RH: Cretaceous-Palaeogene extinction does not affect mammalian disparity.

Cretaceous-Palaeogene extinction does not affect mammalian disparity.

THOMAS GUILLERME[1,2*], AND NATALIE COOPER[1,2,3]

[1]*School of Natural Sciences, Trinity College Dublin, Dublin 2, Ireland.*

[2]*Trinity Centre for Biodiversity Research, Trinity College Dublin, Dublin 2, Ireland.*

[3]*Department of Life Sciences, Natural History Museum, Cromwell Road, London, SW7 5BD, UK.*

*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

(Keywords: disparity, diversity, punctuated equilibrium, gradualism, time slicing)

INTRODUCTION

Throughout history, life on Earth has suffered a series of mass extinction events resulting in drastic declines in global biodiversity (e.g. Raup, 1979; Benton and Twitchett, 2003; Renne et al., 2013; Brusatte et al., 2015). However, the long-term effects of mass extinctions are more varied (Erwin, 1998), and include species richness increases in some clades (Friedman, 2010) or declines in others (Benton, 1985), changes in morphological diversity (Ciampaglio et al., 2001; Ciampaglio, 2004; Korn et al., 2013) and shifts in ecological dominance (e.g. Brusatte et al., 2008b; Toljagic and Butler, 2013; Benson and Druckenmiller, 2014). These shifts are characterized by the decline of one clade that is replaced by a different unrelated clade with a similar ecological role (e.g. Brachiopoda and Bivalvia at the end Permian extinction; Sepkoski 1981; Clapham et al. 2006 but see Payne et al. 2014). Shifts in ecological dominance are of particular interest because they are a fairly common pattern observed in the fossil record (e.g. Foraminifera; D'Hondt et al. 1996; Coxall et al. 2006; Ichtyosauria; Thorne et al. 2011; Plesiosauria; Benson and Druckenmiller 2014) and are often linked to major macroevolutionary processes such as adaptive (Losos, 2010) or competitive radiations (Brusatte et al., 2008b).

One classical example of a shift in ecological dominance is at the Cretaceous-Palaeogene (K-Pg) mass extinction 66 million years ago (Renne et al., 2013), where the non-avian dinosaurs went extinct, potentially leading to the "rise of the age of the mammals" (Archibald, 2011; Lovegrove et al., 2014). This is based on the idea

that placental mammals were able to diversify after the extinction of many terrestrial vertebrates at the K-Pg boundary (including the dominant non-avian dinosaur group; Luo 2007; Archibald 2011; O'Leary et al. 2013; Brusatte et al. 2015). Some authors suggest this reflects placental mammals filling the "empty" niches left after the K-Pg event (Archibald, 2011), others suggest it reflects a release from predation and/or competition (Lovegrove et al., 2014). However, evidence for the diversification of placental mammals after the K-Pg event is mixed. Thorough analysis of the fossil record (e.g. Goswami et al., 2011; O'Leary et al., 2013) supports the idea that placental mammals diversified after the K-Pg event as there are no undebated placental mammal fossils before the K-Pg event and many afterwards (Archibald, 2011; Goswami et al., 2011; Slater, 2013; O'Leary et al., 2013; Wilson, 2013; Brusatte et al., 2015). Conversely, evidence from molecular data suggests that the diversification of placental mammals started prior to the K-Pg extinction event without being drastically affected by it (e.g. Douady and Douzery 2003; Bininda-Emonds et al. 2007; Meredith et al. 2011; Stadler 2011 or Beck and Lee 2014 using morphological data as well). Therefore, whether the diversification of placental mammals began before the K-Pg event, or in response to the extinctions at the K-Pg event, is a matter of great debate (dos Reis et al., 2012; O'Leary et al., 2013; Springer et al., 2013; OLeary et al., 2013; dos Reis et al., 2014).

There are two main reasons why there is still debate about the timing of the diversification of placental mammals. In this paper we focus on solving these issues as follows:

4

1. **Palaeontological and neontological data show different patterns**. As mentioned above, conclusions about when placental mammals diversified tend to be split depending on what kind of data are used: palaeontological data generally suggest that placental mammals diversified post K-Pg (e.g. O'Leary et al., 2013), whereas neontological data suggest that K-Pg event had little to no effect on mammalian diversification (Bininda-Emonds et al., 2007; Meredith et al., 2011; Stadler, 2011). Fortunately a recently successfully implemented method allows to use cladistic data for both living and fossil taxa along with molecular data for living taxa (the Total Evidence method; Eernisse and Kluge, 1993; Ronquist et al., 2012). This method can also be combined with the tip-dating method (Ronquist et al., 2012; Wood et al., 2013) to get more accurate estimates of diversification times for both fossil and living species (but see Arcila et al., 2015). Recently, two study have been published using the Total Evidence and tip-dating methods to study (1) variation in mammalian body mass (Slater, 2013) and (2) diversification rates (Beck and Lee, 2014) around the K-Pg boundary. Slater (2013) found good support for a shift in mammalian body mass evolution pattern before and after the K-Pg boundary suggesting a clear effect of the K-Pg boundary on mammalian body mass diversification. Whereas, Beck and Lee (2014) found mixed result on diversification rates supporting both a diversification of placental mammals before ("ancient dates" hypothesis) or after ("accelerated rates" hypothesis) the K-Pg boundary.

2. **Diversity can be defined in different ways.** Diversity is a difficult concept to define. In many studies it is measured as phylogenetic diversity or species richness (Stadler, 2011; Meredith et al., 2011; O'Leary et al., 2013), but often the more interesting aspect of diversity is related to the ecological niches the species occupy (Wesley-Hunt, 2005; Brusatte et al., 2008b; Toljagic and Butler, 2013), particularly if we want to make hypotheses about macroevolutionary processes (Pearman et al., 2008; Olson and Arroyo-Santos, 2009; Losos, 2010; Glor, 2010). Sometimes phylogenetic diversity is used as a proxy for other kinds of diversity, however, species richness can be decoupled from morphological diversity (Slater et al., 2010; Ruta et al., 2013; Hopkins, 2013), so phylogenetic diversity may not be the best proxy for ecological diversity. For example in Slater (2013), the diversity of mammalian body mass is studied instead as the sheer number of species. In this particular case, body mass diversity rather than species diversity can be a better proxy for describing mammalian diversity. One can also use morphological diversity, also known as disparity (e.g. Wills et al., 1994; Erwin, 2007; Hughes et al., 2013), as a way to quantify changes in mammalian diversity that should relate to the ecology of the species. However some methods form measuring disparity are outdated and make inappropriate assumptions. Many studies quantifying changes in morphological diversity were proposed > 20 years ago (Foote, 1994; Wills et al., 1994) and are sometimes used without modifications (e.g., Brusatte et al., 2008a,b; Cisneros and Ruta, 2010; Thorne et al., 2011; Prentice

et al., 2011; Brusatte et al., 2012; Toljagic and Butler, 2013; Ruta et al., 2013; Benton et al., 2014; Benson and Druckenmiller, 2014), even when the statistical assumptions of the methods are violated (see Methods). Additionally, previous methods are based on an underlying assumption that changes in disparity occur by punctuated evolution (e.g. Wesley-Hunt, 2005) which is not always the case (Hunt et al., 2015). Finally, most studies of disparity through time use unequal time units based on biostratigraphy (Brusatte et al., 2008b, 2012; Toljagic and Butler, 2013). This can be tautological as biostratigraphy is already based on changes in fossil assemblages and morphology through time.

Here, we propose an updated approach to test whether mammals diversified before or after K-Pg, using morphological disparity, measured as cladistic disparity (see Methods), as our proxy for diversity. We measured the disparity of living and fossil mammals taken from two previously published studies (Slater, 2013; Beck and Lee, 2014). Using a novel time-slicing approach we produce fine-grain estimates of disparity through time under two different models of morphological character evolution (either gradual or punctuated). Finally, to test whether mammals display significant changes in disparity after the K-Pg boundary, we tested whether there is significant changes in disparity through time. When an effect of time was detected, we ran post-hoc tests to measure if there is any significant difference between the disparity at the end of the Cretaceous and the disparity throughout the Cenozoic. We found no significant changes in mammalian disparity between the end of the Cretaceous and any time

during the Cenozoic. These results suggest that the extinction of non-avian dinosaurs

any many other terrestrial vertebrate clades at the end of the Cretaceous did not affect

mammalian evolution.

Methods

Cladistic data and phylogenies

We used the cladistic morphological matrices and the Total Evidence tip-dated trees

(Ronquist et al., 2012) from Slater (2013, 103 taxa with 446 morphological characters)

and Beck and Lee (2014, 102 taxa with 421 morphological characters). We chose these

two data sets because they have a similar number of taxa and morphological characters.

Slater (2013) ranges from 310 million years ago (Mya; Late Carboniferous) to the

present and focuses on the clade Mammaliaformes at the family-level. Beck and Lee

(2014) ranges from 170 Mya (Middle Jurassic) to the present and focuses on Eutherians

at the genus-level. We used the first and last occurrences reported in Slater (2013) and

Beck and Lee (2014) as the temporal range of each taxon in our analysis. Both

phylogenies are illustrated in the supplementary material (see Figure S1 and S2 @@@).

Estimating ancestral character states

For both datasets we used the re-rooting method (Yang et al., 1996; Garland and Ives,

2000) to get Maximum Likelihood estimates of the ancestral states for each character at

every node in the tree, using the `rerootingMethod` function from the `R` package `phytools` (version 0.4-45; Revell, 2012; R Core Team, 2015). Where there was missing character data for a taxon we followed the method of Lloyd (2015) and treated missing data as any possible observed state for each character. For example, if a character had two observed states (0 and 1) across all taxa, we attributed the multi-state "0&1" value to the taxon with missing data, representing an equal probability of being either 0 or 1. This allows the ancestral node of a taxon with missing data to be estimated with no assumptions other than that the taxon has one of the observed character states. To prevent poor ancestral state reconstructions from biasing our results, especially when a lot of error is associated with the reconstruction, we only included ancestral state reconstructions with a scaled Likelihood ≥ 0.95. Ancestral state reconstructions with scaled Likelihoods below this threshold were replaced by missing data ("?").

Building the cladisto-space

To explore variations in mammalian disparity through time (defined here as the variation in morphologies through time), we use a cladisto-space approach (e.g. Foote, 1994, 1996; Wesley-Hunt, 2005; Brusatte et al., 2008b; Friedman, 2010; Toljagic and Butler, 2013; Hughes et al., 2013). This approach is similar to constructing a morphospace based on continuous morphological data (e.g. Friedman, 2010), except a cladisto-space is an approximation of the morphospace based on cladistic data (i.e. the discrete morphological characters used to build a phylogenetic tree). Mathematically, a cladisto-space is an n dimensional object that summarizes the cladistic distances

9

148 between the taxa present in a cladistic matrix (see details below). Hetherington et al.

149 (2015) have empirically shown inter-taxon distances are not different in a morphospace

150 or a cladisto-space. However, we prefer referring to this object as a cladisto-space to

151 make it clear that this space is estimated using cladistic data and not morphometric

152 data and because both objects have slightly different properties. In fact, because of its

153 inherent combinatory properties, a cladisto-space is a finite theoretical object limited by

154 the product of the number of character states (c.f. the morphospace that is an infinite

155 theoretical object). Thus a cladisto-space will be overloaded if the number of taxa is

156 higher than the product of the number of character states, although this is rarely an

157 issue with empirical data (our cladisto-spaces have maximal capacities of 1.9×10^{181}

158 taxa; Slater, 2013, and 4.5×10^{159} taxa; Beck and Lee, 2014).

159   To estimate the cladisto-spaces for each of our datasets we first constructed

160 pairwise distance matrices of length k, where k is the total number of taxa in the

161 dataset. For each dataset separately, we calculated the $k \times k$ distances using the Gower

162 distance (Gower, 1971), i.e. the Euclidean distance between two taxa divided by the

163 number of shared characters. This allows us to correct for distances between two taxa

164 that share many characters and could be closer to each other than to taxa with fewer

165 characters in common (i.e. because some pairs of taxa share more characters in

166 common than others, they are more likely to be similar). For cladistic matrices, using

167 this corrected distance is preferable to the raw Euclidean distance because of its ability

168 to deal with discrete or/and ordinated characters as well as with missing data

(Anderson and Friedman, 2012). However, the Gower distance cannot calculate

distances when taxa have no overlapping data. Therefore, we used the

`TrimMorphDistMatrix` function from the `Claddis` R package (Lloyd, 2015) to remove

pairs of taxa with no cladistic characters in common. This led us to remove 11 taxa

from Slater (2013) but none from Beck and Lee (2014).

After calculating our distance matrices we transformed them using classical

multidimensional scaling (MDS; Torgerson, 1965; Gower, 1966; Cailliez, 1983). This

method (also referred to as PCO; e.g. Brusatte et al. 2015; or PCoA; e.g. Paradis et al.

2004) is an eigen decomposition of the distance matrix. Because we used Gower

distances instead of raw Euclidean distances, negative eigenvalues can be calculated. To

avoid this problem, we first transformed the distance matrices by applying the Cailliez

correction (Cailliez, 1983) which adds a constant c^* to the values in a distance matrix

(apart from the diagonal) so that all the Gower distances become Euclidean

($d_{Gower} + c^* = d_{Euclidean}$; Cailliez 1983). We were then able to extract n eigenvectors for

each matrix (representing the n dimensions of the cladisto-space) where n is equal to

$k - 2$, i.e. the number of taxa in the matrix (k) minus the two last eigenvector which are

always null after applying the Cailliez correction. Contrary to previous studies (e.g

Brusatte et al., 2008a; Cisneros and Ruta, 2010; Prentice et al., 2011; Anderson and

Friedman, 2012; Hughes et al., 2013; Benton et al., 2014), we use all n dimensions of our

cladisto-spaces and not a sub-sample representing the majority of the variance in the

distance matrix (e.g. selecting only m dimensions that represent up to 90% of the

11

variance in the distance matrix; Brusatte et al. 2008b; Toljagic and Butler 2013).

Note that our cladisto-spaces represent an ordination of all possible mammalian morphologies coded in each study through time. It is unlikely that all morphologies will co-occur at each time point, therefore, the disparity of the whole cladisto-space is expected to be ≥ the disparity at any specific point in time.

Calculating disparity

Disparity can be estimated in many different ways (e.g. Wills et al., 1994; Ciampaglio, 2004; Thorne et al., 2011; Hopkins, 2013; Huang et al., 2015), however most studies estimate disparity using four metrics: the sum and products of ranges and variances, each of which gives a slightly different estimate of how the data fits within the cladisto-space (Foote, 1994; Wills et al., 1994; Brusatte et al., 2008a,b; Cisneros and Ruta, 2010; Thorne et al., 2011; Prentice et al., 2011; Brusatte et al., 2012; Toljagic and Butler, 2013; Ruta et al., 2013; Benton et al., 2014; Benson and Druckenmiller, 2014). The sum and products of ranges and variances are based on the ranges and variances of the eigenvectors calculated from a distance matrix. However, these metrics do not take into account the covariance among eigenvectors. This is only statistically valid if the eigenvectors are independent. In multidimensional scaling, all n eigenvectors are calculated from the same distance matrix and are therefore not independent, thus covariances among eigenvectors should be included when estimating disparity. Furthermore, range metrics are also strongly affected by the uneven sampling of the fossil record (Butler et al., 2012) Additionally, because we include all n dimensions in

12

the analysis (see above), the products of ranges and variances will tend towards zero

since the scores of the last dimension are usually really close to zero themselves. These

features make using the sum and products of ranges and variances unfeasible in our

study. Instead, we use a different metric that comes with no statistical assumptions for

measuring the dispersion of the data in the cladisto-space: the median distance from

centroid (similar but not equivalent to Wills et al. 1994; Korn et al. 2013; Huang et al.

2015) calculated as:

$$Disparity = Md \sqrt{\sum_{i=1}^{k} (\mathbf{v}_{n_i} - Centroid_n)^2} \tag{1}$$

where:

$$Centroid_n = \frac{\sum_{i=1}^{k} (\mathbf{v}_{n_i})}{k} \tag{2}$$

k is the size of the distance matrix (i.e. the total number of taxa), $\mathbf{v}_n$ is any of the n

eigenvectors (i.e. any of the n dimension of the cladisto-space), $Centroid_n$ is the centroid

euclidean distance of the nth eigenvector (equation 2) and Md is the median value of

the distance from centroid (equation 1). Note that we also calculated the sum and

products of ranges and variances in the supplementary material (@@@).

Estimating disparity through time

Changes in disparity through time are generally investigated by calculating the

disparity of taxa that occupy the cladisto-space during specific time intervals (e.g

Cisneros and Ruta, 2010; Prentice et al., 2011; Hughes et al., 2013; Hopkins, 2013;

₂₂₈ Benton et al., 2014; Benson and Druckenmiller, 2014). These time intervals are usually

₂₂₉ defined based on biostratigraphy (e.g. Cisneros and Ruta, 2010; Prentice et al., 2011;

₂₃₀ Hughes et al., 2013; Benton et al., 2014) but can also be arbitrarily chosen time periods

₂₃₁ of equal duration (Butler et al., 2012; Hopkins, 2013; Benson and Druckenmiller, 2014).

₂₃₂ However, this approach suffers from two main biases. First, if biostratigraphy is used to

₂₃₃ determine the time intervals, disparity may be distorted towards higher differences

₂₃₄ between time intervals because biostratigraphical periods are geologically defined

₂₃₅ based on differences in the morphology of fossils found in the different strata. Second,

₂₃₆ this approach assumes that all characters evolve following a punctuated equilibrium

₂₃₇ model, because disparity is only estimated once for each interval resulting in all

₂₃₈ changes in disparity occurring between intervals, rather than also allowing for gradual

₂₃₉ changes within intervals (Hunt et al., 2015).

₂₄₀   To address these issues, we use a "time-slicing" approach that considers subsets

₂₄₁ of taxa in the cladisto-space at specific equidistant points in time, as opposed to

₂₄₂ considering subsets of taxa between two points in time. This results in even-sampling

₂₄₃ of the cladisto-space across time and permits us to define the underlying model of

₂₄₄ character evolution (such as punctuated or gradual). In practice, time-slicing considers

₂₄₅ the disparity of any element present in the phylogeny (branches, nodes and tips) at any

₂₄₆ point in time. When the phylogenetic elements are nodes or tips, the eigenvector scores

₂₄₇ for the nodes (estimated using ancestral state reconstruction as described above) or tips

₂₄₈ are directly used for estimating disparity. When the phylogenetic elements are branches

₁₄

we chose the eigenvector score for the branch using one of two evolutionary models:

1. **Punctuated evolution.** This model selects the eigenvector score from either the ancestral node or the descendant node/tip of the branch regardless of the position of the slice along the branch. Similarly to the time interval approach, this reflects a model of punctuated evolution where changes in disparity occur either at the start or at the end of a branch over a relatively short time period and clades undergo a long stasis period during their evolution (Gould and Eldredge, 1977; Hunt, 2007). We applied this model in three ways:

   (i) selecting the eigenvector score of the ancestral node of the branch

   (ii) selecting the eigenvector score of the descendant node/tip of the branch

   (iii) randomly selecting either the eigenvector score of the ancestral node or the descendant node/tip of the branch

   Method (i) assumes that changes always occurs early on the branch (accelerated transition, ACCTRAN) and (ii) assumes that changes always occur later (delayed transition, DELTRAN). We prefer not to make either assumption so we report the results from (iii), although the ACCTRAN and DELTRAN results are available in the Supplementary Information @@@.

2. **Gradual evolution.** This model also selects the eigenvector score from either the ancestral node or the descendant node/tip of the branch, but the choice depends on the distance between the sampling time point and the end of the branch. If the

15

²⁶⁹ sampling time point falls in the first half of the branch length the eigenvector

²⁷⁰ score is taken from the ancestral node, conversely, if the sampling time point falls

²⁷¹ in the second half of the branch length the eigenvector score is taken from the

²⁷² descendant node/tip. This reflects a model of gradual evolution where changes in

²⁷³ disparity are gradual and cumulative along the branch. Under this model, the

²⁷⁴ gradual changes could be either directional or random, however, directional

²⁷⁵ evolution have been empiricaly shown to be rare (only 5% of the time Hunt, 2007).

²⁷⁶ We therefore considered that changes from a character state A to B where just

²⁷⁷ dependant on the branch length.

²⁷⁸ We applied our time-slicing approach to the two cladisto-spaces calculated from Slater

²⁷⁹ (2013) and Beck and Lee (2014), time-slicing the phylogeny every five million years

²⁸⁰ from 170 Mya to the present resulting in 35 sub-samples of the cladisto-space. For each

²⁸¹ sub-sample, we estimated its disparity assuming punctuated (ACCTRAN, DELTRAN

²⁸² and random) and gradual evolution as described above. To reduce the influence of

²⁸³ outliers on our disparity estimates, we bootstrapped each disparity measurement by

²⁸⁴ randomly re-sampling with replacement a new sub-sample of taxa from the observed

²⁸⁵ taxa in the sub-sample 1000 times. We then calculated the median disparity value for

²⁸⁶ each sub-sample along with the 50% and the 95% confidence intervals.

²⁸⁷ We also reported the number of phylogenetic elements (branches, nodes and

²⁸⁸ tips) in each sub-sample representing the observed taxonomic diversity. Disparity may

²⁸⁹ be higher in sub-samples with more phylogenetic elements simply because there are

16

more taxa represented. To test whether our analyses were biased in this way, we rarefied our sub-samples during the bootstrap procedure by randomly re-sampling a fix number of taxa across each sub-sample. In Slater (2013), the minimum number of taxa in each sub-sample from 170 to present was 8. In Beck and Lee (2014), the minimum number of taxa however was 3, however, from 150 Mya until the present, the minimum number of taxa is 8. To make both data sets comparable, we used 8 as a minimum number of taxa for the rarefied bootstrap measurements in Beck and Lee (2014) ignoring therefore the sub-sample between 170 and 150 Mya. We report both results of the bootstrapped measurements and the rarefied bootstrap measurements.

To compare our results to previous studies we also repeated our analyses using the time interval approache based on biostratigraphy (e.g. Cisneros and Ruta, 2010; Prentice et al., 2011; Hughes et al., 2013; Benton et al., 2014) using each geological stage from the Middle Jurassic to the present. We report the results of these analyses in the Supplementary Materials (@@@).

Finally, to assess if the K-Pg boundary had a significant effect on mammals disparity, we performed Anderson (2001)'s Permutational Analysis of Variance (also referred to as PERMANOVA or NPANOVA; e.g. Brusatte et al., 2008a; Ruta et al., 2013) to test whether there was a significant effect of time on our calculated disparity. We calculated the euclidean distance of the ordinated data with 1000 permutations on both data sets and on both evolutionary scenario using the `adonis` function from the `R` package `vegan` (Oksanen et al., 2015). When a significant effect of time on disparity was

17

measured, we ran a series of *post-hoc* t-test between the time sub-samples (Anderson and Friedman, 2012; Zelditch et al., 2012; Smith et al., 2014) to test whether there was a significant effect between of the K-Pg boundary. We measured the difference between the last sub-sample of the Cretaceous (65 Mya) to all the slices of the Cenozoic to test whether there was either:

1. no effect of the K-Pg event: no significant difference between the last Cretaceous sub-sample and any of the Cenozoic sub-samples.

2. a direct effect of the K-Pg event: significant difference between the last Cretaceous sub-sample and the first Cenozoic sub-sample.

3. a lag effect of the K-Pg event: a significant difference between the last Cretaceous sub-sample and some sub-samples during the Cenozoic.

Because these *post-hoc* t-tests involve multiple p-values (13 comparisons), we corrected each p-value by multiplying them by the number of comparisons (Holm-Bonferonni correction; Holm, 1979).

Results

Both data sets do display a change in taxonomic diversity after and before the K-Pg boundary (Figure 1): respectively a decrease in Eutherians (data from Beck and Lee, 2014) and an increase in Mammaliaformes (data from Slater, 2013). However, these variations in taxonomic diversity are not linked with changes in disparity (Figure 2). In

18

both data sets, we measured similar changes in morphological disparity in the last 170

Million years for both the full data sets and the rarefied data sets (Figure 1 and 2) We

measured a significant effect of time on disparity in Eutherians under both gradual and

punctuated evolution model and in Mammaliaformes only under the gradual evolution

model only (table 1). For the Eutherians data set, we detected at least one significant

difference between the disparity in the last sub-sample of the Cretaceous and any

sub-sample of the Cenozoic suggesting an effect of the K-Pg boundary on disparity

(table 2). However, once corrected for taxonomic diversity, we did not detect any

significant difference, either under gradual or punctuated evolution model (table 3). For

the Mammaliaformes data set under gradual evolution model, there was no significant

difference between the disparity in the last sub-sample of the Cretaceous and any

sub-sample of the Cenozoic in both the raw and the rarefied data (table 4).
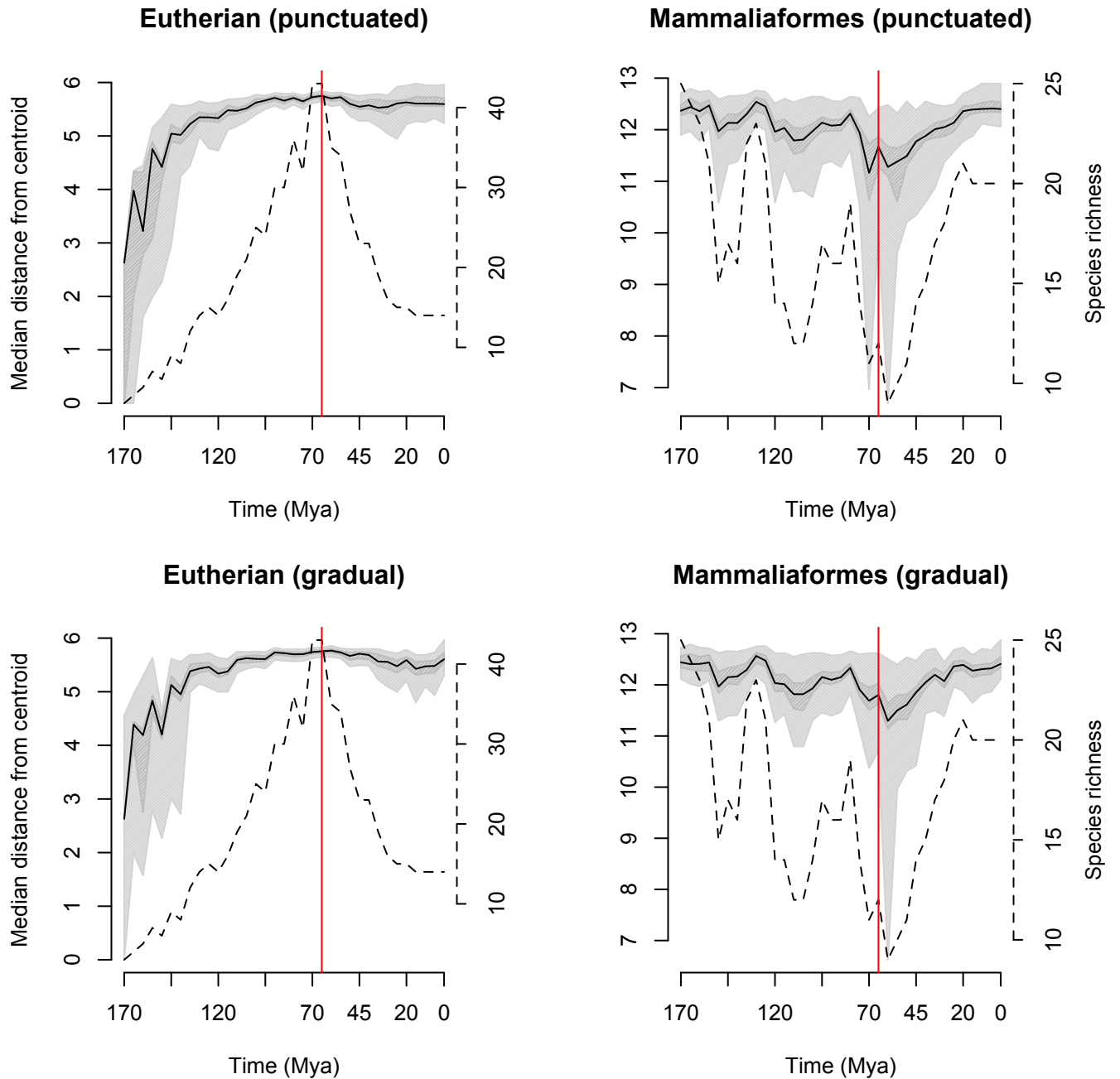
Figure 1: Observed variations of disparity through time among Eutherian and Mammaliaformes under punctuated or gradual evolution model. The x axis represents the time in Million of years ago (Mya). The y axis represents the disparity measured as the median distance from centroid per sub-sample. The solid black lines is the mean disparity from the bootstrapped pseudo-replicates, the confidence intervals (CI) are represent by the grey polygons (50% CI in dark grey and 95% CI in light grey). The dashed line represent the species richness in each sub-sample (values are reported on the right hand
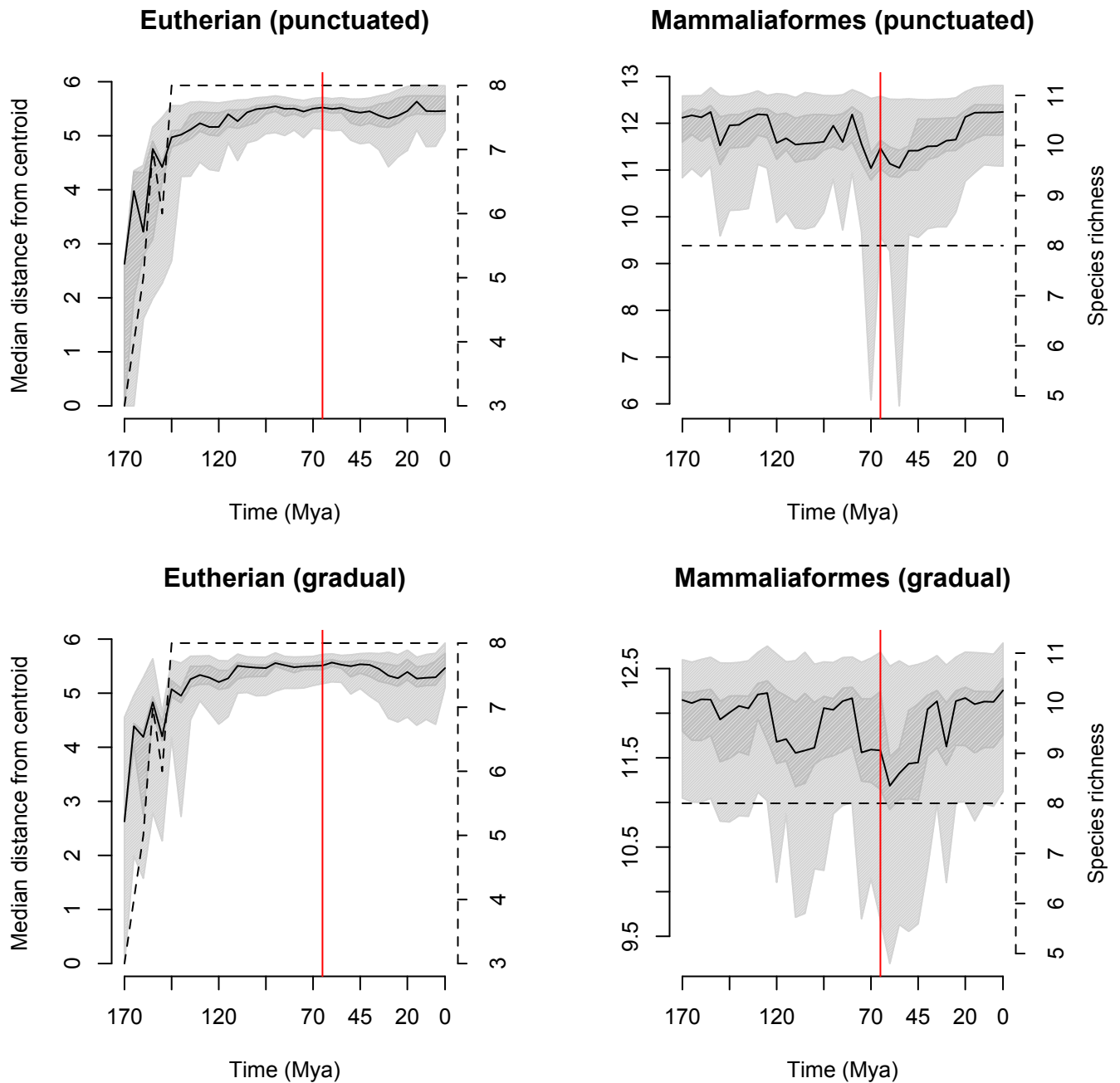
Figure 2: Rarefied variations of disparity through time among Eutherian and Mammaliaformes under punctuated or gradual evolution model. The x axis represents the time in Million of years ago (Mya). The y axis represents the disparity measured as the median distance from centroid per sub-sample. The solid black lines is the mean disparity from the bootstrapped pseudo-replicates; the confidence intervals (CI) are represent by the grey polygons (50% CI in dark grey and 95% CI in light grey). The dashed line represent the species richness in each sub-sample (values are reported on the right hand side of

Table 1: Permanova results of testing the effect of time on the ordinated distance matrix using euclidean distance with 1000 permutations. Data: Eutherian (data from Beck and Lee, 2014); Mammaliaformes, (data from Slater, 2013). Model: evolutionary model. Significant effects are highlighted in bold: one star (*) signifies a p-value between 0.05 and 0.005; two starts between 0.005 and 0.0005 and three stars < 0.0005.

Data	model	terms	Df	Sum of squares	Mean sum of squares	F Model	R^2	p-value	
Eutherian	gradual	time	34	1825.92	53.703	1.5784	0.0769	**0.0009**	***
		residuals	644	21911.65	34.024		0.9231		
	punctuated	time	34	1597.07	46.973	1.3693	0.0674	**0.0009**	***
		residuals	644	22092.28	34.305		0.9326		
Mammaliaformes	gradual	time	34	6525.61	191.930	1.1660	0.0663	**0.0009**	***
		residuals	558	91852.55	164.610		0.9337		
	punctuated	time	34	5741.25	168.860	1.0167	0.0583	0.2248	
		residuals	558	92672.75	166.080		0.9417		

Discussion

Our results show that there is a significant effect of time on changes in disparity under the assumption of gradual evolution in Mammaliaformes and Eutherians as well as under the assumption of punctuated evolution for Eutherians (table 1). However, regardless the taxonomic level (i.e. family *vs.* genus) and regardless the evolutionary model (i.e. gradual or punctuated evolution), there is no significant difference in disparity between the latests Cretaceous sub-sample and any of the Cenozoic sub-samples (Figure 2). In fact the disparity seems to reach a plateau at the end of the Jurassic (150 Mya) for Eutherians and during the late Triassic (Norian; 220 Mya) for the Mammaliaformes and stays relatively constant after that (see Figure 2 and S4 @@@). These results shows that, within the frame of our data-sets, we did not detect any short term or long term effect of the K-Pg event on mammalian disparity. Therefore, we argue that the extinction of the many terrestrial vertebrates (namely non-avian dinosaurs) at the K-Pg boundary did not directly affected mammals evolution during the Cenozoic, or at least their morphological diversity.

Global pattern of disparity

The observed global patterns of changes in disparity are consistent with previous studies on mammalian disparity and on disparity in metazoans in general. The patterns of disparity seems to plateau at their maximum disparity early in their history (approximatively after 25% and 12% of their history in respectively Mammaliaformes

23

362 and Eutherians). In fact, this quick increase in disparity early in history is consistent

363 with disparity patterns in metazoans (Hughes et al., 2013). Additionally, as showed

364 previously, disparity display a pattern decoupled from taxonomic diversity (Slater

365 et al., 2010; Ruta et al., 2013; Hopkins, 2013).

Effect of the K-Pg boundary on mammalian disparity

367 The results for the Mammaliaformes data-set disparity variation are consistent with the

368 most recent studies of Mammalian disparity, showing and early peak of disparity in the

369 Early to Middle Jurassic (Close et al., 2015). However, our results for Eutherians differ

370 from Grossnickle and Polly (2013) that showed a significant decrease in disparity

371 during the late Cretaceous (but see Wilson et al., 2012). These differences can be due to

372 the different input data to calculate disparity (morphometric data in Grossnickle and

373 Polly 2013; and cladistic data in the present study; but see Hetherington et al. 2015), the

374 different method to calculate disparity through time (time bin in Grossnickle and Polly

375 2013; and time slicing in the present study; see below) or the different focal

376 morphological aspect (dental morphology in Grossnickle and Polly 2013; and overall -

377 including dental - morphology in the present study). Furthermore, we found a

378 fundamental difference with Slater (2013) which shows solid evidences for a change in

379 mode of body mass evolution at the K-Pg boundary using the same data set as in this

380 study. We argue that this difference can be due to the number of traits used in Slater

381 (2013) and the present study. In this study we look at an aggregate of discrete traits (the

382 446 morphological characters) in opposition of one continuous trait (body mass in

24

Slater, 2013). Both variation in morphology (i.e. from cladistic data) and variation in body mass are two different and aspects of diversity and can likely be decoupled in the same way as taxonomic diversity is decoupled from disparity. However, we believe our results show some robustness because the same absence of signal of an eventual effect the K-Pg boundary as been found in two independent data sets (i.e. Slater, 2013; Beck and Lee, 2014). Furthermore, our results only suggest that disparity is not affected by the K-Pg boundary but they do not allow us to assess the effect of the K-Pg boundary on changes in body mass evolution. Besides, few caveats can be underlined.

Firstly, both our data sets are limited. They do not represent the full known mammalian taxonomic diversity, especially during the Neogene (23–2.58 Mya) where no fossils were represented in both data sets. However, this might not cause a serious under-sampling problem, at least in the Mammaliaformes dataset, since their diversity peaked during the late Cretaceous (Campanian; 72.1–83.6 Mya; Newham et al., 2014). Additionally Raia et al. (2012) have shown that mammalian diversification rates declined throughout the whole Cenozoic. In our study, these findings could suggest that an effect of the K-Pg boundary would be more likely detected during the Palaeogene when mammalian diversification rates where still high. Also, both data sets contains cladistic data for only a few living mammals which can also have a effect of topological accuracy (Guillerme and Cooper, In review, 2015). However, both original phylogenies where built using strong topological constraint to avoid such a caveat (Slater, 2013; Beck and Lee, 2014).

₄₀₄   Secondly, the core of the debate in mammalian evolution is weather *placental*

₄₀₅ mammals originated before or after the K-Pg boundary (dos Reis et al., 2012; O'Leary

₄₀₆ et al., 2013; Springer et al., 2013; OLeary et al., 2013; dos Reis et al., 2014). The infraclass

₄₀₇ Placentalia can be defined as "the least inclusive clade that includes all extant

₄₀₈ placentals" (Beck and Lee, 2014). However, part of the dating debate might be due to

₄₀₉ the lack of clear characters that can be used to define early placental mammals

₄₁₀ (Bininda-Emonds et al., 2012; Beck and Lee, 2014). Cartmill (2012) also argues that the

₄₁₁ use of higher taxa definition in general might be obsolete since "there is only a long,

₄₁₂ geologically slow cascade of accumulating small apomorphies". Therefore, in this

₄₁₃ study, we made the deliberate choice to focus on the taxonomic levels (genus *vs.* family)

₄₁₄ rather than on the higher clades definitions. We argue that if a significant change in

₄₁₅ disparity occurred at the K-Pg boundary in any of such infraclass (Placentalia,

₄₁₆ Marsupialia, etc...) it would be detectable even at a higher level (i.e. changes in

₄₁₇ Placentalia correspond by definition to changes in Eutheria and Mammaliaformes).

₄₁₈ Also, using Total Evidence tip-dated trees provides more accurate estimates of

₄₁₉ diversification times (Ronquist et al. 2012; Wood et al. 2013; Beck and Lee 2014; but see

₄₂₀ Arcila et al. 2015) and allows the opportunity to look at changes in disparity for both

₄₂₁ living and fossil species.

₄₂₂   *Methodology improvements for measuring disparity*

₄₂₃ Additionally, throughout this paper, we propose several incremental changes to the

₄₂₄ classic ways to measure disparity. This is how we believe they improve disparity

through time analysis:

1. **Using all the axis of the cladisto-space.** Previous studies focusing on disparity
   have used various ways to select a sub-sample of the the full cladisto-space (i.e. a
   sub-sample of the ordinated distance matrix) arguing that the m first axis of the
   cladisto-space usually bear most of the data-set's variance (e.g Brusatte et al.,
   2008a; Cisneros and Ruta, 2010; Prentice et al., 2011; Anderson and Friedman,
   2012; Hughes et al., 2013; Benton et al., 2014). For example in Brusatte et al.
   (2008b) and Toljagic and Butler (2013), the authors decided to select only the m
   first dimensions that represent up to 90% of the variance in the distance matrix.
   The cut-off value is either given arbitrary or by visually detecting a substantial
   break in the slop of a scree plot of the variance per axis (Wills et al., 1994). We
   argue that even if the last dimensions of the cladisto-space bears a trivial amount
   of variance, there is no statistical justification to exclude them. However, by doing
   so, we included dimensions of the cladisto-space with a near 0 variance and range
   (variance of 2×10^{-14} and 1.15×10^{-15} and range of 7.31×10^{-7} and 3.33×10^{-7}
   in respectively Slater 2013 and Beck and Lee 2014) This makes the calculation of
   certain disparity metrics impossible (see below). An alternative method to avoid
   these near 0 dimensional axis problem is to simply not ordinated the data and
   measure disparity just from the k distance matrix (e.g. Benson and Druckenmiller,
   2014; Close et al., 2015).

2. **Using the median distance from centroid as a disparity metric.** As stated above

27

(see methods section), we deliberately chose to use the median distance from centroid as a metric from measuring disparity among many other (e.g. Wills et al., 1994; Ciampaglio, 2004; Thorne et al., 2011; Hopkins, 2013; Huang et al., 2015). Using this metric gave use several advantages upon the four classic sum and products of ranges and variance. First, this metric comes with no special statistical assumptions (c.f. sum and products of variance *and* covariance). Secondly, this metric is not affected by the last dimensions of the cladisto-space problem (see above). And thirdly, this metric seems less coupled with taxonomic diversity (especially for the products of ranges and variance and the sum of ranges, see supplementary figures @@@).

3. **Using time slicing method.** Contra to numerous studies focusing on disparity (e.g Cisneros and Ruta, 2010; Prentice et al., 2011; Hughes et al., 2013; Hopkins, 2013; Benton et al., 2014; Benson and Druckenmiller, 2014), we chose not to bin our data in time intervals but rather to look it as a continuous process. We argue that by doing so, we can avoid two caveats of such method:

   (i) firstly, using time intervals based on biostratigraphy is tautologic. In fact, such method is likely to artefactualy emphasize disparity differences between time intervals because the same time intervals are based on notable differences in fossil fauna and flora (see supplementary figures @@@ where differences in disparity through time are much more contrasted in the interval methods than in the slicing methods). But note that not all studies

use biostratigraphy and sometimes arbitrarily time bins of equal duration are used which also fix this caveat (Butler et al., 2012; Hopkins, 2013; Benson and Druckenmiller, 2014).

(ii) secondly, in both cases (time bins based on biostratigraphy or arbitrary durations), such method does not allow to specify assumptions on the evolutionary model. In fact, the underlying assumption to such method is that changes in disparity occur between the time intervals in a punctuated evolutionary model fashion. Although directional gradual evolution has been shown to be rare, punctuated (i.e. stasis) and gradual (i.e. random walk) evolution have been shown to be both relatively common (Hunt, 2007; Hunt et al., 2015). Therefore, assuming that evolution is only punctual might be erroneous in some cases and for some traits.

4. **Allowing to choose the evolutionary model.** Finaly, using the time slicing method, allows use to crudely specify the evolutionary model for changes in disparity. Within Eutherians, we showed both support for an effect of time on disparity under both models of evolution. This can reflect the complex combination between the two modes of evolution where morphology (i.e. as inferred from the cladistic data) varies stochastically through time with a mix of random walks (i.e. gradual model) for certain set of characters and stasis (i.e. punctuated model) for others. These results are consistent with previous findings (Hunt, 2007; Hunt et al., 2015). It is also encouraging to see that the distinction

29

between the two modes of evolution can help understanding the patterns of changes in disparity at a finer scale. In fact, for Mammaliaformes, there is no significant effect of time on disparity under the assumption of a punctuated model of evolution but a clear effect of time when evolution is assumed to be gradual (see Table 1@@@). When looking at the details of this results, the same data sets shows also no significant effect of time when using the time bin method (including nodes) or when assuming that disparity evolves under an ACCTRAN model (see supplementary permanova results @@@). This suggests that there is an effect of time on Mammaliaformes with mix between punctuated delay evolution (DELTRAN) and gradual evolution (random walk). This could be interpreted as when a particular morphology (i.e. a set of particular states for cladistic characters) is observed within a clade, this particular morphology will be likely conserved through time. Other common but more complex models could also be implemented such as a combined stasis and random walk (Hunt et al., 2015) or models based on morphological rates rather than just the sheer branch length. For example, one could use a density of probability for choosing the ordinated data for either the descendant or the ancestor based on morphological clocks rather than just branch length.

Two major caveats, however, arise from using such a method. First, this method relies on good estimates of characters states at the nodes of the phylogeny. Estimating discrete ancestral characters can sometimes be tricky and can lead to low scaled

likelihood values supporting any states of a particular character, especially when many data are missing in the observed cladistic matrix. However, in this particular study, we made the methodological choice of selecting only characters with a high scaled likelihood support (> 0.95). Additionally, using trees containing fossil taxa also improves the ability to correctly estimate ancestral characters (David Polly, 2001; Finarelli and Flynn, 2006; Albert et al., 2009; Slater, 2013). Finally, because, this method samples every phylogenetic element (tip, node or edge) through time, disparity calculated close to the root of the tree can exhibit result with large confidence intervals (e.g. when only three phylogenetic elements are sampled see Figure S3 and S@@@). However, it is encouraging to note that measuring disparity from time-slices is decoupled from taxonomic diversity at least after a minimal number of taxa (Slater et al., 2010; Ruta et al., 2013; Hopkins, 2013)

In summary, the majority attempts to solve the debate on whether placental mammals diversified after or before the K-Pg boundary is based on taxonomic diversity and shows unclear evidences on weather the K-Pg extinction event had an effect on mammalian diversification (Meredith et al., 2011; O'Leary et al., 2013; dos Reis et al., 2014; Beck and Lee, 2014). Among the variety of macroevolutionary process proposed to support an effect of the K-Pg boundary on mammalian evolution, some authors proposed the release of ecological niches after the K-Pg boundary (e.g. Archibald, 2011; O'Leary et al., 2013) or a release of competition pressures (e.g. Slater, 2013; Lovegrove et al., 2014). In this study, however, we proposed a different approach looking at

31

morphological diversity (i.e. disparity) through time using a continuous time sampling approach that allows use to specify assumption on the mode of evolution (i.e. punctuated or gradual). We based our analysis on the palaeontological discoveries of the last decade showing an unprecedented and unexpected taxonomic and morphological diversity prior to the Cenozoic (Luo, 2007; Close et al., 2015). We found no evidences for an effect of the K-Pg boundary on changes in mammalian disparity at both the family and the genus level and under both assumption of gradual or punctual evolution. We therefore suggest that, contra to popular believe, the extinction of many terrestrial vertebrates (namely the dominant non-avian dinosaurs) did not significantly affect the evolution of mammals throughout the Cenozoic.

Data availability and reproducibility

Data will be available on Dryad or Figshare. Code for reproducing the analysis is available on GitHub (`ithub.com/TGuillerme/SpatioTemporal_Disparity`).

Acknowledgments

Funding

*

References

Albert, J. S., D. M. Johnson, and J. H. Knouft. 2009. Fossils provide better estimates of ancestral body size than do extant taxa in fishes. Acta Zoologica 90:357–384.

Anderson, M. J. 2001. A new method for non-parametric multivariate analysis of variance. Austral Ecology 26:32–46.

Anderson, P. S. and M. Friedman. 2012. Using cladistic characters to predict functional variety: experiments using early gnathostomes. Journal of Vertebrate Paleontology 32:1254–1270.

Archibald, J. D. 2011. Extinction and radiation: how the fall of dinosaurs led to the rise of mammals. JHU Press.

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

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.

Benson, R. B. J. and P. S. Druckenmiller. 2014. Faunal turnover of marine tetrapods during the Jurassic-Cretaceous transition. Biological Reviews 89:1–23.

33

566   Benton, M. J. 1985. Mass extinction among non-marine tetrapods. Nature 316:811–814.

567   Benton, M. J., J. Forth, and M. C. Langer. 2014. Models for the rise of the dinosaurs.

568   Current biology : CB 24:R87–R95.

569   Benton, M. J. and R. J. Twitchett. 2003. How to kill (almost) all life: the end-permian

570   extinction event. Trends in Ecology and Evolution 18:358 – 365.

571   Bininda-Emonds, O., R. Beck, and R. D. MacPhee. 2012. Rocking clocks and clocking

572   rocks: a critical look at divergence time estimation in mammals. From bone to clone:

573   the synergy of morphological and molecular tools in paleobiology Pages 38–82.

574   Bininda-Emonds, O. R., M. Cardillo, K. E. Jones, R. D. MacPhee, R. M. Beck, R. Grenyer,

575   S. A. Price, R. A. Vos, J. L. Gittleman, and A. Purvis. 2007. The delayed rise of

576   present-day mammals. Nature 446:507–512.

577   Brusatte, S., R. J. Butler, A. Prieto-Márquez, and M. A. Norell. 2012. Dinosaur

578   morphological diversity and the end-Cretaceous extinction. Nature Communications

579   3:804–804.

580   Brusatte, S. L., M. J. Benton, M. Ruta, and G. T. Lloyd. 2008a. The first 50Łmyr of

581   dinosaur evolution: macroevolutionary pattern and morphological disparity. Biology

582   Letters 4:733–736.

583   Brusatte, S. L., M. J. Benton, M. Ruta, and G. T. Lloyd. 2008b. Superiority, competition,

584   and opportunism in the evolutionary radiation of dinosaurs. Science 321:1485–1488.

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.

Butler, R. J., S. L. Brusatte, B. Andres, and R. B. J. Benson. 2012. How do geological sampling biases affect studies of morphological evolution in deep time? a case study of pterosaur (reptilia: Archosauria) disparity. Evolution 66:147–162.

Cailliez, F. 1983. The analytical solution of the additive constant problem. Psychometrika 48:305–308.

Cartmill, M. 2012. Primate origins, human origins, and the end of higher taxa. Evolutionary Anthropology: Issues, News, and Reviews 21:208–220.

Ciampaglio, C. N. 2004. Measuring changes in articulate brachiopod morphology before and after the permian mass extinction event: do developmental constraints limit morphological innovation? Evolution & Development 6:260–274.

Ciampaglio, C. N., M. Kemp, and D. W. McShea. 2001. Detecting changes in morphospace occupation patterns in the fossil record: characterization and analysis of measures of disparity. Paleobiology 27:695–715.

Cisneros, J. C. and M. Ruta. 2010. Morphological diversity and biogeography of procolophonids (amniota: Parareptilia). Journal of Systematic Palaeontology 8:607–625.

604 Clapham, M. E., D. J. Bottjer, C. M. Powers, N. Bonuso, M. L. Fraiser, P. J. Marenco,

605   S. Q. Dornbos, and S. B. Pruss. 2006. Assessing the ecological dominance of

606   phanerozoic marine invertebrates. PALAIOS 21:431–441.

607 Close, R., M. Friedman, G. Lloyd, and R. Benson. 2015. Evidence for a mid-jurassic

608   adaptive radiation in mammals. Current Biology .

609 Coxall, H. K., S. D'Hondt, and J. C. Zachos. 2006. Pelagic evolution and environmental

610   recovery after the cretaceous-paleogene mass extinction. Geology 34:297–300.

611 David Polly, P. 2001. Paleontology and the comparative method: Ancestral node

612   reconstructions versus observed node values. The American Naturalist 157:pp.

613   596–609.

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

615   asteroids, and marine production: Death and recovery at the cretaceous-tertiary

616   boundary. Geological Society of America Special Papers 307:303–317.

617 dos Reis, M., P. C. J. Donoghue, and Z. Yang. 2014. Neither phylogenomic nor

618   palaeontological data support a palaeogene origin of placental mammals. Biology

619   Letters 10.

620 dos Reis, M., J. Inoue, M. Hasegawa, R. J. Asher, P. C. J. Donoghue, and Z. Yang. 2012.

621   Phylogenomic datasets provide both precision and accuracy in estimating the

622   timescale of placental mammal phylogeny. Proceedings of the Royal Society of

623   London B: Biological Sciences .

36

[624] Douady, C. J. and E. J. Douzery. 2003. Molecular estimation of eulipotyphlan divergence [625] times and the evolution of insectivora. Molecular Phylogenetics and Evolution 28:285 [626] – 296.

[627] Eernisse, D. and A. Kluge. 1993. Taxonomic congruence versus total evidence, and [628] amniote phylogeny inferred from fossils, molecules, and morphology. Molecular [629] Biology and Evolution 10:1170–1195.

[630] Erwin, D. H. 1998. The end and the beginning: recoveries from mass extinctions. Trends [631] in Ecology and Evolution 13:344 – 349.

[632] Erwin, D. H. 2007. Disparity: Morphological pattern and developmental context. [633] Palaeontology 50:57–73.

[634] Finarelli, J. A. and J. J. Flynn. 2006. Ancestral state reconstruction of body size in the [635] caniformia (carnivora, mammalia): The effects of incorporating data from the fossil [636] record. Systematic Biology 55:301–313.

[637] Foote, M. 1994. Morphological disparity in ordovician-devonian crinoids and the early [638] saturation of morphological space. Paleobiology 20:320–344.

[639] Foote, M. 1996. Ecological controls on the evolutionary recovery of post-paleozoic [640] crinoids. Science 274:1492–1495.

[641] Friedman, M. 2010. Explosive morphological diversification of spiny-finned teleost [642] fishes in the aftermath of the end-Cretaceous extinction. Proceedings of the Royal [643] Society B: Biological Sciences 277:1675–1683.

Garland, J., Theodore and A. R. Ives. 2000. Using the past to predict the present: Confidence intervals for regression equations in phylogenetic comparative methods. The American Naturalist 155:346–364.

Glor, R. E. 2010. Phylogenetic insights on adaptive radiation. Annual Review of Ecology, Evolution, and Systematics 41:251–270.

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.

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

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

Gower, J. C. 1971. A general coefficient of similarity and some of its properties. Biometrics 27:pp. 857–871.

Grossnickle, D. M. and P. D. Polly. 2013. Mammal disparity decreases during the cretaceous angiosperm radiation. Proceedings of the Royal Society of London B: Biological Sciences 280.

Guillerme, T. and N. Cooper. 2015. Assessment of cladistic data availability for living mammals. bioRxiv .

Guillerme, T. and N. Cooper. In review. Effects of missing data on topological inference using a total evidence approach,. Molecular Phylogenetics and Evolution .

Hetherington, A. J., E. Sherratt, M. Ruta, M. Wilkinson, B. Deline, and P. C. Donoghue. 2015. Do cladistic and morphometric data capture common patterns of morphological disparity? Palaeontology 58:393–399.

Holm, S. 1979. A simple sequentially rejective multiple test procedure. Scandinavian journal of statistics Pages 65–70.

Hopkins, M. 2013. Decoupling of taxonomic diversity and morphological disparity during decline of the cambrian trilobite family pterocephaliidae. Journal of Evolutionary Biology 26:1665–1676.

Huang, S., K. Roy, and D. Jablonski. 2015. Origins, bottlenecks, and present-day diversity: Patterns of morphospace occupation in marine bivalves. Evolution .

Hughes, M., S. Gerber, and M. A. Wills. 2013. Clades reach highest morphological disparity early in their evolution. Proceedings of the National Academy of Sciences 110:13875–13879.

Hunt, G. 2007. The relative importance of directional change, random walks, and stasis in the evolution of fossil lineages. Proceedings of the National Academy of Sciences 104:18404–18408.

Hunt, G., M. J. Hopkins, and S. Lidgard. 2015. Simple versus complex models of trait

evolution and stasis as a response to environmental change. Proceedings of the National Academy of Sciences 112:4885–4890.

Korn, D., M. J. Hopkins, and S. A. Walton. 2013. Extinction spacea method for the quantification and classification of changes in morphospace across extinction boundaries. Evolution 67:2795–2810.

Lloyd, G. T. 2015. Claddis: Measuring Morphological Diversity and Evolutionary Tempo. R package version 0.1.

Losos, J. B. 2010. Adaptive radiation, ecological opportunity, and evolutionary determinism. The American Naturalist 175:pp. 623–639.

Lovegrove, B. G., K. D. Lobban, and D. L. Levesque. 2014. Mammal survival at the cretaceous–palaeogene boundary: metabolic homeostasis in prolonged tropical hibernation in tenrecs. Proceedings of the Royal Society of London B: Biological Sciences 281.

Luo, Z.-X. 2007. Transformation and diversification in early mammal evolution. Nature 450:1011–1019.

    KEY: luo2007

     ANNOTATION: 10.1038/nature06277

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.

Newham, E., R. Benson, P. Upchurch, and A. Goswami. 2014. Mesozoic mammaliaform diversity: The effect of sampling corrections on reconstructions of evolutionary dynamics. Palaeogeography, Palaeoclimatology, Palaeoecology 412:32 – 44.

Oksanen, J., F. G. Blanchet, R. Kindt, P. Legendre, P. R. Minchin, R. B. O'Hara, G. L. Simpson, P. Solymos, M. H. H. Stevens, and H. Wagner. 2015. vegan: Community Ecology Package. R package version 2.3-0.

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

Olson, M. E. and A. Arroyo-Santos. 2009. Thinking in continua: beyond the adaptive radiation metaphor. BioEssays 31:1337–1346.

OLeary, 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. 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. Response to

722 comment on the placental mammal ancestor and the postk-pg radiation of placentals.

723 Science 341:613.

724 Paradis, E., J. Claude, and K. Strimmer. 2004. APE: analyses of phylogenetics and

725 evolution in R language. Bioinformatics 20:289–290.

726 Payne, J. L., N. A. Heim, M. L. Knope, and C. R. McClain. 2014. Metabolic dominance

727 of bivalves predates brachiopod diversity decline by more than 150 million years.

728 Proceedings of the Royal Society B: Biological Sciences 281.

729 Pearman, P. B., A. Guisan, O. Broennimann, and C. F. Randin. 2008. Niche dynamics in

730 space and time. Trends in Ecology and Evolution 23:149 – 158.

731 Prentice, K. C., P. Ruta, and M. J. Benton. 2011. Evolution of morphological disparity in

732 pterosaurs. Journal of Systematic Palaeontology 9:337–353.

733 R Core Team. 2015. R: a language and environment for statistical computing. R

734 Foundation for Statistical Computing Vienna, Austria.

735 Raia, P., F. Carotenuto, F. Passaro, P. Piras, D. Fulgione, L. Werdelin, J. Saarinen, and

736 M. Fortelius. 2012. Rapid action in the palaeogene, the relationship between

737 phenotypic and taxonomic diversification in coenozoic mammals. Proceedings of the

738 Royal Society of London B: Biological Sciences 280.

739 Raup, D. M. 1979. Size of the permo-triassic bottleneck and its evolutionary

740 implications. Science 206:217–218.

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

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

743 Cretaceous-Paleogene boundary. Science 339:684–687.

744 Revell, L. J. 2012. phytools: An R package for phylogenetic comparative biology (and

745 other things). Methods in Ecology and Evolution 3:217–223.

746 Ronquist, F., S. Klopfstein, L. Vilhelmsen, S. Schulmeister, D. Murray, and A. Rasnitsyn.

747 2012. A total-evidence approach to dating with fossils, applied to the early radiation

748 of the Hymenoptera. Systematic Biology 61:973–999.

749 Ruta, M., K. D. Angielczyk, J. Fröbisch, and M. J. Benton. 2013. Decoupling of

750 morphological disparity and taxic diversity during the adaptive radiation of

751 anomodont therapsids. Proceedings of the Royal Society of London B: Biological

752 Sciences 280.

753 Sepkoski, J., J. John. 1981. A factor analytic description of the phanerozoic marine fossil

754 record. Paleobiology 7:pp. 36–53.

755 Slater, G. J. 2013. Phylogenetic evidence for a shift in the mode of mammalian body size

756 evolution at the cretaceous-palaeogene boundary. Methods in Ecology and Evolution

757 4:734–744.

758 Slater, G. J., S. A. Price, F. Santini, and M. E. Alfaro. 2010. Diversity versus disparity and

759 the radiation of modern cetaceans. Proceedings of the Royal Society of London B:

760 Biological Sciences .

43

761 Smith, A. J., M. V. Rosario, T. P. Eiting, and E. R. Dumont. 2014. Joined at the hip:

762   Linked characters and the problem of missing data in studies of disparity. Evolution

763   68:2386–2400.

764 Springer, M. S., R. W. Meredith, E. C. Teeling, and W. J. Murphy. 2013. Technical

765   comment on the placental mammal ancestor and the postk-pg radiation of placentals.

766   Science 341:613.

767 Stadler, T. 2011. Mammalian phylogeny reveals recent diversification rate shifts.

768   Proceedings of the National Academy of Sciences 108:6187–6192.

769 Thorne, P. M., M. Ruta, and M. J. Benton. 2011. Resetting the evolution of marine

770   reptiles at the Triassic-Jurassic boundary. Proceedings of the National Academy of

771   Sciences 108:8339–8344.

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

773   mesozoic radiation of crocodylomorphs. Biology Letters 9.

774 Torgerson, W. S. 1965. Multidimensional scaling of similarity. Psychometrika

775   30:379–393.

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

777   america. Paleobiology 31:35–55.

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

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

Wilson, G. P. 2013. Mammals across the k/pg boundary in northeastern montana, u.s.a.: dental morphology and body-size patterns reveal extinction selectivity and immigrant-fueled ecospace filling. Paleobiology 39:429–469.

Wilson, G. P., A. R. Evans, I. J. Corfe, P. D. Smits, M. Fortelius, and J. Jernvall. 2012. Adaptive radiation of multituberculate mammals before the extinction of dinosaurs. Nature 483:457–460.

Wood, H. M., N. J. Matzke, R. G. Gillespie, and C. E. Griswold. 2013. Treating fossils as terminal taxa in divergence time estimation reveals ancient vicariance patterns in the palpimanoid spiders. Systematic Biology 62:264–284.

Yang, Z., S. Kumar, and M. Nei. 1996. A new method of inference of ancestral nucleotide and amino acid sequences. Genetics 141:1641–50.

Zelditch, M. L., D. L. Swiderski, and H. D. Sheets. 2012. Geometric morphometrics for biologists: a primer. Academic Press.

Table 2: Results of the *post-hoc* t-tests for comparing the disparity at the last sub-sample of the Cretaceous (65 Mya) to all the sub-samples of the Cenozoic for the Eutherians (data from Beck and Lee, 2014). Sub-samples: reference sample (65 Million years ago; Mya) to Cenozoic sample (from 60 Mya to present). Gradual: gradual evolution; punctuated: punctuated evolution. Difference: mean sub-sample difference; Df: degrees of freedom; T: T statistic; p-value: adjusted p-value using Holm-Bonferroni correction. Significant differences are highlighted in bold: one star (*) signifies a p-value between 0.05 and 0.005; two starts between 0.005 and 0.0005 and three stars < 0.0005.

Sub-samples	Gradual					Punctuated				
(Mya)	Difference	Df	T	p.value		Difference	Df	T	p.value	
65:60	0.06	76	1.055	1		0.04	76	0.760	1	
65:55	0.05	75	0.999	1		0.16	75	3.145	**0.0310**	*
65:50	0.15	68	2.412	0.2413		0.08	68	1.403	1	
65:45	0.21	64	3.016	**0.0478**	*	0.18	64	2.685	0.1200	
65:40	0.18	64	2.579	0.1590		0.13	64	2.173	0.4354	
65:35	0.23	60	2.840	0.0800	.	0.21	60	2.962	0.0568	.
65:30	0.27	57	2.927	0.0639	.	0.29	57	3.810	**0.0044**	**
65:25	0.22	56	2.500	0.1999		0.28	56	3.544	**0.0104**	*
65:20	0.16	56	1.922	0.7762		0.25	56	3.117	**0.0374**	*
65:15	0.14	55	1.819	0.9670		0.30	55	3.567	**0.0098**	**
65:10	0.14	55	1.843	0.9203		0.42	55	4.540	**0.0004**	***
65:5	0.14	55	1.790	1		0.30	55	3.377	**0.0176**	*
65:0	0.14	55	1.818	0.9692		0.17	55	2.250	0.3705	

Table 3: Results of the *post-hoc* t-tests for comparing the disparity at the last sub-sample of the Cretaceous (65 Mya) to all the sub-samples of the Cenozoic for the rarefied Eutherians (data from Beck and Lee, 2014). Column heads explained same as given in table 2.

Sub-samples	Gradual				Punctuated			
(Mya)	Difference	Df	T	p.value	Difference	Df	T	p.value
65:60	0.04	76	0.218	1	0.01	76	0.064	1
65:55	0.04	75	0.213	1	0.14	75	0.797	1
65:50	0.11	68	0.553	1	0.04	68	0.224	1
65:45	0.15	64	0.716	1	0.13	64	0.600	1
65:40	0.11	64	0.544	1	0.07	64	0.358	1
65:35	0.15	60	0.627	1	0.12	60	0.572	1
65:30	0.15	57	0.636	1	0.17	57	0.772	1
65:25	0.10	56	0.423	1	0.16	56	0.697	1
65:20	0.03	56	0.131	1	0.13	56	0.555	1
65:15	0	55	0.005	1	0.16	55	0.674	1
65:10	-0.01	55	-0.034	1	0.27	55	1.129	1
65:5	0.01	55	0.029	1	0.15	55	0.640	1
65:0	0	55	0.005	1	0.02	55	0.071	1

Table 4: Results of the *post-hoc* t-tests for comparing the disparity at the last sub-sample of the Cretaceous (65 Mya) to all the sub-samples of the Cenozoic for the Mammaliaformes (data from Slater, 2013) under gradual evolution model. Raw data: data without correcting for taxonomic diversity; Rarefied data: rarefied bootstrapped data. Other column heads explained same as given in table 2.

Sub-samples	Raw data				Rarefied data			
(Mya)	Difference	Df	T	p.value	Difference	Df	T	p.value
65:60	0.49	19	0.826	1	0.26	19	0.365	1
65:55	0.45	20	0.734	1	0.31	20	0.428	1
65:50	0.13	21	0.267	1	0.03	21	0.042	1
65:45	-0.05	24	-0.109	1	0.03	24	0.051	1
65:40	-0.22	25	-0.543	1	-0.08	25	-0.118	1
65:35	-0.33	27	-0.858	1	-0.19	27	-0.321	1
65:30	-0.37	28	-0.973	1	-0.21	28	-0.335	1
65:25	-0.48	30	-1.358	1	-0.25	30	-0.394	1
65:20	-0.69	31	-2.030	0.6625	-0.44	31	-0.711	1
65:15	-0.76	30	-2.201	0.4620	-0.53	30	-0.906	1
65:10	-0.86	30	-2.666	0.1593	-0.66	30	-1.241	1
65:5	-0.85	30	-2.668	0.1585	-0.63	30	-1.197	1
65:0	-0.86	30	-2.678	0.1548	-0.62	30	-1.133	1