

NEW PHYTOLOGIST SUPPORTING INFORMATION

ARTICLE TITLE:

Island colonization in flowering plants is determined by the interplay of breeding system, lifespan, floral symmetry, and arrival opportunity

AUTHORS:

Annie N. Zell, Charlotte H. Miranda, Erin L. Grady, Dena L. Grossenbacher & Boris Igić

ARTICLE ACCEPTANCE DATE:

7 October 2024.

Methods S1

. Supporting methods.

Data Collection, Taxonomy, and Classification

We assembled a large database of traits and geographic occurrences. It is comprised of previously collected breeding system databases, beginning with Raduski *et al.* (2012), which was selectively enlarged by targeting three families in Grossenbacher *et al.* (2017), now merged, and greatly expanded by new literature surveys. We also expanded the focus on new reproductive traits, lifespan and floral symmetry, which may influence propensity for island arrival and establishment. In order to ensure that trait observations are documented for identical taxa, and are not confused by taxonomic changes, we conducted species-level name synonym resolution for all data records.

For each datum, we recorded the associated taxon name, as provided in the original source, and used Taxonomic Name Resolution Service (Boyle *et al.* 2013, TNRS v4.1) to update and synonymize scientific names. We checked names in the following order of precedence among three taxonomic name databases. We checked World Checklist of Vascular Plants (Govaerts *et al.* 2021) as the primary database, passing unrecognized names to Tropicos (<https://www.tropicos.org>), and finally World Flora Online (<https://www.worldfloraonline.org/>). Updated names that resulted in generic epithet changes were verified by checking the primary literature (taxonomic revision).

We manually collected additional trait and geographic occurrence data for species in our breeding systems database. We recorded data on lifespan and floral symmetry following searches of each species in a variety of topical references. Our main sources included taxon-specific primary literature: Plants of the World Online (POWO 2023), the Global Biodiversity Information Facility (GBIF 2023), eFloras (eFloras 2023), Solanaceae Source (SolSource 2023), Encyclopedia of Life (EOL 2023), and Flora of North America (FNA 2023). If we failed to find the data in these sources, we filled gaps using scanned herbarium sheets and botanical literature (Azani *et al.* 2017; Funk *et al.* 2009; Kubitzki 2004, 2007, 2014; Simpson 2019; Stern 2014). Occasionally, the source data conflicted, so we validated them by additionally searching taxonomic monographs. Lastly, we obtained species-level data for each trait from the TRY database (Kattge *et al.* 2020), the World Checklist of Vascular Plants (accessed via POWO 2023), and BIEN (<https://bien.nceas.ucsb.edu/bien/>; Maitner *et al.* 2018), and compared it to our manually collected trait data to flag any discrepancies and fix errors.

Breeding system

We used two previous literature surveys as a starting point for collection of breeding systems. In Raduski *et al.* (2012) and Grossenbacher *et al.* (2017), we surveyed the literature (as well as provided some original data) accumulating two types of data: (i) closely examined reports of qualitative estimates of breeding systems for 1560 plant species in three families (Asteraceae, Brassicaceae and Solanaceae; see Grossenbacher *et al.* (2017) for quality checking details and quality scores), and (ii) quantitative estimates for 1238 species from 144 families, with transparent calculation procedures from which relative success of selfed and outcrossed hand-pollinations could be calculated (see below). We then extended these databases, by collecting additional published quantitative breeding system data.

We searched Google Scholar (<https://scholar.google.com>) using dozens of search terms related to plant breeding systems, in Bosnian-Croatian-Serbian, English, German, Portuguese, Russian, Spanish, and Standard Chinese. We expanded the search to references cited within each recovered source, which included books, manuscripts, monographs, theses, and personal communications. We include and report results based on data found and entered before July 1st, 2021.

We calculated ISI as the relative fruit set following controlled self- and cross-pollinations:

$$\text{ISI} = 1 - \frac{\text{self-pollination fruit set}}{\text{cross-pollination fruit set}}.$$

This formulation (ISI=1.0 indicates SI) is closely related but practically reversing the one used in Lloyd (1965), who termed it “the relative fraction of positive results on self-pollination” (index of self-compatibility, ISC, where ISC=1.0 indicates SC).

ISI is an imperfect proxy of SI because many unrelated factors can affect fruit set along with various sources of experimental error and bias, particularly when the reported index instead uses the relative seed set (detailed in Raduski *et al.* 2012). Although it is strongly bimodal, species with consistently intermediate values of ISI do occur (Raduski *et al.* 2012). To facilitate their inclusion in the dataset, we discretized them using a traditionally employed cut-off. SI taxa were defined by having $\text{ISI} \geq 0.8$, and species with $\text{ISI} < 0.8$ were considered to be SC. This conservative cut-off may lump SC expressing inbreeding depression with SI species. Therefore, as in Grossenbacher *et al.* (2017), we conducted analyses using a stricter threshold ($\text{ISI} \geq 0.9$), which re-assigns 6% of species in range $0.8 \leq \text{ISI} < 0.9$ to SI, but did not significantly influence any of our main findings (Figure S1).

Such breeding assignment errors very likely conservatively reduce power of our tests, because

they have the effect of randomizing a small proportion of the observed response variables.

Finally, we only included species with consistent reports as either SI or SC. This omits species with conflicting reports on breeding system, which could arise from experimental errors or polymorphic breeding systems. We excluded 53 species with inconsistent qualitative breeding system data and 13 species with conflicting quantitative data from our analyses. An index of self-incompatibility (ISI) could be calculated for 2212 species from 977 genera and 176 families, while 1217 species were scored with binary values (SI or SC).

Lifespan

We recorded lifespan as a proxy for duration of lifetime reproductive bouts. The basic rationale is that many reproductive bouts may ease mate limitation on islands by increasing the probability of arrival and overlap of reproductively compatible individuals over a lifetime. Species were scored as three broad lifespan categories: annual, biennial or perennial. For analyses, annual and biennial species were grouped together to represent shorter lifespans, leaving perennials to represent longer lifespans. We found that 24 species we scored in a third category, polymorphic annual/perennial, and therefore not included in analyses.

Floral symmetry

We collected floral symmetry data as a proxy for generalized and specialized pollination syndromes, because radial flowers are associated with a greater diversity of floral visitors relative to bilateral flowers (Yoder *et al.* 2020). Species with radially symmetric flowers, indicative of a more generalized pollination syndrome, may have an advantage when establishing on islands, where the functional diversity of pollinators is typically lower compared to mainlands (Inoue 1993; Hiraiwa and Ushimaru 2017). Floral symmetry was recorded from species- or genus-level taxonomic descriptions, or assessed visually from images of flowers from multiple sources. If flowers were borderline bilateral or radially symmetric, we scored symmetry as “slightly bilateral” (e.g., radially symmetric perianth with bilaterally symmetric androecium or gynoecium). For species with tight clusters of small flowers (e.g., composite capitula), we considered the symmetry of the entire head rather than each individual flower. Species initially categorized as “slightly bilateral” (155 species) and “asymmetric” (six species) were subsequently reclassified under “bilateral” for all analyses. This grouping was used as a broad approximation of a specialized pollination syndrome. We omitted 96 wind-pollinated species from our analyses, because their reproductive success is not affected by the scarcity of animal pollinators on islands.

Geographic distribution

Island-mainland classification. Following the same classification scheme as Grossenbacher *et al.* (2017), we categorized plant geographical regions as either mainlands or islands using the TDWG encoding. As a way to provide a standard for recording geographic occurrence for each plant species (Brummit 2001, <https://www.tdwg.org/standards/109>), TDWG divides the global terrestrial surfaces into geographic units at four nested levels, with unique codes for each region and nested sub-region. Mainlands were categorized at the TDWG continental Level 1 (North America, South America, Europe, Asia, or Australia). Islands were categorized at a finer resolution, generally corresponding to the TDWG Level 4 (see below). Due to missing precise geographic information, we excluded island-presence or -absence data for 21 species from analyses.

Occurrence data were downloaded from BIEN using the R package BIEN v1.2.3 (Maitner *et al.* 2018; R Core Team. 2023). Searches for 3222 species, using the updated and resolved taxonomy, yielded records for 2500 species. Of the remaining 722 taxa, searches of GBIF yielded data for further 478, resulting in a comprehensive geospatial dataset of 2978 taxa. The 244 species left without location data were mostly island endemics (170) and species with otherwise scarce geographic data. The final geographic distribution dataset used for our key analyses omitted 47 island endemics, and contained 2931 taxa.

Species-island distance. Dispersal rates likely vary with distance between mainland species' ranges and closest islands. Therefore, we calculated a novel proxy for island arrival opportunity based on the distances between each species' native mainland occurrences and the nearest island. For each species, we report the median of these distances to account for skewness or outliers in occurrence data. Island-occurring taxa with high species-island distances likely underwent long-distance dispersal events, while island-occurring taxa with lower species-island distances underwent relatively shorter dispersal events. Species with extensive ranges, especially along continental coastlines, tend to be closer to islands, affording them greater opportunities for island dispersal compared to species with ranges localized to deep interior areas.

We calculated the species-island distance metric using **Generate Near Table** (ArcGIS Pro v3.0.0; Environmental Systems Research Institute, Redlands, California, U.S.A.) to obtain the shortest distance between each species mainland occurrence record and the coastline of any island (Sayre *et al.* 2019). We used a USGS global island shapefile from Sayre *et al.* (2019) to match island definitions in our TDWG Level 4 selection. This metric only represents a rough proxy of

island arrival opportunity as it has several potential flaws. Given the biased sampling of species occurrence data, the location where a particular species is most commonly documented may not reflect where it is most common. Mainland-only species with low species-island distances may have certain traits or climatic requirements that prevent successful island establishment, despite their close proximity to islands. Moreover, oceanic currents, prevailing winds, and direction of cyclones, among many other factors, may affect the relationship between distance and arrival opportunity. Despite these caveats, we included this species-island distance metric in a subset of our models to account for the variable dispersal opportunity between species' mainland ranges and islands, and to evaluate its effect on the strength of island colonization filtering. Specifically, we expect greater effects of mate limitation with increasing distances between species and islands.

Island geological history. The strength of a colonization filter may also depend on the geological history of an island (Grossenbacher *et al.* 2017). We included this proxy in our models to assess the effect of potential dispersal from nearby mainland on the strength of island colonization filtering. For each designated island group, as defined by TDWG Level 4, we recorded whether the group is a part of continental plate, surmising that such islands may have been intermittently connected to the nearby mainland at some point in recent geological history.

On the other hand, oceanic islands (commonly of volcanic origin) are unlikely to have had recent land bridges connecting them to mainland regions. We expect stronger colonization filtering for oceanic islands given that taxa on these islands evidently colonized through long-distance dispersal. The detailed island classification scheme is described in Grossenbacher *et al.* (2017).

To determine the strength of an island colonization filter imposed by mate limitation, we assessed whether SC species are more likely to occur on islands relative to SI species. We used generalized linear mixed models (GLMM), closely following the approach of Grossenbacher *et al.* (2017). All analyses were performed in R (v4.2.1; R Core Team. 2023). GLMMs were implemented in the `lme4` package (v1.1-30; Bates *et al.* 2015), using `glmer` with a logistic link function. To maximize speed and improve model fit, we changed the optimizer from the `lme4` default to BOBQA (Powell *et al.* 2009), after testing among all available optimizers with the `allfit` function in the `lme4` package (Bolker 2008). We used the Akaike Information Criterion (AIC) scores to guide model fitting, comparing models with the `aictab` function in the `aicmodavg` package (v2.3-3).

For all models, the categorical response variable was island “presence” (species that occur on both mainlands and islands; including those endemic to islands in some analyses) or “absence” (species recorded as occurring exclusively on mainlands). The predictor in the each model was

breeding system (SI or SC). We included plant family or genus as a random effect to account for taxonomic non-independence of observations (species).

We replicated the same GLMM analyses to evaluate the robustness of our results with the outlined modifications. To account for the possibility of secondary trait evolution on islands, Model 1b excludes island endemics. To assess the effects of increased taxonomic resolution, Model 1c includes genus nested within family as a random effect. In Model 1d we increased the cutoff for classifying breeding system by changing the ISI threshold from 0.8 to 0.9. To test whether the geological origin influenced our results for island occurring species, the response variable for Model 1e was changed to oceanic island versus continental island. For Model 2 through Model 8, we added variables to the model that we hypothesized interact with breeding system to predict island presence. Model 2 includes breeding system interacting with two traits that we used as a proxy to assess the strength of a mate limitation filter on islands: lifespan and floral symmetry. To account for island arrival opportunity or variable proximities to islands across each species range, Model 3 includes the interaction between breeding system and species-island distance. Models 4-7 build on Model 3, including or excluding trait interactions with breeding system and species-island distance. Model 8 includes a three-way interaction between breeding system, lifespan and species-island distance. Our most complex model, Model 6, could not converge without increasing the maximum number of iterations to 10,000. This adjustment allows more time for the optimization algorithm to find the best fit for model parameters, but risks masking underlying issues within the model (Bates *et al.* 2023).

For each fixed effect in Models 1-8, we extracted the p-value and conservatively accounted for multiple test comparisons ($n = 51$) with the Bonferroni correction using the `p.adjust` function in core R package **STATS**. To obtain the probability of island presence for breeding system, as well as interacting categorical traits, we extracted predicted probabilities and their 95% confidence intervals using the **emmeans** package (v1.8.11) function `emip` (Lenth 2023). We then performed pairwise comparisons of these probabilities using the **emmeans** package function `pairs` with Tukey's adjustment for multiple testing. For the continuous distance-to-island variable, we calculated the probability of island presence using the **cars** package (v 3.1-0) function `predict`, and bootstrapping to generate 95% confidence intervals (**boot** package (1.3-28) function `boot`). Although there was no significant interaction between breeding system and island distance metric, we extracted the probability of island presence separately for SC and SI species, to effectively illustrate the relative advantage of SC over SI for probability of island colonization with increasing distance to nearest island. We used the

MuMIn package (v1.47.1) function `r.squaredGLMM` to calculate the marginal R^2 to obtain the proportion of variance explained by only the fixed effects, and conditional R^2 to obtain the proportion of variance explained by both fixed and random effects (Johnson 2014; Grossenbacher *et al.* 2017). Specifically, to compare the relative magnitudes of each fixed effect as a predictor in our selected model, Model 5, we estimated the proportion of variance explained by breeding system, lifespan, floral symmetry, and species-island distance (`partR2` function in the **partR2** package, v0.9.1; Stoffel *et al.* 2021). Following methods outlined in Stoffel *et al.* (2021), we pooled breeding system interactions with each main effect. The full model R^2 calculated from the **MuMIn** does not exactly match the full model R^2 calculated from the **partR2** package due to differences in calculation methods, but they are well within the range of the 95% confidence intervals from the **partR2** bootstrapping.

Notes S1

. Supporting results.

We assembled a trait database with breeding system, lifespan and floral symmetry for 3222 species. Of these, 1015 (32%) species are recorded on islands, including 217 island endemics, and 2207 (68%) are exclusively recorded on mainlands. Among the island endemics, 151 (70%) are SC and 66 (29%) are SI. Out of the island-occurring species, 471 (46%) are exclusively found on continental islands and 544 (54%) on at least on one oceanic island. Among the 544 oceanic island-occurring species, 375 (69%) are SC and 169 (31%) are SI.

Most of our analyses (Models 1c-5) exclude 217 island endemics and 74 species with missing occurrence data. The remaining 2931 species are from 1019 genera and 162 families. Of this final total, 785 (27%) species are recorded on both islands and mainlands, while the remaining 2146 (73%) are exclusively recorded on mainlands. For breeding system, 1569 (54%) are SC and 1362 (46%) are SI (see Fig. 1a in paper for illustrated summary statistics). Using a more conservative $ISI \geq 0.9$ threshold for SC classification, 1710 are SC (58%) and 1221 (42%) are SI. For lifespan, 577 (20%) are annual or biennial species, while the remaining 2354 are perennial (80%). As expected, a larger proportion of the annual or biennial species are SC (69%) with the remainder SI (31%) (Henslow 1879; Barrett *et al.* 1996; Friedman 2020). For floral symmetry, 715 species (24%) are bilaterally symmetric, and 2216 (76%) are radially symmetric. We also found that a larger proportion of species with bilateral symmetry are SC (65%) with the remainder SI (35%). Additional tabular data summaries, as well as model output and parameters are presented in Tables S1-S14.

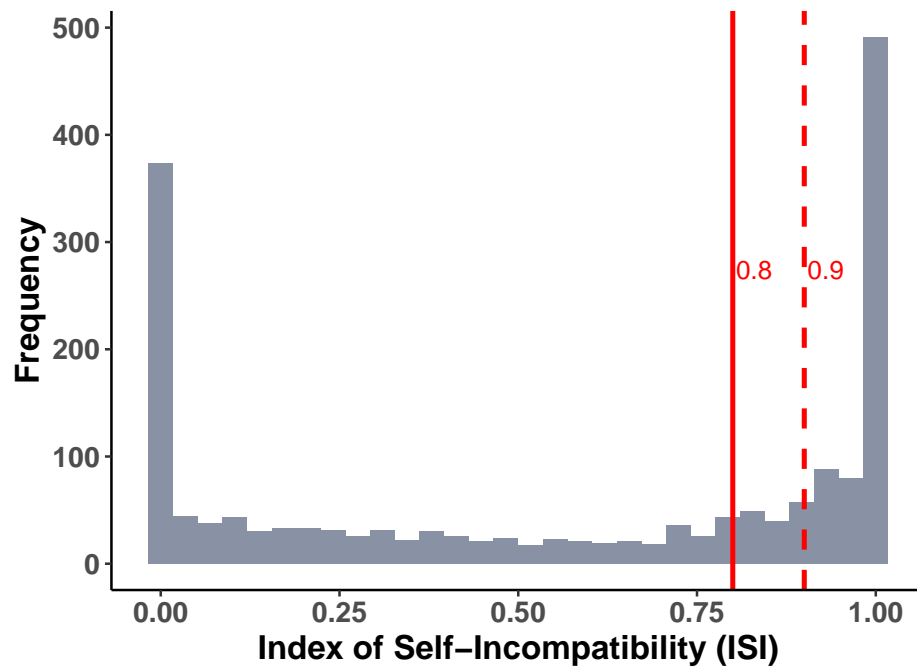


Figure S1: The distribution of the Index of Self-incompatibility (ISI) for all plant species for which selfing and outcrossing data allow its calculation. Imposing either of the conventional cut-off values ($ISI > 0.8$ or $ISI > 0.9$) to arbitrarily define self-incompatible (SI) species, does not strongly affect any results involving breeding system traits.

Table S1: Sample counts representing each listed taxonomic rank.

Taxonomic Rank	Count
Family	162
Genus	1019
Species	2931

Table S2: Breeding system discretization counts by threshold ISI values (2931 species).

Threshold ISI	State	Count	Percentage
0.8	SI	1362	46%
0.8	SC	1569	54%
0.9	SI	1221	42%
0.9	SC	1710	58%

Table S3: Island absence (Mainland Only) and presence (Mainland+Island) data and filtering. Starting with 3222 species with trait information, we removed from our geographic analyses 217 island *endemic* species and 74 species for which occurrence data were not available. (See main text for details.)

Category	Count	Percentage
Island Endemics	217	7%
Missing Data	74	2%
Mainland Only	2146	73%
Mainland+Island	785	27%

Table S4: Island classification (geological origin) for all 1015 island-occurring species in our dataset, including island endemics.

Island Origin	Count	Percentage
Continental	308	30%
Oceanic	707	70%

Table S5: Lifespan trait species count for 2931 sampled species.

Life Span	Count	Percentage
Annual	577	20%
Perennial	2354	80%

Table S6: Breeding system by lifespan joint trait count for 2931 sampled species.

Breeding System	Life Span	Count	Percentage
SI	Annual	176	6%
SI	Perennial	1186	40%
SC	Annual	401	14%
SC	Perennial	1168	40%

Table S7: Floral symmetry trait species count for 2931 sampled species.

Floral Symmetry	Count	Percentage
Bilateral	715	24%
Radial	2216	76%

Table S8: Breeding System by floral symmetry species count for 2931 sampled species.

Breeding System	Floral Symmetry	Count	Percentage
SI	Bilateral	247	8%
SI	Radial	1115	38%
SC	Bilateral	468	16%
SC	Radial	1101	38%

Table S9: Breeding system by island presence trait joint coverage data table for 2931 species in the dataset.

Breeding System	Island Presence	Count	Percentage
SI	Absent	1075	37%
SC	Absent	1071	37%
SI	Present	287	10%
SC	Present	498	17%

Table S10: Breeding system, lifespan, and island presence trait joint coverage data table for 2931 species in the dataset.

Breeding System	Life Span	Island Presence	Count	Percentage
SI	Annual	Absent	146	5%
SI	Perennial	Absent	929	32%
SC	Annual	Absent	242	8%
SC	Perennial	Absent	829	28%
SI	Annual	Present	30	1%
SI	Perennial	Present	257	9%
SC	Annual	Present	159	5%
SC	Perennial	Present	339	12%

Table S11: Breeding system, floral symmetry, and island presence trait joint coverage data table for 2931 species in the dataset.

Breeding System	Floral Symmetry	Island Presence	Count	Percentage
SI	Bilateral	Absent	204	7%
SI	Radial	Absent	871	30%
SC	Bilateral	Absent	375	13%
SC	Radial	Absent	696	24%
SI	Bilateral	Present	43	1%
SI	Radial	Present	244	8%
SC	Bilateral	Present	93	3%
SC	Radial	Present	405	14%

Model (Number, Description)	Species	Families	R^2		Predictor	P	P-adjusted	K	AICc	$\Delta AICc$	ModelLik	AICcWt	lnL	CumWt
			conditional	marginal										
1a. Breeding system (all species)	3222	169	0.26	0.04	Breeding System	9.95E-21	5.08E-19							
	2831	162	0.23	0.03	Breeding System	2.04E-13	1.04E-11	3	3175.352242	164.0174842	2.42E-36	1.17E-36	-1584.672021	1
	2831	1181	0.41	0.03	Breeding System	2.15E-11	1.10E-09							
	2831	162	0.23	0.03	Breeding System	1.71E-12	8.71E-11	3	3179.416104	168.0813459	3.17E-37	1.53E-37	-1586.703952	1
	1015	112	0.17	0.01	Breeding System	0.002915347	0.148682697							
2. Breeding system - trait interactions	2831	162	0.21	0.05	Breeding System	0.014317083	0.730171245	7	3148.073242	136.7384837	2.03E-30	9.77E-31	-1567.017462	1
					Life Span	0.022181667	1							
					Floral Symmetry	0.157025318	1							
					Breeding System:Life Span	0.009241314	0.471307008							
					Breeding System:Floral Symmetry	0.013515626	0.689296906							
3. Breeding system interaction with distance	2831	162	0.31	0.1	Breeding System	8.68E-15	4.43E-13	5	3059.977483	48.64272466	2.74E-11	1.32E-11	-1524.978485	1
					Distance	2.21E-12	1.13E-10							
					Breeding System:Distance	0.470490965	1							
	2831	162	0.29	0.13	Breeding System	0.006127135	0.31248391	9	3027.465102	16.13034413	0.000314297	1.513E-4	-1504.70174	0.99999
					Life Span	0.015270969	0.77881943							
4. Breeding system, trait, and distance interaction					Floral Symmetry	0.072614658	1							
					Distance	6.44E-13	3.28E-11							
					Breeding System:Life Span	0.006184119	0.315390081							
					Breeding System:Floral Symmetry	0.012965021	0.661216083							
					Breeding System:Distance	0.445653233	1							
5. Including lifespan and distance interaction	2831	162	0.29	0.15	Breeding System	0.000899553	0.045877215	10	3011.334758	0	1	0.481321728	-1495.629708	0.4813
					Life Span	0.00089028	0.045404282							
					Floral Symmetry	0.068873269	1							
					Distance	1.13E-12	5.76E-11							
					Breeding System:Life Span	0.000730087	0.037234457							
6. Including floral symmetry and distance interaction					Breeding System:Floral Symmetry	0.01514827	0.772561765							
					Breeding System:Distance	0.094093157	1							
					Life Span:Distance	5.20E-05	0.002649756							
	2831	162	0.29	0.15	Breeding System	0.000985657	0.050268497	11	3013.157789	1.823030711	0.401914719	0.193450287	-1495.533673	0.99985
					Life Span	0.000923443	0.047095592							
7. Removing breeding system and distance interaction					Floral Symmetry	0.080544885	1							
					Distance	1.07E-07	5.45E-06							
					Breeding System:Life Span	0.000761137	0.038818006							
					Breeding System:Floral Symmetry	0.013965885	0.712260137							
					Breeding System:Distance	0.1120433	1							
8. Three-way breeding system-lifespan-distance interaction					Life Span:Distance	7.58E-05	0.003865566							
					Floral Symmetry:Distance	0.657687207	1							
	2831	162	0.29	0.15	Breeding System	0.002126228	0.108437604	9	3012.119709	0.784951305	0.67538279	0.325076411	-1497.029043	0.80640
					Life Span	0.001460211	0.074470774							
					Floral Symmetry	0.081711782	1							
					Distance	9.41E-14	4.80E-12							
					Breeding System:Life Span	0.001270236	0.064782016							
					Breeding System:Floral Symmetry	0.010909649	0.556392111							
					Breeding System:Distance	0.000153803	0.007843935							
	2831	162	0.31	0.13	Breeding System	2.45E-07	1.25E-05	9	3039.934642	28.59988418	6.16E-07	2.97E-07	-1510.93651	1
					Life Span	0.002741901	0.139836958							
					Distance	0.000336392	0.017156017							
					Breeding System:Life Span	0.002899136	0.147855922							
					Breeding System:Distance	0.554405357	1							
					Life Span:Distance	0.062079403	1							
					Breeding System:Life Span:Distance	0.941076858	1							

Table S12: Extended version of Table 1 from the main text showing generalized linear mixed models assessing the effect of breeding system (BS) on island presence or absence. See main text for details on Models 1-8. Marginal (Mar.) r^2 values denote the proportion of variance explained by predictors (fixed effects). For each species, plant family was included as a random effect (genus level analysis model 1c). Conditional (Cond.) r^2 values denote the proportion of variance explained by breeding system and the random factor (family membership). Both P-values (P) and adjusted P-values from Bonferroni correction (P-adjusted) are included. Model selection output, including the number of parameters (K), AIC score (AICc), delta-AIC score, Maximum likelihood (ModelLik), AIC weight (AICcWt), log-likelihood estimate (lnL), and sum of AIC score weights (CumWt) are also provided.

Table S13: Model output with trait interaction probabilities from the generalized linear mixed model (Model 5). LCL and UCL denote lower- and upper 95% confidence limit, respectively.

Breeding System	Trait	Probability	SE	df	LCL	UCL	Interaction
SI	annual	0.0766	0.0204	Inf	0.0451	0.1273	SI-annual
SC	annual	0.2646	0.0351	Inf	0.2017	0.3387	SC-annual
SI	perennial	0.1645	0.0204	Inf	0.1283	0.2085	SI-perennial
SC	perennial	0.2452	0.0239	Inf	0.2013	0.2950	SC-perennial
SI	bilateral	0.0938	0.0217	Inf	0.0591	0.1459	SI-bilateral
SC	bilateral	0.1701	0.0266	Inf	0.1241	0.2286	SC-bilateral
SI	radial	0.1363	0.0213	Inf	0.0996	0.1836	SI-radial
SC	radial	0.3631	0.0320	Inf	0.3031	0.4278	SC-radial

Table S14: Model output parameters: pairwise comparisons for breeding system, lifespan, and floral symmetry traits from the generalized linear mixed model (Model 5).

Contrast	estimate	SE	z.ratio	p.value
SI annual - SC annual	-1.4672	0.2870	-5.1127	1.90E-06
SI annual - SI perennial	-0.8645	0.2602	-3.3231	0.0049
SI annual - SC perennial	-1.3650	0.2812	-4.8547	7.18E-06
SC annual - SI perennial	0.6027	0.1764	3.4175	0.0035
SC annual - SC perennial	0.1022	0.1560	0.6550	0.9138
SI perennial - SC perennial	-0.5005	0.1330	-3.7643	0.0010
SI bilateral - SC bilateral	-0.6830	0.2580	-2.6474	0.0405
SI bilateral - SI radial	-0.4212	0.2316	-1.8193	0.2642
SI bilateral - SC radial	-1.7060	0.2583	-6.6054	2.38E-10
SC bilateral - SI radial	0.2617	0.2244	1.1663	0.6482
SC bilateral - SC radial	-1.0230	0.1872	-5.4645	2.78E-07
SI radial - SC radial	-1.2847	0.1524	-8.4299	3.97E-14

Literature Cited

- Azani N, Babineau M, Bailey CD, Banks H, Barbosa AR, Pinto RB, Boatwright JS, Borges LM, Brown GK, Bruneau A, *et al.* 2017. A new subfamily classification of the Leguminosae based on a taxonomically comprehensive phylogeny: The Legume Phylogeny Working Group (LPWG). *Taxon* **66**: 44–77.
- Barrett SCH, Harder LD, Worley AC. 1996. The comparative biology of pollination and mating in flowering plants. *Philosophical Transactions of the Royal Society of London Series B: Biological Sciences* **351**: 1271–1280.
- Bates D, Mächler M, Bolker B, Walker S. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* **67**: 1–48.
- Bates D, Maechler M, Bolker B, Walker S, Christensen R, Singmann H, Dai B, Scheipl F, Grothendieck G, Green P, *et al.* 2023. lme4: Linear Mixed-Effects Models using ‘Eigen’ and S4 (v1.1-34).
- Bolker BM. 2008. *Ecological models and data in R*. Princeton University Press.
- Boyle B, Hopkins N, Lu Z, Raygoza Garay JA, Mozzherin D, Rees T, Matasci N, Narro ML, Piel WH, Mckay SJ, *et al.* 2013. The taxonomic name resolution service: an online tool for automated standardization of plant names. *BMC Bioinformatics* **14**: 1–15.
- Brummit R. 2001. World geographic scheme for recording plant distributions, edition 2. international working group on taxonomic databases for plant sciences.
- eFloras . 2023. Missouri Botanical Garden (St. Louis, MO) and Harvard University Herbaria (Cambridge, MA).
- EOL . 2023. Encyclopedia of Life [online].
- FNA . 2023. Flora of North America North of Mexico [online].
- Friedman J. 2020. The evolution of annual and perennial plant life histories: ecological correlates and genetic mechanisms. *Annual Review of Ecology, Evolution, and Systematics* **51**: 461–481.
- Funk VA, Susanna A, Steussy TF, Robinson HE. 2009. Classification of Compositae. *Systematics, Evolution, and Biogeography of Compositae* .
- GBIF . 2023. GBIF: The Global Biodiversity Information Facility.
- 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**: 215.
- Grossenbacher DL, Brandvain Y, Auld JR, Burd M, Cheptou P-O, Conner JK, Grant AG, Hovick SM, Pannell JR, Pauw A, *et al.* 2017. Self-compatibility is over-represented on islands. *New Phytologist* **215**: 469–478.
- Henslow G. 1879. On the self-fertilisation of plants. *The Popular Science Review* **3**: VI–XIV.
- Hiraiwa MK, Ushimaru A. 2017. Low functional diversity promotes niche changes in natural island pollinator communities. *Proceedings of the Royal Society B: Biological Sciences* **284**: 20162218.
- Inoue K. 1993. Evolution of mutualism in plant-pollinator interactions on islands. *Journal of Biosciences* **18**: 525–536.
- Johnson PC. 2014. Extension of Nakagawa & Schielzeth’s R2GLMM to random slopes models. *Methods in Ecology and Evolution* **5**: 944–946.
- Kattge J, Bönisch G, Díaz S, Lavorel S, Prentice IC, Leadley P, Tautenhahn S, Werner GD, Aakala T, Abedi M, *et al.* 2020. TRY plant trait database–enhanced coverage and open access. *Global Change Biology* **26**: 119–188.
- Kubitzki K. 2004. *Flowering Plants. Dicotyledons: Celastrales, Oxalidales, Rosales, Cornales, Ericales*, vol. VI. Berlin and Heidelberg: Springer.
- Kubitzki K. 2007. *Flowering Plants. Eudicots: Berberidopsidales, Buxales, Crossosomatales, Fabales p.p., Geraniales, Gunnerales, Myrtales p.p., Proteales, Saxifragales, Vitales, Zygophyllales, Clusiaceae Alliance, Passifloraceae Alliance, Dilleniaceae, Huaceae, Picramniaceae, Sabiaceae*, vol. IX. Berlin and Heidelberg: Springer.