Seed size drives species diversification across angiosperms

J. IGEA[1], E. F. MILLER[1,2], A. S. T. PAPADOPULOS[3], A. J. TANENTZAP[1]

[1] *Department of Plant Sciences, University of Cambridge, Downing St, Cambridge, CB2 3EA, UK*

[2] *Department of Zoology, University of Cambridge, Downing St, Cambridge, CB2 3EJ, UK*

[3] *Royal Botanic Gardens, Kew, Richmond, TW9 3AB, UK.*

Species diversity varies greatly across the different taxonomic groups that comprise the Tree of Life (ToL). This imbalance is particularly conspicuous within angiosperms, but is largely unexplained [1]. Seed mass is one factor that may help some lineages diversify more than others by influencing key life history traits, such as dispersal, colonisation, environmental tolerance and reproductive success [2]. However, the extent and direction of these effects have not been assessed across the angiosperm ToL. Here, we show for the first time that absolute seed size and the rate of change in seed size are both associated with variation in diversification rates. Based on an unequalled phylogenetic tree that included 4105 angiosperm genera, we found that smaller-seeded plants had higher rates of diversification, possibly due to improved colonisation potential [3]. The rate of phenotypic change in seed size was also strongly positively correlated with speciation rates, supporting emerging evidence that rapid morphological change is associated with species divergence [4]. Our study now reveals that

variation in morphological traits, as well as the rate at which traits evolve, contribute significantly to the extremely uneven distribution of diversity across the ToL.

Angiosperms are one of the most species-rich clades on Earth and have dominated terrestrial plant communities since the Late Cretaceous [5]. The astounding diversity of flowering plants is, however, extremely unevenly distributed across the ToL. Each of the five most species-rich angiosperm families contains >10,000 species while more than 200 families contain <100 species each. An enduring pursuit in evolutionary biology is to explain this uneven distribution of biodiversity, not only in angiosperms, but also across the whole ToL [1].

Seed mass influences individual life history in ways that can ultimately shape angiosperm diversification [6]. Along with adult plant size, seed mass affects survival, reproductive lifespan and dispersal [7]. These traits contribute to fitness and adaptation, which are the ultimate determinants of whether lineages diversify or go extinct [8]. In support of this idea, seed mass has been shown to correlate negatively with diversification in the Polygoneaceae [9], but this has not been investigated across large taxonomic scales. As seed mass varies over ten orders of magnitude in angiosperms, from the minute 1 μg seeds of some orchids to the >18 kg seeds of the sea coconut *Lodoicea maldivica*, this huge variation may coincide with variation in species diversity. Generalising the direction and magnitude of an effect of seed mass on diversification across taxonomic scales has however proved difficult. Some life history traits encapsulated by seed mass are expected to promote speciation or

51 extinction, while others may simultaneously counteract such effects.

52

53 The rate of change in key life history traits such as seed size can be as important in
54 driving macroevolutionary dynamics as the absolute values of the traits themselves [4].
55 This is because phenotypic divergence may cause reproductive isolation that results in
56 speciation [10]. Nevertheless, few empirical studies have detected a correlation between
57 rates of phenotypic evolution and lineage diversification [11]. This correlation may be
58 expected where a trait can change more rapidly in some species than others in
59 response to selective pressures (i.e. high "evolvability" [12]); this enables greater access
60 to new ecological niches or quicker establishment of reproductive isolation, and
61 thereby increasing the rate of speciation (λ) [13]. In the case of seed mass, the ability to
62 switch rapidly, for example, from small seeds with high dispersal ability to larger
63 seeds with lower dispersal ability, might promote cycles of rapid colonisation and
64 isolation or permit adaptation to new dispersal vectors in novel environments. Rapid
65 evolution of new phenotypes may also allow individuals to escape harsh
66 environmental conditions and competitive interactions [14], thereby decreasing
67 extinction rates (μ). The net outcome of these processes on diversification ($r = \lambda - \mu$)
68 will ultimately depend upon which of these rates responds more strongly to
69 phenotypic change.

70

71 Here we show for the first time that both seed mass and its phenotypic rate of
72 evolution influence speciation and extinction across the angiosperm ToL. Our
73 approach combined the most comprehensive phylogenetic timetree available [15] with
74 an unparalleled dataset of seed mass measurements to obtain mean seed mass values
75 for 24% of all described angiosperm genera (n = 4105). We estimated rates of

speciation and extinction using Bayesian Analysis of Macroevolutionary Mixtures (BAMM) [16], which models rate heterogeneity through time and lineages, and accounts for incomplete taxon sampling. Similarly, we analysed the rates of change in seed size across the phylogeny. We then estimated the correlations of the macroevolutionary dynamics with absolute seed size and the rate of seed size evolution.

As expected, given the high degree of taxonomic imbalance observed in the angiosperm phylogeny, we found strong support for multiple shifts in the rates of diversification. The median number of rate shifts for the speciation/extinction analysis was 123 (95% confidence interval (CI): 95-155) and 160 (95% CI: 62-110) for the seed size evolution analysis (Extended Data Fig. 1). There was marked heterogeneity in the rates of seed size evolution (Fig. 1), which varied over two orders of magnitude (Extended Data Fig. 2). We then estimated whether shifts in macroevolutionary dynamics (λ, μ and r) were significantly correlated with absolute mean genus seed size and rates of seed size evolution by comparing the existing correlations to a null distribution of correlations using STructured Rate Permutations on Phylogenies (STRAPP), which is robust to phylogenetic pseudoreplication (see Methods for details) [17].

For the first time, we were able to explain major differences in diversity across the angiosperm ToL with a single trait that integrates multiple aspects of life history. We specifically found evidence that increased speciation was associated with smaller seed sizes (Spearman's ρ = -0.22, p-value = 0.016; Fig. 2a). Increased extinction rate was similarly associated with smaller seeds (ρ = -0.20, p-value = 0.045), but given its relatively weaker effect, the net outcome of λ-μ was that diversification rates

101 increased with decreasing seed size (ρ =-0.19, p-value = 0.049). We also identified a
102 stronger positive association between the rate of seed size evolution and both
103 speciation and extinction (ρ = 0.48, p-value < 0.0001 and ρ = 0.37, p-value = 0.003,
104 respectively; Fig. 2b). Again, as the effect of speciation was greater than that of
105 extinction, rates of diversification and phenotypic change were positively correlated
106 (ρ = 0.46, p-value = 0.0002; Fig. 2b). Generally, the observed correlations arose from
107 many phenotypically fast-evolving clades distributed across the phylogeny (Extended
108 Data Fig. 2).
109
110 Our results were unaffected by intra-generic variation in seed mass for two reasons.
111 First, there was no systematic bias in intra-generic variation across the phylogenetic
112 tree. We detected no correlation between the mean and the coefficient of variation
113 (CV) for seed mass of each genus (Extended Data Fig. 3, PGLS: $F_{1,131}$ = 0.67, p-value
114 = 0.416). Second, we could disregard intra-generic variation in rates of seed mass
115 evolution as influencing our results as we found that genera with larger variation in
116 seed size did not have different macroevolutionary dynamics (Extended Data Fig. 4,
117 Spearman's ρ_λ = -0.05, p-value = 0.114; ρ_μ = -0.03, p-value = 0.406; ρ_r = -0.067, p-
118 value = 0.058).
119
120 Our study supports the idea that variation in seed mass can explain disparity in
121 diversification across the angiosperm ToL by influencing dispersal and habitat
122 interactions. We specifically found that smaller-seeded genera had faster speciation
123 rates. This may be because they are capable of dispersing over larger distances [3] that
124 can result in isolated populations and eventually lead to speciation [18]. Large dispersal
125 distances may be especially important for isolation to occur in continuous habitats.

126 However, the positive effect on speciation from long-distance dispersal may be
127 dampened in highly fragmented landscapes where isolation is more common over
128 shorter distances [19]. Dispersal syndromes may also alter the effect of seed size on
129 speciation. Species with larger seeds are generally associated with biotic dispersal that
130 distributes seeds over greater distances than wind or gravity dispersal [7]. However,
131 broad scale predictions on the effects of dispersal syndromes on diversification may
132 be inaccurate, since the former depend on landscape connectivity [20] and can
133 sometimes be inconsistent, e.g. a wind-dispersed seed might be transported by an
134 animal. Detailed contextual data will be necessary to expand upon the mechanisms
135 underlying our findings. We also found that smaller-seeded genera had higher rates of
136 extinction, possibly due to smaller nutritional reserves that constrain establishment,
137 environmental tolerance and access to limiting resources [21].
138
139 Seed mass is associated with other traits that can affect diversification, but there is
140 little evidence that these better explain our observed correlations. For example,
141 genome size positively correlates with seed mass [22] and faster rates of genome size
142 evolution have been linked to increased speciation in angiosperms [23]. Shorter,
143 smaller-seeded plants also tend to have faster life cycles, which may accelerate
144 mutation rates [24] and promote diversification [25]. By comparing the effects of genome
145 size and life cycle across a subset of >900 genera in our dataset, we found that only
146 the distinction between strictly annual versus perennial genera influenced
147 macroevolutionary dynamics in a similar way as seed size (Supplementary
148 Information, Extended Data Fig. 5). Unlike other traits [23] both absolute seed size as
149 well as its rate of change were correlated with speciation and extinction. Thus, while
150 we cannot exclude unobserved traits as drivers of diversification [26], we argue that seed

151 mass plays a central role in plant life history both on its own and by integrating across
152 traits that should predictably shape macroevolutionary dynamics (Extended Data Fig.
153 6).
154
155 Our finding that high rates of phenotypic change correlate with diversification (Fig.
156 2b) has recently been observed in other taxonomic groups [4,11], but never across the
157 whole of the angiosperm ToL. Accelerated morphological evolution may allow
158 radiating lineages to occupy more complex adaptive landscapes [27]. Similarly, species
159 with greater rate of change in their seed mass (i.e., higher evolvability) could shift
160 between adaptive peaks or develop reproductive barriers more rapidly. However,
161 current methods do not allow us to distinguish whether speciation is responding to
162 morphological change or vice versa when reconstructing 250 million years of
163 evolutionary history [4].
164
165 The approach applied here can help to unravel the processes responsible for
166 generating large-scale asymmetries in biodiversity. It also offers the potential to test
167 how widely-varying traits influence other aspects of the evolution and adaptation of
168 flowering plants (e.g. [15]). Clade-specific exceptions arising from local interactions
169 with non-focal traits [29] and specific spatio-temporal contexts will undoubtedly interact
170 with broad-scale macroevolutionary patterns and may modulate the effects of seed
171 mass on diversification. Regardless, our results clearly demonstrate that seed size, and
172 its rate of change, drive speciation and extinction and help to explain why some
173 clades are much more species-rich than others.
174
175 **Methods**

176

177 *Seed mass and phylogenetic dataset*

178 Seed mass data for 31,932 species were obtained from the Royal Botanic Gardens
179 Kew Seed Information Database [30]. Species names were standardised with The Plant
180 List (TPL) nomenclature [31] and cleaned using the *Taxonstand* R package [32]. Further
181 processing at the genus-level was carried out with the *taxonlookup* R package [33],
182 which is a complete genus-family-order mapping for vascular plants that draws from
183 TPL, the Angiosperm Phylogeny website [34] and a higher-level manually curated
184 taxonomic lookup [15]. Seed mass mean values for each genus were calculated for a
185 total of 4763 genera.

186

187 We used the most comprehensive phylogenetic tree for land plants [15,35] that comprises
188 31,389 species. Taxonomic information for our phylogenetic tree was run through
189 *Taxonstand* and *taxonlookup* as described above to make it as comparable as possible
190 to the seed mass dataset. Monophyly of the genera in the tree was assessed using the
191 *Monophy* package [36] The initial estimate was 16%, but we removed taxa classified as
192 outliers by *Monophy* (i.e., taxa that lay outside of an established "core clade" for each
193 genus) and this resulted in only 8% of the genera not being monophyletic. The *Phyndr*
194 [33] package was then used to generate a genus-level tree with as much overlap as
195 possible between the phylogeny and the trait data. The final phylogenetic tree
196 included representatives from 303 plant families (Extended Data Fig. 7).

197

198 *Diversification and phenotypic evolution analysis*

199 Speciation, extinction and net diversification rates and rates of seed size evolution
200 were estimated separately on the phylogeny using BAMM version 2.5.0 [16]. BAMM

201 models shifts in macroevolutionary regimes across a phylogenetic tree using
202 reversible-jump Markov chain Monte Carlo (rjMCMC) sampling. Initial prior settings
203 were calculated with the *setBAMMpriors* function in *BAMMtools* [37], and the
204 *expectedNumberOfShifts* parameter was set at 25 and 10 for the speciation/extinction
205 and trait evolution analyses, respectively. We incorporated non-random incomplete
206 sampling information following BAMM protocols by calculating the proportion of
207 genera sampled inside each family and estimated the backbone sampling as the
208 overall proportion of sampled genera. *Taxonlookup* was used as a reference for these
209 calculations.
210
211 All analyses were run for 50 million generations. We verified convergence by plotting
212 chain traces and ensuring that the effective sample sizes of all relevant parameters
213 exceeded 200. The first 10 million generations were discarded as burn-in.
214
215 *Correlation of diversification and trait evolution*
216 We used STRAPP to test for multiple associations between macroevolutionary
217 dynamics and each of seed mass (using genus mean values as character states) and
218 seed mass rate of evolution (using seed mass evolutionary rates at the tips of the
219 phylogeny as character states). STRAPP compares the correlation between a focal
220 trait and a macroevolutionary parameter (λ, μ or r) to a null distribution of
221 correlations. The null correlations are generated by permuting the evolutionary rates
222 in the tips of the phylogenetic tree while maintaining the location of rate shift events
223 in the phylogeny. In each case, we calculated the absolute difference between the
224 observed correlation of the macroevolutionary rate and the trait state and the null
225 correlation obtained by the structured permutations across 5000 samples from the

BAMM posterior. The reported p-value was the proportion of replicates where the null correlation coefficient was greater than the observed correlation. We found a low type I error associated with our STRAPP correlation analysis (p-value = 0.11, Extended Data Fig. 8).

Code availability

Scripts used to carry out the analysis described in the paper and generate the figures will be deposited in Github.

References

1. Davies, T. & Barraclough, T. in *Reconstructing the Tree of Life: Taxonomy and Systematics of Species Rich Taxa* **72,** 149–163 (2007).
2. Willson, M. & Traveset, A. in *Seeds: the ecology of regeneration in plant communities.* 85–110 (2000). doi:10.1146/annurev.es.13.110182.001221
3. Coomes, D. A. & Grubb, P. J. Colonization, tolerance, competition and seed-size variation within functional groups. *Trends Ecol. Evol.* **18,** 283–291 (2003).
4. Rabosky, D. L. *et al.* Rates of speciation and morphological evolution are correlated across the largest vertebrate radiation. *Nat. Commun.* **4,** 1958 (2013).
5. Crane, P. R., Friis, E. M. & Pedersen, K. R. The origin and early diversification of angiosperms. *Nature* **374,** 27–33 (1995).
6. Moles, A. T. *et al.* A Brief History of Seed Size. *Science (80-.).* **307,** 576–580 (2005).
7. Moles, A. & Leishman, M. in *Seedling ecology and evolution* (eds. Leck, M. A., Parker, V. T. & Simpson, R. L.) 215–235 (Cambridge University Press, 2008).
8. Salguero-Gómez, R. *et al.* Fast-slow continuum and reproductive strategies structure plant life-history variation worldwide. *Proc. Natl. Acad. Sci. U. S. A.* **113,** 230–5 (2016).
9. Kostikova, A., Salamin, N. & Pearman, P. B. The role of climatic tolerances and seed traits in reduced extinction rates of temperate polygonaceae. *Evolution (N. Y).* **68,** 1856–1870 (2014).
10. Coyne, J. A. & Orr, H. A. *Speciation*. (Sinauer Associates, 2004).
11. Price, S. L., Etienne, R. S. & Powell, S. Tightly congruent bursts of lineage and phenotypic diversification identified in a continental ant radiation. *Evolution* (2016). doi:10.1111/evo.12894
12. Pigliucci, M. Is evolvability evolvable? *Nat. Rev. Genet.* **9,** 75–82 (2008).
13. Ricklefs, R. E. & Renner, S. S. Species Richness Within Families of Flowering Plants. *Evolution (N. Y).* **48,** 1619–1636 (1994).
14. Rees, M. & Westoby, M. Game-Theoretical Evolution of Seed Mass in Multi-Species Ecological Models. *Oikos* **78,** 116–126 (1997).

15. Zanne, A. E. *et al.* Three keys to the radiation of angiosperms into freezing environments. *Nature* **506,** 89–92 (2014).
16. Rabosky, D. L. Automatic detection of key innovations, rate shifts, and diversity-dependence on phylogenetic trees. *PLoS One* **9,** e89543 (2014).
17. Rabosky, D. L. & Huang, H. A Robust Semi-Parametric Test for Detecting Trait-Dependent Diversification. *Syst. Biol.* **65,** 181–93 (2016).
18. Kisel, Y. & Barraclough, T. G. Speciation has a spatial scale that depends on levels of gene flow. *Am. Nat.* **175,** 316–34 (2010).
19. Claramunt, S., Derryberry, E. P., Remsen, J. V & Brumfield, R. T. High dispersal ability inhibits speciation in a continental radiation of passerine birds. *Proc. Biol. Sci.* **279,** 1567–74 (2012).
20. Givnish, T. J. Ecology of plant speciation. *Taxon* **59,** 1326–1366 (2010).
21. Leishman, M. R., Wright, I. J., Moles, A. T. & Westoby, M. in *The Ecology of Regeneration in Plant Communities* 31–58 (2000).
22. Beaulieu, J. M. *et al.* Correlated evolution of genome size and seed mass. *New Phytol.* **173,** 422–37 (2007).
23. Puttick, M. N., Clark, J. & Donoghue, P. C. J. Size is not everything: rates of genome size evolution, not C-value, correlate with speciation in angiosperms. *Proc. Biol. Sci.* **282,** 20152289– (2015).
24. Lanfear, R., Ho, S. Y. W., Love, D. & Bromham, L. Mutation rate is linked to diversification in birds. *Proc. Natl. Acad. Sci. U. S. A.* **107,** 20423–8 (2010).
25. Lanfear, R. *et al.* Taller plants have lower rates of molecular evolution. *Nat. Commun.* **4,** 1879 (2013).
26. Beaulieu, J. M. & O'Meara, B. C. Detecting hidden diversification shifts in models of trait-dependent speciation and extinction. *Syst. Biol.* syw022–(2016). doi:10.1093/sysbio/syw022
27. Lovette, I. J., Bermingham, E. & Ricklefs, R. E. Clade-specific morphological diversification and adaptive radiation in Hawaiian songbirds. *Proc. Biol. Sci.* **269,** 37–42 (2002).
28. Eldredge, N. & Gould, S. J. in *Models in Paleobiology* (ed. Schopf, T.) 82–115 (Freeman Cooper, 1972).
29. Donoghue, M. J. & Sanderson, M. J. Confluence, synnovation, and depauperons in plant diversification. *New Phytol.* **207,** 260–274 (2015).
30. Kew Seed Information Database. *Royal Botanic Gardens Kew* (2016). at <http://data.kew.org/sid/>
31. The Plant List. Version 1.1. (2013). at <http://www.theplantlist.org/>
32. Cayuela, L., Granzow-de la Cerda, Í., Albuquerque, F. S. & Golicher, D. J. taxonstand: An r package for species names standardisation in vegetation databases. *Methods Ecol. Evol.* **3,** 1078–1083 (2012).
33. Pennell, M. W., FitzJohn, R. G. & Cornwell, W. K. A simple approach for maximizing the overlap of phylogenetic and comparative data. *Methods Ecol. Evol.* n/a–n/a (2015). doi:10.1111/2041-210X.12517
34. Stevens, P. . Angiosperm Phylogeny Website. *Angiosperm Phylogeny Website. Version 12* (2012). at <http://www.mobot.org/MOBOT/research/APweb/>
35. Qian, H. & Jin, Y. An updated megaphylogeny of plants, a tool for generating plant phylogenies and an analysis of phylogenetic community structure. *J. Plant Ecol.* rtv047 (2015). doi:10.1093/jpe/rtv047
36. Schwery, O. & O'Meara, B. C. MonoPhy: A simple R package to find and visualize monophyly issues. (2015). doi:10.7287/peerj.preprints.1600v1
37. Rabosky, D. L. *et al.* BAMMtools: an R package for the analysis of

evolutionary dynamics on phylogenetic trees. *Methods Ecol. Evol.* **5,** 701–707 (2014).
38. Bennett, M. & Leitch, I. J. Plant DNA C-values database (release 6.0, Dec 2012). (2012). at <http://www.kew.org/cvalues/>

Acknowledgments

We thank D. A. Coomes, A. J. Helmstetter, T. Jucker and W. G. Lee for useful comments that helped improve the manuscript. J.I. and A.J.T. thank the Gatsby Charitable Foundation, Wellcome Trust and Newton Trust for funding. E.F.M was funded by the BBSRC DTP at the University of Cambridge.

Author contributions

J.I and A.S.T.P conceived the study. J.I. and E.F.M. performed the analysis. J.I. and A.J.T interpreted the analysis and wrote the manuscript. All authors edited the manuscript.

Author information

Scripts used to carry out the analyses described in the manuscript will be uploaded to Github. Correspondence and requests for materials should be sent to J.I. (ji247@cam.ac.uk)

Figure Legends

Figure 1. Seed mass and macroevolutionary rates across the angiosperm tree of life**.** Phylogenetic tree of 4105 genera of plants with seed mass, rate of seed mass change, and speciation (λ), extinction (μ) and net diversification (r) rates estimated by BAMM. Seed mass and rate data were standardised to Z-scores so that variation could

be directly compared.

Figure 2. Correlation of seed mass and seed mass rate with macroevolutionary dynamics. STRAPP correlations between speciation (λ), extinction (μ), and net diversification (r) and each of a) seed mass and b) rate of seed mass change. Coloured lines are Spearman correlations between λ (shown in red), μ (blue), r (purple) and each trait for one sample of the BAMM posterior distribution, bold line is the median. The insets show the density plots of the absolute difference between the observed and null correlation calculated across 5000 structured permutations of the evolutionary rates on the phylogenetic tree (myr = million years).

Extended Data Figure Legends

Extended Data Fig. 1. Posterior distribution of the number of rate shifts in BAMM for a) the speciation/extinction and b) phenotypic evolution analyses.

Extended Data Fig. 2. Phylogenetic tree of 4105 angiosperm genera with branches scaled by diversification rate as determined by BAMM. The colours of the branches indicate the rate of seed mass evolution, also estimated with BAMM.

Extended Data Fig. 3. Mean genus seed mass and coefficient of variation for 2613 angiosperm genera. The two variables were not significantly correlated (PGLS test, $F_{1,131} = 0.67$, p-value = 0.416).

369 **Extended Data Fig. 4.** Correlation of coefficient of variation in seed mass with
370 macroevolutionary dynamics. STRAPP correlations with speciation (λ), extinction (μ)
371 and net diversification (r) and the genus-level coefficient of variation in seed mass for
372 2613 angiosperm genera (i.e. each with seed mass data for >1 species in the Kew
373 SID). Correlations and insets calculated as in Fig. 2.
374
375 **Extended Data Fig. 5.** STRAPP correlations of diversification and phenotypic traits
376 for 938 angiosperm genera. The distribution of the absolute difference in the observed
377 correlation minus the null correlation is plotted for each trait. The coloured dotted
378 lines indicate the mean of that distribution, and the black dotted line indicates 0; a
379 distribution with mean = 0 would show no association between a focal trait and
380 speciation. STRAPP correlation of seed mass (shown in blue), C-value (shown in red)
381 and life cycle (shown in green) with a) speciation rate (λ), b) extinction rate (μ), and
382 c) net diversification rate (r).
383
384 **Extended Data Fig. 6.** Proposed effects of seed mass and other life history traits on
385 diversification (solid lines). Dashed lines indicate correlations between life history
386 traits. Numbers indicate reference where the link is proposed.
387
388 **Extended Data Fig. 7.** Phylogenetic tree of 303 angiosperm families with
389 representatives in this study. The red bars indicate the levels of sampling for each
390 family.
391
392 **Extended Data Fig. 8.** Type I error analysis. We estimated the type I error rate of our
393 analysis by simulating neutral traits on the angiosperm phylogenetic tree. We

394 performed 1000 simulations and then ran 1000 STRAPP tests with each simulated
395 dataset. We estimated the corresponding p-values for the association between traits
396 and diversification and calculated the type I error as the proportion of datasets where
397 a significant association (p-value < 0.05) was detected.
398
399 **Extended Data Fig. 9.** Mean genus seed mass of strict annual (n = 106) and perennial
400 (n = 832) genera. No significant difference between the means of the two groups was
401 found when phylogeny was accounted for (PGLS, $F_{1,937} < 0.01$, p-value = 0.972).
402
403 **Extended Data Table 1.** STRAPP correlations for 938 genera of angiosperms with
404 seed mass, genome size (i.e., C-value) and life cycle data. Macroevolutionary
405 parameters were obtained from an analysis with 4105 genera (see Main text).
406
407 **Supplementary Information**
408
409 Seed mass is central to a network of inter-correlated life history traits that can impact
410 diversification. Two such traits are genome size or plant C-value (measured as
411 picograms of DNA per haploid nucleus) and life cycle. We compared the correlation
412 between macroevolutionary parameters (λ, μ, r) and each of seed mass, C-value and
413 life cycle across a dataset of 938 plant genera (106 strictly annuals and 832 strictly
414 perennials). Genome content data and life cycle data were downloaded from the Plant
415 DNA C-values database [38]. Surprisingly, mean seed mass did not differ between
416 strictly annual and perennial plants when accounting for phylogenetic relationships
417 (Extended Data Fig. 9, PGLS: $F_{1,937} <0.01$, p-value = 0.972).
418

419 STRAPP requires large phylogenetic trees (n > 1000 species) to detect correlations [17].
420 In this reduced dataset, we ran STRAPP correlations for each focal trait with the
421 diversification parameters calculated from our 4105 genera BAMM analysis. We then
422 calculated the absolute differences in the observed and the null correlations between
423 the macroevolutionary parameters and seed mass, C-value and "annuality" (a binary
424 variable specifying whether the genus was annual or not). These differences were
425 similar and positive for seed mass and life cycle, while smaller for C-value (Extended
426 Data Fig. 5). However, with this reduced dataset, only life-cycle showed significant
427 correlations with macroevolutionary dynamics (Extended Data Table 1). These
428 correlations were weaker than those observed with seed mass in the main text (Fig. 2),
429 and considerably weaker than correlations with the rate of seed size evolution.
430

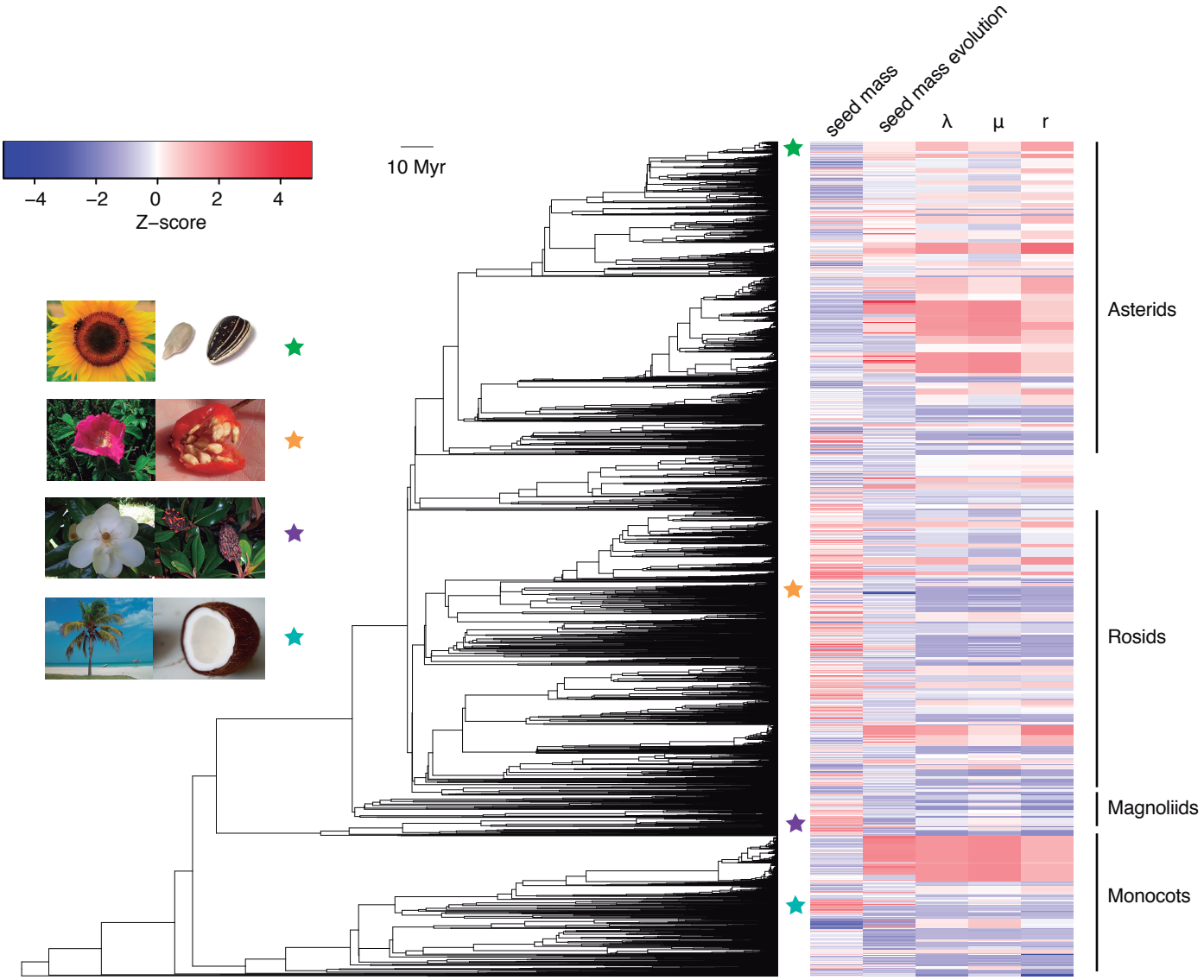
seed mass
seed mass evolution
λ
μ
r
−4
−2
0
2
4
Z−score
10 Myr
Asterids
Rosids
Magnoliids
Monocots

Figure 1

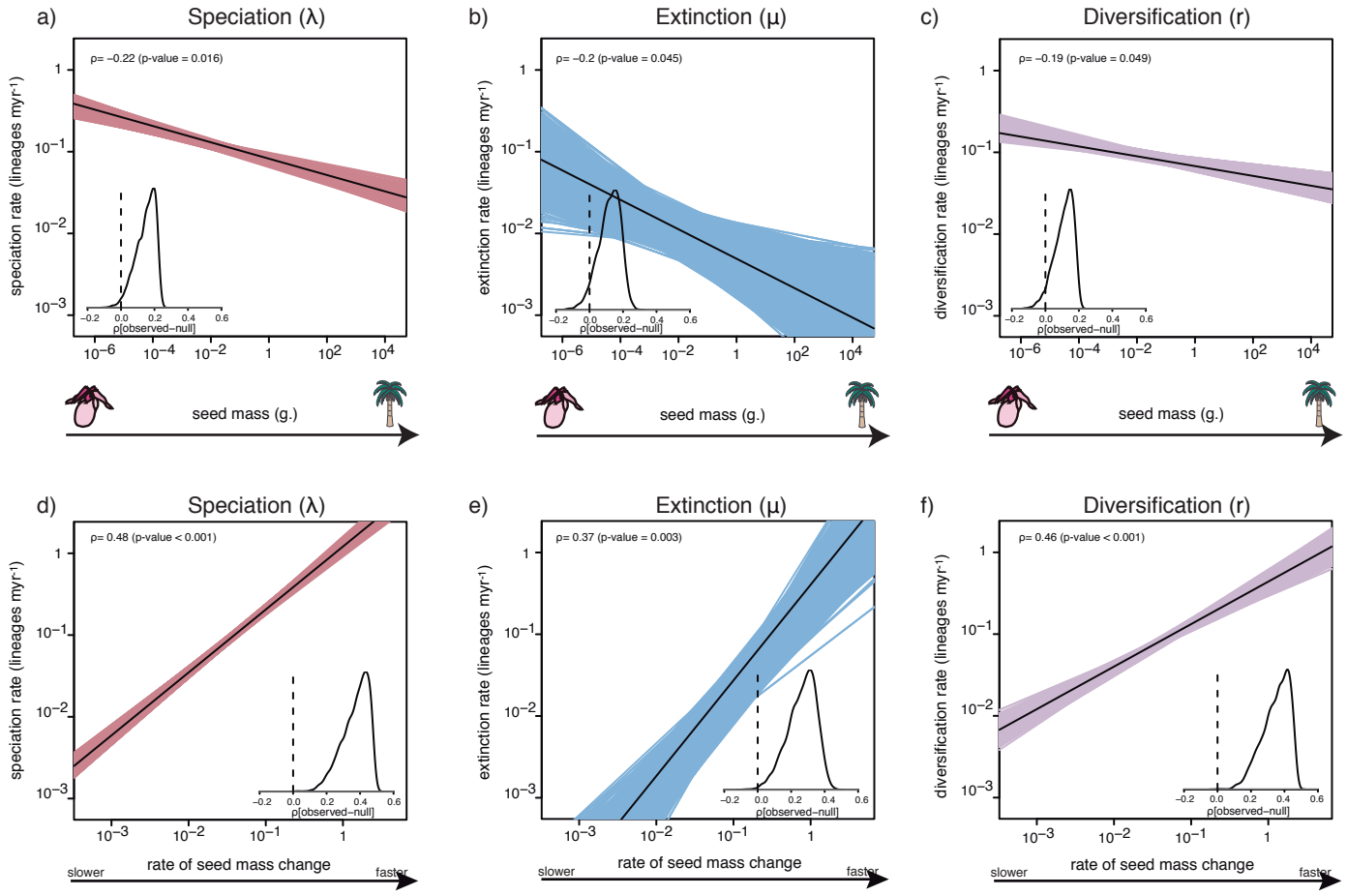
a)
Speciation (λ)
ρ= −0.22 (p-value = 0.016)
speciation rate (lineages myr⁻¹)
ρ[observed−null]
seed mass (g.)
b)
Extinction (μ)
ρ= −0.2 (p-value = 0.045)
extinction rate (lineages myr⁻¹)
ρ[observed−null]
seed mass (g.)
c)
Diversification (r)
ρ= −0.19 (p-value = 0.049)
diversification rate (lineages myr⁻¹)
ρ[observed−null]
seed mass (g.)
d)
Speciation (λ)
ρ= 0.48 (p-value < 0.001)
speciation rate (lineages myr⁻¹)
ρ[observed−null]
slower
rate of seed mass change
faster
e)
Extinction (μ)
ρ= 0.37 (p-value = 0.003)
extinction rate (lineages myr⁻¹)
ρ[observed−null]
slower
rate of seed mass change
faster
f)
Diversification (r)
ρ= 0.46 (p-value < 0.001)
diversification rate (lineages myr⁻¹)
ρ[observed−null]
slower
rate of seed mass change
faster

Figure 2