**Supplementary Online Materials for “Quaternary megafauna extinctions altered body size distribution in tortoises”**

**Authors:** Julia Joos, Catalina Pimiento, Donald B. Miles, Johannes Müller

**Journal Name:** Proceedings of the Royal Society B

**Article DOI:** 10.1098/rspb

*List of Figures:*

Figure S1: Map displaying all localities of fossil records in the body size data set.

Figure S2: Straight carapace length estimation based on plastron length via multivariate imputation by chained equations (R package “mice”).

Figure S3: Accumulation curves based on references used to collect body size measurements.

Figure S4: Data points indicate the size difference in SCL between two sampled time bins bracketing a gap in the fossil record of a genus plotted across the duration of the gap.

Figure S5: Comparison of absolute sample sizes of large and small tortoises in each time bin.

Figure S6: Comparison of relative sample sizes of large and small tortoises in each time bin.

Figure S7. Temporal and latitudinal trends in body size (SCL) based on a generalized additive model (GAM).

Figure S8: Body size distribution in the late (A) and early (B) Quaternary.

Figure S9: Comparison of tortoise body size based on temporal and spatial status, separately.

Figure S10: Illustration of how co-occurring trends of gigantism and miniaturization can lead to stasis on the family level.

Figure S11: Presence of genera and their body sizes throughout the time bins.

Tables

Table S1. Fossil Tortoises.xlsx

Table S2. Extant Tortoises.xlsx

Table S3. Summary of Extant Tortoises.xlsx

Table S4. Multiple comparison test after Kruskal-Wallis.

**Sections**

Background on data selection

A. Gaps in fossil occurrences

B. Preservation bias and pull of the recent

C. Analysis of age and latitudinal trends in body size

References Cited.



Figure S1: Map displaying all localities of fossil records in the body size data set. Color indicates age, size of data points indicates carapace length, transparency of points reflects sample size. Large bodied tortoises were abundant throughout the distribution area of Testudinidae, both on the mainland and on islands.



Figure S2: Straight carapace length (CL) estimation based on plastron length (PL) via multivariate imputation by chained equations (R package “mice”). Blue data points are records where both carapace and plastron measurements were available, red data points are imputed values via a Bayesian linear regression.



Figure S3: Accumulation curves based on references used to collect body size measurements. A) Accumulation curve on the species level, curve is far from reaching asymptote indicating insufficient sampling. B) Accumulation curve on the genus level, curve is close to reaching asymptote and exhibits a shape we can reasonably expect for a sufficiently sampled fossil record with rare species.

1. **Background on data selection**
2. Gaps in fossil occurrences

In our data set, 13 of 31 genera have gaps in their fossil record (the range in gaps is 1-5 time bins; the mean number of bins with missing body size data is ~ 3 time bins). It makes intuitive sense to assume the presence of a given genus within a time bin where it was not sampled despite occurring in adjacent time bins. However, assuming the presence of a genus is not the same as inferring its body size. There is a high degree of uncertainty in placing an estimate of SCL when we do not know the temporal trend in the evolution of body size. We found that the difference in body size between the occurrence of a genus ranged from 2 - 817 mm, with an average of 354 mm (Fig. S4). Filling in this data would require assumptions about the evolutionary tempo of body size we cannot make. For example, does body size decrease/increase linearly or are there evolutionary jumps? Therefore*,* we feel more comfortable accepting the gaps in the data on body size. We believe that because our analysis is based on Testudinidae as a whole the missing body size values do not influence the results. We also should like to point out our analysis is at the familial-level and we included genera to reduce potential sampling bias if some genera are over-represented.



Figure S4: Data points indicate the size difference in SCL between two sampled time bins bracketing a gap in the fossil record of a genus plotted across the duration of the gap. The color of the points indicates whether the younger record at the most recent end of the gap has decreased (red) or increased (blue) from the next oldest record. Point size indicates SCL of the younger record.

Genus level analysis

Since there are fewer genera than species in a clade, it is to be expected that genera reach an asymptote earlier than species [1]. Although the accumulation curve for the entire data set does not completely converge to an asymptote, considering the large area covered [2] and the high number of rare genera in the dataset (which are to be expected in a fossil dataset), it can be considered well enough sampled for the present study [1]. The remaining analyses are conducted on the generic level because generic level identifications in the fossil record are more robust than species level identification and genera are better sampled in my data set [3].

1. Preservation bias and pull of the recent

*Material and Methods*

Smaller tortoise shells are less likely to be preserved in the fossil record and found during excavation than larger tortoise shells [4], especially with increasing age of the fossil record. To determine whether our data set was affected by a preservation/sampling bias, we split the dataset into two size classes and compared sample sizes within each time bin.

*Results*

We found that smaller tortoises (< 1000 mm) have a higher sample size than larger tortoises (> 1000 mm) in each time bin (Fig. S5) with more than 50% of each sample being comprised of small tortoise shells (Fig. S6). Langhian, Messinian, and Zanclean epochs have the smallest difference in relative sample sizes between large and small tortoise shells: however, the pattern is specific to those three time bins. Note: these three epochs also have the lowest diversity in terms of sampled species and genera. Moreover, the presence of large versus small specimens is not a temporal pattern as the oldest time bin (B/A) as well as Serravallian and Tortonian are well-sampled (also in terms of diversity). The above comparison of size classes does not support the presence of pull of the recent.

*Discussion*

In general, smaller organisms have a lower chance of being preserved in the fossil record as well as being discovered by paleontologists [5]. In our data set, there is the potential for smaller tortoises to be affected by preservation bias. However, the smallest species in our data set measures 9 cm in straight carapace length and, while there is still a difference in preservation likelihood between a small tortoise of 9 cm and a large tortoise of 1-2 m, size may not play as much of a factor as it does in other groups as the hard tortoise shell preserves well in the fossil record. Further, Rhodin et al. [6] compiled 121 species of Testudinidae in the fossil record since the beginning of the Pleistocene and 117 (97%) of those species are represented in our body size data set. Therefore, we consider our data set on body size, at least for the time period since the Pleistocene provides a robust representation of the actual fossil record of tortoises.



Figure S5: Comparison of absolute sample sizes of large and small tortoises in each time bin. Small tortoises have a higher sample size than large tortoises in every time bin.



Figure S6: Comparison of relative sample sizes of large and small tortoises in each time bin. Small tortoises make up at a minimum 55% of the sample size of each time bin.

1. **Analysis of age and latitudinal trends in body size**.

We applied a generalized additive model (GAM) to assess the temporal and latitudinal patterns in body size. We used a GAM because body size exhibited a nonlinear pattern with age. GAMs is a quadratic, penalized generalized linear model that uses smoother terms to fit the data and the smoothers are penalized regression splines. We fit the data using Gaussian distribution and the REML method.



Figure S7. Temporal and latitudinal trends in body size (SCL) based on a generalized additive model (GAM). The two smoothers were statistically significant (Age – F6.87,4.99 = 4.07, P = 0.0012,;Latitude – F6.87,7.63 = 7.27, P < 0.001)



Figure S8: Body size distribution in the late (A) and early (B) Quaternary. Colors indicate whether the taxa occur on islands or on the mainland. Boxplots show the median, interquartile range, and outliers. The dashed line at 1400 mm (and corresponding log value for the density plots) shows that several outliers in the early Quaternary exceed the size of extant giant tortoises (> 1500 mm) while all outliers in the late Quaternary reach at maximum 1300 mm, supporting our conclusion of the alteration of body size patterns in Testudinidae due to late Quaternary megafauna extinctions.



Figure S9: Comparison of tortoise body size based on temporal and spatial status, separately. Bold lines indicate medians, boxes indicate lower and upper quartiles, whiskers indicate largest and smallest observations and outliers represent extreme values. Mean straight carapace length per genera are depicted as grey circles with error bars indicating the respective standard deviation. A) Comparison of body size in extant and fossil taxa. Extant tortoises have a smaller mean body size than fossil taxa. B) Comparison of body size in mainland and insular taxa. Mainland tortoises have a smaller mean body size than insular taxa.



Figure S10: Illustration of how co-occurring trends of gigantism and miniaturization can lead to stasis on the family level. Our data set has many small taxa and fewer large taxa (see histogram), which result in a mean body size closer towards the smaller-bodied tortoise taxa rather than the giant tortoise taxa while also keeping fluctuations of mean body size low (see time-scale analysis). The body sizes displayed here are an example roughly based on the range of body sizes present in our data set to visualize the interaction between body size and abundance of different taxa.



Figure S11: Presence of genera and their body sizes throughout the time bins. A circle indicates the exact dating of a record in Ma whereas a line indicates a record was dated to occur within a time interval. Color indicates the mean body size of the genus, while size of the circle/width of the line indicates the sample size. Some genera are well sampled throughout several time bins, e.g. Hesperotestudo, Cheirogaster, while other are scarce, e.g. Impregnochelys, Ergilemys. In terms of mean body size, some genera exhibit large variation in body size over time, for example Megalochelys, ranging from smaller than 500 mm up to 2000 mm in SCL. Other genera are rather consistent in mean body size over time, e.g. Paleotestudo.

Supplemental Tables

Table S4: Multiple comparison test after Kruskal-Wallis. Adjacent time bin pairs are bolded. The last column indicates whether the pairwise comparison detected a significant difference (p < 0.05 = TRUE) between the time bins. The observed difference is the observed difference in median body size, the critical difference is the expected difference.

| Compared time bins | Significance level | Observed difference | Critical difference | Significant difference |
| --- | --- | --- | --- | --- |
| **Extant-Late Pleistocene** | **0.05** | **118.6398413** | **94.89971472** | **TRUE** |
| Extant-Middle Pleistocene | 0.05 | 84.14719976 | 92.6314415 | FALSE |
| Extant-Early Pleistocene | 0.05 | 69.31370092 | 89.90591171 | FALSE |
| Extant-Gelasian | 0.05 | 1.527905786 | 116.6698954 | FALSE |
| Extant-Piacencian | 0.05 | 115.4126984 | 139.2252844 | FALSE |
| Extant-Zanclean | 0.05 | 209.2506105 | 126.2647612 | TRUE |
| Extant-Messinian | 0.05 | 146.9516595 | 188.8112943 | FALSE |
| Extant-Tortonian | 0.05 | 57.7281746 | 96.53536573 | FALSE |
| Extant-Serravallian | 0.05 | 5.310803891 | 116.6698954 | FALSE |
| Extant-Langhian | 0.05 | 195.484127 | 168.3151615 | TRUE |
| Extant-Burdigalian/Aquitanian | 0.05 | 41.9781746 | 109.2172988 | FALSE |
| **Late Pleistocene-Middle Pleistocene** | **0.05** | **34.49264151** | **120.8488114** | **FALSE** |
| Late Pleistocene-Early Pleistocene | 0.05 | 49.32614035 | 118.7725736 | FALSE |
| Late Pleistocene-Gelasian | 0.05 | 117.1119355 | 140.1274982 | FALSE |
| Late Pleistocene-Piacencian | 0.05 | 3.227142857 | 159.3973999 | FALSE |
| Late Pleistocene-Zanclean | 0.05 | 90.61076923 | 148.2114745 | FALSE |
| Late Pleistocene-Messinian | 0.05 | 28.31181818 | 204.1410202 | FALSE |
| Late Pleistocene-Tortonian | 0.05 | 60.91166667 | 123.8665738 | FALSE |
| Late Pleistocene-Serravallian | 0.05 | 123.9506452 | 140.1274982 | FALSE |
| Late Pleistocene-Langhian | 0.05 | 76.84428571 | 185.3479022 | FALSE |
| Late Pleistocene-Burdigalian/Aquitanian | 0.05 | 76.66166667 | 133.9860799 | FALSE |
| **Middle Pleistocene- Early Pleistocene** | **0.05** | **14.83349884** | **116.968168** | **FALSE** |
| Middle Pleistocene-Gelasian | 0.05 | 82.61929397 | 138.6013848 | FALSE |
| Middle Pleistocene-Piacencian | 0.05 | 31.26549865 | 158.0574553 | FALSE |
| Middle Pleistocene-Zanclean | 0.05 | 125.1034107 | 146.7694426 | FALSE |
| Middle Pleistocene-Messinian | 0.05 | 62.80445969 | 203.0964899 | FALSE |
| Middle Pleistocene-Tortonian | 0.05 | 26.41902516 | 122.137448 | FALSE |
| Middle Pleistocene-Serravallian | 0.05 | 89.45800365 | 138.6013848 | FALSE |
| Middle Pleistocene-Langhian | 0.05 | 111.3369272 | 184.1968321 | FALSE |
| Middle Pleistocene-Burdigalian/Aquitanian | 0.05 | 42.16902516 | 132.3891903 | FALSE |
| **Early Pleistocene-Gelasian** | **0.05** | **67.78579513** | **136.7948569** | **FALSE** |
| Early Pleistocene-Piacencian | 0.05 | 46.09899749 | 156.4757112 | FALSE |
| Early Pleistocene-Zanclean | 0.05 | 139.9369096 | 145.0646693 | FALSE |
| Early Pleistocene-Messinian | 0.05 | 77.63795853 | 201.8679599 | FALSE |
| Early Pleistocene-Tortonian | 0.05 | 11.58552632 | 120.0834927 | FALSE |
| Early Pleistocene-Serravallian | 0.05 | 74.62450481 | 136.7948569 | FALSE |
| Early Pleistocene-Langhian | 0.05 | 126.1704261 | 182.8413573 | FALSE |
| Early Pleistocene-Burdigalian/Aquitanian | 0.05 | 27.33552632 | 130.4966924 | FALSE |
| **Gelasian-Piacencian** | **0.05** | **113.8847926** | **173.2438735** | **FALSE** |
| Gelasian-Zanclean | 0.05 | 207.7227047 | 163.0108886 | TRUE |
| Gelasian-Messinian | 0.05 | 145.4237537 | 215.1266249 | FALSE |
| Gelasian-Tortonian | 0.05 | 56.20026882 | 141.240351 | FALSE |
| Gelasian-Serravallian | 0.05 | 6.838709677 | 155.6972202 | FALSE |
| Gelasian-Langhian | 0.05 | 193.9562212 | 197.3822522 | FALSE |
| Gelasian-Burdigalian/Aquitanian | 0.05 | 40.45026882 | 150.1938023 | FALSE |
| **Piacencian-Zanclean** | **0.05** | **93.83791209** | **179.8453923** | **FALSE** |
| Piacencian-Messinian | 0.05 | 31.53896104 | 228.1474963 | FALSE |
| Piacencian-Tortonian | 0.05 | 57.68452381 | 160.3765945 | FALSE |
| Piacencian-Serravallian | 0.05 | 120.7235023 | 173.2438735 | FALSE |
| Piacencian-Langhian | 0.05 | 80.07142857 | 211.4983896 | FALSE |
| Piacencian-Burdigalian/Aquitanian | 0.05 | 73.43452381 | 168.3151615 | FALSE |
| **Zanclean-Messinian** | **0.05** | **62.29895105** | **220.477641** | **FALSE** |
| Zanclean-Tortonian | 0.05 | 151.5224359 | 149.2640685 | TRUE |
| Zanclean-Serravallian | 0.05 | 214.5614144 | 163.0108886 | TRUE |
| Zanclean-Langhian | 0.05 | 13.76648352 | 203.20108 | FALSE |
| Zanclean-Burdigalian/Aquitanian | 0.05 | 167.2724359 | 157.7628083 | TRUE |
| **Messinian-Tortonian** | **0.05** | **89.22348485** | **204.9065082** | **FALSE** |
| Messinian-Serravallian | 0.05 | 152.2624633 | 215.1266249 | FALSE |
| Messinian-Langhian | 0.05 | 48.53246753 | 246.9769095 | FALSE |
| Messinian-Burdigalian/Aquitanian | 0.05 | 104.9734848 | 211.1776944 | FALSE |
| **Tortonian-Serravallian** | **0.05** | **63.03897849** | **141.240351** | **FALSE** |
| Tortonian-Langhian | 0.05 | 137.7559524 | 186.1906706 | FALSE |
| Tortonian-Burdigalian/Aquitanian | 0.05 | 15.75 | 135.149512 | FALSE |
| **Serravallian-Langhian** | **0.05** | **200.7949309** | **197.3822522** | **TRUE** |
| Serravallian-Burdigalian/Aquitanian | 0.05 | 47.28897849 | 150.1938023 | FALSE |
| **Langhian-Burdigalian/Aquitanian** | **0.05** | **153.5059524** | **193.0707315** | **FALSE** |

**References Cited**

1. Gotelli NJ, Colwell RK. 2001 Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. *Ecol. Lett.* **4**, 379–391. (doi:10.1046/j.1461-0248.2001.00230.x)

2. Thompson GG, Withers PC. 2003 Effect of species richness and relative abundance on the shape of the species accumulation curve. , 6.

3. Jass CN, Cobb TP, Bell CJ. 2014 Regional, depositional, and chronologic comparisons of Pleistocene turtle richness in North America. *Chelonian Conserv. Biol.* **13**, 16–26. (doi:10.2744/CCB-1027.1)

4. Cooper RA, Maxwell PA, Crampton JS, Beu AG, Jones CM, Marshall BA. 2006 Completeness of the fossil record: Estimating losses due to small body size. *Geology* **34**, 241. (doi:10.1130/G22206.1)

5. Jablonski D, Roy K, Valentine JW, Price RM, Anderson PS. 2003 The impact of the pull of the Recent on the history of marine diversity. *Science* **300**, 1133–1135. (doi:10.1126/science.1083246)

6. Rhodin AGJ *et al.* 2015 Turtles and tortoises of the world during the rise and global spread of humanity: first checklist and review of extinct Pleistocene and Holocene chelonians. In *Conservation biology of freshwater turtles and tortoises: a compilation project of the IUCN/SSC tortoise and freshwater turtle specialist group*, pp. 1–66. (doi:10.3854/crm.5)