



**HAL**  
open science

## Trait-dependent diversification in angiosperms: Patterns, models and data

Andrew J. Helmstetter, Rosana Zenil-Ferguson, Hervé Sauquet, Sarah P Otto,  
Marcos Méndez, Mario Vallejo-Marín, Jürg Schönenberger, Concetta  
Burgarella, Bruce Anderson, Hugo De Boer, et al.

### ► To cite this version:

Andrew J. Helmstetter, Rosana Zenil-Ferguson, Hervé Sauquet, Sarah P Otto, Marcos Méndez, et al..  
Trait-dependent diversification in angiosperms: Patterns, models and data. *Ecology Letters*, 2023, 26  
(4), pp.640-657. 10.1111/ele.14170 . hal-04013691

**HAL Id: hal-04013691**

**<https://hal.science/hal-04013691v1>**

Submitted on 5 May 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

**Article type:** Synthesis

**Running title:** Flowering plant diversification

# Trait-dependent diversification in angiosperms: patterns, models and data

Andrew J. Helmstetter<sup>1,\*</sup>, Rosana Zenil-Ferguson<sup>2</sup>, Hervé Sauquet<sup>3,4</sup>, Sarah P. Otto<sup>5</sup>, Marcos Méndez<sup>6</sup>, Mario Vallejo-Marin<sup>7</sup>, Jürg Schönenberger<sup>8</sup>, Concetta Burgarella<sup>9</sup>, Bruce Anderson<sup>10</sup>, Hugo de Boer<sup>11</sup>, Sylvain Glémin<sup>12</sup>, Jos Käfer<sup>13</sup>

<sup>1</sup>*Fondation pour la Recherche sur la Biodiversité - Centre for the Synthesis and Analysis of Biodiversity, 34000 Montpellier, France. email: andrew.j.helmstetter@gmail.com*, <sup>2</sup>*Department of Biology, University of Kentucky, 101 T.H. Morgan Building, Lexington, KY 40506 USA. email: roszenil@uky.edu*, <sup>3</sup>*National Herbarium of New South Wales, Royal Botanic Gardens and Domain Trust, Sydney, New South Wales, 2000, Australia*, <sup>4</sup>*Evolution and Ecology Research Centre, School of Biological, Earth and Environmental Sciences, University of New South Wales, Sydney, Australia. email: herve.sauquet@gmail.com*, <sup>5</sup>*Department of Zoology, 6270 University Blvd., University of British Columbia, Vancouver BC, V6T 1Z4, Canada. email: otto@zoology.ubc.ca*, <sup>6</sup>*Area of Biodiversity and Conservation, Universidad Rey Juan Carlos, 28933 Móstoles (Madrid), Spain. email: marcos.mendez@urjc.es*, <sup>7</sup>*Department of Ecology and Genetics, Uppsala University, Uppsala, 752 36, Sweden. email: mario.vallejo-marin@ebc.uu.se*, <sup>8</sup>*Department of Botany and Biodiversity Research, University of Vienna, Rennweg 14, 1030 Vienna, Austria. email: juerg.schoenenberger@univie.ac.a*

<sup>9</sup>*Department of Organismal Biology, University of Uppsala, Uppsala, 75236, Sweden. email: concetta.burgarella@gmail.com*, <sup>10</sup>*Department of Botany and Zoology, University of Stellenbosch, Private Bag X1, Matieland 7602, South Africa. email: banderso.bruce@gmail.com*, <sup>11</sup>*Natural History Museum, University of Oslo, 0318 Oslo, Norway. email: h.d.boer@nhm.uio.no*, <sup>12</sup>*CNRS, Ecosystèmes Biodiversité Evolution (Université de Rennes), 35000 Rennes, France. email: sylvain.glemin@univ-rennes1.fr*, <sup>13</sup>*Université de Lyon, Université Lyon 1, CNRS, Laboratoire de Biométrie et Biologie Evolutive UMR 5558, F-69622 Villeurbanne, France; DIADE, Université Montpellier, IRD, CIRAD, Montpellier, France. email: jos.kafer@cnrs.fr*

<sup>9</sup>*Department of Organismal Biology, University of Uppsala, Uppsala, 75236, Sweden. email: concetta.burgarella@gmail.com*, <sup>10</sup>*Department of Botany and Zoology, University of Stellenbosch, Private Bag X1, Matieland 7602, South Africa. email: banderso.bruce@gmail.com*, <sup>11</sup>*Natural History Museum, University of Oslo, 0318 Oslo, Norway. email: h.d.boer@nhm.uio.no*, <sup>12</sup>*CNRS, Ecosystèmes Biodiversité Evolution (Université de Rennes), 35000 Rennes, France. email: sylvain.glemin@univ-rennes1.fr*, <sup>13</sup>*Université de Lyon, Université Lyon 1, CNRS, Laboratoire de Biométrie et Biologie Evolutive UMR 5558, F-69622 Villeurbanne, France; DIADE, Université Montpellier, IRD, CIRAD, Montpellier, France. email: jos.kafer@cnrs.fr*

<sup>10</sup>*Department of Botany and Zoology, University of Stellenbosch, Private Bag X1, Matieland 7602, South Africa. email: banderso.bruce@gmail.com*, <sup>11</sup>*Natural History Museum, University of Oslo, 0318 Oslo, Norway. email: h.d.boer@nhm.uio.no*, <sup>12</sup>*CNRS, Ecosystèmes Biodiversité Evolution (Université de Rennes), 35000 Rennes, France. email: sylvain.glemin@univ-rennes1.fr*, <sup>13</sup>*Université de Lyon, Université Lyon 1, CNRS, Laboratoire de Biométrie et Biologie Evolutive UMR 5558, F-69622 Villeurbanne, France; DIADE, Université Montpellier, IRD, CIRAD, Montpellier, France. email: jos.kafer@cnrs.fr*

<sup>11</sup>*Natural History Museum, University of Oslo, 0318 Oslo, Norway. email: h.d.boer@nhm.uio.no*, <sup>12</sup>*CNRS, Ecosystèmes Biodiversité Evolution (Université de Rennes), 35000 Rennes, France. email: sylvain.glemin@univ-rennes1.fr*, <sup>13</sup>*Université de Lyon, Université Lyon 1, CNRS, Laboratoire de Biométrie et Biologie Evolutive UMR 5558, F-69622 Villeurbanne, France; DIADE, Université Montpellier, IRD, CIRAD, Montpellier, France. email: jos.kafer@cnrs.fr*

<sup>12</sup>*CNRS, Ecosystèmes Biodiversité Evolution (Université de Rennes), 35000 Rennes, France. email: sylvain.glemin@univ-rennes1.fr*, <sup>13</sup>*Université de Lyon, Université Lyon 1, CNRS, Laboratoire de Biométrie et Biologie Evolutive UMR 5558, F-69622 Villeurbanne, France; DIADE, Université Montpellier, IRD, CIRAD, Montpellier, France. email: jos.kafer@cnrs.fr*

<sup>13</sup>*Université de Lyon, Université Lyon 1, CNRS, Laboratoire de Biométrie et Biologie Evolutive UMR 5558, F-69622 Villeurbanne, France; DIADE, Université Montpellier, IRD, CIRAD, Montpellier, France. email: jos.kafer@cnrs.fr*

**Correspondence:** Andrew J. Helmstetter, FRB-CESAB Institut Bouisson Bertrand, 5, rue de l'École de médecine - 34000 MONTPELLIER. Tel: +33 4 11 28 20 58. Email: andrew.j.helmstetter@gmail.com

**Authorship statement:** SG, JK and AJH conceived the study with input from all authors. AJH collated data with assistance from all authors. AJH performed data analyses with input from JK, SG, HS, RZF and SHO. AJH wrote the first draft of the manuscript, and all authors contributed substantially to revisions.

**Data accessibility statement:** Data and code associated with this manuscript are available from the Dryad repository: <https://doi.org/10.5061/dryad.8gtht76t2> and the GitHub repository: <http://github.com/ajhelmstetter/sseReview>.

**Keywords:** Macroevolution, flowering plants, speciation, extinction, BiSSE, trait evolution, phylogenetics

**Abstract word count:** 199 words, **Main text word count:** 7330 words, **Box 1:** 563, **Box 2:** 556, **Box 3:** 477 **No. references:** 88, **No. figures and tables:** 5

## 1 **Abstract**

2       Variation in species richness across the tree of life, accompanied by the incredible variety of  
3 ecological and morphological characteristics found in nature, has inspired many studies to link  
4 traits with species diversification. Angiosperms are a highly diverse group that has fundamentally  
5 shaped life on earth since the Cretaceous, and illustrate how species diversification affects ecosystem  
6 functioning. Numerous traits and processes have been linked to differences in species richness within  
7 this group, but we know little about their relative importance and how they interact. Here, we  
8 synthesized data from 152 studies that used state-dependent speciation and extinction (SSE) models  
9 on angiosperm clades. Intrinsic traits related to reproduction and morphology were often linked to  
10 diversification but a set of universal drivers did not emerge as traits did not have consistent effects  
11 across clades. Importantly, SSE model results were correlated to dataset properties - trees that  
12 were larger, older, or less well-sampled tended to yield trait-dependent outcomes. We compared  
13 these properties to recommendations for SSE model use and provide a set of best practices to follow  
14 when designing studies and reporting results. Finally, we argue that SSE model inferences should  
15 be considered in a larger context incorporating species' ecology, demography and genetics.

## 16 **Introduction**

17 Species diversity is unevenly distributed across the tree of life and while substantial research has  
18 investigated why some clades are more species-rich than others, many fundamental questions remain  
19 unanswered. The causes behind this unevenness can be diverse from catastrophic mass extinctions  
20 that decimate diversity (Raup & Sepkoski, 1982) to key innovations that spur on rapid speciation  
21 (Hunter, 1998) or ecological factors such as competition that shape species co-existence (Drury  
22 et al., 2016; Rabosky, 2013). A greater understanding of the drivers of species diversification can  
23 provide insights into the assembly of communities and their phylogenetic structure, the evolution of  
24 functional traits that underpin a species' role in its environment, the formation of species interaction  
25 networks, and simply how biodiversity has evolved through time (Morlon, 2014). In practice,  
26 phylogenetic diversity metrics can be used as a surrogate for functional diversity (Tucker et al.,  
27 2018), and both traits and diversification history can help predicting ecological processes, such as  
28 extinction threats and naturalization patterns (Schmidt et al., 2021).

29 Traits can themselves affect speciation and extinction because they determine how species  
30 interact with their environment and with other species. Morphological and physiological traits  
31 can affect species' competitive and colonization abilities, defence against predators, herbivores and  
32 parasites; reproductive success, gene flow and adaptation to certain environments. In a wider  
33 sense, extrinsic characteristics, such as species' habitat, climatic zone or geographical area can  
34 also be considered traits (Jablonski, 2008). An important goal of modern macroevolutionary  
35 research is to discover which traits actually cause differences in diversification (through ecological  
36 processes) rather than being linked to species richness purely by common ancestry (phylogenetic  
37 inertia (Felsenstein, 1985)).

38 This research has been boosted over the last decade by the increasing availability of large  
39 phylogenetic trees (Jetz et al., 2012; Rabosky et al., 2018; Smith & Brown, 2018; Upham et al.,

2019) and the continuing development of a range of statistical models to infer patterns in species  
diversification and the drivers behind them (see Box 1). The growing amount of empirical knowledge  
provides an opportunity to synthesise what we know so far about a wide range of ecologically diverse  
and species-rich clades to try to uncover how traits have driven their diversification. Here we focus  
on synthesizing results from state-dependent speciation and extinction (SSE) models (Fig. 1) used  
to investigate trait-dependent diversification (TDD) in flowering plants, or angiosperms. There  
have been more than 150 such studies, providing an opportunity for an updated perspective on the  
role different traits have played in angiosperm diversification (Vamosi et al., 2018). Angiosperms  
form a clade of more than 350,000 extant species, making it by far the most species-rich division  
of plants, even when compared to much older clades (Crepet & Niklas, 2009; Sauquet & Magallón,  
2018). Almost all of terrestrial life is linked, directly or indirectly, to angiosperms (Benton et al.,  
2022) and their success makes them an ideal study group for uncovering the intrinsic traits and  
extrinsic factors driving their diversification.

Previous work has suggested that the origins of angiosperm diversity can neither be tied to  
major global events nor the evolution of a single key innovation (Sauquet & Magallón, 2018; Vamosi  
et al., 2018). Instead various combinations of traits, environment and ecology may have acted to  
stimulate diversification in different groups (Davies et al., 2004; Magallón & Castillo, 2009), creating  
a landscape of macroevolutionary dynamics that varies substantially across different angiosperm  
clades (Magallón et al., 2019). One hypothesis proposes that the traits driving the differences in  
diversification are a range of vegetative and reproductive characteristics, some of which are unique  
to angiosperms (Stebbins, 1974). For example, dioecy originated in 890-5000 independent instances  
(Renner, 2014) but these appear to have led to quite different macroevolutionary dynamics (Käfer  
et al., 2014; Sabath et al., 2016; Wang et al., 2021), probably because of their association with other  
traits (Renner & Ricklefs, 1995; Vamosi et al., 2003). However, we do not know how pervasive

64 such differences are: although most studies agree that angiosperm diversification is driven by a  
65 combination of traits, we know relatively little about which traits are most influential and the  
66 extent to which their effects on diversification are consistent across different clades.

67 In this study we bring together the latest empirical knowledge on angiosperm diversification to  
68 compare the effects of traits in different contexts and identify those that have repeatedly stimulated  
69 diversification. Limiting our analysis to SSE approaches allows us to directly compare results  
70 across studies in a standardized framework, but does leave scope for the influence of known SSE  
71 model shortcomings (Davis et al., 2013; Maddison & FitzJohn, 2015; Rabosky & Goldberg, 2015).  
72 As a result, we also investigate the relationships between dataset properties (e.g. tree size and  
73 sampling fraction), methodological innovation (e.g. hidden states) and the results of published  
74 studies, highlighting how biases in our use of SSE models can affect our search for general trends.  
75 To help mitigate these effects we provide a set of best practices including recommendations for  
76 result-reporting. Finally, we identify gaps in our current knowledge and suggest how to better  
77 incorporate ecology into the study of trait-dependent diversification.

## 78 **Materials & Methods**

### 79 **Data collection**

80 We collected published studies that cited SSE methods papers (Beaulieu & O’Meara, 2016;  
81 Caetano et al., 2018; FitzJohn, 2010, 2012; Freyman & Höhna, 2018; Goldberg & Igić, 2012;  
82 Goldberg et al., 2011; Herrera-Alsina et al., 2019; Maddison et al., 2007; Magnuson-Ford & Otto,  
83 2012; Nakov et al., 2018; Rabosky & Goldberg, 2017; Verboom et al., 2020) using Google Scholar,  
84 last accessed 18th May 2021. To facilitate data collection from papers using SSE models, we  
85 developed an R package called ‘papieRmache’ (<https://github.com/ajhelmstetter/papieRmache>).

86 This package has two main purposes (1) to classify papers into different categories based on the  
87 frequency of term use in the text and (2) to pull out sections of the main text that contain a keyword  
88 or a pair of keywords while highlighting relevant information. We identified the SSE studies on  
89 angiosperms by using the keywords ‘angiosperm’, ‘flowering’ and ‘plant’ subsequently validating  
90 the subset of papers by hand and removing any studies on groups other than angiosperms. We  
91 then collected up to 31 different properties from each paper relating to the trait investigated, the  
92 group studied, the phylogenetic tree and the outcome of the SSE model used (see appendix S1),  
93 incorporating previously collated information on these properties where available (Onstein, 2019).  
94 In cases where there was uncertainty in how to interpret or collect data from a study we contacted  
95 the authors for their assistance and clarification, where possible.

## 96 **Trait classification**

97 While some sets of character states were the same among studies (e.g. annual vs perennial;  
98 diploid vs polyploid), many of them did not overlap. The nature of a trait can be difficult to  
99 define so here we use the SSE convention that considers that traits or characters (e.g. pollination)  
100 have two or more states (e.g. biotic pollination, abiotic pollination, hummingbird pollination,  
101 etc). We classified traits into different categories to facilitate comparisons among different trait  
102 types. At the broadest classification these were intrinsic (traits belonging to the species), extrinsic  
103 (environmental or geographic traits), interaction (traits explicitly related to other species), and  
104 combination (multiple traits belonging to different categories that were grouped, e.g. species that  
105 have small fruits and are found on islands). To allow for analyses at different grouping levels we  
106 developed a trait ontology (Table S1) starting at the broadest level as just described and becoming  
107 more specific, up to level six.

## 108 Data analysis

109 To examine the effect of particular traits on diversification we used the trait categories defined  
110 above and calculated the proportion of models in which TDD was inferred. Studies often consider  
111 the effect of a single trait of interest in several different clades, or the effect of several traits on  
112 diversification in a single clade of interest. We considered each combination of phylogenetic tree,  
113 SSE model type and trait separately here with an outcome of 1 (TDD detected) or 0 (no effect of  
114 trait detected). Whether or not TDD was detected was typically based on significance in model  
115 comparisons and/or posterior distributions of rates among states. However, if significance was not  
116 inferred or reported, we tried to follow the study narrative and statements made in the text. If  
117 model comparisons were conducted and reported, only the best-fitting model was considered, unless  
118 other models were explicitly referred to in the study. We examined patterns at different levels of  
119 trait categorization and among different model types. To facilitate comparison, we mainly consider  
120 whether or not a trait has an effect on diversification, irrespective of the direction of the effect, as  
121 the direction is only defined at the state level. When both the effect and the absence of an effect of  
122 a trait on diversification have been inferred in different models, we consider the effect of the trait  
123 ‘inconsistent’. Note that inconsistency does not necessarily imply a contradiction, as the differences  
124 in the results of SSE models might be caused by differences in statistical power, type I errors, or  
125 biological differences between clades, as we will discuss.

126 Where possible, we extracted net diversification rates (per lineage per million years) associated  
127 with each character state. Then, relative differences in net diversification rates were calculated as  
128  $(r_{max} - r_{min})/r_{max}$  (where  $r_{max}$  and  $r_{min}$  are the net diversification rates of the states with the  
129 highest and lowest rates respectively). These were used to represent the magnitude of the effect of  
130 a given trait on diversification, while taking into account general variation in diversification rates  
131 among clades. Comparisons were then made across the broadest trait categories, levels 1 and 2. For



132 ease of interpretation, these analyses were restricted to those models where all net diversification  
133 rates were positive. At a narrower scale, we identified eight traits for which there was enough  
134 replication to be able to assess whether one character state was consistently inferred to have higher  
135 net diversification rates than the other. We ensured these traits had been tested at least five times,  
136 in at least two different studies and two different clades.

137 We examined the relationship between SSE model inferences and several dataset properties:  
138 number of tips in the tree, root age, number of genetic markers used to build the tree, sampling  
139 fraction (here referring to global sampling fraction unless stated otherwise) and tip bias (here  
140 calculated as the number of tips with the most common state divided by the number of tips with  
141 rarest state). For each of these we constructed two density plots representing the distributions of  
142 values in cases where TDD was, and was not, inferred and compared the overlap between densities.  
143 We also fitted generalized additive models (GAM, Hastie and Tibshirani, 2017) to the continuous  
144 dataset properties with the SSE model result as a binary response variable (TDD vs no effect).  
145 Variables were log-transformed (or arcsine in the case of sampling fraction) to conform better to a  
146 normal distribution. Initially, we constructed a GAM using all five variables and assigned the mean  
147 of the known values to any missing values. We also assessed each variable individually to determine  
148 the shape of each relationship when examined in isolation. In all cases we used smoothing functions  
149 (cubic regression splines) and the dimension of the basis used to represent the smooth term was  
150 set to  $k = 5$  to avoid overfitting.

### 151 **Predicting results based on dataset properties**

152 After collecting information from all studies, we found that the outcome of the SSE model,  
153 that is, whether TDD was inferred or not, was sometimes associated with the dataset properties.  
154 We therefore attempted to predict SSE model results (inference of TDD vs no effect) from dataset

155 properties alone, and identify those properties with the largest predictive power. We used all  
156 available dataset properties except for highly-specific categorical variables (e.g. trait levels 5-6,  
157 clade, family) and those that varied among different states (putative root state, sampling per state,  
158 samples per state). We used a machine learning approach, extreme gradient boosting, with the  
159 R package ‘xgboost’ (Chen & Guestrin, 2016), a supervised learning approach based on gradient  
160 boosting machines. It uses a labelled dataset (the outcome is known) and an ensemble of weak  
161 prediction models (e.g. decision trees) whereby new models are added on to existing models per  
162 iteration to minimize error. Categorical variables with more than two categories were converted into  
163 binary, dummy variables with one-hot encoding (i.e. categorical features in a column are converted  
164 into binary variables) to facilitate model building. We trained models on a random selection of 80%  
165 of our dataset and tested them on the other 20%. After a parameter optimisation step we repeated  
166 this process 500 times to produce a range of accuracy values, the percentage of cases where the real  
167 outcome matched the classification, to account for stochasticity in the test and training datasets.  
168 For each iteration we also recovered the relative importance of each variable, which allowed us to  
169 determine which dataset properties had the most influence on average. This method is specifically  
170 useful when there are many inter-dependent variables, i.e. when the inference of TDD depends  
171 on certain combinations of variables rather than major determinants. Given the inter-dependency  
172 of variables across the different decision trees, it is difficult to uncover whether a given property  
173 generally leads to TDD or not with xgboost. We avoid interpreting the results in this way, focusing  
174 on how accurate prediction can be and the types of variables that are most important to the model’s  
175 predictive ability.

## 176 **Results**

### 177 **Traits studied and their effects on diversification**

178 We collated information on trait-based diversification from 152 studies using a total of 629 SSE  
179 models to study angiosperm diversification. We found that 127 studies were conducted on a single  
180 clade, the rest examined diversification patterns across multiple, independent clades that were  
181 analysed separately. Variation in breadth of different traits investigated was also observed within  
182 studies. In total, 92 studies examined just a single trait (i.e. one level 6 category, see Table S1),  
183 while 40 studies looked at diversification patterns in sets of traits that belonged to more than one  
184 trait category at the highest level (e.g. extrinsic and intrinsic traits). In terms of taxonomic level,  
185 SSE models were most often run on genera or families (Fig. S1) and study clades were relatively  
186 evenly-distributed across the angiosperm tree of life (Fig. S2). Studies have focused on 36 out of  
187 64 angiosperm orders, and 81 out of 416 families. As expected, diversification interest is generally  
188 proportional to the amount of species diversity in different parts of the angiosperm tree of life.  
189 There was a clear, positive correlation between the number of species in a clade (order, family) and  
190 the number of TDD studies applied to species within the clade (Fig. S3, S4).

191 At the highest level of classification, intrinsic traits (i.e. those belonging to the plant species  
192 itself) were tested more often (291 models or 46% of models run) than extrinsic traits (i.e. those  
193 related to the species' habitat and geography, 260 models or 41%). Researchers tended to study  
194 intrinsic traits relating to reproduction (e.g. flower morphology, fruit morphology, breeding system),  
195 traits related to species' biogeography (e.g. biome, geographic range) and vegetative traits (e.g. life  
196 form, leaf morphology), investigating less often those traits related to the genome (e.g. ploidy level)  
197 or species interactions (e.g. symbiosis, defense or dispersal) (Fig. 2). In general, intrinsic traits  
198 were found to be associated with diversification slightly more often than extrinsic traits (59.8% vs

199 52.7%).

200 Our collation of data showed that results inferred with SSE models were inconsistent at both  
201 broad and narrow scales (grey vs coloured portions of bars in Fig. 2). For example, traits such as  
202 lifespan (Azani et al., 2019; Drummond et al., 2012; Salariato et al., 2016; Soltis et al., 2013) and  
203 ploidy level (Folk & Freudenstein, 2014; Han et al., 2020; Landis et al., 2018; Zenil-Ferguson et al.,  
204 2019) yielded different results depending on the angiosperm group studied. Among broad trait  
205 categories (level 2; see Table S1) that have been tested using >25 models, vegetative traits yielded  
206 TDD more often than any other trait type, while pollination traits had the lowest proportion of  
207 TDD (Fig. 2a).

208 Given that we were assessing a range of SSE models, we wondered if using different models  
209 led to different patterns in TDD inference. Of those SSE models used frequently (> 30 times),  
210 MuSSE, GeoSSE and QuaSSE yielded TDD at high relative frequencies (> 60%) while BiSSE and  
211 HiSSE detected TDD less often (~ 50%; Fig. S5). Models with hidden states will decrease the  
212 proportion of false positives so may provide more robust general trends. We compared results  
213 from only hidden-state models to those from the entire dataset and found that the proportion of  
214 trait-dependent outcomes decreased notably for vegetative, habitat and biogeography traits (Fig.  
215 S6). Outcomes for hidden state models remained inconsistent within categories that were tested  
216 five times or more, reflecting patterns in the entire dataset. We then took a closer look at results  
217 from HiSSE and BiSSE for three commonly-tested, narrow trait categories (level 4) and found  
218 mixed patterns. TDD was detected by BiSSE and HiSSE at about the same frequency in traits  
219 related to flower morphology (50% vs 42%), BiSSE detected TDD more often in traits related to  
220 life form (70% vs 25%) and HiSSE detected TDD more often in traits related to breeding system  
221 (75% vs 46%).

222 We examined eight traits in detail, three of which (lifespan, sexual system and woodiness)

223 rarely exhibited TDD while in the other five (epiphytism, biome, ploidy, photosynthesis and  
224 self-compatibility) TDD was often detected (Fig. S7). We found conflict in which states increased  
225 diversification among different models in all traits except epiphyte form, photosynthesis and self-compatibility.  
226 Examining the absolute net diversification rates among states of seven traits (sexual system was  
227 not assessed as most rates were relative, i.e. not time-calibrated) we found that patterns across  
228 clades reflected those detailed above (Fig. S8). For example, net diversification rates in traits  
229 rarely associated with diversification (e.g. woodiness or lifespan) were generally similar among the  
230 different states. To understand the effect of major trait categories (levels 1 and 2) on diversification  
231 we plotted the distribution of relative differences in net diversification rates for 208 models from  
232 78 studies, grouping rates by trait category (Fig. S9). Generally we found that there was a wide  
233 range of relative differences in each trait category but no significant differences among categories.

## 234 **The evolution of SSE model use and methodological innovation**

235 As SSE models themselves have diversified, the relative frequency of model use has evolved.  
236 We collated data on the types of SSE model used in each study, and plotted their use by year  
237 of publication (Fig. 3). BiSSE has remained popular even as newer, more complex models have  
238 emerged. Models with multiple states, predominantly MuSSE, have also been commonly used  
239 showing that researchers are interested in the effects of more complex traits or trait groups with  
240 more than two states. There has also been a consistent focus on using SSE approaches related to  
241 geography in models like GeoSSE and GeoHiSSE. When examining the number of studies that use  
242 SSE models each year we find a rapid increase since the first study we found that used of BiSSE on  
243 angiosperms in 2009 until a conspicuous slowdown and slight drop in 2015 (Fig. 3). This appears to  
244 coincide with the publication of a number of influential papers that criticised the propensity of SSE  
245 methods for false positives (Maddison & FitzJohn, 2015; Rabosky & Goldberg, 2015) and pointed

246 out power limitations (Davis et al., 2013). After this, SSE model use continued with a greater  
247 variety of models owing to the development of models with hidden states (Beaulieu & O’Meara,  
248 2016), which have since spread to all aspects of SSE model use (Box 1; Fig. 1) becoming the  
249 dominant set of models by 2020 (Fig. 3).

## 250 **Dataset properties are highly influential**

251 The input data for macroevolutionary studies have grown in size and quality, in parallel with the  
252 innovations in the SSE models. For example, we found evidence that trees used with SSE models  
253 have gradually grown larger over time (Fig. S10). We then examined the relationship between tree  
254 size and whether or not TDD was inferred, regardless of the trait investigated. We found that,  
255 in general, TDD was detected less often when trees had smaller numbers of tips (Fig. 4a, S11a).  
256 The number of tips in a tree is important for robustness of SSE model results and guidelines for  
257 adequate power were put forward by Davis et al. (2013) who suggested that results from models  
258 using trees with fewer than 300 tips should be treated with caution. But has this recommendation  
259 shaped SSE model use? The proportion of models run on trees with fewer than 300 tips was initially  
260 very high (94% of 139 models for which we found information) in studies published up until 2013.  
261 It then decreased to 57% (277 of 482 models) in studies published from 2014 onwards. Despite this  
262 reduction, more than 60 models were run on trees with fewer than 50 tips after 2013.

263 Tree size and root age are closely linked because trees with larger numbers of tips are generally  
264 older (Fig. S12). Indeed, we found that TDD was detected more often when trees with an older  
265 root age were used (Fig. 4b, S11b). Regardless of their size or age, trees that more accurately  
266 represent the true phylogeny of a group will allow us to more reliably estimate its diversification  
267 history. We used information on the total number of molecular markers as a proxy for tree quality.  
268 We found a difference in the distributions indicating that models with trait-dependent outcomes

269 usually had better quality trees than those that did not (Fig. 4c, S11c).

270 Another issue that has been repeatedly brought up in simulation studies is the potential effect  
271 of inflated tip bias (Davis et al., 2013; Maddison et al., 2007). Tip bias increases when there is a  
272 higher frequency of one state than the others across the tips of the tree. Upon examining the data  
273 used with SSE models we found substantial overlap between densities (Fig. 4d) except for extreme  
274 values of tip bias where SSE models tended to find no effect of the trait studied (Fig. S11d). Tip  
275 ratio bias recommendations were also made by Davis et al. (2013), who cast doubt on inferences  
276 made when the ratio exceeds 10:1. Prior to 2014, 85% of SSE models (56 of 66) had suitable tip  
277 ratios and this figure remained similar (87%, 314 of 360) for the studies that came after.

278 Global sampling fraction is the proportion of known species that are present in the tree. If  
279 the sampling fraction is low it can drastically affect diversification rate estimation (Chang et al.,  
280 2020; FitzJohn et al., 2009; Sun et al., 2020). The sampling fraction was explicitly modeled in  
281 SSE methods by FitzJohn et al. (2009), who recommended that the sampling fraction should be  
282 at least 25% to adequately capture diversification dynamics. Approximately 62% of 616 models  
283 had sampling fractions greater or equal to 25%. Sampling fractions were initially high but became  
284 more varied over time (Fig. S13), which probably reflects easing of assumptions on complete species  
285 sampling, but also indicates that high levels of incomplete sampling are not uncommon in recent  
286 literature.

287 Furthermore, we found that models that used trees in which sampling fraction was low ( $< 50\%$ )  
288 generally yielded TDD (Fig. 4e, Fig. S11e). Conversely, high sampling fraction was more often  
289 associated with a lack of TDD. Given that the inference of TDD varies with tree size (Fig. 4a),  
290 we wondered whether there might also be a relationship between sampling fraction and tree size.  
291 However, upon examination we found only a weak, negative trend where trees with more tips  
292 had slightly lower sampling fractions (Fig. S14). We then looked at the relationship between

293 sampling fraction and the total number of species in the clade and found a steeper, stronger  
294 negative relationship (Fig. S15) meaning that the larger the clade is, the less well-sampled it tends  
295 to be. Datasets of small clades with low sampling fraction generally do not exist (as they should  
296 not be studied) and large clades with high sampling are currently very rare (Figure S15). These  
297 negative trends remain similar regardless of whether TDD is inferred or not.

298 Better quality data will lead to more robust results, so we compared trends taken from all  
299 studies, irrespective of dataset quality, to those from medium quality or better (at least 10% taxon  
300 sampling, 100 tips and a maximum tip bias ratio of 20:1) and high-quality datasets only (at least  
301 25% taxon sampling, 300 tips and a maximum tip bias ratio of 10:1; Davis et al., 2013; FitzJohn  
302 et al., 2009). We found that generally as data quality improves then the proportion of BiSSE models  
303 with TDD increases substantially (Fig. S16). Conversely, the proportion of HiSSE models with  
304 TDD remains relatively stable regardless of dataset quality. As HiSSE was conceived to leverage  
305 the issue of false positive rates in BiSSE, this could indicate that the higher TDD rates in BiSSE  
306 and MuSSE are due to false positives. However, we found a relatively low number of studies (20)  
307 with datasets that meet all recommendations, so more examples with high quality data are needed  
308 to determine whether these trends are robust. We then examined whether dataset quality affected  
309 general trends for major trait types (level 2 category). We found that increasing dataset quality  
310 had little effect on extrinsic and interaction traits (e.g. biogeography, habitat or pollination) while  
311 frequency of TDD for some intrinsic traits such as those related to reproduction tended to increase  
312 slightly with dataset quality (Fig. S17), but again the number of studies was low.

### 313 **How predictable is the inference of trait-dependent diversification?**

314 Empirical results in angiosperms clearly exhibit strong relationships between various dataset  
315 properties and whether TDD is inferred by the SSE model. To assess the importance of the



316 continuous dataset properties together we fit a GAM including the number of tips, root age, the  
317 number of markers, the sampling fraction and the tip bias (Fig. S18). We found that all variables  
318 except tree age were significant when predicting SSE model outcome ( $r^2 = 0.211$ , see Table S2 for  
319 full details).

320 If we had comprehensive information about studies and their data, including dataset properties,  
321 taxonomy and traits, could we predict whether TDD would be inferred? Using a machine learning  
322 approach, extreme gradient boosting (Chen & Guestrin, 2016), we were able to correctly predict,  
323 with approximately 71% accuracy (56-81%, Fig. S19), whether SSE models would infer TDD. The  
324 most important features were the information-dense, continuous variables (e.g. number of tips,  
325 tip bias, sampling fraction; Fig. 5), further reinforcing earlier observations about their potential  
326 influence on SSE model outcomes (Fig. 4, S11, Table S2). Generally, categorical variables, such as  
327 those related to the trait studied (e.g. biogeography or fruit morphology) or SSE model used (e.g.  
328 BiSSE or HiSSE), played a smaller role in the model's predictive ability (Fig. 5).

## 329 Discussion

### 330 No consistent drivers of angiosperm diversification

331 Previous broad-scale empirical studies (Davies et al., 2004; Hernández-Hernández & Wiens,  
332 2020; Magallón & Castillo, 2009) and reviews (de Queiroz, 2002; Donoghue, 2005; Donoghue  
333 & Sanderson, 2015) have proposed that diversity in different angiosperm groups may have been  
334 caused by various combinations of intrinsic and extrinsic traits. In our synthesis of studies using  
335 SSE models, we found that while there certainly are particular traits correlated with higher  
336 diversification rates, general trends are mostly limited. It may be unrealistic to expect that all  
337 angiosperm diversification has been driven by similar responses to a handful of traits but it is

338 useful and interesting to investigate the extent to which such traits exist. Indeed, our compilation  
339 of the results of 152 studies on TDD in angiosperms indicates that the factors driving angiosperm  
340 diversification are more complex than a set of universal drivers. When we compared studies  
341 investigating the same traits we found that conclusions generally differed, with some indicating  
342 that the trait does have an effect on diversification and others concluding there is no effect. When  
343 only considering results from models with hidden states, which should be more reliable, we found  
344 that the proportion of TDD changed substantially for some traits, though inconsistency was still  
345 common (Figs. 2, S6). As more studies with hidden states are conducted, we will find out whether  
346 these trends are general. We note that the inconsistency observed might reflect real trends in the  
347 data, or be due to dataset properties (lack of power, model mis-specification). In the following,  
348 we will first discuss the biological conclusions of our study, before considering dataset and model  
349 properties.

350 Our analyses and results centered around how traits (e.g. pollination-related traits) rather  
351 than the character states of these traits (e.g. bee vs bird pollination) affect diversification. After  
352 grouping traits into several levels of categories (Table S1), we found that some types of traits were  
353 more often found to affect angiosperm diversification than others. It may come as no surprise that  
354 floral traits are among the most investigated and influential (Fig. 2). Indeed, the flower contains  
355 the organs needed for sexual reproduction, making it central to the biology, ecology and evolution  
356 of angiosperms, and flower characteristics certainly have a large role in determining differences in  
357 diversification (Vamosi et al., 2018). In particular, reproductive systems are highly variable in  
358 angiosperms (Barrett, 2013) and are again thought to be closely linked to their success (Barrett  
359 et al., 1996). Results from SSE models lend some support to this idea - for example, TDD was  
360 commonly inferred when mating system traits were investigated (e.g. self-compatibility; Fig. S7).  
361 Even so, we found that breeding system (the higher level trait classification including all aspects of

362 mating and sexual systems) did not yield TDD as often, due to variability in the effects of sexual  
363 systems. Other intrinsic traits including those related to photosynthesis, non-floral morphology  
364 and the genome have received less attention than traits related to reproduction (Fig. 2). However,  
365 they were regularly found to be associated with TDD in the limited number of studies in which  
366 they were assessed.

367 Unfortunately, we could say little about which state was generally advantageous for a given  
368 trait because of a lack of overlap among states across the 152 studies. Nevertheless, we were able to  
369 examine how particular states affected diversification in eight traits. Five of these demonstrated how  
370 different states of the same trait (e.g. woody and herbaceous) can increase diversification in different  
371 groups (Fig. S7) while just three traits showed patterns where one state was consistently associated  
372 with elevated diversification rates (epiphytism, non-C3 photosynthesis and self-incompatibility).  
373 However, it is difficult to say whether these have truly consistent effects on diversification as they  
374 have only been investigated a handful of times in a relatively small proportion of angiosperm species.

375 Despite some general patterns in those traits that are more often influencing angiosperm  
376 diversification, the overarching trend is that the effect of a trait on diversification is clade-dependent.  
377 Therefore, the main question remains open: what drives differences in diversification among angiosperms?  
378 The fact that a definitive answer has yet to be found suggests that it is the complex interplay  
379 between trait evolution, biotic interactions and geography that matters. For example, (Hernández-Hernández  
380 & Wiens, 2020) identified geography (range size, biome) as an important factor but they could not  
381 resolve whether this is a cause or a consequence of differences in diversification. Others have  
382 suggested that it is not the presence or absence of a trait that determines the evolutionary success  
383 of a clade, but rather the capacity to change (Onstein, 2019; Ricklefs & Renner, 1994). This could  
384 partly explain the inconsistency we observe, but again, trait diversity could be both a cause and a  
385 consequence of species richness. Furthermore, the choice of clades and traits, as well as the quality

386 of the input data, also influence whether or not differences in diversification are detected, and  
387 therefore our conclusions.

### 388 **The importance of evolutionary history and ecological context**

389 Users of models of trait-based diversification face an important challenge - choosing the context  
390 in which to conduct analyses. In the simplest scenario, where a trait only evolved once in the study  
391 clade, its effect on diversification cannot be tested (Maddison & FitzJohn, 2015) so this should  
392 be avoided. At the intermediate scale a trait may have evolved multiple times in closely-related  
393 clades but their shared evolutionary history makes them much more similar than distantly-related  
394 ones. So, associations between states and rates cannot be interpreted as a general pattern in this  
395 limited phylogenetic scope. Broadening the scope of the analysis, by way of either a larger tree, or  
396 multiple trees in a meta-analytic framework (Sabath et al., 2016) can help to reveal these general  
397 patterns but leads to different challenges. At larger phylogenetic scales, trees with many taxa are  
398 more likely to contain a range of branching patterns where lineage accumulation is faster in some  
399 parts of the tree than in others. In older clades there has been more time for macroevolutionary  
400 processes to have an impact on the trees we infer and the distribution of traits we observe today.  
401 As we found, larger, older trees more commonly yield TDD (Fig. 4).

402 Yet at the same time, these trees often contain considerable variation in traits. Many of these  
403 traits could have effects on diversification, either contrasting or additive. Thus, the diversification  
404 rate of a clade with a certain trait also depends on the presence or absence of other traits. This leads  
405 to considerable heterogeneity in large trees that makes analyses more susceptible to false-positive  
406 errors caused by factors other than the focal trait (other traits, the environment, time) influencing  
407 diversification (Maddison & FitzJohn, 2015; Rabosky & Goldberg, 2015). While hidden-state  
408 models help to mitigate this issue, they still assume that any hidden states are categorical and

409 have constant transition rates, which does not capture all sources of heterogeneity. Furthermore,  
410 they cannot handle all cases of possibly misleading inferences, e.g. when the state impacting  
411 diversification has only evolved once or a handful of times (Beaulieu & O'Meara, 2016; Maddison  
412 & FitzJohn, 2015). These issues make the choice of study clade and how to apply SSE models  
413 to it an important consideration (Box 2). One way to deal with the oversimplification of reality  
414 inherent to SSE models (reducing diversification rate variation to a few categories) is to adopt a  
415 model averaging procedure as suggested by Caetano et al. (2018). Diversification rate estimates  
416 are calculated taking model uncertainty into account, and thus allow a richer interpretation of  
417 rate differences than simply whether or not there is a difference in the best fitting model. The  
418 estimation of branch-specific diversification rates (Maliet et al., 2019) would be another avenue to  
419 explore, but to our knowledge this has not yet been adapted to explicitly test for trait-dependent  
420 diversification.

421 It is not only the evolutionary history of the species in the study clade and their traits that is  
422 important, but also how these species interact with their environment, and each other. The effect  
423 of each trait on diversification should be considered in its own ecological context - a trait may  
424 stimulate diversification but does it have the same effect in a temperate vs tropical environment?  
425 This ecological context can tell us more about why a trait affects diversification but is insufficiently  
426 taken into account as many studies test single traits, are limited to two traits (e.g Zenil-Ferguson  
427 et al., 2019), or use hidden states that are difficult to interpret biologically. Thus, to test in what  
428 contexts a trait affects diversification, several traits (intrinsic and extrinsic) must be assessed in  
429 tandem. Multi-state approaches (ideally including hidden states (e.g. Herrera-Alsina et al., 2019))  
430 could be useful for testing for the effect of ecological context (e.g. Onstein et al. (2017)). However,  
431 current models only allow the investigation of a small number of traits simultaneously before models  
432 become intractable due to the large number of parameters. This could be somewhat mitigated by

433 careful choice on which parameters, notably transition rates, are allowed to vary, which could also  
434 help with biological interpretation if unlikely and impossible transitions between states are limited.  
435 An alternative approach was used by O’Meara et al. (2016) who tested the effect of six binary traits  
436 by fitting binary models for all trait combinations separately.

### 437 **Best practices for SSE model use**

438        Though a number of recommendations have been made for accurate inference with SSE models,  
439 most empirical studies do not meet them. When using the strict thresholds suggested in the  
440 literature we found that just 20 studies (out of 127 for which we retrieved information about each  
441 property) involve datasets that meet all strict criteria. We also found that, in some cases, meeting  
442 recommendations can affect how often TDD is detected (Fig. S16) and general patterns when  
443 trying to determine which traits have been important for angiosperm diversification (Fig. S17) so  
444 we encourage users to meet recommendations when possible and to clearly report the properties of  
445 their dataset (Box 3). The apparent relationship between sampling fraction and inference of TDD  
446 (Fig. 4e) should invite us to be cautious when sampling fractions are low, particularly as better  
447 sampled trees yield more accurate estimates of diversification rates (Chang et al., 2020; FitzJohn  
448 et al., 2009). One explanation may be that as sampling fraction goes down, rate estimates become  
449 more reliant on the model rather than the data, though further work is needed to understand why  
450 this might favour TDD. Publication bias may also be playing a role. If no TDD is detected in a  
451 poorly sampled clade this may be attributed to a lack of power that ultimately prevents publication  
452 thereby inflating the number of studies with low sampling that detect TDD.

453        To clarify these observations, simulation studies should be undertaken to investigate the influence  
454 of dataset properties (e.g. sampling fraction) and model inadequacy on SSE model inference  
455 simultaneously. While false-positive error rates have been explored for HiSSE (Caetano et al.,

2018; Rabosky & Goldberg, 2017), we still know little about how different dataset properties can affect the model's inferences. A study similar to Davis et al. (2013) for HiSSE would be beneficial going forward, perhaps being extended to trees with thousands of tips to match the growing size of datasets. For the moment, the guidelines put forward for BiSSE are probably appropriate starting points to ensure that models are robust, though these may change in the future as the interaction between dataset properties and hidden states is better understood. In the light of these observations, we present recommendations for reporting dataset properties and model results in Box 3, which will help researchers assess the robustness of results, identify their possible limitations and better compare them to other studies. We have focused here on SSE models as they have been applied many times and have themselves been extensively studied. However, our recommendations apply to other diversification models as well because they typically use the same underlying birth-death model, and criticisms such as those in Maddison and FitzJohn (2015) apply to evolutionary correlation analyses in general, not just SSE models.

### **Knowledge gaps and future avenues**

Our review allowed us to identify clades that are understudied and therefore good focal points for future research to gain a more well-rounded picture of angiosperm macroevolutionary dynamics. One of the most obvious is Asteraceae, species-rich yet subject to relatively few trait-based diversification studies (Fig. S4), or Alismatales, an order that has more than 4,500 species (Fig. S3) but just a single study on their trait-based diversification (Canal et al., 2019). In addition, some families with more than 1,000 species, such as Phyllanthaceae or Orobanchaceae have yet to be studied in this way.

High-quality phylogenetic trees are not the only ingredient for SSE studies; trait data also need to be available. We highlight understudied traits related to lifespan, dispersal and symbiosis as

479 ripe avenues for future work that have potential to unearth important patterns in TDD. However,  
480 apart from a few traits such as geographical range or climatic preferences, gathering high-quality  
481 data for several traits in large numbers of species is a time-consuming activity. We encourage the  
482 integration of trait data generated from SSE studies (and others) into large, global trait databases  
483 such as eFLOWER (Sauquet et al., 2017), TRY (Kattge et al., 2020) or more focused databases  
484 (e.g. AusTraits (Falster et al., 2021)). These will act as important resources as researchers consider  
485 several traits in tandem when testing for context-dependent effects of traits. Most importantly,  
486 studies should be conducted on traits where clear hypotheses can be generated about their effect  
487 on diversification in the chosen study clade.

488 SSE methods are statistical tools that are used to uncover correlations, and cannot themselves  
489 discover causal relationships. By definition, macroevolutionary models try to capture the result of  
490 many aggregated small-scale processes in a few high-level parameters. Speciation is an instantaneous  
491 split of one branch into two in most macroevolutionary models, although in reality there might be a  
492 wide range of different dynamics depending on environmental heterogeneity, biotic interactions, and  
493 intrinsic traits (e.g. breeding systems, genomic incompatibilities) (Coyne & Orr, 2004). Thus, if a  
494 trait is predicted to affect speciation and extinction, high-quality inferences of diversification rates,  
495 for which our synthesis provides some guidelines, should be able to detect any signal. However, this  
496 signal is only a piece of the puzzle, and the various underlying ecological and genetic processes can  
497 also be put to the test. For example, Park et al. (2018) compared sister species with contrasted  
498 mating systems (selfing vs. outcrossing) and showed that niche breadth tended to decline over time  
499 in selfing lineages, in agreement with the dead-end scenario proposed for this trait and detected in  
500 macroevolutionary analyses (Freyman & Höhna, 2019; Goldberg & Igić, 2012). Additionally, we  
501 can identify traits that have an effect on ecological and genetic mechanisms that control speciation  
502 and extinction, such as traits affecting coexistence and niche partitioning (Adler et al., 2013) (e.g.



503 specific leaf area or seed mass), genetic differentiation between populations or species (Gamba  
504 & Muchhala, 2020) (e.g. pollination mode, mating system, growth form) or those associated with  
505 commonness and rarity (Murray et al., 2002) (e.g. seed production). Such traits come with *a priori*  
506 hypotheses and could be ideal candidates for macroevolutionary studies exploring their effect on  
507 diversification.

508 Furthermore, in SSE methods, information about interactions between species is taken into  
509 account via the traits that species possess. There is no direct consideration of ecological context  
510 (Morlon, 2014), i.e., whether species actually interact, or whether the diversification of a clade  
511 in a certain area is influenced by the diversity of other clades that are present there. Recently,  
512 an integrated model including trait evolution, species interaction (competition) and diversification  
513 has been studied in the context of evolutionary radiation (Aristide & Morlon, 2019) and showed  
514 that competition can partly decouple trait diversity and species richness. The effect of competition  
515 and diversity-dependence in SSE models has also recently been analysed theoretically (Kaj et al.,  
516 2021). However, inference tools based on these more integrated models still remain to be developed.  
517 This would require integrating multiple types of data, such as species' distributions and ecological  
518 preferences, possibly from different kingdoms (plants, animals) and it requires closer collaboration  
519 between modellers, statisticians, data analysts and specialists of the studied taxa.

## 520 **Conclusions**

521 When bringing together more than a decade of study on TDD in angiosperms, it is the  
522 inconsistent effects of traits that stand out, rather than the importance of a particular set of  
523 universal drivers. The absence of consistent effects highlights the important role the evolutionary  
524 history and ecological context of a clade play in determining how a particular trait affects diversification.  
525 Future research will need to focus on how traits interact to influence diversification and identify the

526 ecological, genetic and demographic processes that these traits affect with the help of promising  
527 new models that are in development. We have only touched the surface of what we can learn about  
528 TDD in angiosperms. The production of new datasets that meet recommendations for robust  
529 inference, future methodological developments enabling studies at wider scopes and the potential  
530 for new discoveries in understudied traits and clades will continue to shape our understanding of  
531 angiosperm diversification. Though our study focused on angiosperms, the conclusions we draw  
532 about consistency, context dependence and diversification model use apply to studies of TDD across  
533 the entire tree of life.

## 534 **Acknowledgements**

535 This research is supported by the Fondation pour la Recherche sur la Biodiversité (FRB) through  
536 the CESAB project ‘DiveRS’. We thank the members of the ‘DiveRS’ group for the ideas and  
537 discussions that led to this manuscript. We thank Nicolas Casajus and Aurore Receveur for advice  
538 on data analyses.

## 539 **Supplementary Material**

540 A template for result reporting can be found in supplementary data 1. PapieRmache can be  
541 found at <http://github.com/ajhelmstetter/papieRmache>.

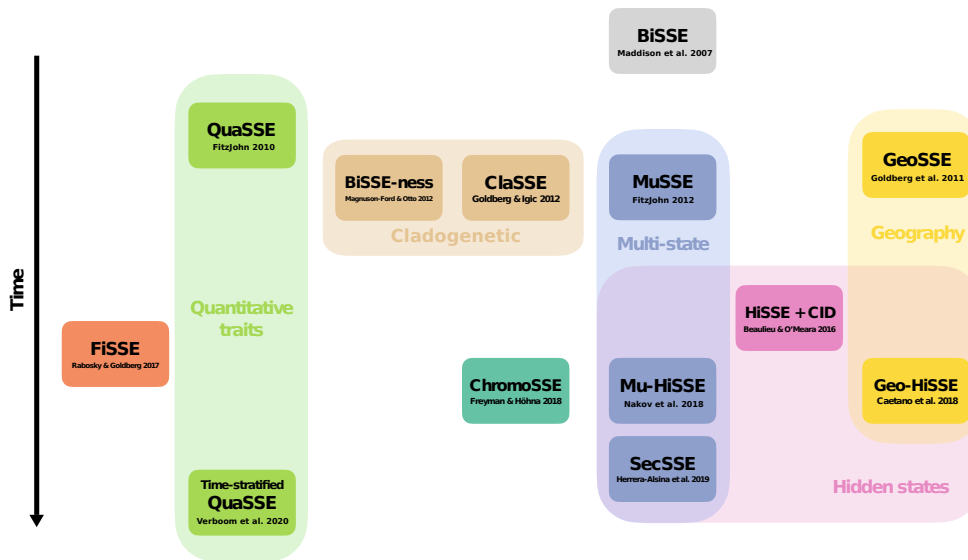


Figure 1: The development of state-dependent speciation and extinction (SSE) models. The original binary-state speciation and extinction model (BiSSE) model (Maddison et al., 2007) is shown at the top of the diagram with all other models depicted below, in the order of their publication. Acronyms are defined as follows: Binary-State Speciation and Extinction–node enhanced state shift (BiSSE-ness; Magnuson-Ford and Otto, 2012), Cladogenetic and Anagenetic Models of Chromosome Number Evolution (ChromoSSE; Freyman and Höhna, 2018), Character-Independent Diversification (CID Beaulieu and O’Meara, 2016), Cladogenetic State change Speciation and Extinction (ClaSSE; Goldberg and Igić, 2012), Fast, intuitive State-dependent Speciation-Extinction (FiSSE; Rabosky and Goldberg, 2017, Geographic State Speciation and Extinction (GeoSSE; Goldberg et al., 2011), Hidden Geographic State Speciation and Extinction (GeoHiSSE; Caetano et al., 2018), Hidden State Speciation and Extinction (HiSSE; Beaulieu and O’Meara, 2016), Multi-State Speciation and Extinction (MuSSE; FitzJohn, 2012), Multicharacter Hidden State Speciation and Extinction (Mu-HiSSE; Nakov et al., 2018), Quantitative State Speciation and Extinction (QuaSSE; FitzJohn, 2010; Verboom et al., 2020), Several Examined and Concealed States-dependent Speciation and Extinction (SecSSE; Herrera-Abina et al., 2019). Each box shows the name of the model and the associated citation. Models that share similar attributes (e.g. those with hidden states) are colour coded and grouped with boxes. This is not an exhaustive list of SSE models and does not include, for example, models used in epidemiology that allow tips to be sampled at various points in time (Scire et al., 2020).

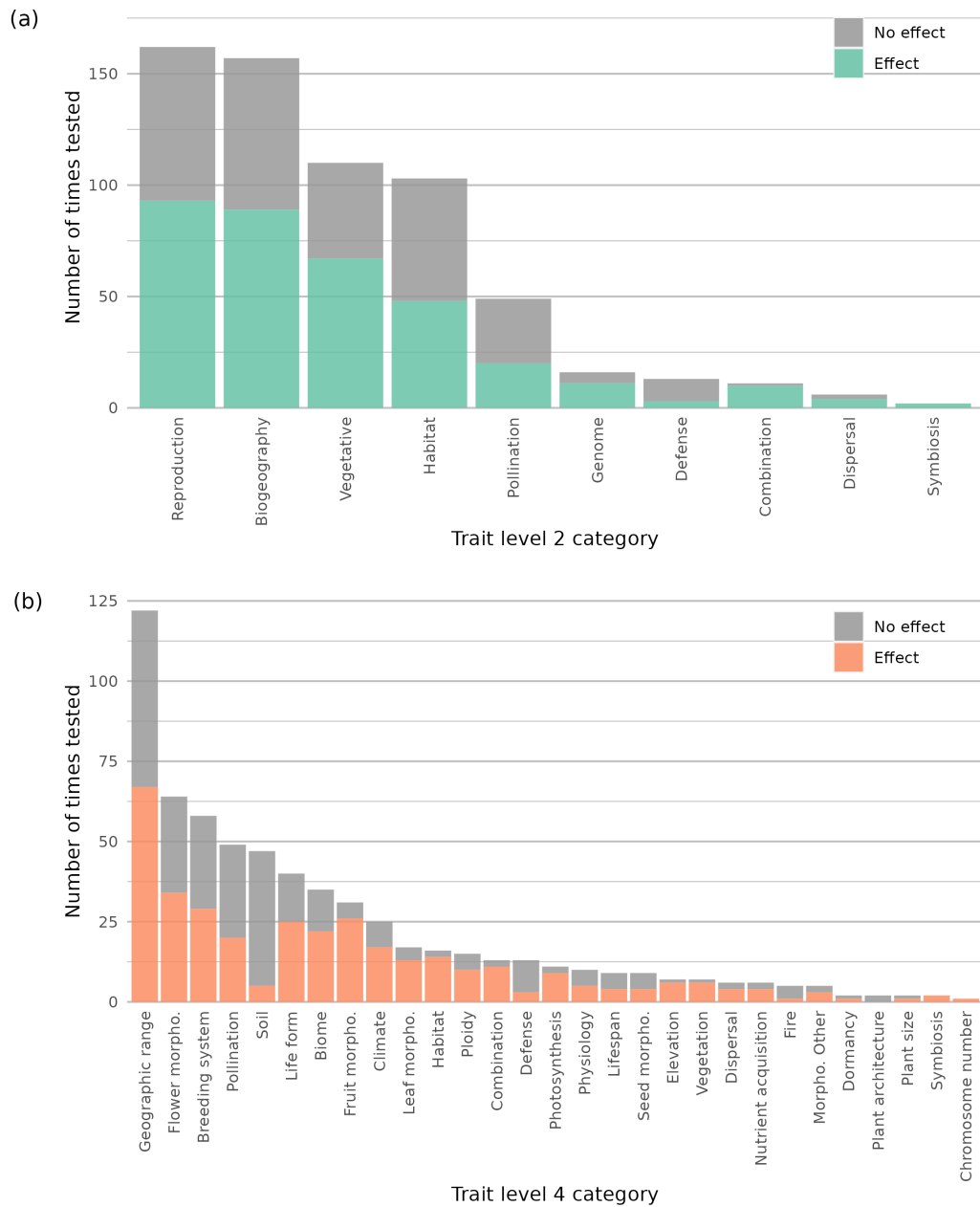


Figure 2: Stacked barplots showing how often particularly trait types were tested with state-dependent speciation and extinction (SSE) models. Bars are coloured to depict how often Trait-dependent diversification was detected per trait type. If results from multiple SSE models were reported in a single study they were considered separately i.e. each model contributed one result to the totals for each trait category. Two plots are shown, (a) one with relatively broad trait categories (level 2) and (b) one with narrower categories (level 4). An ontology depicting how different trait classification levels are connected can be found in Table S1.

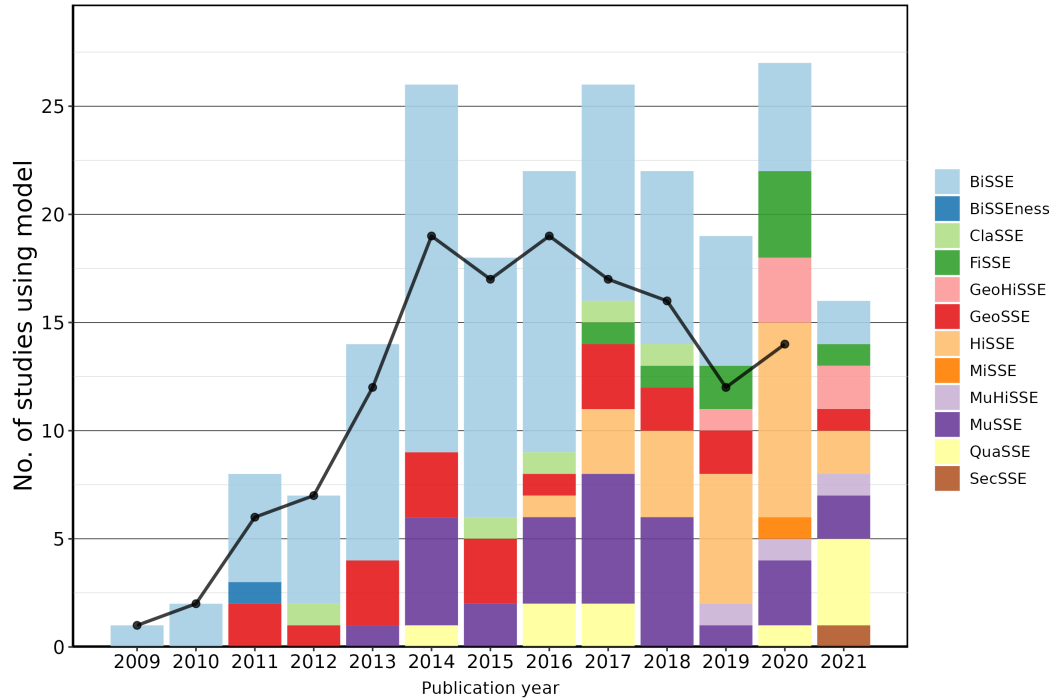


Figure 3: A stacked barplot showing the change in state-dependent speciation and extinction (SSE) models used on angiosperm clades over time. Each bar indicates the number of unique SSE model types per study totaled over the publication year. Bars are coloured according to the proportion of each SSE model type published in that year (see legend on the right of the plot). If the same SSE model was used multiple times in a single study it is only counted once (e.g. if BiSSE was used four times in a study published in 2012 this contributes an increase of one to the BiSSE portion of the 2012 bar). The black line shows the total number of studies using SSE models on angiosperms per year. Note that the year 2021 is incomplete.

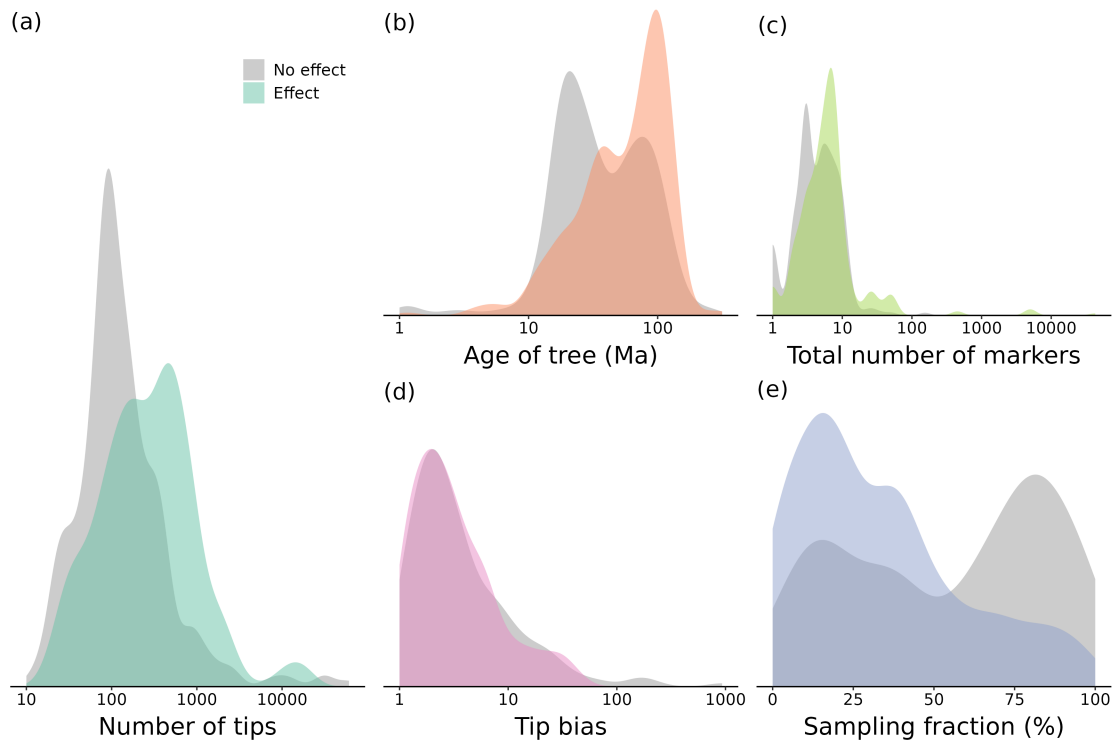


Figure 4: A set of densities depicting the distribution of values for five dataset properties in SSE models that infer trait dependent diversification (coloured densities) and those that do not (grey densities). The dataset properties shown are (a) number of tips in the phylogenetic tree used with the SSE model (data taken from  $n = 621$  models), (b) the age of the tree used with the SSE model ( $n = 529$ ), (c) the total number of genetic markers (nuclear + plastid + mitochondrial) used to infer the phylogenetic tree used with the SSE model ( $n = 615$ ), (d) the tip bias, here calculated as the largest tip frequency divided by the smallest ( $n = 426$ ) and (e) the global sampling fraction ( $n = 616$ ).

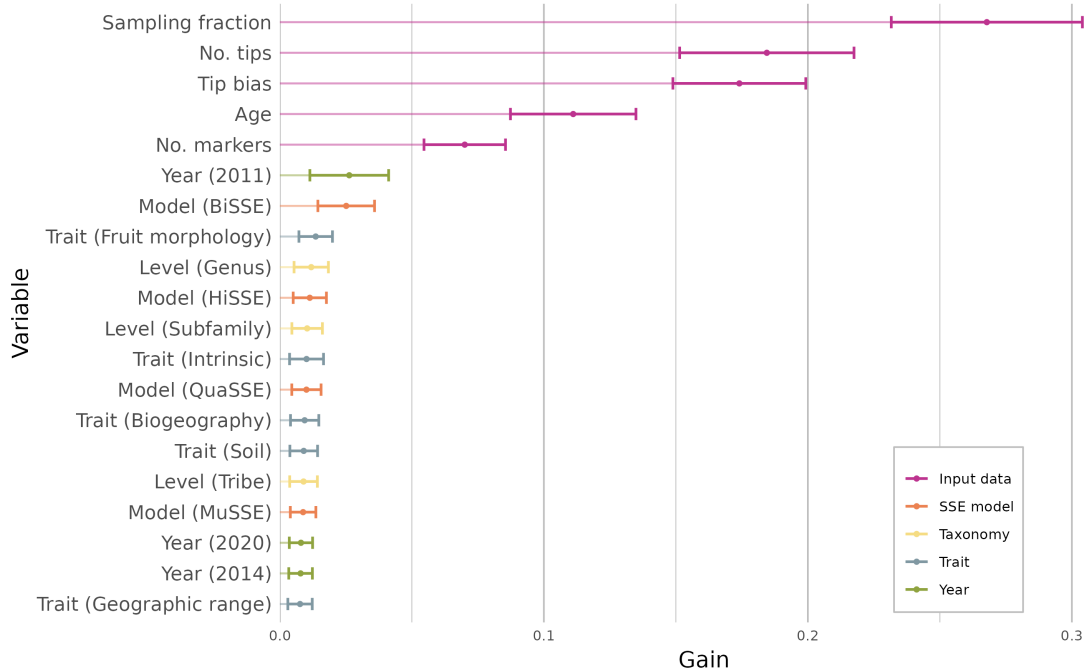


Figure 5: A horizontal barplot showing the relative influence of the 20 most important features included in the xgboost model used to predict the outcomes of SSE models, whether or not trait-dependent diversification is inferred, using input data properties and other characteristics of each study. All categorical variables are binary (e.g. ‘Intrinsic’ provides the information that traits that are intrinsic (1) and not intrinsic (0)). Points are the mean gain values calculated from the 500 iterations that were run. Error bars represent one standard deviation around the mean. Bars are coloured based on the type of variable they represent.



## 542 **Box 1: Studying trait-dependent diversification with SSE models**

543 There are many different ways to explore differences in diversification (Ricklefs, 2007). The simplest  
544 way is to quantify species richness, but as species richness is highly dependent on time, either one should  
545 only compare equally old clades (sister clade comparisons) or correct for the ages of the clades (e.g. the  
546 method-of-moments estimator (Magallon & Sanderson, 2001)). More information can be obtained by  
547 modelling the shape of the phylogenetic tree (see Morlon, 2014). Some methods are aimed at identifying  
548 shifts in the diversification rate along the phylogenetic tree (e.g. Rabosky, 2014) after which it is up to the  
549 user to interpret these shifts and test whether they correlate with species' characteristics. Other more recent  
550 methods allow users to estimate diversification rates at the tips of the tree and then to test whether rate  
551 differences are associated with species' traits (Rabosky & Huang, 2016).

552 Species diversification can be explicitly linked to traits via state-dependent speciation and extinction  
553 (SSE) models. This popular family of models is based on birth-death processes where the diversification  
554 rates (birth is speciation, and death is extinction) are dependent on character states, and where transition  
555 rates between states define how state changes occur. The simplest SSE model is the Binary-State Speciation  
556 and Extinction (BiSSE) model (Maddison et al., 2007) that takes as input a phylogenetic tree and state  
557 values (0 or 1) for each species in the tree. This allows users to uncover whether lineages with one state  
558 diversify faster than those with the other. SSE models can also be used to test whether the transition rates  
559 between states in one direction (0 to 1) are faster than the other (1 to 0). A first important extension to  
560 BiSSE was the possibility to use incomplete trees (FitzJohn et al., 2009).

561 The original model has been extended in various ways to address different types of macroevolutionary  
562 questions, as depicted in Fig. 1. ClaSSE (Goldberg & Igić, 2012) and BiSSE-ness (Magnuson-Ford & Otto,  
563 2012) are extensions of BiSSE that include cladogenetic events (speciation simultaneously associated with  
564 change in state). ChromoSSE (Freyman & Höhna, 2018) is an extension of ClaSSE specially designed to  
565 study changes in chromosome numbers such as polyploidization. GeoSSE (Goldberg et al., 2011) explicitly  
566 models how diversification differs between geographic regions. Other developments include the extension to  
567 more than two character states in MuSSE (FitzJohn, 2012), and QuaSSE (FitzJohn, 2010; Verboom et al.,  
568 2020), which can examine the effect of quantitative traits such as mean temperature.

569 Early SSE models were criticised for their propensity to infer false positives due to model inadequacy:  
570 the models were based on the assumption that only the trait of interest would influence diversification, so  
571 any kind of heterogeneity would lead to the rejection of the null hypothesis (Maddison & FitzJohn, 2015;  
572 Rabosky & Goldberg, 2015). FiSSE (Rabosky & Goldberg, 2017) is one solution for that problem, but  
573 perhaps the most important innovation after the initial wave of SSE models is the introduction of hidden  
574 states. These allowed diversification rates to be influenced by the focal traits as well as unobserved factors  
575 and provided a new set of more complex null hypotheses: the character independent diversification (CID)  
576 models (Beaulieu & O'Meara, 2016)). Hidden states have been incorporated into most initial SSE methods:

577 HiSSE (Beaulieu & O'Meara, 2016) as an extension to BiSSE, Mu-HiSSE (Nakov et al., 2018), GeoHiSSE  
578 (Caetano et al., 2018) and SecSSE (Herrera-Alsina et al., 2019).

## 579 **Box 2: Choosing appropriate clades and traits**

580 Given that dataset properties, evolutionary history and ecological context appear to be important for  
581 understanding TDD, what strategies can be used to select clades and traits that will produce the most reliable  
582 results? Trait choice can be helped by preliminary knowledge of the phylogenetic tree and ancestral state  
583 reconstruction, which could be used to ensure that the derived state(s) arose multiple times and that the  
584 ratio among different states is not extreme ( $<10:1$ ). However, exactly how many times a trait needs to evolve  
585 to be suitable for study remains a difficult question (Maddison & FitzJohn, 2015) that should be further  
586 investigated with simulations. With the uncertainties inherent to many studies due to less-than-optimal  
587 datasets, we argue there is a large avenue for the investigation of the effect of previously studied traits on  
588 diversification using high-quality data and the most powerful methods. Although our expectation is that  
589 the results will still be inconsistent due to context dependence, precise knowledge about which traits are  
590 most often important, and in which clades, will set the stage for more detailed studies on the causes of  
591 inconsistency.

592 In terms of choosing a clade, it is first important to adhere, as best as possible, to the recommendations  
593 for using SSE models e.g. avoid clades much smaller than 300 taxa and focus on those that are well sampled  
594 ( $> 25\%$ ). If recommendations cannot be followed, because of natural limitations in clade size, for example,  
595 these should be stated clearly as caveats. Working at a large scale, e.g. angiosperm-level analyses, is certainly  
596 appealing but creates a range of issues related to confounding factors that current models will find difficult to  
597 disentangle. These range from evolutionary history and ecological context, as discussed above, but also other  
598 analytical issues such as molecular clock rate variation (Shafir et al., 2020). To better learn about the factors  
599 that influence angiosperm diversity, we suggest that studies focus on multiple intermediate-sized clades, i.e.  
600 large, well-defined genera, families or tractable orders. However, if these clades are well-sampled they would  
601 approach the limit of our current computational feasibility (but see Louca and Pennell, 2020b). Working  
602 with many smaller clades may therefore be more feasible in the near future and also yield important insights  
603 via the comparison of diversification patterns among many different groups (e.g. Sabath et al., 2016), which  
604 we think is an acceptable tradeoff for reduced power in standalone analyses.

605 Focusing on many smaller clades can also help to study rare traits, i.e. those that would lead to strong  
606 tip ratio bias ( $> 10 : 1$ ) in large clades, by selecting a subset of smaller clades with a higher frequency of  
607 the trait. Care should be taken that this could potentially bias the results, as those clades could be the ones  
608 where the trait is most successful. However, in the case of dioecy, which has an angiosperm-wide frequency  
609 of  $\sim 5\%$ , this SSE approach applied to  $\sim 40$  genera by Sabath et al. (2016) lead to similar conclusions  
610 as angiosperm-wide sister clade comparisons (Käfer et al., 2014). In general, examining the effect of the  
611 same trait in multiple clades would allow researchers to account for the unique and shared aspects of their  
612 biology (e.g. through the use of hidden states or trait combinations), and then to combine results (Rabosky  
613 & Goldberg, 2015) to uncover broad patterns about the trait of interest.

### Box 3: Recommendations for reporting SSE model results

In most studies, at least part of the information we consider crucial for the interpretation of the results was lacking, or it was difficult to access. Collecting data for many properties required us to count from figures or extract statistics from archived raw data, which were not always available. For example, even if the tree has been published in a different study, we recommend that SSE model users clearly and succinctly describe it in their own study, notably the number and nature of the markers used to construct it, the number of tips, the sampling fraction and the age of the root. An estimation of the total number of species in each state would allow users to assess whether there are differences in tip state frequencies or state-specific sampling fraction that could bias results. Knowledge about the root state or assumptions made about it when running the model should also be reported. Surprisingly, the results about diversification and transition rates were often not reported in an easily-accessible and standardized manner, or in some cases, not at all. These should be reported, preferably with confidence metrics around rate estimates e.g. Bayesian credible intervals.

We were unable to extract and use the number of independent origins of each character state. Robust estimates of associations between traits and diversification rates necessitate multiple independent origins (but not too many (Rabosky & Goldberg, 2015)) and corresponding rate changes (Maddison & FitzJohn, 2015), so an idea of this value per study, inferred using ancestral state reconstructions, would be useful for interpretation of the robustness of results. This could be done by combining stochastic mapping of traits with an SSE model (Freyman & Höhna, 2019), though this is generally not available for SSE approaches.

Louca and Pennell (2020a) recently pointed out how diversification rate estimation can be susceptible to issues of unidentifiability. Though SSE models are not directly implicated (Helmstetter et al., 2021), it is known that extinction rate is especially difficult to estimate correctly (Beaulieu & O'Meara, 2016; Maddison et al., 2007). One potential way to help 'future proof' analyses from unidentifiability caused by overfitting would be to avoid reporting and assessing speciation and extinction rates separately, focusing instead on compound parameters such as net diversification rate ( $\lambda - \mu$ ), turnover rate ( $\lambda + \mu$ ) and extinction fraction ( $\mu/\lambda$ ) that are typically used in more recent SSE models (Beaulieu & O'Meara, 2016). However, testing explicitly whether diversification rates, which can be robustly estimated and are often the focus of the studies, are significantly different (irrespective of the exact values of the speciation and extinction rates) is not often done explicitly. Researchers should thus carefully inspect (and report) the bootstrap, confidence, or credible intervals of the state-dependent diversification rates they obtain in order to conclude whether or not they are really different. To encourage standardized result reporting we propose an initial set of characteristics that should be made available in all future studies using SSE models (Supplementary Data 1).

## References

- Adler, P. B., Fajardo, A., Kleinhesselink, A. R., & Kraft, N. J. B. (2013). Trait-based tests of coexistence mechanisms. *Ecology Letters*, *16*(10), 1294–1306. <https://doi.org/10.1111/ele.12157>
- Aristide, L., & Morlon, H. (2019). Understanding the effect of competition during evolutionary radiations: An integrated model of phenotypic and species diversification. *Ecology Letters*, *22*(12), 2006–2017. <https://doi.org/10.1111/ele.13385>
- Azani, N., Bruneau, A., Wojciechowski, M. F., & Zarre, S. (2019). Miocene climate change as a driving force for multiple origins of annual species in *Astragalus* (Fabaceae, Papilionoideae). *Molecular Phylogenetics and Evolution*, *137*, 210–221. <https://doi.org/10.1016/j.ympev.2019.05.008>
- Barrett, S. C. H. (2013). The evolution of plant reproductive systems: How often are transitions irreversible? *Proceedings of the Royal Society B: Biological Sciences*, *280*(1765), 20130913. <https://doi.org/10.1098/rspb.2013.0913>
- Barrett, S. C. H., Harder, L. D., & Worley, A. C. (1996). The comparative biology of pollination and mating in flowering plants. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *351*(1345), 1271–1280. <https://doi.org/10.1098/rstb.1996.0110>
- Beaulieu, J. M., & O'Meara, B. C. (2016). Detecting Hidden Diversification Shifts in Models of Trait-Dependent Speciation and Extinction. *Systematic Biology*, *65*(4), 583–601. <https://doi.org/10.1093/sysbio/syw022>
- Benton, M. J., Wilf, P., & Sauquet, H. (2022). The Angiosperm Terrestrial Revolution and the origins of modern biodiversity. *New Phytologist*, *233*(5), 2017–2035. <https://doi.org/10.1111/nph.17822>
- Caetano, D. S., O'Meara, B. C., & Beaulieu, J. M. (2018). Hidden state models improve state-dependent diversification approaches, including biogeographical models. *Evolution*, *72*(11), 2308–2324. <https://doi.org/10.1111/evo.13602>
- Canal, D., Köster, N., Celis, M., Croat, T. B., Borsch, T., & Jones, K. E. (2019). Out of Amazonia and Back again: Historical Biogeography of the Species-Rich Neotropical Genus *Philodendron* (Araceae)1. *Annals of the Missouri Botanical Garden*, *104*(1), 49–68. <https://doi.org/10.3417/2018266>

- 675 Chang, J., Rabosky, D. L., & Alfaro, M. E. (2020). Estimating Diversification Rates on Incompletely  
676 Sampled Phylogenies: Theoretical Concerns and Practical Solutions (T. Buckley, Ed.).  
677 *Systematic Biology*, 69(3), 602–611. <https://doi.org/10.1093/sysbio/syz081>
- 678 Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. *Proceedings of the*  
679 *22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*,  
680 785–794. <https://doi.org/10.1145/2939672.2939785>
- 681 Coyne, J. A., & Orr, H. A. (2004). *Speciation*. Oxford University Press, Incorporated.
- 682 Crepet, W. L., & Niklas, K. J. (2009). Darwin’s second “abominable mystery”: Why are there so  
683 many angiosperm species? *American Journal of Botany*, 96(1), 366–381.
- 684 Davies, T. J., Barraclough, T. G., Chase, M. W., Soltis, P. S., Soltis, D. E., & Savolainen, V. (2004).  
685 Darwin’s abominable mystery: Insights from a supertree of the angiosperms. *Proceedings*  
686 *of the National Academy of Sciences*, 101(7), 1904–1909. [https://doi.org/10.1073/pnas.](https://doi.org/10.1073/pnas.0308127100)  
687 0308127100
- 688 Davis, M. P., Midford, P. E., & Maddison, W. (2013). Exploring power and parameter estimation of  
689 the BiSSE method for analyzing species diversification. *BMC Evolutionary Biology*, 13(1),  
690 38. <https://doi.org/10.1186/1471-2148-13-38>
- 691 de Queiroz, A. (2002). Contingent Predictability in Evolution: Key Traits and Diversification.  
692 *Systematic Biology*, 51(6), 917–929.
- 693 Donoghue, M. J. (2005). Key innovations, convergence, and success: Macroevolutionary lessons from  
694 plant phylogeny. *Paleobiology*, 31(sp5), 77–93. [https://doi.org/10.1666/0094-8373\(2005\)](https://doi.org/10.1666/0094-8373(2005)031[0077:KICASM]2.0.CO;2)  
695 031[0077:KICASM]2.0.CO;2
- 696 Donoghue, M. J., & Sanderson, M. J. (2015). Confluence, synnovation, and depauperons in plant  
697 diversification. *New Phytologist*, 207(2), 260–274. <https://doi.org/10.1111/nph.13367>
- 698 Drummond, C. S., Eastwood, R. J., Miotto, S. T. S., & Hughes, C. E. (2012). Multiple Continental  
699 Radiations and Correlates of Diversification in *Lupinus* (Leguminosae): Testing for Key  
700 Innovation with Incomplete Taxon Sampling. *Systematic Biology*, 61(3), 443–460. <https://doi.org/10.1093/sysbio/syr126>  
701
- 702 Drury, J., Clavel, J., Manceau, M., & Morlon, H. (2016). Estimating the Effect of Competition on  
703 Trait Evolution Using Maximum Likelihood Inference. *Systematic Biology*, 65(4), 700–710.  
704 <https://doi.org/10.1093/sysbio/syw020>

- 705 Falster, D., Gallagher, R., Wenk, E. H., Wright, I. J., Indiarto, D., et al. (2021). AusTraits, a  
706 curated plant trait database for the Australian flora. *Scientific Data*, 8(1), 254. <https://doi.org/10.1038/s41597-021-01006-6>  
707
- 708 Felsenstein, J. (1985). Phylogenies and the Comparative Method. *The American Naturalist*, 125(1),  
709 1–15.
- 710 FitzJohn, R. G. (2010). Quantitative Traits and Diversification. *Systematic Biology*, 59(6), 619–633.  
711 <https://doi.org/10.1093/sysbio/syq053>
- 712 FitzJohn, R. G. (2012). Diversitree : Comparative phylogenetic analyses of diversification in R:  
713 *Diversitree. Methods in Ecology and Evolution*, 3(6), 1084–1092. [https://doi.org/10.1111/](https://doi.org/10.1111/j.2041-210X.2012.00234.x)  
714 [j.2041-210X.2012.00234.x](https://doi.org/10.1111/j.2041-210X.2012.00234.x)
- 715 FitzJohn, R. G., Maddison, W. P., & Otto, S. P. (2009). Estimating Trait-Dependent Speciation  
716 and Extinction Rates from Incompletely Resolved Phylogenies. *Systematic Biology*, 58(6),  
717 595–611. <https://doi.org/10.1093/sysbio/syp067>
- 718 Folk, R. A., & Freudenstein, J. V. (2014). Phylogenetic relationships and character evolution in  
719 Heuchera (Saxifragaceae) on the basis of multiple nuclear loci. *American Journal of Botany*,  
720 101(9), 1532–1550. <https://doi.org/10.3732/ajb.1400290>
- 721 Freyman, W. A., & Höhna, S. (2018). Cladogenetic and Anagenetic Models of Chromosome Number  
722 Evolution: A Bayesian Model Averaging Approach. *Systematic Biology*, 67(2), 195–215.  
723 <https://doi.org/10.1093/sysbio/syx065>
- 724 Freyman, W. A., & Höhna, S. (2019). Stochastic Character Mapping of State-Dependent Diversification  
725 Reveals the Tempo of Evolutionary Decline in Self-Compatible Onagraceae Lineages. *Systematic*  
726 *Biology*, 68(3), 505–519. <https://doi.org/10.1093/sysbio/syy078>
- 727 Gamba, D., & Muchhala, N. (2020). Global patterns of population genetic differentiation in seed  
728 plants. *Molecular Ecology*, 29(18), 3413–3428. <https://doi.org/10.1111/mec.15575>
- 729 Goldberg, E. E., & Igić, B. (2012). TEMPO AND MODE IN PLANT BREEDING SYSTEM  
730 EVOLUTION: TEMPO AND MODE IN PLANT BREEDING SYSTEM EVOLUTION.  
731 *Evolution*, 66(12), 3701–3709. <https://doi.org/10.1111/j.1558-5646.2012.01730.x>
- 732 Goldberg, E. E., Lancaster, L. T., & Ree, R. H. (2011). Phylogenetic Inference of Reciprocal  
733 Effects between Geographic Range Evolution and Diversification. *Systematic Biology*, 60(4),  
734 451–465. <https://doi.org/10.1093/sysbio/syr046>

- 735 Han, T.-S., Zheng, Q.-J., Onstein, R. E., Rojas-Andrés, B. M., Hauenschild, F., et al. (2020).  
736 Polyploidy promotes species diversification of *Allium* through ecological shifts. *New Phytologist*,  
737 *225*(1), 571–583. <https://doi.org/10.1111/nph.16098>
- 738 Hastie, T. J., & Tibshirani, R. J. (2017). *Generalized Additive Models*. Routledge. [https://doi.org/](https://doi.org/10.1201/9780203753781)  
739 [10.1201/9780203753781](https://doi.org/10.1201/9780203753781)
- 740 Helmstetter, A. J., Glemin, S., Käfer, J., Zenil-Ferguson, R., Sauquet, H., et al. (2021). Pulled  
741 Diversification Rates, Lineages-Through-Time Plots and Modern Macroevolutionary Modelling.  
742 *Systematic Biology*, syab083. <https://doi.org/10.1093/sysbio/syab083>
- 743 Hernández-Hernández, T., & Wiens, J. J. (2020). Why Are There So Many Flowering Plants?  
744 A Multiscale Analysis of Plant Diversification. *The American Naturalist*, *195*(6), 948–963.  
745 <https://doi.org/10.1086/708273>
- 746 Herrera-Alsina, L., van Els, P., & Etienne, R. S. (2019). Detecting the Dependence of Diversification  
747 on Multiple Traits from Phylogenetic Trees and Trait Data (M. Alfaro, Ed.). *Systematic*  
748 *Biology*, *68*(2), 317–328. <https://doi.org/10.1093/sysbio/syy057>
- 749 Hunter, J. P. (1998). Key innovations and the ecology of macroevolution. *Trends in Ecology &*  
750 *Evolution*, *13*(1), 31–36. [https://doi.org/10.1016/S0169-5347\(97\)01273-1](https://doi.org/10.1016/S0169-5347(97)01273-1)
- 751 Jablonski, D. (2008). Species Selection: Theory and Data. *Annual Review of Ecology, Evolution, and*  
752 *Systematics*, *39*(1), 501–524. <https://doi.org/10.1146/annurev.ecolsys.39.110707.173510>
- 753 Jetz, W., Thomas, G. H., Joy, J. B., Hartmann, K., & Mooers, A. O. (2012). The global diversity of  
754 birds in space and time. *Nature*, *491*(7424), 444–448. <https://doi.org/10.1038/nature11631>
- 755 Käfer, J., Boer, H. J. d., Mousset, S., Kool, A., Dufay, M., & Marais, G. a. B. (2014). Dioecy  
756 is associated with higher diversification rates in flowering plants. *Journal of Evolutionary*  
757 *Biology*, *27*(7), 1478–1490. <https://doi.org/https://doi.org/10.1111/jeb.12385>
- 758 Kaj, I., Glémin, S., Tahir, D., & Lascoux, M. (2021). Analysis of diversity-dependent species  
759 evolution using concepts in population genetics. *Journal of Mathematical Biology*, *82*(4),  
760 22. <https://doi.org/10.1007/s00285-021-01559-5>
- 761 Kattge, J., Bönisch, G., Díaz, S., Lavorel, S., Prentice, I. C., et al. (2020). TRY plant trait database  
762 – enhanced coverage and open access. *Global Change Biology*, *26*(1), 119–188. [https://doi.](https://doi.org/10.1111/gcb.14904)  
763 [org/10.1111/gcb.14904](https://doi.org/10.1111/gcb.14904)
- 764 Landis, J. B., Soltis, D. E., Li, Z., Marx, H. E., Barker, M. S., et al. (2018). Impact of whole-genome  
765 duplication events on diversification rates in angiosperms. *American Journal of Botany*,  
766 *105*(3), 348–363. <https://doi.org/10.1002/ajb2.1060>



- 767 Louca, S., & Pennell, M. W. (2020a). Extant timetrees are consistent with a myriad of diversification  
768 histories. *Nature*, *580*(7804), 502–505. <https://doi.org/10.1038/s41586-020-2176-1>
- 769 Louca, S., & Pennell, M. W. (2020b). A General and Efficient Algorithm for the Likelihood of  
770 Diversification and Discrete-Trait Evolutionary Models. *Systematic Biology*, *69*(3), 545–556.  
771 <https://doi.org/10.1093/sysbio/syz055>
- 772 Maddison, W. P., & FitzJohn, R. G. (2015). The unsolved challenge to phylogenetic correlation  
773 tests for categorical characters. *Systematic Biology*, *64*(1), 127–136. [https://doi.org/10.](https://doi.org/10.1093/sysbio/syu070)  
774 [1093/sysbio/syu070](https://doi.org/10.1093/sysbio/syu070)
- 775 Maddison, W. P., Midford, P. E., & Otto, S. P. (2007). Estimating a Binary Character’s Effect  
776 on Speciation and Extinction (T. Oakley, Ed.). *Systematic Biology*, *56*(5), 701–710. <https://doi.org/10.1080/10635150701607033>
- 777 [//doi.org/10.1080/10635150701607033](https://doi.org/10.1080/10635150701607033)
- 778 Magallon, S., & Sanderson, M. J. (2001). Absolute Diversification Rates in Angiosperm Clades.  
779 *Evolution*, *55*(9), 1762–1780.
- 780 Magallón, S., & Castillo, A. (2009). Angiosperm diversification through time. *American Journal of*  
781 *Botany*, *96*(1), 349–365. <https://doi.org/10.3732/ajb.0800060>
- 782 Magallón, S., Sánchez-Reyes, L. L., & Gómez-Acevedo, S. L. (2019). Thirty clues to the exceptional  
783 diversification of flowering plants. *Annals of Botany*, *123*(3), 491–503. [https://doi.org/10.](https://doi.org/10.1093/aob/mcy182)  
784 [1093/aob/mcy182](https://doi.org/10.1093/aob/mcy182)
- 785 Magnuson-Ford, K., & Otto, S. P. (2012). Linking the Investigations of Character Evolution and  
786 Species Diversification. *The American Naturalist*, *180*(2), 225–245. [https://doi.org/10.](https://doi.org/10.1086/666649)  
787 [1086/666649](https://doi.org/10.1086/666649)
- 788 Maliet, O., Hartig, F., & Morlon, H. (2019). A model with many small shifts for estimating  
789 species-specific diversification rates. *Nature Ecology & Evolution*, *3*(7), 1086–1092. <https://doi.org/10.1038/s41559-019-0908-0>
- 790 [//doi.org/10.1038/s41559-019-0908-0](https://doi.org/10.1038/s41559-019-0908-0)
- 791 Morlon, H. (2014). Phylogenetic approaches for studying diversification (A. Mooers, Ed.). *Ecology*  
792 *Letters*, *17*(4), 508–525. <https://doi.org/10.1111/ele.12251>
- 793 Murray, B. R., Thrall, P. H., Gill, A. M., & Nicotra, A. B. (2002). How plant life-history and  
794 ecological traits relate to species rarity and commonness at varying spatial scales. *Austral*  
795 *Ecology*, *27*(3), 291–310. <https://doi.org/10.1046/j.1442-9993.2002.01181.x>
- 796 Nakov, T., Beaulieu, J. M., & Alverson, A. J. (2018). Accelerated diversification is related to  
797 life history and locomotion in a hyperdiverse lineage of microbial eukaryotes (Diatoms,  
798 Bacillariophyta). *New Phytologist*, *219*(1), 462–473. <https://doi.org/10.1111/nph.15137>

- 799 O'Meara, B. C., Smith, S. D., Armbruster, W. S., Harder, L. D., Hardy, C. R., et al. (2016).  
800 Non-equilibrium dynamics and floral trait interactions shape extant angiosperm diversity.  
801 *Proceedings of the Royal Society B: Biological Sciences*, 283(1830), 20152304. <https://doi.org/10.1098/rspb.2015.2304>
- 802
- 803 Onstein, R. E. (2019). Darwin's second 'abominable mystery': Trait flexibility as the innovation  
804 leading to angiosperm diversity. *New Phytologist*, nph.16294. <https://doi.org/10.1111/nph.16294>
- 805
- 806 Onstein, R. E., Baker, W. J., Couvreur, T. L. P., Faurby, S., Svenning, J.-C., & Kissling, W. D.  
807 (2017). Frugivory-related traits promote speciation of tropical palms. *Nature Ecology &*  
808 *Evolution*, 1(12), 1903–1911. <https://doi.org/10.1038/s41559-017-0348-7>
- 809 Park, D. S., Ellison, A. M., & Davis, C. C. (2018). Mating system does not predict niche breath.  
810 *Global Ecology and Biogeography*, 27(7), 804–813. <https://doi.org/10.1111/geb.12740>
- 811 Rabosky, D. L. (2013). Diversity-Dependence, Ecological Speciation, and the Role of Competition  
812 in Macroevolution. *Annual Review of Ecology, Evolution, and Systematics*, 44(1), 481–502.  
813 <https://doi.org/10.1146/annurev-ecolsys-110512-135800>
- 814 Rabosky, D. L. (2014). Automatic Detection of Key Innovations, Rate Shifts, and Diversity-Dependence  
815 on Phylogenetic Trees (S.-O. Kolokotronis, Ed.). *PLoS ONE*, 9(2), e89543. <https://doi.org/10.1371/journal.pone.0089543>
- 816
- 817 Rabosky, D. L., Chang, J., Title, P. O., Cowman, P. F., Sallan, L., et al. (2018). An inverse  
818 latitudinal gradient in speciation rate for marine fishes. *Nature*, 559(7714), 392–395. <https://doi.org/10.1038/s41586-018-0273-1>
- 819
- 820 Rabosky, D. L., & Goldberg, E. E. (2015). Model Inadequacy and Mistaken Inferences of Trait-Dependent  
821 Speciation. *Systematic Biology*, 64(2), 340–355. <https://doi.org/10.1093/sysbio/syu131>
- 822 Rabosky, D. L., & Goldberg, E. E. (2017). FiSSE: A simple nonparametric test for the effects  
823 of a binary character on lineage diversification rates: TEST FOR TRAIT-DEPENDENT  
824 DIVERSIFICATION. *Evolution*, 71(6), 1432–1442. <https://doi.org/10.1111/evo.13227>
- 825 Rabosky, D. L., & Huang, H. (2016). A Robust Semi-Parametric Test for Detecting Trait-Dependent  
826 Diversification. *Systematic Biology*, 65(2), 181–193. <https://doi.org/10.1093/sysbio/syv066>
- 827 Raup, D. M., & Sepkoski, J. J. (1982). Mass Extinctions in the Marine Fossil Record. *Science*,  
828 215(4539), 1501–1503. <https://doi.org/10.1126/science.215.4539.1501>

- 829 Renner, S. S. (2014). The relative and absolute frequencies of angiosperm sexual systems: Dioecy,  
830 monoecy, gynodioecy, and an updated online database. *American Journal of Botany*, *101*(10),  
831 1588–1596. <https://doi.org/10.3732/ajb.1400196>
- 832 Renner, S. S., & Ricklefs, R. E. (1995). Dioecy and Its Correlates in the Flowering Plants. *American*  
833 *Journal of Botany*, *82*(5), 596–606. <https://doi.org/10.2307/2445418>
- 834 Ricklefs, R. E. (2007). Estimating diversification rates from phylogenetic information. *Trends in*  
835 *Ecology & Evolution*, *22*(11), 601–610. <https://doi.org/10.1016/j.tree.2007.06.013>
- 836 Ricklefs, R. E., & Renner, S. S. (1994). Species Richness Within Families of Flowering Plants.  
837 *Evolution*, *48*(5), 1619–1636. <https://doi.org/10.2307/2410252>
- 838 Sabath, N., Goldberg, E. E., Glick, L., Einhorn, M., Ashman, T.-L., et al. (2016). Dioecy does not  
839 consistently accelerate or slow lineage diversification across multiple genera of angiosperms.  
840 *New Phytologist*, *209*(3), 1290–1300. <https://doi.org/10.1111/nph.13696>
- 841 Salariato, D. L., Zuloaga, F. O., Franzke, A., Mummenhoff, K., & Al-Shehbaz, I. A. (2016).  
842 Diversification patterns in the CES clade (Brassicaceae tribes Cremolobaeae, Eudemeae,  
843 Schizopetaleae) in Andean South America. *Botanical Journal of the Linnean Society*, *181*(4),  
844 543–566. <https://doi.org/10.1111/boj.12430>
- 845 Sauquet, H., & Magallón, S. (2018). Key questions and challenges in angiosperm macroevolution.  
846 *New Phytologist*, *219*(4), 1170–1187. <https://doi.org/10.1111/nph.15104>
- 847 Sauquet, H., von Balthazar, M., Magallón, S., Doyle, J. A., Endress, P. K., et al. (2017). The  
848 ancestral flower of angiosperms and its early diversification. *Nature Communications*, *8*(1),  
849 16047. <https://doi.org/10.1038/ncomms16047>
- 850 Schmidt, J. P., Davies, T. J., & Farrell, M. J. (2021). Opposing macroevolutionary and trait-mediated  
851 patterns of threat and naturalisation in flowering plants. *Ecology Letters*, *24*(6), 1237–1250.  
852 <https://doi.org/10.1111/ele.13740>
- 853 Scire, J., Barido-Sottani, J., Kühnert, D., Vaughan, T. G., & Stadler, T. (2020). *Improved multi-type*  
854 *birth-death phylodynamic inference in BEAST 2* (preprint). *Evolutionary Biology*. <https://doi.org/10.1101/2020.01.06.895532>
- 856 Shafir, A., Azouri, D., Goldberg, E. E., & Mayrose, I. (2020). Heterogeneity in the rate of molecular  
857 sequence evolution substantially impacts the accuracy of detecting shifts in diversification  
858 rates. *Evolution*, *n/a*(n/a). <https://doi.org/10.1111/evo.14036>
- 859 Smith, S. A., & Brown, J. W. (2018). Constructing a broadly inclusive seed plant phylogeny.  
860 *American Journal of Botany*, *105*(3), 302–314. <https://doi.org/10.1002/ajb2.1019>

- 861 Soltis, D. E., Mort, M. E., Latvis, M., Mavrodiev, E. V., O'Meara, B. C., et al. (2013). Phylogenetic  
862 relationships and character evolution analysis of Saxifragales using a supermatrix approach.  
863 *American Journal of Botany*, *100*(5), 916–929. <https://doi.org/10.3732/ajb.1300044>
- 864 Stebbins, G. L. (1974). *Flowering Plants: Evolution Above the Species Level*. Harvard University  
865 Press.
- 866 Sun, M., Folk, R. A., Gitzendanner, M. A., Soltis, P. S., Chen, Z., et al. (2020). Estimating rates and  
867 patterns of diversification with incomplete sampling: A case study in the rosids. *American*  
868 *Journal of Botany*, *107*(6), 895–909. <https://doi.org/10.1002/ajb2.1479>
- 869 Tucker, C. M., Davies, T. J., Cadotte, M. W., & Pearse, W. D. (2018). On the relationship between  
870 phylogenetic diversity and trait diversity. *Ecology*, *99*(6), 1473–1479. [https://doi.org/10.](https://doi.org/10.1002/ecy.2349)  
871 [1002/ecy.2349](https://doi.org/10.1002/ecy.2349)
- 872 Upham, N. S., Esselstyn, J. A., & Jetz, W. (2019). Inferring the mammal tree: Species-level sets  
873 of phylogenies for questions in ecology, evolution, and conservation (A. J. Tanentzap, Ed.).  
874 *PLOS Biology*, *17*(12), e3000494. <https://doi.org/10.1371/journal.pbio.3000494>
- 875 Vamosi, J. C., Otto, S. P., & Barrett, S. C. H. (2003). Phylogenetic analysis of the ecological  
876 correlates of dioecy in angiosperms. *Journal of Evolutionary Biology*, *16*(5), 1006–1018.  
877 <https://doi.org/10.1046/j.1420-9101.2003.00559.x>
- 878 Vamosi, J. C., Magallón, S., Mayrose, I., Otto, S. P., & Sauquet, H. (2018). Macroevolutionary  
879 Patterns of Flowering Plant Speciation and Extinction. *Annual Review of Plant Biology*,  
880 *69*(1), 685–706. <https://doi.org/10.1146/annurev-arplant-042817-040348>
- 881 Verboom, G. A., Boucher, F. C., Ackerly, D. D., Wootton, L. M., & Freyman, W. A. (2020). Species  
882 Selection Regime and Phylogenetic Tree Shape. *Systematic Biology*, *69*(4), 774–794. <https://doi.org/10.1093/sysbio/syz076>
- 883
- 884 Wang, Y., Luo, A., Lyu, T., Dimitrov, D., Xu, X., et al. (2021). Global distribution and evolutionary  
885 transitions of angiosperm sexual systems. *Ecology Letters*, *24*(9), 1835–1847. [https://doi.](https://doi.org/10.1111/ele.13815)  
886 [org/10.1111/ele.13815](https://doi.org/10.1111/ele.13815)
- 887 Zenil-Ferguson, R., Burleigh, J. G., Freyman, W. A., Igić, B., Mayrose, I., & Goldberg, E. E. (2019).  
888 Interaction among ploidy, breeding system and lineage diversification. *New Phytologist*,  
889 *224*(3), 1252–1265. <https://doi.org/10.1111/nph.16184>