# 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 Cros[i], Pierre-Alexandre Delecroix[j], Mathilde Dufay[k], John R. Pannell[l], Dessislava Savova Bianchi[l], Daniel J. Schoen[m], Mario Vallejo-Marin[n], Rosana Zenil-Ferguson[o], Hervé Sauquet[pq], Sylvain Glémin[jn*], Jos Käfer[b*]

[a]FRB-CESAB, Montpellier, France
[b]ISEM, Univ Montpellier, CNRS, IRD, Montpellier, France
[c]Area of Biodiversity and Conservation, Universidad Rey Juan Carlos, 28933 Móstoles (Madrid), Spain
[d]Department of Botany and Biodiversity Research, University of Vienna, Rennweg 14, 1030 Vienna, Austria
[e]Department of Organismal Biology, University of Uppsala, Uppsala, 75236, Sweden
[f]Department of Botany and Zoology, University of Stellenbosch, Private Bag X1, Matieland 7602, South Africa
[g]University Lille, CNRS, UMR 8198 – Evo-Eco-Paleo, 59000 Lille, France
[h]Natural History Museum, University of Oslo, 0318 Oslo, Norway
[i]Université de Lyon, Université Lyon 1, CNRS, Laboratoire de Biométrie et Biologie Evolutive UMR 5558, Villeurbanne, France
[j]CNRS, Ecosystèmes Biodiversité Evolution (Université de Rennes), 35000 Rennes, France
[k]CEFE, University Montpellier, CNRS, University Paul Valéry Montpellier 3, Ephe, IRD, Montpellier, France
[l]Department of Ecology and Evolution, University of Lausanne, 1015 Lausanne, Switzerland
[m]Department of Biology, McGill University, Montreal, QC, Canada
[n]Department of Ecology and Genetics, Uppsala University, Uppsala, 752 36, Sweden
[o]Department of Biology, University of Kentucky, 101 T.H. Morgan Building, Lexington, KY 40506 USA
[p]National Herbarium of New South Wales, Royal Botanic Gardens and Domain Trust, Sydney, New South Wales, 2000, Australia
[q]Evolution 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.

# Abstract

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


# Significance statement

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

# Introduction

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

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

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

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

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

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

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

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

# Materials & Methods

## Data collection

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

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

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

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

## Trait encoding

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

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

## Filtering, transformations and missing data imputation

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

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

## Trait spaces

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

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

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

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

## Clustering

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

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

## Phylogenetic signal, transition rates and ancestral state estimation

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

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

## Simulated data sets

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

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

# Results

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

## Exploring reproductive trait space

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

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

## Comparison of reproductive and vegetative trait spaces

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

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

## Describing major reproductive strategies

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

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

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

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

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

## Trait space and phylogeny

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

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

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

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

# Discussion

## Pollination is a major source of plant trait variation

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

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

## The major reproductive strategies of flowering plants

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

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

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

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

## Sexual and mating systems contribute differently to reproductive strategies

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

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

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

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

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

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

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

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

# Acknowledgements

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


# Figures

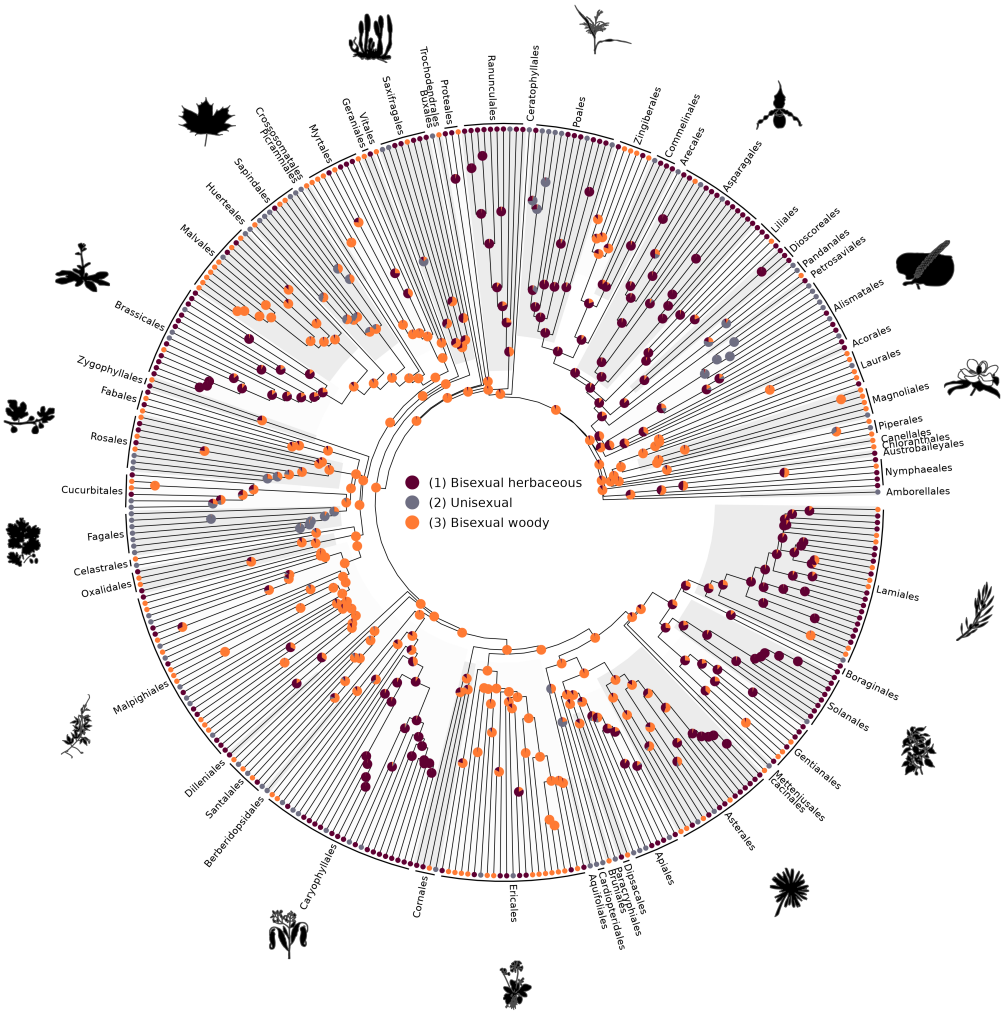


Figure 1: Phylogenetic tree representing the 360 species used in our analyses. Tips are labelled with circles indicating which group they were assigned to in our hierarchical clustering analysis (k = 3). Broadly, the three groups are summarised as: (1) Herbaceous species with bisexual flowers (2) Species 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.

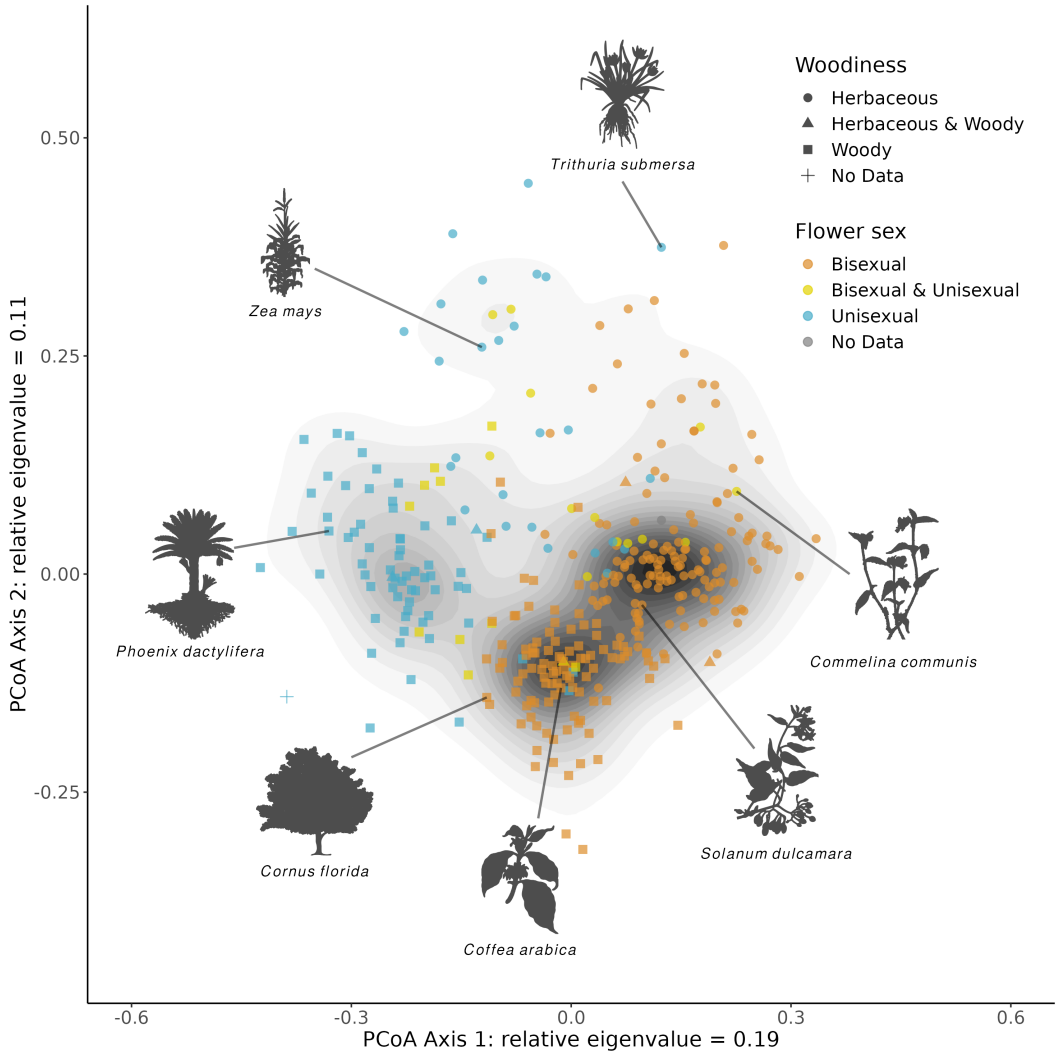


Figure 2: Location of 360 angiosperm species on the first two axes of our principal coordinate analysis (PCoA) on the original data set. These two axes represent approximately 30% of the variation in the data set. Points are coloured by the flower sex trait and shapes represent the woodiness trait. Polygons 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.

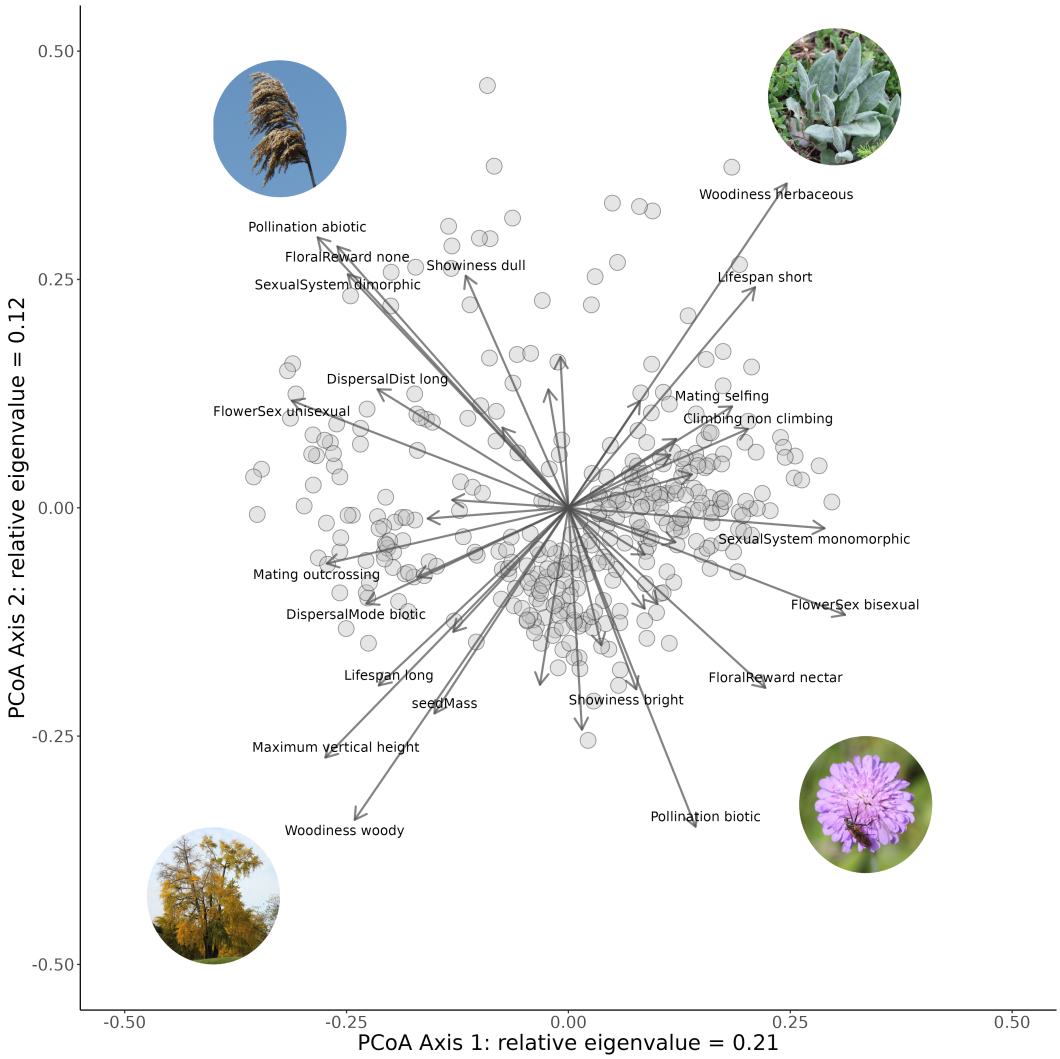


Figure 3: Principal coordinate analysis performed on the one-hot encoded data set. The first two axes represent approximately 33% of variation. Using one-hot encoding allowed us to plot arrows indicating how different states and traits affect the trait space, as well as which traits are acting in a similar 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.

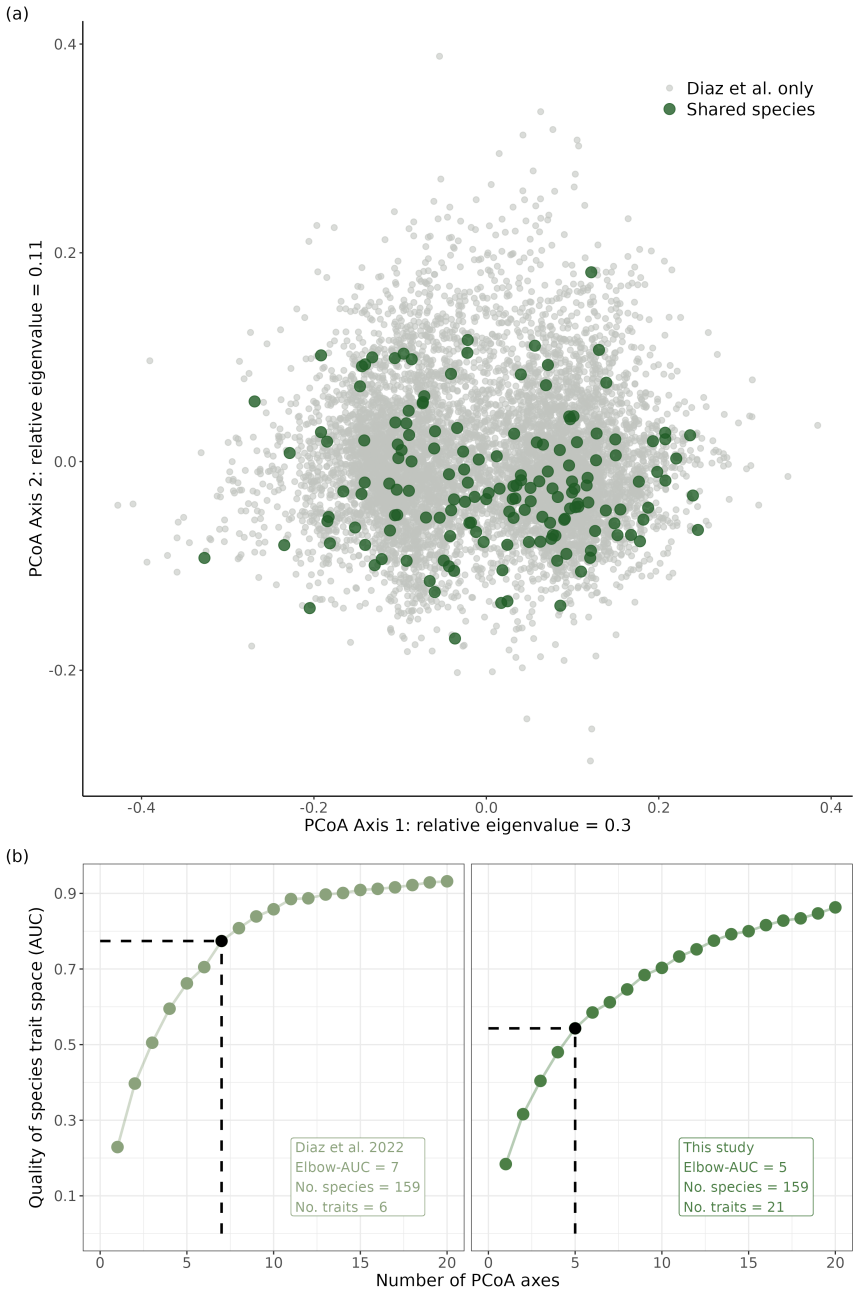


Figure 4: Comparison of the species sampled in the current study and the obtained trait space with species and trait spaces derived from Díaz et al. (2022). Panel (a) shows a scatterplot of a subset of approximately 8,000 species from Díaz et al. (2022) on the first two axes of a principal coordinate 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.

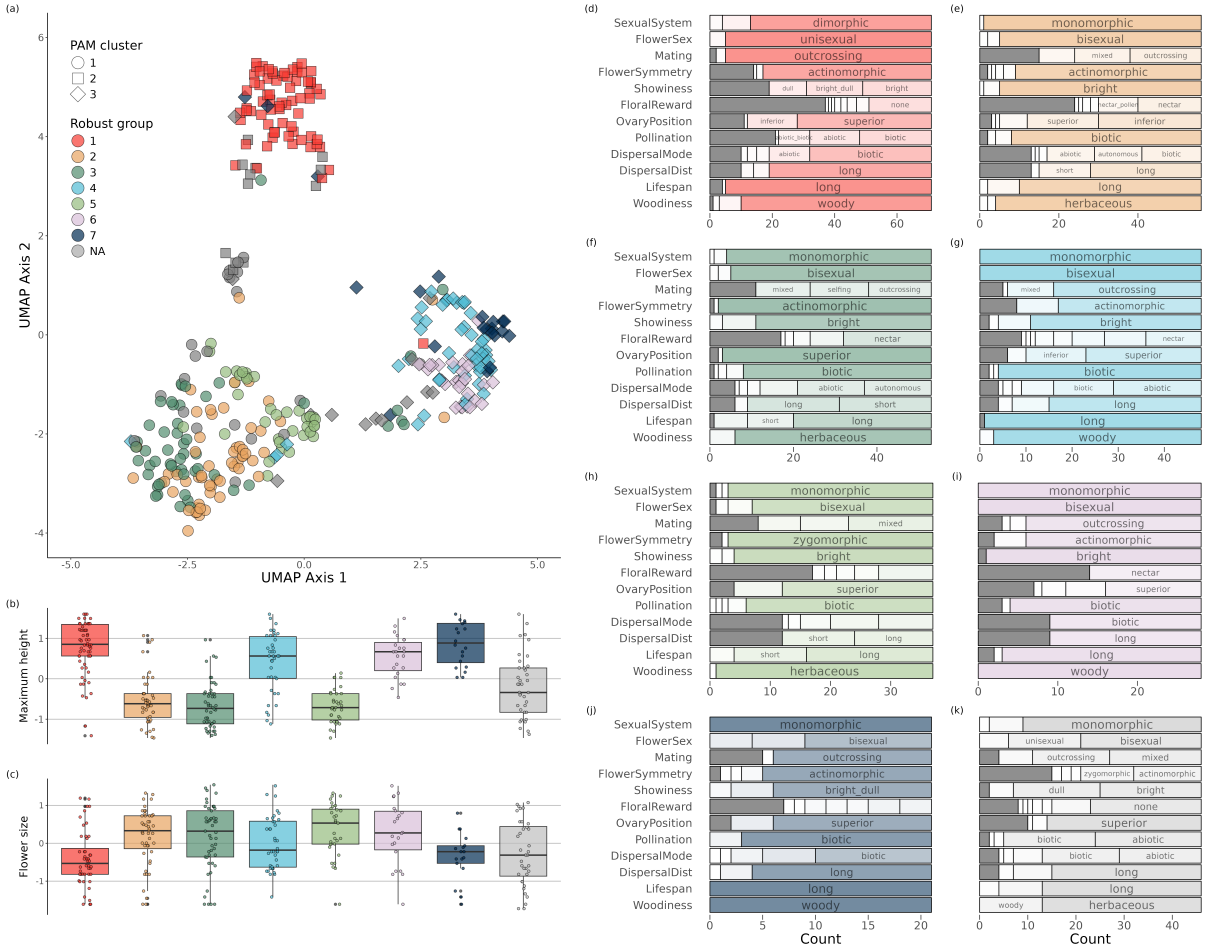


Figure 5: Characterizing robust groups identified using the Partitioning Around Mediods (PAM) clustering approach. Seven different robust groups were defined, and a final group was made with those species that did not fall into a robust group. Panel (a) shows the distribution of species along the first 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.

# References


Anderson, B., Pannell, J., Billiard, S., Burgarella, C., Boer, H. d., et al. (2023). Opposing effects of plant traits on diversification. *iScience*, *26*(4). https://doi.org/10.1016/j.isci.2023.106362

APGIV. (2016). An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. *Botanical Journal of the Linnean Society*, *181*(1), 1–20. https://doi.org/10.1111/boj.12385

Ashman, T.-L., Knight, T. M., Steets, J. A., Amarasekare, P., Burd, M., et al. (2004). Pollen limitation of plant reproduction: Ecological and evolutionary causes and consequences. *Ecology*, *85*(9), 2408–2421. https://doi.org/10.1890/03-8024

Barrett, S. C. H. (2002). The evolution of plant sexual diversity. *Nature Reviews Genetics*, *3*(4), 274–284. https://doi.org/10.1038/nrg776

Barrett, S. C. H. (2013). The evolution of plant reproductive systems: How often are transitions irreversible? *Proceedings of the Royal Society B: Biological Sciences*, *280*(1765), 20130913. https://doi.org/10.1098/rspb.2013.0913

Barrett, S. C. H., Harder, L. D., & Worley, A. C. (1997). The comparative biology of pollination and mating in flowering plants. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *351*(1345), 1271–1280. https://doi.org/10.1098/rstb.1996.0110

Bawa, K. S. (1980). Evolution of dioecy in flowering plants. *Ann. Rev. Ecol. Syst.*, *11*, 15–39. https://doi.org/10.1146/annurev.es.11.110180.000311

Blomberg, S. P., Garland JR., T., & Ives, A. R. (2003). Testing for Phylogenetic Signal in Comparative Data: Behavioral Traits Are More Labile. *Evolution*, *57*(4), 717–745. https://doi.org/10.1111/j.0014-3820.2003.tb00285.x

Borges, R., Machado, J. P., Gomes, C., Rocha, A. P., & Antunes, A. (2019). Measuring phylogenetic signal between categorical traits and phylogenies. *Bioinformatics*, *35*(11), 1862–1869. https://doi.org/10.1093/bioinformatics/bty800

Boyko, J. D., & Beaulieu, J. M. (2021). Generalized hidden Markov models for phylogenetic comparative datasets. *Methods in Ecology and Evolution*, *12*(3), 468–478. https://doi.org/10.1111/2041-210X.13534

Burgarella, C., Brémaud, M.-F., Hirschheydt, G. V., Viader, V., Ardisson, M., et al. (2023). Mating systems and recombination landscape strongly shape genetic diversity and selection in wheat relatives. *bioRxiv*, 2023.03.16.532584. https://doi.org/10.1101/2023.03.16.532584

Burgarella, C., & Glémin, S. (2017). Population genetics and genome evolution of selfing species. In J. Wiley & L. Sons (Eds.), *Els* (1st ed., pp. 1–8). Wiley. https://doi.org/10.1002/9780470015902.a0026804

Cardoso, J. C. F., Viana, M. L., Matias, R., Furtado, M. T., Caetano, A. P. D., et al. (2018). Towards a unified terminology for angiosperm reproductive systems. *Acta Bot. Bras.*, *32*(3), 329–348. https://doi.org/10.1590/0102-33062018abb0124

Charlesworth, D. (1993). Why are unisexual flowers associated with wind pollination and unspecialized pollinators? *The American Naturalist*, *141*(3), 481–490.

Charlesworth, D., & Charlesworth, B. (1978). Population-genetics of partial male-sterility and evolution of monoecy and dioecy. *Heredity*, *41*(OCT), 137–153.

Charnov, E. L., Maynard Smith, J., & Bull, J. J. (1976). Why be an hermaphrodite? *Nature*, *263*(5573), 125–126. https://doi.org/10.1038/263125a0

Cheptou, Imbert, Lepart, & Escarre. (2000). Effects of competition on lifetime estimates of inbreeding depression in the outcrossing plant Crepis sancta (Asteraceae). *Journal of Evolutionary Biology*, *13*(3), 522–531. https://doi.org/10.1046/j.1420-9101.2000.00175.x

Craig, D. M., & Mertz, D. B. (1994). Inbreeding effects on competition inTribolium. *Researches on Population Ecology*, *36*(2), 251–254. https://doi.org/10.1007/BF02514941

Crnokrak, P., & Roff, D. A. (1999). Inbreeding depression in the wild. *Heredity*, *83*(33), 260–270. https://doi.org/10.1038/sj.hdy.6885530

Debastiani, V. J., Bastazini, V. A. G., & Pillar, V. D. (2021). Using phylogenetic information to impute missing functional trait values in ecological databases. *Ecological Informatics*, *63*, 101315. https://doi.org/10.1016/j.ecoinf.2021.101315

Díaz, S., Kattge, J., Cornelissen, J. H. C., Wright, I. J., Lavorel, S., et al. (2016). The global spectrum of plant form and function. *Nature*, *529*(7585), 167–171. https://doi.org/10.1038/nature16489

Díaz, S., Kattge, J., Cornelissen, J. H. C., Wright, I. J., Lavorel, S., et al. (2022). The global spectrum of plant form and function: Enhanced species-level trait dataset. *Scientific Data*, *9*(1), 755. https://doi.org/10.1038/s41597-022-01774-9

Donoghue, M. J., & Sanderson, M. J. (2015). Confluence, synnovation, and depauperons in plant diversification. *New Phytologist*, *207*(2), 260–274. https://doi.org/10.1111/nph.13367

E-Vojtkó, A., de Bello, F., Durka, W., Kühn, I., & Götzenberger, L. (2020). The neglected importance of floral traits in trait-based plant community assembly. *Journal of Vegetation Science*, *31*(4), 529–539. https://doi.org/10.1111/jvs.12877

E-Vojtkó, A., Junker, R. R., de Bello, F., & Götzenberger, L. (2022). Floral and reproductive traits are an independent dimension within the plant economic spectrum of temperate central Europe. *New Phytologist*, *236*(5), 1964–1975. https://doi.org/10.1111/nph.18386

Goldberg, E. E., Kohn, J. R., Lande, R., Robertson, K. A., Smith, S. A., & Igić, B. (2010). Species Selection Maintains Self-Incompatibility. *Science*, *330*(6003), 493–495. https://doi.org/10.1126/science.1194513

Goldberg, E. E., Otto, S. P., Vamosi, J. C., Mayrose, I., Sabath, N., et al. (2017). Macroevolutionary synthesis of flowering plant sexual systems. *Evolution*, *71*(4), 898–912. https://doi.org/10.1111/evo.13181

Govaerts, R., Nic Lughadha, E., Black, N., Turner, R., & Paton, A. (2021). The World Checklist of Vascular Plants, a continuously updated resource for exploring global plant diversity. *Scientific Data*, *8*(1), 215. https://doi.org/10.1038/s41597-021-00997-6

Gower, J. C. (1971). A General Coefficient of Similarity and Some of Its Properties. *Biometrics*, *27*(4), 857–871. https://doi.org/10.2307/2528823

Grime, J. P. (1974). Vegetation classification by reference to strategies. *Nature*, *250*(5461), 26–31. https://doi.org/10.1038/250026a0

Harris, D. M., & Harris, S. L. (2013). 1 - from zero to one. In D. M. Harris & S. L. Harris (Eds.), *Digital design and computer architecture (second edition)* (pp. 2–52). Morgan Kaufmann. https://doi.org/10.1016/B978-0-12-394424-5.00001-X

Helmstetter, A. J., Zenil-Ferguson, R., Sauquet, H., Otto, S. P., Méndez, M., et al. (2023). Trait-dependent diversification in angiosperms: Patterns, models and data. *Ecology Letters*, *26*(4), 640–657. https://doi.org/10.1111/ele.14170

Huang, Z. (1998). Extensions to the k-Means Algorithm for Clustering Large Data Sets with Categorical Values. *Data Mining and Knowledge Discovery*, *2*(3), 283–304. https://doi.org/10.1023/A:1009769707641

Igic, B., & Kohn, J. R. (2006). The Distribution of Plant Mating Systems: Study Bias Against Obligately Outcrossing Species. *Evolution*, *60*(5), 1098–1103. https://doi.org/10.1111/j.0014-3820.2006.tb01186.x

Jin, Y., & Qian, H. (2022). V.PhyloMaker2: An updated and enlarged R package that can generate very large phylogenies for vascular plants. *Plant Diversity*, *44*(4), 335–339. https://doi.org/10.1016/j.pld.2022.05.005

Johnson, C. A., Dutt, P., & Levine, J. M. (2022). Competition for pollinators destabilizes plant coexistence. *Nature*, *607*(7920), 721–725. https://doi.org/10.1038/s41586-022-04973-x

Joly, S., & Schoen, D. J. (2021). Repeated evolution of a reproductive polyphenism in plants is strongly associated with bilateral flower symmetry. *Current Biology*, *31*(7), 1515–1520.e3. https://doi.org/10.1016/j.cub.2021.01.009

Käfer, J., Marais, G. A. B., & Pannell, J. R. (2017). On the rarity of dioecy in flowering plants. *Molecular Ecology*, *26*(5), 1225–1241. https://doi.org/10.1111/mec.14020

Kattge, J., Bönisch, G., Díaz, S., Lavorel, S., Prentice, I. C., et al. (2020). TRY plant trait database – enhanced coverage and open access. *Global Change Biology*, *26*(1), 119–188. https://doi.org/10.1111/gcb.14904

Kaufman, L., & Rousseeuw, P. (1990). Partitioning Around Medoids (Program PAM). In *Finding Groups in Data* (pp. 68–125). John Wiley & Sons, Ltd. https://doi.org/10.1002/9780470316801.ch2

Lamanna, C., Blonder, B., Violle, C., Kraft, N. J. B., Sandel, B., et al. (2014). Functional trait space and the latitudinal diversity gradient. *Proceedings of the National Academy of Sciences*, *111*(38), 13745–13750. https://doi.org/10.1073/pnas.1317722111

Lanuza, J. B., Rader, R., Stavert, J., Kendall, L. K., Saunders, M. E., & Bartomeus, I. (2023). Covariation among reproductive traits in flowering plants shapes their interactions with pollinators. *Functional Ecology*, *37*, 2072–2084. https://doi.org/10.1111/1365-2435.14340

Lesaffre, T., & Billiard, S. (2020). The joint evolution of lifespan and self-fertilization. *Journal of Evolutionary Biology*, *33*(1), 41–56. https://doi.org/10.1111/jeb.13543

Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M., & Hornik, K. (2022). *Cluster: Cluster Analysis Basics and Extensions*.

Magneville, C., Loiseau, N., Albouy, C., Casajus, N., Claverie, T., et al. (2022). mFD: An R package to compute and illustrate the multiple facets of functional diversity. *Ecography*, *2022*(1). https://doi.org/10.1111/ecog.05904

McInnes, L., Healy, J., & Melville, J. (2020). Umap: Uniform manifold approximation and projection for dimension reduction. (arXiv:1802.03426). https://doi.org/10.48550/arXiv.1802.03426

Mouillot, D., Loiseau, N., Grenié, M., Algar, A. C., Allegra, M., et al. (2021). The dimensionality and structure of species trait spaces. *Ecology Letters*, *24*(9), 1988–2009. https://doi.org/10.1111/ele.13778

Munoz, F., Violle, C., & Cheptou, P.-O. (2016). Csr ecological strategies and plant mating systems: Outcrossing increases with competitiveness but stress-tolerance is related to mixed mating. *Oikos*, *125*(9), 1296–1303. https://doi.org/10.1111/oik.02328

Murtagh, F., & Legendre, P. (2014). Ward's Hierarchical Agglomerative Clustering Method: Which Algorithms Implement Ward's Criterion? *Journal of Classification*, *31*(3), 274–295. https://doi.org/10.1007/s00357-014-9161-z

Pagel, M. (1999). Inferring the historical patterns of biological evolution. *Nature*, *401*(6756), 877–884. https://doi.org/10.1038/44766

Paradis, E., Claude, J., & Strimmer, K. (2004). APE: Analyses of Phylogenetics and Evolution in R language. *Bioinformatics*, *20*(2), 289–290. https://doi.org/10.1093/bioinformatics/btg412

Paradis, E., & Schliep, K. (2019). Ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*, *35*, 526–528. https://doi.org/10.1093/bioinformatics/bty633

Pennell, M. W., Eastman, J. M., Slater, G. J., Brown, J. W., Uyeda, J. C., et al. (2014). Geiger v2.0: An expanded suite of methods for fitting macroevolutionary models to phylogenetic trees. *Bioinformatics*, *30*(15), 2216–2218. https://doi.org/10.1093/bioinformatics/btu181

Perez-Harguindeguy, N., Diaz, S., Garnier, E., Lavorel, S., Poorter, H., et al. (2013). New handbook for standardised measurement of plant functional traits worldwide. *Australian Journal of Botany*, *61*, 167–234. https://doi.org/10.1071/BT12225_CO

Petrone Mendoza, S., Lascoux, M., & Glémin, S. (2018). Competitive ability of Capsella species with different mating systems and ploidy levels. *Annals of Botany*, *121*(6), 1257–1264. https://doi.org/10.1093/aob/mcy014

Renner, S. S. (2014). The relative and absolute frequencies of angiosperm sexual systems: Dioecy, monoecy, gynodioecy, and an updated online database. *American Journal of Botany*, *101*(10), 1588–1596. https://doi.org/10.3732/ajb.1400196

Revell, L. J. (2012). Phytools: An R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution*, *3*(2), 217–223. https://doi.org/10.1111/j.2041-210X.2011.00169.x

Roddy, A. B., Martínez-Perez, C., Teixido, A. L., Cornelissen, T. G., Olson, M. E., et al. (2021). Towards the flower economics spectrum. *New Phytologist*, *229*(2), 665–672. https://doi.org/10.1111/nph.16823

Römer, G., Dahlgren, J. P., Salguero-Gómez, R., Stott, I. M., & Jones, O. R. (2023). Plant demographic knowledge is biased towards short-term studies of temperate-region herbaceous perennials. *Oikos*, *e10250*. https://doi.org/10.1111/oik.10250

Rose, J. P., & Sytsma, K. J. (2021). Complex interactions underlie the correlated evolution of floral traits and their association with pollinators in a clade with diverse pollination systems. *Evolution*, *75*(6), 1431–1449. https://doi.org/10.1111/evo.14220

Salguero-Gómez, R., Jones, O. R., Jongejans, E., Blomberg, S. P., Hodgson, D. J., et al. (2016). Fast-slow continuum and reproductive strategies structure plant life-history variation worldwide. *Proceedings of the National Academy of Sciences*, *113*(1), 230–235. https://doi.org/10.1073/pnas.1506215112

Salguero-Gómez, R. (2017). Applications of the fast–slow continuum and reproductive strategy framework of plant life histories. *New Phytologist*, *213*(4), 1618–1624. https://doi.org/10.1111/nph.14289

Sato, H. (2002). Invasion of unisexuals in hermaphrodite populations of animal-pollinated plants: Effects of pollination ecology and floral size-number trade-offs. *Evolution.*, *56*(12), 2374–2382. https://doi.org/10.1111/j.0014-3820.2002.tb00163.x

Sauquet, H. (2019). PROTEUS: A database for recording morphological data and fossil calibrations.

Sauquet, H., von Balthazar, M., Magallón, S., Doyle, J. A., Endress, P. K., et al. (2017). The ancestral flower of angiosperms and its early diversification. *Nature Communications*, *8*(1), 16047. https://doi.org/10.1038/ncomms16047

Schönenberger, J., von Balthazar, M., López Martínez, A., Albert, B., Prieu, C., et al. (2020). Phylogenetic analysis of fossil flowers using an angiosperm-wide data set: Proof-of-concept and challenges ahead. *American Journal of Botany*, *107*(10), 1433–1448. https://doi.org/10.1002/ajb2.1538

Schuldt, A., Ebeling, A., Kunz, M., Staab, M., Guimarães-Steinicke, C., et al. (2019). Multiple plant diversity components drive consumer communities across ecosystems. *Nature Communications*, *10*(1), 1460. https://doi.org/10.1038/s41467-019-09448-8

Scofield, D. G., & Schultz, S. T. (2006). Mitosis, stature and evolution of plant mating systems: Low-Phi and high-Phi plants. *Proceedings of the Royal Society B: Biological Sciences*, *273*(1584), 275–282. https://doi.org/10.1098/rspb.2005.3304

Sicard, A., & Lenhard, M. (2011). The selfing syndrome: A model for studying the genetic and evolutionary basis of morphological adaptation in plants. *Annals of Botany*, *107*(9), 1433–1443. https://doi.org/10.1093/aob/mcr023

Sicard, A., Stacey, N., Hermann, K., Dessoly, J., Neuffer, B., et al. (2011). Genetics, evolution, and adaptive significance of the selfing syndrome in the genus capsella. *The Plant Cell*, *23*(9), 3156–3171. https://doi.org/10.1105/tpc.111.088237

Smith, S. A., & Brown, J. W. (2018). Constructing a broadly inclusive seed plant phylogeny. *American Journal of Botany*, *105*(3), 302–314. https://doi.org/10.1002/ajb2.1019

Society for Ecological Restoration, International Network for Seed Based Restoration, & Royal Botanic Gardens Kew. (2023). *Seed Information Database (SID)*. https://ser-sid.org/

Stekhoven, D. J., & Bühlmann, P. (2012). MissForest—non-parametric missing value imputation for mixed-type data. *Bioinformatics*, *28*(1), 112–118. https://doi.org/10.1093/bioinformatics/btr597

Szepannek, G. (2018). clustMixType: User-Friendly Clustering of Mixed-Type Data in R. *The R Journal*, *10*(2), 200–208.

Vamosi, J. C., & Vamosi, S. M. (2004). The role of diversification in causing the correlates of dioecy. *Evolution*, *58*(4), 723–731.

Vamosi, J. C., Magallón, S., Mayrose, I., Otto, S. P., & Sauquet, H. (2018). Macroevolutionary Patterns of Flowering Plant Speciation and Extinction. *Annual Review of Plant Biology*, *69*(1), 685–706. https://doi.org/10.1146/annurev-arplant-042817-040348

Wei, N., Kaczorowski, R. L., Arceo-Gómez, G., O'Neill, E. M., Hayes, R. A., & Ashman, T. L. (2021). Pollinators contribute to the maintenance of flowering plant diversity. *Nature*, *597*(7878), 688–692. https://doi.org/10.1038/s41586-021-03890-9

Wright, I. J., Reich, P. B., Westoby, M., Ackerly, D. D., Baruch, Z., et al. (2004). The worldwide leaf economics spectrum. *Nature*, *428*(6985), 821–827. https://doi.org/10.1038/nature02403

Zanne, A. E., Tank, D. C., Cornwell, W. K., Eastman, J. M., Smith, S. A., et al. (2014). Three keys to the radiation of angiosperms into freezing environments. *Nature*, *506*(7486), 89–92. https://doi.org/10.1038/nature12872

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