



HAL
open science

Pollination and mating traits underlie diverse reproductive strategies in flowering plants

Andrew Helmstetter, Marcos Méndez, Jürg Schönenberger, Concetta Burgarella, Bruce Anderson, Maria von Balthazar, Sylvain Billiard, Hugo de Boer, Johanne Cros, Pierre-Alexandre Delecroix, et al.

► To cite this version:

Andrew Helmstetter, Marcos Méndez, Jürg Schönenberger, Concetta Burgarella, Bruce Anderson, et al.. Pollination and mating traits underlie diverse reproductive strategies in flowering plants. 2024. hal-04793336

HAL Id: hal-04793336

<https://hal.science/hal-04793336v1>

Preprint submitted on 20 Nov 2024

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 - NoDerivatives 4.0 International License

Pollination and mating traits underlie diverse reproductive strategies in flowering plants

Andrew J. Helmstetter^{ab*}, Marcos Méndez^c, Jürg Schönenberger^d, Concetta Burgarella^e, Bruce Anderson^f, Maria von Balthazar^d, Sylvain Billiard^g, Hugo de Boer^h, Johanne Crosⁱ, Pierre-Alexandre Delecroix^j, Mathilde Dufay^k, John R. Pannell^l, Dessislava Savova Bianchi^l, Daniel J. Schoen^m, Mario Vallejo-Marinⁿ, Rosana Zenil-Ferguson^o, Hervé Sauquet^{pq}, Sylvain Glémin^{jn*}, Jos Käfer^{b*}

^aFRB-CESAB, Montpellier, France

^bISEM, Univ Montpellier, CNRS, IRD, Montpellier, France

^cArea of Biodiversity and Conservation, Universidad Rey Juan Carlos, 28933 Móstoles (Madrid), Spain

^dDepartment of Botany and Biodiversity Research, University of Vienna, Rennweg 14, 1030 Vienna, Austria

^eDepartment of Organismal Biology, University of Uppsala, Uppsala, 75236, Sweden

^fDepartment of Botany and Zoology, University of Stellenbosch, Private Bag X1, Matieland 7602, South Africa

^gUniversity Lille, CNRS, UMR 8198 – Evo-Eco-Paleo, 59000 Lille, France

^hNatural History Museum, University of Oslo, 0318 Oslo, Norway

ⁱUniversité de Lyon, Université Lyon 1, CNRS, Laboratoire de Biométrie et Biologie Evolutive UMR 5558, Villeurbanne, France

^jCNRS, Ecosystèmes Biodiversité Evolution (Université de Rennes), 35000 Rennes, France

^kCEFE, University Montpellier, CNRS, University Paul Valéry Montpellier 3, Ephe, IRD, Montpellier, France

^lDepartment of Ecology and Evolution, University of Lausanne, 1015 Lausanne, Switzerland

^mDepartment of Biology, McGill University, Montreal, QC, Canada

ⁿDepartment of Ecology and Genetics, Uppsala University, Uppsala, 752 36, Sweden

^oDepartment of Biology, University of Kentucky, 101 T.H. Morgan Building, Lexington, KY 40506 USA

^pNational Herbarium of New South Wales, Royal Botanic Gardens and Domain Trust, Sydney, New South Wales, 2000, Australia

^qEvolution and Ecology Research Centre, School of Biological, Earth and Environmental Sciences, University of New South Wales, Sydney, Australia

* **Correspondence:** Andrew J. Helmstetter, FRB-CESAB Institut Bouisson Bertrand, 5, rue de l'École de médecine - 34000 Montpellier. Email: andrew.j.helmstetter@gmail.com

Sylvain Glémin, CNRS, Ecosystèmes Biodiversité Evolution (Université de Rennes), 35000 Rennes, France. Email: sylvain.glemin@univ-rennes1.fr

Jos Käfer, ISEM, Univ Montpellier, CNRS, IRD, Montpellier, France. Email: jos.kafer@cnrs.fr

Authorship statement: SG and JK initiated the project; BA, SB, CB, HdB, MD, SG, JK, MM, SPO, JP, DR, HS, DS, JS, MVM, RZF, AH conceived the project; MvB, HS and JS provided data management; MvB, CB, JC, PAD, SG, JK, MM, DSB, HS and JS entered data; SG, AH, JK and HS performed analyses; AH visualized results; SG, AH, JK drafted the first version of the manuscript; all authors contributed to and approved of the final version.

Classification: Biological Sciences/Ecology

Keywords: angiosperms, floral morphology, reproductive systems, functional ecology, trait evolution

Data accessibility statement: The data set we compiled is available on the PROTEUS/eFLOWER website: <https://eflower.myspecies.info/publications>. Data and all scripts are available at the following github repository: <https://github.com/divers-it/rs-traitspace>.

1 **Abstract**

2 Reproductive and floral traits are strikingly diverse in angiosperms. However, general patterns
3 in how these traits are associated and how these associations shape species diversity are still poorly
4 known. To date, major trait associations, or plant ecological strategies, have largely been described
5 using vegetative traits and are underpinned by size and growth. However, these do not consider
6 traits relating to pollination, mating and sexual systems, which are critical to species' reproductive
7 success. Here, we characterize the strategies plants have evolved to achieve reproduction using 21
8 traits from an original set of 360 species sampled in 259 families across the angiosperm phylogeny.
9 While outcrossing rates were associated with a well-known trait axis involving plant size and growth
10 form, we revealed that pollination-related traits, including flower sex and floral reward, contribute to
11 an equally important axis of variation. Surprisingly, this pollination axis clearly sets species with
12 unisexual flowers apart from those with bisexual flowers, likely the result of selection for sexual
13 specialisation at the flower and individual levels. Three angiosperm-scale reproductive strategies
14 emerged, all containing considerable diversity that cannot be easily reduced, in contrast to strategies
15 derived from vegetative traits. Our study demonstrates that often-overlooked, pollination-related
16 traits represent an important dimension of plant trait variation that provides new insights into the
17 ecology and evolution of mating and sexual systems. The reproductive strategies outlined here
18 are the result of long-term evolutionary processes and ecological interactions between plants and
19 pollinators, highlighting the importance of considering both fields when disentangling the origins
20 of reproductive diversity.

21 **Significance statement**

22 Flowering plants (angiosperms) possess a spectacular diversity of flower morphologies, colours,
23 modes of pollination and systems of reproduction. These traits are known to have evolved together
24 through interactions with pollinators and have contributed to species diversification. However, we
25 lack a general view of how these traits are associated across the angiosperm tree of life. To this
26 end we sampled species representing all major angiosperm groups to uncover the main associations
27 among reproductive and life-history traits. We described an important pollination-related axis
28 of trait variation that is largely independent of a well-known size/lifespan axis, leading to the
29 identification of three broad reproductive strategies. Our study provides a framework to explore
30 how evolutionary history and ecology shape reproductive strategies at finer scales.

31 Introduction

32 Angiosperms, or flowering plants, are by far the most species-rich group of plants today. They
33 also present an exceptional diversity of modes of reproduction, with variation in mode of pollination
34 (from abiotic to biotic); variation in sex allocation and the sexual system (from hermaphrodites
35 with bisexual flowers to unisexual individuals); and variation in the mating system (from obligate
36 outcrossing to fully self-fertilizing). The combinations of traits that plants adopt along these three
37 dimensions are not evenly distributed across angiosperms - most species are biotically pollinated,
38 hermaphroditic and outcrossing (Barrett, 2002; Igc & Kohn, 2006; Renner, 2014). Furthermore,
39 it has long been recognised that different modes of reproduction are often associated, in one way
40 or another, to the morphological traits of flowers (Barrett, 2002; J. C. Vamosi et al., 2018). For
41 example, the flowers of wind-pollinated species, which have reduced petals, exerted stamens, no
42 nectar, etc., differ strikingly and predictably from those pollinated by animals. Unisexual flowers,
43 such as those of dioecious species, tend to be morphologically simple, small, and white or green,
44 in contrast to the often more ornate bisexual flowers of hermaphrodites. And the flowers of
45 self-pollinating species (which are small, produce little pollen and lack nectar) differ from the
46 more attractive and often highly rewarding flowers of their outcrossing counterparts.

47 Trait associations are the product of evolutionary processes. Natural selection can cause the
48 joint evolution of trait, as is the case of lifespan and mating system, a well-known association in
49 plants (Barrett et al., 1997) that leads to either long-lived outcrossers or short-lived selfers (Scofield
50 & Schultz, 2006; Lesaffre & Billiard, 2020). Alternatively, after a first trait has evolved, selection
51 pressures could favour a second one. For instance, specialised pollinators that more efficiently
52 deliver pollen would lead to higher outcrossing rates, but also put species at the risk of pollinator
53 failure. Species could thus evolve other mechanisms to ensure pollination, as is probably the
54 case with cleistogamy, the occurrence of closed and obligate self-pollinated flowers, which is often
55 associated with bilaterally symmetric (zygomorphic) flowers adapted to specialized pollinators (Joly
56 & Schoen, 2021). Traits can also affect speciation and extinction (cf Helmstetter et al., 2023) and
57 thus influence lineage selection, which can in turn affect how often certain trait combinations are
58 found in nature. This process is thought to cause the association of dioecy with other traits, such as
59 wind pollination or fleshy fruits (J. C. Vamosi & Vamosi, 2004), leading to higher extinction rates
60 when dioecy evolves in lineages that do not have these traits. Though macroevolutionary studies
61 of trait-dependent diversification have often focused on traits in isolation (Helmstetter et al., 2023)

62 it is likely that ‘key innovations’ that promote diversification result from a combination of traits
63 (Donoghue & Sanderson, 2015). The net effect of a trait thus depends on ecological and evolutionary
64 context, as no single trait confers evolutionary success in all cases (Anderson et al., 2023).

65 Reproductive traits are no exception as indicated by the variety of plant reproductive trait
66 associations found in nature. However, are these associations so much linked to the evolutionary
67 and ecological context of a clade that they are different in every case? Alternatively, can we discover
68 certain combinations of reproductive traits, or ‘reproductive strategies’, that consistently appear
69 across the angiosperm tree of life? Such questions have been difficult to address due to a lack of
70 an adequate sampling of reproductive traits across the phylogeny of flowering plants.

71 Instead, much research has focused on using traits related to growth, survival and reproduction
72 that are often present in large-scale databases (Díaz et al., 2016; Salguero-Gómez et al., 2016) to
73 describe the main axes of trait variation and identify trait combinations forming plant ecological
74 strategies. Following seminal work by Grime (1974), these traits are predicted to influence plant
75 survival according to the levels of stress, competition and disturbance an individual experiences,
76 and can be interpreted in terms of ecological processes, such as the colonization-exploitation
77 (r-K) continuum, the leaf economic spectrum (Díaz et al., 2016), or the fast-slow continuum
78 (Salguero-Gómez, 2017). Apart from vegetative traits (e.g. height, leaf area and mass), reproduction
79 is only characterized by reproductive output (e.g. seed mass and number) in such analyses.

80 Traits associated with pollination, mating and sexual systems, which play a role in the reproductive
81 processes that happen before seed production, have largely been ignored in the description of plant
82 ecological strategies (cf Munoz et al., 2016; E-Vojtkó et al., 2020) despite being crucial for the
83 reproductive success of a plant. They influence the way individuals exchange genetic material, and
84 thus have many demographic and genetic consequences (Anderson et al., 2023). For example, the
85 number and quality of seeds that are produced can be influenced by pollen limitation and inbreeding
86 (Ashman et al., 2004; Crnokrak & Roff, 1999) and could thus affect plants’ abilities to compete
87 and cope with stress (Craig & Mertz, 1994; Cheptou et al., 2000; Petrone Mendoza et al., 2018).
88 Likewise, pollination and reproduction affect genetic diversity and adaptive capacity, and thus most
89 likely play an important role in long-term population and species survival (Burgarella & Glémin,
90 2017). Thus, these traits may also play an important and underappreciated role in dictating the
91 ecological strategies of flowering plants.

92 Here, we bring together a range of traits related to different aspects of reproduction, life history
93 and ecology for a more synthetic understanding of plant form and function. We built a novel data

94 set of 21 traits for 360 species sampled from across the angiosperm tree of life to characterise major
95 reproductive strategies, i.e. sets of co-occurring traits that promote ecological and evolutionary
96 success. Furthermore, we investigate the extent to which reproductive trait variation is linked to
97 variation in other plant life-history characteristics. Unlike studies based on pre-existing databases
98 (e.g. Kattge et al., 2020) or datasets (e.g. from community surveys as in E-Vojtkó et al., 2022;
99 Lanuza et al., 2023), we designed our sampling to maximize coverage of major taxa in angiosperms,
100 including all families containing at least 100 species, and almost all orders. This sampling strategy
101 also allowed us to jointly explore the ecology and evolution of mating and sexual systems, the two
102 main components of plant reproductive systems.

103 **Materials & Methods**

104 **Data collection**

105 We collated trait data for angiosperm species using the PROTEUS collaborative database
106 (Sauquet, 2019). Mating system traits are the least well documented in the literature, so we
107 first selected a list of 207 species for which information about the mating system (e.g., outcrossing
108 rates, self-(in)compatibility) was available in the literature, and that already had some floral trait
109 data in PROTEUS. Next, we added dioecious species that already had some data in PROTEUS or
110 belonged to completely dioecious families, as the first list contained only hermaphroditic species.
111 Finally, we adjusted the species sampling by adding at least one species from each family with
112 more than 100 species, and adding more species for the most species-rich families (e.g., Asteraceae,
113 Orchidaceae) choosing species in the main clades of the family to best represent its diversity. This
114 lead to an initial set of 363 species.

115 We selected a list of traits based on prior knowledge of how reproduction-related and associated
116 traits influence evolutionary success (Helmstetter et al., 2023; Anderson et al., 2023). These traits
117 were selected primarily to encompass the main aspects of angiosperm reproduction and include
118 those related to mating system, sexual system, floral morphology (flower sex, ovary position, flower
119 colour, flower size, flower symmetry) as well as dispersal distance/mode and pollination mode.
120 We also included several vegetative traits related to growth form and lifespan, which can also
121 be related to reproduction. For each trait, detailed scoring instructions were provided; for traits
122 already in PROTEUS, we used the instructions from Sauquet et al. (2017) and Schönenberger
123 et al. (2020), while for newly added traits, we compiled instructions based on already available

124 guidelines (Perez-Harguindeguy et al., 2013; Cardoso et al., 2018). These instructions are available
125 in Appendix S1. Seed mass was added outside PROTEUS as the average value for each of the species
126 according to the Seed Information Database (Society for Ecological Restoration et al., 2023).

127 To compare our results with classical functional trait spaces, we also analysed a data set of
128 six plant traits (leaf area, leaf mass per area, leaf nitrogen per mass, diaspore mass, stem specific
129 density and plant height) from > 45,000 species (Díaz et al., 2016; Díaz et al., 2022). This data set
130 was filtered to leave only those species that had measurements for at least four traits ($n = 7968$)
131 to ensure distances could be calculated between all species pairs while increasing computational
132 feasibility and accuracy of distance calculations.

133 **Trait encoding**

134 To facilitate downstream analyses, we modified the initial trait encoding as recorded initially
135 to obtain a set of traits that could be used in our analyses (Table S1). For categorical traits, we
136 reduced the number of states to 2 to 4 for comparisons to be informative. In some cases we split
137 an initially complex trait into multiple different ones for easier interpretation of results (e.g., habit
138 was recoded into woodiness, climbing and aquatic). For quantitative traits, if several values were
139 available for a species/trait combination (either several measurements or indication of minimum
140 and maximum values), we took the mean of these. The outcrossing rate was transformed into a
141 categorical trait using three bins: selfing (< 0.2), mixed ($0.2 - 0.8$) and outcrossing (> 0.8), which
142 allowed us to combine species with a proper estimate of outcrossing rate with species for which a
143 simple, qualitative classification was available (“Phenotypic mating system” trait in PROTEUS).

144 We used two different ways of encoding the categorical traits. Either we had one variable
145 per trait with as many values as there are states in the trait, plus separate values for cases in
146 which the trait was polymorphic for a species (we refer to this encoding as the “original” data
147 set). Alternatively, categorical traits were encoded using a one-hot approach, where each category
148 of a trait is treated as a distinct binary variable (e.g., woody vs non-woody; Harris and Harris,
149 2013). While this may introduce some redundancy (e.g., most species that are herbaceous are not
150 woody and vice-versa), it is an alternative way of dealing with polymorphic states while keeping
151 the relations between values. For example, in the original encoding, a species that can be both
152 woody and herbaceous is assigned to a separate category with no explicit relation to other woody
153 and herbaceous species, while in the one-hot encoding, such a species is similar to both herbaceous
154 and woody species.

155 **Filtering, transformations and missing data imputation**

156 To limit the impact of missing data on our analyses, traits were removed from the original
157 and one-hot data sets if they contained more than 50% missing data in the original data set.
158 Likewise, species were removed if they contained more than 50% missing data. We log-transformed
159 quantitative traits to conform better to normality expectations, except for fusion of ovaries that is
160 coded as a proportion. We also scaled and centered these variables to limit biases when building
161 the trait spaces.

162 After filtering, many of the traits still contained missing data. We conducted imputation with
163 ‘missForest’ (Stekhoven & Bühlmann, 2012) to determine how this affected distances between
164 species and the trait space. We followed the approach outlined in Debastiani et al. (2021). Briefly,
165 a pairwise phylogenetic distance matrix (see below for how the phylogenetic tree was generated)
166 containing all species was decomposed to a set of eigenvectors using the function *PVRdecomp* from
167 the ‘PVR’ R package. The first 10 eigenvectors were added to the trait data as additional, complete
168 traits to conduct imputation of missing data with the *missForest* function.

169 **Trait spaces**

170 To build trait spaces we first calculated pairwise distance matrices among species using Gower’s
171 distance (Gower, 1971) with the function *daisy* from R package ‘cluster’ (Maechler et al., 2022).
172 Gower’s distance was used because it can deal with missing data and mixed data types (e.g.
173 categorical and continuous). We then performed principal coordinate analysis (PCoA) on the
174 resultant distance matrix with the *pcoa* function of the R package ‘ape’ (Paradis & Schliep, 2019)
175 to generate a set of orthogonal eigenvectors and their associated eigenvalues. For the one-hot data,
176 we used the *wcmdscale* function in the R package ‘vegan’ for the PCoA, and fitted the individual
177 traits on the resulting trait space with the function *envfit*. We also built an additional trait space
178 using the traits in the Díaz et al. (2022) data set. To compare how this functional trait space
179 differs from our own we extracted the 159 species common to both datasets and re-built the two
180 corresponding trait spaces.

181 We quantified the “quality” of dimensionality reduction of the resultant trait spaces using the
182 method outlined in Mouillot et al. (2021). Briefly, the difference between the initial distance matrix
183 and the distance matrix after dimensionality reduction using PCoA is examined. High-quality trait
184 spaces are those in which a reduced number of PCoA dimensions accurately represents initial

185 distances among species. Quality is quantified using the area under curve (AUC) metric for each
186 number of retained PCoA axes. This approach also provides an indication for how many axes are
187 sufficient to summarize the variation in the initial dataset.

188 PCoA is a linear dimension reduction approach that does not account for more complex non-linear
189 patterns. We thus also used an alternative dimensional reduction approach, Uniform Manifold
190 Approximation and Projection (UMAP, McInnes et al., 2020), to visualise non-linear patterns in
191 our data. UMAP is based on manifold learning techniques and allows the user to assess patterns at
192 local and global scales simultaneously, depending on the size of the neighbourhood ('n_neighbours')
193 chosen.

194 Clustering

195 We performed three different clustering approaches in order to account for variation in output
196 by different algorithms. First, we conducted clustering using a hierarchical approach with the
197 *hclust* function of the 'stats' R package and the ward.D2 algorithm (Murtagh & Legendre, 2014).
198 This approach uses a dissimilarity matrix (in this case calculated using Gower's distance) and
199 progressively groups all species into clusters, starting with a cluster per species and joining the most
200 similar clusters until a single cluster is reached. Second, we clustered species using k-prototypes
201 (Huang, 1998) implemented in the 'clustMixType' package (Szepannek, 2018). This method
202 is similar to k-means but allows for mixed data types. Euclidean distances are used between
203 quantitative traits and the number of matching categories is used for qualitative traits. Third,
204 we used partition around medoids (PAM) (Kaufman & Rousseeuw, 1990) as implemented in the
205 'cluster' R package (Maechler et al., 2022). This method is based on determining a set of medoids
206 (points) that represent the structure of the data. Each species is assigned to a medoid, with the
207 goal of minimizing the sum of the dissimilarities of each species to their closest medoid. Gower's
208 distance matrices were used as input for this approach.

209 We used different approaches to determine the number of clusters (k) depending on the clustering
210 method. For hierarchical clustering, values of k were calculated based on changes in within-clusters
211 sum of squares. For the remaining two methods (PAM and k-prototypes), we calculated 'robust'
212 groups of species - those that remained clustered together as values of k were increased from $k = 2$
213 to $k = 7$. Groups were only considered if they contained > 20 (PAM) or > 10 (k-prototypes)
214 species to facilitate downstream interpretation. Lastly, functional diversity indices were calculated
215 using the R package mFD (Magneville et al., 2022) to compare cluster characteristics.

216 **Phylogenetic signal, transition rates and ancestral state estimation**

217 We built a phylogenetic tree among our species using V.PhyloMaker2 (Jin & Qian, 2022). We
218 used the default ‘GBOTB.extended.TPL’ tree that is derived from a large phylogenetic tree of
219 all seed plants (Smith & Brown, 2018) and built the tree using the default approach described as
220 ‘scenario 3’ (Jin & Qian, 2022). Prior to building the tree we standardized genus and species epithets
221 using U.Taxonstand (Zhang & Qian, 2023) and a database derived from the World Checklist of
222 Vascular Plants (Govaerts et al., 2021) found at <https://github.com/nameMatch/Database>. Using
223 this tree we calculated the level of phylogenetic signal for each one-hot encoded trait to examine
224 how well phylogeny, or the relatedness of species, explains the distribution of trait values and
225 states. For categorical traits we used the δ statistic (Borges et al., 2019), a phylogenetic analog of
226 Shannon entropy. Significance was tested by randomly shuffling states 100 times, recalculating δ
227 and recording the percentage that were lower than the observed value for δ . For continuous traits
228 we used the *phylosig* function in the ‘phytools’ package (Revell, 2012) to calculate λ (Pagel, 1999)
229 and Blomberg’s K (Blomberg et al., 2003) with 1000 simulations to calculate p-values.

230 We then examined how transition rates varied among the different categorical traits in our data
231 set. Transition rates and ancestral states at nodes in the phylogenetic tree were calculated using
232 corHMM (Boyko & Beaulieu, 2021). We built a model of trait evolution for each one-hot encoded
233 trait independently. We estimated a transition rate (per million years) of change between the two
234 states using an equal rates (‘ER’) model with one rate category in which the transition rate is the
235 same in each direction, thus providing a single rate per binary trait that can be compared to others.
236 We also used cluster membership as states and built models of trait evolution in which transition
237 rates were free to vary (all-rates-different or ‘ARD’).

238 **Simulated data sets**

239 To determine how phylogeny influences trait space for our set of species and traits we simulated
240 trait data using the phylogenetic tree of our species with models of trait evolution. For continuous
241 traits we fitted Ornstein-Uhlenbeck (OU) models using the *fitContinuous* function in the R package
242 ‘geiger’ (Pennell et al., 2014) to estimate OU model parameters and root state values. For
243 categorical traits we fitted continuous-time Markov (Mk) models to generate transition rate matrices
244 and ancestral likelihoods for the root state. We then used the estimated parameters and the
245 phylogenetic tree to simulate new datasets with *rTraitCont* from the ‘ape’ R package (Paradis

246 et al., 2004) and *sim.history* from ‘phytools’ (Revell, 2012). Traits were simulated independently
247 and then combined into a single simulated dataset, from which we calculated distance matrices and
248 ran PCoAs, as above.

249 Results

250 After recoding and filtering, our final data set consisted of 21 traits (Table S1) for 360 species
251 and 13% missing data, down from 25% initially (Fig. S1). With 259 out of 416 families, and 61 out
252 of 64 orders (Fig. 1, Table S2; APGIV, 2016), our dataset included a broad range of angiosperm
253 diversity. When comparing distance matrices (Fig. S2) we found that imputing data only slightly
254 changed Gower’s distances between species, and so would likely have little impact on the following
255 results (Mantel statistic $r = 0.9133$). In the following we did not use the imputed data unless
256 explicitly stated. Using one-hot encoding had a minor effect on distances, which remained highly
257 correlated (Mantel statistic $r = 0.9466$).

258 Exploring reproductive trait space

259 From the distance matrix, we conducted a PCoA on our data sets to summarise and visualize
260 variation (Figs. 2 and 3). The first two PCoA axes explained 30 to 34% of the variation, depending
261 on the data set used. These first two axes were correlated with a combination of general life-history
262 and floral traits. The most important life-history traits were woodiness, lifespan, seed size and plant
263 height, and they were distributed along a diagonal in this 2D trait space. Many pollination-related
264 traits were more or less orthogonal to this diagonal: flower sex, floral reward, biotic vs. abiotic
265 pollination, showiness and plant sexual system (Fig. 3). The third PCoA axis was mainly associated
266 with dispersal (mode and distance), and the fourth with ovary position and flower symmetry (Fig.
267 S3). By calculating the correlation of the traits to the PCoA axes (Table S3) and to one another
268 (Fig. S4), we confirmed that there are two main groups of traits, one containing vegetative traits,
269 mating and dispersal, and another containing flower anatomy, pollination and sexual system.

270 We identified three main density peaks in the trait space formed by the first two PCoA axes.
271 The two most dense peaks corresponded to herbaceous species with bisexual flowers and woody
272 species with bisexual flowers (Fig. 2), while a less dense peak corresponded to woody species with
273 unisexual flowers. Species outside of these high density areas (e.g., at the top of Figure 2) tended
274 to be abiotically pollinated, herbaceous species.

275 **Comparison of reproductive and vegetative trait spaces**

276 We compared this trait space to the one obtained by Díaz et al. (2016), built with six traits
277 classically used in plant functional ecology (Díaz et al., 2022). First, we found that the species
278 in our dataset are scattered throughout the mainly vegetative trait space derived from Díaz et al.
279 (2016), as shown in Fig. 4a. Second, we found a weak correlation between the Gower's distances
280 obtained for the 159 shared species using the six functional ecological traits and the distances
281 obtained from the same set of species and the trait set used in this study (Mantel statistic $r =$
282 0.2827; Fig. S5).

283 Then, using the approach of Mouillot et al. (2021), we calculated statistics allowing comparison
284 of our trait space to others. The AUC criterion, which indicates how well the first n axes summarize
285 the total variation in the data set, indicated that to get a good representation of our trait space, we
286 must keep more dimensions than for the data set derived from Díaz et al. (2022) as shown in Fig.
287 4b. This is equally the case when comparing our results to those of the other trait spaces analyzed
288 by Mouillot et al. (2021). This means that, in our data set, traits are weakly redundant and most
289 of them contribute small but significant amounts of variation that cannot be reduced to variation
290 in other, more structurally important traits. We did not see notable differences in the trait space
291 quality analyses between our different data sets (original, imputed, one-hot; Fig. S6).

292 **Describing major reproductive strategies**

293 UMAP (Fig. 5a), which uses an alternative decomposition approach that allows the visualisation
294 of non-linear and local patterns, demonstrated a clear separation of different groups across the trait
295 space. Three major groups gradually separated in the first two axes as the size of the neighbourhood
296 (knn) decreased (Fig. S7). These groups corresponded to the woody bisexual, herbaceous bisexual,
297 and unisexual species groups identified using standard PCoA, the difference being that herbaceous
298 unisexual species were associated more closely to the woody unisexual ones in the UMAP trait
299 space.

300 The number of clusters (i.e. values of k) that could be selected ranged from two to seven
301 depending on the clustering method used and evaluation criteria. To tackle this subjectivity issue,
302 we examined how cluster membership changed as values of k were changed using Sankey plots (Figs.
303 S8 and S9). We then identified which species stayed consistently together despite a changing k ,
304 which we term 'robust groups'. Taking the results from the original data set as examples, we found

305 that the number of robust groups was eight for k-prototypes (Fig. S10) and seven for PAM (Fig.
306 5). The differences between the two methods were rather subtle, mainly the number of species that
307 did not belong to a robust group, which was greater for k-prototypes. These species tended to fall
308 between two robust groups (Fig. 5a).

309 The overarching pattern in our clustering tended to reflect the first two axes of our trait space
310 decompositions, that is, three different combinations of flower sex (or sexual system) and woodiness.
311 The majority of species belonging to each cluster were (1) herbaceous species with bisexual flowers,
312 (2) species with unisexual flowers (i.e. dioecious and monoecious, regardless of woodiness) and (3)
313 woody species with bisexual flowers, as illustrated by hierarchical clustering on the original data
314 set with three clusters (Figs. S11). Other traits that exemplify these clusters, and those that vary
315 within them can be seen in Figure S12. This structuring was present to at least some extent in the
316 results of each of our clustering approaches (see supplementary data for full cluster/robust group
317 membership). Those approaches that yielded more complex results tended to split the three major
318 clusters by other traits including dispersal distance/mode, pollination mode, mating system and
319 floral symmetry (e.g. Figs. 5b-k and S10). This substructure is clearly seen when examining the
320 distribution of robust groups in the UMAP representation of the trait space (Fig. 5a).

321 Functional diversity indices for the three clusters from our hierarchical clustering (Table S4)
322 revealed that species in cluster 2 (typically with unisexual flowers) were furthest from the centre of
323 the trait space while species in cluster 3 (woody with bisexual flowers) were the closest (functional
324 dispersion, FDis). Likewise, species in cluster 3 were also the most tightly packed in the trait
325 space (functional mean nearest neighbour distance, FNND, and functional mean pairwise distance,
326 FMPD) while cluster 2 and cluster 1 (the herbaceous species with bisexual flowers) were generally
327 further apart. Functional divergence (FDiv) was highest for cluster 1 indicating that species were
328 more often at the extremes (the edge of the cluster), and that there is still a lot of diversity in this
329 cluster. The unisexual cluster had the highest values of functional originality (FOri) and functional
330 specialisation (FSpe), which means its species were among the most distinct in the global species
331 pool. Species in cluster 3 had the lowest values of these indices. All in all, these indices confirm the
332 results inferred by visual inspection of the PCoA trait space (Fig. 2) with respect to the density
333 of the groups of species with bisexual flowers, while the species with unisexual flowers were more
334 widely scattered.

335 Trait space and phylogeny

336 Almost half of the qualitative traits (17/35 one-hot traits) showed significant levels ($p < 0.05$) of
337 phylogenetic signal, though the majority of these were weak (Table S5). Those qualitative traits
338 with the strongest phylogenetic signal were biotic/abiotic pollination, lack of floral reward and
339 aquatic habit. All quantitative traits had significant phylogenetic signal, with values of λ ranging
340 from 0.74 for flower size to 1.09 for fusion of ovaries (Table S6).

341 We fitted models of trait evolution and extracted the transition rates for each trait. Among the
342 traits that had the highest rates were biotic/abiotic dispersal mode and outcrossing mating system
343 (Table S7). Some of the least common traits such as aquatic habit, climbing habit and rare types
344 of symmetry had the lowest rates as they evolve infrequently.

345 We then repeated the ancestral state reconstruction with the original dataset and allowing all
346 rates to be different (ARD models) to produce transition rate matrices that were used to simulate
347 new data sets where all traits evolved independently. As expected, distance matrices from these
348 simulated data sets were not correlated to the real data sets (Mantel statistic $r = 0.003$) and
349 maximum distances were greater when using real data (Fig. S2). When we ran PCoA on simulated
350 data sets we found that the first three to four axes explained substantially less variation than any
351 of the real data sets (Fig. S13). Thus, at least the first three PCoA axes from the analysis of our
352 original data set are due to co-evolution of traits rather than neutral phylogenetic co-occurrence.

353 When reconstructed on the phylogeny with an ARD model, the three clusters from our hierarchical
354 analysis showed many transitions across angiosperms (Fig. 1). There were only few larger clades
355 for which all species belonged to the same cluster, and the largest among them did not contain more
356 than 10 species. Due to the high number of missing species and the sampling strategy (maximizing
357 the representation of angiosperm diversity) the estimated transition rates and inferred ancestral
358 states should not be directly interpreted. Interestingly however, the ancestor of the angiosperms
359 was inferred to be a woody species with bisexual flowers, which is in accordance with other studies
360 (Zanne et al., 2014; Sauquet et al., 2017).

361 Discussion

362 Pollination is a major source of plant trait variation

363 The search to identify trait-based strategies in the global distribution of plant biodiversity
364 represents a major goal in contemporary plant ecology and evolution. Previous work successfully
365 identified trait axes in terms of plant size, investment in biomass and reproductive output, which
366 are thought to be linked through trade-offs in optimization of resource use under differential stress,
367 given the physiological and morphological constraints of plant growth. Optimization and economics
368 rationales have also been applied to plant organs to define a ‘leaf economic spectrum’ in terms of
369 costly vs. cheap leaves, based on the ratio of leaf area and mass or nitrogen and carbon content
370 (Wright et al., 2004; Díaz et al., 2016). More recently, the characterisation of a ‘flower economic
371 spectrum’ sorts species in terms of the relative costs and longevity of their flowers (Roddy et al.,
372 2021). All these perspectives view reproduction as a function of seed production and thus ignore
373 strategy dimensions that pertain to pollination (cf E-Vojtkó et al., 2020).

374 Our study has extended this research by examining the distribution of reproductive traits in
375 terms of flowers, sexual systems and mating systems. First, these traits represent an important
376 source of variation not easily reducible to a small number of variables (Fig. 4b). As expected
377 based on the sheer diversity of flowers in angiosperms, the trait space has many dimensions and
378 is thus much more complex than one based on vegetative characteristics alone. Nevertheless, we
379 identify two major axes in this trait space (Fig. 3). One is the well-documented size/growth form
380 axis, correlated with outcrossing. A second, almost orthogonal axis, is related to the way pollen is
381 transferred, and includes flower sex, sexual system, pollination mode, floral reward, and to a lesser
382 extent, flower size. Strategies that include this second axis thus add further dimensions to trait
383 space variation beyond, for example, a ‘fast-slow’ axis (Salguero-Gómez et al., 2016).

384 The major reproductive strategies of flowering plants

385 Species are not equally distributed in the reproductive trait space (Fig. 2), and we can
386 distinguish three major reproductive strategies that have evolved frequently. Most species are
387 clustered around two density peaks in the bisexual part of the trait space, largely corresponding
388 to herbs and trees. Among trees, a distinct albeit less dense peak corresponds to species with
389 unisexual flowers, often wind-pollinated, and containing many dioecious species. These species
390 occupy a relatively large part of the trait space, despite this group being less species-rich. Among

391 herbs, wind-pollinated species occupy an even larger part of the trait space, at even lower densities.
392 These species often (but not uniquely) have unisexual flowers and many aquatic species are found
393 among them. They represent rather unique strategies that did not evolve frequently (and do not
394 consistently cluster together; Fig. S14).

395 Equally important for our understanding of plant reproduction is the observation that there still
396 is a large amount of variation within these strategies. The third dimension in our trait space was
397 associated with dispersal mode and distance, which thus seem largely independent from variation
398 in woodiness/size and pollination/flower sex. Furthermore, other floral characteristics, e.g., flower
399 symmetry, explain part of the variation within the strategies. This variation, which is secondary
400 at the angiosperm level, will likely become more important as one focuses on smaller clades or
401 specific floras, as some of the variation is reduced (clades with only trees or herbs, all animal-
402 or wind-pollinated, etc.). This is partly observed on the non-linear projection where sub-clusters
403 appear within at least two of the three main clusters (Fig. 5). Other species that are difficult to
404 assign to a cluster could have trait combinations that are favourable in particular contexts only.

405 Although we can distinguish major reproductive strategies that would not have arisen if traits
406 evolved independently from one another, we find that these strategies are truly multifactorial, and
407 it remains challenging to summarize them using a combination of two or three traits. Indeed, not all
408 unisexual species are wind-pollinated and the showy, biotically pollinated monoecious or dioecious
409 species cluster with the bisexual species. Examples in our dataset are *Dalechampia spathulata*
410 which has unisexual flowers united in showy bisexual inflorescences, but also other insect-pollinated
411 monoecious species such as *Cucurbita pepo* or *Akebia quinata*. Equally, grasses (Poaceae) with
412 bisexual flowers are often clustered with the unisexual species as they are wind-pollinated.

413 **Sexual and mating systems contribute differently to reproductive strategies**

414 While dioecy has often been viewed largely as an outcrossing mechanism that is easier to evolve
415 than (genetic) self-incompatibility (Barrett, 2013), we found it is associated with a number of
416 other pollination-related traits, and not primarily with plant size and lifespan, which are otherwise
417 associated with outcrossing. Indeed, according to our results, monoecy and dioecy have very similar
418 trait associations despite the fact that monoecy doesn't impede selfing. This suggests that flower
419 unisexuality, including both monoecy and dioecy, mainly evolves through ecological interactions
420 with pollinators, or the absence thereof (cf Bawa, 1980; Charlesworth, 1993). Resource allocation
421 models are capable of taking such interactions into account (e.g Charnov et al., 1976; Sato, 2002),

422 but, just like models based on outcrossing dynamics (Charlesworth & Charlesworth, 1978), they
423 mostly consider hermaphroditism and monoecy as similar strategies. Transitions between monoecy
424 and dioecy seem to occur rather frequently (Renner, 2014; Goldberg et al., 2017), which is plausible
425 if both occur in otherwise similar species as we find here. However, it is not clear what factors
426 might favour monoecy in some cases and dioecy in others. The rarity of dioecy in comparison with
427 hermaphroditism has often been attributed to the specific handicaps of having unisexual individuals
428 (cf Käfer et al., 2017) but these do not apply to monoecy which occurs in similar frequencies. Our
429 results thus point out the need for more detailed studies on the differences between these sexual
430 systems.

431 By contrast, the variation from selfing to outcrossing is not consistently linked to floral traits.
432 Indeed, selfing is known to occur primarily in annual species, while the vast majority of trees
433 reproduce through outcrossing. While the so-called ‘selfing syndrome’ might be readily observed
434 among closely related species, with a reduction in corolla, stamen or pistil size (Sicard & Lenhard,
435 2011), our sampling does not allow the detection of such covariation at the angiosperm scale. A
436 transition to predominant selfing can arise in very different pollination contexts, e.g., wind-pollinated
437 grasses (Burgarella et al., 2023), small-flowered herbs with generalist pollinators (Sicard et al.,
438 2011), or in groups where specialist pollination syndromes have evolved (Rose & Sytsma, 2021).
439 Furthermore, selfing is not expected to persist at long evolutionary timescales, as it is often
440 associated with higher extinction rates (Goldberg et al., 2010). The scope for co-evolution of
441 multiple traits is thus rather limited and we expect that traits associated to selfing are specific to
442 each clade.

443 **Integrating floral and mating traits in the study of plant functional diversity**

444 It is increasingly recognized that floral traits should be considered when characterizing plant
445 functional diversity. Here we aimed to incorporate as much diversity as possibly by working with
446 an original dataset chosen to represent all major angiosperm clades. This is complementary to
447 the approach taken in other studies dealing with specific floras or datasets, and explains some of
448 the differences with previous results. For example, E-Vojtkó et al. (2022) studied two datasets
449 of European species (central Europe and Alps), and found that floral trait variation is largely
450 independent of variation in vegetative traits, a conclusion similar to ours. They did not seem to
451 find any clusters or density peaks in their trait spaces, which might be due to the exclusion of trees
452 and grasses in their analyses, and acknowledge that variation is reduced at the European scale they

453 studied. Similarly, Lanuza et al. (2023) analyzed plant reproductive strategies in the light of their
454 interactions with pollinators, and thus excluded wind-pollinated species. Again, no clear clusters
455 were found, but in agreement with our results, variation in selfing and outcrossing was only weakly
456 correlated with variation in flower traits. These studies, together with ours, confirm that floral and
457 pollination-related traits vary in ways that are not reducible to vegetative and seed characteristics.

458 A challenge for the inclusion of floral traits in large-scale evolutionary and ecological studies
459 seems to be their lack of availability in databases. The largest plant database today, TRY,
460 contains limited data about flowers as compared to vegetative and seed traits (Kattge et al., 2020).
461 Many studies and floras contain such data, but collecting them requires botanical knowledge in
462 order to correctly interpret the terminology. Although floras are increasingly available online,
463 they still represent a small fraction of known plant species, with tropical floras typically being
464 underrepresented (cf Römer et al., 2023). We argue that publicly available data with well-described
465 trait standards are the most convenient way to ensure that datasets can be combined and extended.

466 The inclusion of floral and pollination-related traits in the description of the plant functional
467 diversity is necessary to improve our understanding of ecosystem functioning. Indeed, vegetative
468 functional diversity has proven instrumental in testing theories of diversity (e.g. Lamanna et al.,
469 2014; Schuldt et al., 2019). However, interactions with pollinators have also been identified as
470 playing a role in the maintenance of diversity in plant communities (Wei et al., 2021) and the
471 decrease of pollinator abundance can destabilize the mechanisms promoting species coexistence
472 (Johnson et al., 2022). As we are only starting to standardize floral trait data and make them
473 available, we still have a very incomplete picture about how they influence pollinator diversity and
474 abundance, and how these feedback on plant community composition. However, such understanding
475 is urgently needed as pollinators are declining rapidly in many agricultural and semi-natural
476 landscapes, and this angiosperm-wide study provides a framework for future studies.

477 **Acknowledgements**

478 This research is supported by the Fondation pour la Recherche sur la Biodiversité (FRB)
479 through the CESAB project ‘DiveRS’. We thank the DiveRS members Sally Otto and Denis Roze
480 for constructive discussions and feedback on the manuscript. We thank Nicolas Casajus, Nicolas
481 Loiseau and Marion Chartier for advice on data analyses, Solenn Sarton for contributing data, and
482 Maud Calmet for the organisation of the workshops of the DiveRS group. We thank the University
483 of Vienna for hosting the PROTEUS database and eFLOWER server.

484 **Figures**

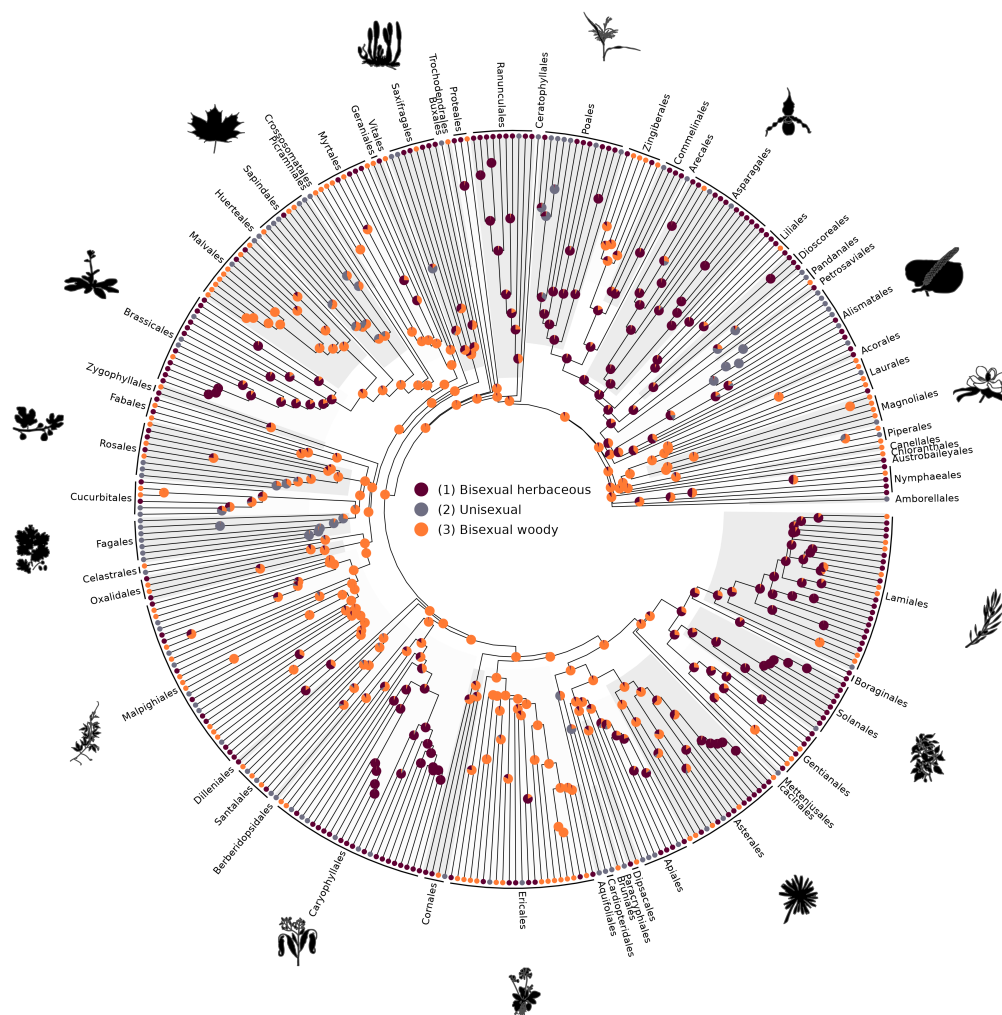


Figure 1: Phylogenetic tree representing the 360 species used in our analyses. Tips are labelled with circles indicating which group they were assigned to in our hierarchical clustering analysis ($k = 3$). Broadly, the three groups are summarised as: (1) Herbaceous species with bisexual flowers (2) Species with unisexual flowers and (3) Woody species with bisexual flowers. Note that these trait combinations are used to represent the majority of species in these clusters, but that not all species in these clusters possess these traits. Figure S12 shows the frequency of states per qualitative trait and the distribution of quantitative trait values for species in each cluster. An associated principal coordinate analysis (PCoA) can be found in Figure S11. Pie charts at nodes depict the results of an ancestral state estimation analysis using an all-rates-different (ARD) model using cluster assignment as the trait. Pies are coloured based on the ancestral state probability of each node. Orders that are represented in our set of species are highlighted around the outside of the tree alongside a selection of species' silhouettes (which are not necessarily of species in our data set) from phylopic.org.

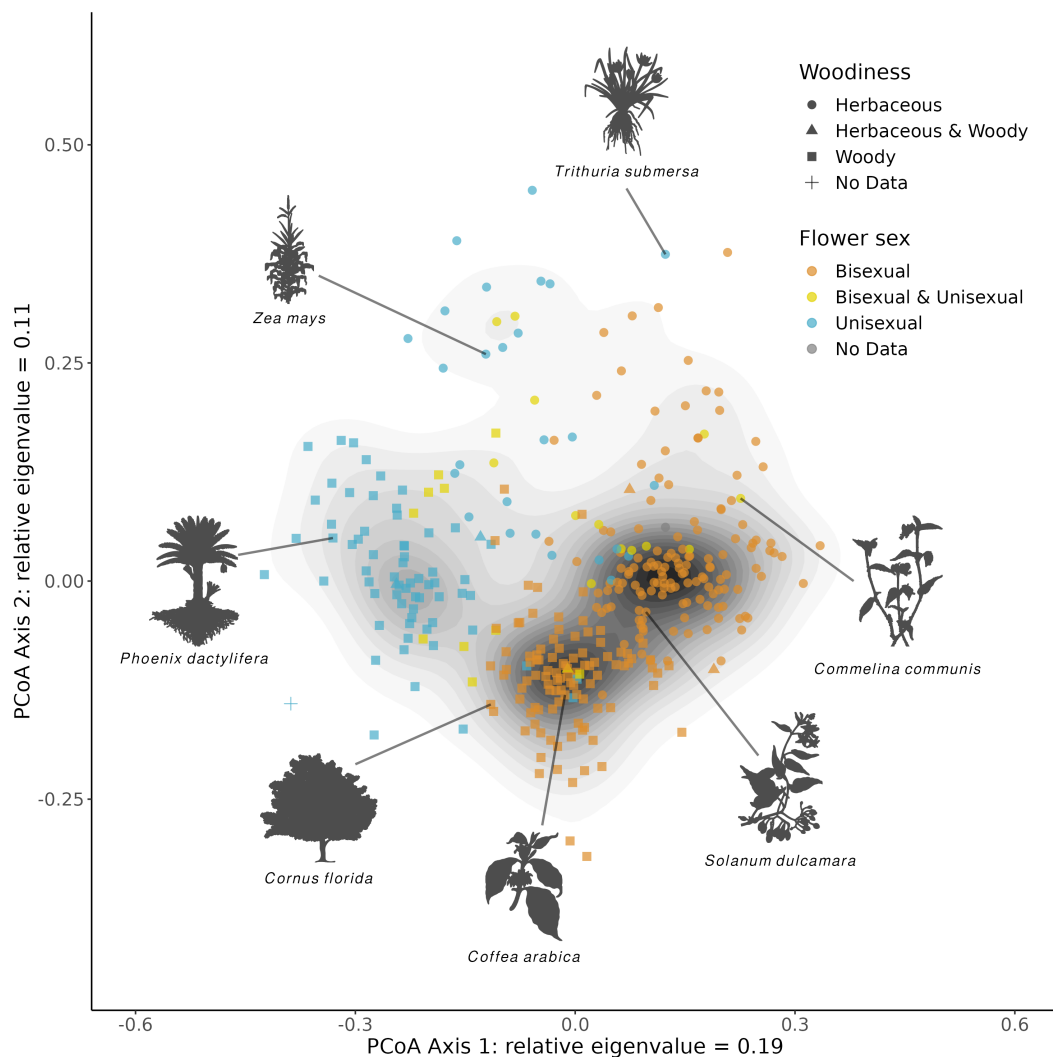


Figure 2: Location of 360 angiosperm species on the first two axes of our principal coordinate analysis (PCoA) on the original data set. These two axes represent approximately 30% of the variation in the data set. Points are coloured by the flower sex trait and shapes represent the woodiness trait. Polygons are used to show the density of points (= species), with darker sections indicating higher density. Silhouettes of various species found in our data set (taken from phylopic.org) are shown with lines linking them to their position in the trait space.

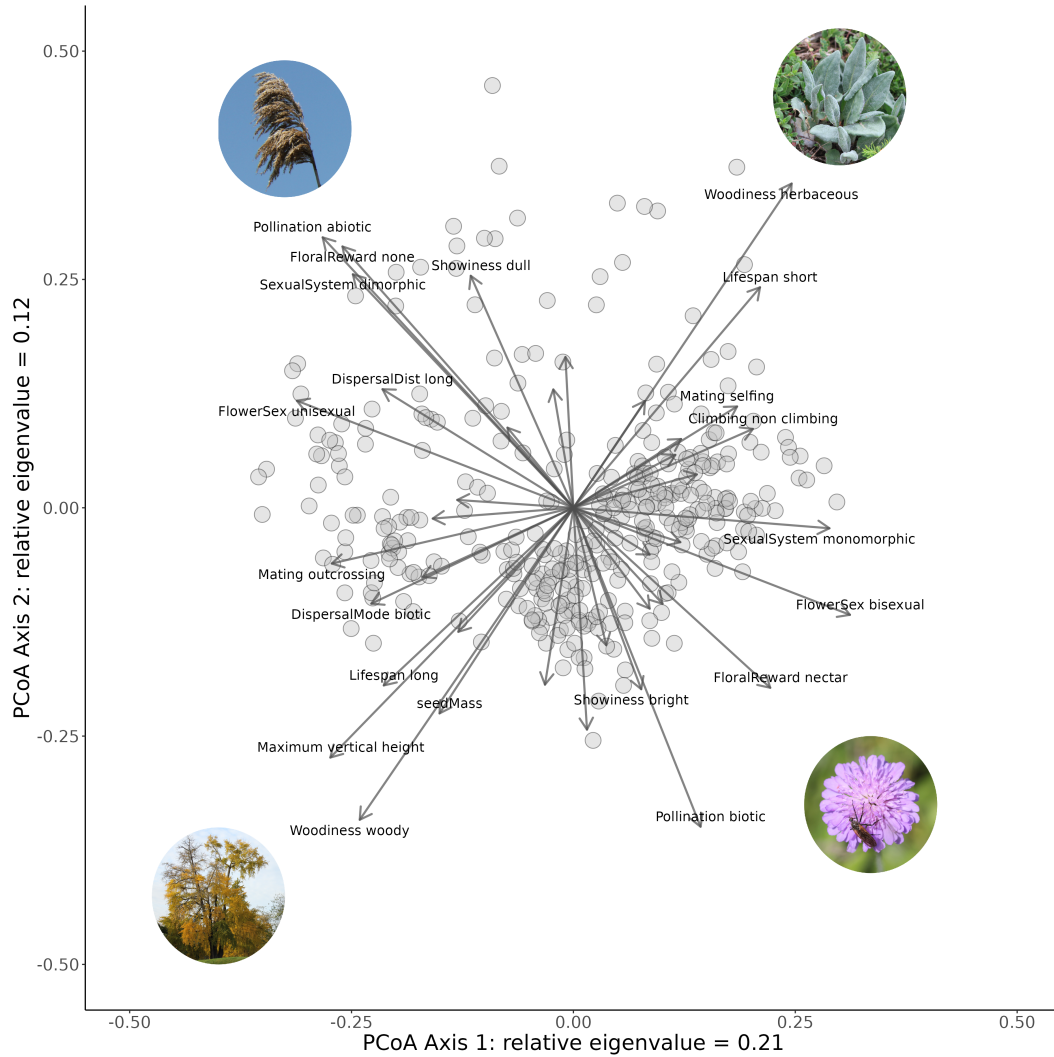


Figure 3: Principal coordinate analysis performed on the one-hot encoded data set. The first two axes represent approximately 33% of variation. Using one-hot encoding allowed us to plot arrows indicating how different states and traits affect the trait space, as well as which traits are acting in a similar manner. The length of the arrows indicates the strength of the effect of the state and labels were added to arrows representing those states with large effects. Images representing aspects of the trait space can be found in the four corners of the plot.

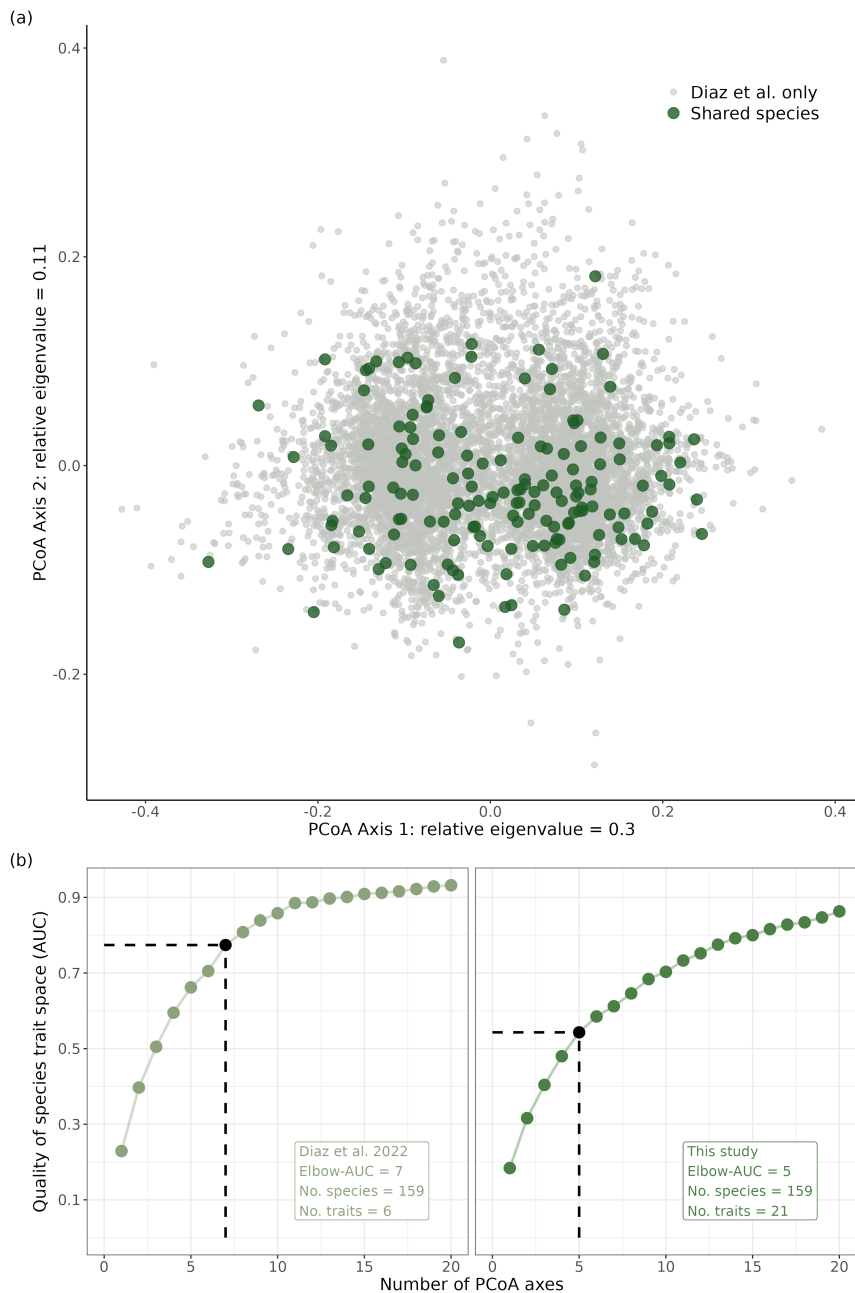


Figure 4: Comparison of the species sampled in the current study and the obtained trait space with species and trait spaces derived from Díaz et al. (2022). Panel (a) shows a scatterplot of a subset of approximately 8,000 species from Díaz et al. (2022) on the first two axes of a principal coordinate analysis (PCoA). Species that overlap with those in this study are highlighted in green. Panel (b) shows two line plots depicting how trait space quality (measured using AUC) changes as PCoA axes are added. Both graphs are made using the same set of 159 species but with trait data from Díaz et al. (2022) (left) and this study (right). Dashed lines show the number of axes to be kept, as determined by the elbow method, when optimising the trade-off between trait space quality and operationality. See Mouillot et al. (2021) for further details.

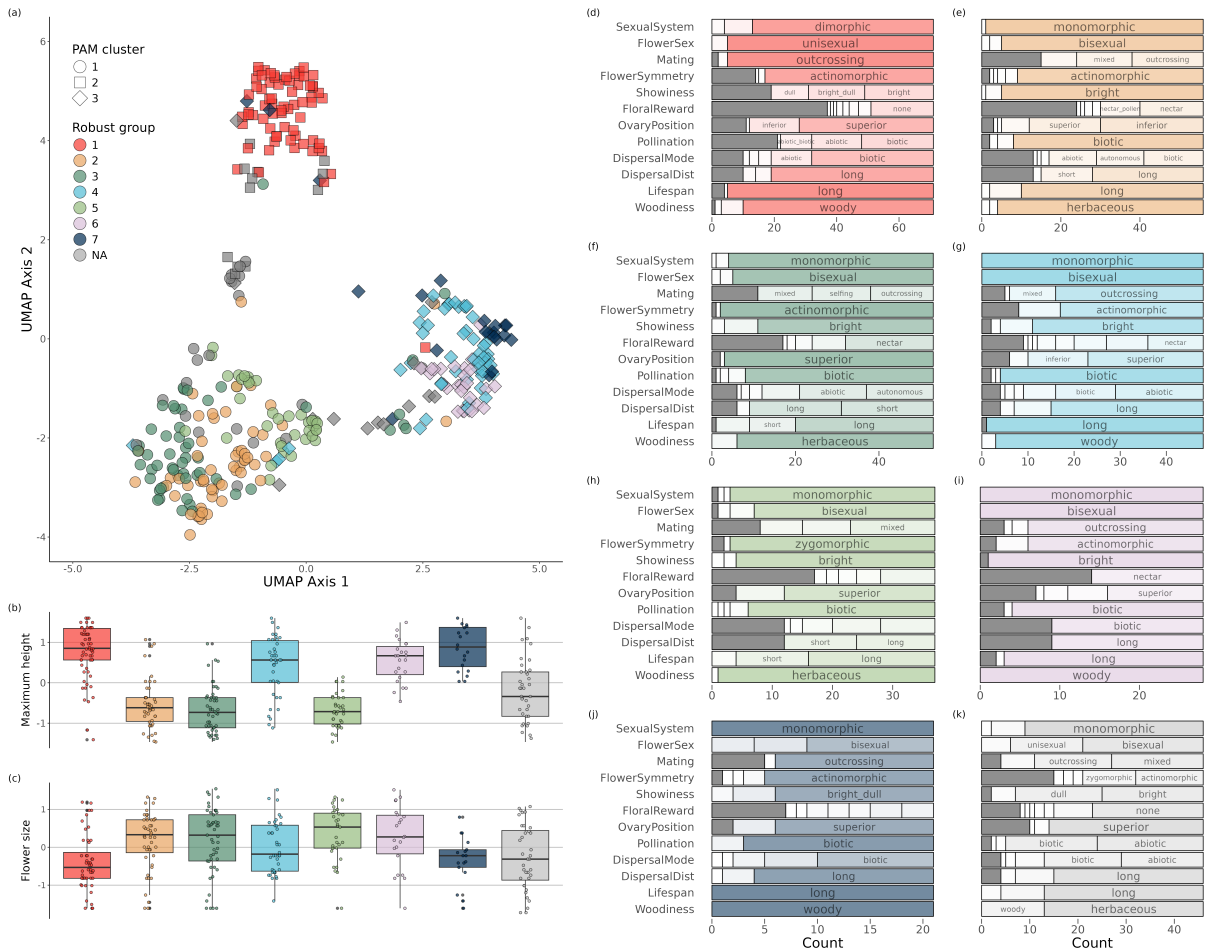


Figure 5: Characterizing robust groups identified using the Partitioning Around Medoids (PAM) clustering approach. Seven different robust groups were defined, and a final group was made with those species that did not fall into a robust group. Panel (a) shows the distribution of species along the first two axes of a UMAP (uniform manifold approximation) analysis with a neighbourhood size of 10. Points are coloured by robust group, and their shape indicates their PAM cluster assignment ($k = 3$). Boxplots in the bottom left show the distribution (after log transformation and scaling) of values for two traits per cluster: (b) maximum vertical height and (c) flower size. Values of species within each cluster are overlain and outliers were removed for display purposes. The stacked barplots (panels (d-k)) show the frequency of states for 12 traits for each of the eight groups. Colours match those in the other panels and bars are split into sections depending on the frequency of each state. Sections representing those states with high frequencies are labelled and the dark grey sections correspond to missing data.

485 References

- 486 Anderson, B., Pannell, J., Billiard, S., Burgarella, C., Boer, H. d., et al. (2023). Opposing effects of
487 plant traits on diversification. *iScience*, *26*(4). <https://doi.org/10.1016/j.isci.2023.106362>
- 488 APGIV. (2016). An update of the Angiosperm Phylogeny Group classification for the orders and
489 families of flowering plants: APG IV. *Botanical Journal of the Linnean Society*, *181*(1),
490 1–20. <https://doi.org/10.1111/boj.12385>
- 491 Ashman, T.-L., Knight, T. M., Steets, J. A., Amarasekare, P., Burd, M., et al. (2004). Pollen
492 limitation of plant reproduction: Ecological and evolutionary causes and consequences.
493 *Ecology*, *85*(9), 2408–2421. <https://doi.org/10.1890/03-8024>
- 494 Barrett, S. C. H. (2002). The evolution of plant sexual diversity. *Nature Reviews Genetics*, *3*(4),
495 274–284. <https://doi.org/10.1038/nrg776>
- 496 Barrett, S. C. H. (2013). The evolution of plant reproductive systems: How often are transitions
497 irreversible? *Proceedings of the Royal Society B: Biological Sciences*, *280*(1765), 20130913.
498 <https://doi.org/10.1098/rspb.2013.0913>
- 499 Barrett, S. C. H., Harder, L. D., & Worley, A. C. (1997). The comparative biology of pollination and
500 mating in flowering plants. *Philosophical Transactions of the Royal Society of London. Series*
501 *B: Biological Sciences*, *351*(1345), 1271–1280. <https://doi.org/10.1098/rstb.1996.0110>
- 502 Bawa, K. S. (1980). Evolution of dioecy in flowering plants. *Ann. Rev. Ecol. Syst.*, *11*, 15–39.
503 <https://doi.org/10.1146/annurev.es.11.110180.000311>
- 504 Blomberg, S. P., Garland JR., T., & Ives, A. R. (2003). Testing for Phylogenetic Signal in Comparative
505 Data: Behavioral Traits Are More Labile. *Evolution*, *57*(4), 717–745. <https://doi.org/10.1111/j.0014-3820.2003.tb00285.x>
- 506
- 507 Borges, R., Machado, J. P., Gomes, C., Rocha, A. P., & Antunes, A. (2019). Measuring phylogenetic
508 signal between categorical traits and phylogenies. *Bioinformatics*, *35*(11), 1862–1869. <https://doi.org/10.1093/bioinformatics/bty800>
- 509
- 510 Boyko, J. D., & Beaulieu, J. M. (2021). Generalized hidden Markov models for phylogenetic
511 comparative datasets. *Methods in Ecology and Evolution*, *12*(3), 468–478. <https://doi.org/10.1111/2041-210X.13534>
- 512
- 513 Burgarella, C., Brémaud, M.-F., Hirschheydt, G. V., Viader, V., Ardisson, M., et al. (2023). Mating
514 systems and recombination landscape strongly shape genetic diversity and selection in wheat
515 relatives. *bioRxiv*, 2023.03.16.532584. <https://doi.org/10.1101/2023.03.16.532584>

- 516 Burgarella, C., & Glémin, S. (2017). Population genetics and genome evolution of selfing species.
517 In J. Wiley & L. Sons (Eds.), *Els* (1st ed., pp. 1–8). Wiley. [https://doi.org/10.1002/](https://doi.org/10.1002/9780470015902.a0026804)
518 [9780470015902.a0026804](https://doi.org/10.1002/9780470015902.a0026804)
- 519 Cardoso, J. C. F., Viana, M. L., Matias, R., Furtado, M. T., Caetano, A. P. D., et al. (2018). Towards
520 a unified terminology for angiosperm reproductive systems. *Acta Bot. Bras.*, *32*(3), 329–348.
521 <https://doi.org/10.1590/0102-33062018abb0124>
- 522 Charlesworth, D. (1993). Why are unisexual flowers associated with wind pollination and unspecialized
523 pollinators? *The American Naturalist*, *141*(3), 481–490.
- 524 Charlesworth, D., & Charlesworth, B. (1978). Population-genetics of partial male-sterility and
525 evolution of monoecy and dioecy. *Heredity*, *41*(OCT), 137–153.
- 526 Charnov, E. L., Maynard Smith, J., & Bull, J. J. (1976). Why be an hermaphrodite? *Nature*,
527 *263*(5573), 125–126. <https://doi.org/10.1038/263125a0>
- 528 Cheptou, Imbert, Lepart, & Escarre. (2000). Effects of competition on lifetime estimates of inbreeding
529 depression in the outcrossing plant *Crepis sancta* (Asteraceae). *Journal of Evolutionary*
530 *Biology*, *13*(3), 522–531. <https://doi.org/10.1046/j.1420-9101.2000.00175.x>
- 531 Craig, D. M., & Mertz, D. B. (1994). Inbreeding effects on competition in *Tribolium*. *Researches on*
532 *Population Ecology*, *36*(2), 251–254. <https://doi.org/10.1007/BF02514941>
- 533 Crnokrak, P., & Roff, D. A. (1999). Inbreeding depression in the wild. *Heredity*, *83*(33), 260–270.
534 <https://doi.org/10.1038/sj.hdy.6885530>
- 535 Debastiani, V. J., Bastazini, V. A. G., & Pillar, V. D. (2021). Using phylogenetic information to
536 impute missing functional trait values in ecological databases. *Ecological Informatics*, *63*,
537 101315. <https://doi.org/10.1016/j.ecoinf.2021.101315>
- 538 Díaz, S., Kattge, J., Cornelissen, J. H. C., Wright, I. J., Lavorel, S., et al. (2016). The global
539 spectrum of plant form and function. *Nature*, *529*(7585), 167–171. [https://doi.org/10.](https://doi.org/10.1038/nature16489)
540 [1038/nature16489](https://doi.org/10.1038/nature16489)
- 541 Díaz, S., Kattge, J., Cornelissen, J. H. C., Wright, I. J., Lavorel, S., et al. (2022). The global
542 spectrum of plant form and function: Enhanced species-level trait dataset. *Scientific Data*,
543 *9*(1), 755. <https://doi.org/10.1038/s41597-022-01774-9>
- 544 Donoghue, M. J., & Sanderson, M. J. (2015). Confluence, synnovation, and depauperons in plant
545 diversification. *New Phytologist*, *207*(2), 260–274. <https://doi.org/10.1111/nph.13367>

- 546 E-Vojtkó, A., de Bello, F., Durka, W., Kühn, I., & Götzenberger, L. (2020). The neglected importance
547 of floral traits in trait-based plant community assembly. *Journal of Vegetation Science*,
548 *31*(4), 529–539. <https://doi.org/10.1111/jvs.12877>
- 549 E-Vojtkó, A., Junker, R. R., de Bello, F., & Götzenberger, L. (2022). Floral and reproductive traits
550 are an independent dimension within the plant economic spectrum of temperate central
551 Europe. *New Phytologist*, *236*(5), 1964–1975. <https://doi.org/10.1111/nph.18386>
- 552 Goldberg, E. E., Kohn, J. R., Lande, R., Robertson, K. A., Smith, S. A., & Igić, B. (2010). Species
553 Selection Maintains Self-Incompatibility. *Science*, *330*(6003), 493–495. [https://doi.org/10.](https://doi.org/10.1126/science.1194513)
554 [1126/science.1194513](https://doi.org/10.1126/science.1194513)
- 555 Goldberg, E. E., Otto, S. P., Vamosi, J. C., Mayrose, I., Sabath, N., et al. (2017). Macroevolutionary
556 synthesis of flowering plant sexual systems. *Evolution*, *71*(4), 898–912. [https://doi.org/10.](https://doi.org/10.1111/evo.13181)
557 [1111/evo.13181](https://doi.org/10.1111/evo.13181)
- 558 Govaerts, R., Nic Lughadha, E., Black, N., Turner, R., & Paton, A. (2021). The World Checklist
559 of Vascular Plants, a continuously updated resource for exploring global plant diversity.
560 *Scientific Data*, *8*(1), 215. <https://doi.org/10.1038/s41597-021-00997-6>
- 561 Gower, J. C. (1971). A General Coefficient of Similarity and Some of Its Properties. *Biometrics*,
562 *27*(4), 857–871. <https://doi.org/10.2307/2528823>
- 563 Grime, J. P. (1974). Vegetation classification by reference to strategies. *Nature*, *250*(5461), 26–31.
564 <https://doi.org/10.1038/250026a0>
- 565 Harris, D. M., & Harris, S. L. (2013). 1 - from zero to one. In D. M. Harris & S. L. Harris (Eds.),
566 *Digital design and computer architecture (second edition)* (pp. 2–52). Morgan Kaufmann.
567 <https://doi.org/10.1016/B978-0-12-394424-5.00001-X>
- 568 Helmstetter, A. J., Zenil-Ferguson, R., Sauquet, H., Otto, S. P., Méndez, M., et al. (2023). Trait-dependent
569 diversification in angiosperms: Patterns, models and data. *Ecology Letters*, *26*(4), 640–657.
570 <https://doi.org/10.1111/ele.14170>
- 571 Huang, Z. (1998). Extensions to the k-Means Algorithm for Clustering Large Data Sets with
572 Categorical Values. *Data Mining and Knowledge Discovery*, *2*(3), 283–304. [https://doi.](https://doi.org/10.1023/A:1009769707641)
573 [org/10.1023/A:1009769707641](https://doi.org/10.1023/A:1009769707641)
- 574 Igić, B., & Kohn, J. R. (2006). The Distribution of Plant Mating Systems: Study Bias Against
575 Obligately Outcrossing Species. *Evolution*, *60*(5), 1098–1103. [https://doi.org/10.1111/j.](https://doi.org/10.1111/j.0014-3820.2006.tb01186.x)
576 [0014-3820.2006.tb01186.x](https://doi.org/10.1111/j.0014-3820.2006.tb01186.x)

- 577 Jin, Y., & Qian, H. (2022). V.PhyloMaker2: An updated and enlarged R package that can generate
578 very large phylogenies for vascular plants. *Plant Diversity*, *44*(4), 335–339. [https://doi.org/](https://doi.org/10.1016/j.pld.2022.05.005)
579 [10.1016/j.pld.2022.05.005](https://doi.org/10.1016/j.pld.2022.05.005)
- 580 Johnson, C. A., Dutt, P., & Levine, J. M. (2022). Competition for pollinators destabilizes plant
581 coexistence. *Nature*, *607*(7920), 721–725. <https://doi.org/10.1038/s41586-022-04973-x>
- 582 Joly, S., & Schoen, D. J. (2021). Repeated evolution of a reproductive polyphenism in plants is
583 strongly associated with bilateral flower symmetry. *Current Biology*, *31*(7), 1515–1520.e3.
584 <https://doi.org/10.1016/j.cub.2021.01.009>
- 585 Käfer, J., Marais, G. A. B., & Pannell, J. R. (2017). On the rarity of dioecy in flowering plants.
586 *Molecular Ecology*, *26*(5), 1225–1241. <https://doi.org/10.1111/mec.14020>
- 587 Kattge, J., Bönisch, G., Díaz, S., Lavorel, S., Prentice, I. C., et al. (2020). TRY plant trait database
588 – enhanced coverage and open access. *Global Change Biology*, *26*(1), 119–188. <https://doi.org/10.1111/gcb.14904>
- 589
- 590 Kaufman, L., & Rousseeuw, P. (1990). Partitioning Around Medoids (Program PAM). In *Finding*
591 *Groups in Data* (pp. 68–125). John Wiley & Sons, Ltd. [https://doi.org/10.1002/](https://doi.org/10.1002/9780470316801.ch2)
592 [9780470316801.ch2](https://doi.org/10.1002/9780470316801.ch2)
- 593 Lamanna, C., Blonder, B., Violle, C., Kraft, N. J. B., Sandel, B., et al. (2014). Functional trait space
594 and the latitudinal diversity gradient. *Proceedings of the National Academy of Sciences*,
595 *111*(38), 13745–13750. <https://doi.org/10.1073/pnas.1317722111>
- 596 Lanuza, J. B., Rader, R., Stavert, J., Kendall, L. K., Saunders, M. E., & Bartomeus, I. (2023).
597 Covariation among reproductive traits in flowering plants shapes their interactions with
598 pollinators. *Functional Ecology*, *37*, 2072–2084. <https://doi.org/10.1111/1365-2435.14340>
- 599 Lesaffre, T., & Billiard, S. (2020). The joint evolution of lifespan and self-fertilization. *Journal of*
600 *Evolutionary Biology*, *33*(1), 41–56. <https://doi.org/10.1111/jeb.13543>
- 601 Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M., & Hornik, K. (2022). *Cluster: Cluster Analysis*
602 *Basics and Extensions*.
- 603 Magneville, C., Loiseau, N., Albouy, C., Casajus, N., Claverie, T., et al. (2022). mFD: An R package
604 to compute and illustrate the multiple facets of functional diversity. *Ecography*, *2022*(1).
605 <https://doi.org/10.1111/ecog.05904>
- 606 McInnes, L., Healy, J., & Melville, J. (2020). Umap: Uniform manifold approximation and projection
607 for dimension reduction. (arXiv:1802.03426). <https://doi.org/10.48550/arXiv.1802.03426>

- 608 Mouillot, D., Loiseau, N., Grenié, M., Algar, A. C., Allegra, M., et al. (2021). The dimensionality
609 and structure of species trait spaces. *Ecology Letters*, *24*(9), 1988–2009. [https://doi.org/](https://doi.org/10.1111/ele.13778)
610 [10.1111/ele.13778](https://doi.org/10.1111/ele.13778)
- 611 Munoz, F., Violle, C., & Cheptou, P.-O. (2016). Csr ecological strategies and plant mating systems:
612 Outcrossing increases with competitiveness but stress-tolerance is related to mixed mating.
613 *Oikos*, *125*(9), 1296–1303. <https://doi.org/10.1111/oik.02328>
- 614 Murtagh, F., & Legendre, P. (2014). Ward’s Hierarchical Agglomerative Clustering Method: Which
615 Algorithms Implement Ward’s Criterion? *Journal of Classification*, *31*(3), 274–295. [https://doi.org/](https://doi.org/10.1007/s00357-014-9161-z)
616 [10.1007/s00357-014-9161-z](https://doi.org/10.1007/s00357-014-9161-z)
- 617 Pagel, M. (1999). Inferring the historical patterns of biological evolution. *Nature*, *401*(6756), 877–884.
618 <https://doi.org/10.1038/44766>
- 619 Paradis, E., Claude, J., & Strimmer, K. (2004). APE: Analyses of Phylogenetics and Evolution in R
620 language. *Bioinformatics*, *20*(2), 289–290. <https://doi.org/10.1093/bioinformatics/btg412>
- 621 Paradis, E., & Schliep, K. (2019). Ape 5.0: An environment for modern phylogenetics and evolutionary
622 analyses in R. *Bioinformatics*, *35*, 526–528. <https://doi.org/10.1093/bioinformatics/bty633>
- 623 Pennell, M. W., Eastman, J. M., Slater, G. J., Brown, J. W., Uyeda, J. C., et al. (2014). Geiger
624 v2.0: An expanded suite of methods for fitting macroevolutionary models to phylogenetic
625 trees. *Bioinformatics*, *30*(15), 2216–2218. <https://doi.org/10.1093/bioinformatics/btu181>
- 626 Perez-Harguindeguy, N., Diaz, S., Garnier, E., Lavorel, S., Poorter, H., et al. (2013). New handbook
627 for standardised measurement of plant functional traits worldwide. *Australian Journal of*
628 *Botany*, *61*, 167–234. https://doi.org/10.1071/BT12225_CO
- 629 Petrone Mendoza, S., Lascoux, M., & Glémin, S. (2018). Competitive ability of *Capsella* species
630 with different mating systems and ploidy levels. *Annals of Botany*, *121*(6), 1257–1264.
631 <https://doi.org/10.1093/aob/mcy014>
- 632 Renner, S. S. (2014). The relative and absolute frequencies of angiosperm sexual systems: Dioecy,
633 monoecy, gynodioecy, and an updated online database. *American Journal of Botany*, *101*(10),
634 1588–1596. <https://doi.org/10.3732/ajb.1400196>
- 635 Revell, L. J. (2012). Phytools: An R package for phylogenetic comparative biology (and other
636 things). *Methods in Ecology and Evolution*, *3*(2), 217–223. [https://doi.org/10.1111/j.2041-](https://doi.org/10.1111/j.2041-210X.2011.00169.x)
637 [210X.2011.00169.x](https://doi.org/10.1111/j.2041-210X.2011.00169.x)

- 638 Roddy, A. B., Martínez-Perez, C., Teixido, A. L., Cornelissen, T. G., Olson, M. E., et al. (2021).
639 Towards the flower economics spectrum. *New Phytologist*, *229*(2), 665–672. <https://doi.org/10.1111/nph.16823>
640
- 641 Römer, G., Dahlgren, J. P., Salguero-Gómez, R., Stott, I. M., & Jones, O. R. (2023). Plant
642 demographic knowledge is biased towards short-term studies of temperate-region herbaceous
643 perennials. *Oikos*, *e10250*. <https://doi.org/10.1111/oik.10250>
- 644 Rose, J. P., & Sytsma, K. J. (2021). Complex interactions underlie the correlated evolution of floral
645 traits and their association with pollinators in a clade with diverse pollination systems.
646 *Evolution*, *75*(6), 1431–1449. <https://doi.org/10.1111/evo.14220>
- 647 Salguero-Gómez, R., Jones, O. R., Jongejans, E., Blomberg, S. P., Hodgson, D. J., et al. (2016).
648 Fast-slow continuum and reproductive strategies structure plant life-history variation worldwide.
649 *Proceedings of the National Academy of Sciences*, *113*(1), 230–235. [https://doi.org/10.1073/](https://doi.org/10.1073/pnas.1506215112)
650 [pnas.1506215112](https://doi.org/10.1073/pnas.1506215112)
- 651 Salguero-Gómez, R. (2017). Applications of the fast–slow continuum and reproductive strategy
652 framework of plant life histories. *New Phytologist*, *213*(4), 1618–1624. [https://doi.org/10.](https://doi.org/10.1111/nph.14289)
653 [1111/nph.14289](https://doi.org/10.1111/nph.14289)
- 654 Sato, H. (2002). Invasion of unisexuales in hermaphrodite populations of animal-pollinated plants:
655 Effects of pollination ecology and floral size-number trade-offs. *Evolution.*, *56*(12), 2374–2382.
656 <https://doi.org/10.1111/j.0014-3820.2002.tb00163.x>
- 657 Sauquet, H. (2019). PROTEUS: A database for recording morphological data and fossil calibrations.
658 Sauquet, H., von Balthazar, M., Magallón, S., Doyle, J. A., Endress, P. K., et al. (2017). The
659 ancestral flower of angiosperms and its early diversification. *Nature Communications*, *8*(1),
660 16047. <https://doi.org/10.1038/ncomms16047>
- 661 Schönenberger, J., von Balthazar, M., López Martínez, A., Albert, B., Prieu, C., et al. (2020).
662 Phylogenetic analysis of fossil flowers using an angiosperm-wide data set: Proof-of-concept
663 and challenges ahead. *American Journal of Botany*, *107*(10), 1433–1448. [https://doi.org/](https://doi.org/10.1002/ajb2.1538)
664 [10.1002/ajb2.1538](https://doi.org/10.1002/ajb2.1538)
- 665 Schuldt, A., Ebeling, A., Kunz, M., Staab, M., Guimarães-Steinicke, C., et al. (2019). Multiple plant
666 diversity components drive consumer communities across ecosystems. *Nature Communications*,
667 *10*(1), 1460. <https://doi.org/10.1038/s41467-019-09448-8>

- 668 Scofield, D. G., & Schultz, S. T. (2006). Mitosis, stature and evolution of plant mating systems:
669 Low-Phi and high-Phi plants. *Proceedings of the Royal Society B: Biological Sciences*,
670 *273*(1584), 275–282. <https://doi.org/10.1098/rspb.2005.3304>
- 671 Sicard, A., & Lenhard, M. (2011). The selfing syndrome: A model for studying the genetic and
672 evolutionary basis of morphological adaptation in plants. *Annals of Botany*, *107*(9), 1433–1443.
673 <https://doi.org/10.1093/aob/mcr023>
- 674 Sicard, A., Stacey, N., Hermann, K., Dessoly, J., Neuffer, B., et al. (2011). Genetics, evolution, and
675 adaptive significance of the selfing syndrome in the genus capsella. *The Plant Cell*, *23*(9),
676 3156–3171. <https://doi.org/10.1105/tpc.111.088237>
- 677 Smith, S. A., & Brown, J. W. (2018). Constructing a broadly inclusive seed plant phylogeny.
678 *American Journal of Botany*, *105*(3), 302–314. <https://doi.org/10.1002/ajb2.1019>
- 679 Society for Ecological Restoration, International Network for Seed Based Restoration, & Royal
680 Botanic Gardens Kew. (2023). *Seed Information Database (SID)*. <https://ser-sid.org/>
- 681 Stekhoven, D. J., & Bühlmann, P. (2012). MissForest—non-parametric missing value imputation for
682 mixed-type data. *Bioinformatics*, *28*(1), 112–118. [https://doi.org/10.1093/bioinformatics/](https://doi.org/10.1093/bioinformatics/btr597)
683 [btr597](https://doi.org/10.1093/bioinformatics/btr597)
- 684 Szepannek, G. (2018). clustMixType: User-Friendly Clustering of Mixed-Type Data in R. *The R*
685 *Journal*, *10*(2), 200–208.
- 686 Vamosi, J. C., & Vamosi, S. M. (2004). The role of diversification in causing the correlates of dioecy.
687 *Evolution*, *58*(4), 723–731.
- 688 Vamosi, J. C., Magallón, S., Mayrose, I., Otto, S. P., & Sauquet, H. (2018). Macroevolutionary
689 Patterns of Flowering Plant Speciation and Extinction. *Annual Review of Plant Biology*,
690 *69*(1), 685–706. <https://doi.org/10.1146/annurev-arplant-042817-040348>
- 691 Wei, N., Kaczorowski, R. L., Arceo-Gómez, G., O'Neill, E. M., Hayes, R. A., & Ashman, T. L.
692 (2021). Pollinators contribute to the maintenance of flowering plant diversity. *Nature*, *597*(7878),
693 688–692. <https://doi.org/10.1038/s41586-021-03890-9>
- 694 Wright, I. J., Reich, P. B., Westoby, M., Ackerly, D. D., Baruch, Z., et al. (2004). The worldwide leaf
695 economics spectrum. *Nature*, *428*(6985), 821–827. <https://doi.org/10.1038/nature02403>
- 696 Zanne, A. E., Tank, D. C., Cornwell, W. K., Eastman, J. M., Smith, S. A., et al. (2014). Three
697 keys to the radiation of angiosperms into freezing environments. *Nature*, *506*(7486), 89–92.
698 <https://doi.org/10.1038/nature12872>

699 Zhang, J., & Qian, H. (2023). U.Taxonstand: An R package for standardizing scientific names of
700 plants and animals. *Plant Diversity*, 45(1), 1–5. <https://doi.org/10.1016/j.pld.2022.09.001>