

1 **Title page**

2 **Article title:** The Effect of Phylogenetic Uncertainty and Imputation on EDGE Scores

3 **Running Head:** Effects of Phylogenetic Uncertainty and Imputation on EDGE

4 **Authors:** K. Bodie Weedop<sup>1\*</sup>, Arne Ø. Mooers<sup>2</sup>, Caroline M. Tucker<sup>3</sup>, and William D. Pearse<sup>1</sup>

5 <sup>1</sup> Department of Biology & Ecology Center, Utah State University, 5305 Old Main Hill, Logan UT, 84322

6 <sup>2</sup> Department of Biological Sciences, Simon Fraser University, Burnaby, British Columbia, Canada

7 <sup>3</sup> Department of Biology, University of North Carolina–Chapel Hill

8 \*To whom correspondence should be addressed: K. Bodie Weedop ([kbweedop@gmail.com](mailto:kbweedop@gmail.com))

## 9 Abstract

10 Faced with the challenge of saving as much diversity as possible given financial and time constraints, conser-  
11 vation biologists are increasingly prioritizing species on the basis of their overall contribution to evolutionary  
12 diversity. Metrics such as EDGE (Evolutionary Distinct and Globally Endangered) have been used to set  
13 such evolutionarily-based conservation priorities for a number of taxa, such as mammals, birds, corals, am-  
14 phibians, and sharks. Each application of EDGE has required some form of correction to account for species  
15 whose position within the tree of life are unknown. Perhaps the most advanced of these corrections is phy-  
16 logenetic imputation, but to date there has been no systematic assessment of both the sensitivity of EDGE  
17 scores to a phylogeny missing species, and the impact of using imputation to correct for species missing from  
18 the tree. Here we perform such an assessment, by simulating phylogenies, removing some species to make  
19 the phylogeny incomplete, imputating the position of those species, and measuring (1) how robust ED scores  
20 are for the species that are not removed and (2) how accurate the ED scores are for those removed and  
21 then imputed. We find that the EDGE ranking for species on a tree is remarkably robust to missing species  
22 from that tree, but that phylogenetic imputation for missing species, while unbiased, does not accurately  
23 reconstruct species' evolutionary distinctiveness. On the basis of these results, we provide clear guidance for  
24 EDGE scoring in the face of phylogenetic uncertainty.

25 **Keywords:** conservation prioritization, evolutionary distinctiveness, EDGE, phylogenetic imputation

## 26 Introduction

27 Evidence from the fossil record and present-day studies argue we are in the midst of, or entering, a sixth  
28 mass extinction (Barnosky et al., 2011; Ceballos et al., 2015), such that more populations than ever are  
29 declining and species face heightened danger of extinction (Wake and Vredenburg, 2008; Thomas et al.,  
30 2004). Habitat destruction (Brooks et al., 2002), invasive species (Molnar et al., 2008), climate change  
31 (Pounds et al., 2006), and disease (Lips et al., 2006) are some of the leading causes of species declines  
32 globally. Conservation biologists seek to reduce these detrimental effects on species populations, but in  
33 reality they have limited resources with which to do so. This challenge, termed the “Noah’s Ark problem”  
34 (Weitzman, 1998), has driven conservation biologists to identify different ways by which to prioritize, or  
35 triage, their resource allocation (Bottrill et al., 2008).

36 Conservation triage, like all sound decision-making, requires a method to quantify the relative urgency or  
37 importance for conservation among a set of options. This allows scientists and policy-makers to use data to  
38 quantify need and inform conservation decision-making and management activities. One triage strategy uses  
39 the EDGE metric to identify and prioritize species that are Evolutionarily Distinct and Globally Endangered  
40 (Isaac et al., 2007). Evolutionary Distinctiveness (ED) measures the relative contributions made by each  
41 species within a particular clade to phylogenetic diversity, assigning each branch length equally to all the  
42 subtending species (Redding, 2003; Isaac et al., 2007). Global Endangerment (GE), assigns numerical values  
43 to each of the International Union for Conservation of Nature (IUCN) Red List Categories. As species  
44 become increasingly threatened and are placed into categories of increasing concern (*e.g.* from Vulnerable  
45 to Endangered), the GE numerical value increases. A species’ EDGE score is an aggregate value intended  
46 to equally reflect a species’ evolutionary distinctiveness and conservation status (even if it does not always  
47 in practice; see Pearse et al., 2015).

48 Usage of the EDGE metric has expanded greatly. First used to prioritize global mammals (Isaac et al.,  
49 2007), EDGE scores are now available for a variety of taxonomic groups, including amphibians (Isaac et al.,  
50 2012), birds (Jetz et al., 2014), corals (Curnick et al., 2015), squamate reptiles (Tonini et al., 2016), sharks  
51 (Stein et al., 2018), and all tetrapods (Gumbs et al., 2018). Related metrics are also now available, each  
52 subtly emphasizing different things, such as the expected contribution of each species to future phylogenetic  
53 diversity (HEDGE, I-HEDGE; Steel, Mimoto, and Mooers, 2007; Jensen et al., 2016) and our uncertainty  
54 over a species’ future (EDAM; Pearse et al., 2015) The development and expansion of EDGE-like metrics  
55 mirrors progress in other areas of conservation biology, and the likelihood of success in conservation (Wilson  
56 et al., 2007; McBride et al., 2007), the relative cost of certain interventions (Naidoo et al., 2006), and

57 complementarity of interventions (Pressey et al., 1993; Myers et al., 2000) can now be considered in its  
58 calculation. The EDGE index was developed explicitly with the intention of informing conservation triage,  
59 and is now the basis of the global EDGE of Existence Program (<http://www.edgeofexistence.org/>). The  
60 successful application of EDGE highlights the potential for phylogenetic conservation prioritization metrics  
61 to provide actionable insights while quantitatively measuring the evolutionary history a species represents.  
62 Nonetheless, almost every application of an EDGE-type approach must address uncertainty resulting from  
63 missing data. Addressing, and hopefully improving, our ability to handle uncertainty should be a continual  
64 effort to increase the support for such approaches. **However, in an effort to delineate between science and  
65 policy, it is important to note that the implications of missing data on policy making will vary depending  
66 upon the demands and goals of a particular person or organization.**

67 Missing data can affect EDGE scores in several ways. First, the IUCN identifies some species as Data  
68 Deficient (IUCN, 2001; IUCN, 2008), which affects the GE component of a species' EDGE score. Fortunately,  
69 the IUCN provides guidance for using any available contextual data to assign some threat status to such  
70 species. A number of studies illustrate how to assign threat categories to Data Deficient species, which  
71 in turn should reduce the uncertainty in GE (Good, Zjhra, and Kremen, 2006; Butchart and Bird, 2010;  
72 Morais et al., 2013; Dulvy et al., 2014). The issue of missing phylogenetic data is arguably more complicated  
73 because not only does the focal species have no ED score, but its absence from the phylogeny may affect the  
74 ED scores of related species. Species of conservation concern are almost by definition rare, and frequently  
75 lack sufficient DNA (or even morphological) data to be placed with certainty on a phylogeny. In most cases,  
76 taxonomic information rather than sequence data alone has been used to place species in the tree of life  
77 when constructing EDGE lists (see Isaac et al., 2007; Collen et al., 2011; Isaac et al., 2012; Jetz et al., 2014;  
78 Curnick et al., 2015; Stein et al., 2018; Gumbs et al., 2018; Forest et al., 2018). Yet, to our knowledge,  
79 there has been no systematic study of the effect of imputation on species' EDGE scores, despite this practice  
80 having received attention in other areas of comparative biology (Kuhn, Mooers, and Thomas, 2011; Thomas  
81 et al., 2013; Rabosky, 2015). Thus we do not know how accurate EDGE scores are when species are missing,  
82 or when species are added to phylogenies by imputation, nor do we know how accurate EDGE scores for  
83 imputed species might be. As interest in using EDGE-type measures and phylogenies for conservation triage  
84 grows, the need for consensus on how to resolve cases of phylogenetic uncertainty becomes increasingly  
85 urgent.

86 Here we attempt to quantify the effect of one sort of phylogenetic uncertainty—the effect of missing species  
87 on EDGE rankings—and assess the degree to which subsequent imputation affects the accuracy of EDGE  
88 scores. We do so by simulating phylogenies and then removing species either at random, or with bias, across

89 those phylogenies. By contrasting the ED scores of the species before and after the loss of other species from  
90 the phylogenies, we measure the impact of missing species on ED scores. We then assess the extent to which  
91 phylogenetic imputation can accurately estimate the EDGE scores of missing species in simulated data. We  
92 also examine the extent to which such imputation affects the scores of species for which we have data. In  
93 doing so, we hope to provide clear guidance as to the applicability of phylogenetic imputation as a solution  
94 for species missing phylogenetic data. From our results, we argue that species' ED values are remarkably  
95 robust to missing species, and that phylogenetic imputation does not reliably reconstruct the true ranking  
96 of those missing species.

## 97 **Methods**

98 We use a simulation approach to test the effect of having missing species on a phylogeny (through species re-  
99 moval from simulated phylogenies) and then imputing species for species' ED (Evolutionary Distinctiveness)  
100 scores. We focus exclusively on the ED-component of the EDGE metric, since uncertainty in species GE  
101 scores has already been addressed by the IUCN's proposal to assign Data Deficient species scores (IUCN,  
102 2001; IUCN, 2008). Because EDGE is the product of both ED and GE components, even perfectly accurate  
103 GE values could be associated with imperfect EDGE scores if the ED scores were inaccurate.

104 All trees (both starting and imputed) were simulated under a pure-birth Yule model using 'gieger::sim.bdtree'  
105 (setting parameters  $b=1$  and  $d=0$ ; Pennell et al., 2014). This model was chosen because it is the simplest model  
106 possible: speciation rates are constant across the entire tree of life and there is no extinction. We suggest  
107 that imputation under a simple model that is identical to that used to simulate the data is a low, and fair,  
108 benchmark for a method to meet. However, we acknowledge that more complex and/or biologically realistic  
109 models of diversification could potentially improve the performance of imputation. We used 'caper::ed.calc'  
110 to calculate ED values (Orme et al., 2013). All simulations and analyses were performed using R (version  
111 3.4.0; R Core Team, 2017). We performed 100 replicate simulations of each parameter combination. All  
112 our analysis code is available online (<https://github.com/bweedop/edgeSims>) and in the [supplementary](#)  
113 [materials](#).

## 114 **The impact of missing species on EDGE scores**

115 Our first set of simulations assess the impact of missing species data on the ED scores of remaining species,  
116 considering data missing either in a random or phylogenetically-biased fashion. We simulated phylogenies  
117 of different sizes (number of species: 64, 128, 256, ..., 2048, 4096) and then removed constant fractions of  
118 tips from the tree (0%, 1%, 2%, ..., 19%, ..., 99%). To simulate species missing at random throughout the  
119 phylogeny, we used 'sample()' to select the relevant fraction of species (rounded to the nearest whole number)  
120 without replacement. To remove species in a phylogenetically-biased manner, we used Felsenstein (2005)'s  
121 threshold model. We simulated a trait under a constant rate Brownian-motion model ( $\sigma=0.5$ , starting root  
122 value = 1) (using 'geiger::sim.char' Pennell et al., 2014). Species were then removed from the tree if their  
123 simulated trait was in the upper quantile matching the fraction of species to be removed. For example,  
124 if 10% of species were to be removed from the tree, the species with the highest 10% of values would be  
125 removed. This results in closely related species being removed more often than expected by chance.

126 To quantify the effect of these manipulations, we calculated the ED values of species that are not removed  
127 from a tree both before and after removal. We then correlated these ED scores: if missing species do not  
128 affect ED values of the remaining species, we would expect a strong, positive correlation between the ED  
129 scores of the remaining species calculated before and after species were removed from the phylogeny. We  
130 emphasize that species removed from the phylogeny are omitted from this comparison. We outline our  
131 approach in figure 1.

## 132 **The impact of phylogenetic imputation on EDGE scores**

133 Our second set of simulations tested the impact of imputation on ED scores within an imputed clade. We  
134 used relatively small clades (5, 6, 7, ..., 30, 31, 32 species) from phylogenies of different sizes [128 ( $2^7$ ), 147  
135 ( $2^{7.2}$ ), 168 ( $2^{7.4}$ ), ... , 776 ( $2^{9.6}$ ), 891 ( $2^{9.8}$ ), 1024 ( $2^{10}$ ) species). We first randomly selected a clade to be  
136 removed from the 'true' tree and then simulated a new phylogeny of the same size as the removed clade.  
137 This newly simulated clade was generated under the same pure-birth model as the original phylogeny. We  
138 then placed the newly simulated clade in the full phylogeny, in the same location as the removed clade. If a  
139 newly simulated clade was so old that it was not possible to graft it into place, we discarded that clade and  
140 simulated another. In an empirical study the model of evolution under which the phylogeny had evolved  
141 would have to be estimated, which is an additional source of error not considered here. We simulated each  
142 combination of clade and total phylogeny sizes 100 times when using a pure-birth Yule model and 5 times

143 when simulating under models with past extinction. An overview of our approach is given in figure 2.

144 To assess whether clades, once imputed, had similar ED scores to their true values, we correlated the imputed  
145 ED scores with the true ED scores. We also calculated the sum of the absolute change in ranked ED for all  
146 species, which is particularly relevant for EDGE-listing as conservation actions are often focused around the  
147 top 100, 200, etc., species. Moreover, the correlation of imputed and real scores are bounded by the depth of  
148 the imputed clade, and therefore a high correlation could still produce inaccurate imputed scores, and a low  
149 correlation could still not be important (*e.g.* they could be anticorrelated but still differ in rank by a max  
150 of the size of the subclade). We modeled both of these metrics (the change in ranking and the correlation)  
151 as a function of a number of potential explanatory variables. Specifically, we included in our models: the  
152 estimated speciation rate of the original phylogeny (using ‘ape::yule’; Paradis, Claude, and Strimmer, 2004),  
153 the sum of all phylogenetic branch-lengths in the original phylogeny (Faith’s PD; Faith, 1992), the sum of  
154 all phylogenetic branch-lengths in the original focal clade (Faith’s PD; Faith, 1992), the value of  $\gamma$  in the  
155 original phylogeny (using ‘phytools::gammatest’; Pybus and Harvey, 2000; Revell, 2012), Colless’ index of  
156 the original phylogeny (using ‘apTreeshape::as.treeshape’; Colless, 1982; Bortolussi et al., 2009), the kurtosis  
157 of species’ ED values in the original phylogeny (using ‘moments::kurtosis’; Komsta and Novomestky, 2015),  
158 the skew of species’ ED values in the original phylogeny (using ‘moments::skew’; Komsta and Novomestky,  
159 2015), the total number of species in the original phylogeny, the total number of species within the imputed  
160 clade, and the depth (age) of the imputed clade in the phylogeny. Although the expectations of many of  
161 these explanatory variables are known for Yule trees, in each simulation they are expected to vary somewhat  
162 by chance.

163 Recently, there has been interest in assigning missing species the mean ED score of the most exclusive  
164 clade which contain the species (see Gumbs et al., 2018). To test the efficacy of such methods, we assigned  
165 the average ED of the selected clade to each of its’ species and calculated (as above) the mean change in  
166 absolute ranking under this scheme. Note that we could not correlate ED scores (as we do above), since such  
167 a correlation would require variation in species’ scores and under this approach a single score (the mean ED)  
168 is assigned to all imputed species.

169 We present, in the supplementary materials, two additional sets of analyses intended to examine the impact  
170 that past extinction rates may have played on our analyses. These simulations incorporate conditions of  
171 past extinction at low and high rates using ‘gieger::sim.bdtree’ (setting parameters for low extinction at  
172  $b=1$  and  $d=0.5$  and high extinction  $b=1$  and  $d=0.95$ ; Pennell et al., 2014). These models represent large  
173 departures from our main simulations (which have  $b=1$  and  $d=0$ ), and so we performed only 5 replicates per

174 set of parameter combinations as our only aim was to detect any major differences in our results stemming  
175 from these changes. Otherwise, these simulations were identical to those we present in the main text.

## 176 Results

177 We asked how robust ED scores were for species with known positions on the phylogeny, when other species  
178 were missing from the phylogeny. In fact, when there were increasing numbers of missing species, ED  
179 scores for the remaining species' became less accurate (table 1; figure 3). When species were missing from  
180 the tree in a phylogenetically-based fashion, ED values were less robust as compared to when species are  
181 randomly missing from the tree. However, the effect of missing species is not necessarily severe; even if  
182 20% of species are missing from the tree, the average correlation coefficient between true and estimated ED  
183 scores for the remaining species is 0.88 and 0.94 for phylogenetically-biased and random missing species,  
184 respectively.

185 We also considered the impact of imputation on the accuracy of ED scores for imputed species. When  
186 clades were imputed on the tree, we found a weak (if any) average positive correlation between the imputed  
187 ED and true ED values for species within the imputed clades (overall mean correlation of 0.197 in a statistical  
188 model with an  $r^2$  of 0.5%; figure 4, table 2). We also found no explanatory variables that explained significant  
189 variation in this relationship (table 2; see Appendix S1 in Supporting Information). However, we did find  
190 evidence that, when imputing larger clades, the variation in the correlation between true and imputed ED  
191 scores decreases, although we emphasise the effect is weak (see table 2). When considering rankings rather  
192 than raw scores, we found that imputation can introduce sizable error into the estimation of species' ED  
193 values (figure 5 and table 3). This ranking error increased with the size of the imputed clade and phylogeny  
194 (table 3), and can affect ranking error within the top 100 and 250 species (see Appendix S2 in Supporting  
195 Information). To give an example of the magnitude of the effect, within a phylogeny of 1024 species, the  
196 members of an imputed clade of 30 species are, on average,  $\pm 315$  rankings from their true rankings. We  
197 found similar effects in ranking error when using the average ED value of clade for a missing species (see  
198 Appendix S3 in Supporting Information). [As we show in the supplementary materials, the simulation \(and](#)  
199 [subsequent imputation\) of phylogenies under models incorporating extinction rates \(\*i.e.\*, not Yule models\)](#)  
200 [had qualitatively identical results. We do not, therefore, discuss them in detail here.](#)



## 201 Discussion

202 Phylogenies are playing an increasing role in conservation prioritization, decision-making, and policy (Vézquez  
203 and Gittleman, 1998; Veron et al., 2017). A major obstacle to a more widespread adoption of phylogenetic  
204 prioritization methods such as EDGE is phylogenetic uncertainty (Collen, 2015). There is a tension between  
205 a purported need to make decisions to preserve biodiversity—including evolutionary history—now, and the  
206 reality that we rarely have complete information about the phylogenetic placement of many species of con-  
207 servation concern (Isaac and Pearse, 2018). The intention of our study is to provide concrete information  
208 about the impact of one source of phylogenetic uncertainty - missing species - on conservation prioritization.  
209 To address this uncertainty, we addressed two key issues: (1) the extent to which species that are missing  
210 from the tree of life impact the ED scores of species for which we do have data, and (2) the extent to which  
211 phylogenetic imputation can accurately estimate ED scores for taxa with no phylogenetic data. First, we  
212 found that missing species had a surprisingly small impact on the ED scores of other species, particularly if  
213 species are missing at random from the tree of life. Second, we found that phylogenetic imputation generally  
214 fails to accurately reconstruct species’ ED scores and rankings.

215 [In this study, we have examined imputation under three separate models of diversification: pure-birth Yule](#)  
216 [models \(presented in the main text\), and models with relatively high and low rates of extinction \(both](#)  
217 [in the Supporting Information\).](#) We acknowledge that lineages evolve in more complex ways, although  
218 we suggest that focusing on these fundamental models of diversification makes our results more broadly  
219 applicable. We suggest that a method should perform well under basic conditions, and as such these results  
220 form an appropriate benchmark, particularly given we can see no reason to suppose that more complex  
221 models should increase model performance. Further, we focus here solely on the results from a single  
222 imputation in each simulation, despite, empirically, biologists reporting average ED scores calculated across  
223 pseudo-posterior distributions of many imputed phylogenies (Kuhn et al., 2011). Thus our results show  
224 that the variation within these pseudo-posterior distributions is likely very large. It is well-known that  
225 such imputation methods are not biased (indeed, this was originally shown by Kuhn et al., 2011): here we  
226 emphasize that the uncertainty they introduce is sufficiently large such that they may be less informative  
227 than previously has been thought.

228 [Conservation prioritization and triage have been controversial: to some triage represents an unacceptable](#)  
229 [defeat by accepting that some species will go extinct \(Jachowski and Kesler, 2009; Parr et al., 2009\), while](#)  
230 [to others it is either efficient resource allocation or a grim necessity \(Bottrill et al., 2008\).](#) The debate over  
231 [the implications of triage, both philosophically and practically, is an important one, but this study does not](#)

232 address it. Conservation biology has been described as a crisis discipline where it is often necessary to act  
233 with imperfect information and, ultimately, tolerate and manage uncertainty (Soulé, 1985). Our intention  
234 here is to shine a light on how phylogenetic uncertainty and imputation can impact species ED(GE) scores.  
235 While we feel that EDGE and related approaches are worthwhile for conservation biologists, every user of  
236 any triage method must weigh the potential benefits and drawbacks associated with that method.

## 237 **ED scores are relatively robust to missing species**

238 Missing species and poor phylogenetic resolution have been identified as causes of uncertainty when calculat-  
239 ing ED (Isaac et al., 2007), but we were unable to find a quantitative assessment of how missing species might  
240 affect ED values of species for which data is available. Empirically in corals and gymnosperms, incomplete  
241 phylogenies produced similar results as later, more complete trees (Curnick et al., 2015; Forest et al., 2018).  
242 Our results support this finding. Indeed, our analysis suggests that, on average (and we emphasize that  
243 there is a good amount of variation about that average; see figure 3), a phylogeny missing 20% of species at  
244 random will still have ED scores for the remaining species that are strongly correlated (mean  $\rho = 0.94$ )  
245 with the true ED scores.

246 We did find that missing species are more problematic when those species are non-randomly distributed  
247 across the phylogeny. Our simulations do not examine extreme phylogenetic patterning, such as if an entire  
248 clade were missing. This is notable because clades that are geographically restricted to difficult-to-reach  
249 regions are both difficult to sequence and not uncommon (as is seen with 27 coral species in the Indian  
250 Ocean; Arrigoni et al., 2012). We also do not attempt to comprehensively simulate all of the different ways  
251 in which species could be missing from a phylogeny. We emphasize that we have not demonstrated, and  
252 do not argue, that missing species cannot affect ED scores. We simply demonstrate that, compared to  
253 a scenario in which species are missing at random, phylogenetically patterned missing species can have a  
254 greater effect on the ED scores of species for which we have data, and that (in our opinion) ED scores are  
255 remarkably robust to missing species. Other patterns and scenarios for species to be missing could easily  
256 lead to systematic biases of ED scores, and so very effort should be made to gather accurate phylogenetic  
257 information for all species within a clade before prioritisation is carried out.

258 **Imputation does not reconstruct the ED values of missing species with great**  
259 **precision**

260 Our results show that neither imputation (figures 4 and 5), nor clade-averages of ED (see Appendix S3  
261 in Supporting Information), accurately recover the true ED values or the true ED rank of missing species.  
262 Thus we argue that, even though imputation allows missing species to be incorporated into EDGE lists, their  
263 associated EDGE scores may not accurately reflect their true scores. We acknowledge these are averages  
264 and may change depending on particular phylogeny, but we can find no statistically significant predictors of  
265 that variation.

266 While we did not assess clades with fewer than five species (we do not consider correlations or averages to  
267 be reliable with so few data-points), we cannot think why smaller clades would necessarily be more reliable  
268 (and this would require a large deviation from the trend in figure 4). Indeed, in the smallest possible clade  
269 (two species), imputation is essentially sampling a terminal branch length from an exponential distribution  
270 (Kuhn et al., 2011); such a process should still lead to a great degree of uncertainty.

271 It is, perhaps, unsurprising that imputed ED values do not correlate with their true values (see figure 4),  
272 but we were surprised at the degree of ranking error. Indeed, larger phylogenies showed *greater* ranking  
273 error; we naïvely would have expected the opposite. We would expect the the upper bound on the age of  
274 the imputed clade, which should have expected be relatively younger in larger phylogenies, would partially  
275 controlled the range of the ranks for the imputed species. ED is known to be driven mostly by terminal  
276 branch length (Isaac et al., 2007; Steel et al., 2007; Redding et al., 2008); our results therefore emphasize  
277 this.

278 Imputation is not the only way to incorporate missing species into EDGE-like frameworks (see Collen et al.,  
279 2011; Gumbs et al., 2018), but it is likely the most common. 3,330 of the birds (~30%; Jetz et al., 2014),  
280 250 of the mammals (~5.6%; Collen et al., 2011), and 610 of the sharks (~49%; Stein et al., 2018) in recent  
281 EDGE lists were imputed. It is well-known that phylogenetic imputation can cause biases in other statistical  
282 methods, such as the estimation of evolutionary phylogenetic signal (Rabosky, 2015). We emphasize that  
283 we are not suggesting that imputation *biases* ED scores: we are, instead, suggesting that it is less precise  
284 than has previously been acknowledged.

## 285 Guidelines for the use of imputation

286 The impact of imputation on EDGE scores is almost certainly less than its impact on ED scores, because  
287 EDGE scores are a product of both ED and IUCN status ('GE'). However, the goal of EDGE-like measures  
288 is to incorporate phylogeny, and if imputed EDGE scores are driven by their GE component because of  
289 uncertainty introduced by imputation, this essentially creates another metric of IUCN status. With this in  
290 mind, we hope to provide clear guidelines, along with the benefits and drawbacks, when using imputation in  
291 EDGE-based approaches to scientists and policy makers.

292 Our results further suggest that incomplete phylogenies can be used to estimate ED scores with remarkably  
293 high degrees of accuracy. Instead of using imputation to account **solely** for the relatively minor impact of miss-  
294 ing species, we suggest that conservation biologists should ~~without accounting for phylogenetic uncertainty,~~  
295 ~~address focus on~~ the **phylogenetic uncertainty** of species for which they have data. While we have not ex-  
296 plored this uncertainty here, evolutionary biologists commonly work with distributions of trees generated  
297 from genetic data (reviewed in Huelsenbeck et al., 2001; Bollback, 2005), since the precise topology and  
298 dating of a phylogeny is almost always uncertain. This uncertainty has, indeed, already been shown to affect  
299 EDGE scores and rankings (Pearse et al., 2015). If biologists are concerned about the impact of missing  
300 species on known species' ED(GE) scores we see no harm in being precautionary and using imputation. It  
301 is important, however, to focus on known sources of potential error, and so we would encourage biologists  
302 to incorporate uncertainty in species with phylogenetic data as a priority.

303 Our results suggest that prioritizing species whose phylogenetic structure has been imputed should be done  
304 with extreme care, if at all. In the case that an species is imputed to be below a threshold set for conservation  
305 (most EDGE studies focus on the 'top 100' species or something similar), then the path forward is clear:  
306 that species should not have conservation funds allocated to it at this time. The case where a species, on  
307 average, passes a threshold is more complex, but the theory underlying imputation can give some guidance.  
308 Imputed distributions of trees essentially represent Bayesian posterior distributions (Kuhn et al., 2011), and  
309 so the 95% posterior densities of these distributions' ED values represent a range within which we can be  
310 95% certain the true ED scores lie (if the model assumptions are met). Thus we suggest that conservation  
311 action should only be initiated for a species if there is a 95% (or 80%, or whatever confidence is deemed  
312 appropriate) probability that it is above that threshold. For example, a species whose ranking is estimated to  
313 have a 20% probability of being between the 1<sup>st</sup> and 100<sup>th</sup> highest-ranked species could not, with confidence,  
314 be called a top-100 species. Our results suggest that, on average, very few imputed species will meet such a  
315 criterion. Regardless, the calculations of such probabilities is trivial with the data users of imputation have

316 in hand already.

317 Ultimately, we are currently fighting a losing battle to preserve the tree of life. Our results are good news:  
318 they suggest that we can start right away using the (incomplete) phylogenies we already have. The effect of  
319 missing species is negligible enough that we often do not need time-consuming imputation, and imputation  
320 rarely gives us sufficiently precise estimates of species' ED scores anyway. We suggest that, given we do not  
321 have the resources to save everything, we should consider focusing our efforts on those species whose ED  
322 scores we can know with greater certainty: those for which we have data.

## 323 **Acknowledgments**

324 We thank the Department of Biology and Ecology Center of Utah State University for their support. We are  
325 grateful to E. Simpson, M. Sneddon, J. Stachewicz and J. Rosindell for ongoing discussion and-or constructive  
326 feedback on this manuscript.

## References

- 327
- 328 Arrigoni, Roberto et al. (2012). Molecular phylogeny of the robust clade (Faviidae, Mussidae, Merulinidae,  
329 and Pectiniidae): an Indian Ocean perspective. *Mol. Phylogenet. Evol.* 65.1, 183–193.
- 330 Barnosky, Anthony D et al. (2011). Has the Earth’s sixth mass extinction already arrived? *Nature* 471.7336,  
331 51–57.
- 332 Bollback, JonathanP. (2005). Posterior Mapping and Posterior Predictive Distributions. *Statistical Methods*  
333 *in Molecular Evolution*. Ed. by Rasmus Neilsen. Springer New York. Chap. 16, 439–462.
- 334 Bortolussi, Nicolas et al. (2009). Package ‘apTreeshape’.
- 335 Bottrill, Madeleine C et al. (2008). Is conservation triage just smart decision making? *Trends. Ecol. Evol.*  
336 23.12, 649–654.
- 337 Brooks, Thomas M et al. (2002). Habitat loss and extinction in the hotspots of biodiversity. *Conserv. Biol.*  
338 16.4, 909–923.
- 339 Butchart, Stuart HM and Jeremy P Bird (2010). Data deficient birds on the IUCN Red List: what don’t we  
340 know and why does it matter? *Biol. Conserv.* 143.1, 239–247.
- 341 Ceballos, Gerardo et al. (2015). Accelerated modern human–induced species losses: Entering the sixth mass  
342 extinction. *Sci. Adv.* 1.5, e1400253.
- 343 Collen, Ben (2015). Conservation prioritization in the context of uncertainty. *Anim. Conserv.* 18.4, 315–317.
- 344 Collen, Ben et al. (2011). Investing in evolutionary history: implementing a phylogenetic approach for mam-  
345 mal conservation. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* 366.1578, 2611–2622.
- 346 Colless, Donald H. (Mar. 1982). Review of phylogenetics: the theory and practice of phylogenetic systematics.  
347 *Syst. Zool.* 31.1, 100–104.
- 348 Curnick, DJ et al. (2015). Setting evolutionary-based conservation priorities for a phylogenetically data-poor  
349 taxonomic group (Scleractinia). *Anim. Conserv.* 18.4, 303–312.
- 350 Dulvy, Nicholas K et al. (2014). Extinction risk and conservation of the world’s sharks and rays. *Elife* 3,  
351 e0059.
- 352 Faith, Daniel P (1992). Conservation evaluation and phylogenetic diversity. *Biol. Conserv.* 61.1, 1–10.
- 353 Felsenstein, Joseph (2005). Using the quantitative genetic threshold model for inferences between and within  
354 species. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* 360.1459, 1427–1434.
- 355 Forest, Félix et al. (2018). Gymnosperms on the EDGE. *Sci. Rep.* 8.1, 6053.
- 356 Good, Tatjana C, Michelle L Zjhra, and Claire Kremen (2006). Addressing data deficiency in classifying  
357 extinction risk: a case study of a radiation of Bignoniaceae from Madagascar. *Conserv. Biol.* 20.4, 1099–  
358 1110.

359 Gumbs, Rikki et al. (2018). Tetrapods on the EDGE: Overcoming data limitations to identify phylogenetic  
360 conservation priorities. *PloS One* 13.4, e0194680.

361 Huelsenbeck, John P et al. (2001). Bayesian inference of phylogeny and its impact on evolutionary biology.  
362 *Science* 294.5550, 2310–2314.

363 Isaac, Nick JB and William D Pearse (2018). The Use of EDGE (Evolutionary Distinct Globally Endangered)  
364 and EDGE-Like Metrics to Evaluate Taxa for Conservation. *Phylogenetic Diversity*. Springer, 27–39.

365 Isaac, Nick J et al. (2007). Mammals on the EDGE: conservation priorities based on threat and phylogeny.  
366 *PloS One* 2.3, e296.

367 Isaac, Nick J et al. (2012). Phylogenetically-informed priorities for amphibian conservation. *PloS One* 7.8,  
368 e43912.

369 IUCN (2001). *IUCN Red List categories and criteria*. International Union for Conservation of Nature.

370 — (2008). *Guidelines for Using the IUCN Red List Categories and Criteria, Version 7.0*. International Union  
371 for Conservation of Nature.

372 Jachowski, David S and Dylan C Kesler (2009). Allowing extinction: should we let species go? *Trends in*  
373 *Ecology & Evolution* 24.4, 180.

374 Jensen, Evelyn L et al. (2016). I-HEDGE: determining the optimum complementary sets of taxa for conser-  
375 vation using evolutionary isolation. *PeerJ* 4, e2350.

376 Jetz, Walter et al. (2014). Global distribution and conservation of evolutionary distinctness in birds. *Curr.*  
377 *Biol.* 24.9, 919–930.

378 Komsta, Lukasz and Frederick Novomestky (2015). Moments, cumulants, skewness, kurtosis and related  
379 tests. *R package version 0.14*.

380 Kuhn, Tyler S, Arne Ø Mooers, and Gavin H Thomas (2011). A simple polytomy resolver for dated phylo-  
381 genies. *Methods Ecol. Evol.* 2.5, 427–436.

382 Lips, Karen R et al. (2006). Emerging infectious disease and the loss of biodiversity in a Neotropical am-  
383 phibian community. *Proc. Natl. Acad. Sci. U.S.A.* 103.9, 3165–3170.

384 McBride, Marissa F et al. (2007). Incorporating the effects of socioeconomic uncertainty into priority setting  
385 for conservation investment. *Conserv. Biol.* 21.6, 1463–1474.

386 Molnar, Jennifer L et al. (2008). Assessing the global threat of invasive species to marine biodiversity. *Front.*  
387 *Ecol. Environ.* 6.9, 485–492.

388 Morais, Alessandro R et al. (2013). Unraveling the conservation status of Data Deficient species. *Biol.*  
389 *Conserv.* 166, 98–102.

390 Myers, Norman et al. (2000). Biodiversity hotspots for conservation priorities. *Nature* 403.6772, 853–858.

391 Naidoo, Robin et al. (2006). Integrating economic costs into conservation planning. *Trends. Ecol. Evol.* 21.12,  
392 681–687.

393 Orme, David et al. (2013). *caper: Comparative Analyses of Phylogenetics and Evolution in R*. R package  
394 version 0.5.2. URL: <https://CRAN.R-project.org/package=caper>.

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

397 Parr, Michael J et al. (2009). Why we should aim for zero extinction. *Trends in Ecology & Evolution* 24.4,  
398 181.

399 Pearse, William D et al. (2015). Beyond the EDGE with EDAM: prioritising British plant species according  
400 to evolutionary distinctiveness, and accuracy and magnitude of decline. *PloS one* 10.5, e0126524.

401 Pennell, Matthew W et al. (2014). geiger v2. 0: an expanded suite of methods for fitting macroevolutionary  
402 models to phylogenetic trees. *Bioinformatics* 30.15, 2216–2218.

403 Pounds, J Alan et al. (2006). Widespread amphibian extinctions from epidemic disease driven by global  
404 warming. *Nature* 439.7073, 161–167.

405 Pressey, RL et al. (1993). Beyond opportunism: key principles for systematic reserve selection. *Trends. Ecol.*  
406 *Evol.* 8.4, 124–128.

407 Pybus, Oliver G and Paul H Harvey (2000). Testing macro-evolutionary models using incomplete molecular  
408 phylogenies. *Proc. R. Soc. Lond., B, Biol. Sci.* 267.1459, 2267–2272.

409 R Core Team (2017). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical  
410 Computing. Vienna, Austria. URL: <https://www.R-project.org/>.

411 Rabosky, Daniel L (2015). No substitute for real data: a cautionary note on the use of phylogenies from  
412 birth–death polytomy resolvers for downstream comparative analyses. *Evolution* 69.12, 3207–3216.

413 Redding, D (2003). Incorporating genetic distinctness and reserve occupancy into a conservation prioritisation  
414 approach. *Master’s thesis, University of East Anglia, Norwich OpenURL*.

415 Redding, David W et al. (2008). Evolutionarily distinctive species often capture more phylogenetic diversity  
416 than expected. *J. Theor. Biol.* 251.4, 606–615.

417 Revell, Liam J. (2012). phytools: An R package for phylogenetic comparative biology (and other things).  
418 *Methods Ecol. Evol.* 3, 217–223.

419 Soulé, Michael E. (1985). What is conservation biology? *BioScience* 35.11, 727–734.

420 Steel, Mike, Aki Mimoto, and Arne Ø Mooers (2007). Hedging our bets: the expected contribution of species  
421 to future phylogenetic diversity. *Evol. Bioinform.* 3, 237.

422 Stein, R William et al. (2018). Global priorities for conserving the evolutionary history of sharks, rays and  
423 chimaeras. *Nat. Ecol. Evol.* 1.



424 Thomas, Chris D et al. (2004). Extinction risk from climate change. *Nature* 427.6970, 145–148.

425 Thomas, Gavin H et al. (2013). PASTIS: an R package to facilitate phylogenetic assembly with soft taxonomic  
426 inferences. *Methods Ecol. Evol.* 4.11, 1011–1017.

427 Tonini, João Filipe Riva et al. (2016). Fully-sampled phylogenies of squamates reveal evolutionary patterns  
428 in threat status. *Biol. Conserv.* 204, 23–31.

429 Veron, Simon et al. (2017). Predicting loss of evolutionary history: Where are we? *Biol. Rev.* 92.1, 271–291.

430 Vézquez, Diego P and John L Gittleman (1998). Biodiversity conservation: does phylogeny matter? *Curr.*  
431 *Biol.* 8.11, R379–R381.

432 Wake, David B and Vance T Vredenburg (2008). Are we in the midst of the sixth mass extinction? A view  
433 from the world of amphibians. *Proc. Natl. Acad. Sci. U.S.A.* 105.Supplement 1, 11466–11473.

434 Weitzman, Martin L (1998). The Noah’s ark problem. *Econometrica*, 1279–1298.

435 Wilson, Kerrie A et al. (2007). Conserving biodiversity efficiently: what to do, where, and when. *PLoS Biol.*  
436 5.9, e223.

## 437 **Data Accessibility Statement**

438 All simulation and analysis code, along with underlying data, generated for this study are in the supplemen-  
439 tary materials and online at: <https://github.com/bweedop/edgeSims>

## 440 **Tables**

	Estimate	Std. Error	t value	Pr(> t )
Reference—Phylogenetically biased				
Intercept	1.0315	0.0013	821.39	<0.0001
Fraction of species removed	-0.4696	0.0020	-233.16	<0.0001
Number of species overall	$2.500 \times 10^{-6}$	$2.984 \times 10^{-7}$	7.89	<0.0001
Contrast—Random				
Intercept	0.0630	0.0018	35.47	<0.0001
Fraction of species removed	-0.2774	0.0028	-97.45	<0.0001
Number of species overall	$5.013 \times 10^{-6}$	$4.219 \times 10^{-7}$	-4.38	<0.0001

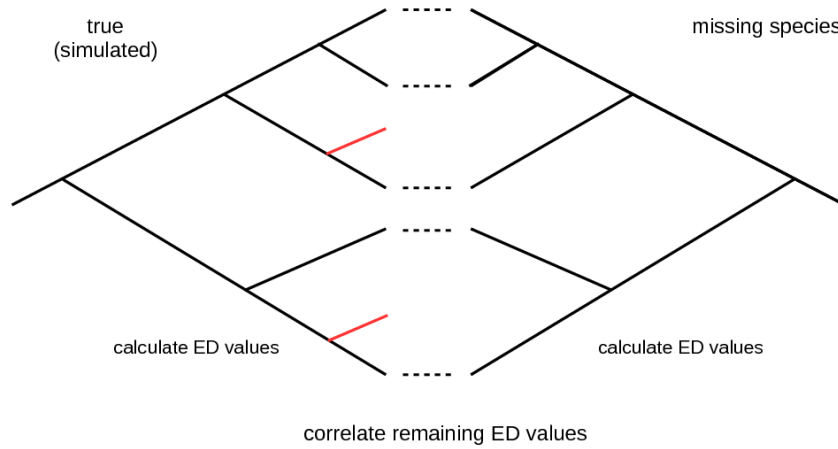
**Table 1: Statistical model of the effect of missing data on the calculation of the remaining species’ ED values.** Results of a multiple regression fit to the data shown in figure 3, regressing the correlation coefficient of (remaining) species’ ED scores before and after other species were removed from the phylogeny ( $F_{139696,5} = 40,350$ ,  $r^2 = 0.5908$ ,  $p < 0.0001$ ). We emphasize that these are simulated data, and so, as the extremely large sample sizes are likely driving the low standard errors of the model terms, we encourage the reader to focus on the magnitudes of the effects and the overall variance explained by our model ( $r^2 = 0.5908$ ). The first three rows refer to the overall intercept, effect of the fraction of species removed from the phylogeny, and the overall size of the phylogeny when species were removed in a phylogenetically biased fashion. The last three rows are contrasts, [reporting whether there is a difference in each coefficient when the simulations were conducted with random, or phylogenetically biased, species loss.](#) ~~reporting the difference (contrast) of each parameter when species were removed at random from the phylogeny, whether random loss of species has a statistically different effect.~~ The correlation of ED scores appears affected by an interaction between the number of species removed from the tree and whether those species were removed at random or in a phylogenetically-biased fashion. The overall size of the phylogeny has little discernible effect, and its statistical significance is likely driven by the large number of simulations we performed (139,700).

	Estimate	Std. Error	t value	Pr(> t )
Intercept	0.1974	0.0501	3.94	0.0001
Size of Focal Clade	-0.0036	0.0005	-7.60	< 0.0001
Size of Phylogeny	0.0001	0.0001	0.60	0.5497
PD	-0.0001	0.0001	-0.64	0.5241
Estimated speciation rate	-0.0199	0.0493	-0.40	0.6865
Colless’ Index	-0.0000	0.0000	-0.08	0.9380
Skew	0.0022	0.0083	0.27	0.7885
Kurtosis	-0.0001	0.0008	-0.16	0.8736
Depth of Imputed Clade	0.0006	0.0005	1.27	0.2045

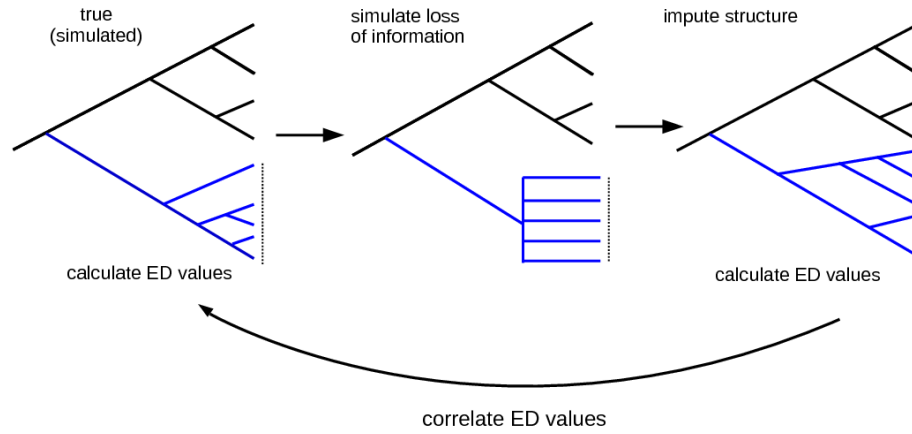
**Table 2: Statistical model of the potential drivers of the correlation between imputed and true ED values.** Results of a multiple regression fitted to the data shown in figure 4, showing a relatively poor correlation between imputed and true ED scores ( $F_{44791,8} = 29.1$ ,  $r^2 = 0.005$ ,  $p < 0.0001$ ). Given the extremely low predictive power of this statistical model we are reticent to make strong claims about drivers of the correlation between imputed and observed ED. [Each coefficient refers to a measured variable in our simulations, as described in the text.](#)

	Estimate	Std. Error	t value	Pr(> t )
Intercept	-1.6344	0.0332	-49.29	0.0001
Size of focal (imputed) clade	0.0900	0.0010	91.22	<0.0001
Size of phylogeny	0.5179	0.0013	383.99	<0.0001

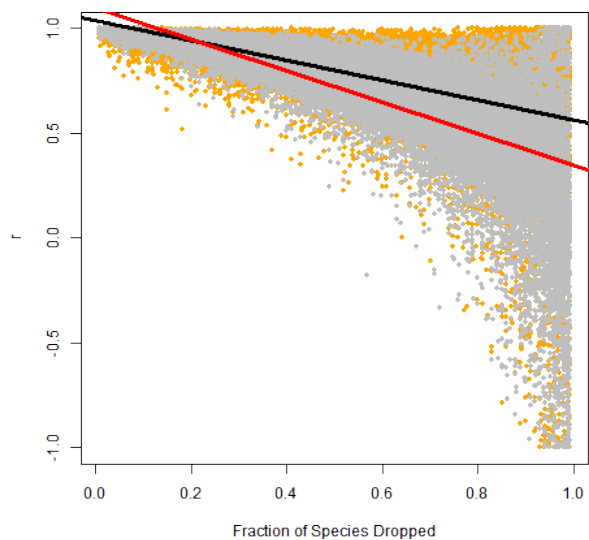
**Table 3: Statistical model of the effect of clade and phylogeny size on ranking error.** Model of the raw data underlying figure 5, regressing the ranking error of imputed species against the number of species in the imputed clade and the entire phylogeny ( $F_{47997,2} = 77890$ ,  $r^2 = 0.7644$ ,  $p < 0.0001$ ). As can be seen in figure 5, the average ranking error is positively correlated with the size of the clade being imputed and the entire phylogeny. Square-root transformations ~~werehave-been~~ applied to both ranking error and size of phylogeny [prior to fitting this model.](#)



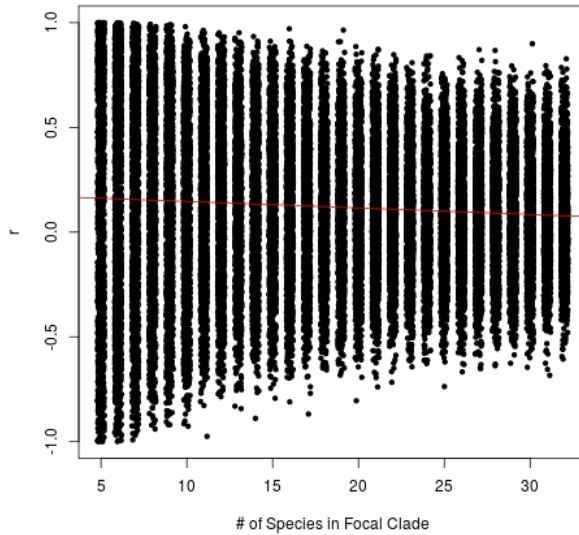
**Figure 1: Conceptual overview of the missing-species simulations in this study.** The simulated tree on the left is the true tree prior to removal of missing species. On the right is the same tree after missing species have been removed. Species that are removed are shown in red. To compare the ED values of the remaining species, we correlate their ED values before (left) and after (right) removal of the missing species. Dashed lines can be seen for the species which would have ED scores compared.



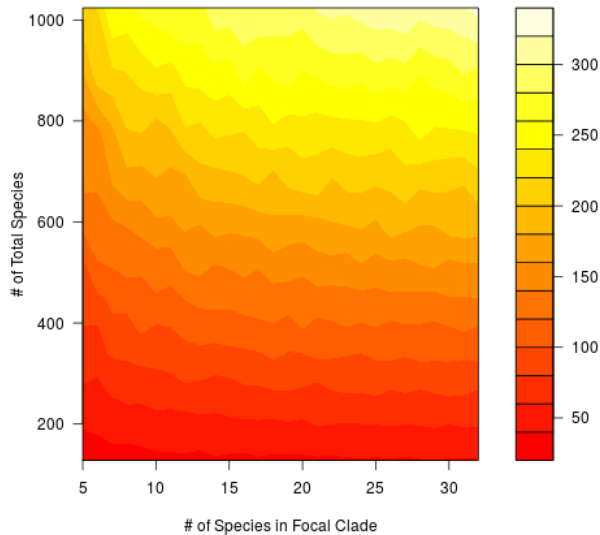
**Figure 2: Conceptual overview of the imputation simulations conducted in this study.** The simulated tree on the left is the ‘true tree’. We selected a clade to treat as ‘missing’ (highlighted with a dashed line and in blue) by treating it as a polytomy (middle panel), and then imputed the ‘missing’ species to produce the imputed clade in the right panel. To compare true and imputed ED values within the imputed clade, we correlated ED values calculated for the true clade (left) with those for the imputed clade (right).



**Figure 3: The effect of missing data on the calculation of the remaining species' ED values.** The correlation coefficient of species' ED values in full (simulated) phylogenies, comparing values before and after the random loss of (other) species from the tree. The color of data points denote whether the species were removed from the phylogeny completely at random (orange) or in a phylogenetically biased fashion (see text; grey). Lines show regressions for random (red) or phylogenetically biased (black) species loss; see table 1 for model coefficients. This plot shows that the accuracy of estimation of ED values is inversely proportional to the number of species missing from the phylogeny, and that phylogenetically-biased species loss has a greater impact on accuracy.



**Figure 4: The correlation between species' imputed and true ED scores plotted as a function of the number of species imputed (focal clade size from all sizes of phylogenies used ( $n = 128, \dots, 1024$ )).** Each data point represents the correlation between ED values within the focal clades where imputation has occurred, comparing species' true ED values with their imputed ED values. This plot, and the statistical analysis of it in table 2, show limited support for an association between true and imputed ED values.



**Figure 5: Mean ranking error of imputed species.** An interpolated heat-map of the mean ranking error of imputed species as a function of the total number of species in the phylogeny (vertical axis) and number of species in the focal (imputed) clade (horizontal axis). Table 3 gives statistical support for the trend of increased error in larger phylogenies and imputed clades. [This figure shows a tendency for an increase in error in larger phylogenies and imputed clades.](#)