1

TITLE: How the Past Impacts the Future: Modelling the Performance of Evolutionarily Distinct

Mammals through Time.

KEYWORDS: living fossil; evolutionary distinctness; evolutionary potential; Mammalia; fossil

record; phylogeny

AUTHORS: D.J. Bennett[1, 2], M.D. Sutton[3] and S.T. Turvey[4]

AFFILIATIONS:

1. Gothenburg Global Biodiversity Centre, Box 461, SE-405 30 Gothenburg, Sweden

2. Department of Biological and Environmental Sciences, University of Gothenburg, Box 461, SE-405 30 Gothenburg, Sweden

3. Department of Earth Sciences and Engineering, Imperial College London, London, SW7 2BP, UK.

4. Institute of Zoology, Zoological Society of London, London NW1 4RY, UK.

EMAILS:

dominic.john.bennett@gmail.com, m.sutton@imperial.ac.uk and samuel.turvey@ioz.ac.uk

ABSTRACT

How does past evolutionary performance impact future evolutionary performance? This is an important question not just for macroevolutionary biologists who wish to chart the phenomena that describe deep-time changes in biodiversity but also for conservation biologists, as evolutionarily distinct species – which may be deemed "low-performing" in our current era – are increasingly the focus of conservation efforts. Contrasting hypotheses exist to account for the history and future of evolutionarily distinct species: on the one hand they may be relics of large radiations, potentially "doomed" to extinction; or they may be slow-evolving, "living fossils", likely neither to speciate nor go extinct; or they may be seeds of future radiations. Here we attempt to test these hypotheses in Mammalia by combining a molecular phylogenetic supertree with fossil record occurrences and measuring change in evolutionary distinctness (ED) at different time slices. With these time slices, we modelled future ED as a function of past ED. We find that past evolutionary performance does indeed have an impact on future evolutionary performance: the most evolutionarily isolated clades tend to become more evolutionarily distinct with time, indicating that low-performing clades tend to remain low-performing throughout their evolutionary history.

200 words max

INTRODUCTION

Evolutionary distinctness (ED; [1]) is a measure of the isolation of a species in a phylogenetic tree, expressed in millions of years. Many conservation biologists are interested in the conservation of species that are highly evolutionarily distinct as these species represent a disproportionately large amount of evolutionary history [1, 2]; they may be likened to "living fossils" [3]. Furthermore, ED has often been correlated (albeit contentiously; [4]) with trait distinctiveness. Consequently, it is often argued (and disputed [5]) that by targeting many evolutionarily distinct species, we would likely be capturing greater trait diversity, and would then be preserving ecosystem functions and services [6]. However, interpretation of the "conservation value" of distinct species is not straightforward, and targeting them for conservation attention may be legitimately questioned if they are deemed more likely to become extinct regardless of human activity [4, 7]. Being members of species-poor lineages, we might consider the evolutionarily distinct as potentially "doomed" to extinction if they are simply the tail-ends of once-diverse clades [8, 9], and/or evolutionary "dead-ends" [10]. Alternatively, the distinctness of such species may garner them with unique adaptations that would allow them to respond differently to future environmental or ecological change [1, 11, 12], allowing them novel opportunities to radiate into recently vacated ecological space [13]. In this scenario, evolutionary distinctness may be related to G.G. Simpson's "phylogenetic fuse": the idea that lineages may persist for long periods of time "tinkering" at low species number before exploding in diversity at a later time point [14]. Yet another interpretation is that evolutionarily distinct species may be members of long-lasting lineages that neither go extinct nor diversify in great numbers ("panchronic" *sensu* [15]). Finally, a neutral model could be proposed in which the origin and future potential of evolutionarily distinct species is not influenced by *any* overarching processes or phenomena.

Little is known about the future evolutionary potential of evolutionarily distinct lineages. Multiple studies have found that lineage age does not correlate with extinction risk [16-18].

61   Additionally, we know that extinction risk today is non-randomly distributed across the tree of life [19-

62   22], although this does not appear to translate into higher extinction risks for evolutionarily distinct

63   species [23-26]. However, these latter studies are mainly based on recent rates and patterns of

64   extinction vulnerability and resilience, and species and clades at risk of extinction can radically vary

65   over longer periods of geological time, such as during mass extinctions [27]. In particular, although the

66   fossil record represents a uniquely important long-term archive that can be used to trace patterns of

67   lineage diversity and diversification through time, few studies have hitherto explicitly attempted to test

68   how past evolutionary performance has impacted future evolutionary performance using the fossil

69   record, particularly from a conservation perspective (see [28] for one such recent example).

70   ED is most often calculated at the species-level [1], but we can also define a clade's ED as the

71   average ED of all descendant species. It is therefore possible to assess the relationship between the ED

72   of a clade at a given time point (ED_{t0}) and its subsequent ED at a later time point (ED_{t1}). We might

73   expect a linear relationship between the ED values of the two time points if ED does not influence

74   diversification (dotted line, figure 1). Such a model, however, does not take into account differences in

75   the ages of trees between different time points and the number of species in a tree. A better null model

76   is able to take into account these differences and may not be linear. In figure 1, we outline a series of

77   possible scenarios for the relationship between past and future ED based on simulated birth-death tree

78   models. The null model is generated from simulations where evolutionary distinctness has no impact

79   on future speciation or extinction rates ("Null", figure 1). If future evolutionary potential of species is

80   influenced by how evolutionarily isolated a clade is, the relationship should diverge from this line. For

81   example, under a strict "panchronic" scenario, clades with low ED are more likely to contain species

82   that will go on to speciate more in the future, so should have lower mean ED at the next time point as

83   they will then share a similar proportion of the tree among more species, whereas clades with high ED

84   are less likely to speciate and should hence show an increase in their mean ED over time ("Pan.", figure

85   1). Alternatively, under a strict "evolutionary relict" scenario, species with mid-ranging ED become

86   increasingly more evolutionarily distinct up to the point of extinction and therefore show ED that is

87   increasingly higher than expected in successive time points, while high ED should tend towards

88   expectation as the whole clade goes extinct ("Rel.", figure 1). Finally, we might consider the less

89   discussed idea that highly evolutionarily distinct clades are more likely to diversify due to their

90   differences from other taxa, "phylogenetic fuse" [14] ("P.F.", figure 1).

91

92   [INSERT FIGURE 1 HERE]

93   FIGURE 1. Polynomial, linear modelling of possible general relationships between ED at time points in the past (t0, x-axis) and in the

94   future (t1, y-axis) as determined from simulated birth-death trees under different scenarios. ED is natural logged to create a normal

95   distribution. Trees were rescaled to make their branch lengths comparable to the Mammalian tree of life. Dark-grey indicates 95%

96   confidence interval. Dotted lines indicate past ED matching future ED and are added for reference. "Null" indicates the model where

97   speciation and extinction rates are independent from evolutionary distinctness. This model is repeated in subsequent panels as the solid,

98   black line. Tendencies above this line indicate higher than expected evolutionary distinctness at the next time step, and below the line,

99   lower than expected evolutionary distinctness at the next time step. In the panchronic scenario (Pan.), mid to high values lead to higher

100   values. In the evolutionary relict scenario (Rel.), mid to high ED leads to increasing ED up to a maximum value at which point the whole

101   clade becomes extinct. In the phylogenetic fuse scenario (P.F.), low to mid ED leads to higher values, and high values lead to low values.

102   These curves were generated from birth-death tree-growth simulations where biases can be introduced based on tip evolutionary

103   distinctness. Scenarios are based on those described in [3]. For more details on how these simulations were created see electronic

104   supplementary material, methods and results.

105

106   Here, we compare ED between mammalian clades at different time points to investigate how

107   evolutionary distinctness changes at the clade level through time. We use mammals due to their well-

108   established phylogenetic tree [29] and the availability of relatively well-sampled fossil data [30]. We

109   generate large molecular-fossil phylogenetic trees by taking a large mammalian supertree and adding

110   fossil tips using a taxonomically constrained stochastic process. We then use these large trees to take

111   time slices at comparable time points, and track the evolutionary distinctnesses of identifiable clades

112   and species across these time points to generate an ED_{t0} and ED_{t1} dataset. With this dataset we then test

113   whether the ED_{t1}~ ED_{t0} relationship is non-linear and whether it corresponds to any previously

114   identified scenarios of evolutionary diversification in relation to ED as outlined above. Our results

115   establish a robust new baseline for understanding the likely potential of evolutionarily distinct species

116   today in terms of their future contribution to global biodiversity and its implications on their

117   conservation.

METHODS AND MATERIALS

118

119 All of our analyses were performed in the R environment (v3.5) and the scripts for reproducing these

120 results can be found on the primary author's GitHub page (https://github.com/DomBennett/Project-

121 karenina).

122   Mammalian fossil data were downloaded from the Paleobiological Database (PBDB) [31] using

123 the R package paleobioDB [32] on 25/09/2018. All records were constrained to 'Mammalia'. In total

124 109,536 fossil occurrence records were downloaded. These records were converted to species temporal

125 records by merging records with shared species names and identifying minimum and maximum species

126 ages. Taxonomic named ranks were determined for every species record, and records lacking

127 taxonomic information below the family level were removed (438 records). Records with temporal

128 ranges extending beyond the age of the mammalian supertree (166.2 MY) were also excluded (121

129 records). In total 19,028 species records remained, representing 128 families and 6,659 genera. The

130 temporal distribution was skewed towards the recent; most observations occurred in the first 18 MY

131 (0.0059 MY – 0%, 7.0 MY – 25%, 17.7 MY – 50%, 39.7 MY – 75%, 164.8 MY – 100%).

132

133 *Phylogenetic Placement of Fossils*

134 Fossil species were added to a time-calibrated, molecular-based phylogenetic supertree of 4,510 extant

135 mammal species [29]. Taxonomic information was added to the tree by identifying the most likely

136 taxonomic group for every node, by matching descendants to named entries in the NCBI taxonomic

137 database [33] via the Global Names Resolver [34]. For every node, the lowest shared taxonomic group

138 was selected from all taxonomic named ranks of matching named entries. In the NCBI taxonomy under

139 mammals there is up to 12 taxonomic ranks (from superorder to subgenus) for each species. Taxonomic

140 groups were identified for 6,603 of the 6,618 nodes in the tree (>99%).

141   Tips representing fossil species were added to the supertree using the pinTips() function in the

142   R package treeman [35]. The position for each fossil species was constrained to edges directly parent to

143   or descending from the node(s) with the lowest-ranked matching taxonomic label (with any edge or

144   part of an edge younger than the minimum age of the fossil species excluded). Where these constraints

145   left multiple possible attachment points, one was chosen at random. For example, if a fossil is matched

146   to a given genus name, then the fossil tip could be added to the edge leading to this genus or to any of

147   the edges that make up the genus. The time of extinction for fossil species was determined as a random

148   point between its split from the tree and its youngest possible age (electronic supplementary material,

149   figures S1-S2). Fossil species added to the tree thus all became extinct within their estimated age range,

150   and their origination occurred before or during their estimated age range.

151   The stochastic fossil-adding process was iterated to generate a distribution of possible

152   molecular-fossil trees; 100 iterations were deemed sufficient as initial analyses demonstrated that trees

153   generated through this process had similar distributions of ED values. In total a mean $18,964 \pm 16$

154   fossil tips were added to the original 4,510 tipped mammalian supertree for each iteration. To assess

155   whether the stochastic fossil-adding process was biasing results, the process was repeated with fossil

156   lineages and age ranges randomly assorted, so that fossil placement was not constrained by their actual

157   ages or taxonomy. Before analysing changes in ED between time points, we first tested whether the

158   random set of trees differed significantly from the non-random set (electronic supplementary material,

159   methods and results). If the random set were to not differ, then this would indicate that the fossil-record

160   informed pinning is no different from random placement.

161

162   *Determining Change in ED*

163   For each tree generated through fossil-addition, a 'time slice' was taken to generate an ultrametric tree

164   representing a phylogeny up to a given time point (electronic supplementary material, figure S2). For

165 each of these slices, EDs were calculated for all tips using the method of [1], with ED of internal nodes

166 calculated as the mean ED of all descendants of that node. In order to generate ultrametric trees that

167 were as equally sampled as possible, time slices were taken at the midpoints of each epoch (Upper

168 Jurassic to Recent; electronic supplementary material, table S1). ED_{t0} and ED_{t1} datasets were generated

169 from the two distributions by matching ED values from one epoch to the next for all species/clades that

170 spanned two or more epochs. Mean ED values were calculated for species/clades that occurred more

171 than once across the tree distributions (later termed *shared nodes*). All values were converted to their

172 natural logarithm before analysis, to convert the skewed distribution of ED to a normal distribution.

173 Our datasets of species/clades were supplemented with taxonomic information (genus and order), the

174 number of species present in the tree at *t0*, and the time (in MY) between *t0* and *t1*.

175

176 *Modelling*

177 **Mixed-Effects Models**

178 ED values of epoch-to-epoch transitions are not straightforward to compare: the ED values in one

179 epoch will not be independent of those from another because many of the taxa are the same; transitions

180 differ in tree age and starting number of tips at *t0;* the number of fossil records differ between each

181 epoch (older have fewer than recent); and transition time-steps are not the same (electronic

182 supplementary, table S1). To account for these factors and control for differences between epochs, we

183 used linear mixed-effects models (LMEMs) [36, 37]. We modelled future ED as a function of past ED

184 ($ED_{t1} \sim ED_{t0}$) with the option of including a random-effects structure that was able to incorporate both

185 epoch and taxonomic information. We used an iterative model approach, starting with a basic LMEM

186 model and adding additional terms to the random effect structures until there was no longer any

187 significant or notable gain in explained variance. We compared models using ANOVA and the Akaike

188   Information Criterion (AIC) [38]. For more details on LMEMs and our approach see electronic

189   supplementary information, methods and results.

190

191

192

193   **The Three Models**

194   Using the LMEM framework, we generated two models to investigate how past ED affects future ED:

195   a linear observed model (*obs1*), a non-linear observed model (*obs2*), and an expected model (*exp*). We

196   generated *obs1* using the modelling framework as described above to generate the best-fitting linear

197   model of $ED_{t1}\sim ED_{t0}$ given different random effects structures. We then took the same random effect

198   structure to generate *obs2* by adding increasing numbers of orthogonal polynomial degrees to explore

199   increasingly complex non-linear relationships, until there was no longer any increase in explained

200   variance.

201   An expected model, akin to the "Null" birth-death model in figure 1., was built for comparison,

202   to represent $ED_{t1}\sim ED_{t0}$ in a 'neutral' scenario where past average ED for a clade has no impact on

203   future ED other than simply determining the starting points from which future ED can diverge. This

204   model was a function of ED_{t1} and the factors that influence ED between epochs: the number of tips in

205   the tree at *t0* (*n*), the time difference between epochs (*tm*), and an ED_{t0} dummy variable (*dummy*). The

206   dummy variable was a coarsened representation of ED_{t0} calculated by rounding ED_{t0} to the nearest

207   integer and dividing the values by the maximum to generate values between 0 and 1. The variable

208   provided limited categorical information on ED_{t0} to the model (e.g. high ED, mid ED and low ED) and

209   acted as a basis from which ED_{t1} can then be estimated in conjunction with the other factors.

210   With these models we first tested whether $ED_{t1}\sim ED_{t0}$ is non-linear by comparing the goodness

211   of fit to the data of *obs1* and *obs2*, and then tested which of the simulated scenarios (figure 1) best

212   describe $ED_{t1}\sim ED_{t0}$ by visually comparing the best observed (*obs1* or *obs2*) to *exp*.

RESULTS

213

Estimating the Linear Model

214

215   The generated dataset consisted of 115,810 species/clade ED values, recorded across the nine epoch-to-

216   epoch transitions. The distribution of shared nodes across the iterated trees was bimodal, with the

217   majority of species/clades occurring 12 or fewer times (1 - 0%, 3 - 25%, 12 - 50%, 39 - 75%, 100 -

218   100%). There was a positive non-linear relationship between ED_{t0} and ED_{t1} and differences between

219   the epoch-to-epoch transitions were substantial, particularly at low ED values. Two epoch-to-epoch

220   transitions (Jurassic Upper to Cretaceous Lower [JU-CL], Cretaceous Lower to Cretaceous Upper [CL-

221   CU]) showed a different relationship from the others (electronic supplementary material, figure S3),

222   probably due to limited availability of data points for these epochs (6,321 and 2,288 for CL-CU and

223   JU-CL, respectively, versus a mean of 15,314 for all other epoch transitions) and the much longer time

224   separating them (39.5 MY and 31.5 MY for CL-CU and JU-CL, respectively, versus a mean of 11.9

225   MY for all other transitions). We consequently removed JU-CL and CL-CU transitions from all

226   subsequent analyses.

227   We determined the best linear model of ED_{t1}~ED_{t0} to be *m2i* (table 1). This model incorporated

228   a random effects structure that consisted of independent slopes for both epochs and a hierarchical

229   taxonomic random effect (order/genus). However, we did not consider model choice to be crucial in

230   interpretation of results, as estimated slopes were similar across all models (0.63-0.73).

231

232 TABLE 1. Increasingly complex models for estimating the best observed linear model, *obs1*. Formulae (*f*) [37], intercepts (Int), slopes

233 (Slp), degrees of freedom (DF), Akaike Information Criterion (AIC), and significance (P) are indicated. Significance is indicated by

234 comparing the current row's model to the last significant model (*** p < 0.001, ** p < 0.01, * p < 0.05). The selected linear model (*m2i*)

235 is highlighted in bold.

Id.	f	Int	Slp	DF	AIC	P
m0	$ED_{t1} \sim ED_{t0}$	1.1068	0.7342	1	49,418	
m1a	$ED_{t1} \sim ED_{t0} + (1|epoch)$	1.2101	0.7148	4	21,337	
m1b	$ED_{t1} \sim ED_{t0} + (ED_{t0}|epoch)$	1.2274	0.7061	6	17492	***
m2a	$ED_{t1} \sim ED_{t0} + (ED_{t0}|epoch) + (1|order)$	1.2219	0.7100	7	16568	***
m2b	$ED_{t1} \sim ED_{t0} + (ED_{t0}|epoch) + (1|genus)$	1.3483	0.6642	7	10,459	***
m2c	$ED_{t1} \sim ED_{t0} + (ED_{t0}|epoch) + (1|order/genus)$	1.3562	0.6638	8	10,296	***
m2d	$ED_{t1} \sim ED_{t0} + (ED_{t0}|epoch) + (1|id)$	1.2646	0.6920	7	16,372	
m2e	$ED_{t1} \sim ED_{t0} + (ED_{t0}|epoch) + (1|order/id)$	1.2626	0.6951	8	15,557	
m2f	$ED_{t1} \sim ED_{t0} + (ED_{t0}|epoch) + (ED_{t0}|order)$	1.2111	0.7144	9	16,378	
m2g	$ED_{t1} \sim ED_{t0} + (ED_{t0}|epoch) + (ED_{t0}|genus)$	1.4242	0.6274	9	7,255	***
m2h	$ED_{t1} \sim ED_{t0} + (ED_{t0}|epoch) + (ED_{t0}|id)$	1.2726	0.6882	9	16,379	
m2i	$\mathbf{ED_{t1} \sim ED_{t0} + (ED_{t0}|epoch) + (ED_{t0}|order/genus)}$	1.4010	0.6405	12	7,106	***

236

237

238   *Estimating the Non-Linear Model*

239   We compared the best linear model (*m2i*) with a range of non-linear models, using the same random

240   effects structure, based on orthogonal polynomials generated from ED_{t0} for different exponents. Using

241   ANOVA and AIC values, as expected, we found all the polynomial models to have significantly better

242   fits than the linear model, indicating that evolutionarily distinct clades do have differences in

243   diversification potential (table 2). Of the non-linear models, we determined *m3b*, the 3rd order

244   polynomial model, to be the best; this model had one of the lowest AICs and a low $\chi 2$ P-value

245   indicating goodness-of-fit. While there were significant improvements in AIC in higher orders of

246   polynomial model, the drops in AIC were small ($< 1\%$), and were not considered sufficient to warrant

247   the greater model complexity. Additionally, we repeated the final plotting analysis (figure 2) and found

248   similar trends for all higher-order polynomial (> 3) models, indicating that any conclusions are not

249   affected by our choice of the lower order polynomial model.

250

251

252 TABLE 2. Increasingly complex models for estimating the best observed non-linear model, *obs2*. Order of polynomial (Poly), degrees of

253 freedom (DF), Akaike Information Criterion (AIC), and significance (P) are indicated. All models share the same model formula as *m2i*.

254 Significance is indicated by comparing the current row's model to the last significant model (*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$). The

255 selected non-linear model (*m3b*) is highlighted in bold.

Id.	Poly	DF	AIC	P
m2i	1	12	7,106	
m3a	2	13	4,727	***
m3b	**3**	**14**	**4,513**	***
m3c	4	15	4,501	***
m3d	5	16	4,493	**
m3e	6	17	4,473	***
m3f	7	15	4,674	

256
257

258 *Estimating the Expected Model*

259 We found the best-expected model to be *n2b*, which incorporates the ED_{t0} dummy variable (*dummy*)

260 with three orders of polynomial, difference in time between epochs (*tm*) and the starting number of

261 species in a Mammalian tree at the beginning of an epoch (*n*), as well as a genus random effects

262 structure of random slopes for *dummy* (table 3). Although it is likely that a model with random slopes

263 for all three fixed effects would have produced a better fit, this model and models with multiple random

264 effects did not converge indicating model over-fitting. Again, we found only marginal improvement

265 with the use of a hierarchical taxonomic random effects structure (order/genus). Consequently, we

266 limited our random-effects structure to one non-hierarchal factor.

267

268 TABLE 3. Expected models, *exp*, of ED_{t1} against difference of time between epochs (*tm*), number of species in the tree at *t0* (*n*), and an

269 ED_{t0} dummy variable generated from the rounded figures of ED_{t0}. Formulae (*f*) [37], degrees of freedom (DF), Akaike Information

270 Criterion (AIC), and significance (P) are indicated. Significance is indicated by comparing the current row's model to the last significant

271 model (*** p < 0.001, ** p < 0.01, * p < 0.05). The '-' indicates the separation of linear models and linear mixed-effects models. The

272 selected expected model (*n2b*) is highlighted in bold.

Id.	f	DF	AIC	P
n0a	ED_{t1}~tm	3	118,248	
n0b	ED_{t1}~n	3	114,116	
n0c	ED_{t1}~tm+n	4	113,934	***
n0d	ED_{t1}~dummy	3	71,796	
n0e	ED_{t1}~dummy+tm+n	5	49,229	***
n1a	ED_{t1}~dummy+tm+n+(1|order)	6	48,182	-
n1b	ED_{t1}~dummy+tm+n+(1|genus)	6	35,783	***
n1c	ED_{t1}~dummy+tm+n+(1|id)	6	47,644	
n1d	ED_{t1}~dummy+tm+n+(dummy|genus)	8	31,649	***
n1e	ED_{t1}~dummy+tm+n+(tm|genus)	8	32,141	
n1f	ED_{t1}~dummy+tm+n+(n|genus)	8	31,262	
n2a	ED_{t1}~poly(dummy, 2)+tm+n+(dummy|genus)	9	30,721	***
n2b	**ED_{t1}~poly(dummy, 3)+tm+n+(dummy|genus)**	**10**	**30,496**	*******
n2c	ED_{t1}~poly(dummy, 4)+tm+n+(dummy|genus)	11	30,472	***

273

274

275

276   *Comparing the Expected and the Observed*

277   We compared the expected model (*exp*, *n2b*) with the best observed polynomial model (*obs2*, *m3b*) by

278   plotting the parameterised lines to create figures analogous to figure 1. Because these models generate

279   tens of thousands of individual lines (one line per genus per epoch transition), plotting all outputs

280   directly would make visual interpretation of trends difficult, so we constructed a representative dataset

281   consisting of 100 equally spaced ED_{t0} values spanning from the observed minimum to observed

282   maximum for 100 random selections of all genera across all epochs. This representative dataset

283   consisted of 70,000 rows. We used the dataset to predict ED_{t1} for *obs2* and plotted the predicted values

284   by calculating the median and 95% CIs across genera and epochs (figure 2).

285   For both the best non-linear model when compared to the expected model, we found that, on

286   average, high averaged clade ED_{t0} values (above ~ $e^{2.5}$ or ~12 MY) tended to lead to higher than

287   expected ED_{t1} values, a pattern that was most consistent with the "panchronic" scenario; showing

288   increasing divergence from the expected model at higher ED_{t0} values. For low values of averaged clade

289   ED_{t0} (below ~ $e^{1.5}$ or ~5 MY) we found that ED_{t1} was lower than expected, indicating clades with low

290   ED tended to speciate more than clades with high ED. These patterns were consistent across the

291   different epochs (electronic supplementary material, figure S6). Additionally, upon repeating the above

292   analyses for non-taxonomically informed stochastic fossil placement in the molecular-fossil trees, we

293   were able to rule out that our results were the product of random fossil placement (electronic

294   supplementary material, methods and results).

295

296   [INSERT FIGURE 2 HERE]

297   FIGURE 2. Predicted *log(ED$_{t1}$)* values generated from the best-observed linear (*obs1*, solid red line) and expected model (i.e. "dummy

298   variable", n2b, solid black line) for a representative dataset of a range of *log(ED$_{t0}$)* values, a random subset of one hundred genera, and all

299   epoch-to-epoch transitions. When compared to figure 1, the observed relationship most closely resembles the "panchronic" scenario.

300   Dotted line indicates *log(ED$_{t0}$)* = *log(ED$_{t1}$)*. Estimates across the different genera and epoch-to-epoch transitions are median averaged and

301   95% CIs were calculated, with values representing natural logged millions of years (1 = 2.7, 2 = 7.4, 3 = 20.1, 4 = 54.6, 5 = 148.4).

302

DISCUSSION

In order to determine the future evolutionary potential of clades composed of evolutionarily distinct species, we modelled future evolutionary distinctness as a function of past evolutionary distinctness using data generated from the mammalian fossil record. We find that taxa that are or have been evolutionarily distinct are more likely to become more evolutionarily distinct in the future. These findings fit closest to a "panchronic" scenario for the evolution of the evolutionarily distinct; evolutionarily distinct taxa in this scenario are disproportionately composed of "living fossils" that experience both reduced rates of extinction and speciation [15]. Observed results differ from what would be expected from an "evolutionarily relict" ("dead-end") scenario, in which species most commonly become evolutionarily distinct by being the sole survivors of a once-large clade (figures 1-2). The observed model shows a slow gain in ED for clades that are *already* very evolutionarily distinct; however, in an "evolutionary relict" scenario, clades with mid to low ED – below e^2 or 7 MY – at *t0* should tend to reach high-ED at *t1*. Instead, these clades have lower ED at *t1*. Equally, our results are not compatible with a "phylogenetic fuse" scenario in which evolutionarily distinct clades tend to be the seeds of future diversity, because taxa with high ED were not observed to lead to taxa with low ED over time.

Potential Biases

Our results indicate that the randomness of the fossil-adding process is likely to have *reduced* (not generated) the non-linearity of the observed results. Firstly, upon comparing the models fitted to the taxonomically informed versus randomly added distribution of molecular-fossil trees, we showed that the pattern for high ED values in *t0* leading to even higher ED values in *t1* was stronger in the taxonomically informed distribution. Secondly, the observed non-linearity was greatest for the data points in which we have the greatest confidence, as high ED values in *t0* led to even higher ED values

328   in *t1* for the species/clades that were shared more than fifty times across tree iterations (electronic

329   supplementary material, "Comparing Model Outcomes between the Real and Random").

330   The level of detail of the fossil taxonomy limits the placement of fossils (electronic

331   supplementary material, figure S1). This issue is more likely to affect the ED estimates of lower-level

332   clades (e.g. at the genus or subfamily level) than higher-level clades, because although fossil species

333   may be added to the wrong subfamily due to lack of taxonomic resolution, they are very unlikely to be

334   added to the wrong family. Nonetheless, the taxonomic misplacement of fossil species will impact

335   estimates of ED. Another source of error in the placement of fossil species is their age of speciation and

336   extinction. Ages of appearance are likely to be underestimated due to the Signor-Lipps effect [39].

337   Although this effect is not directly accounted for using any statistical or mechanistic models, the

338   effect's impact should be mitigated by our fossil-adding process, in which origination of fossil lineages

339   was set to occur before or during the estimated age ranges. Both of these sources of error could

340   theoretically introduce bias that might account for some or all of the non-linearity observed in our

341   model outputs, but it is difficult to conceive of a mechanism by which taxonomic or dating errors could

342   produce the strong observed patterns. Additionally, we note that phylogenetic trees that are not wholly

343   accurate can still produce metrics more similar to true values than would be expected [40].

344   Taphonomic (preservational) inconsistencies in the fossil record are potentially another source

345   of bias [41]. To produce the observed non-linearity in our model outputs, a bias where species with

346   high ED were sampled less frequently than species with low ED would be required. One mechanism

347   that could achieve a fossil bias of this kind might be differential sampling of evolutionarily distinct

348   species if they have smaller body sizes and/or population sizes [42]. However, no studies have found

349   significant correlations between evolutionary distinctness and population size, and many studies

350   suggest that evolutionarily distinct species tend to have *greater* body mass for a range of taxonomic

351   groups [43-46], which may be associated with increased likelihood of representation in the fossil

352   record, e.g. [47]. In mammals body mass is only weakly or not at all correlated with ED [12].

353   Additionally, taphonomic biases are likely to be strongly taxonomically controlled, for example

354   because related species are likely to occupy similar habitats [48]. However, the non-linear relationship

355   in our model outputs is detected even when using a random effects structure that controls for inter-

356   relatedness. We therefore consider it very unlikely that this potential scenario for bias is responsible for

357   generating the patterns that we observe.

358

359   **Implications for Conservation Biology**

360   Palaeontology and macroevolution have an important potential role to play in informing conservation,

361   and our study demonstrates how fossil data can be used to provide unique insights for conservation

362   prioritisation and decision-making. Funds are limited and not all species can be saved; prioritisation is

363   therefore key to conservation strategy [49, 1, 12]. To date, most species-level conservation

364   prioritisation has been focused on well-known charismatic taxa such as tigers, rhinos and polar bears

365   [50, 51], as well as on highly threatened species in urgent need of conservation [52]. One approach to

366   move away from this bias has been to place greater conservation effort on evolutionarily distinct

367   species (e.g. [53-55]), based in part upon the supposition that such taxa might contribute to future

368   evolutionary potential; that is, they represent "cradles" of future diversity, as opposed to "museums" of

369   past diversity [56]. One such widely publicised effort has been to rank species for conservation

370   prioritisation based on their "EDGE" (Evolutionarily Distinct and Globally Endangered) score [1, 12].

371   However, arguments arise over the usefulness of investing efforts into species that may already be

372   evolutionarily "doomed" [7, 4]. Indeed a recent study found that conservation strategies that prioritise

373   the preservation of evolutionarily distinct species led to a reduction in the number of future lineages

374   [28].

375   The findings of our investigation of large-scale fossil data provide an important new perspective

376   on this conservation question. It is clear that, at a broad global scale and based on data for mammals,

377   focusing conservation attention on evolutionarily distinct species will not safeguard the overall future

378 "evolvability" of life on Earth, as implied by a "phylogenetic fuse" scenario [14]. However, our results

379 are equally not indicative of an evolutionary "dead-end" scenario for the origin and future of

380 evolutionarily distinct species over geological time [8, 9]. These taxa are neither doomed nor the likely

381 seeds of future radiations; our findings demonstrate that they are likely to remain evolutionarily

382 distinct, and appear to be merely the slow evolving ends of the tree of life. Although the exact

383 mechanism for what may cause the increasing distinctness of the distinct is not tested here, the results

384 corroborate with studies demonstrating the widespread phenomenon of age-dependant speciation [57-

385 59].

386   Our findings may appear to argue against special prioritisation of evolutionarily distinct species,

387 contrary to the philosophy of the "EDGE" approach. However, we definitely do not advocate

388 withdrawing conservation attention from these species. While our study provides a new overview of

389 the likely persistence and future diversification of evolutionarily distinct species and clades, thus

390 challenging one of the general justifications that has been proposed for conserving such species, these

391 findings provide a broad pattern across the Mammalia as a whole; we are unable to predict whether any

392 specific species or clade within this major animal group may indeed diversify in the future, to what

393 extent, or at what time, in response to any number of possible future ecological scenarios. Indeed, we

394 note that whereas our analyses are based upon data for speciation and extinction events across deep

395 time caused by non-anthropogenic processes, current-day global biodiversity loss is driven by

396 anthropogenic processes and is likely to be associated with different patterns of species vulnerability or

397 resilience [60, 61], including the generation of "artificially evolutionarily distinct" or "neo-relict"

398 clades which were species-rich before the Late Quaternary but have experienced disproportionate

399 levels of human-caused extinction (e.g., proboscideans, sloths) [22].

400   Furthermore, future potential is only one criterion for targeting conservation effort, both

401 quantifiable and intangible factors should also be considered. Conserving evolutionary history and the

402 breadth of mammalian biodiversity is an important goal in and of itself, for a variety of biological

403   reasons (e.g. trait diversity and its relationship with ecosystem function and complexity; [46]) and other

404   reasons (e.g. ethical, financial, cultural), irrespective of the possible future evolutionary trajectory of a

405   given species or clade. Evolutionarily distinct lineages have just as much of a right to survive the

406   Anthropocene as do any others, and their unique representation of more independent evolutionary

407   history than 'normal' lineages may still be considered worthy of particular attention.

408

409

ACKNOWLEDGEMENTS

FUNDING

417

418 # REFERENCES

419
420 1. Isaac, N. J. B., Turvey, S.T., Collen, B., Waterman, C. and Baillie, J.E.M. 2007. Mammals on the EDGE:
421   conservation priorities based on threat and phylogeny. PloS One, 2(3):e296.

422 2. Rosauer, D.F. and Mooers, A.Ø. 2013. Nurturing the use of evolutionary diversity in nature conservation. Trends
423   in Ecology and Evolution, 28(6):322–3.

424 3. Bennett, D.J., Sutton, M.D., and Turvey, S.T. 2018. Quantifying the living fossil concept. Palaeontologia
425   Electronica, 21.1.14A:25

426 4. Winter, M., Devictor, V. and Schweiger, O. 2013. Phylogenetic diversity and nature conservation: where are we?
427   Trends in Ecology and Evolution, 28(4):199–204

428 5. Mazel, F., Pennell, M.W., Cadotte, M.W., Diaz, S., Dalla Riva, G.V., Grenyer, R., Leprieur, F., Mooers, A.Ø.,
429   Mouillot, D., Tucker, C.M. and Pearse, W.D. 2018. Prioritizing phylogenetic diversity captures functional diversity
430   unreliably. *Nature Communications*, 9, 2888

431 6. Millennium Ecosystem Assessment. 2005. Ecosystems and Human Well-being: Biodiversity Synthesis.

432 7. Krajewski, C. 1991. Phylogeny and diversity. Science, 254(5034):918–919.

433 8. Nagalingum, N. S., Marshall, C.R., Quental, T.B., Rai, H.S., Little, D.P. and Mathews, S. 2011. Recent
434   synchronous radiation of a living fossil. Science, 334:796–799.

435 9. Mouquet, N. Devictor, V., Meynard, C.N., Munoz, F., Bersier, L-F. and Chave, J. 2012. Ecophylogenetics:
436   advances and perspectives. Biological reviews of the Cambridge Philosophical Society, 87(4):769–85.

437 10. Parsons, P. 1994. Habitats, stress, and evolutionary rates. Journal of Evolutionary Biology, 397:387–397.

438 11. Redding, D. W., DeWolff, C. V. and Mooers, A. Ø. 2010 Evolutionary distinctiveness, threat status and
439   ecological oddity in primates. Conserv. Biol. 24, 1052–1058.

440 12. Collen, B., Turvey, S.T., Waterman, C., Meredith, H.M.R., Kuhn, T.S., Baillie, J.E.M. and Isaac, N.J.B. 2011.
441   Investing in evolutionary history: implementing a phylogenetic approach for mammal conservation. Proceedings of
442   the Royal Society of London B - Biological Sciences, 366(1578):2611–22.

443 13. Schluter, D. 2000. The Ecology of Adaptive Radiation. Oxford University Press, Oxford.

444 14. Cooper, A. and Fortey, R. 1998. Evolutionary explosions and the phylogenetic fuse. Trends in Ecology and
445   Evolution, 13(4):151–156.

15. Bennett, D.J., Sutton, M.D. and Turvey, S.T. 2017. Evolutionarily distinct "living fossils" require both lower speciation and lower extinction rates. Paleobiology, 43(1):34–48.

16. Van Valen, L. 1973. A new evolutionary law. Evolutionary Theory, 1:1–30.

17. Alroy, J. 1996. Constant extinction, constrained diversification, and uncoordinated stasis in North American mammals. Palaeogeography, Palaeoclimatology and Palaeoecology,127:285– 311.

18. Ezard, T.H.G., Aze, T., Pearson, P.N., and Purvis, A. 2011. Interplay between changing climate and species' ecology drives macroevolutionary dynamics. Science 332:349–51

19. Nee, S. and May, R. 1997. Extinction and the Loss of Evolutionary History. Science, 278(5338):692–694.

20. von Euler, F. 2001. Selective extinction and rapid loss of evolutionary history in the bird fauna. Proceedings of the Royal Society of London B - Biological Sciences, 268(1463):127–130.

21. Purvis, A. Agapow, P-M. Gittlemann, J.L. and Mace, G.M. 2000. Nonrandom Extinction and the Loss of Evolutionary History. Science, 288(5464):328–330.

22. Mooers AØ, Goring SJ, Turvey ST, Kuhn TS. 2009. Holocene extinctions and the loss of feature diversity. In Turvey ST (Ed.) Holocene Extinctions: 263-277. Oxford University Press, Oxford.

23. Verde Arregoitia, L.D., Blomberg, S.P. and Fisher, D.O. 2013. Phylogenetic correlates of extinction risk in mammals: species in older lineages are not at greater risk. Proceedings of the Royal Society of London B - Biological Sciences, 280(1765):20131092.

24. Jetz, W., Thomas, G.H., Joy, J.B., Redding, D.W., Hartmann, K. and Mooers, A.Ø. 2014. Global Distribution and Conservation of Evolutionary Distinctness in Birds. *Current Biology*, 24(9):919-930.

25. Tonini, J.F.R., Beard, K.H., Ferreira, R.B., Jetz, W. and Pyron, R.A. 2016. Fully-sampled phylogenies of squamates reveal evolutionary patterns in threat status. *Biological Conservation*, 204(A):23-31.

26. Stein, R.W., Mull, C.G., Kuhn, T.S., Aschliman, N.C., Davidson, L.N.K., Joy, J.B., Smith, G.J., Dulvy, N.K. and Mooers, A.Ø. 2018. Global priorities for conserving the evolutionary history of sharks, rays and chimaeras. *Nature Ecology & Evolution*, 2:288-298.

27. Jablonski, D. 1995. Extinctions in the fossil record. In Extinction Rates. 25–44.

28. Cantalapiedra, J.L., Aze T., Cadotte M.W., Dalla Riva G.V., Huang D., Mazel F., Pennell M.W., Ríos M. and Mooers A.Ø. 2019. Conserving evolutionary history does not result in greater diversity over geological time scales. Proc. R. Soc. B, 286(1904).

29. Bininda-Emonds, O. R. P., Cardillo, M., Jones, K. E., MacPhee, R. D. E., Beck, R. M. D., Grenyer, R., … Purvis, A. 2007. The delayed rise of present-day mammals. Nature, 446(7135), 507-12.

476   30. Liow, L.H. Fortelius, M., Bingham, E., Lintulaakso, K., Mannila, H., Flynn, L. and Stenseth, N.C. 2008. Higher

477   origination and extinction rates in larger mammals. Proceedings of the National Academy of Sciences of the

478   United States of America, 105(16):6097–6102.

479   31. The Paleobiology Database 2018. The Paleobiology Database. [https://paleobiodb.org] (25/09/2018)

480   32. Varela, S., González-Hernández, J., Sgarbi, L.F., Marshall, C., Uhen, M.D., Peters, S. and McClennen, M. 2015.

481   paleobioDB: An R package for downloading, visualizing and processing data from the Paleobiology Database.

482   Ecography, 38(4):419–425.

483   33. Federhen, S. 2012. The NCBI taxonomy database. Nucleic Acids Research, 40(Database issue), D136-43.

484   34. Global Names Architecture 2018. Global Names Resolver. [https://resolver.globalnames.org/] (25/09/2018).

485   35. Bennett, D.J., Sutton, M.D. and Turvey, S.T. 2017. treeman: an R package for efficient and intuitive manipulation

486   of phylogenetic trees. BMC Research Notes, 10(1):30.

487   36. Fisher, R.A. 1919. XV. The Correlation between relatives on the supposition of Mendelian inheritance.

488   Transactions of the Royal Society of Edinburgh, 52(2):399–433.

489   37. Bates, D. Mächler, M., Bolker, B., and Walker, S. 2015. Fitting Linear Mixed-Effects Models Using lme4. Journal

490   of Statistical Software, 67(1):48.

491   38. Akaike, H. 1974. A new look at the statistical model identification. IEEE Transactions on Automatic Control,

492   19(6), 716–723.

493   39. Signor III, P. W. and Lipps, J. H. (1982) "Sampling bias, gradual extinction patterns, and catastrophes in the fossil

494   record", in Geological implications of impacts of large asteroids and comets on the Earth (ed. L. T. Silver and P. H.

495   Schultz), Geological Society of America Special Publication, vol. 190, pp. 291–296.

496   40. Rodrigues, A.S.L., Grenyer, R., Baillie, J.E.M., Bininda-Emonds, O.R.P., Gittlemann, J.L., Hoffmann, M., Safi,

497   K., Schipper, J., Stuart, S.N. and Brooks, T. 2011. Complete, accurate, mammalian phylogenies aid conservation

498   planning, but not much. Proceedings of the Royal Society of London B - Biological Sciences, 366(1578):2652–60.

499   41. Shipman, P. 1981 Life History of a Fossil: An Introduction to Taphonomy and Paleoecology, Harvard University

500   Press, Cambridge, Massachusetts.

501   42. Cooper, R.A. Maxwell, P.A., Crampton, J.S., Beu, A.G., Jones, C.M. and Marshall, B.A. 2006. Completeness of

502   the fossil record: Estimating losses due to small body size. Geology, 34(4):241–244.

503   43. Ricklefs, R.E. 2005. Small clades at the periphery of passerine morphological space. The American Naturalist,

504   165(6):651–9.

505   44.   Latiolais, J.M., Taylor, M.S., Roy, K. and Hellberg, M.E. 2006. A molecular phylogenetic analysis of strombid

506   gastropod morphological diversity. Molecular Phylogenetics and Evolution, 41(2):436–444.

507   45.   Magnuson-Ford, K., Ingram, T., Redding, D.W. and Mooers, A.Ø. 2009. Rockfish (Sebastes) that are

508   evolutionarily isolated are also large, morphologically distinctive and vulnerable to overfishing. Biological

509   Conservation, 142(8):1787–1796.

510   46.   Redding, D.W., DeWolff C.V. and Mooers, A.Ø. 2010. Evolutionary distinctiveness, threat status, and ecological

511   oddity in Primates. Conservation Biology, 24(4):1052–1058.

512   47.   Turvey ST, Blackburn TM. 2011. Determinants of species abundance in the Quaternary vertebrate fossil record.

513   Paleobiology 37: 537-546.

514   48.   Foote, M. and Raup, D.M. 1996. Fossil preservation and the stratigraphic ranges of taxa. Paleobiology, 22(2):121–

515   140.

516   49.   Vane-Wright, R. I., Humphries, C. J. and Williams, P. H. 1991 What to protect—systematics and the agony of

517   choice. Biol. Conserv. 55, 235–254.

518   50.   Mace, G., and Collar, N. 2002. Priority-setting in species conservation. In K. Norris and D. Pain (Eds.),

519   Conserving Bird Biodiversity: General Principles and their Application (Conservation Biology, pp. 61-73).

520   Cambridge: Cambridge University Press.

521   51.   Entwistle, A. and Dunstone, N. 2000. Priorities for the Conservation of Mammalian Diversity: Has the Panda had

522   its Day? Cambridge: Cambridge University Press.

523   52.   Brooks, T. 2010. Conservation planning and priorities   . In N.S. Sodhi and P.R. Ehrlich (Eds.), Conservation

524   Biology for All. Oxford Scholarship Online.

525   53.   Celis-Diez, J.L., Hetz, J., Marín-Vial, P.A., Fuster, G., Necochea, P., Vásquez, R.A. Jaksic, F.M., Armesto, J.J.

526   2012. Population abundance, natural history, and habitat use by the arboreal marsupial Dromiciops gliroides in

527   rural Chiloé Island, Chile. Journal of Mammalogy, 93(1):134–148.

528   54.   Biton, R. Geffen, E., Vences, M., Cohen, O., Bailon, S., Rabinovich, R., Malka, Y., Oron, T., Boistel, R. Brumfeld,

529   V. and Gafny S. 2013. The rediscovered Hula painted frog is a living fossil. Nature Communications, 4:1959.

530   55.   Li, Y-Y., Tsang, E.P.K., Cui, M-Y. and Chen, X-Y. 2012. Too early to call it success: An evaluation of the natural

531   regeneration of the endangered Metasequoia glyptostroboides. Biological Conservation, 150(1):1–4.

532   56.   Mace, G.M., Gittleman, J.J. and Purvis, A. 2003 Preserving the Tree of Life. Science, 300, 1707–1709.

533   57. Purvis, A., Fritz, S.A, Rodríguez, J., Harvey, P.H. and Grenyer, R. 2011. The shape of mammalian phylogeny:

534   patterns, processes and scales. Proceedings of the Royal Society of London B - Biological Sciences,

535   366(1577):2462–77.

536   58. Hagen, O., Hartmann, K., Steel, M. and Stadler, T. 2015. Age-dependent speciation can explain the shape of

537   empirical phylogenies. Systematic Biology, 64(3):432–440.

538   59. Rabosky, D.L. and Goldberg, E.E. 2015. Model inadequacy and mistaken inferences of trait- dependent speciation.

539   Systematic Biology, 64(2):340–355.

540   60. Cardillo, M., Mace, G.M., Jones, K.E., Bielby, J., Bininda-Emonds , O.R.P., Sechrest, W., Orme, C.D.L., Purvis,

541   A. 2005 Multiples causes of high extinction risk in large mammal species. Science, 30912391241

542   61. Turvey ST, Crees JJ, Li Z, Bielby J, Yuan J. 2017. Long-term archives reveal shifting extinction selectivity in

543   China's postglacial mammal fauna. Proceedings of the Royal Society B 284: 20171979.

544