1 **Death is on Our Side: Paleontological Data Drastically Modify Phylogenetic**

2 **Hypotheses**

3

4 Nicolás Mongiardino Koch[a*] & Luke A. Parry[a]

5

6 *[a] Department of Geology & Geophysics, Yale University, 210 Whitney Avenue, New Haven, CT 06510,*

7 *USA.*

8 ** Corresponding author. Email: nicolas.mongiardinokoch@yale.edu. ORCID ID: https://orcid.org/0000-*

9 *0001-6317-5869.*

N Mongiardino Koch & LA Parry

10 **ABSTRACT**

11   Fossils are the only remaining evidence of the majority of species that have ever existed,

12   providing a direct window into events in evolutionary history that shaped the diversification of life on

13   Earth. Phylogenies underpin our ability to make sense of evolution but are routinely inferred only from

14   data available from living organisms. Although extinct taxa have been shown to add crucial information

15   for inferring macroevolutionary patterns and processes including ancestral states, paleobiogeography and

16   diversification dynamics, the role that fossils play in inferring the tree of life itself is controversial. Since

17   the early years of phylogenetic systematics, different studies have dismissed the impact of fossils due to

18   their incompleteness, championed their ability to overturn phylogenetic hypotheses or concluded that

19   their behavior is indistinguishable from that of extant taxa. Here we show paleontological data has a

20   remarkable effect in phylogenetic inference. Fossils often have higher levels of topological influence than

21   extant taxa, while inducing unique topological rearrangements. Previous studies have proposed a suite of

22   explanations for the topological behavior of fossils, such as their retention of unique morphologies or

23   their ability to break long branches. We develop predictive models that demonstrate that the possession of

24   distinctive character state combinations is the primary predictor of the degree of induced topological

25   change, and that the relative impact of taxa (fossil and extant) can be predicted to some extent before any

26   analysis. Our results bolster the consensus of recent empirical studies by showing the unique role of

27   paleontological data in phylogenetic inference, and provide the first quantitative assessment of its

28   determinants, with broad consequences for the design of taxon sampling in both morphological and total-

29   evidence analyses.

30

31   **Keywords:** phylogeny, morphology, fossils, parsimony, Bayesian.

32   The overwhelming majority of species produced through the diversification of life on Earth are

33   now extinct (Simpson 1952; Raup 1992). Even though much of this diversity is lost, a significant

34   proportion has been preserved in the fossil record, which often provides the most direct evidence of

35   evolution in deep-time. Although evolutionary inferences are routinely performed using data from extant

36   taxa alone, this can often provide only a partial, or even misleading, view of evolutionary processes and

37   patterns. Consequently, paleontological data can not only expand the range of evolutionary questions

38   accessible to inquiry, but also drastically improve estimates of evolutionary phenomena. Both simulations

39   and empirical case studies attest to the positive effect that incorporating extinct diversity can have on the

40   reconstruction of ancestral states (Finarelli and Flynn 2006; Finarelli and Goswami 2013; Puttick 2016),

41   rates and modes of macroevolution (Slater et al. 2012; Bokma et al. 2015; Mitchell 2015; Schnitzler et al.

42   2017), diversification dynamics (Liow et al. 2010; Quental and Marshall 2010; Rabosky 2010; Mitchell et

43   al. 2018) and historical biogeography (Wood et al. 2012; Field and Hsiang 2018). Fossils also provide the

44   most direct and widely employed evidence used to time-calibrate phylogenies (Laurin 2012; Dos Reis et

45   al. 2016), a key first step in most of modern comparative biology.

46   Despite the key role of the fossil record in understanding evolutionary history, the degree to

47   which extinct taxa contribute to the inference of phylogenetic relationships has been much more

48   controversial. Although this is often discussed in the context of the merits (and caveats) of reconstructing

49   phylogeny using morphology (e.g., Scotland et al. 2003; Jenner 2004), the debate has a longer history.

50   Hennig (1966) first suggested that the higher proportion of missing data in fossils should compromise

51   their usefulness for elucidating phylogenetic relationships, a view shared by other early systematists

52   (Løvtrup 1977; Ax 1987). From this perspective, fossils hold a subsidiary role, and their significance

53   should only be discussed in light of phylogenies built from extant taxa (Patterson 1977; Nelson 1978).

54   This assumes, either implicitly or explicitly, that fossils do not modify tree topology (Hennig 1981;

55   Patterson 1981; Goodman 1989), and can be grafted onto phylogenies inferred using data from living

56   species.

57   However, extinct organisms have several characteristics that should make them especially

58   important for inferring accurate trees from morphological data. Fossils allow taxon sampling to be

59   extended beyond the reach of molecular data, and preserve character state combinations not present

60   among living clades, potentially modifying homology statements, character polarity and tree rooting

61   (Doyle and Donoghue 1987; Marshall and Schultze 1992; Novacek 1992a; Wilson 1992; Smith 1994,

62   1998; Forey and Fortey 2001; Edgecombe 2010). Fossil terminals can also occupy unique phylogenetic

63   positions, lying close to divergence events, in the midst of ancient and rapid radiations, or subdividing the

64   long branches that often separate morphologically distant extant lineages (Doyle and Donoghue 1987;

65   Gauthier et al. 1988; Donoghue et al. 1989; Huelsenbeck 1991; Sumrall 1997; Smith 1998; Wills and

66   Fortey 2000; Smith and Turner 2005; Mayr 2006). Their morphology may often resemble that of the

67   common ancestors from which extant clades originated, being less modified by subsequent evolutionary

68   history (Beck and Baillie 2018; Asher et al. 2019). Furthermore, the proportion of missing data does not

69   necessarily compromise the phylogenetic placement of terminals, nor the overall resolution of

70   phylogenetic analyses (Kearney and Clark 2003; Wiens 2003a, b; Prevosti and Chemisquy 2010;

71   Pattinson et al. 2014), and incomplete terminals can in fact increase topological accuracy (Huelsenbeck

72   1991; Wiens 2005). Fossils are therefore expected to have a strong topological impact, and early claims

73   for their dismissal were rapidly falsified by several case studies (Gauthier et al. 1988; Donoghue et al.

74   1989; Doyle and Donoghue 1992; Novacek 1992b; Wilson 1992; Cloutier and Ahlberg 1995; Smith

75   1998).

76   Nevertheless, it is difficult to draw general conclusions regarding the impact of paleontological

77   data from individual studies, and it further remains unclear whether fossil terminals modify phylogenetic

78   trees above a baseline of expected change given increased taxon sampling. The only empirical study

79   addressing these issues analyzed 45 empirical morphological matrices and concluded that there was no

80   significant difference in the degree of induced topological change between fossil and extant terminals

81   (Cobbett et al. 2007). This conclusion was supported by first-order taxon jackknifing experiments (i.e.,

82   comparison of topologies obtained with and without a focal taxon). Although the authors interpreted this

83   result as strongly supporting the inclusion of fossils in phylogenetic analyses, this was mostly justified in

84   the lack of a distinctive behavior by paleontological data. This conclusion not only conflicts with the

85   literature cited above, much of which considers paleontological data to be unique (in either a negative or

86   positive way), but also with a number of more recent studies where prominent and long-standing cases of

87   conflict between morphological and molecular trees have been claimed to be resolved through the

88   addition of key fossils (Legg et al. 2013; Parry et al. 2016; Coiro et al. 2018; Simões et al. 2018;

89   Miyashita et al. 2019). However, two points should be noted regarding this discrepancy: 1) Given their

90   experimental design, Cobbett et al. (2007) never tested the degree to which fossils overturn relationships

91   inferred *exclusively* from living taxa, and 2) recent case studies do not necessarily claim that fossils have a

92   strong topological effect, rather that the type of change induced is not generated by increasing sampling

93   among extant taxa. These two aspects of the interaction between paleontological and neontological data in

94   phylogenetic studies have never been systematically explored.

95   Given that relationships among living clades are now routinely inferred from molecular data, this

96   discussion has been deemed obsolete (Scotland et al. 2003). However, even in the genomic era,

97   morphology will remain the only means to resolve the relationships among extinct species, as well as

98   their position relative to extant clades (Giribet 2015; Lee and Palci 2015). Incorporating the information

99   preserved in the fossil record into phylogenetic frameworks not only greatly improves the behavior of

100   phylogenetic comparative methods (Slater and Harmon 2013; Goswami et al. 2016; see above), but also

101   allows the use of tip-dating approaches to divergence time estimation, which require fewer assumptions

102   and make better use of stratigraphic data than more traditional node-dating methods (Ronquist et al.

103   2012a; Heath et al. 2014; Lee and Palci 2015; Zhang et al. 2016). Furthermore, it has been shown that

104   morphological data has the power to modify tree topology in total-evidence analyses, even when

105   constituting a minimal fraction of the data (Wiens et al. 2010; Bapst et al. 2018; Cascini et al. 2019).

106   Moreover, phylogenies inferred from molecular data are far from stable for all nodes in the tree of life,

107   with several recent examples of phylogenomic data generating conflicting topologies that have sparked

108   controversy (e.g., Ballesteros and Sharma 2019; Philippe et al. 2019). Therefore, morphological data

109   remains an important source of independent data, and understanding the impact of including fossil taxa in

110   phylogenetic analyses remains paramount to obtaining a complete and accurate picture of evolutionary

111   history.

112   Here we employ multiple empirical large-scale morphological matrices to explore the degree and

113   type of topological change exerted by fossils on trees of extant lineages. As probabilistic approaches to

114   morphological inference have become increasingly common, we extend previous efforts by analyzing

115   results obtained under both maximum parsimony (MP) and Bayesian inference (BI). Finally, we evaluate

116   for the first time several potential determinants of the topological impact of taxa, allowing us to build a

117   framework that can predict whether a terminal will have a strong effect on tree topology.

118

119   **MATERIALS & METHODS**

120   **Dataset Selection and Subsampling Procedure**

121   Datasets were selected for their size, large number of fossil and extant taxa, and relatively small

122   proportions of missing data. We focused on large-scale empirical morphological matrices (also known as

123   'phenomic' matrices (O'Leary and Kaufman 2011; O'Leary et al. 2013)), as these are expected to

124   generate better constrained distributions of optimal topologies. Datasets were also required to contain

125   reasonably large numbers of both fossils and extant taxa, therefore allowing for the comparison of their

126   topological effects. Given that some of our analyses explored the topological changes induced on trees

127   built *only* from extant terminals, high numbers of these were especially important. Finally, fossil taxa had

128   to be coded for a significant fraction of the total number of characters. We translated these requirements

129   into a set of rules, employing matrices that had: 1) a number of characters larger than the number of taxa;

130   2) at least 40 extant terminals; 3) at least 20 fossil terminals; 4) a fraction of missing data among fossils

131   less than 80%. Six datasets (shown in Table 1 and available as SI File 1) satisfied these criteria, and were

132   thus used for all analyses. Preliminary analyses with a larger number of datasets revealed that either the

133   subsampling procedure or the subsequent statistical analyses (see below) could not be performed if these

134   requirements were not enforced. In each case, a single outgroup was included to root all trees. Characters

135   considered ordered by the authors were analyzed as such. If datasets were modified in any other way,

136   details can be found in SI File 2.

137   Datasets were imported into the R statistical environment (R Core Team 2019) using function

138   ReadMorphNexus from package *Claddis* v. 0.3 (Lloyd 2016). For each dataset, 25 initial

139   pseudoreplicated matrices composed of n randomly selected extant taxa were generated (Fig. 1, step 1),

140   and phylogenetic inference was performed before and after the incorporation of further terminals. Unlike

141   previous efforts that measured the topological impact of adding individual terminals (Cobbett et al. 2007),

142   we explored the topological effects induced by the simultaneous addition of groups of terminals (of size

143   m) to these initial replicates (Fig. 1, step 2). We believe this approach to more accurately reflect the way

144   in which morphological datasets grow with time, as well as providing greater subsampling flexibility and

145   increased statistical power. The values of n and m were determined for each matrix following the

146   approach described in SI File 2.

147   The groups of added terminals were of one of three different types: fossil, extant and

148   pseudoextinct, with the last two providing different bases with which to compare the effects induced by

149   fossils. The first two of these groups were generated by selecting at random among the fossil or extant

150   taxa left unsampled. However, the direct comparison of the topological effect of fossil and extant

151   terminals might be confounded by the systematic difference in the amount of data coded between them

152   (see Table 1). Therefore, a third group of terminals (pseudoextinct, following the nomenclature of

153   Springer et al. 2007) was generated by selecting at random m of the unsampled extant taxa and pairing

154   each with a randomly selected fossil. Characters missing in the fossil were then deleted from the extant

155   terminal with which it was paired (see Pattinson et al. 2014 for an equivalent approach). This procedure

7

156   preserves the pattern of missing data found in fossils, and is expected to generate better fossil analogs

157   compared to the deletion of random characters given how morphological structures differ in both their

158   preservation potential and phylogenetic signal (Sansom and Wills 2013; Mounce et al. 2016; Sansom and

159   Wills 2017; Sansom et al. 2017). Note however that the total amount of data in pseudoextinct taxa is

160   expected to be slightly lower than that of fossils, as extant taxa already have missing data that might be

161   coded in some fossils. Comparison of the amounts of missing data between these three groups of

162   terminals can be found in Figure S1 (SI File 3). While fossils have on average 10.5%-30.9% more

163   missing data than extant taxa (depending on the dataset), pseudoextinct taxa have on average 4.6%-8.5%

164   more missing data than fossils. This represents a 2.3-4.6 times reduction in the discrepancy of missing

165   data between compared groups. Given how missing data is known to affect the topological impact of

166   terminals (Huelsenbeck 1991; Wiens 2003b, 2005), pseudoextinct terminals likely provide a better

167   baseline for the topological effect that can be expected from fossils.

168   For each of the 25 initial replicates per dataset, terminals of the different types were added in a

169   stepwise manner in groups of size m until unsampled terminals were exhausted. Furthermore, three

170   iterations of this procedure were performed in order to estimate the average effect of taxon addition to a

171   given taxonomic composition. This entire subsampling procedure generated between 2,275 and 3,625

172   morphological matrices per dataset.

173

Phylogenetic Inference

175   Phylogenetic inference on these matrices was performed under both MP and BI (Fig. 1, step 3).

176   Although MP was historically favored as a method of inference from morphological data, several studies

177   have argued probabilistic approaches of inference, and more specifically BI under variants of the Mk

178   model (Lewis 2001), might outperform parsimony-based methods (Wright and Hillis 2014; O'Reilly et al.

179   2016; Puttick et al. 2017). These results have received different criticisms (Goloboff et al. 2018, 2019;

180   Smith 2019), but ultimately rely on analyses performed on simulated data, under conditions where the

181   accuracy of the inference can be evaluated. It still remains unclear whether these results can be

182   extrapolated to the analysis of empirical matrices (Sansom et al. 2018; Schrago et al. 2018; Goloboff et al.

183   2019), but it can be argued that they have established BI under the Mk model as a valid alternative to MP.

184   Furthermore, given the flexibility of Bayesian methods to combine different data sources and calibrate

185   divergence times, BI of morphological datasets is likely to become common in the future. We here treat

186   these as valid methods of inference from morphological data and explore whether any of our results varies

187   depending on method choice, extending previous attempts which had relied exclusively on MP (Cobbett

188   et al. 2007).

189   Inference under MP was performed using TNT 1.5 (Goloboff and Catalano 2016) under equal

190   weights, using driven tree searches with five initial replicates that were subject to new technology search

191   heuristics (Goloboff 1999; Nixon 1999). Search was continued until minimum length was found twenty

192   times. TBR branch swapping was then performed on the topologies in memory, retaining a maximum of

193   10,000 maximum parsimony trees (an example TNT batch script to perform tree searches can be found as

194   SI File 4). Bayesian analyses were performed in MrBayes 3.2.6 (Ronquist et al. 2012b) under the Mk+Γ

195   model with a correction for coding only parsimony-informative characters (MK$_{parsinf}$). Two runs of four

196   Metropolis-coupled MCMC chains were continued for either one million generations or until a standard

197   deviation of split frequencies < 0.01 was attained. We considered this condition to represent an accurate

198   sampling of the posterior distribution of topologies (as have other authors; e.g., Puttick et al. 2019). Trees

199   were sampled every 100 generations and the initial 50% was discarded as burn-in, therefore retaining a

200   maximum of 10,000 posterior topologies.

201

202   **Estimating and Analyzing Topological Impact**

203   Optimal topologies from MP and BI runs were imported into R using functions from packages

204   *ape* (Paradis and Schliep 2018) and *TNTR* (Matzke 2015), and a random sample of up to 100 topologies

205   was drawn from each analysis. Topologies before and after the addition of groups of taxa were compared

206   directly by calculating their normalized Robinson-Foulds (RF) distance (Robinson and Foulds 1981) after

207   pruning to the same taxonomic composition (Fig. 1, step 4 and 5). All trees analyzed were thus fully

208   bifurcating, avoiding the issue of computing similarity when topologies differ in their level of resolution

209   (Smith 2019). For pairs of unrooted bifurcating trees the RF distance is equivalent to the number of

210   bipartitions not shared (Penny and Hendy 1985), and was normalized to the total number of bipartitions in

211   both. RF distances were computed using package *phangorn* v. 2.5.5 (Schliep 2011). In order to derive a

212   single metric that reflects the distance between sets of trees, we measured both the mean and minimum

213   RF distance of a given tree to all trees in the other set, and took the average of the resulting values. Given

214   the strong linear correlation between these two estimates (Fig. S2 of SI File 3; all Pearson's $r > 0.89$, $P <$

215   10^{-259}), analyses employed the average of distances to the nearest neighbor. This metric was favored as it

216   attains a value of 0 if the sets being compared are identical (Cobbett et al. 2007), and allows for a more

217   straightforward comparison of sets of trees derived from MP and BI, which are expected to systematically

218   differ in their within-set distances (O'Reilly et al. 2016; Schrago et al. 2018).

219   We explored the degree of topological change generated by taxon addition by summarizing RF

220   distances across iterations and analyzing them with generalized linear models. The effect produced by the

221   incorporation of fossils to topologies of extant taxa was compared in multiple ways with that introduced

222   by further increasing sampling of extant taxa. First, we used the number (log-transformed) and type

223   (fossil/extant/pseudoextinct) of added terminals as predictors of topological change. Extant and

224   pseudoextinct terminals provided two alternative baselines of topological change introduced by increasing

225   taxon sampling, and so the relative effect of these was compared to that of fossils by fitting separate linear

226   models. However, as already explained, this approach can be problematic as the proportion of missing

227   data differs between these types of terminals (Fig. S1 of SI File 3). To more directly account for this, a

228   second analysis was performed using the amount of added information (i.e., the sum of the number of

229   coded characters across added terminals, also log-transformed) as the continuous predictor in the linear

230   model. Given that this approach fully accounts for differences in the amount of coded characters, only

231   data for fossil and extant terminals was analyzed in such a way.

232   In each case, models with and without an interaction term were further compared using likelihood

233   ratio tests (LRTs) with a significance level (α) of 0.1. In case of a significant interaction term, regions of

234   significance were estimated using the Johnson-Neyman approach with package *jnt* (Middleton 2016);

235   otherwise results presented are those of ANCOVAs. The Johnson-Neyman procedure finds the values of

236   the continuous predictor for which there is a transition from significant to non-significant differences

237   among the groups of the categorical predictor. In cases of non-significant differences between the

238   different types of added terminals, a power analysis using the method of Borm et al. (2007) was

239   performed (with a standard power level of 0.8).

240   Finally, we also relied on the construction of treespace (Hillis et al. 2005) to explore the type of

241   topological change exerted by the addition of extant and fossil taxa. This approach is based on estimating

242   all pairwise RF distances and decomposing them into a two-dimensional space using principal coordinate

243   analysis (PCoA) (Jombart et al. 2017), providing a summary of the diversity of optimal topologies that

244   are obtained as different types of terminals are incorporated into the inference procedure. Given the

245   combinatorial nature of this method, we further subsampled at most 20 trees per analysis, and did not

246   include topologies obtained after the addition of pseudoextinct terminals. Different treespaces were built

247   for each combination of dataset, inference method and initial replicate. After obtaining the location of

248   each tree in the first two PCoA axes, sets of trees from different phylogenetic analyses were collapsed to

249   their centroid, and the relative position of centroids generated by the addition of fossil and extant

250   terminals were compared using convex hulls and permutational (non-parametric) MANOVAs. Convex

251   hulls provide a straightforward way in which to estimate the degree to which fossil and extant taxon

252   addition results in the exploration of similar regions of treespace. Topological similarity was expressed as

253   the area of overlap between the two convex hulls divided by the sum of the total area covered by the two,

254   calculated using package *geometry* (Habel et al. 2019). Permutational MANOVA (Anderson 2001) is an

255   analysis of variance using distance matrices, and employs a permutational test to evaluate significant

256   sources of variation. The test was run for each treespace individually with package *vegan* (Oksanen et al.

257   2019), using Euclidean distances and evaluating significance with 10,000 replicates.

258   Given that the 25 initial replicates of each dataset do not share the same taxonomic composition,

259   and therefore cannot be placed in a common treespace, a visual summary per dataset was obtained by

260   using Procrustes superimposition. The different treespaces of each dataset were rotated so as to minimize

261   the sum of squared distances between centroids obtained after adding the same number and type of

262   terminals across replicates. In order to further facilitate interpretation of the plots, the different treespaces

263   were then translated so that the average position of trees produced before taxon addition was at the origin.

264

Estimating and Analyzing Taxon Influence

266   We also analyzed the topological effect induced by individual taxa using first-order taxon

267   jackknifing. Unlike previous efforts however (Cobbett et al. 2007; Mariadassou et al. 2012; Denton and

268   Goolsby 2018), we first standardized the size of the matrices from which taxa were jackknifed. We

269   consider this approach to be superior for multiple reasons. First of all, it is evident from our results (as

270   well as from previous analyses; e.g., Huelsenbeck 1991; Wiens 2003b, 2005; Prevosti and Chemisquy

271   2010) that the topological impact of a taxon will strongly depend on the size of the dataset to which it is

272   added (see Results). If differences in size across datasets are not accounted for, these are bound to distort

273   the impact of taxa when combining results from different datasets. Furthermore, standardizing matrix size

274   has the added benefit that the impact of taxa can be measured against different taxonomic backgrounds.

275   The resulting average impact is likely to be a better predictor of the effect a taxon will exert if

276   incorporated into a new morphological matrix. Finally, it also allows for more intense replication and

277   increased statistical power, while avoiding prohibitive computational times.

278   For each taxon, 30 matrices composed of the outgroup and 38 other randomly selected ingroup

279   taxa were built. Unlike the approach described above, this subsampling scheme produces matrices

280   composed of a mix of extinct and extant terminals. Phylogenetic inference was performed before and after

281   the addition of the focal taxon to each of these matrices. A total of 27,930 MP and BI analyses were

282   performed. Both phylogenetic inference and estimation of topological impact were performed as

283   explained above. The impact of taxa under BI and MP was significantly correlated (Pearson's r = 0.40, *p*

284    $< 10^{-16}$, Figs. S3 of SI File 3). Topological impact was then averaged across replicates to obtain a single

285   measure per taxon. Given that we found a very strong structuring of topological impact by dataset (Figs.

286   S4 and S5 of SI File 3), we subtracted from the estimated impact of each taxon the average impact for its

287   dataset of provenance. This rescaling produces a metric that expresses deviations from an average impact

288   and accounts for the difference in stability across datasets.

289   Generalized linear models were again employed to survey potential determinants of topological

290   impact. We attempted to quantify all possible determinants previously hypothesized in the literature to

291   affect a taxon's topological impact. Some of these properties were estimated directly from the

292   morphological matrices, while others were measured on BI posterior topologies (as they required branch

293   lengths). From the matrices, we measured the proportion of missing/inapplicable data and two measures

294   of morphological distance, which we refer to as 'distinctiveness' and 'uniqueness'. These correspond to

295   the minimum and average morphological distance to all other taxa, respectively, and were estimated using

296   the maximum observable rescaled distance (Lloyd 2016). Among the tree-based metrics we quantified—

297   for each focal taxa—the patristic distance to the root of the tree ('primitiveness'), the length of its

298   terminal branch ('autapomorphies'), as well as three measures designed to capture different conditions

299   that might lead to long-branch attraction scenarios. These included the change induced by the focal taxon

300   on the mean and variance of branch lengths, as well as the variance of root-to-tip distances ('clock-

301 likeness'). Models also included the taxon type (fossil/extant) as an additional predictor. Further details

302 on how these measures were quantified can be found in SI File 2. In order to explore the relevance of

303 these predictors, stepwise variable selection in both directions was performed, once again comparing

304 models using LRTs with $\alpha = 0.1$.

305 Finally, we also explored the determinants of topological impact using random forest models.

306 Random forests were chosen given their robustness to multicollinearity and deviations from normality, as

307 well as their ability to accommodate non-linear and local relationships. Both the entire set of independent

308 variables and the subset that can be estimated without running a phylogenetic analysis (i.e., all three

309 matrix-based variables and the taxon's type) were employed. These were used as predictors of both the

310 topological impact score (regression forest), as well as a recoded binary variable grouping taxa into those

311 with above and below average impact (classification forest). Models were fit using package

312 *randomForest* (Liaw and Wiener 2002), growing 5,000 trees and estimating accuracy of prediction using

313 out-of-bag errors. The number of predictors tried at each step was automatically determined by the

314 algorithm.

315 R code to perform all of the procedures detailed above (other than phylogenetic inference) and

316 perform the statistical analyses is available as SI File 5. It can be used to replicate all stages of the study

317 alongside SI File 1 (original datasets) and SI File 6 (which contains all results).

318

319 **RESULTS**

320 Our results show that the average degree of induced topological change strongly depends on the

321 logarithm of the amount of information added to the analysis, measured as either the number of taxa or

322 the number of coded cells incorporated to the matrix (Figs. 2, 3 and S6 of SI File 3; linear regressions, all

323 $R^2 > 0.86$, $P < 8e$-5). As the taxonomic sampling of phylogenetic analyses increases, it becomes

324 progressively less likely that new data will further modify tree topology. The relative effect of adding new

325   extant or fossil terminals proved highly dependent on the dataset, method of inference and number of

326   added taxa (Fig. 2), although there was an overall stronger evidence for the addition of extant taxa

327   generating higher topological impacts. Six of twelve analyses performed (i.e., combinations of datasets

328   and inference procedures) showed a significantly higher impact of extant taxa across most/all of the range

329   analyzed, while the opposite was only true for one analysis (echinoid dataset under BI; see also Fig. S6 of

330   SI File 3). In stark contrast, fossil terminals had a significantly higher topological effect than

331   pseudoextinct terminals, across all matrices and methods of inference (Fig. 3). This discordance reveals

332   the strong relationship between topological impact and levels of missing data, and questions whether

333   fossil and extant terminals are directly comparable.

334   For this comparison to be meaningful, statistical models need to account for the fact that fossil

335   and extant terminals do not contain equivalent amounts of information (see Table 1 and Fig. S1 of SI File

336   3). One possible way of achieving this is to compare the topological effect induced by the addition of

337   paleontological and neontological data (rather than the number of terminals) to the inference procedure.

338   We defined this variable as the sum of the number of characters that were neither missing nor

339   inapplicable across the added terminals. From this perspective, differences in the behavior of these types

340   of data are evident, with four matrices showing a stronger topological impact of paleontological data and

341   two (Coccomorpha and Panarthropoda) a stronger effect of neontological data (Fig. 4). No systematic

342   difference was found between the results obtained for MP and BI, which were very similar across

343   datasets. The sole difference found was the behavior of the different types of data for the lemuriform

344   dataset, which proved significant under BI but not under MP. This likely stems from a lack of statistical

345   power, as this is the dataset with the smallest number of taxa, and therefore the fewest datapoints. In fact,

346   a power analysis revealed significance should be attained with an increase of 16% in sample size. The

347   lack of significance is therefore not given much weight, especially as the relative effects of

348   paleontological and neontological data are the same for both inference procedures.

349   We further compared the generated topologies through the use of treespaces, allowing us to

350   explore the type of topological change induced by fossils and test if it systematically differed from that

351   generated through the addition of further extant terminals (Fig. 5). As the trees obtained from the multiple

352   replicates performed on each dataset do not share the same terminals (see Fig. 1 and Materials &

353   Methods) and therefore cannot be represented in a common treespace, each panel of Figure 5 consists of

354   the optimal juxtaposition of multiple distinct treespaces. The almost complete lack of overlap between the

355   clouds of dots seen in Figure 5 is therefore even more pronounced if individual treespaces are scrutinized

356   individually. Among these, convex hulls for topologies obtained after extant and fossil addition were

357   completely disjunct 19.3% of the time, and otherwise showed an average overlap of only 9.8% of the total

358   area covered. In fact, the relative positions occupied by these topologies were significantly different in

359   95% of the cases (permutational MANOVA, $p < 0.05$), with 13 of the 15 non-significant differences

360   restricted to the MP analysis of the panarthropod dataset.

361   It is evident from this suite of analyses that fossil taxa have a unique behavior in phylogenetic

362   inference (Figs. 2-5). However, this result raises the question of what determines the topological impact

363   of a taxon. To address this question, we resorted to first-order taxon jackknifing experiments, extending

364   previous efforts by assessing the topological effect of each terminal on multiple smaller matrices of the

365   same size and composed of randomly sampled extant and fossil taxa (see Materials & Methods). In

366   complete agreement with previous results (Cobbett et al. 2007), we found very small differences between

367   fossil and extant terminals in their relative effects on topology. The type of terminal (fossil/extant) was

368   not a significant predictor of topological impact for BI (t-test, $p = 0.91$), while fossils had a slightly

369   smaller than expected impact under MP ($p = 0.04$). This result was mostly driven by the mammal dataset,

370   the only individual matrix for which differences were significant ($p = 0.001$). However, this approach

371   once again relies on comparing groups of terminals that systematically differ in their proportion of

372   missing data (Table 1, Fig. S1 of SI File 3). After accounting for this difference, fossils were shown to

373   impact topology significantly more than extant terminals above a threshold of 43-45% missing data

374 (depending on the method of inference, Fig. 6a and Fig. S7 of SI File 3), a condition satisfied by over

375 78% of fossils in our dataset. Extant taxa are expected to have a significantly higher impact than fossils

376 only in the very narrow condition of no missing data (although this requires extrapolation, and holds true

377 only for MP, see Fig. S7 of SI File 3).

378   We quantified several other matrix and tree-based properties of terminals and explored their

379 usefulness as predictors of topological impact using generalized linear models. The obtained best-fit

380 models under BI and MP were extremely similar, yet explained widely different proportions of total

381 variance (adjusted R^2 of 0.06 and 0.31, respectively). For this reason we focus here only on the results for

382 MP. All of the results obtained under BI are placed in SI File 3, and a discussion of the differences

383 between methods can be found in SI File 2.

384   Our analysis revealed the presence of atypical morphologies as a major driver in a taxon's

385 topological impact (Figs. 6b and S8 of SI File 3). Organisms displaying character state combinations

386 distinct from those of all other taxa strongly modified topology, while completely unique morphologies

387 had a significantly lower impact. The amount of missing data and the degree of evolutionary change also

388 had negative effects, with taxa displaying 'primitive' morphologies (i.e., those inferred to have undergone

389 the least amount of change since the last common ancestor of the ingroup) and possessing fewer

390 autapomorphies having stronger effects on topology. Furthermore, terminals that increased the average

391 branch length of trees and decreased the clock-likeness of morphological evolution also showed a higher

392 impact. It is possible that these taxa are generating conditions conducive to long-branch attractions.

393 Finally, even after accounting for the aforementioned variables, fossils still induced stronger topological

394 changes than extant taxa, suggesting that the peculiarities of paleontological data are still not completely

395 captured by our model.

396   The topological impact of terminals has been claimed to be unpredictable (Cobbett et al. 2007),

397 yet to our knowledge no attempt at directly predicting it has been performed. Given how random forest

398 models using the same set of predictors were able to explain an even higher proportion of variance than

399 linear models (44% in the case of MP, see Table S1 of SI File 7), we tested the accuracy of random

400 forests to predict whether a taxon had a higher or lower-than-average topological impact. Prediction was

401 carried out using only the set of four variables that are tree-independent and can therefore be estimated

402 before running a phylogenetic analysis (distinctiveness, uniqueness, proportion of missing data and type

403 of taxon). Even using this reduced set of variables, the model attained a classification accuracy of 72%

404 (Table S2 of SI File 7), indicating that the topological impact of taxa in morphological phylogenetic

405 analysis has a strongly predictable component.

406

407 **DISCUSSION**

408   For decades, systematists have disagreed on the relevance that the fossil record has for inferring

409 phylogenetic relationships. By focusing on just a few of the many ways in which paleontological and

410 neontological data differ, fossils were portrayed to be either vital or irrelevant for phylogenetic inference,

411 as well as everything in between. However, much of this discussion was based on either first principles or

412 individual case studies. The only attempt to systematically assess the relevance of fossils suggested that

413 differences in the behavior of extant and extinct taxa might have been exaggerated, and that fossils do not

414 differ systematically from extant terminals, at least not in terms of their average topological effect.

415   Our results are able to reconcile these seemingly incompatible claims about the nature of the

416 fossil record and its relevance for systematics. All of our analyses reinforce the idea that the information

417 preserved in the morphology of extinct organisms is unlike data that can be obtained from the study of

418 living species, and that its inclusion in phylogenetic inference has strong consequences. However, this is

419 evident only after the fragmentary nature of fossils is accounted for. Missing data does in fact affect

420 topological impact (Fig. 6a), something that had already been demonstrated through simulations

421 (Huelsenbeck 1991; Wiens 2003b, 2005), although not always confirmed using empirical datasets

422 (Denton and Goolsby 2018). While previous studies had directly compared fossil and extant taxa, a better

423   understanding of their behavior required accounting for the different amount of information they contain.

424   When this is done, paleontological data tends to have stronger topological effects than neontological data

425   (Figs. 2-4, 6, S6-S8). Furthermore, adding fossils induces topological changes that are never obtained

426   through better taxonomic sampling of extant taxa (Fig. 5), a phenomenon that likely underlies their ability

427   to resolve conflicts between morphological and molecular phylogenies (Legg et al. 2013; Beck and

428   Baillie 2018; Simões et al. 2018; Asher et al. 2019).

429   Why do fossils show such a remarkable effect on phylogenetic analysis? Our results show

430   morphological distinctiveness to be the main predictor of topological impact (Figure 6b). The fossil

431   record contains a wealth of examples of extinct organisms with morphologies wholly unlike those we see

432   in the modern biosphere, and whose appearance we would not be able to predict even with the closest

433   scrutiny of the morphology of extant taxa. Examples of such organisms are found throughout geological

434   time, from the 'weird wonders' of the Cambrian to the diversity of extinct archosaurs that populate the

435   stem lineage of birds. The unique combination of characters in such organisms has long been recognized

436   to modify character polarity, reveal hidden homoplasies (or cryptic homologies) and break long branches

437   by revealing the sequence of character acquisition (Gauthier et al. 1988; Donoghue et al. 1989; Smith

438   1998; Edgecombe 2010). However, taxa with completely unique morphologies exhibit decreased

439   topological impact, likely attaching to a topology without modifying character optimization. In such cases

440   of extreme morphological modifications, such as those separating hexapods and crustaceans (Legg et al.

441   2013), or cetaceans from their ungulate relatives (Spaulding et al. 2009), fossil stem groups with

442   intermediate morphologies are required to correctly articulate these clades. Finally, the fossil record

443   provides our only access to morphologies that lie close to the origin of ancient lineages and are less

444   burdened by character state changes imposed by subsequent evolutionary history, properties that further

445   increase the topological impact of individual taxa.

446   Fossils not only play a critical role when inferring evolutionary processes from phylogenetic trees

447   (Slater et al. 2012; Goswami et al. 2016), but are clearly also influential in the construction of the trees

19

448   themselves. Beyond recognizing the unique role of paleontological data in phylogenetic inference, our

449   analyses prove that topological impact has strong determinants. Consideration of these properties can

450   provide new ways for optimizing the phylogenetic design of total-evidence analyses, aiding in the

451   generation of phylogenies that better incorporate the vast records of evolutionary history preserved in the

452   morphology of extant and extinct lineages. Given that genomic data is yet to resolve a stable phylogeny

453   for all branches in the tree of life, phylogenetic analyses of morphology incorporating fossil taxa remain

454   an important and independent framework for unraveling evolutionary history.

455

SUPPLEMENTARY MATERIAL

457   Data available from the Dryad Digital Repository: http://dx.doi.org/10.5061/dryad.[NNNN].

458

ACKNOWLEDGEMENTS

463

AUTHOR CONTRIBUTIONS

465   NMK designed the study and wrote all code. NMK and LAP gathered the original datasets and ran

466   the analyses. NMK analyzed the data and wrote the manuscript with contributions from LAP.

467

COMPETING INTERESTS

469   The authors declare no competing financial interests.

470   **LITERATURE CITED**

471   Anderson M.J. 2001. A new method for non-parametric multivariate analysis of variance. Austral Ecol.

472   26:32-46.

473   Asher R.J., Smith M.R., Rankin A., Emry R.J. 2019. Congruence, fossils and the evolutionary tree of

474   rodents and lagomorphs. Roy. Soc. Open Sci. 6:190387.

475   Ax P. 1987. The phylogenetic system: the systematization of organisms on the basis of their

476   phylogenesis. New York, John Wiley & Sons.

477   Ballesteros J.A., Sharma P.P. A Critical Appraisal of the Placement of Xiphosura (Chelicerata) with

478   Account of Known Sources of Phylogenetic Error. Syst. Biol. syz011,

479   https://doi.org/10.1093/sysbio/syz011.

480   Bapst D.W., Schreiber H.A., Carlson S.J. 2018. Combined Analysis of Extant Rhynchonellida

481   (Brachiopoda) Using Morphological and Molecular Data. Syst. Biol. 67:32-48.

482   Beck R.M., Baillie C. 2018. Improvements in the fossil record may largely resolve current conflicts

483   between morphological and molecular estimates of mammal phylogeny. Proc. Roy. Soc. B Biol. Sci.

484   285:20181632.

485   Bokma F., Godinot M., Maridet O., Ladevèze S., Costeur L., Solé F., Gheerbrant E., Peigné S., Jacques

486   F., Laurin M. 2015. Testing for Depéret's rule (body size increase) in mammals using combined extinct

487   and extant data. Syst. Biol. 65:98-108.

488   Borm G.F., Fransen J., Lemmens W.A. 2007. A simple sample size formula for analysis of covariance in

489   randomized clinical trials. J. Clin. Epidemiol. 60:1234-1238.

490   Cascini M., Mitchell K.J., Cooper A., Phillips M.J. 2019. Reconstructing the Evolution of Giant Extinct

491   Kangaroos: Comparing the Utility of DNA, Morphology, and Total Evidence. Syst. Biol. 68:520-537.

492   Cloutier R., Ahlberg P.E. 1995. Sarcopterygian interrelationships: how far are we from a phylogenetic

493   consensus? Geobios 28:241-248.

494 Cobbett A., Wilkinson M., Wills M.A. 2007. Fossils impact as hard as living taxa in parsimony analyses

495 of morphology. Syst. Biol. 56:753-766.

496 Coiro M., Chomicki G., Doyle J.A. 2018. Experimental signal dissection and method sensitivity analyses

497 reaffirm the potential of fossils and morphology in the resolution of the relationship of angiosperms and

498 Gnetales. Paleobiology 44:490-510.

499 Denton J.S., Goolsby E.W. 2018. Measuring inferential importance of taxa using taxon influence indices.

500 Ecol. Evol. 8:4484-4494.

501 Donoghue M.J., Doyle J.A., Gauthier J.A., Kluge A.G., Rowe T. 1989. The importance of fossils in

502 phylogeny reconstruction. Annu. Rev. Ecol. Syst. 20:431-460.

503 Dos Reis M., Donoghue P.C., Yang Z. 2016. Bayesian molecular clock dating of species divergences in

504 the genomics era. Nat. Rev. Genet. 17:71.

505 Doyle J.A., Donoghue M.J. 1987. The importance of fossils in elucidating seed plant phylogeny and

506 macroevolution. Rev. Palaeobot. Palynol. 50:63-95.

507 Doyle J.A., Donoghue M.J. 1992. Fossils and seed plant phylogeny reanalyzed. Brittonia 44:89.

508 Edgecombe G.D. 2010. Palaeomorphology: fossils and the inference of cladistic relationships. Acta Zool.

509 91:72-80.

510 Field D.J., Hsiang A.Y. 2018. A North American stem turaco, and the complex biogeographic history of

511 modern birds. BMC Evol. Biol. 18:102.

512 Finarelli J.A., Flynn J.J. 2006. Ancestral state reconstruction of body size in the Caniformia (Carnivora,

513 Mammalia): the effects of incorporating data from the fossil record. Syst. Biol. 55:301-313.

514 Finarelli J.A., Goswami A. 2013. Potential pitfalls of reconstructing deep time evolutionary history with

515 only extant data, a case study using the Canidae (Mammalia, Carnivora). Evolution 67:3678-3685.

516 Forey P., Fortey R. 2001. Fossils in the reconstruction of phylogeny. In: Palaeobiology II. Briggs DEG,

517 Crowther PR. Eds. Blackwell Science, p. 515-519.

518 Gauthier J., Kluge A.G., Rowe T. 1988. Amniote phylogeny and the importance of fossils. Cladistics

519 4:105-209.

520   Gauthier J.A., Kearney M., Maisano J.A., Rieppel O., Behlke A.D. 2012. Assembling the squamate tree

521   of life: perspectives from the phenotype and the fossil record. Bull. Peabody Mus. Nat. Hist. 53:3-308.

522   Giribet G. 2015. Morphology should not be forgotten in the era of genomics–a phylogenetic perspective.

523   Zool. Anz. 256:96-103.

524   Goloboff P.A. 1999. Analyzing large data sets in reasonable times: solutions for composite optima.

525   Cladistics 15:415-428.

526   Goloboff P.A, Catalano S.A. 2016. TNT version 1.5, including a full implementation of phylogenetic

527   morphometrics. Cladistics 32:221-238.

528   Goloboff P.A., Pittman M., Pol D., Xu X. 2019. Morphological Data Sets Fit a Common Mechanism

529   Much More Poorly than DNA Sequences and Call Into Question the Mkv Model. Syst. Biol. 68:494-504.

530   Goloboff P.A., Torres A., Arias J.S. 2018. Weighted parsimony outperforms other methods of

531   phylogenetic inference under models appropriate for morphology. Cladistics 34:407-437.

532   Goodman M. 1989. Emerging alliance of phylogenetic systematics and molecular biology: a new age of

533   exploration. In: The Hierarchy of Life. Fernholm B., Bremer K., Jornvall H. Eds. Amsterdam: Elsevier

534   Science Publishers, p. 43-61.

535   Goswami A., Mannion P.D., Benton M.J. 2016. Radiation and extinction: investigating clade dynamics in

536   deep time. Biol. J. Linn. Soc. 118:6-12.

537   Habel K., Grasman R., Gramacy R.B., Mozharovskyi P., Sterratt D.C. 2019. geometry: Mesh Generation

538   and Surface Tessellation. R package version 0.4.1. https://CRAN.R-project.org/package=geometry.

539   Heath T.A., Huelsenbeck J.P., Stadler T. 2014. The fossilized birth–death process for coherent calibration

540   of divergence-time estimates. Proc. Natl. Acad. Sci USA 111:E2957-E2966.

541   Hennig W. 1966. Phylogenetic Systematics. Urbana: Univ. Illinois Press.

542   Hennig W. 1981. Insect phylogeny. John Wiley & Sons Ltd.

543   Herrera J.P., Dávalos L.M. 2016. Phylogeny and divergence times of lemurs inferred with recent and

544   ancient fossils in the tree. Syst. Biol. 65:772-791.

545   Hillis D.M., Heath T.A., John K.S. 2005. Analysis and visualization of tree space. Syst. Biol. 54:471-482.

546   Huelsenbeck J.P. 1991. When are fossils better than extant taxa in phylogenetic analysis? Syst Biol.

547   40:458-469.

548   Jenner R.A. 2004. Accepting partnership by submission? Morphological phylogenetics in a molecular

549   millennium. Syst. Biol. 53:278-298.

550   Jombart T., Kendall M., Almagro-Garcia J., Colijn C. 2017. treespace: statistical exploration of

551   landscapes of phylogenetic trees. Mol. Ecol. Res. 17:1385-1392.

552   Kearney M., Clark J.M. 2003. Problems due to missing data in phylogenetic analyses including fossils: a

553   critical review. J. Vert. Paleontol. 23:263-274.

554   Kroh A., Smith A.B. 2010. The phylogeny and classification of post-Palaeozoic echinoids. Journal of

555   Syst. Palaeontol. 8:147-212.

556   Laurin M. 2012. Recent progress in paleontological methods for dating the Tree of Life. Front Genet.

557   3:130.

558   Lee M.S, Palci A. 2015. Morphological phylogenetics in the genomic age. Curr. Biol. 25:R922-R929.

559   Legg D.A., Sutton M.D., Edgecombe G.D. 2013. Arthropod fossil data increase congruence of

560   morphological and molecular phylogenies. Nat Comm. 4:2485.

561   Lewis P.O. 2001. A likelihood approach to estimating phylogeny from discrete morphological character

562   data. Syst. Biol. 50:913-925.

563   Liaw A., Wiener M. 2002. Classification and Regression by randomForest. R News, 2:18-22.

564   Liow L.H., Quental T.B., Marshall C.R. 2010. When can decreasing diversification rates be detected with

565   molecular phylogenies and the fossil record? Syst. Biol. 59:646-659.

566   Lloyd G.T. 2016. Estimating morphological diversity and tempo with discrete character-taxon matrices:

567   implementation, challenges, progress, and future directions. Biol. J. Linn. Soc. 118:131-151.

568   Løvtrup S. 1977. The phylogeny of Vertebrata. London: John Wiley & Sons.

569   Mariadassou M., Bar-Hen A., Kishino H. 2012. Taxon influence index: assessing taxon-induced

570   incongruities in phylogenetic inference. Syst. Biol. 61:337-345.

571   Marshall C., Schultze H.-P. 1992. Relative importance of molecular, neontological, and paleontological

572   data in understanding the biology of the vertebrate invasion of land. J. Mol. Evol. 35:93-101.

573   Matzke N.J. 2015. TNTR: R functions to aid analyses in the cladistics program TNT. Available online at

574   PhyloWiki: http://phylo.wikidot.com/tntr.

575   Mayr G. 2006. The contribution of fossils to the reconstruction of the higher-level phylogeny of birds.

576   Species, Phylogeny and Evolution 1:59-64.

577   Middleton K. 2016. jnt: Johnson-Neyman Technique. R package version 0.4.

578   Mitchell J.S. 2015. Extant-only comparative methods fail to recover the disparity preserved in the bird

579   fossil record. Evolution 69:2414-2424.

580   Mitchell J.S., Etienne R.S., Rabosky D.L. 2018. Inferring diversification rate variation from phylogenies

581   with fossils. Syst. Biol. 68:1-18.

582   Miyashita T., Coates M.I., Farrar R., Larson P., Manning P.L., Wogelius R.A., Edwards N.P., Anné J.,

583   Bergmann U., Palmer A.R. 2019. Hagfish from the Cretaceous Tethys Sea and a reconciliation of the

584   morphological–molecular conflict in early vertebrate phylogeny. Proc. Natl. Acad. Sci. USA 116:2146-

585   2151.

586   Mongiardino Koch N., Gauthier J.A. 2018. Noise and biases in genomic data may underlie radically

587   different hypotheses for the position of Iguania within Squamata. PLoS One 13:e0202729.

588   Mounce R.C., Sansom R., Wills M.A. 2016. Sampling diverse characters improves phylogenies:

589   craniodental and postcranial characters of vertebrates often imply different trees. Evolution 70:666-686.

590   Nelson G. 1978. Ontogeny, phylogeny, paleontology, and the biogenetic law. Syst. Zool. 27:324-345.

591   Nixon K.C. 1999. The parsimony ratchet, a new method for rapid parsimony analysis. Cladistics 15:407-

592   414.

593   Novacek M.J. 1992a. Fossils as critical data for phylogeny. In: Extinction and Phylogeny. Novacek M.,

594   Wheeler Q. Eds. Columbia University Press, p. 46-88.

595   Novacek M.J. 1992b. Fossils, topologies, missing data, and the higher level phylogeny of eutherian

596   mammals. Syst. Biol. 41:58-73.

597 O'Leary M.A., Bloch J.I., Flynn J.J., Gaudin T.J., Giallombardo A., Giannini N.P., Goldberg S.L., Kraatz

598 B.P., Luo Z.-X., Meng J. 2013. The placental mammal ancestor and the post–K-Pg radiation of placentals.

599 Science 339:662-667.

600 O'Leary M.A., Kaufman S. 2011. MorphoBank: phylophenomics in the "cloud". Cladistics 27:529-537.

601 O'Reilly J.E., Puttick M.N., Parry L.A., Tanner A.R., Tarver J.E., Fleming J., Pisani D., Donoghue P.C.

602 2016. Bayesian methods outperform parsimony but at the expense of precision in the estimation of

603 phylogeny from discrete morphological data. Biol. Lett. 12:20160081.

604 Oksanen J., Blanchet F.G., Friendly M., Kindt R., Legendre P., McGlinn D., Minchin P.R., O'Hara R.B.,

605 Simpson G.L., Solymos P., Stevens M.H.H., Szoecs E., Wagner H. 2019. vegan: Community Ecology

606 Package. R package version 2.5-4. https://CRAN.R-project.org/package=vegan.

607 Paradis E., Schliep K. 2018. ape 5.0: an environment for modern phylogenetics and evolutionary analyses

608 in R. Bioinformatics, 35:526-528.

609 Parry L.A., Edgecombe G.D., Eibye-Jacobsen D., Vinther J. 2016. The impact of fossil data on annelid

610 phylogeny inferred from discrete morphological characters. Proc. Roy. Soc. B Biol. Sci. 283:20161378.

611 Patterson C. 1977. The contribution of paleontology to teleostean phylogeny. In: Major patterns in

612 vertebrate evolution. Hecht M.K., Goody P.C., Hecht B.M. Eds. New York: Plenum Press, p. 579-643.

613 Patterson C. 1981. Significance of fossils in determining evolutionary relationships. Annu. Rev. Ecol.

614 Syst. 12:195-223.

615 Pattinson D.J., Thompson R.S., Piotrowski A.K., Asher R.J. 2014. Phylogeny, paleontology, and

616 primates: do incomplete fossils bias the tree of life? Syst. Biol. 64:169-186.

617 Penny D., Hendy M. 1985. The use of tree comparison metrics. Syst. Zool. 34:75-82.

618 Philippe H., Poustka A.J., Chiodin M., Hoff K.J., Dessimoz C., Tomiczek B., Schiffer P.H., Müller S.,

619 Domman D., Horn M. 2019. Mitigating anticipated effects of systematic errors supports sister-group

620 relationship between Xenacoelomorpha and Ambulacraria. Curr. Biol. 29:1818-1826.

621   Prevosti F.J., Chemisquy M.A. 2010. The impact of missing data on real morphological phylogenies:

622   influence of the number and distribution of missing entries. Cladistics 26:326-339.

623   Puttick M.N. 2016. Partially incorrect fossil data augment analyses of discrete trait evolution in living

624   species. Biol. Lett. 12:20160392.

625   Puttick M.N., O'Reilly J.E., Pisani D., Donoghue P.C. 2019. Probabilistic methods outperform parsimony

626   in the phylogenetic analysis of data simulated without a probabilistic model. Palaeontology 62:1-17.

627   Puttick M.N., O'Reilly J.E., Tanner A.R., Fleming J.F., Clark J., Holloway L., Lozano-Fernandez J., Parry

628   L.A., Tarver J.E., Pisani D. 2017. Uncertain-tree: discriminating among competing approaches to the

629   phylogenetic analysis of phenotype data. Proc. Roy. Soc. B Biol. Sci. 284:20162290.

630   Quental T.B., Marshall C.R. 2010. Diversity dynamics: molecular phylogenies need the fossil record.

631   Trends Ecol. Evol. 25:434-441.

632   Rabosky D.L. 2010. Extinction rates should not be estimated from molecular phylogenies. Evolution

633   64:1816-1824.

634   Raup D.M. 1992. Extinction: bad genes or bad luck? WW Norton & Company.

635   Robinson D.F, Foulds L.R. 1981. Comparison of phylogenetic trees. Math. Biosci. 53:131-147.

636   Ronquist F., Klopfstein S., Vilhelmsen L., Schulmeister S., Murray D.L., Rasnitsyn A.P. 2012a. A Total-

637   Evidence Approach to Dating with Fossils, Applied to the Early Radiation of the Hymenoptera. Syst.

638   Biol. 61:973-999.

639   Ronquist F., Teslenko M., Van Der Mark P., Ayres D.L., Darling A., Höhna S., Larget B., Liu L.,

640   Suchard M.A., Huelsenbeck J.P. 2012b. MrBayes 3.2: efficient Bayesian phylogenetic inference and

641   model choice across a large model space. Syst. Biol. 61:539-542.

642   Sansom R.S., Choate P.G., Keating J.N., Randle E. 2018. Parsimony, not Bayesian analysis, recovers

643   more stratigraphically congruent phylogenetic trees. Biol. Lett. 14:20180263.

644   Sansom R.S., Wills M.A. 2013. Fossilization causes organisms to appear erroneously primitive by

645   distorting evolutionary trees. Sci. Rep. 3:2545.

646   Sansom R.S., Wills M.A. 2017. Differences between hard and soft phylogenetic data. Proc. Roy. Soc. B

647   Biol. Sci. 284:20172150.

648   Sansom R.S., Wills M.A., Williams T. 2017. Dental data perform relatively poorly in reconstructing

649   mammal phylogenies: morphological partitions evaluated with molecular benchmarks. Syst. Biol. 66:813-

650   822.

651   Schliep K.P. 2011. phangorn: phylogenetic analysis in R. Bioinformatics 27:592.

652   Schnitzler J., Theis C., Polly P.D., Eronen J.T. 2017. Fossils matter–understanding modes and rates of

653   trait evolution in Musteloidea (Carnivora). Evol. Ecol. Res. 18:187-200.

654   Schrago C.G., Aguiar B.O., Mello B. 2018. Comparative evaluation of maximum parsimony and

655   Bayesian phylogenetic reconstruction using empirical morphological data. J. Evol. Biol. 31:1477-1484.

656   Scotland R.W., Olmstead R.G., Bennett J.R. 2003. Phylogeny reconstruction: the role of morphology.

657   Syst. Biol. 52:539-548.

658   Simões T.R., Caldwell M.W., Tałanda M., Bernardi M., Palci A., Vernygora O., Bernardini F., Mancini

659   L., Nydam R.L. 2018. The origin of squamates revealed by a Middle Triassic lizard from the Italian Alps.

660   Nature 557:706.

661   Simpson G.G. 1952. How many species? Evolution 6:342-342.

662   Slater G.J., Harmon L.J. 2013. Unifying fossils and phylogenies for comparative analyses of

663   diversification and trait evolution. Methods Ecol. Evol. 4:699-702.

664   Slater G.J., Harmon L.J., Alfaro M.E. 2012. Integrating fossils with molecular phylogenies improves

665   inference of trait evolution. Evolution 66:3931-3944.

666   Smith A.B. 1994. Systematics and the fossil record: documenting evolutionary patterns. John Wiley &

667   Sons.

668   Smith A.B. 1998. What does palaeontology contribute to systematics in a molecular world? Mol.

669   Phylogenet. Evol. 9:437-447.

670   Smith M.R. 2019. Bayesian and parsimony approaches reconstruct informative trees from simulated

671   morphological datasets. Biol. Lett. 15:20180632.

672   Smith N.D., Turner A.H. 2005. Morphology's role in phylogeny reconstruction: perspectives from

673   paleontology. Syst. Biol. 54:166-173.

674   Spaulding M., O'Leary M.A., Gatesy J. 2009. Relationships of Cetacea (Artiodactyla) among mammals:

675   increased taxon sampling alters interpretations of key fossils and character evolution. PLoS One 4:e7062.

676   Springer M.S., Burk-Herrick A., Meredith R., Eizirik E., Teeling E., O'Brien S.J., Murphy W.J. 2007. The

677   adequacy of morphology for reconstructing the early history of placental mammals. Syst. Biol. 56:673-

678   684.

679   Sumrall C.D. 1997. The role of fossils in the phylogenetic reconstruction of Echinodermata. The

680   Paleontological Society Papers, 3:267-288.

681   R Core Team. 2019. R: A Language and Environment for Statistical Computing. Vienna, Austria, R

682   Foundation for Statistical Computing. https://www.R-project.org/.

683   Vea I.M., Grimaldi D.A. 2015. Diverse new scale insects (Hemiptera: Coccoidea) in amber from the

684   Cretaceous and Eocene with a phylogenetic framework for fossil Coccoidea. Am. Mus. Nov. 2015:1-16.

685   Wiens J.J. 2003a. Incomplete taxa, incomplete characters, and phylogenetic accuracy: is there a missing

686   data problem? J. Vert. Paleontol. 23:297-310.

687   Wiens J.J. 2003b. Missing data, incomplete taxa, and phylogenetic accuracy. Syst. Biol. 52:528-538.

688   Wiens J.J. 2005. Can incomplete taxa rescue phylogenetic analyses from long-branch attraction? Syst.

689   Biol. 54:731-742.

690   Wiens J.J., Kuczynski C.A, Townsend T., Reeder T.W., Mulcahy D.G., Sites Jr J.W. 2010. Combining

691   phylogenomics and fossils in higher-level squamate reptile phylogeny: molecular data change the

692   placement of fossil taxa. Syst. Biol. 59:674-688.

693   Wills M.A., Fortey R.A. 2000. The shape of life: how much is written in stone? BioEssays 22:1142-1152.

694   Wilson M.V. 1992. Importance for phylogeny of single and multiple stem-group fossil species with

695   examples from freshwater fishes. Syst. Biol. 41:462-470.

696   Wood H.M., Matzke N.J., Gillespie R.G., Griswold C.E. 2012. Treating fossils as terminal taxa in

697   divergence time estimation reveals ancient vicariance patterns in the palpimanoid spiders. Syst. Biol.

698   62:264-284.

699   Wright A.M., Hillis D.M. 2014. Bayesian analysis using a simple likelihood model outperforms

700   parsimony for estimation of phylogeny from discrete morphological data. PLoS One 9:e109210.

701   Zhang C., Stadler T., Klopfstein S., Heath T.A., Ronquist F. 2016. Total-evidence dating under the

702   fossilized birth–death process. Systematic biology, 65:228-249.

703   **Table 1.** Morphological matrices employed differ in taxonomic scope, character coding strategies and

704   size.

Dataset	Number of ingroup taxa	Percentage of fossils	Number of characters	Percentage of multistate – ordered characters	Percentage of missing data (extant – fossil)
Coccomorpha (Vea and Grimaldi 2015)	116	37.1	174	50.6 – 0.0	23.8 – 50.2
Echinoidea (Kroh and Smith 2010)	163	55.8	303	34.0 – 5.3	17.5 – 28.0
Lemuriformes (Herrera and Dávalos 2016)	62	32.3	421	69.1 – 0.0	44.0 – 71.1
Mammalia (O'Leary et al. 2013)	83	44.6	4541	21.2– 0.0	39.4 – 65.8
Panarthropoda (Legg et al. 2013)	309	69.2	753	14.2 – 0.0	55.2 – 76.1
Squamata (unpublished, based on (Gauthier et al. 2012, Mongiardino Koch and Gauthier 2018))	198	28.3	767	35.3 – 23.9	32.0 – 62.6

705

706

707   **Figure 1:** Protocol for the assessment of topological impact. **1.** Matrices were subsampled to a fixed

708   number (*n*) of extant taxa plus one outgroup (OG). 25 such initial replicates were built from each dataset.

709   **2.** A given number (*m*) of terminals were added to these initial replicates. These were composed entirely

710   of extant (top, blue), fossil (bottom, orange), or pseudoextinct taxa (not shown). Three iterations of taxon

711   addition were performed. **3.** Matrices were subject to phylogenetic analysis under maximum parsimony

712   (MP) and Bayesian inference (BI), resulting in a set of optimal topologies from which at most 100 were

713   sampled at random (represented here by a single tree). **4.** Topologies obtained after taxon addition were

714   pruned to match the taxon sampling of the initial replicate. **5.** Topologies were compared using Robinson-

715   Foulds (RF) distances. Further rounds of addition of *m* terminals were performed on the matrices

716   resulting from step 2, followed by the same subsequent steps. The values of *n* and *m* were determined for

717   each dataset as explained in the SI File 2. The same color scheme is used throughout.

718

719   **Figure 2:** Comparison of the topological impact of fossil (orange) and extant (blue) terminals. Each dot is

720   the topological distance (average RF distance to nearest neighbor) between trees before and after the

721   addition of groups of either extant or fossil terminals, averaged across iterations. Regression lines are

722   shown for the best-fit linear model. When the lines are parallel, the interaction term was not significant.

723   Regions of significance are shown above each plot (the entire range is marked in the absence of a

724   significant interaction). Silhouettes were taken from PhyloPic (http://phylopic.org/), and will be used

725   throughout.

726

727   **Figure 3**: Comparison of the topological impact of fossil (orange) and pseudoextinct (green) terminals.

728   Analysis of the data is performed as in Figure 2.

729

730   **Figure 4:** Comparison of the topological impact of paleontological (orange) and neontological (blue)

731   data. Analysis of the data is performed as in Figure 2, although the continuous predictor employed is the

732   logarithm of the amount of information added to the inference procedure (averaged across iterations).
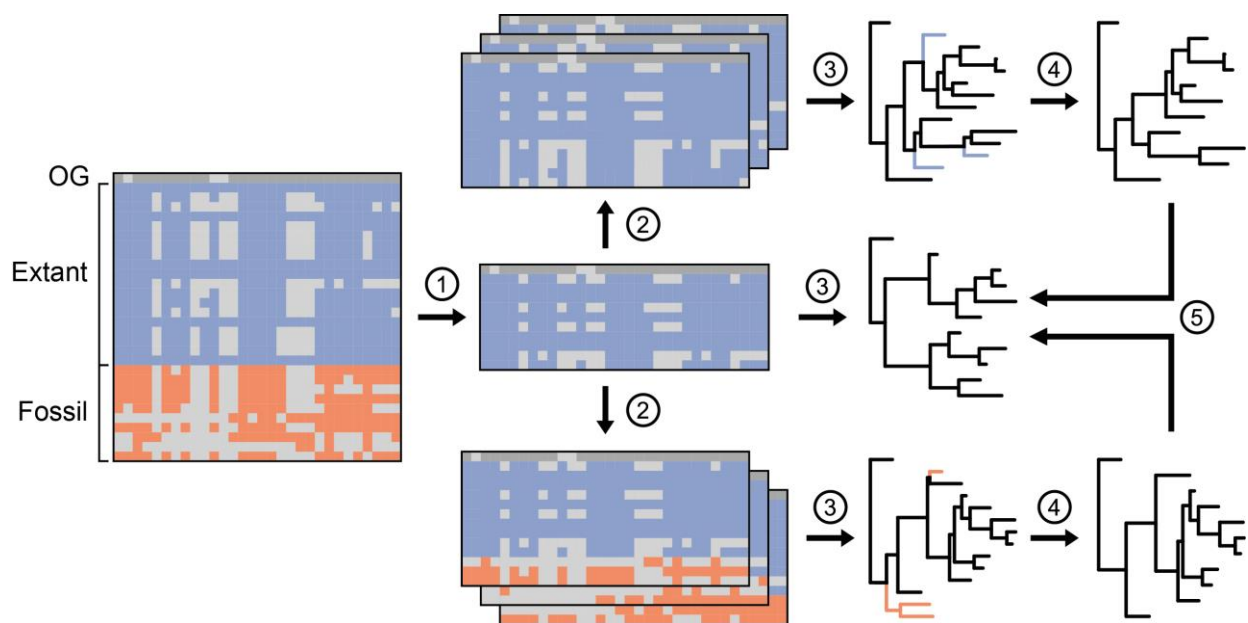
733

734   **Figure 5**: Topological changes induced by the addition of paleontological and neontological data to trees

735   of extant terminals. A single treespace per replicate was generated, summarizing the position of trees

736   obtained after incorporating increasing numbers of extant (blue) and fossil (orange) terminals. Treespaces

737   for each dataset were then rotated to maximize superimposition, and translated so that the topologies

738   product of the initial replicate (that obtained before any terminal is added, see Fig. 1, step 2) is at the

739   origin. Dot size scales with the number of added terminals.

740

741   **Figure 6:** Exploring the determinants of topological impact under MP. **a)** Both missing data, type of taxa

742   (fossil/extant) and their interaction plays a role in determining the topological effect of individual

743   terminals. Across most of the region of missing data for which extant (blue) and fossil (orange) taxa

744   overlap, fossils have a significantly higher topological impact (see also Fig. S7 of SI File 3). **b)** Effect size

745   of the significant determinants of topological impact. Variables are ordered (top to bottom) according to

746   the order in which they are incorporated into the stepwise linear model. Variables with an asterisk were

747   also found to be significant determinants under BI (Fig. S8 of SI File 3). Effect sizes are standardized to

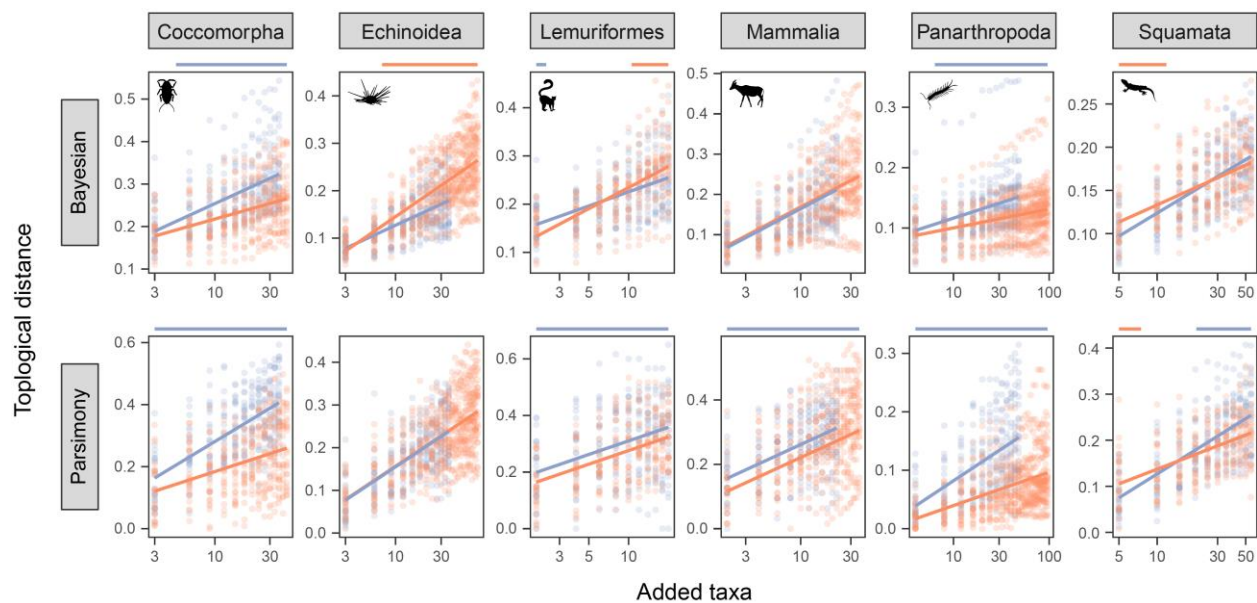748   reflect expected changes generated by a difference of one standard deviation.

749

750

751

752

753

754

755

756
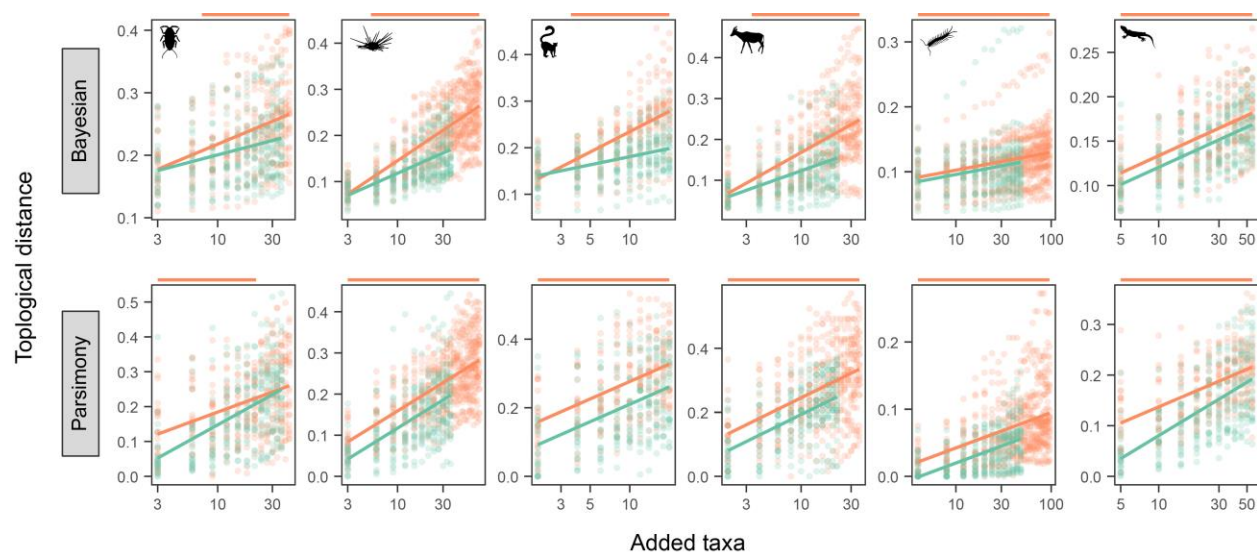
757

758

759

760