Geographical and life-history traits associated with low and high species richness across angiosperm families

Miriam Monserrat Ferrer[1], Marilyn Vásquez-Cruz[2†], Tania Hernández-Hernández[3] and Sara V. Good[4*]

[1]Departamento de Manejo y Conservación de Recursos Naturales Tropicales, Universidad Autónoma de Yucatán, Mérida Yucatán, Mexico, [2]Tecnológico Nacional de México/ITS Irapuato, Guanajuato, Mexico, [3]Research and Collections, Desert Botanical Garden, Phoenix, AZ, United States, [4]Department of Biology, The University of Winnipeg, Winnipeg, MB, Canada

Introduction: The phenomenal expansion of angiosperms has prompted many investigations into the factors driving their diversification, but there remain significant gaps in our understanding of flowering plant species diversity.

Methods: Using the crown age of families from five studies, we used a maximum likelihood approach to classify families as having poor, predicted or high species richness (SR) using strict consensus criteria. Using these categories, we looked for associations between family SR and i) the presence of an inferred familial ancestral polyploidization event, ii) 23 life history and floral traits compiled from previously published datasets and papers, and iii) sexual system (dioecy) or genetically determined self-incompatibility (SI) mating system using an updated version of our own database and iv) geographic distribution using a new database describing the global distribution of plant species/families across realms and biomes and inferred range.

Results: We find that more than a third of angiosperm families (65%) had predicted SR, a large proportion (30.2%) were species poor, while few (4.8%) had high SR. Families with poor SR were less likely to have undergone an ancestral polyploidization event, exhibited deficits in diverse traits, and were more likely to have unknown breeding systems and to be found in only one or few biomes and realms, especially the Afrotropics or Australasia. On the other hand, families with high SR were more likely to have animal mediated pollination or dispersal, are enriched for epiphytes and taxa with an annual life history, and were more likely to harbour sporophytic SI systems. Mapping the global distribution of georeferenced taxa by their family DR, we find evidence of regions dominated by taxa from lineages with high vs low SR.

Discussion: These results are discussed within the context of the literature describing "depauperons" and the factors contributing to low and high biodiversity in angiosperm clades.

1 Introduction

Plants ventured into terrestrial environments for the first time ca. 420 million years before the present (mybp) (Kenrick and Crane, 1997; Magallon et al., 2013; Nie et al., 2020; Su et al., 2021), and proliferated to establish modern land floras. The majority of extant land plant species are angiosperms (90% of species), followed by mosses (ca. 4%), ferns (ca. 3%), and liverworts (ca. 2%), with the remaining species being hornworts, lycophytes, and gymnosperms (Fiz-Palacios et al., 2011; Magallon et al., 2013). Recent studies indicate that the origin of flowering plant orders occurred predominantly in the Early Cretaceous (145–100 mybp) while most extant families radiated during the Paleocene (66–56 mybp) (Magallón et al., 2015; Ramírez-Barahona et al., 2020), although there remains uncertainty in the divergence time of many lineages (Li et al., 2019). Nevertheless, the phenomenal expansion of angiosperm species richness (SR) in that relatively short time span has prompted many investigations into the factors influencing their diversification over geological time (Stanley, 1979; Raup and Sepkoski, 1984; Heard and Hauser, 1995; Nee, 2006; Jablonski, 2007).

Macroevolutionary studies of angiosperm diversification have been dominated by a focus on radiations (Hernández-Hernández, 2019), specifically by studies exploring the possible biotic and abiotic traits that have promoted diversification in lineages that significantly exceed some background level of SR (Magallón and Sanderson, 2001; Magallon et al., 2013; Ferrer et al., 2014; Magallón et al., 2015; Vamosi et al., 2018; Magallón et al., 2019; Soltis et al., 2019; Hernández-Hernández and Wiens, 2020; Ramírez-Barahona et al., 2020). These studies suggest a contextual importance of, for example, innovations in floral morphology and reproduction efficiency that facilitated interactions with pollinators and seed dispersers (Kay et al., 2006a; Sauquet et al., 2017); coevolution with animals (especially pollinators and herbivores; (Xiao et al., 2021); new photosynthetic capabilities (Benton et al., 2022); whole genome duplication (WGD) events (Tank et al., 2015; Clark and Donoghue, 2018); or geographic distribution patterns (Vamosi and Vamosi, 2011). However, significant gaps in our understanding of angiosperm evolution remain; it has been particularly difficult to find universal drivers of increased DR that can explain SR patterns in angiosperms, because traits do not seem to have a consistent effect across clades (Doyle and Donoghue, 1986; Friis et al., 2006; Crepet and Niklas, 2009; Vamosi and Vamosi, 2010; Sauquet and Magallón, 2018; Soltis et al., 2019; Hernández-Hernández and Wiens, 2020; Helmstetter et al., 2023).

A relatively unexplored macroevolutionary pattern ubiquitous along the Tree of Life is the opposite of radiations, the so-called "depauperons"; or lineages showing lower SR than expected (Donoghue and Sanderson, 2015; Caron and Pie, 2022). Depauperons have been recognized since the pre-molecular era, and "living fossils", or the phenomenon known as "arrested evolution", are a special case. Living fossils are extant species anatomically similar to a fossil one that occurred early in the history of a lineage (Eldredge and Stanley, 1984). In plants, the genus *Ginkgo* is a textbook example of a living fossil, because it is thought to have appeared at least from the Jurassic period 170

million years (Myr) ago with very little changes in its morphology since (Zhou and Zheng, 2003). Other types of depauperons are the so-known "dead clades walking", or fossil groups that suffer major drops in their biodiversity at a mass extinction but do not completely disappear (Jablonski, 2007; Barnes et al., 2021). The existence of depauperons represents one of the most interesting puzzles in evolutionary biology, because under normal circumstances, a lineage is expected to either diversify or go extinct, and although theoretical explanations for their existence have been postulated, their drivers remain uncertain (Strathmann and Slatkin, 1983; Barnes et al., 2021). Although overlooked, the study of the evolutionary mechanisms sustaining depauperons in contrast to those driving radiations or angiosperm "success stories" will likely illuminate explanations of SR patterns across flowering plants (Caron and Pie, 2022).

To explore SR patterns in angiosperms from a depauperons perspective, our study aims at identifying plant traits and geographic distribution patterns associated with angiosperm families that have unusually low SR. For this purpose, we followed a consensus maximum-likelihood (ML)-based approach to model the diversification rate process in angiosperm families, and identified families with significantly higher or lower DR/SR than expected by using five relevant calibrated angiosperm reconstruction datasets that vary in the sampling, calibration, and statistical approach. Then, using only those families categorized as having low, expected, or high SR across all five studies (strict consensus), we assessed if family SR was associated with traits that have been recurrently reported as having an impact on angiosperm DR: (i) the presence of an ancestral WGD event; (ii) 23 intrinsic traits from published studies including traits related to flowering morphology, mode of pollination or fruit dispersal, and growth habit; (iii) variation in sexual system, specifically the presence of dioecy, or self-incompatibility (SI) mating system; and/or (iv) the geographical distribution of families. Our traits dataset compiled an updated version of a previous one (Ferrer and Good, 2012) with previous datasets (Hernández-Hernández and Wiens, 2020), and we included a curated version of georeference dataset for all angiosperm families studied. Our results reveal important patterns in the distribution of depauperons, which, although are not a constitutive component of floras around the world, compose ~30% of angiosperm diversity at a family level, and include several vulnerable taxa that are highly endemic. Our results are discussed within the context of the literature describing "depauperons" (sensu Donoghue and Sanderson, 2015), and the wealth of literature underscoring the importance of breeding systems and animal–plant interactions in biodiversity conservation.

2 Materials and methods

2.1 Crown age and species richness of angiosperm families

Since the taxonomic delineation of families has changed since the oldest study we included (Wikström et al., 2001), and families have been split, merged, etc. over time, the family names and

relationship included in this study reflect the most recent classification recognized by the APGIV (The Angiosperm Phylogeny Group et al., 2016) (Table 1). The SR for 432 of the 434 families recognized by the APG IV was obtained from the angiosperm phylogeny site (Stevens, 2021); when the number of species in a family was presented as a range, the mean was used as the SR. The crown age of families was obtained from five calibrated angiosperm phylogenies: Wikström et al. (2001); Bell et al. (2010) (exponential), Hernández-Hernández and Wiens (2020); Li et al. (2019), and Ramírez-Barahona et al. (2020) (RC_complete_MCCv_2.tre as recommended by the authors). The crown age of families not included in a study was set to that of its' sister clade according to the APGIV, or to the stem age in the case of monotypic families or those in which only one species has been genotyped to represent the family (Stevens, 2021). The dataset names were abbreviated throughout the paper using the last name of the first author as follows: (1) Wikström (Wikström et al., 2001), (2) Bell (Bell et al., 2010), (3) Hernández-Hernández (Hernández-Hernández and Wiens, 2020), (4) Li (Li et al., 2019), and (5) Ramírez-Barahona (Ramírez-Barahona et al., 2020).

2.2 Shifts on the diversification rate of angiosperms

Since diversification patterns can change through global events in different geological epochs, we divided the data into time intervals during which the diversification rate followed a uniform distribution (following Ferrer et al., 2014). To identify significant shifts in DR, the original trees from each dataset were trimmed to exclude outgroups and linearized to be ultrametric with respect to time (if not already ultrametric) using an R script (available from https://hcliedtke.github.io/R-scrapheap/be_ultrametric.html). Subsequently, potential shifts in the rate of speciation or extinction were modeled using reversible jump MCMC as implemented in the program TESS (Höhna et al., 2016) (Supplementary Data File 1); shift points were inferred when 2lnBayesFactor ≥ 2 and the median age of the peak was used to divide the data into periods with approximately uniform DR. Once divided, we employed an MLE approach (below) to estimate the DR parameters and their 95% CI within each period.

2.3 Estimates of the speciation, extinction, and diversification rate of angiosperm families

We used the SR/clade age data from each of the five datasets to estimate the average absolute diversification rate of angiosperm families using a birth–death model. For all families originating within each geological interval, MLEs of $\hat{r}_{\lambda-\mu}$, λ, μ, ε and their 95% CIs were estimated from the function β (see equation 2b, from Magallón and Sanderson, 2001) as:

$$\beta = \frac{\lambda \left[e^{[\lambda-\mu]^t} - 1 \right]}{\lambda \left[\lambda e^{[\lambda-\mu]^t} - \varepsilon \right]}$$

and then the MLE of $\hat{r}_{\lambda-\mu}$, λ, μ, ε obtained using the following likelihood equation (Bokma, 2003).

$$L(\lambda - \mu) = lnL \sum_i^k \ln(1 - \beta_i) + \sum_i^k [(n-1)ln\beta_i]$$

For each time interval per study, the likelihood of the model (pSp) was estimated using diversification (r) rate values from 0.0001 to 0.2, and speciation rates (λ) from 0.1 to 100 species (per million years), in increments of 0.0002 for r and of 0.01 for λ using the divergence age (t_i) and SR of each family (n_i) following Bokma (2003) with the aid of Maple (see Supplementary Data File 1).

Next, two sets of CIs were estimated: CIs for the MLEs of $\hat{r}_{\lambda-\mu}$, λ, and μ, were estimated from the likelihood surface and a second set of CIs were obtained to account for error in the expected SR of families under a stochastic diversification process. In the first case, we used the fact that when estimating related parameters via ML (i.e., $\hat{r} = \lambda - \mu$), approximate CIs are given by ±1 unit of the likelihood surface (Meeker and Escobar, 1995). The ML surface generated following Eq. (2) was explored to consider all models within one likelihood unit of the best model, and then the 95% CIs were taken to be the lowest and highest values of λ and r, and by deduction, μ. Secondly, CIs associated with the number of species per clade under the birth–death model were obtained by solving equations 10a and 10b from Magallón and Sanderson (2001) using the following equations:

$$k_{(upper)t} = \frac{\ln(0.025)}{\ln(\beta)+1}$$

TABLE 1　Genes, number of families, outgroup taxa, and fossils included in the five calibrated phylogenies employed in this study: **Wikström et al. (2001); Bell et al. (2010); Hernández-Hernández and Wiens (2020); Li et al. (2019),** and **Ramírez-Barahona et al. (2020).**

Phylogeny	Genes	No. of families			Outgroup taxa	Fossils
		Original work	APGIV	Inferred ages		
Wikström	*rbcl, atpB, 18S rDNA*	340	369	65	7	1
Bell	*rbcl, atpB, 18S rDNA*	335	295	137	7	36
Hernández-Hernández	*rbcl, atpB, matK, 18S, 26S rDNA*	363	402	30	7	137
Li	*80 (plastid)*	353	353	79	163	62
Ramírez-Barahona	*rbcl, atpB, matK, ndhF, 18S, 26S, 5.8S, rDNA*	435	402	30	7	238

Given the changes in the identity of angiosperm phylogeny group (APG) recognized families since 2001, the families considered in the original work were lifted over to APGIV, and the age of missing sister clades inferred following APGIV.

$$k_{(lower)t} = \frac{\ln(0.975)}{\ln(\beta)+1}$$

where α and β were obtained by solving α_t and β_t as given in Magallón and Sanderson (2001) from equation 8.47 of Bailey (1964), in which $\hat{r}_{\lambda-\mu}$ and ϵ were taken as the value of the net diversification rate and extinction fraction that maximized the likelihood given in equation 2, where t was the age of divergence of the family.

2.4 Assigning families as having rich, predicted, or poor species richness

Using the maximum width of the CI, families were classified as having low, predicted, or high SR if the number of species in the family fell below, within, or above the maximum or minimum CI, respectively. To assess whether family classification was independent of the time period in which the family originated, a χ^2 with Haberman residual analysis was performed (Haberman, 1973). Finally, if families were identified as having low, predicted, or high SR based on strict (5/5) consensus across datasets, they were classified as having "poor", "expected", or "high" SR overall, while families lacking consistency across datasets were labeled as "undefined".

2.5 Traits associated with families differing in species richness

We tested whether the presence of a WGD event within the evolutionary history (crown or stem age) of a family based on the data from Landis et al. (2018), and/or the presence of fleshy fruits or floral symmetry classification, following Vamosi and Vamosi (2011) and Reyes et al. (2016), respectively, were associated with SR category. A χ^2 contingency analysis followed by a Haberman residual analysis (Haberman, 1973) was used to test for significance. Next, we looked for associations between family SR category and data on 21 traits compiled by Hernández-Hernández and Wiens (2020). The traits were divided into those related to the proportion of species in a family with a given: mode of dispersal (biotic, insect, vertebrate, water, or wind), fertilization (via biotic pollinator, insects, vertebrate pollinator, water, or wind), growth form (climbing, herbaceous, shrub, or tree), habitat (aquatic, epiphytic, or terrestrial), mean seed weight, or the proportion presenting with an annual life-history form, nectar spurs, polyploidy, or dioecy. Kruskal–Wallis tests followed by Dunn's post-hoc comparisons were used to test if the distribution of trait values between SR categories differed using a global FDR of 0.1.

2.6 Association of mating and sexual systems among families differing in species richness

2.6.1 Presence of SI

We used an updated version of the database from Ferrer and Good (2012) describing evidence for the presence of a self-sterility (SS), self-compatibility (SC), or one or more types of genetically controlled self-incompatibility (SI) systems in angiosperm families to test if there was an association between mating systems and family SR category. Families were classified as self-compatible when >80% of the species examined in a family exhibited SC, as SC-SI when >20% but<80% of species were classified as SI or SS, or as self-incompatible when >80% of the species in a family were classified as having SI or SS.

2.6.2 Presence of dioecy

Information about sexual system (dioecy) was included using the proportion of dioecious taxa obtained from Renner (2014) and Hernández-Hernández and Wiens (2020): only families in which all of the species exhibited dioecy were considered to be dioecious clades.

2.6.3 Type of SI

To assess if there are differences in the type of SI across families, we updated the database of Ferrer and Good (2012), to assign families with SI as having GSI (gametophytic self-incompatibility), SSI (sporophytic self-incompatibility), LAS (late-acting self-incompatibility), HET (heterostyly), HET + SI, or poly (multiple forms of SI in the same family), or as being unclassified (when there was evidence of self-sterility but unknown SI system). Inclusion criteria for designating a species as having genetically controlled SI required evidence from (unsuccessful) cross-pollinations between individuals putatively harboring the same S-genotype, microscopic analyses of pollen tube arrest in the stigma (SSI), pistil (GSI), or ovary (LAS), and/or expression of known S-alleles. In total, data on 6,441 species were extracted from manuscripts published between 1940 and 2022, and data on the type of SI in the 248 angiosperm families with an SR category based on strict consensus across studies included in the analyses. After removing 88 families with unknown mating systems, 111 families with some form of SI or SS, 38 families with SC, and 11 families with dioecious taxa were identified.

2.7 Geographic distribution of families differing in species richness

To evaluate the geographic distribution and range of families, we used the Global Biodiversity Information Facility database (GBIF, gbif.org). Only records from preserved specimens and literature with georeferenced coordinates were downloaded. All family names were corroborated and standardized according to The Plant List (The Plant List, available now http://www.worldfloraonline.org/) using the taxonstand package in R software (Cayuela et al., 2012). We removed all records lacking full genus and species names, and eliminated species records with duplicated georeferenced data, and those from common crops, cultivated specimens, botanical gardens, or from greenhouses or markets. The geographical distribution of taxa was finally corroborated manually at the family level, based on information for each family in the Kubitzki system (Kubitzki, 1990 onward), and online platforms (https://www.tropicos.org/home, http://www.mobot.org/MOBOT/research/APweb/; http://

www.worldfloraonline.org/). This curated georeferenced database was used to map the mean DR of taxa belonging to families with high vs. poor SR. To do this, we first calculated the mean DR of the 75 families with poor SR and the 12 families with high SR (Supplementary Data File 4). Next, we divided the globe into grids consisting of 1 degree squared. Then, within each grid, the average DR of all the families to which each georeferenced data point (taxa) belonged was calculated and plotted. Thus, the color of each pixel on the map represents the global mean DR of the families represented by the taxa georeferenced to that location.

2.8 Distribution of families with different SR categories in global biomes and realms

The same curated databases were used to score the presence of families in global biomes and realms, as defined by Olson and Dinerstein (1998). We used the bioregionalization categories from the Terrestrial Ecoregions of the World (TEOW), with a subdivision of the terrestrial world into 14 biomes and eight biogeographic realms. For each family, we counted the percentage of records falling within each realm and biome; thus, the maps generated from this analysis depict the density of taxa from families with low, expected, or high SSR across the globe. Next, to analyze the distributional data using contingency analysis, the distributional range of families with poor, predicted, or high SR were categorized as highly localized, localized, widespread, and cosmopolitan by visual arbitrary assessments with the following criteria: highly localized (when all records were concentrated in a single country or small region), localized (when all records were concentrated in a larger area within a continent), widespread (when the distribution of records span more than one continent), and cosmopolitan (when the records were distributed in several continents). In addition, and by visual assessment, the distributional pattern for families was categorized as disjunct (distributed in clusters but in two separate continents, or clusters separated by a considerable large area without continental contiguity) or continuous (no disjunctions).

3 Results

3.1 Crown age and species richness of angiosperm families

The five studies employed overlapping, but not identical, sets of molecular markers, primarily derived from the chloroplast genome, but differed markedly in the number of fossils used to calibrate the phylogeny and the methods used to calibrate the phylogeny (Table 1). Of the 432 angiosperm families, 235 were included in all five studies, and crown ages were available for 295 (Bell) to 402 (Hernández-Hernández and Wiens, 2020, and (Ramírez-Barahona et al., 2020) families recognized by APGIV, such that ages were inferred using the (missing) family's age for 137 to 30 families (Table 1).

3.2 Shifts on the diversification rate of angiosperms

Bayesian analysis identified shifts in the DR of angiosperm families in all of the ultrametric calibrated phylogenies, and inferred either four (Li), five (Wikström), six (Hernández-Hernández and Ramírez-Barahona), or seven (Bell) shifts in DR (Supplementary Data File 2, Table S1; Figure 1).

3.3 Estimates of the speciation, extinction, and diversification rate of angiosperm families

The MLE of the DR in most time intervals and for most phylogenies was low ($r \sim 0.0001$, Supplementary Data File 2, Table S2), although, in some cases, the CIs were high (e.g., Bell). In four of the five studies, r was highest for families that originated in the Albian-Turonian (113–100.5 mybp), while in the Li et al. study, r was the highest for families originating during the Neogene (Figure 1). The speciation rate (λ) estimated from the Wikström, Bell, and Hernández-Hernández datasets increased towards the present and ranged from 25.1 to 132.60 during the Cenozoic (Figure 1). On the other hand, λ (Figure 1) and μ decreased towards the present based on the Li and Ramírez-Barahona datasets (Supplementary Data File 2, Table S2).

The MLEs of the DR parameters based on 432 angiosperm families were similar to those estimated from the original number of families included in each phylogeny (Supplementary Data File 2, Tables S2, S3, respectively), although the diversification and speciation rates were higher in the original phylogeny due to the lack of inclusion of families with low SR.

3.4 Assigning families as having rich, predicted or poor species richness

Using strict consensus criteria across the five studies, 248 angiosperm families were classified as having either predicted SR (161 families, 65% representing 73,803 total species), poor SR (75 families 30.2%, respectively representing 196 species), or high SR (12 families, 4.8%, representing 133,802 total species) (Figures 2A–C; Supplementary Data File 4). Families with poor SR typically had<10 species, while families with high SR had >4,000 species (Supplementary Data File 2, Table S4). Using only those families that had a crown age estimated in the original phylogeny, the trend was similar, though the number of families in the poor category was lower since families with low SR were more likely to be not included in the phylogenetic reconstructions (Supplementary Data File 2, Table S4).

There was no evidence that family SR category was associated with the crown age of the family (Figure 2D). Based on the independence test, only the Bell and Ramírez-Barahona datasets had more families than expected that had high (99–128 mybp
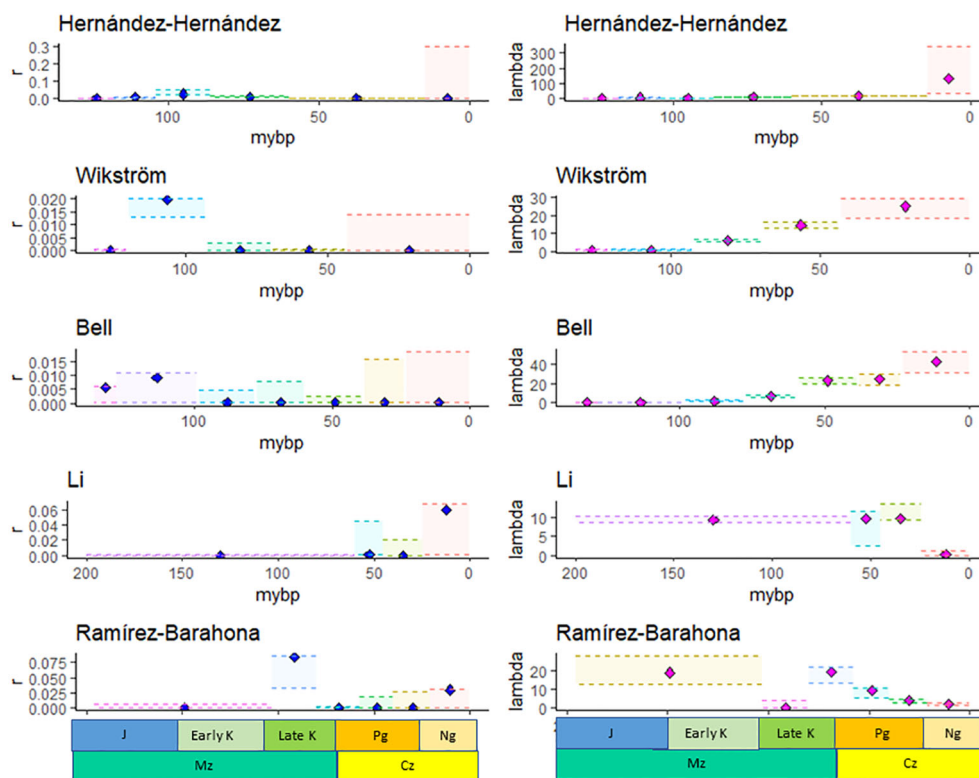
FIGURE 1
The maximum likelihood estimate, MLE (diamonds), and minimum and maximum confidence interval for *the net diversification rate*, r, and the *speciation rate, λ, estimated using the method of* Bokma (2003) *based on the species richness of families given by APG IV and the crown age of families estimated by* Hernández-Hernández and Wiens (2020); Wikström et al. (2001); Bell et al. (2010); Li et al. (2019), *and* Ramírez-Barahona et al. (2020).

period) or low (103.1–196 mybp period) SR within a given geologic interval (Supplementary Data File 2, Tables S5, S6). For the remaining studies, the inferred SR of a family was independent of the period in which the family originated, and was predominantly so for the Bell and Ramírez-Barahona datasets (Supplementary Data File 2, Table S5, S6). As expected, DRs increased towards the present (Figure 3, right panels), but the number of families with low, predicted, or high SR appeared evenly distributed over time when plotted on a log scale (Figure 3, left panels). The Bell dataset showed the same pattern, while in the Ramírez-Barahona dataset, the DR of families tended to increase towards the present, but the speciation rate declined (Supplementary Data File 3, Figure S1). In agreement with this, the Kolmogorov–Smirnov test found no difference in the distribution of median ages of families with poor, predicted, or high SR based on strict consensus criterion ($D = 0.8$, $p > 0.1$; Figure 2D). Furthermore, there is no evidence of phylogenetic clustering of families with poor or high SR (Figure 4).

3.5 Traits associated with families having different species richness

Given that previous studies have shown that lineages experiencing an ancestral WGD event are more likely to exhibit sustained and elevated DRs, we performed an independence test of

the relationship between SR category and the presence of one or more WGD events near the crown age of a family, using the inferred WGD events from Landis et al. (2018). This indicated that families with high SR were more likely to have more than two WGD events, while families with poor SR were more likely to have no WGD events ($\chi^2 = 73.65$, $p < 0.0001$) (Table 2).

Next, we investigated whether angiosperm family SR category was associated with specific traits. We find that SR classification was associated with floral symmetry patterns ($\chi^2 = 35.91$, $p = 0.0047$), and families with poor SR were less likely to have zygomorphic flowers and more likely to have spiral floral symmetry (Data File 2, Table S8). Using the data on the distribution of 20 angiosperm traits from Hernández-Hernández and Wiens (2020) and a global FDR = 0.1 (p-value threshold = 0.005), we found that 8 of the 20 traits exhibited a significant difference in trait value between family SR category, while one showed a trend (Figure 5). Families with poor SR were less likely to have insect pollen dispersal, vertebrate fertilization, and epiphytic taxa; reside in terrestrial habitats; or have an annual life history relative to families with high SR, and were less likely to have a climbing growth habit and an annual life history, and had lower mean seed weight compared to families with predicted SR. There was a clear but non-significant trend that the percent of polyploid species in families increased with SR category (Figure 5). The remaining 11 traits exhibited no significant differences among family SR class (Supplementary Data File 3,
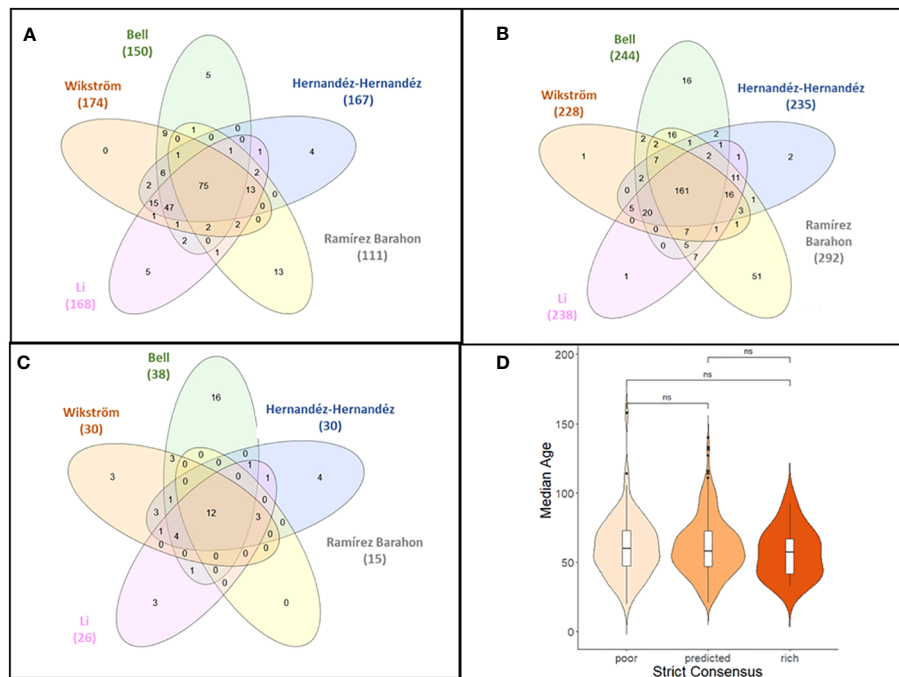
FIGURE 2
(A–C) Venn diagrams showing the levels of agreement between classification of angiosperm families with low **(A)**, expected **(B)**, or high **(C)** levels of species richness as inferred from five datasets. The strict consensus is given in the center. **(D)** Violin plot of the median age of angiosperm families across the five phylogenetic studies employed in the analysis by family classification category (poor, predicted, or rich) based on strict consensus criterion. The distribution of the median crown ages was not significantly different for families with poor, predicted, or high SR using the Kolmogorov–Smirnov test.
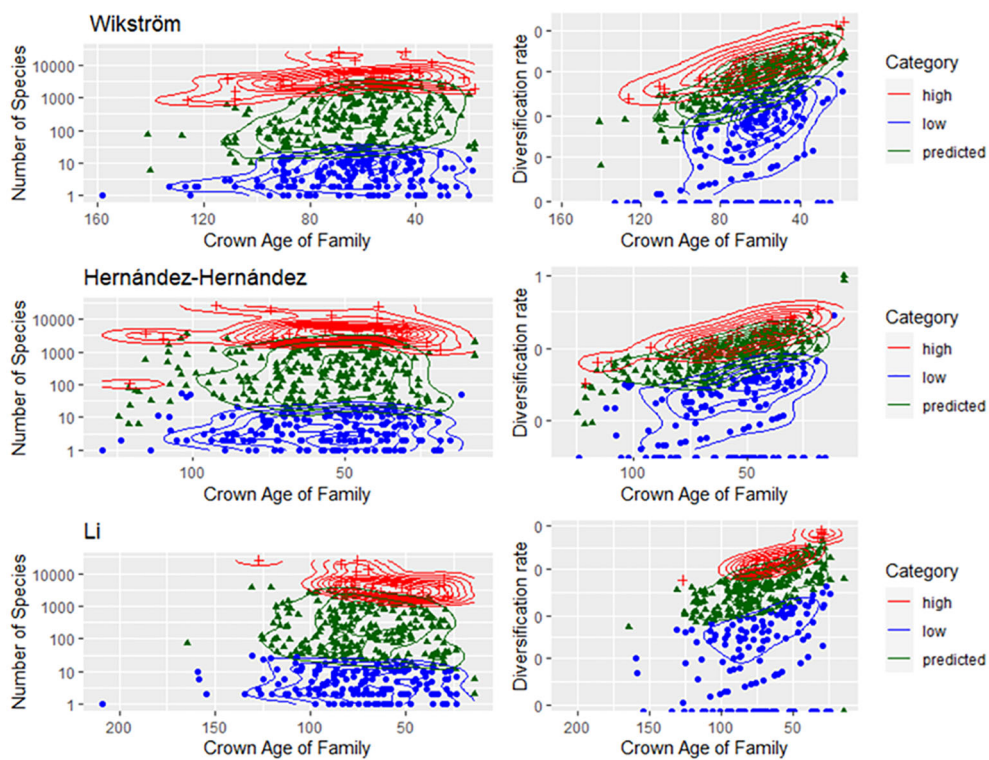

FIGURE 3
Families categorized as having high, predicted, or low SR (left panel) or DR (right panel) by family crown age based on the datasets from Wikström et al. (2001); Hernández-Hernández and Wiens (2020), and Li et al. (2019).
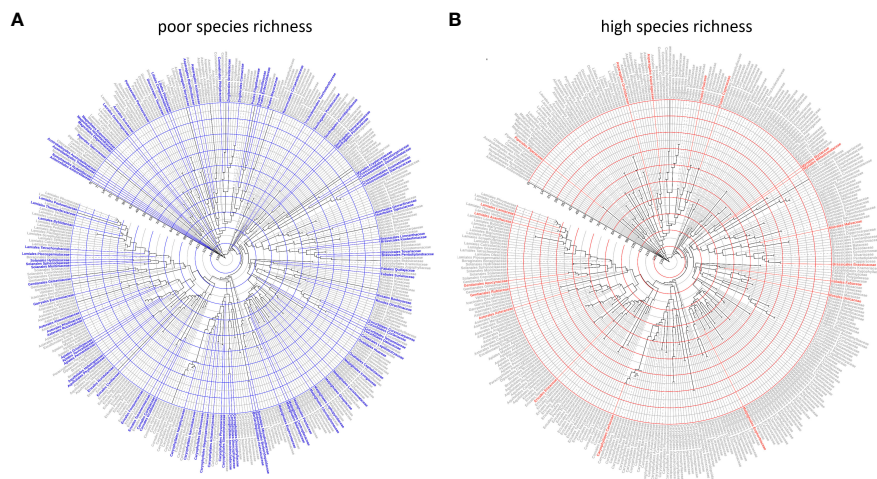
FIGURE 4
Ultrametric molecular phylogeny of angiosperm families based on Hernández-Hernández and Wiens (2020). The family name and external branches are colored by the inferred species richness status, estimated using the MLE modeling described in the methods. The root of the phylogeny (Amborellaceae) is 132 mybp; concentric circles are placed every 14 mybp to the present. **(A)** Families with poor species richness based on majority consensus (4/5 studies in agreement) criterion. **(B)** Families with high species richness based on majority consensus (4/5 studies in agreement) criterion. Solid lines indicate families for which the SR category was declared by strict consensus, and dotted lines denote families for which the SR category was declared based on majority consensus.

Figure S2), and there was no association of family SR with the presence of fleshy fruits (χ^2 = 0.081, P = 0.667, Data File 2, Table S7).

3.6 Association of different mating and sexual systems among angiosperm families with different species richness

Comparison of the sexual and mating systems across the 248 angiosperm families revealed a strong association between family SR and mating and sexual system (χ^2 = 101.72, p< 0.0001). Families with poor SR are more likely to be composed of exclusively dioecious taxa and (for those families that are not dioecious) more likely to have unknown mating systems, while families with predicted SR are less likely to be dioecious (Table 3). Interestingly, families with high SR are more likely to harbor intermediate numbers of taxa with SI systems (SC-SI, Table 3), and are more likely to have SSI (Supplementary Data File 2, Table S9).

TABLE 2 Frequency of families with poor, predicted, rich, or mixed species richness based on strict consensus tabulated by the absence or presence of a WGD event (data from Landis et al., 2018).

WGD	Poor	Predicted	High	Total
None	71	119	2	192
One	3	39	5	47
> Two	1	3	5	9
Total	75	161	12	248

Results are presented for the strict consensus categories across five datasets from angiosperm phylogenies (χ^2 = 73.65, p< 0.0001). Frequencies in blue font are those displaying fewer families than the expected, while those in red are those with more families than expected according to Haberman's residual test (Haberman, 1973).

3.7 Geographic distribution of families differing in species richness

We plotted the mean DR of all georeferenced data for families with poor SR (Figure 6A, top) and high SR (Figure 6A, bottom). The mean DR of georeferenced taxa from families with poor SR varied from −0.04 to +0.04 and families with poor SR appear to be non-randomly distributed in the globe. Keeping in mind that the GBIF database does not equally sample the world's flora, families with poor SR are more represented in Japan and Southern China, parts of Indonesia, Western and Southern Africa, Northern Europe and Spain, and throughout MesoAmerica and North America. On the other hand, the mean DR of georeferenced taxa from families with high SR ranged from 0.11 to 0.249. Families with high SR were dominant in grasslands (particularly in Central Canada and Patagonia) consistent with two families in the Poales having high SR (the Poaceae and Cyperaceae), as well as many mountainous regions, including the Rockies, Andes, Atlas, Himalayas, and others mountain ranges.

Lastly, we used independence tests to assess the association between family SR category and four metrics of geographic distribution: distributional range and pattern, presence in biome, and realm. There was a strong association between family SR and distributional range (χ^2 = 238.67, p< 0.0001), but a modest one for distribution pattern (χ^2 = 6.47, p = 0.0393); families with poor SR were more likely to exhibit highly localized or localized ranges, and were less likely to be widespread, while families with predicted SR exhibited the opposite pattern and those with high SR were more likely to have a cosmopolitan distribution (Data File 2, Tables S10, S11, Data File 3, Figure S3). Similarly, families with poor SR were more likely to be present in a single realm, while families with predicted SR were more likely to found in five to seven (of the eight) realms, and families with high SR were found in all eight realms
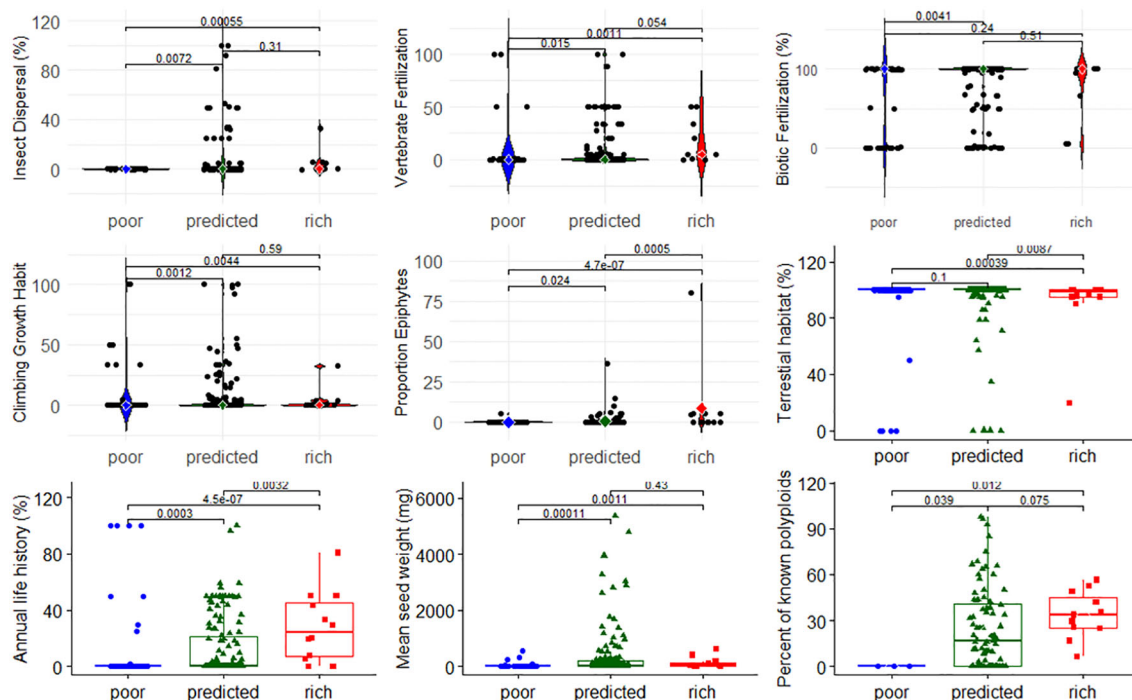
FIGURE 5
Distribution of nine traits across families with poor, predicted, or high species richness based on strict consensus. Differences in the proportion of species per family between SR categories were analyzed using a Kruskal−Wallis, non-parametric rank test, at an FDR = 0.1. The remaining 11 traits with non-significant differences among SR categories are shown in the Supplementary Data File 3 (Figure S2).

(Data File 2, Table S12, Figure 6B). There was a non-random association of family SR category and realm presence (χ^2 = 53.97, p< 0.0001, Data File 2, Tables S13, S14), and families with poor SR were more likely to be located in the Afrotropics and Australasia while families with predicted SR were less likely to be found in Australasia (Figure 6B, Data File 2, Tables S13, S14). Turning to biome number, not surprisingly, families with poor SR were also more likely to be present in 1 or 2–4 biomes while predicted and high SR families were more likely to be found in 11–13 (predicted

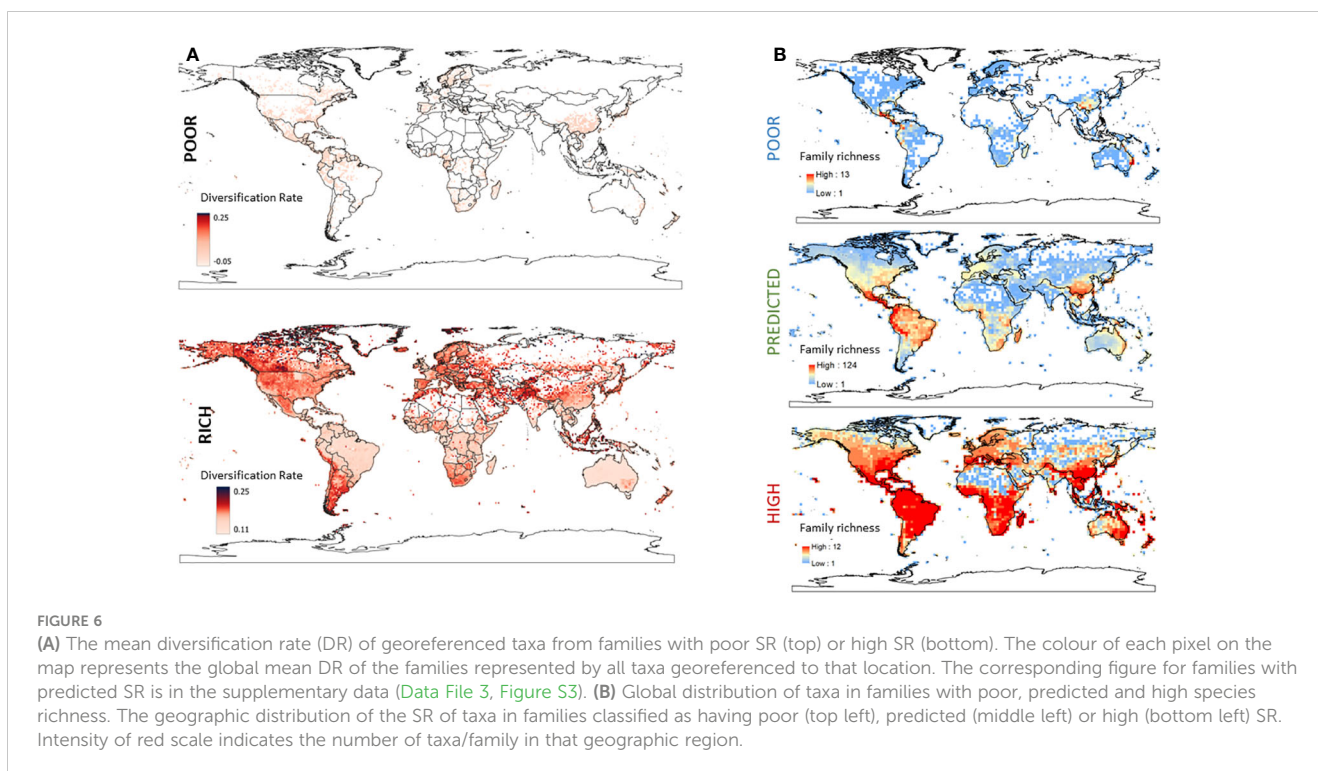SR families) or all 14 biomes (high SR families) (Data File 2, Table S15–S17).

4 Discussion

The study of depauperons, or lineages having a lower-than-expected SR, has received much less attention than the opposite pattern, the evolutionary radiations (Donoghue and Sanderson, 2015; Caron and Pie, 2022). This is partly because monotypic and/or low species-rich families have not been consistently included in molecular phylogenetic studies owing to the difficulty of obtaining specimens. However, much progress has been made to include them in the Tree of life (To), and ongoing efforts to sequence under-represented taxa in the angiosperm tree, such as the oneKB initiative (Leebens-Mack et al., 2019), make the study of depauperons feasible. Another reason why there has been less attention to depauperons lies in the difficulties to assess whether plant families have unusually low SR due to the complexity of diversification rate models, the existence of broad shifts in the inferred DR over evolutionary time, and the wide variation in the estimated crown and stem age of taxa across studies that employ different markers, statistical methods, and fossil data (Sauquet and Magallón, 2018; Benton et al., 2022; Helmstetter et al., 2023). In this study, we aimed to account for these limitations by (1) estimating the age of significant shifts in the DR rate of angiosperm families using Bayesian methods in each of five large phylogenetic studies and then dividing the data into geological intervals within which DRs could be expected to be uniform and then (2) employing the

TABLE 3 Frequency of families with poor, predicted, rich, or mixed species richness based on strict consensus tabulated by the sexual system (data from Renner, 2014) and mating system type (data from Ferrer and Good, 2012).

SI status	Poor	Predicted	Rich	Total
Dioecious	10	2	0	11
SC	8	35	0	47
SC-SI	4	57	12	70
SI	3	30	0	32
Unknown	50	37	0	88
Total	75	161	12	248

Results are presented for the strict consensus categories across five datasets from angiosperm phylogenies (χ^2 = 104.62, p< 0.0001). Frequencies in blue font are those displaying less families than the expected, while those in red are those with more families than expected according to Haberman's residual test (Haberman, 1973).
SC, Self-compatible (>80% of species within the family are self-compatible); SI, Self-incompatible (>80% of species within the family are self-incompatible); and SC-SI, (<80% and >20% of species within the family are self-compatible).

FIGURE 6

(A) The mean diversification rate (DR) of georeferenced taxa from families with poor SR (top) or high SR (bottom). The colour of each pixel on the map represents the global mean DR of the families represented by all taxa georeferenced to that location. The corresponding figure for families with predicted SR is in the supplementary data (Data File 3, Figure S3). **(B)** Global distribution of taxa in families with poor, predicted and high species richness. The geographic distribution of the SR of taxa in families classified as having poor (top left), predicted (middle left) or high (bottom left) SR. Intensity of red scale indicates the number of taxa/family in that geographic region.

approach of Bokma (2003) to obtain MLEs of the DR parameters and their 95% CIs within each geological interval including two factors contributing to measurement uncertainty, (3) using families as the taxonomic unit of choice since plant families continue to be robust to insights from molecular phylogenetic analysis (Soltis et al., 2005), and (4) using strict consensus criteria to categorize family SR, so that our inferences regarding trait evolution were robust to the uncertainties in the inferred crown age of families. In contrast to many other studies that examine the relationship between DR/SR and other variables, our approach is not dependent on the topology of any phylogeny.

Based on the 248 families that met the strict consensus criterion, we find that the majority of them have an expected SR (65%), while interestingly, 30.2% have lower SR than expected, while only 4.8% have an SR that is higher than expected. This points to important inferences from this analysis, namely, that (i) almost 1/3 of angiosperm families have low SR and (ii) despite finding strong evidence for shifts in the mean rate of diversification of angiosperms over geological time, angiosperm family SR is not strongly correlated with clade age, as previously observed in both plants (Magallón and Sanderson, 2001; Vamosi and Vamosi, 2011) and animal lineages (Rabosky et al., 2012).

4.1 Shifts in angiosperm diversification rate

All five of the phylogenetic studies employed in our analyses showed evidence of a shift in the rate of extinction or speciation between 90 and 110 mybp, corresponding to the Albian-Turonian, and usually an increase in the rate of speciation between 50 and 70

mybp, corresponding to the Cretaceous/Paleocene (K/T) boundary, recently called the Angiosperm Terrestrial Revolution (Benton et al., 2022). These results corroborate other studies indicating that angiosperm families radiated mainly during the early Cretaceous and the Cenozoic (Magallón et al., 2015; Sauquet and Magallón, 2018; Li et al., 2019; Onstein, 2020; Ramírez-Barahona et al., 2020), while the diversification of orders (or stem ages) of angiosperms occurred during the Albian-Turonian (Magallón et al., 2015; Ramírez-Barahona et al., 2020). Fossil data show that ~80% of plant species exhibited a sudden extinction at the K/T boundary, which was paralleled by the extinction of many insect species (Johnson et al., 1989; Johnson, 2002; Labandeira et al., 2002). Subsequently, DRs remained low for an estimated 10 my (Wing and Boucher, 1998), but were followed by a radiation during the Paleocene and Eocene (Niklas, 1997; Feild et al., 2004; Soltis et al., 2005).

Cycles in which stasis and sporadic bursts of speciation intermingle are characteristic of angiosperm evolution and have been proposed to be related to the presence of "Greenhouse" and "Icehouse" periods that characterized the earth's global climate before and after the Tertiary period (Willis et al., 2004). During the Paleocene–Eocene transition, high global temperatures reached a maximum at approximately 55.8 mybp (Wolfe, 1978; Zachos et al., 2001), which coincided with an increase in low-latitude palynofloral diversity (Wing and Currano, 2013; Landis et al., 2018), including the diversification and expansion of rainforests worldwide (Morley, 2000; Carvalho et al., 2021; Jaramillo, 2023), and the diversification of biotic taxa that could serve as both pollinators and seed dispersers (Van Der Kooi and Ollerton, 2020; Benton et al., 2022). There is a growing consensus that the increased rate of speciation of

angiosperm families at the Paleocene–Eocene boundary may be associated with (1) a decrease in the number of angiosperm species and therefore an opening of niches after the K/T event, (2) climatic conditions favoring speciation in the early Tertiary to Oligocene, and (3) the diversification of biotic pollinators and seed dispersers that both facilitated and reinforced the evolution of intrinsic plant traits that allowed species to invade new habitats and maintain high genetic diversity while also providing pollinators and seed dispersers access to important resources such as nectar, oils, and pollen for food (Vamosi et al., 2018; Xiao et al., 2021).

4.2 Traits of families with high and low species richness

Based on strict consensus criteria, we identified that 75 angiosperm families are species poor, and 12 had high SR (Supplementary Data File 4). Poor SR families were less likely to have zygomorphic flowers and more likely to have spiral flowers—a finding that is consistent with previous evidence showing that lineages with zygomorphic flowers have elevated rates of diversification potentially driven by plant–pollinator coevolution (Cubas, 2004; Vamosi and Vamosi, 2010). Families with poor SR were also more likely to be dioecious, an association that has been debated, but is consistent with some studies that have argued that the higher extinction and/or lower speciation rates may lead to lower SR of dioecious clades (Heilbuth, 2000; Kay et al., 2006b), an interesting observation given recent evidence that dioecious taxa are more common at higher latitudes (Wang et al., 2021). It should be noted, however, that we only considered families with exclusively dioecious taxa to be dioecious: the relationship between sexual system and DR deserves further attention, especially at lower taxonomic levels (see Sabath et al., 2016). Additionally, we find that families with poor SR were less likely to exhibit biotic fertilization or seed dispersal (lower rates of vertebrate fertilization and insect dispersal); were less likely to have an annual life history, climbing growth habit, or epiphytic taxa; and had a lower mean seed weight (discussed below). Lastly, and consistent with previous studies (Tank et al., 2015; Landis et al., 2018; Ren et al., 2018), our results suggest that families with poor SR are less likely to have undergone an ancestral WGD event while families with high SR are both more likely to have experienced an ancestral WGD and to harbor polyploid taxa. Despite broad evidence that WGD events are associated with lineages exhibiting high DRs in angiosperms, difficulty in placing the timing of WGD events on a phylogenetic tree and tracking how polyploidization directly influenced angiosperm trait evolution continues to be a matter of debate (Clark and Donoghue, 2018; Vamosi et al., 2018; Leebens-Mack et al., 2019).

In contrast to the large number of families with low levels of SR, we find a relatively small number of angiosperm families that exceeded the albeit high rate of diversification in angiosperm families overall. These families, in addition to being cosmopolitan, have a number of interesting characteristics. Families with high SR harbor a higher proportion of annual taxa, which is consistent with the positive relationship between short

generation time and rate of evolution (Aarssen et al., 2006; Crepet and Niklas, 2009). Secondly, families with high SR have a higher proportion of epiphytic taxa, typified by families such as the Orchidaceae and Bromeliaceae; the ability to live in the upper parts of the forest canopy for flowering taxa likely opened new niches with fewer competitors. Third, families with high SR were more likely to have zygomorphic flowers and have SI systems. Families with high SR exhibit large variation in floral morphology and life-history traits (e.g., Fababaeae, Rubiaceae, Asteracee, Poaceae, and Orchidaceae), two factors that have likely been important for their success (Benton et al., 2022). Furthermore, in all of the families with high SR, between 20% and 80% of the taxa for which data were available were identified as self-incompatible or self-sterile and four of these families have SSI, a significantly higher proportion than expected based on the fact that only 17 angiosperm families are exclusively SSI overall (see Ferrer and Good, 2012). Self-incompatible lineages may have historically had high DRs because, in addition to preventing inbreeding, species with SI maintain large effective population sizes due to the negative frequency-dependent selection operating at the S-locus, which serves to maintain connectivity between geographically distant populations and allows for greater opportunities for selection (due to a large N_e) across broad geographic and temporal time scales. These factors are expected to reduce the extinction rate of self-incompatible lineages, and may be associated with high DRs, an association that deserves further attention.

4.3 Biogeographic correlates of angiosperm diversity

In this study, we generated a new database of information regarding the global distribution of flowering plant families and find that families with high SR were exclusively found in 6–8 of the 8 biomes and 11–14 of the 14 realms, while those with poor SR were more likely to be found in a single or a few biomes, and 1–4 of the 14 realms. Thus, biogeographic location appears to be an important factor restricting depauperons. These observations are similar to those made in a study by Vamosi and Vamosi (2011) that found that available area was by far the strongest predictor of SR, explaining ~50% of variation in family SR, while tropical habitat *per se* was not a predictor of SR.

Ramírez-Barahona et al. (2020) examined the stem and crown age of angiosperm families in arid, temperate, and tropical biomes and find that the stem age of angiosperm families in tropical biomes is older than that of families in arid and temperate biomes, a so-called "out of the tropics" scenario of angiosperm diversification. In our investigation of the biomes of 75 families exhibiting poor SR, we find that they were more likely to have localized and highly localized distributions and were predominantly found in the Neotropic, Afrotropic, and Australasia realms. They were also more likely to be found in one or a few biomes, but were only statistically detected to be overrepresented in Tropical and Subtropical Moist Broadleaf Forests biomes. However, visualization of the mean DR of georeferenced taxa belonging to families with poor SR indicates that they are non-randomly distributed in the world.

In a recent study, Hagen et al. (2021) performed a detailed analyses of the factors driving the diversity in tropical moist forests. Tropical moist forests harbor much of the world's biodiversity, but their diversity is not evenly distributed; they are more diverse in the neotropics and Indomalaya than in the Afrotropics. By simulating paleoenvironmental dynamics and macroevolutionary rates, they show that differences in mountain building, aridification (in Africa), and global temperature fluxes shaped historical rates of speciation and extinction resulting in pantropical diversity disparity in both plants and animals. Collectively, this suggests that some of the depauperons we identified belong to families that are less widely distributed and located in ecozones/realms that have overall lower diversity. However, we also find families with poor SR in the neotropics, even though it is home to all of the most species-rich families.

In conclusion, we find that ~1/3 of all angiosperm families are species poor. Intriguingly, the 75 families with poor SR collectively harbor a mere 196 species of the 207,801 species represented by the 248 families included in this study. The families with poor SR have uniformly very low DRs hovering between −0.04 and +0.04 speciation events per million years and appear to be non-randomly distributed around the globe. On the other hand, families with high SR have mean DRs between 0.11 and 0.25 and dominate in mountainous regions and in grasslands (logical since 2 of the 12 high SR families are grasses), though a full analysis is wanting. Although some studies have shown a lack of relationship between SR and DR (Tietje et al., 2022), the manner in which we approached this problem, by first identifying families with high or low SR and then plotting the average DRs of georeferenced taxa belonging to those families, provides a somewhat unique approach to elucidating the evolutionary dynamics of families with unusually low or high DRs. Nevertheless, to understand the evolutionary persistence of depauperons and the broad factors favoring the diversification of angiosperm lineages, further analyses are required. The availability of new geospatial data on the distribution of angiosperms combined with efforts to sequence the ToL should shed light on the particular challenges and roles of plant lineages in responding to future climate change scenarios.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding author.

Author contributions

MF: Conceptualization, Formal analysis, Methodology, Visualization, Writing – original draft, Writing – review & editing, Data curation, Investigation. MV-C: Data curation, Formal analysis, Methodology, Visualization, Writing – review & editing. TH: Data curation, Methodology, Resources, Supervision, Writing – review & editing. SG: Formal analysis, Methodology, Visualization, Writing – review & editing, Conceptualization, Funding acquisition, Project administration, Supervision, Writing – original draft.

Funding

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2023.1276727/full#supplementary-material

SUPPLEMENTARY DATA SHEET 1
Scripts to run R and maple code (.docx).

SUPPLEMENTARY DATA SHEET 2
Supplementary Tables S1–S17 (.docx).

SUPPLEMENTARY DATA SHEET 3
Supplementary Figures S1–S3 (.pdf).

SUPPLEMENTARY DATA SHEET 4
Excel tables describing the date of divergence of angiosperm families inferred from all five studies employed in this analysis, the classification of family species richness (SR), and categorization/value for all traits and geographical distribution examined.

References

Aarssen, L. W., Schamp, B. S., and Pither, J. (2006). Why are there so many small plants? Implications for species coexistence. *J. Ecol.* 94, 569–580. doi: 10.1111/j.1365-2745.2006.01128.x

Bailey, N. T. J. (1964). *The elements of stochastic processes with applications to the natural sciences* (Wiley: New York).

Barnes, B. D., Sclafani, J. A., and Zaffos, A. (2021). Dead clades walking are A pervasive macroevolutionary pattern. *Proc. Natl. Acad. Sci.* 118, E2019208118. doi: 10.1073/pnas.2019208118

Bell, C. D., Soltis, D. E., and Soltis, P. S. (2010). The age and diversification of the angiosperms re-revisited. *Am. J. Bot.* 97, 1296–1303. doi: 10.3732/ajb.0900346

Benton, M. J., Wilf, P., and Sauquet, H. (2022). The angiosperm terrestrial revolution and the origins of modern biodiversity. *New Phytol.* 233, 2017–2035. doi: 10.1111/nph.17822

Bokma, F. (2003). Testing for equal rates of cladogenesis in diverse taxa. *Evolution* 57, 2469–2474.

Caron, F. S., and Pie, M. R. (2022). Arrested diversification? The phylogenetic distribution of poorly-diversifying lineages. *NPJ Biodiversity* 1, 5. doi: 10.1038/s44185-022-00004-0

Carvalho, M. R., Jaramillo, C., de la Parra, F., Caballero-Rodríguez, D., Herrera, F., Wing, S., et al. (2021). Extinction at the end-cretaceous and the origin of modern neotropical rainforests. *Science* 372, 63–68. doi: 10.1126/science.abf1969

Cayuela, L., Granzow-De La Cerda, Í., Albuquerque, F. S., and Golicher, D. J. (2012). Taxonstand: an R package for species names standardisation in vegetation databases. *Methods Ecol. Evol.* 3, 1078–1083. doi: 10.1111/j.2041-210X.2012.00232.x

Clark, J. W., and Donoghue, P. C. J. (2018). Whole-genome duplication and plant macroevolution. *Trends In Plant Sci.* 23, 933–945. doi: 10.1016/j.tplants.2018.07.006

Crepet, W. L., and Niklas, K. J. (2009). Darwin's second "Abominable mystery": why are there so many angiosperm species? *Am. J. Bot.* 96, 366–381. doi: 10.3732/ajb.0800126

Cubas, P. (2004). Floral zygomorphy, the recurring evolution of A successful trait. *Bioessays* 26, 1175–1184. doi: 10.1002/bies.20119

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

Doyle, J. A., and Donoghue, M. J. (1986). Seed plant phylogeny and the origin of angiosperms: an experimental cladistic approach. *Botanical Rev.* 52, 321–431. doi: 10.1007/BF02861082

Eldredge, N., and Stanley, S. M. (1984). "Living fossils: introduction to the casebook," in *Living fossils*. Eds. N. Eldredge and S. M. Stanley (New York, Ny: Springer New York).

Feild, T. S., Arens, N. C., Doyle, J. A., Dawson, T. E., and Donoghue, M. J. (2004). Dark and disturbed: A new image of early angiosperm ecology. *Paleobiology* 30, 82–107. doi: 10.1666/0094-8373(2004)030<0082:DADANI>2.0.CO;2

Ferrer, M. M., and Good, S. V. (2012). Self-sterility in flowering plants: preventing self-fertilization increases family diversification rates. *Ann. Bot.* 110, 535–553. doi: 10.1093/aob/mcs124

Ferrer, M. M., Herbin, J., and Good, S. V. (2014). "Prevalence of self-sterility and other reproductive traits in angiosperm families with high diversification rates," in *Reproductive biology of plants*. Ed. K. G. Ramawat (Boca Raton, Florida, USA, CRC Press). J.-M. M. A. K. S. B. R.

Fiz-Palacios, O., Schneider, H., Heinrichs, J., and Savolainen, V. (2011). Diversification of land plants: insights from A family-level phylogenetic analysis. *BMC Evolutionary Biol.* 11, 341. doi: 10.1186/1471-2148-11-341

Friis, E. M., Pedersen, K. R., and Crane, P. R. (2006). Cretaceous angiosperm flowers: innovation and evolution in plant reproduction. *Palaeogeography Palaeoclimatology Palaeoecol.* 232, 251–293. doi: 10.1016/j.palaeo.2005.07.006

Haberman, S. J. (1973). The analysis of residuals in cross-classified tables. *Biometrics* 29, 205. doi: 10.2307/2529686

Hagen, O., Skeels, A., Onstein, R. E., Jetz, W., and Pellissier, L. (2021). Earth history events shaped the evolution of uneven biodiversity across tropical moist forests. *Proc. Of Natl. Acad. Of Sci.* 118, E2026347118. doi: 10.1073/pnas.2026347118

Heard, S. B., and Hauser, D. L. (1995). Key evolutionary innovations and their ecological mechanisms. *Historical Biol.* 10, 151–173. doi: 10.1080/10292389509380518

Heilbuth, J. C. (2000). Lower species richness in dioecious clades. *Am. Nat.* 156, 221–241. doi: 10.1086/303389

Helmstetter, A. J., Zenil-Ferguson, R., Sauquet, H., Otto, S. P., Méndez, M., Vallejo-Marin, M., et al. (2023). Trait-dependent diversification in angiosperms: patterns, models and data. *Ecol. Lett.* 26, 640–657. doi: 10.1111/ele.14170

Hernández-Hernández, T. (2019). Evolutionary rates and adaptive radiations. *Biol. Philosophy* 34, 41. doi: 10.1007/s10539-019-9694-y

Hernández-Hernández, T., and Wiens, J. J. (2020). Why are there so many flowering plants? A multiscale analysis of plant diversification. *Am. Nat.* 195, 948–963. doi: 10.1086/708273

Höhna, S., May, M. R., and Moore, B. R. (2016). Tess: an R package for efficiently simulating phylogenetic trees and performing bayesian inference of lineage diversification rates. *Bioinformatics* 32, 789–791. doi: 10.1093/bioinformatics/btv651

Jablonski, D. (2007). Scale and hierarchy in macroevolution. *Palaeontology* 50, 87–109. doi: 10.1111/j.1475-4983.2006.00615.x

Jaramillo, C. (2023). The evolution of extant South American tropical biomes. *New Phytol.* 239, 477–493. doi: 10.1111/nph.18931

Johnson, K. R. (2002). *Megaflora of the hell creek and lower fort union formations in the Western Dakotas: vegetational response to climate change, the cretaceous-tertiary boundary event, and rapid marine transgression.* doi: 10.1130/0-8137-2361-2.329

Johnson, K. R., Nichols, D. J., Attrep, M.Jr., and Orth, C. J. (1989). High-resolution leaf-fossil record spanning the cretaceous/tertiary boundary. *Nature* 340, 708–711. doi: 10.1038/340708a0

Kay, K. M., Voelckel, C., Yang, J. Y., Hufford, K. M., Kaska, D. D., and Hodges, S. A. (2006a). "Floral characters and species diversification," in *Ecology and evolution of flowers*. Eds. S.C. H. Barrett and L.D. Harder (Oxford University Press).

Kay, K. M., Voelckel, C., Yang, J. Y., Hufford, K. M., Kaska, D. D., Hodges, S. A. J. E., et al. (2006b). *Floral characters and species diversification.* New York, USA: Oxford University Press. 311–325.

Kenrick, P., and Crane, P. R. (1997). The origin and early evolution of plants on land. *Nature* 389, 33–39. doi: 10.1038/37918

Kubitzki, K. (1990). *Onward. The families and genera of vascular plants, Berlin; Heidelberg* (Springer-Verlag: Germany).

Labandeira, C. C., Johnson, K. R., and Wilf, P. (2002). Impact of the terminal cretaceous event on plant:Insect associations. *Proc. Natl. Acad. Of Sci.* 99, 2061–2066. doi: 10.1073/pnas.042492999

Landis, J. B., Soltis, D. E., Li, Z., Marx, H. E., Barker, M. S., Tank, D. C., et al. (2018). Impact of whole-genome duplication events on diversification rates in angiosperms. *Am. J. Bot.* 105, 348–363. doi: 10.1002/ajb2.1060

Leebens-Mack, J. H., Barker, M. S., Carpenter, E. J., Deyholos, M. K., Gitzendanner, M. A., Graham, S. W., et al. (2019). One thousand plant transcriptomes and the phylogenomics of green plants. *Nature* 574, 679–685. doi: 10.1038/s41586-019-1693-2

Li, H.-T., Yi, T.-S., Gao, L.-M., Ma, P.-F., Zhang, T., Yang, J.-B., et al. (2019). Origin of angiosperms and the puzzle of the jurassic gap. *Nat. Plants* 5, 461–470. doi: 10.1038/s41477-019-0421-0

Magallón, S., Gómez-Acevedo, S., Sánchez-Reyes, L. L., and Hernández-Hernández, T. (2015). A metacalibrated time-tree documents the early rise of flowering plant phylogenetic diversity. *New Phytol.* 207, 437–453. doi: 10.1111/nph.13264

Magallon, S., Hilu, K. W., and Quandt, D. (2013). Land plant evolutionary timeline: gene effects are secondary to fossil constraints in relaxed clock estimation of age and substitution rates. *Am. J. Bot.* 100, 556–573. doi: 10.3732/ajb.1200416

Magallón, S., Sánchez-Reyes, L. L., and Gómez-Acevedo, S. L. (2019). Thirty clues to the exceptional diversification of flowering plants. *Ann. Bot.* 123, 491–503. doi: 10.1093/aob/mcy182

Magallón, S., and Sanderson, M. J. (2001). Absolute diversification rates in angiosperm clades. *Evolution* 55, 1762–1780, 19. doi: 10.1111/j.0014-3820.2001.tb00826.x

Meeker, W. Q., and Escobar, L. A. (1995). Teaching about approximate confidence regions based on maximum likelihood estimation. *Am. Statistician* 49, 48–53. doi: 10.1080/00031305.1995.10476112

Morley, R. J. (2000). *Origin and evolution of tropical rain forests* (John Wiley & Sons: Chichester).

Nee, S. (2006). Birth-death models in macroevolution. *Annu. Rev. Of Ecology Evolution And Systematics* 37, 1–17. doi: 10.1146/annurev.ecolsys.37.091305.110035

Nie, Y., Foster, C. S. P., Zhu, T., Yao, R., Duchene, D. A., Ho, S. Y. W., et al. (2020). Accounting for uncertainty in the evolutionary timescale of green plants through clock-partitioning and fossil calibration strategies. *Syst. Biol.* 69, 1–16. doi: 10.1093/sysbio/syz032

Niklas, K. J. (1997). *The evolutionary biology of plants* (University Of Chicago Press: Chicago).

Olson, D. M., and Dinerstein, E. (1998). The global 200: A representation approach to conserving the earth's most biologically valuable ecoregions. *Conserv. Biol.* 12, 502–515. doi: 10.1046/j.1523-1739.1998.012003502.x

Onstein, R. E. (2020). Darwin's second 'Abominable mystery': trait flexibility as the innovation leading to angiosperm diversity. *New Phytol.* 228, 1741–1747. doi: 10.1111/nph.16294

Rabosky, D. L., Slater, G. J., and Alfaro, M. E. (2012). Clade age and species richness are decoupled across the eukaryotic tree of life. *PloS Biol.* 10, E1001381. doi: 10.1371/journal.pbio.1001381

Ramírez-Barahona, S., Sauquet, H., and Magallón, S. (2020). The delayed and geographically heterogeneous diversification of flowering plant families. *Nat. Ecol. Evol.* 4, 1232–1238. doi: 10.1038/s41559-020-1241-3

Raup, D. M., and Sepkoski, J. J. (1984). Periodicity of extinctions in the geologic past. *Proc. Natl. Acad. Sci.* 81, 801–805. doi: 10.1073/pnas.81.3.801

Ren, R., Wang, H., Guo, C., Zhang, N., Zeng, L., Chen, Y., et al. (2018). Widespread whole genome duplications contribute to genome complexity and species diversity in angiosperms. *Mol. Plant* 11, 414–428. doi: 10.1016/j.molp.2018.01.002

Renner, S. S. (2014). The relative and absolute frequencies of angiosperm sexual systems: dioecy, monoecy, gynodioecy, and an updated online database. *Am. J. Bot.* 101, 1588–1596. doi: 10.3732/ajb.1400196

Reyes, E., Sauquet, H., and Nadot, S. (2016). Perianth symmetry changed at least 199 times in angiosperm evolution. *Taxon* 65, 945–964. doi: 10.12705/655.1

Sabath, N., Goldberg, E. E., Glick, L., Einhorn, M., Ashman, T.-L., Ming, R., et al. (2016). Dioecy does not consistently accelerate or slow lineage diversification across multiple genera of angiosperms. *New Phytol.* 209, 1290–1300. doi: 10.1111/nph.13696

Sauquet, H., and Magallón, S. (2018). Key questions and challenges in angiosperm macroevolution. *New Phytol.* 219, 1170–1187. doi: 10.1111/nph.15104

Sauquet, H., Von Balthazar, M., Magallón, S., Doyle, J. A., Endress, P. K., Bailes, E. J., et al. (2017). The ancestral flower of angiosperms and its early diversification. *Nat. Commun.* 8, 16047. doi: 10.1038/ncomms16047

Soltis, P. S., Endress, P. K., Chase, M. W., and Soltis, D. E. (2005). *Phylogeny and evolution of angiosperms* (Sinauer Associates: Sunderland).

Soltis, P. S., Folk, R. A., and Soltis, D. E. (2019). Darwin review: angiosperm phylogeny and evolutionary radiations. *Proc. R. Soc. B: Biol. Sci.* 286, 20190099. doi: 10.1098/rspb.2019.0099

Stanley, S. M. (1979). Macroevolution. Pattern and process. 332 pp., illustrated. San Francisco: freeman. Price $20.00. Isbn 0 7167 1092 7. *Geological Magazine* 117, 508–509.

Stevens, P. F. (2021) *Angiosperm phylogeny website, version 14*. Available at: http://Www.Mobot.Org/Mobot/Research/Apweb/.

Strathmann, R. R., and Slatkin, M. (1983). The improbability of animal phyla with few species. *Paleobiology* 9, 97–106. doi: 10.1017/S009483730000748X

Su, D., Yang, L., Shi, X., Ma, X., Zhou, X., Hedges, S. B., et al. (2021). Large-scale phylogenomic analyses reveal the monophyly of bryophytes and neoproterozoic origin of land plants. *Mol. Biol. Evol.* 38, 3332–3344. doi: 10.1093/molbev/msab106

Tank, D. C., Eastman, J. M., Pennell, M. W., Soltis, P. S., Soltis, D. E., Hinchliff, C. E., et al. (2015). Nested radiations and the pulse of angiosperm diversification: increased diversification rates often follow whole genome duplications. *New Phytol.* 207, 454–467. doi: 10.1111/nph.13491

The Angiosperm Phylogeny Group, Chase, M. W., Christenhusz, M. J. M., Fay, M. F., Byng, J. W., Judd, W. S., et al. (2016). An update of the angiosperm phylogeny group classification for the orders and families of flowering plants: Apg Iv. *Botanical J. Linn. Soc.* 181, 1–20. doi: 10.1111/boj.12385

Tietje, M., Antonelli, A., Baker, W. J., Govaerts, R., Smith, S. A., and Eiserhardt, W. L. (2022). Global variation in diversification rate and species richness are unlinked in plants. *Proc. Natl. Acad. Sci.* 119, E2120662119. doi: 10.1073/pnas.2120662119

Vamosi, J. C., Magallón, S., Mayrose, I., Otto, S. P., and Sauquet, H. (2018). Macroevolutionary patterns of flowering plant speciation and extinction. *Annu. Rev. Plant Biol.* 69, 685–706. doi: 10.1146/annurev-arplant-042817-040348

Vamosi, J. C., and Vamosi, S. M. (2010). Key innovations within A geographical context in flowering plants: towards resolving darwin's abominable mystery. *Ecol. Lett.* 13, 1270–1279. doi: 10.1111/j.1461-0248.2010.01521.x

Vamosi, J. C., and Vamosi, S. M. (2011). Factors influencing diversification in angiosperms: at the crossroads of intrinsic and extrinsic traits. *Am. J. Bot.* 98, 460–471. doi: 10.3732/ajb.1000311

Van Der Kooi, C. J., and Ollerton, J. (2020). The origins of flowering plants and pollinators. *Science* 368, 1306–1308. doi: 10.1126/science.aay3662

Wang, Y., Luo, A., Lyu, T., Dimitrov, D., Xu, X., Freckleton, R. P., et al. (2021). Global distribution and evolutionary transitions of angiosperm sexual systems. *Ecol. Lett.* 24, 1835–1847. doi: 10.1111/ele.13815

Wikström, N., Savolainen, V., and Chase, M. W. (2001). Evolution of the angiosperms: calibrating the family tree. *Proc. R. Soc. London. Ser. B: Biol. Sci.* 268, 2211–2220. doi: 10.1098/rspb.2001.1782

Willis, K. J., Bennett, K. D., Walker, D., Willis, K. J., and Niklas, K. J. (2004). The role of quaternary environmental change in plant macroevolution: the exception or the rule? *Philos. Trans. R. Soc. London. Ser. B: Biol. Sci.* 359, 159–172. doi: 10.1098/rstb.2003.1387

Wing, S. L., and Boucher, L. D. (1998). Ecological aspects of the cretaceous flowering plant radiation. *Annu. Rev. Earth Planetary Sci.* 26, 379–421. doi: 10.1146/annurev.earth.26.1.379

Wing, S. L., and Currano, E. D. (2013). Plant response to A global greenhouse event 56 million years ago. *Am. J. Bot.* 100, 1234–1254. doi: 10.3732/ajb.1200554

Wolfe, J. A. (1978). A Paleobotanical Interpretation of Tertiary Climates in the Northern Hemisphere: Data from fossil plants make it possible to reconstruct Tertiary climatic changes, which may be correlated with changes in the inclination of the earth's rotational axis. *American Scientist* 66(6), 694–703. https://www.jstor.org/stable/27848958

Xiao, L., Labandeira, C., Dilcher, D., and Ren, D. (2021). Florivory of early cretaceous flowers by functionally diverse insects: implications for early angiosperm pollination. *Proc. R. Soc. B: Biol. Sci.* 288, 20210320.

Zachos, J., Pagani, M., Sloan, L., Thomas, E., and Billups, K. (2001). Trends, rhythms, and aberrations in global climate 65 ma to present. *Science* 292, 686–693. doi: 10.1126/science.1059412

Zhou, Z., and Zheng, S. (2003). The missing link in ginkgo evolution. *Nature* 423, 821–822. doi: 10.1038/423821a