" hypothesis), proposes that

Eutheria underwent an adaptive radiation after the K-Pg boundary (Goswami et al. 2011; Slater 2013; O’Leary et al. 2013): although they originated during the late Cretaceous, they only started diversifying during the Palaeogene due to the extinction of many terrestrial vertebrates at the K-Pg boundary especially non-avian dinosaurs (Brusatte et al. 2015). The second hypothesis (hereafter the “constant evolution” hypothesis) proposes that the K-Pg event had no drastic effect on diversification within Eutheria (Meredith et al. 2011; Stadler 2011; Beck and Lee 2014). These two opposite views on the eutherian radiation can probably be linked to difference in the data and the methods used.

The adaptive radiation view, often proposed by palaeontologists, is mainly based on a thorough study of the fossil record (and especially the absence of observed Placental mammals in the Cretaceous) and cladistic-like methods (e.g. Goswami et al. 2011; O’Leary et al. 2013). On the other hand, the constant evolution view, often proposed by neontologists, is based on molecular data (and indirect estimations of the fossil record) and probabilistic methods (e.g. Bininda-Emonds et al. 2007; Stadler 2011; Meredith et al. 2011). Also, it is important to note that, in the context of testing whether a clade is undergoing an adaptive radiation due to niche vacancy, it is important to use accurate proxies for measuring niche occupancy since the concept remains debated (Pearman et al. 2008). Previous studies have often only used phylogenetic diversity as a way to determine the tempo and mode of evolution in mammals (e.g. Stadler 2011) as opposed to measuring both phylogenetic diversity and morphological disparity (Slater

et al. 2010; Ruta et al. 2013; Price 2013). Therefore, one solution for resolving this apparent conflict is to use all the data available with the state-of-the art methods in phylogenetics along with both phylogenetic diversity and morphological disparity as proxies for determining niche occupancy.

In this study, we propose an updated approach to test whether mammals underwent an adaptive radiation during the Palaeogene due to the K-Pg extinction and niche vacancy. We measured the phylogenetic diversity and the the cladistic morphological disparity from two previously published studies (Slater 2013; Beck and Lee 2014) based on both molecular and morphological data and using the tip-dating method (Ronquist et al. 2012). We measured the variation of both diversity and disparity through time using a new time-slicing approach that provides a fine grain estimation of the diversity and disparity through time and also allows to make different assumption on the evolutionary models underlying the morphological characters evolution (either constant or punctuated). Finally, to test whether mammals display a significant change in diversity and disparity after the K-Pg boundary, we compared the observed changes to a null model to determine whether the group underwent an adaptive radiation. We found that there is no significant increase in cladistic disparity after the K-Pg event and that actually the disparity in mammals piked during the K-Pg event. These results suggest that the shift in dominant terrestrial vertebrate clades in the vertebrate fossil record during the Tertiary (from non-avian dinosaurs to eutherian mammals) is not due to the K-Pg extinctions but follows a constant evolution model

and is probable to be linked to a long term competition with other terrestrial vertebrates (Brusatte et al. 2008).

## METHODS

We tested whether mammals display a significant change in disparity after the K-Pg boundary by using the following protocol (note that each step is explained in detail below this section; Fig. 1).

1. Data: we used the cladistic morphological matrices and the Total Evidence tip-dated trees published in Slater (2013) and Beck and Lee (2014).
2. Estimating ancestral character states: we used the Yang et al. (1995) re-rooting method to estimate the states of each characters in the cladistic morphological matrices at each node and created the reconstructed morphological matrix containing observed morphological characters data for tips in the tree and estimated morphological characters states for nodes.
3. Calculating the cladistic-space: using the reconstructed morphological matrix we defined the cladistic-space by using a principal components ordination of the Gower distance matrix (Gower 1971) representing the  $n$  finite dimensional space that encompasses the total cladistic morphological variation of the taxa present in the analysis.



- 128 4. Time slicing: we then separated the cladistic-space into subdivisions containing  
129 only the edges (nodes or/and tips) present every 5 million years from present  
130 (hereafter called "time-slices").
- 131 5. Disparity through time: we then calculate the phylogenetic diversity as the  
132 number of edges present at each time-slice as well as the cladistic morphological  
133 disparity defined as the spread of the edges in the cladistic-space at each  
134 time-slice (distance from cladistic-space centroid (Finlay and Cooper 2015)).
- 135 6. Null model testing: finally, we compared the observed disparity to the expected  
136 disparity under two different null models: (1) the first one being a completely  
137 stochastic model where disparity is random at each point in time.; and (2) the  
138 second one being a  $Mk_{\eta}$  model (as a proxy for Brownian evolution) where  
139 disparity increase is constant through time.

## 140 *Data*

141 For exploring the changes in mammalian cladistic disparity through time, we used the  
142 cladistic morphological matrices and the Total Evidence tip dated trees (Ronquist et al.  
143 2012) from Slater (2013) (103 taxa and 446 morphological characters) and Beck and Lee  
144 (2014) (102 taxa and 421 morphological characters). Both studies respectively ranges  
145 from the 310 Million years ago (Mya - Late Carboniferous) and 170 Mya (Middle  
146 Jurassic) to the present and focuses on Mammaliaforma up to the family level Slater  
147 (2013) and on Eutheria up to a genus level Beck and Lee (2014). We chose these two

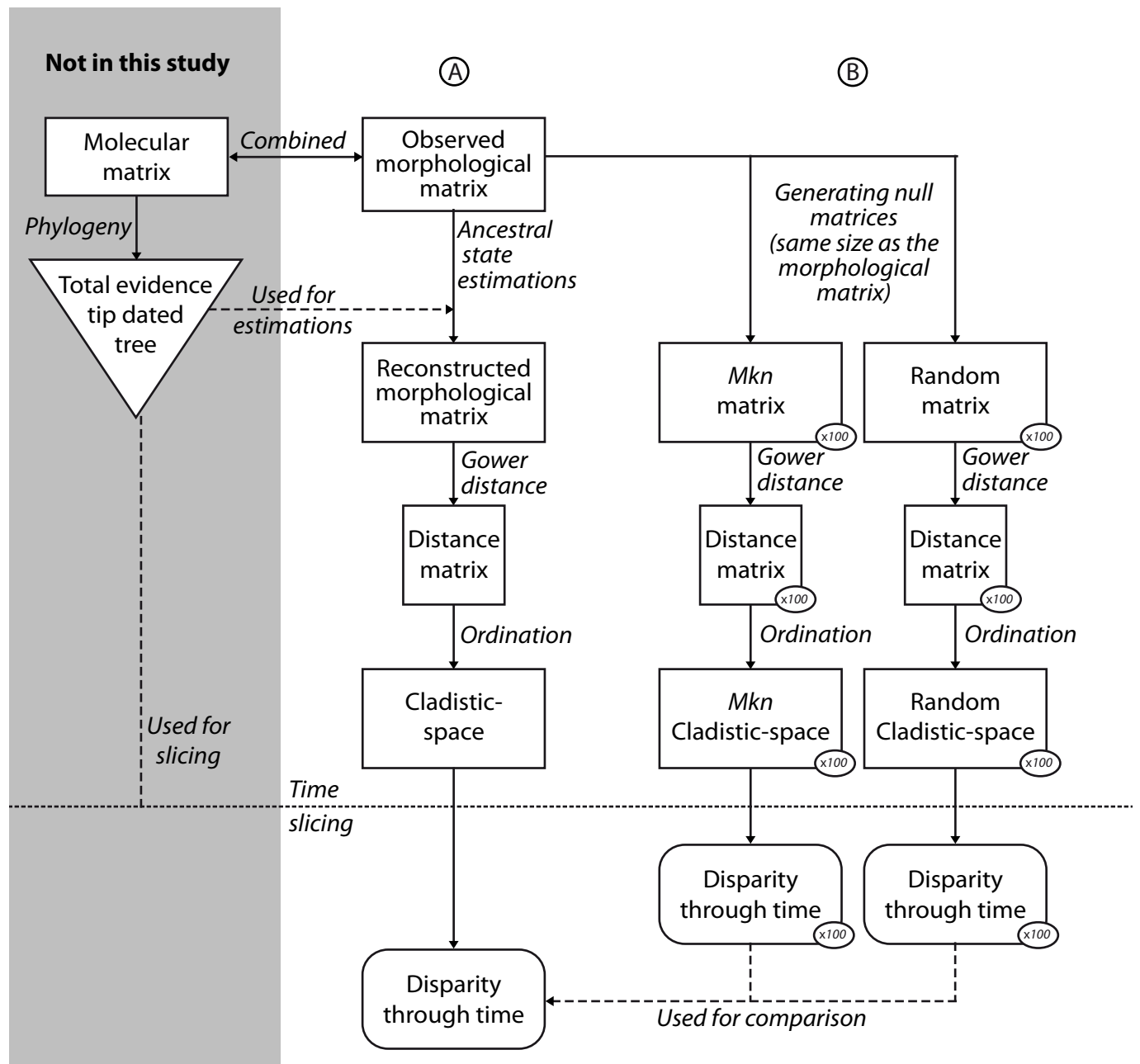


Figure 1: General method outline. The grey section represents the Total Evidence data and tip dated trees from two published studies (Slater 2013; Beck and Lee 2014). **A:** We then estimated the ancestral characters states from the observed morphological matrices from both studies. we then calculated the distance matrix using the Gower distance (Gower 1971) and performed a ordination of the distance matrix to create the cladistic-space. Finally, we divided the cladistic-space matrix using the Time slicing method to measure the changes in morphological disparity through time. **B:** we then generated two sets of a hundred simulated "null" matrices under two null models, the random

data sets because they both have similar cladistic properties (i.e. similar number of taxa and morphological characters) and both use the same phylogenetic method (Total Evidence tip dated trees; Ronquist et al. 2012). However, both matrix differs on their taxonomic levels (Family *vs.* Genus), their taxonomic group (Mammaliaforma *vs.* Eutheria) and their temporal range (310-0 Mya *vs.* 170-0 Mya). To include the temporal span of each taxa in our analysis, we collected all the first and last occurrences datum for each taxa. This data was collected from the original publications (Slater 2013; Beck and Lee 2014).

### *Estimating ancestral character states*

For both matrices (Slater 2013; Beck and Lee 2014), we estimated the ancestral state of each character at every node in the tree. We used a Maximum Likelihood approach to estimate the ancestral states. We preferred Maximum Likelihood methods on Bayesian methods because both methods have been shown to produce similar results Manuela et al. (2013) but Maximum Likelihood methods are several orders of magnitude faster than Bayesian ones (total analysis time 1.2 CPU year). We used the `rerootingMethod` function from the R package `phytools` 0.4-45 (Revell 2012). This method performs a Maximum Likelihood estimation of the ancestral values and the variance of a Brownian motion process based on the re-rooting method of Yang et al. (1995). We followed Lloyd (2015) method for treating missing data at the tips as any possible observed state (e.g. if a character has two observed tips states 0 and 1 and a third tip has missing data (NA),

the third tip will be considered as multi-state character (0&1) and the ancestor of these three tips can be estimated between the two following states: 0 and 1).

### *Calculating the cladistic-space*

To explore the variation in mammals shapes through time, we use a cladistic-space approach Foote (1994, 1996); Wesley-Hunt (2005); Brusatte et al. (2008); Friedman (2010); Toljagic and Butler (2013). Similarly to a morphospace based on continuous morphological data (e.g. Finlay and Cooper (2015)), the cladistic-space is based on cladistic morphological data. Therefore, the cladistic-space is likely to be more influenced by the phylogeny than the morphospace (Foote 1996; Wagner 1997). However, discrete cladistic characters are still the best source for quantifying overall morphology for large and diverse groups (Brusatte et al. 2008).

The first step for creating the cladistic-space is to calculate the pairwise distance matrix between all taxa. We calculated the Gower distance (Gower 1971) rescaled using the arcsine of the square root (Lloyd 2015). We also calculated the Generalised Euclidean Distance (GED) Wills (2001) and the MORD distance (@Lloyd in prep.@) to make our results comparable to other previous studies. The results using both distances are available in the supplementary materials. To avoid problems due to inapplicable characters and missing data and remove NAs from the distance matrix, we used the TrimMorphDistMatrix algorithm from the Claddis R package (Lloyd 2015). This algorithm removes the taxa that contains the most non-calculable distance until the distance matrix contains no NAs. We then ordinated the distance matrix using principal

components ordination (PCO; Gower 1966) to summarize the matrix and create the  $n$  dimensions of our cladistic-space. We used a PCO approach to be able to better handle missing data and inapplicable characters Lofgren et al. (2003); Wesley-Hunt (2005).

### *Time slicing*

We propose a new approach to look at the changes in disparity through time by calculating the spread of the cladistic-space at any time point in the phylogeny rather than considering the spread of the cladistic-space in time intervals (e.g. Brusatte et al. 2008, 2012; Toljagic and Butler 2013). In fact, the time intervals approach can suffer from several biases namely: (1) the time intervals duration is set arbitrarily and can distort the diversity or disparity in that time interval (for example, if stratigraphical ages are used, it is likely to increase the differences between two strata since they are traditionally based on fossils morphological differences); (2) the time intervals method assumes that all characters evolve following a punctuated equilibrium model (i.e. if a character between two time intervals, the character state is assumed to be constant during the whole time of the interval and then evolve quickly to reach the state in the second time interval; Gould and Eldredge 1977).

To address these issues, we propose a time slicing method that only considers subdivisions of the cladistic-space at specific equidistant points in time (as opposed to the time intervals that considers subdivisions of the cladistic-space between two points in time; see Fig. 2). This methods addresses two of the biases induced by the time intervals method: (1) it allows to sample evenly through time elements in the

phylogeny without arbitrarily grouping certain edges together. For example, the slice at 0 Mya will contain only the extant taxa instead of an interval defined as the Quaternary (from 2.5 Mya to the present) that will also include extinct taxa. However, it is highly probable for vertebrate data that no data is available at any point in time. Therefore we propose two models of character evolution to estimate the state of a character at any point in time along a branch, addressing at the same time the second potential bias in the time intervals method: (2) at any specific point in time, a branch is sampled instead of an edge, we propose two different models to determine whether to consider the DATA from the ancestor edge or the descendant one.

1. Punctuated model: this model selects either the data from the ancestor or the descendant edge. Similarly to the time interval method, this model reflects punctuated evolution where the changes occur either at the start or at the end of a branch. In practice, we randomly chose the data from the ancestor or the descendant edge each time. However, we also tested the model with the assumption that changes always occurs early on the branch (accelerate transition, ACCTRAN) or are always delayed towards the descendant edge (delayed transition, DELTRAN). The result of both assumption are available in the supplementary material.

2. Constant model: this models selects either the data from the ancestor or the descendant edge according to the branch length separating the sampling point and the ancestor's edge. If the sampling point along the branch is lower than half

the branch length, then the data from the ancestor edge is selected. Else we selected the data from the descendant edge. This model reflects constant evolution where changes rate are constant and cumulative along the branch (i.e. the hypothetical character state at any point in time will likely to be the same to the closest observed state in time - ancestor or descendant edge).

We applied this time slicing method to both cladistic-spaces (Slater 2013; Beck and Lee 2014) using a time interval of 5 Mya (from present to 170 Mya) slicing each cladistic-space in 35 subdivisions. We also performed a time intervals approach to compare our results to previous publications (e.g. Brusatte et al. 2008, 2012; Toljagic and Butler 2013). The results of this analysis are available in the supplementary materials.

### *Disparity through time*

*Diversity.*— We counted phylogenetic diversity as the number of phylogenetic elements in each subdivision of the cladistic-space (i.e. branches, nodes and tips per slice). We logged our diversity measurements to make the differences additive instead of multiplicatives.

*Disparity.*— We measured the cladistic morphological disparity (i.e. the diversity in discrete morphological characters) by using the distance from centroid (Finlay and Cooper 2015). The distance from centroid is defined as the mean euclidean distance between each edge (nodes and tips) and the mean value of each dimension of the

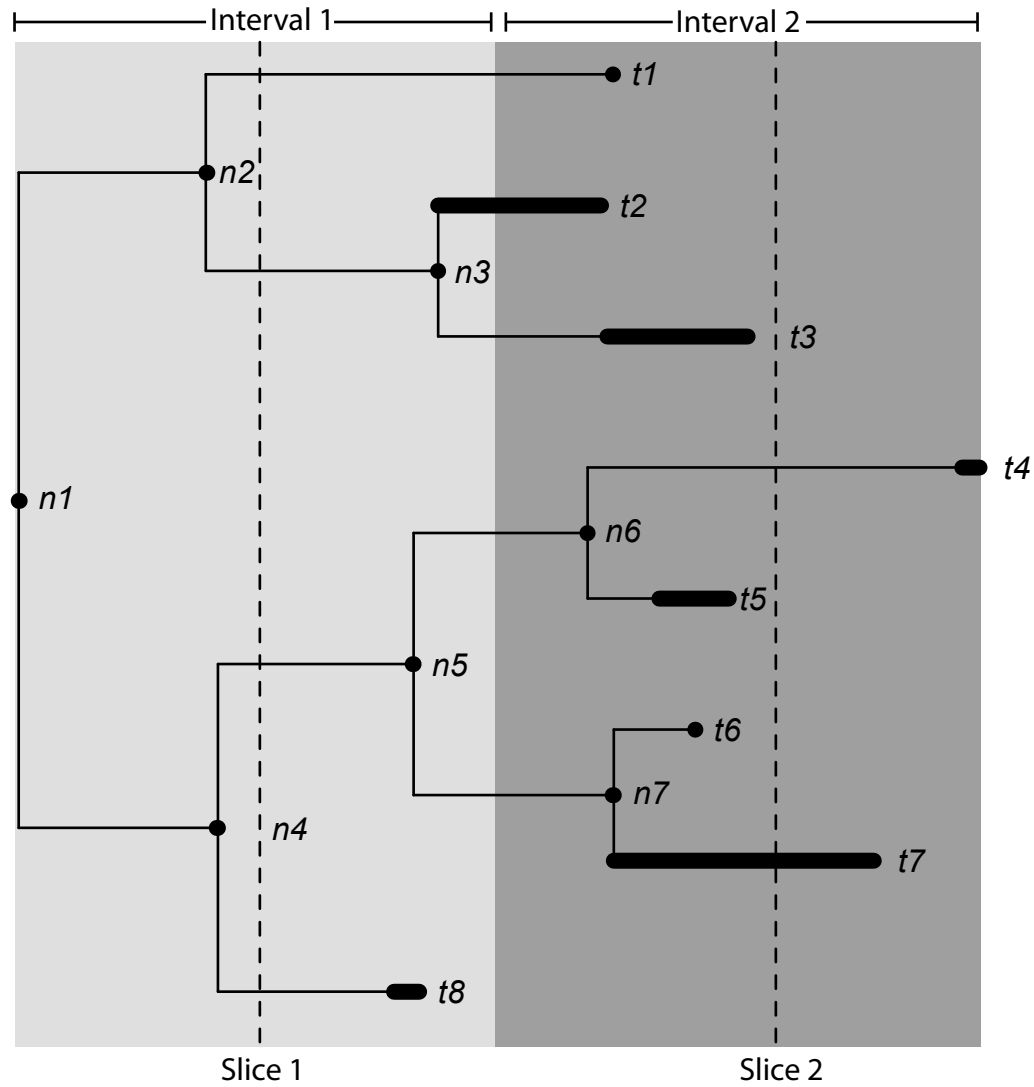


Figure 2: Differences in slicing method and intervals. Solid lines represent the First and Last Appearance Datum span. Interval 1 contains the following elements within the global character-space: taxa t2 and t8 and nodes n1, n2, n3, n4 and n5. Interval 2 contains the following elements within the global character-space: taxa t1, t2, t3, t4, t5, t6 and t7 and nodes n6 and n7. Slice 1 contains the following elements within the global character-space under the proximity model (constant evolution assumption): nodes n2 and n4. Slice 2 contains the following elements within the global character-space under the proximity model (constant evolution assumption): taxa t4 and t7.



cladistic-space:

$$Disparity = \frac{\sqrt{\sum (Edge_n - Centroid_n)^2}}{Number\ of\ edges} \quad (1)$$

Where  $Edge_n$  is any edge value in the  $n^{th}$  dimension of the cladistic-space and  $Centroid_n$  is the average value of the  $n^{th}$  dimension. Note that the disparity is calculated using all the  $n$  dimensions of the cladistic-space (i.e. we used all axis of the ordination).

We used this metric for measuring disparity rather than the classically used sum and product of ranges and variance (e.g. Wills et al. 1994; Foote 1996; Wesley-Hunt 2005; Brusatte et al. 2008; Ruta et al. 2013) because this method is less correlated with diversity (see supplementary) and allows to use the data of all the dimensions of the cladistic-space without being incalculable for the dimensions that bear a little amount of the distance variance. For example, the product of range and variance are always equal to zero if the last dimensions of the character space bare a minimal amount of the variance of the distance matrix (traditionally less than 5%). Also this method is clearer and easier to interpret (Finlay and Cooper 2015). For comparing our results to previous studies, we also calculated the sum and product of ranges and variance, these results are available in the supplementary materials.

For each disparity measurement, we bootstrapped 1000 times the subdivision of the cladistic-space containing only the edges present at each time slice. We then calculated the 50% and 95% confidence interval associated with each disparity measurement.

*Rarefaction.*— To avoid bias due to the number of edges variation in each cladistic-space

subdivision we also re-sampled each subdivision to have an equal number of phylogenetic elements available (edges or branches). Each subdivision was therefore re-sampled to contain the minimal number of taxa in the smallest subdivision. Because this led in certain cases to a absolute minimum of phylogenetic elements to calculate disparity (3) we also arbitrarily re-sampled the subdivisions to contain 10 elements or less. In practice, for both cases, we bootstrapped the disparity measurements by randomly removing  $x$  elements per subdivision, where  $x$  is either (1) the number of edges sampled in each time slice minus the minimum of edges in each time slice or (2) the number of edges to remove so that the number of elements is not higher than 10. Both results are available in the supplementary materials.

### *Null models testing*

Finally, we tested the effect of the K-Pg boundary on mammal disparity, by measuring whether the observed changes were significantly different than expected by chance. We rerun the full disparity analysis (minus the ancestral states reconstruction part) on both phylogenies with the same number of edges and the same time slices. However, we replaced the observed character matrices by two times 100 randomly generated ones following two null models. The first null model (repeated 100 times) was a purely stochastic matrix where each cell value was sampled from a discrete uniform distribution with a number of sates equal to the number of states observed in the the original data. For the second null model, we simulated each characters for each edge using the phylogenetic structure using the `sim.character` function from the

diversitree R package (FitzJohn 2012) and the  $Mk_n$  (Lewis 2001) with an equal transition rate between the number of characters states observed in the original matrix and a unique rate ( $\mu$ ) sampled from a uniform distribution ( $0 < \mu \leq 0.5$ ). We then measured the amount of overlap between both null models and the observed data using the Bhattacharyya Coefficient (Bhattacharyya 1943) which measures the probability of overlap between two distributions (similarly as a two sided *t-test*; Gui. We considered the observed data to be non different than random for a Bhattacharyya Coefficient  $>0.95$  and on the opposite, we considered a significant difference in disparity than expect by chance when the Bhattacharyya Coefficient was  $<0.05$ .

## RESULTS

## DISCUSSION

## CONCLUSION

## DATA AVAILABILITY AND REPRODUCIBILITY

## ACKNOWLEDGMENTS

Graeme Lloyd and Gavin Thomas.

## FUNDING

This work was funded by a European Commission CORDIS Seventh Framework Programme (FP7) Marie Curie CIG grant (proposal number: 321696).

\*

## References

???? Effects of missing data on topological inference using a total evidence approach, author=Guillerme, Thomas and Cooper, Natalie, journal=in review, year=2015, .

Beck, R. M. and M. S. Lee, 2014. Ancient dates or accelerated rates? Morphological clocks and the antiquity of placental mammals. *Proceedings of the Royal Society B: Biological Sciences* 281:1–10. URL <http://dx.doi.org/10.1098/rspb.2014.1278>.

Benton, M. J., J. Forth, and M. C. Langer, 2014. Models for the rise of the dinosaurs. *Current biology* : CB 24:R87–R95. URL <http://dx.doi.org/10.1016/j.cub.2013.11.063>.

Bhattacharyya, A., 1943. On a measure of divergence between two statistical populations defined by their probability distributions. *Bulletin of the Calcutta Mathematical Society* 35:99–109.

Bininda-Emonds, O. R., M. Cardillo, K. E. Jones, R. D. MacPhee, R. M. Beck, R. Grenyer, S. A. Price, R. A. Vos, J. L. Gittleman, and A. Purvis, 2007. The delayed rise of

present-day mammals. *Nature* 446:507–512. URL

<http://dx.doi.org/10.1038/nature05634>.

Brusatte, S., R. J. Butler, A. Prieto-Márquez, and M. A. Norell, 2012. Dinosaur morphological diversity and the end-Cretaceous extinction. *Nature Communications* 3:804–804. URL <http://dx.doi.org/10.1038/ncomms1815>.

Brusatte, S. L., M. J. Benton, M. Ruta, and G. T. Lloyd, 2008. Superiority, competition, and opportunism in the evolutionary radiation of dinosaurs. *Science* 321:1485–1488. URL <http://www.sciencemag.org/content/321/5895/1485.abstract>.

Brusatte, S. L., R. J. Butler, P. M. Barrett, M. T. Carrano, D. C. Evans, G. T. Lloyd, P. D. Mannion, M. A. Norell, D. J. Peppe, P. Upchurch, and T. E. Williamson, 2015. The extinction of the dinosaurs. *Biological Reviews* 90:628–642. URL <http://dx.doi.org/10.1111/brv.12128>.

Clapham, M. E., D. J. Bottjer, C. M. Powers, N. Bonuso, M. L. Fraiser, P. J. Marenco, S. Q. Dornbos, and S. B. Pruss, 2006. Assessing the ecological dominance of phanerozoic marine invertebrates. *PALAIOS* 21:431–441. URL <http://palaaios.sepmonline.org/content/21/5/431.abstract>.

Coxall, H. K., S. D'Hondt, and J. C. Zachos, 2006. Pelagic evolution and environmental recovery after the cretaceous-paleogene mass extinction. *Geology* 34:297–300. URL <http://geology.gsapubs.org/content/34/4/297.abstract>.

D'Hondt, S., T. D. Herbert, J. King, and C. Gibson, 1996. Planktic foraminifera,

asteroids, and marine production: Death and recovery at the cretaceous-tertiary  
boundary. Geological Society of America Special Papers 307:303–317. URL  
<http://specialpapers.gsapubs.org/content/307/303.abstract>.

Erwin, D. H., 1998. The end and the beginning: recoveries from mass extinctions.  
Trends in Ecology and Evolution 13:344 – 349. URL  
<http://www.sciencedirect.com/science/article/pii/S0169534798014360>.

Finlay, S. and N. Cooper, 2015. Morphological diversity in tenrecs (afrosoricida,  
tenrecidae): comparing tenrec skull diversity to their closest relatives. Tech. rep.,  
PeerJ PrePrints.

FitzJohn, R. G., 2012. Diversitree: comparative phylogenetic analyses of diversification  
in R. Methods in Ecology and Evolution 3:1084–1092. URL  
<http://dx.doi.org/10.1111/j.2041-210X.2012.00234.x>.

Foote, M., 1994. Morphological disparity in ordovician-devonian crinoids and the early  
saturation of morphological space. Paleobiology 20:320–344. URL  
<http://paleobiol.geoscienceworld.org/content/20/3/320.short>.

———, 1996. Ecological controls on the evolutionary recovery of post-paleozoic  
crinoids. Science 274:1492–1495. URL  
<http://www.sciencemag.org/content/274/5292/1492.abstract>.

Friedman, M., 2010. Explosive morphological diversification of spiny-finned teleost  
fishes in the aftermath of the end-Cretaceous extinction. Proceedings of the Royal

Society B: Biological Sciences 277:1675–1683. URL

<http://dx.doi.org/10.1098/rspb.2009.2177>.

Goswami, A., G. V. Prasad, P. Upchurch, D. M. Boyer, E. R. Seiffert, O. Verma,  
E. Gheerbrant, and J. J. Flynn, 2011. A radiation of arboreal basal eutherian mammals  
beginning in the late cretaceous of india. Proceedings of the National Academy of  
Sciences of the United States of America 108:16333–16338. URL

<http://dx.doi.org/10.1073/pnas.1108723108>.

Gould, S. J. and N. Eldredge, 1977. Punctuated equilibria: The tempo and mode of  
evolution reconsidered. Paleobiology 3:pp. 115–151. URL

<http://www.jstor.org/stable/2400177>.

Gower, J. C., 1966. Some distance properties of latent root and vector methods used in  
multivariate analysis. Biometrika 53:325–338. URL

<http://biomet.oxfordjournals.org/content/53/3-4/325.abstract>.

———, 1971. A general coefficient of similarity and some of its properties. Biometrics  
27:pp. 857–871. URL <http://www.jstor.org/stable/2528823>.

Lewis, P., 2001. A likelihood approach to estimating phylogeny from discrete  
morphological character data. Systematic Biology 50:913–925. URL

<http://dx.doi.org/10.1080/106351501753462876>.

Lloyd, G. T., 2015. Claddis: an r package for performing disparity and rate analysis on

cladistic-type data sets. <https://github.com/graemetlloyd/Claddis>. URL

<https://github.com/graemetlloyd/Claddis>.

Lofgren, A. S., R. E. Plotnick, Wagner, and P. J., 2003. Morphological diversity of carboniferous arthropods and insights on disparity patterns through the phanerozoic. *Paleobiology* 29:349–368. URL <http://www.psychology.org/doi/abs/10.1666/0094-8373%282003%29029%3C0349%3AMD0CAA%3E2.0.CO%3B2>.

Manuela, R., P. Pontarotti, and G. Didier, 2013. Choosing the best ancestral character state reconstruction method. *Mathematical biosciences* 242:95–9109. URL <http://dx.doi.org/10.1016/j.mbs.2012.12.003>.

Meredith, R., J. Janečka, J. Gatesy, O. Ryder, C. Fisher, E. Teeling, A. Goodbla, E. Eizirik, T. L. Simão, T. Stadler, D. Rabosky, R. Honeycutt, J. Flynn, C. Ingram, C. Steiner, T. Williams, T. Robinson, B. Angela, M. Westerman, N. Ayoub, M. Springer, and W. Murphy, 2011. Impacts of the Cretaceous terrestrial revolution and KPg extinction on mammal diversification. *Science* 334:521–524. URL <http://dx.doi.org/10.1126/science.1211028>.

O’Leary, M. A., J. I. Bloch, J. J. Flynn, T. J. Gaudin, A. Giallombardo, N. P. Giannini, S. L. Goldberg, B. P. Kraatz, Z.-X. Luo, J. Meng, X. Ni, M. J. Novacek, F. A. Perini, Z. S. Randall, G. W. Rougier, E. J. Sargis, M. T. Silcox, N. B. Simmons, M. Spaulding, P. M. Velazco, M. Weksler, J. R. Wible, and A. L. Cirranello, 2013. The placental mammal



402 ancestor and the postK-Pg radiation of placentals. *Science* 339:662–667. URL

403 <http://www.sciencemag.org/content/339/6120/662.abstract>.

404 Paradis, E., J. Claude, and K. Strimmer, 2004. APE: analyses of phylogenetics and  
405 evolution in R language. *Bioinformatics* 20:289–290.

406 Payne, J. L., N. A. Heim, M. L. Knope, and C. R. McClain, 2014. Metabolic dominance  
407 of bivalves predates brachiopod diversity decline by more than 150 million years.

408 *Proceedings of the Royal Society B: Biological Sciences* 281. URL

409 <http://rspb.royalsocietypublishing.org/content/281/1783/20133122.abstract>.

410 Pearman, P. B., A. Guisan, O. Broennimann, and C. F. Randin, 2008. Niche dynamics in  
411 space and time. *Trends in Ecology and Evolution* 23:149 – 158. URL

412 <http://www.sciencedirect.com/science/article/pii/S0169534708000372>.

413 Price, S. A., 2013. Decoupling of taxonomic diversity and morphological disparity  
414 during decline of the cambrian trilobite family pterocephaliidae. *Journal of*

415 *Evolutionary Biology* 26:1665–1676. URL <http://dx.doi.org/10.1111/jeb.12164>.

416 Renne, P. R., A. L. Deino, F. J. Hilgen, K. F. Kuiper, D. F. Mark, W. S. Mitchell, L. E.

417 Morgan, R. Mundil, and J. Smit, 2013. Time scales of critical events around the

418 Cretaceous-Paleogene boundary. *Science* 339:684–687. URL

419 <http://dx.doi.org/10.1126/science.1230492>.

420 Revell, L. J., 2012. phytools: An r package for phylogenetic comparative biology (and  
421 other things). *Methods in Ecology and Evolution* 3:217–223.

422 Ronquist, F., S. Klopstein, L. Vilhelmsen, S. Schulmeister, D. Murray, and A. Rasnitsyn,  
 423 2012. A total-evidence approach to dating with fossils, applied to the early radiation  
 424 of the Hymenoptera. *Systematic Biology* 61:973–999. URL  
 425 <http://dx.doi.org/10.1093/sysbio/sys058>.

426 Ruta, M., K. D. Angielczyk, J. Fröbisch, and M. J. Benton, 2013. Decoupling of  
 427 morphological disparity and taxic diversity during the adaptive radiation of  
 428 anomodont therapsids. *Proceedings of the Royal Society of London B: Biological*  
 429 *Sciences* 280.

430 Sepkoski, J., J. John, 1981. A factor analytic description of the phanerozoic marine fossil  
 431 record. *Paleobiology* 7:pp. 36–53. URL <http://www.jstor.org/stable/2400639>.

432 Slater, G. J., 2013. Phylogenetic evidence for a shift in the mode of mammalian body  
 433 size evolution at the cretaceous-palaeogene boundary. *Methods in Ecology and*  
 434 *Evolution* 4:734–744. URL <http://dx.doi.org/10.1111/2041-210X.12084>.

435 Slater, G. J., S. A. Price, F. Santini, and M. E. Alfaro, 2010. Diversity versus disparity and  
 436 the radiation of modern cetaceans. *Proceedings of the Royal Society of London B:*  
 437 *Biological Sciences* .

438 Stadler, T., 2011. Mammalian phylogeny reveals recent diversification rate shifts.  
 439 *Proceedings of the National Academy of Sciences* 108:6187–6192. URL  
 440 <http://www.pnas.org/content/108/15/6187.abstract>.

441 Toljagic, O. and R. J. Butler, 2013. Triassic-Jurassic mass extinction as trigger for the

mesozoic radiation of crocodylomorphs. *Biology Letters* 9. URL

<http://dx.doi.org/10.1098/rsbl.2013.0095>.

Wagner, P. J., 1997. Patterns of morphologic diversification among the rostroconchia.

*Paleobiology* 23:115–150. URL

<http://paleobiol.geoscienceworld.org/content/23/1/115.short>.

Wesley-Hunt, G. D., 2005. The morphological diversification of carnivores in north

america. *Paleobiology* 31:35–55. URL [http://www.psjournals.org/doi/abs/10.](http://www.psjournals.org/doi/abs/10.1666/0094-8373%282005%29031%3C0035%3ATMDOCI%3E2.0.CO%3B2)

[1666/0094-8373%282005%29031%3C0035%3ATMDOCI%3E2.0.CO%3B2](http://www.psjournals.org/doi/abs/10.1666/0094-8373%282005%29031%3C0035%3ATMDOCI%3E2.0.CO%3B2).

Wills, M., 2001. Morphological disparity: A primer. Pp. 55–144, in J. M. Adrain, G. D.

Edgecombe, and B. S. Lieberman, eds. *Fossils, Phylogeny, and Form, Topics in*

*Geobiology*, vol. 19. Springer US. URL

[http://dx.doi.org/10.1007/978-1-4615-0571-6\\_4](http://dx.doi.org/10.1007/978-1-4615-0571-6_4).

Wills, M. A., D. E. G. Briggs, and R. A. Fortey, 1994. Disparity as an evolutionary index:

A comparison of cambrian and recent arthropods. *Paleobiology* 20:93–130. URL

<http://www.jstor.org/stable/2401014>.

Yang, Z., S. Kumar, and M. Nei, 1995. A new method of inference of ancestral

nucleotide and amino acid sequences. *Genetics* 141:1641–50. URL

<http://www.genetics.org/content/141/4/1641.abstract>.

## SUPPLEMENTARIES

## Ancestral states estimation

We used both the `ace` function from the R package `ape` v. 3.2 Paradis et al. (2004) and the `rerootingMethod` function from the R package `phytools` 0.4-45 Revell (2012). Both methods perform a maximum likelihood estimation of the ancestral values and the variance of a Brownian motion process based on the re-rooting method of Yang et al. (1995). The two methods differ slightly in the calculation of the normalized conditional likelihoods but mainly on the way to treat missing data. We optimised the `ace` function for fast estimation by treating missing data in the matrix as an extra character (e.g. if a character has two observed tip states 0 and 1 and a third tip has missing data (NA), the ancestor of these three tips can be estimated between the three following states: 0, 1 and NA). For the `rerootingMethod`, we followed Lloyd (2015) method and treated the missing in the tips as any possible observed state (e.g. if a character has two observed tip states 0 and 1 and a third tip has missing data (NA), the third tip will be considered as multi-state (0&1) and the ancestor of these three tips can be estimated between the two following states: 0 and 1). Both methods perform similarly but the implementation of the `ace` function has a slightly lower accuracy but is three times faster than the one for the `rerootingMethod` function (see supplementaries).

*Time intervals.*— We then divide our observed cladistic-spaces into sub cladistic-spaces representing the different stages of the character-space filling. For example, if at various points in time. Time intervals from 100Mya (Earliest Cenomanian, Late Cretaceous) to the present. We count all the nodes/tips present in a given time interval. Classic but

482 artificially grouping data. The minimal bin size should contain at least two nodes/tips  
483 and sometime that involves having time intervals spanning accross tens of millions of  
484 years. Such long duration time intervals have no real biological meaning since it is  
485 unlikely that all of the nodes/tips present in the time interval did ever coexisted and  
486 had ever biological interactions together.

### 487 *Diversity*

488 -Diversity in living mammals -Diversity per interval

### 489 *Disparity*

490 -Centroid is less correlated with diversity -Other metrics