bioRxiv preprint doi: https://doi.org/10.1101/016261; this version posted March 9, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

No association between plant mating system & geographic range overlap

Dena Grossenbacher[2,3], Ryan Briscoe Runquist[3,5], Emma E. Goldberg[4,6], and Yaniv Brandvain[3,7]

[3]Department of Plant Biology, University of Minnesota, St. Paul, MN, 55108, USA; [4]Department of Ecology, Evolution and Behavior, University of Minnesota, St. Paul, MN, 55108, USA

[2]*Corresponding author:* Dena Grossenbacher, Department of Plant Biology, University of Minnesota, St. Paul, MN, 55108, USA; 530-204-7824; dgrossen@umn.edu

[5]rbriscoe@umn.edu, [6]eeg@umn.edu, [7]ybrandvain@gmail.com

bioRxiv preprint doi: https://doi.org/10.1101/016261; this version posted March 9, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Abstract

Premise of the Study: Both evolutionary theory and numerous case studies suggest that selfing taxa are more likely to co-occur with outcrossing relatives than are outcrossing taxa. Despite suggestions that this pattern may be general, the extent to which mating system influences range overlap in close relatives has not been tested formally across a diverse group of plant species pairs.

Methods: We test for a difference in range overlap between species pairs where zero, one, or both species are selfers with data from 98 sister species pairs in 20 genera. We also use divergence time estimates from time-calibrated phylogenies to ask how range overlap changes with divergence time and whether this effect depends on mating system.

Key Results: We find no evidence that self-pollination influences range overlap of closely related species. While the extent of range overlap decreased modestly with the divergence time of sister species, this effect did not depend on mating system.

Conclusions: The absence of a strong influence of mating system on range overlap suggests that of the many mechanisms potentially influencing the co-occurrence of close relatives, mating system plays a minor and/or inconsistent role.

Key words*:* age; co-occurrence; geography; mating system; outcrossing; phylogeny; selfing; speciation.

bioRxiv preprint doi: https://doi.org/10.1101/016261; this version posted March 9, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

INTRODUCTION

From the initiation to the completion of speciation and beyond, mating system can dramatically influence the potential for gene exchange, competition for pollinators, and ecological differentiation, (Antonovics, 1968; Levin, 1972; Jain, 1976; Fishman and Wyatt, 1999; Brandvain and Haig, 2005; Martin and Willis, 2007; Smith and Rausher, 2007; Briscoe Runquist and Moeller, 2013). Consequently, autonomous self-pollination may influence the extent of co-occurrence by altering the predominant geographic mode of speciation, the extent of reproductive isolation between incipient species, and the subsequent sympatric persistence of the two species. As such, mating system could be a key trait influencing the co-occurrence of closely related plant species, potentially serving as a model for understanding the role of functional traits in speciation. Although case studies and evolutionary theory both suggest that selfing can allow closely related plant species to co-occur, the generality of the hypothesis that range overlap is greater in pairs of species in which one or both is selfing has not been tested formally across a diverse group of plant species pairs.

There are at least three plausible geographic scenarios under which autonomous self-pollination would influence the geographic mode of speciation, and be reflected in the range overlap of recently diverged selfing-outcrossing and selfing-selfing sister pairs. In the first scenario, selfing species arise following long distance dispersal. Because autonomous selfing allows rare migrants to colonize and establish (Baker's law; Baker, 1955), a sole migrant experiencing a long-distance dispersal event can give rise to an entire selfing species. Baker's law thus suggests a filter by which mating system may influence speciation. This is thought to be the case in *Capsella* (Foxe et

bioRxiv preprint doi: https://doi.org/10.1101/016261; this version posted March 9, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

al., 2009; Guo et al., 2009; but see Brandvain et al., 2013) and in the sea star *Cryptasterina hystera* (Puritz et al., 2009). In the second scenario, selfing may be favored by selection as a means to provide reproductive assurance in marginal habitats (Lloyd, 1992; Schoen, 1996) just outside of the range of an outcrossing relative. This scenario could result in peripatric speciation, as may be the case in *Clarkia* (Lewis and Lewis, 1955; Moeller and Geber, 2005). In the third scenario, selfing may evolve in a population adapted to a novel habitat directly adjacent to an outcrossing population, and may serve as a mechanism to shield locally adaptive genomes from maladaptive introgression (Levin, 2010). This may be the case in several grass species (Antonovics, 1968), *Mimulus* (Ferris et al., 2014), and *Layia* (Baldwin, 2005). In all three scenarios, selfing may either evolve during the processes in question (producing a selfing-outcrossing sister pair), or it may already be the mating system of the parental species (producing a selfing-selfing sister pair). In all but the first scenario, selfing populations and species arise geographically close to outcrossing relatives, so subsequent small range shifts may lead to range overlap and increased amounts of secondary contact. The relationship between mating system and range overlap may thus depend on the time elapsed since speciation.

Autonomous self-pollination limits gene flow, promotes reproductive isolation, and maintains the distinctness of recently diverged lineages in several ways, each of which could facilitate co-existence. First, and perhaps most importantly, the transition toward selfing is generally associated with reductions in pollinator attraction traits and reduced visitation by pollinators (reviewed in Sicard and Lenhard, 2011), decreasing opportunities for heterospecific pollen movement between predominantly selfing and

bioRxiv preprint doi: https://doi.org/10.1101/016261; this version posted March 9, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

predominantly outcrossing taxa (e.g. Fishman and Wyatt, 1999; Martin and Willis, 2007). In fact, in some cases selfing may evolve or be enhanced following secondary contact as a means to prevent the formation of maladaptive hybrids (reinforcement), as is likely in *Clarkia xantiana* (Briscoe Runquist and Moeller, 2013). In cases where pollen transfer does occur, pollen-pistil incompatibilities and abnormal seed development may pleiotropically follow the evolution of selfing (Brandvain and Haig, 2005; Koelling et al., 2011), further reducing the chance of successful hybridization. Together, these barriers could lead to near-complete reproductive isolation between selfing and outcrossing taxa, preventing their fusion.

In addition to potentially promoting reproductive isolation, selfing can facilitate the ecological coexistence of closely related species by reducing pollinator competition. Many studies across angiosperms document that competition for pollinator services can have massive impacts on fitness, population establishment, and persistence (e.g., Waser, 1978; Fishman and Wyatt, 1999; Brown et al., 2002; Bell et al., 2005; Briscoe Runquist, 2012; Grossenbacher and Stanton, 2014). Predominant selfing may eliminate pollinator-mediated competition by reducing reliance on pollinators altogether, allowing species to coexist and preventing competitive exclusion following secondary range shifts. Experimental field transplants have demonstrated the potential importance of this mechanism of co-existence. In *Mimulus ringens* (Bell et al., 2005), competition for pollinator services with an invasive species caused reduced conspecific pollen deposition; plants compensated for the reduction in fitness through a facultative increase in autonomous selfing. In the typically bee-pollinated *Arenaria* (Fishman and Wyatt, 1999) and *Ipomoea* (Smith and Rausher, 2007), competitive interference due to

bioRxiv preprint doi: https://doi.org/10.1101/016261; this version posted March 9, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

heterospecific pollen transfer from congeners generated female fitness costs that favoured increased selfing.

Thus, many biologically plausible models suggest that selfing and outcrossing species will be relatively likely to co-occur. Numerous compelling case studies support this prediction. For example, in Texas, populations of *Phlox drummondii* showed increased self-compatibility in sympatry with its close relative *P. cuspidata* (Levin, 1985). In Mexico, *Solanum grayi* has dramatically reduced flowers and increased selfing rates when it occurs sympatrically with its close relative *S. lumholtzianum*, a pattern that may exist between other closely related species in this clade (Whalen, 1978; Vallejo-Marin et al., 2014). Similarly, populations of *Arenaria* in the southeastern United States and populations of *Clarkia* in southern California have increased selfing rates in sympatry with closely related congeners (Fishman and Wyatt, 1999; Briscoe Runquist and Moeller 2013). In the genus *Mimulus*, sister species that include one selfing species (selfing-outcrossing sister species) are more likely to occur sympatrically than are outcrossing-outcrossing sister species given similar amounts of divergence time (Grossenbacher and Whittall, 2011). Finally, in a study of Bromeliads in southeastern Brazil, self-compatible species co-occurred with significantly more con-familials than did self-incompatible species (Matallana et al., 2010).

Although these case studies suggest that selfing allows closely related species to co-exist, the influence of mating system on range overlap has not been thoroughly tested at a scale larger than focal genera. Here, we test the hypothesis that selfing promotes co-existence of close relatives by asking whether, across many pairs of sister species of flowering plants from around the world, co-occurrence is greater or lesser for

bioRxiv preprint doi: https://doi.org/10.1101/016261; this version posted March 9, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

pairs where one or both species are selfers. We then use divergence time estimates from time-calibrated phylogenies, to explore whether the extent of co-occurrence changes with divergence time, reflecting the extent of post-speciational range shifts, and whether this effect depends on mating system. Surprisingly, we do not find broad support for the anecdotal relationship between mating system and range overlap.

MATERIALS AND METHODS

We identified genera or generic sections that met the following criteria: 1) a published, species-level phylogeny that contained at least one predominantly selfing (or functionally selfing, e.g., asexual) species and one predominantly outcrossing species, and 2) DNA sequence data for at least 50% of the species within the clade available on GenBank (http://www.ncbi.nlm.nih.gov/genbank/). This resulted in 21 clades whose combined native distributions spanned every continent except Antarctica (Appendix S1). On average, clades contained 35 ±7 (±1SE) extant species, 80 ±4.6 percent of which were included in our phylogenies.

For the analyses described below, all data and R scripts will be made available from the Dryad Digital Repository.

Estimating phylogenies-- We generated time-calibrated phylogenies for all 21 genera or generic sections using publicly available sequence data. We reconstructed phylogenies for two reasons: 1) most previously published phylogenies consisted of only a single topology or consensus tree, making it difficult to incorporate uncertainty into our analysis, and 2) most trees were not time calibrated. Prior to estimating the phylogenies, for each clade separately, we downloaded nrITS sequences for species

bioRxiv preprint doi: https://doi.org/10.1101/016261; this version posted March 9, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

within the clade from GenBank and aligned them using the MUSCLE package in R, version 3.8.31-4 (Edgar, 2004). We simultaneously estimated the phylogenetic relationships and the absolute divergence times among species in a Bayesian framework in BEAST version 1.6.2 (Drummond et al., 2012). To estimate absolute divergence times, we used the mean and range substitution rate for herbaceous and woody plants at the nuclear ribosomal internal transcribed spacer locus (nrITS) (Kay et al., 2006), because fossils are not known for any of the clades in the analysis. The substitution rate was set to a normally distributed prior for herbaceous lineages with mean of 4.13×10^{-9} subs/site/yr and standard deviation of 1.81×10^{-9}, and for woody lineages with mean of 2.15×10^{-9} subs/site/yr and standard deviation of 1.85×10^{-9}.

To accommodate heterogeneity in the molecular evolutionary rate among branches, we used an uncorrelated log-normal (UCLN) relaxed clock model. The prior model on branch lengths was set to a Yule process of speciation. The prior model on substitutions and the number of MCMC generations varied by clade (see Appendix S2). Posterior samples of parameter values were summarized and assessed for convergence and mixing using Tracer v. 1.5 (Rambaut et al., 2014). Of the 21 clades identified above, all met convergence criteria except for *Leavenworthia*, which was excluded from subsequent analyses. For the remaining 20 clades, the minimum estimated sum of squares (ESS) for the posterior was >1100, and the minimum ESS across all other parameters was >600 (Table S1).

We identified sister species in a subset of 9000 trees from the posterior distribution for each clade. For each sister species pair, we recorded the average divergence time and the posterior probability of that pair as the proportion of trees that

bioRxiv preprint doi: https://doi.org/10.1101/016261; this version posted March 9, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

contained that pair, providing a measure of phylogenetic uncertainty. Since our phylogenies sampled, on average, only 80% of extant taxa, these sister pairs may not represent "true" extant sisters, but they are recently diverged groups representing independent evolutionary replicates. For all ensuing analyses, we used sister pairs identified from the phylogenies that had (1) the highest posterior probabilities and (2) did not duplicate species already in the dataset to avoid pseudoreplication.

Estimating mating system, ploidy, and lifespan-- We collated 54 studies which described the mating systems of species from the 20 genera or generic sections identified above. Most published studies classified species as predominantly selfing, mixed mating, or predominantly outcrossing. Species were classified as mixed mating when outcrossing rates within an individual or population were between 0.2 and 0.8, or when there was extensive among population variation in outcrossing rates. An exception to this classification scheme, were species in *Oenothera* sect. *oenothera,* which were classified as either functionally asexual, due to a permanent translocation whereby plants self-fertilize but do not undergo segregation and recombination, or sexual reproduction (Johnson et al., 2009). Sexual *Oenothera* sect. *oenothera* species are partially or wholly self-incompatible, and are assumed to be outcrossing relative to the asexual species. Methods for mating system classification varied among clades because different traits are more reliable indicators of mating system in different taxa; within clades methods were generally consistent (Appendix S3). To extend our data set, we occasionally classified taxa that were missing from the primary studies using the same traits and metrics as those used for other species within that clade (Appendix S3). We then assigned previously identified sister pairs to one of three 'sister pair mating

bioRxiv preprint doi: https://doi.org/10.1101/016261; this version posted March 9, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

system' categories: outcrosser-outcrosser, selfer-outcrosser, or selfer-selfer. Pairs that included mixed mating species were excluded from this analysis.

Mating system may coevolve and be correlated with traits such as polyploidy (Stebbins, 1950; Barringer, 2007, Robertson et al., 2011), and lifespans (Barrett et al., 1996). To ensure that these traits did not drive (or obscure) a relationship between mating system shifts and co-occurrence, we gathered published information on ploidy and lifespan when possible. Species were classified as diploid, polyploid, or mixed when both diploid and polyploid individuals were known. Species' lifespans were classified as annual, perennial, or mixed when both annual and perennial individuals were known

Estimating co-occurrence / geographic range overlap-- We downloaded all known species occurrence records for the clades from the Global Biodiversity Information Facility (http://www.gbif.org) and filtered for quality by excluding records with coordinate accuracy <100 km, coordinates failing to match the locality description, and taxonomic misidentifications (verified by the authors and taxonomic specialists of each clade). We checked species' epithets against the most recently published taxonomies and corrected synonyms and spelling errors. We included only coordinates from the native range of species. Coordinates outside the native species range were identified using published monographs and online databases that report native and invaded ranges (e.g., GRIN database, http://www.ars-grin.gov/).

We used the filtered occurrence data to estimate the degree of co-occurrence using a grid approach. In this approach, we divided the world into a series of rectangular cells by grid lines that follow degree longitude and latitude using the "raster" R package

bioRxiv preprint doi: https://doi.org/10.1101/016261; this version posted March 9, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

version 2.3-0 (Hijmans et al., 2011). We calculated co-occurrence as the summed area of grid cells occupied by both species, divided by the summed area of occupied grid cells for the smaller ranged species. Thus, co-occurrence ranges between 0 (no co-occurrence) and 1 (the two species always co-occur, or the smaller ranged species always co-occurs with the larger ranged species) (Barraclough and Vogler, 2000; Fitzpatrick and Turelli, 2006). In order to assess whether the ensuing analyses were sensitive to the spatial scale at which co-occurrence is estimated, co-occurrence was calculated across a range of cell sizes, 0.05, 0.1, 0.5 and 1 decimal degrees, representing grid cells of approximately 25, 100, 2500, and 10000 km^2 respectively.

Analyses-- To explore whether divergence time varied depending on sister species mating system, we used analysis of variance (ANOVA). To meet model assumptions, divergence time was natural log-transformed prior to analysis. The model took the form:

log divergence time ~ 'sister pair mating system'

where 'sister pair mating system' is a categorical variable with states -- outcrosser-outcrosser, selfer-outcrosser, and selfer-selfer. To incorporate phylogenetic uncertainty into our analysis, and all subsequent models, we included a weighting factor for each sister pair that was equal to the posterior probability of the sister pair (the proportion of phylogenetic trees that contained a given sister pair).

To test whether the mating system of species pairs influences co-occurrence, we used beta regression models in the 'betareg' package in R (Cribari-Neto and Zeileis, 2009). Beta regression provides a flexible model for continuous response variables defined on the interval (0,1) that display both heteroscedasticity and skewness, e.g.,

bioRxiv preprint doi: https://doi.org/10.1101/016261; this version posted March 9, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

proportional data with many values close to zero. Prior to analysis, we used a standard transformation on co-occurrence values (y(n-1) + 0.5/n where n is the sample size, Smithson and Verkuilen, 2006) since in some cases co-occurrence assumed values of 0 and 1. The model took the form:

transformed co-occurrence ~ 'sister pair mating system'

We fit this model using maximum likelihood with a bias correction to determine confidence intervals of the estimated coefficients. We used partial Wald tests to compare among the three mating system categories.

To determine whether time since divergence influences co-occurrence, we used beta regression as in the model described above, and the model took the form:

transformed co-occurrence ~ log divergence time

To determine whether the relationship between co-occurrence and divergence time varied by sister pair mating system, we added mating system and its interaction with divergence time to the above model. The expanded model took the form:

transformed co-occurrence ~ 'log divergence time' + 'sister pair mating system' + interaction

To examine whether our results were robust to the spatial scale at which co-occurrence was determined, we performed all analyses four times using the four different co-occurrence estimates described above. We also ran all analyses including only sister pairs that did not differ in ploidy and lifespan to ensure that our results were not driven by these potential correlated traits. Finally, to explore the possibility that certain clades were heavily influencing overall results, we ran all models described above while sequentially dropping individual clades (N=20). We report cases where

bioRxiv preprint doi: https://doi.org/10.1101/016261; this version posted March 9, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

dropping a single clade altered the significance of any model effects.

RESULTS

We identified 98 sister species pairs from the phylogenetic analysis across 20 genera. Of these pairs, 52 were outcrossing-outcrossing, 30 were selfing-outcrossing, and 16 were selfing-selfing.

Divergence time varied across mating system categories, with outcrosser-outcrosser sister species roughly two times older, on average, than selfer-outcrosser sister species (Fig. 1; overall ANOVA, F=4.962$_{2,95}$, P=0.009; Tukey LSM difference test, outcrosser-outcrosser – selfer-outcrosser P = 0.011, outcrosser-outcrosser – selfer-selfer P = 0.482, selfer-outcrosser – selfer-selfer P = 0.506). There was large variation in co-occurrence for all 'sister pair mating system' categories, especially for young sister pairs.

Patterns of co-occurrence between sister species were not strongly influenced by their mating systems. The distribution of co-occurrences between sister species ranged from zero to one, and it was considerably skewed toward zero across all mating system categories (Fig. 2). Only at the finest spatial scale did mating systems of sister pairs explain even a marginally significant proportion of the variation in co-occurrence (Table 1; Fig. 2) -- 'selfing-selfing' sisters had, on average, about two times greater co-occurrence than outcrossing-outcrossing sisters (P = 0.065). However, this result is largely driven by a single clade, *Medicago*, which contained 5 selfer-selfer pairs. When *Medicago* was dropped from the dataset, the effect of 'selfing-selfing' sisters on co-occurrence disappeared (P = 0.504). These results were not qualitatively different after

bioRxiv preprint doi: https://doi.org/10.1101/016261; this version posted March 9, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

excluding sister pairs that differed in ploidy and life span (results not presented).

Although the distribution of divergence times differed between 'sister pair mating systems' (Fig. 1), the relationship between divergence time and range overlap did not obscure the effect of mating system on co-occurrence. Only at the coarsest spatial scale did divergence time explain even a marginally significant proportion of the variation in co-occurrence (Table 2; Fig. 3) -- there was a trend for co-occurrence to decrease with increased divergence time ($P = 0.064$). Additionally, when including divergence time in the model with 'sister pair mating system', mating system is not significant ($P > 0.286$ in all comparisons, Appendix S4) and the interaction between divergence time and mating system did not influence co-occurrence at any spatial scale ($P > 0.295$ in all comparisons, Appendix S4). Excluding sister pairs that differed in ploidy and life span did not alter these qualitative results (results not presented). Together, our results do not support the hypothesis that mating systems influence range overlap.

DISCUSSION

Autonomous selfing is thought to promote the co-occurrence of closely related taxa for two main reasons. First, the geographic modes of speciation that are hypothesized to accompany selfing may also promote early secondary range overlap. Second, the reproductive isolation and reduced competition for pollinators conferred by selfing facilitates the establishment of a new species by preventing their fusion or competitive exclusion upon secondary contact. In an analysis across 20 clades, however, we find no evidence that the ability to autonomously self-pollinate correlates with range overlap with a closely related selfing or outcrossing species. This result contrasts with

bioRxiv preprint doi: https://doi.org/10.1101/016261; this version posted March 9, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

evidence from a handful of case studies and theoretical intuition. There is also no consistent signal in our data that a particular geographic mode of speciation is associated with autonomous self-pollination. Together, these results suggest that (at least at the level of species pairs) the influence of mating system on patterns of range overlap is weak and/or inconsistent. We therefore conclude that current theory is incomplete, and the convincing case-studies represent particularly interesting evolutionary outcomes that are either inconsistent across taxa or that do not scale up to the macroevolutionary level.

The mechanisms that potentially promote increased co-occurrence between selfers and their sister species are diverse. They include hypotheses such as reduced probability of lineage fusion upon secondary contact, the propensity for selfing to involve speciation in peri- or parapatry, and character displacement or reinforcement selection. Yet, we found no consistent signal in selfing-outcrossing or selfing-selfing sister pairs from diverse angiosperm lineages to support these arguments concerning the influence of mating system on range overlap. Why then have these plausible mechanisms not combined to generate a strong influence of mating system on range overlap? We outline below the potential reasons for the discrepancy between expectations and observations.

Although we detected an overall effect of divergence time on range overlap (an age-range correlation), with recently diverged sister species having on average more range overlap than distantly diverged sisters (Fig. 3), we did not detect an overall predominant geographic mode of speciation for any ‘sister pair mating system’ category. The total variation in overlap explained by divergence time is minimal (pseudo r^2 =

bioRxiv preprint doi: https://doi.org/10.1101/016261; this version posted March 9, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

0.05), and the most recently diverged sister pairs display a wide range of overlaps, from completely allopatric to completely sympatric at the coarsest spatial scale. There was also no significant interaction between divergence time and the mating systems of sister pairs. Although this pattern (a negative age-range correlation) has been widely interpreted as evidence of a 'sympatric' mode of speciation (e.g., Barraclough and Vogler, 2000; Fitzpatrick and Turelli, 2006; Anacker and Strauss, 2014), we caution that such patterns are also influenced by the geographic context of extinction. If extinction is more likely for sister species with sympatric/parapatric ranges (e.g., due to competition), older sisters would tend to be the allopatric survivors, and we would expect a negative age-range correlation -- a pattern that is readily visible in our data.

Our findings did not support any particular geographic mode of speciation associated with the transition to selfing, so our data imply that allopatric, parapatric (or peripatric), and sympatric speciation may all occur for selfing-outcrossing sister pairs. Therefore, the evolutionary transition to selfing may have a more complex influence on the geography of speciation than is generally appreciated. In many verbal and quantitative models of the origin of selfing species (e.g., Grant 1971; Jain, 1976; Lloyd, 1992; Schoen, 1996; Moeller and Geber, 2005), selfers are thought to arise via parapatric (or peripatric) speciation in extreme environments at (or beyond) the margins of the range of an outcrossing relative. In these scenarios, slight perturbations in the range could generate high levels of range overlap, a result inconsistent with our data.

Consideration of ecological differences between selfing and outcrossing species could reconcile our results with prevailing wisdom of the geography of speciation in selfers. That is, if a shift in mating system is associated with local adaptation to the

bioRxiv preprint doi: https://doi.org/10.1101/016261; this version posted March 9, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

distinct niches occupied by selfing and outcrossing plants, range shifts may not necessarily result in sympatry because species will follow their separate niches. Selfing species often exhibit a suite of traits, such as early flowering and drought resistance, that reflect changes to a niche not directly related to the mating system *per se* (Guerrant, 1989; Snell and Aarssen, 2005; Sicard and Lenhard, 2011), suggesting that selfing may be associated with a relatively consistent niche. If selfing and outcrossing species have adapted to different environments, environmental filtering will prevent their occurrence in sympatry. Consistent with this explanation, we found modestly greater co-occurrence for selfing-selfing sister pairs at fine spatial scales in our dataset, particularly in the genus *Medicago*. This suggests the potential for environmental filtering, where selfers are adapted to and occupy locations that lack pollinators altogether, or locations with harsh environments (e.g., thin rapidly drying soils) that favor rapid growth.

Studies in several taxa demonstrate that upon secondary contact, selfing can be favored either as a mechanism to prevent maladaptive hybridization (i.e., reinforcement) or to avoid competition for pollinators (character displacement) (e.g., Fishman and Wyatt, 1999; Smith and Rausher, 2007; Briscoe Runquist and Moeller, 2013). Why then did we not observe excess range overlap between selfing-outcrossing pairs? One potential explanation is that focal studies researching mating system's role in species coexistence were not selected at random, but rather, exceptionally compelling cases were chosen because they highlighted interesting biological phenomena. In our larger data set, these few cases in which selfing facilitated coexistence would instead be overwhelmed by the less compelling cases. Alternatively, mating system may play an important role in maintaining species distinctness upon secondary contact, but

bioRxiv preprint doi: https://doi.org/10.1101/016261; this version posted March 9, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

countervailing forces (e.g., niche convergence in selfing species, see above) could overwhelm this signal.

Taxonomic scale may provide another plausible explanation for the discrepancy between our broad species-level results and system-specific studies. Reinforcement or character displacement on the mating system following secondary contact might be common across angiosperms, but the importance of this process might be limited to population level variation at the microevolutionary scale. According to this explanation, population-level analyses would find an excess of selfing populations in sympatry with populations of a closely related species. With selfing limited to these sympatric populations and not spread across the entirety of the species range however, the species would be considered mixed mating and was excluded from our analysis. This pattern of population-level variation in autonomous selfing rate for sympatric versus allopatric populations is found in many taxa. In the more highly selfing subspecies of *Clarkia xantiana*, *C. x. parviflora*, sympatric populations have smaller flowers with higher selfing rates, probably as a result of reinforcement selection, whereas allopatric populations maintain some ability to receive and export outcrossed pollen (Briscoe Runquist and Moeller, 2013). This is also the case in *Arenaria uniflora*, where there is strong selection for autonomous selfing, and selfing populations only occur in areas of sympatry with its close relative *A. glabra* (Fishman and Wyatt, 1999). This would imply that selfing is a potentially important mechanism underlying coexistence, but that this does not generate a discernable macroevolutionary pattern.

Many of the mechanisms that promote mating system divergence between sister species rely on real or potential genetic introgression or competition. In our analysis, the

bioRxiv preprint doi: https://doi.org/10.1101/016261; this version posted March 9, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

finest spatial scale (0.05 decimal degree grid cells) corresponds to roughly 25km^2. Given that most cross-pollination happens locally at scales much smaller than even the smallest spatial scale assessed here, this suggests that at least 60% of sister species (those completely allopatric at the finest scale and perhaps even more of the sympatric pairs; Fig. 2) do not typically have the opportunity for present-day pollen exchange, potential introgression, or reproductive interference competition. Thus, the majority of sister species likely experience limited present-day gene exchange, which may obscure any patterns driven by mating system divergence. It should be noted, however, that the likelihood may be much higher that sister pairs encountered each other during the speciation process.

A final explanation for the lack of a relationship between range overlap and mating system is that mating system is simply one of myriad potential mechanisms that allow close relatives to co-exist. Of the ~40% of sister pairs that co-occur in the same grid cell in our fine-scale analysis, habitat differences, flowering time differences, pollinator shifts, and post-pollination incompatibilities will also prevent hybridization or competition. For example, it is possible that pollinator shifts in outcrossing-outcrossing sister species facilitate their co-existence. Like selfing, pollinator shifts can influence reproductive isolation, competition, and the geography of speciation (reviewed in Kay and Sargent, 2009). For instance, if sympatric outcrossing-outcrossing species pairs are enriched for pollinator shifts, then perhaps both selfing and pollinators may encourage co-existence. Pushing this argument one step further, perhaps all sympatric close relatives have diverged in some key trait that allows their co-existence, and therefore a test of any particular trait across a large number of angiosperm species

bioRxiv preprint doi: https://doi.org/10.1101/016261; this version posted March 9, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

pairs will not uncover a systematic effect of any given trait.

Conclusion: the influence of mating systems on co-occurrence-- Ultimately, we find no evidence for mating system consistently influencing the geography of speciation or secondary range overlap -- although mating system has a major effect on sympatry in some case studies, there is no discernable effect across the 20 genera and generic sections examined here. Instead, co-occurrence of close relatives may be influenced by many mechanisms, of which transitions to selfing are only a small part. It is also possible that the evolution of selfing is associated with reproductive assurance during the adaptation to marginal or mate-limited habitats and is therefore concomitant with other adaptations that preclude general co-occurrence. Alternatively, selection for selfing in secondary contact may be only a population level phenomenon that does not influence species-level patterns of co-occurrence. Greater understanding of the evolutionary causes of the transition to selfing is necessary to determine the general influence of mating system on co-occurrence.

References

ANTONOVICS, J. 1968. Evolution in closely adjacent plant populations V. evolution of self-fertility. *Heredity* 23: 219–238.

BAKER, H.G. 1955. Self compatibility and establishment after long distance dispersal. *Evolution* 9: 347–349.

BALDWIN, B.G. 2005. Origin of the serpentine-endemic herb Layia discoidea from the widespread L. glandulosa (Compositae). *Evolution* 59: 2473–9.

bioRxiv preprint doi: https://doi.org/10.1101/016261; this version posted March 9, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

BARRACLOUGH, T.G., AND A.P. VOGLER. 2000. Detecting the Geographical Pattern of Speciation from Species-Level Phylogenies. *The American Naturalist* 155: 419–434.

BARRETT, S.C.H., L.D. HARDER, AND A.C. WORLEY. 1996. The comparative biology of pollination and mating in flowering plants. *Philosophical Transactions of the Royal Society B: Biological Sciences* 351: 1271–1280.

BARRINGER, B. 2007. Polyploidy and self-fertilization in flowering plants. *American Journal of Botany* 94: 1527–1533.

BELL, J.M., J.D. KARRON, AND R.J. MITCHELL. 2005. Interspecific competition for pollination lowers seed production and outcrossing in Mimulus ringens. *Ecology* 86: 762–771.

BRANDVAIN, Y., AND D. HAIG. 2005. Divergent mating systems and parental conflict as a barrier to hybridization in flowering plants. *The American Naturalist* 166: 330–338.

BRANDVAIN, Y., T. SLOTTE, K.M. HAZZOURI, S.I. WRIGHT, AND G. COOP. 2013. Genomic identification of founding haplotypes reveals the history of the selfing species Capsella rubella. *PLoS genetics* 9: .

BRISCOE RUNQUIST, R.D. 2012. Pollinator-mediated competition between two congeners, Limnanthes douglasii subsp. rosea and L. alba (Limnanthaceae). *American journal of botany* 99: 1125–32.

BRISCOE RUNQUIST, R.D., AND D.A. MOELLER. 2013. Floral and mating system divergence in secondary sympatry: testing an alternative hypothesis to reinforcement in Clarkia. *Annals of botany*.

BROWN, B.J., R.J. MITCHELL, AND S.A. GRAHAM. 2002. Competition for pollination between an invasive species (Purple Loosestrife) and a native congener. *Ecology* 83: 2328–

bioRxiv preprint doi: https://doi.org/10.1101/016261; this version posted March 9, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

2336.

CRIBARI-NETO, F., AND A. ZEILEIS. 2009. Beta regression in R.

DRUMMOND, A.J., M.A. SUCHARD, D. XIE, AND A. RAMBAUT. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular biology and evolution* 29: 1969–73.

EDGAR, R.C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic acids research* 32: 1792–1797.

FERRIS, K.G., J.P. SEXTON, AND J.H. WILLIS. 2014. Speciation on a local geographic scale: the evolution of a rare rock outcrop specialist in Mimulus. *Philosophical Transactions of the Royal Society B-Biological Sciences* 369: 27–29.

FISHMAN, L., AND R. WYATT. 1999. Pollinator-mediated competition, reproductive character displacement, and the evolution of selfing in Arenaria uniflora (Caryophyllaceae). *Evolution* 53: 1723–1733.

FITZPATRICK, B.M., AND M. TURELLI. 2006. The geography of mammalian speciation: Mixed signals from phylogenies and range maps. *Evolution* 60: 601–615.

FOXE, J.P., T. SLOTTE, E.A. STAHL, B. NEUFFER, H. HURKA, AND S.I. WRIGHT. 2009. Recent speciation associated with the evolution of selfing in Capsella. *Proceedings of the National Academy of Sciences* 106: 5241–5245.

GRANT, V. 1971. Plant speciation. *New York*.

GROSSENBACHER, D.L., AND M.L. STANTON. 2014. Pollinator-mediated competition influences selection for flower-color displacement in sympatric monkeyflowers. *American journal of botany* 101: 1915–24.

GROSSENBACHER, D.L., AND J.B. WHITTALL. 2011. Increased floral divergence in sympatric monkeyflowers. *Evolution* 65: 2712–2718.

bioRxiv preprint doi: https://doi.org/10.1101/016261; this version posted March 9, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

GUERRANT, E.O. 1989. Early maturity, small flowers and autogamy: a developmental connection. *The evolutionary ecology of plants* 61: 84.

GUO, Y.-L., J.S. BECHSGAARD, T. SLOTTE, B. NEUFFER, M. LASCOUX, D. WEIGEL, AND M.H. SCHIERUP. 2009. Recent speciation of Capsella rubella from Capsella grandiflora, associated with loss of self-incompatibility and an extreme bottleneck. *Proceedings of the National Academy of Sciences* 106: 5246–5251.

HIJMANS, R.J., S. PHILLIPS, J. LEATHWICK, AND J. ELITH. 2011. dismo: Species distribution modeling. R package version 0.7–23.

JAIN, S. 1976. The evolution of inbreeding in plants. *Annual Review of Ecology and Systematics* 7: 469–495.

JOHNSON, M.T.J., S.D. SMITH, AND M.D. RAUSHER. 2009. Effects of plant sex on range distributions and allocation to reproduction. *The New phytologist* 186: 769–79.

KAY, K.M., AND R.D. SARGENT. 2009. The role of animal pollination in plant speciation: integrating ecology, geography, and genetics. *Annual Review of Ecology Evolution and Systematics* 40: 637–656.

KAY, K.M., J.B. WHITTALL, AND S.A. HODGES. 2006. A survey of nuclear ribosomal internal transcribed spacer substitution rates across angiosperms: an approximate molecular clock with life history effects. *BMC Evolutionary Biology* 6: 36.

KOELLING, V.A., J.L. HAMRICK, AND R. MAURICIO. 2011. Genetic diversity and structure in two species of Leavenworthia with self-incompatible and self-compatible populations. *Heredity* 106: 310–318.

LEVIN, D. 1972. Competition for pollinator service: A stimulus for the evolution of autogamy. *Evolution* 26: 668–669.

bioRxiv preprint doi: https://doi.org/10.1101/016261; this version posted March 9, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

LEVIN, D. 1985. Reproductive Character Displacement in Phlox. *Evolution* 39: 1275–1281.

LEVIN, D. A. 2010. Environment-enhanced self-fertilization: implications for niche shifts in adjacent populations. *Journal of Ecology* 98: 1276–1283.

LEWIS, H., AND M.E. LEWIS. 1955. The genus Clarkia. University of California Press Berkeley, California, USA.

LLOYD, D. 1992. Self- and cross-fertilization in plants. II. The selection of self-fertilization. *International Journal of Plant Sciences* 153: 370–380.

MARTIN, N.H., AND J.H. WILLIS. 2007. Ecological divergence associated with mating system causes nearly complete reproductive isolation between sympatric Mimulus species. *Evolution* 61: 68–82.

MATALLANA, G., M. GODINHO, F. GUILHERME, M. BELISARIO, T. COSER, AND T. WENDT. 2010. Breeding systems of Bromeliaceae species: evolution of selfing in the context of sympatric occurrence. *Plant Systematics and Evolution* 289: 57–65.

MOELLER, D., AND M. GEBRE. 2005. Ecological context of the evolution of self-pollination in Clarkia xantiana: population size, plant communities, and reproductive assurance. *Evolution* 59: 786–799.

PURITZ, J.B., C.C. KEEVER, J. A ADDISON, M. BYRNE, M.W. HART, R.K. GROSBERG, AND R.J. TOONEN. 2012. Extraordinarily rapid life-history divergence between Cryptasterina sea star species. *Philosophical Transactions of the Royal Society B-Biological Sciences* 279: 3914–22.

RAMBAUT, A., M. SUCHARD, D. XIE, AND A. DRUMMOND. 2014. Tracer v1.6. *Available from http://beast.bio.ed.ac.uk/Tracer*.

ROBERTSON, K., E.E. GOLDBERG, AND B. IGIĆ. 2011. Comparative evidence for the

bioRxiv preprint doi: https://doi.org/10.1101/016261; this version posted March 9, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

correlated evolution of polyploidy and self-compatibility in Solanaceae. *Evolution* 65: 139–155.

SCHOEN, D. 1996. How does self-pollination evolve? Inferences from floral ecology and molecular genetic variation. *Philosophical Transactions of the Royal Society B-Biological Sciences* 351: 1281–1290.

SICARD, A., AND M. LENHARD. 2011. The selfing syndrome: a model for studying the genetic and evolutionary basis of morphological adaptation in plants. *Annals of Botany* 107: 1433–14443.

SMITH, R.A., AND M.D. RAUSHER. 2007. Close clustering of anthers and stigma in Ipomoea hederacea enhances prezygotic isolation from Ipomoea purpurea. *New Phytologist* 173: 641–647.

SMITHSON, M., AND J. VERKUILEN. 2006. A better lemon squeezer? Maximum-likelihood regression with beta-distributed dependent variables. *Psychological methods* 11: 54.

SNELL, R., AND L.W. AARSSEN. 2005. Life history traits in selfing versus outcrossing annuals: exploring the “time-limitation” hypothesis for the fitness benefit of self-pollination. *BMC ecology* 5: 2.

STEBBINS, G.L. 1950. Variation and evolution in plants. Columbia University Press, New York, London.

VALLEJO-MARÍN, M., C. WALKER, P. FRISTON-REILLY, L. SOLIS-MONTERO, AND B. IGIC. 2014. Recurrent modification of floral morphology in heterantherous Solanum reveals a parallel shift in reproductive strategy. *Philosophical Transactions of the Royal Society B-Biological Sciences* 369: 20130256.

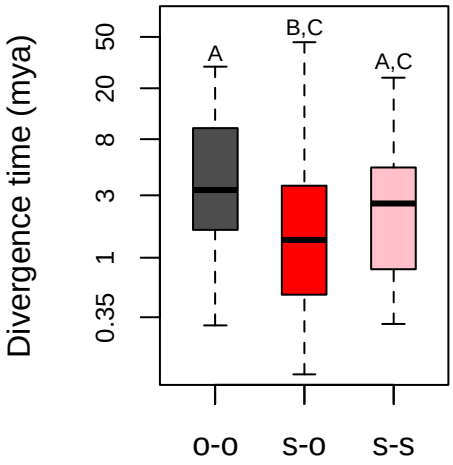
bioRxiv preprint doi: https://doi.org/10.1101/016261; this version posted March 9, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

WASER, N. 1978. Competition for hummingbird pollination and sequential flowering in two Colorado wildflowers. *Ecology* 59: 934–944.

WHALEN, M.D. 1978. Reproductive Character Displacement and Floral Diversity in Solanum Section Androceras. *Systematic Botany* 3: 77–86.

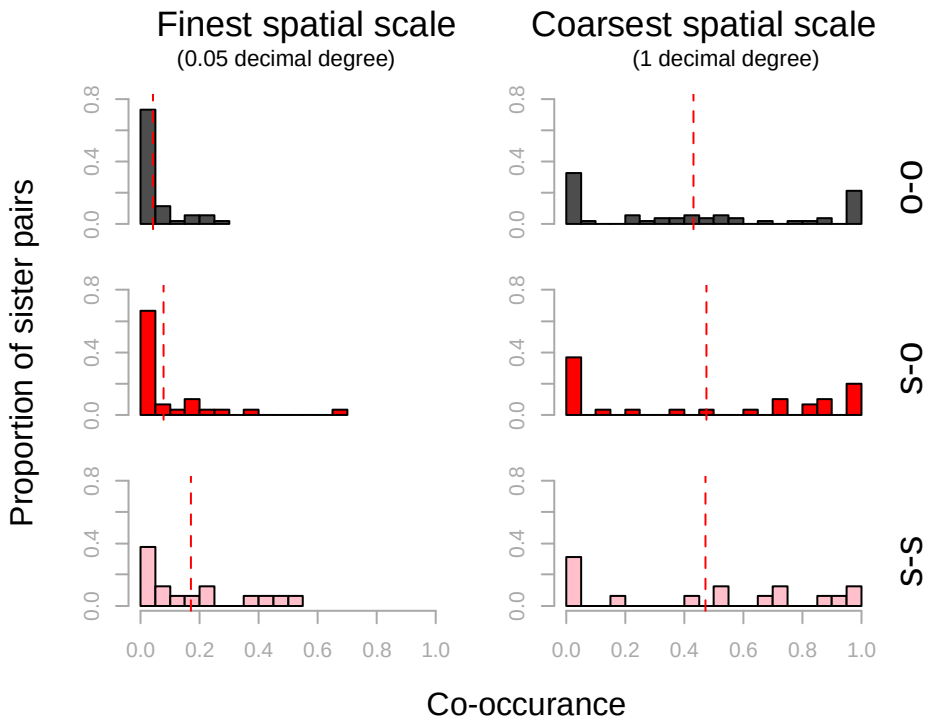
bioRxiv preprint doi: https://doi.org/10.1101/016261; this version posted March 9, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 1. Box plots of sister pair divergence times by mating system category: outcrossing-outcrossing (o-o, dark gray), selfing-outcrossing (s-o, red), selfing-selfing (s-s, pink). Letters represent *a posteriori* Tukey groupings; see text for ANOVA summary. Divergence time axis is natural logarithmic scale (back-transformed).



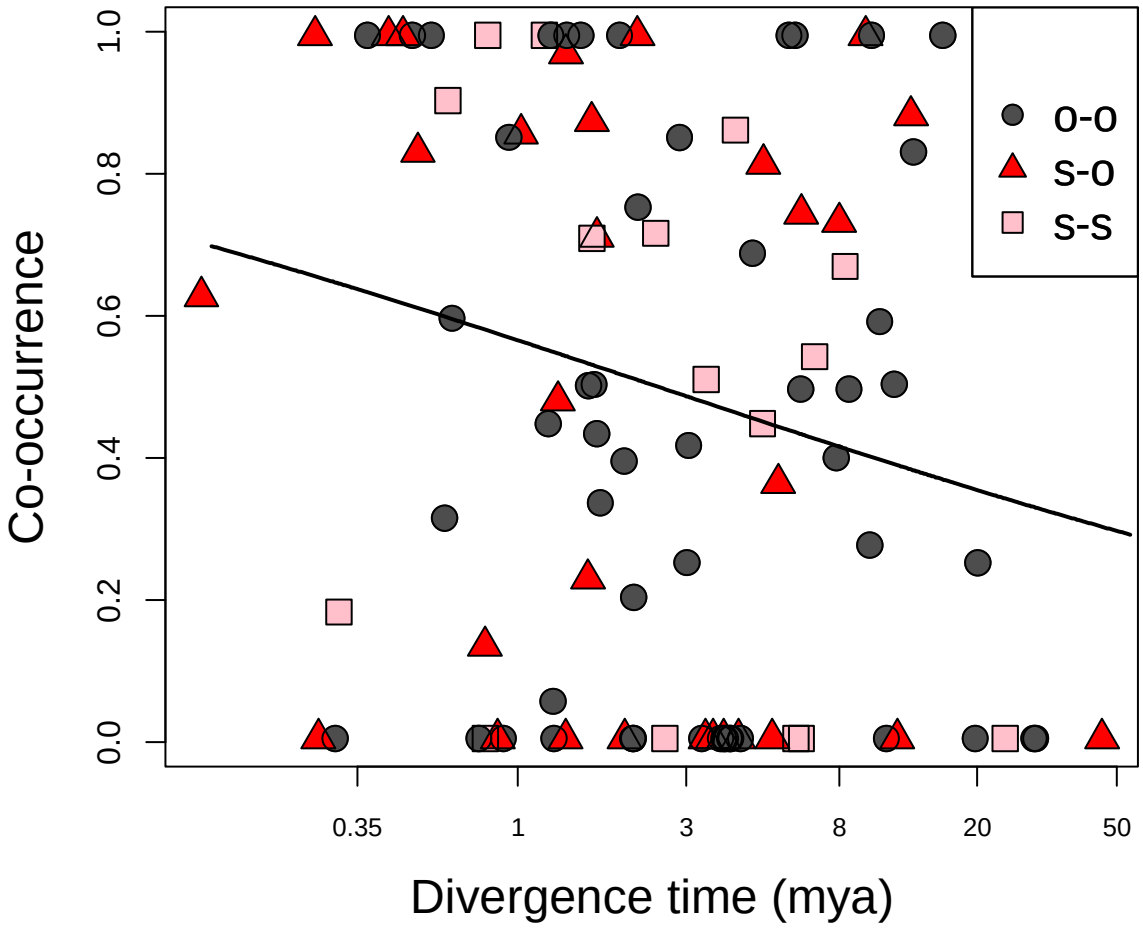
bioRxiv preprint doi: https://doi.org/10.1101/016261; this version posted March 9, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 2. Histograms of sister pair co-occurrence by mating system category: outcrossing-outcrossing (o-o, dark gray), selfing-outcrossing (s-o, red), selfing-selfing (s-s, pink). Dashed vertical lines indicate mean co-occurrence. See Table 1 for statistical results.



bioRxiv preprint doi: https://doi.org/10.1101/016261; this version posted March 9, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 3. Co-occurrence at the coarsest spatial scale (1 decimal degree) by divergence time for 98 sister species across 20 clades. The line segment represents the predicted slope from beta regression. Divergence time axis is natural logarithmic scale (back-transformed). See Table 2 for statistical results.



bioRxiv preprint doi: https://doi.org/10.1101/016261; this version posted March 9, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table 1. Results of beta regression models analyzing the effect of 'sister pair mating system' on co-occurrence, estimated at four spatial scales (**A-D**).

	Coefficient (s.e.)	*Wald's z-value*	*P*
A. 0.05 decimal degree (~25 km^2)			
Intercept	-2.371 (0.253)	-9.377	<0.001
selfer-outcrosser	0.031 (0.315)	0.098	0.922
selfer-selfer	0.646 (0.350)	1.844	0.065
B. 0.1 decimal degree (~100 km^2)			
Intercept	-2.003 (0.254)	-7.878	<0.001
selfer-outcrosser	0.209 (0.326)	0.640	0.522
selfer-selfer	0.606 (0.369)	1.643	0.100
C. 0.5 decimal degree (~2500 km^2)			
Intercept	-0.581 (0.268)	-2.168	0.030
selfer-outcrosser	0.482 (0.413)	1.168	0.243
selfer-selfer	0.211 (0.462)	0.457	0.648
D. 1 decimal degree (~10,000 km^2)			
Intercept	-0.071 (0.271)	-0.263	0.792
selfer-outcrosser	0.220 (0.420)	0.523	0.601
selfer-selfer	0.084 (0.476)	0.177	0.860

Note: The categorical coefficient estimates are log-odds ratios, measured as departures from the 'outcrosser-outcrosser' category.

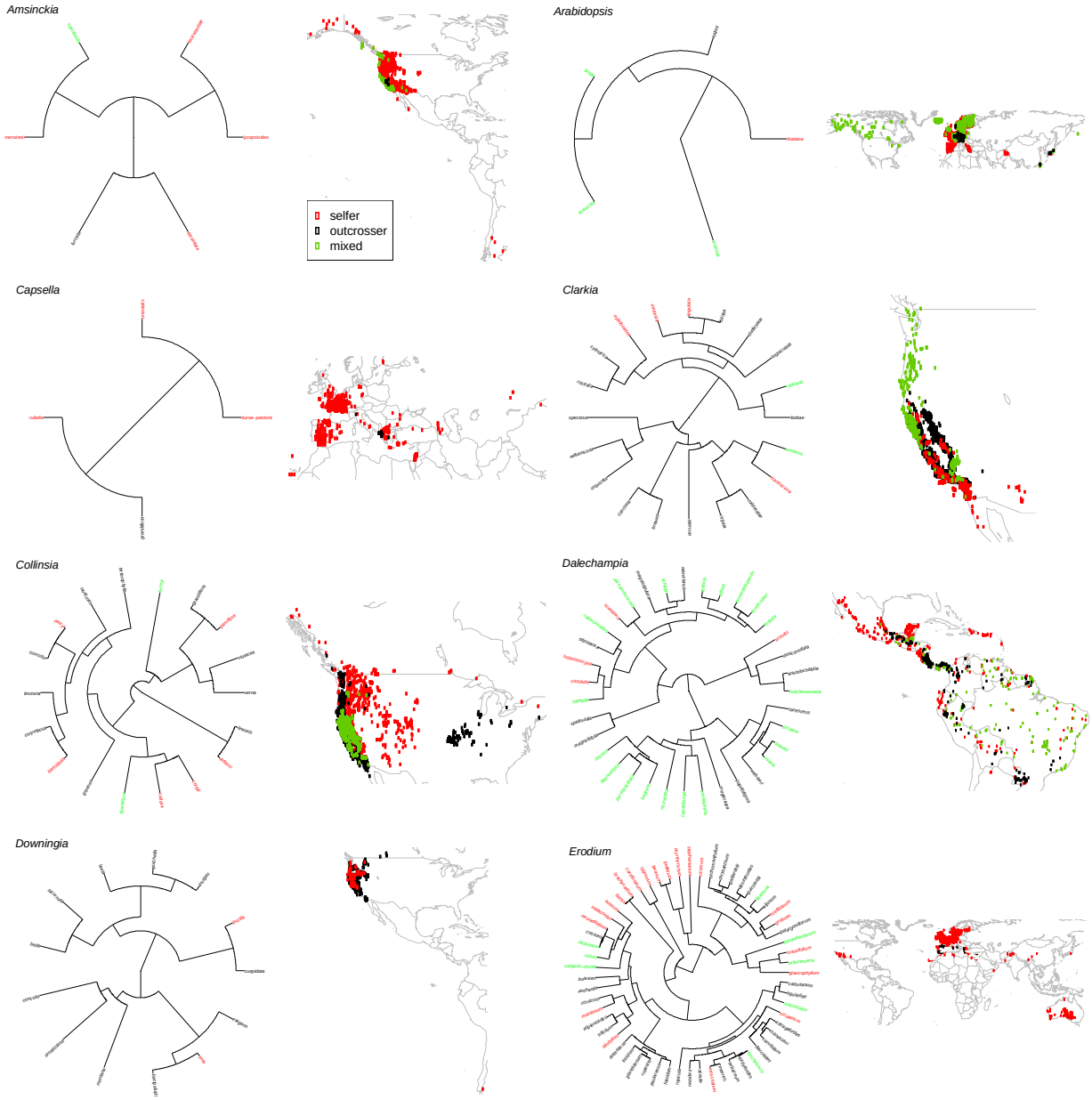
bioRxiv preprint doi: https://doi.org/10.1101/016261; this version posted March 9, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table 2. Results of beta regression models analyzing the effect of divergence time on co-occurrence, estimated at four spatial scales (**A-D**).

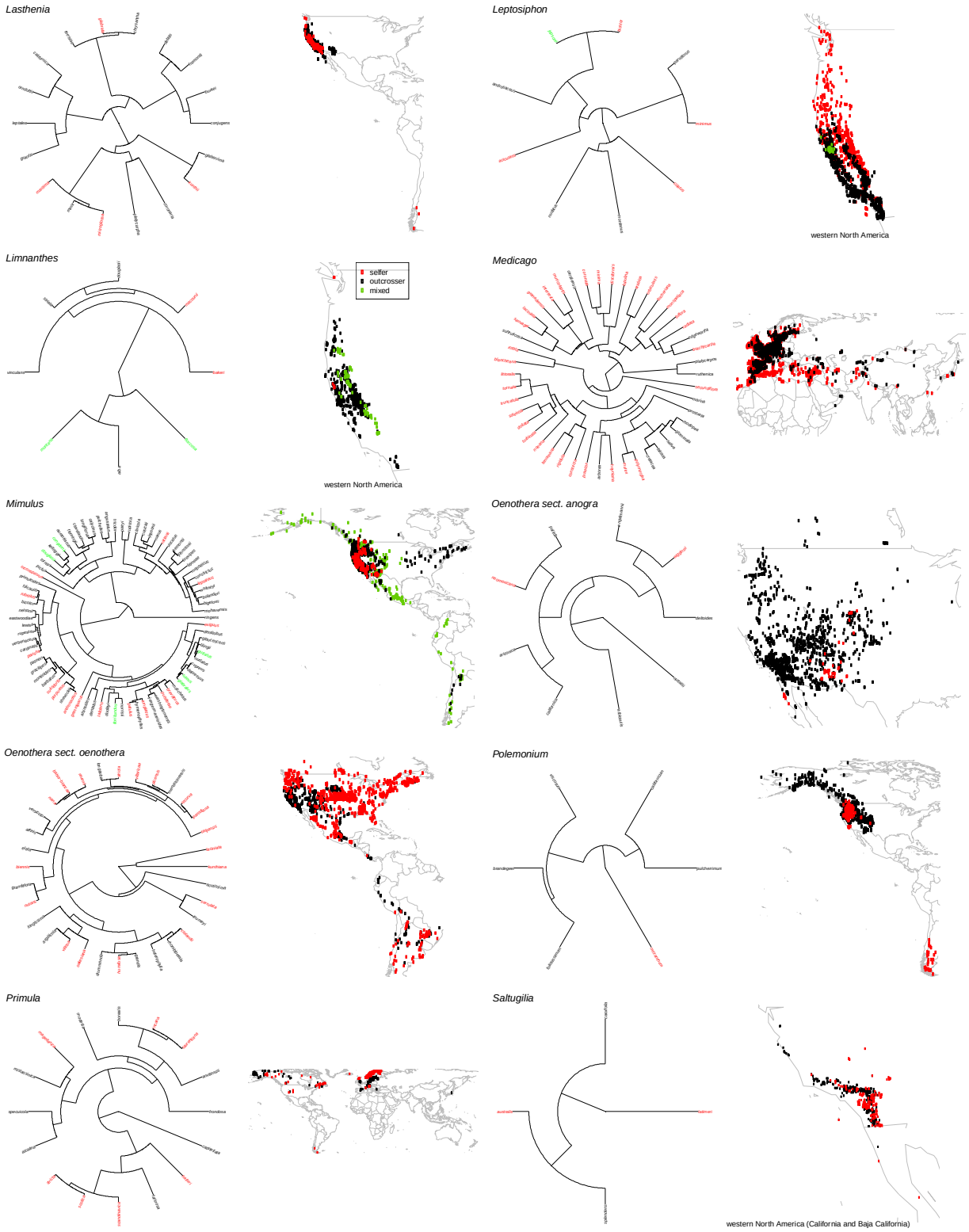
	Coefficient (s.e.)	*Wald's z-value*	*P*
A. 0.05 decimal degree (~25 km^2)			
Intercept	-2.118 (0.218)	-9.734	<0.001
Log divergence time	-0.100 (0.117)	-0.853	0.393
B. 0.1 decimal degree (~100 km^2)			
Intercept	-1.650 (0.212)	-7.800	<0.001
Log divergence time	-0.181 (0.121)	-1.493	0.135
C. 0.5 decimal degree (~2500 km^2)			
Intercept	-0.166 (0.222)	-0.748	0.455
Log divergence time	-0.257 (0.152)	-1.694	0.090
D. 1 decimal degree (~10,000 km^2)			
Intercept	-0.263 (0.224)	1.173	0.241
Log divergence time	-0.287 (0.155)	-1.852	0.064

bioRxiv preprint doi: https://doi.org/10.1101/016261; this version posted March 9, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

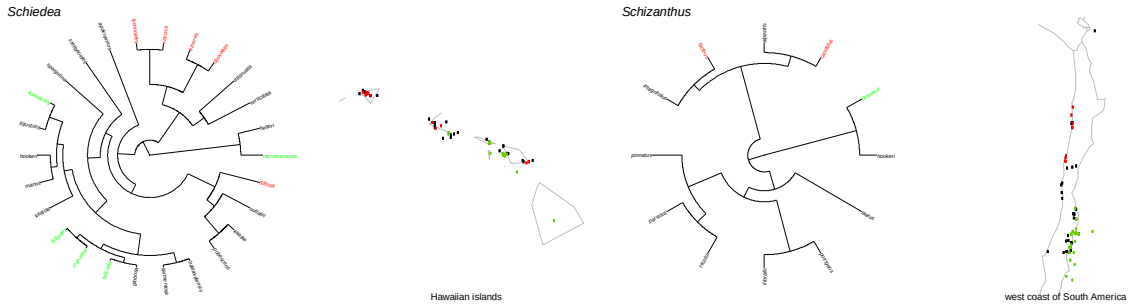
Appendix S1. Evolutionary relationships and native distributions of 20 clades. Trees represent bayesian consensus phylogenies with tips colored by mating system (red selfers, black outcrossers, green mixed). Geographic distributions represent species' occurrences, obtained from the global biodiversity information facility (www.gbif.org).



bioRxiv preprint doi: https://doi.org/10.1101/016261; this version posted March 9, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



bioRxiv preprint doi: https://doi.org/10.1101/016261; this version posted March 9, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



bioRxiv preprint doi: https://doi.org/10.1101/016261; this version posted March 9, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Appendix S2. Phylogenetic information for 20 clades included in our study.

Clade total species (proportion in our phylogeny)	**Previously published phylogeny [number of loci]**	**Predominant life history strategy: herbaceous(H) woody(W)**	**Subst. model**	**Chain length X 10^6**	**min. ESS**	**posterior ESS**
Amsinckia 14(0.57)	Schoen et al., 1997 [0nuc,1cp]	H	GTR +gamma	200	3207	6259
Arabidopsis 10(0.6)	Beck et al., 2007 [1nuc,0cp]	H	GTR + gamma	20	790	1367
Capsella 5 (1.0)	Hurka et al., 2012 [1nuc,4cp]	H	GTR + gamma	100	1066	1128
Clarkia ±41(0.51)	Kay et al., in prep [2nuc,0cp]	H	GTR + gamma	20	856	3257
Collinsia ±20(0.95)	Baldwin et al., 2011 [2nuc,1cp]	H	GTR + gamma	20	766	3929
Dalechampia 120 (0.5)	Armbruster et al., 2009 [2nuc,2cp]	W	GTR + gamma	200	4179	7435
Downingia 15 (0.87)	Schultheis, 2001 [1nuc,1cp]	H	GTR + gamma	20	884	2709
Erodium ±74(0.80)	Fiz-Palacios et al., 2010 [1nuc,1cp]	H	GTR + gamma	20	734	1713
Lasthenia 18 (1.0)	Chan et al., 2001 [2nuc,1cp]	H	GTR + gamma	20	697	2680
Limnanthes 8 (1.0)	Meyers et al., 2010 [1nuc,2cp]	H	HKY + gamma	100	2588	6690
Leptosiphon 30 (0.83)	Goodwillie, 1999 [1nuc,0cp]	H	GTR + gamma	20	602	3610
Medicago 83 (0.70)	Maureira-Butler et al., 2008 [2nuc, 1mito]	H	GTR + gamma	20	880	3148
Mimulus ±120(0.94)	Beardsley et al., 2004 [2nuc,1cp]	H	GTR + gamma	152	1974	4892
Oenothera sect. Anogra 10 (0.9)	Theiss et al., 2010 [1nuc,5cp]	H	GTR + gamma	200	602	2243
*Oenothera sect. Oenothera and Calylophus** 65 (0.51)	Johnson et al., 2009 [2nuc,3cp]	H	GTR + gamma	200	602	2243
Polemonium 30 (0.63)	Worley et al., 2009 [aflp]	H	GTR + gamma	20	834	3343
Primula sect. Aleuritia 21 (0.81)	Guggisberg et al., 2006 [0,4]	H	GTR + gamma	20	882	2949
Saltugilia	Johnson et al., 2007	H	HKY +	200	426	6655

bioRxiv preprint doi: https://doi.org/10.1101/016261; this version posted March 9, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

4 (1.0)	[1nuc,2cp]		gamma		4	
Schiedea 34 (0.76)	Soltis et al., 1996 [1nuc,1cp]	W	HKY + gamma	20	740	5438
Schizanthus 12 (1.0)	Perez and Arroyo, 2006 [2nuc,1cp]	H	GTR + gamma	20	674	2223

ARMBRUSTER, W.S., J. LEE, AND B.G. BALDWIN. 2009. Macroevolutionary patterns of defense and pollination in Dalechampia vines: adaptation, exaptation, and evolutionary novelty. *Proceedings of the National Academy of Sciences* 106: 18085–90.

BALDWIN, B.G., S. KALISZ, AND W.S. ARMBRUSTER. 2011. Phylogenetic perspectives on diversification, biogeography, and floral evolution of Collinsia and Tonella (Plantaginaceae). *American journal of botany* 98: 731–53.

BEARDSLEY, P. 2004. Patterns of evolution in western North American Mimulus (Phrymaceae). *American Journal of Botany* 91: 474–489.

BECK, J.B., I. A AL-SHEHBAZ, S.L. O'KANE, AND B. A SCHAAL. 2007. Further insights into the phylogeny of Arabidopsis (Brassicaceae) from nuclear Atmyb2 flanking sequence. *Molecular phylogenetics and evolution* 42: 122–30.

CHAN, R., B. BALDWIN, AND R. ORNDUFF. 2001. Goldfields revisited□: A molecular phylogenetic perspective on the evolution of Lasthenia (Compositae□: Heliantheae sensu lato). *International Journal of Plant Sciences* 162: 1347–1360.

FIZ-PALACIOS, O., P. VARGAS, R. VILA, A.S.T. PAPADOPULOS, AND J.J. ALDASORO. 2010. The uneven phylogeny and biogeography of Erodium (Geraniaceae): radiations in the Mediterranean and recent recurrent intercontinental colonization. *Annals of botany* 106: 871–84.

GOODWILLIE, C. 1999. Multiple origins of self-compatibility in Linanthus section leptosiphon (Polemoniaceae): Phylogenetic evidence from Internal-Transcribed-Spacer sequence data. *Evolution* 53: 1387–1395.

GUGGISBERG, A., G. MANSION, S. KELSO, AND E. CONTI. 2006. Evolution of biogeographic patterns, ploidy levels, and breeding systems in a diploid-polyploid species complex of Primula. *The New phytologist* 171: 617–32.

HURKA, H., N. FRIESEN, D. A GERMAN, A. FRANZKE, AND B. NEUFFER. 2012. "Missing link" species Capsella orientalis and Capsella thracica elucidate evolution of model plant genus Capsella (Brassicaceae). *Molecular ecology* 21: 1223–38.

JOHNSON, L. 2007. Transfer of the western north American species Gilia splendens to Saltugilia (Polemoniaceae), and the taxonomic affinities of Gilia scopulorum, Gilia stellata, and Gilia. *Novon: A Journal for Botanical Nomenclature* 17: 193–197.

JOHNSON, M.T.J., S.D. SMITH, AND M.D. RAUSHER. 2009. Plant sex and the evolution of plant defenses against herbivores. *Proceedings of the National Academy of Sciences* 106: 18079–84.

MAUREIRA-BUTLER, I.J., B.E. PFEIL, A. MUANGPROM, T.C. OSBORN, AND J.J. DOYLE. 2008. The reticulate history of Medicago (Fabaceae). *Systematic biology* 57: 466–82.

bioRxiv preprint doi: https://doi.org/10.1101/016261; this version posted March 9, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

MEYERS, S., A. LISTON, AND R. MEINKE. 2010. A molecular phylogeny of Limnanthes (Limnanthaceae) and investigation of an anomalous Limnanthes population from California , U . S . A . *Systematic botany* 35: 552–558.

PÉREZ, F., AND M. ARROYO. 2006. Ancestral reconstruction of flower morphology and pollination systems in Schizanthus (Solanaceae). *American Journal of Botany* 93: 1029–1038.

SCHOEN, D., M. JOHNSTON, A. L'HEUREUX, AND J. MARSOLAIS. 1997. Evolutionary history of the mating system in Amsinckia (Boraginaceae). *Evolution* 51: 1090–1099.

SCHULTHEIS, L. 2001. Systematics of Downingia (Campanulaceae) based on molecular sequence data: implications for floral and chromosome evolution. *Systematic Botany* 26: 603–621.

SOLTIS, P., D. SOLTIS, S. WELLER, A. SAKAI, AND W. WAGNER. 1996. Molecular phylogenetic analysis of the Hawaiian endemics Schiedea and Alsinidendron (Caryophyllaceae). *Systematic Botany* 21: 365–379.

THEISS, K.E., K.E. HOLSINGER, AND M.E.K. EVANS. 2010. Breeding system variation in 10 evening primroses (Oenothera sections Anogra and Kleinia; Onagraceae). *American journal of botany* 97: 1031–9.

WORLEY, A., H. GHAZVINI, AND D. SCHEMSKE. 2009. A phylogeny of the genus Polemonium based on amplified fragment length polymorphism (AFLP) markers. *Systematic botany* 34: 149–161.

bioRxiv preprint doi: https://doi.org/10.1101/016261; this version posted March 9, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Appendix S3. Description of how mating system was determined for each clade. The proportion of species assigned as outcrossers, mixed maters, and selfers are included in brackets, [outcrosser/mixed/selfer].

Amsinckia	Outcrossing rates were estimated with molecular markers and shown to be correlated with style type for a subset of species (e.g. Schoen et al., 1997). Distylous species are predominantly outcrossing or mixed mating; homostylous species are predominantly selfing. Style type was thus used to estimate mating system for additional species in this clade. [0.64/0.00/0.36]
Arabidopsis	Outcrossing and self-incompatibility rates were estimated using molecular data and experimental hand-pollinations (e.g., Clauss and Koch, 2006). [0.20/0.40/0.40]
Capsella	Outcrossing and self-incompatibility rates were estimated using molecular data and experimental hand pollinations (e.g., Brandvain et al., 2013, Hurka et al., 2012). [0.33/0.00/0.66]
Clarkia	Automatic selfing rates and outcrossing rates (estimated with molecular data for a few species) were correlated with herkogamy (e.g., Lewis and Lewis, 1955). Herkogamy was used to estimate mating system for the remainder of species. [0.48/0.08/0.45]
Collinsia	Outcrossing rates were estimated with molecular markers for all species and shown to be correlated with timing of stigma receptivity and flower size (e.g., Kalisz et al., 2012). [0.58/0.05/0.37]
Dalechampia	Automatic seed set for bagged flowers was shown to correlate with herkogamy for a subset of species (e.g., Armbruster, 1988 and 1993). Herkogamy was used to estimate mating system for the remainder (Armbruster, 1993). [0.33/0.51/0.16]
Downingia	Automatic seed set in the greenhouse was correlated with stigma exertion and flower size in a subset of species (Schultheis, 2001). Stigma exertion was used to estimate mating system for the remainder. [0.85/0.00/0.15]
Erodium	Automatic fruit set (% fruit production on bagged plants; unbagged plants were used as a control) was highly correlated with pollen ovule ratios for a subset of species (Alarcon et al., 2011). Pollen ovule ratios were used to estimate mating system for the remaining species. [0.53/0.12/0.35]
Lasthenia	Self-incompatibility rates were determined for all species presumably using experimental hand pollinations (e.g., Orundorf, 1966). For those species that were self compatible, rates of automatic selfing were high (presumably when flowers were bagged) on a subset of species (Orundorf, 1966; Chan et al., 2001). [0.79/0.00/0.21]
Limnanthes	Protandry and gynodioecy was correlated with high outcrossing rates, while cleistogamy was correlated with high selfing rates in a subset of species (e.g., McNeil and Jain, 1983). A range of field pollination studies and molecular studies have been

bioRxiv preprint doi: https://doi.org/10.1101/016261; this version posted March 9, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

	used to confirm this across the majority of taxa. [0.44/0.33/0.22]
Leptosiphon	Self-incompatibility was determined using experimental hand pollinations for all species (e.g., Goodwillie, 1999). For a subset of the self-compatible species, outcrossing rates were estimated with molecular markers and were either low (consistent with predominant selfing) or intermediate (consistent with mixed mating) (e.g., Goodwillie, 2001). [0.50/0.00/0.50]
Medicago	Degree of automatic fruit set in the greenhouse was used to categorize species as "selfers" or "outcrossers" (Maureira-Butler et al., 2008) and was found to be consistent with previously published reports for a subset of the species. [0.72/0.00/0.28]
Mimulus	Outcrossing rates were estimated with molecular markers and were correlated with the degree of herkogamy for a subset of species (e.g. Ritland and Ritland, 1989). The degree of herkogamy was then used to estimate mating system across other species (e.g. Grossenbacher and Whittall, 2011). [0.71/0.07/0.23]
Oenothera sect. Anogra	Self-incompatibility rates were estimated using experimental hand pollinations for all species (e.g. Theiss et al., 2010). For self-compatible species, herkogamy was used to estimate whether species were predominantly selfing or outcrossing. [0.90/0.00/0.10]
*Oenothera sect. Oenothera and Calylophus**	Species were classified as either sexual, or functionally asexual due to a permanent translocation whereby plants self-fertilize but do not undergo segregation and recombination (Johnson et al., 2009). For species defined as sexual, experimental crosses showed that about half were self-incompatible and thus outcrossing. The remaining half displayed partial self-incompatibility and may be mixed mating. For the purposes of the present study, sexual species are assumed to be outcrossing relative to asexual species. [0.42/0.00/0.58]
Polemonium	Self-incompatibility rates were determined using experimental hand pollinations on a subset of species (Worley pers. com.). Species were classified as outcrossing when the combined fruit x seed set of selfed flowers was <25% of that in outcrossed flowers. For species that were self-compatible, automatic selfing rates were assessed in the field or greenhouse (e.g., Worley pers. com., Hill et al., 2008). [0.85/0.00/0.15]
Primula sect. Aleuritia	Self-incompatibility was correlated with style type for a subset of species (discussed in Guggisberg et al., 2006). Distylous species were found to be self-incompatible, homostylous species self-compatible and autogamous (e.g., Tremayne and Richards, 2000). [0.56/0.00/0.44]
Saltugilia	Self-incompatibility and autogamy rates were assessed for all 4 species (Grant and Grant, 1965). [0.50/0.00/0.50]
Schiedea	Species were determined to be dioecious, subdioecious, gyodioecious, or hermaphroditic (Weller et al., 1995). Dioecious and subdioecious species were reported as mostly outcrossing. For gynodioecious and hermaphroditic species, outcrossing and automatic selfing rates (measured using molecular data and presumably bagging experiments) ranged from mostly outcrossing to mostly selfing.

bioRxiv preprint doi: https://doi.org/10.1101/016261; this version posted March 9, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

	[0.66/0.17/0.17]
Schizanthus	Automatic selfing rates in the field were correlated with pollen dehiscence and other reproductive characters for the majority of species (e.g., Perez et al., 2009). These correlated characters were used to estimate mating system in the remaining 3 species. [0.75/0.08/0.17]

ALARCÓN, M.L., C. ROQUET, AND J.J. ALDASORO. 2011. Evolution of pollen/ovule ratios and breeding system in Erodium (Geraniaceae). *Systematic Botany* 36: 661–676.

ARMBRUSTER, W. 1993. Evolution of plant pollination systems□: Hypotheses and tests with the neotropical vine Dalechampia. *Evolution* 47: 1480–1505.

ARMBRUSTER, W. 1988. Multilevel comparative analysis of the morphology, function, and evolution of Dalechampia blossoms. *Ecology* 69: 1746–1761.

BRANDVAIN, Y., T. SLOTTE, K.M. HAZZOURI, S.I. WRIGHT, AND G. COOP. 2013. Genomic identification of founding haplotypes reveals the history of the selfing species Capsella rubella. *PLoS genetics* 9: .

CHAN, R., B. BALDWIN, AND R. ORNDUFF. 2001. Goldfields revisited□: A molecular phylogenetic perspective on the evolution of Lasthenia (Compositae□: Heliantheae sensu lato). *International Journal of Plant Sciences* 162: 1347–1360.

CLAUSS, M.J., AND M. A KOCH. 2006. Poorly known relatives of Arabidopsis thaliana. *Trends in plant science* 11: 449–59.

GOODWILLIE, C. 1999. Multiple origins of self-compatibility in Linanthus section leptosiphon (Polemoniaceae): Phylogenetic evidence from Internal-Transcribed-Spacer sequence data. *Evolution* 53: 1387–1395.

GOODWILLIE, C. 2001. Pollen limitation and the evolution of self□compatibility in Linanthus (Polemoniaceae). *International Journal of Plant Sciences* 162: 1283–1292.

GRANT, V., AND K.A. GRANT. 1965. Flower pollination in the phlox family. Columbia University Press, New York.

GROSSENBACHER, D.L., AND J.B. WHITTALL. 2011. Increased floral divergence in sympatric monkeyflowers. *Evolution* 65: 2712–2718.

GUGGISBERG, A., G. MANSION, S. KELSO, AND E. CONTI. 2006. Evolution of biogeographic patterns, ploidy levels, and breeding systems in a diploid-polyploid species complex of Primula. *The New phytologist* 171: 617–32.

HILL, L.M., A.K. BRODY, AND C.L. TEDESCO. 2008. Mating strategies and pollen limitation in a globally threatened perennial Polemonium vanbruntiae. *Acta Oecologica* 33: 314–323.

HURKA, H., N. FRIESEN, D. A GERMAN, A. FRANZKE, AND B. NEUFFER. 2012. “Missing link” species Capsella orientalis and Capsella thracica elucidate evolution of model plant genus Capsella (Brassicaceae). *Molecular ecology* 21: 1223–38.

bioRxiv preprint doi: https://doi.org/10.1101/016261; this version posted March 9, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

JOHNSON, M.T.J., S.D. SMITH, AND M.D. RAUSHER. 2009. Plant sex and the evolution of plant defenses against herbivores. *Proceedings of the National Academy of Sciences* 106: 18079–84.

KALISZ, S., A. RANDLE, D. CHAIFFETZ, M. FAIGELES, A. BUTERA, AND C. BEIGHT. 2012. Dichogamy correlates with outcrossing rate and defines the selfing syndrome in the mixed-mating genus Collinsia. *Annals of botany* 109: 571–82.

LEWIS, H., AND M.E. LEWIS. 1955. The genus Clarkia. University of California Press Berkeley, California, USA.

MAUREIRA-BUTLER, I.J., B.E. PFEIL, A. MUANGPROM, T.C. OSBORN, AND J.J. DOYLE. 2008. The reticulate history of Medicago (Fabaceae). *Systematic biology* 57: 466–82.

MCNEILL, C., AND S. JAIN. 1983. Genetic differentiation studies and phylogenetic inference in the plant genus Limnanthes (section Inflexae). *Theoretical and applied genetics*257–269.

ORNDUFF, R. 1966. A biosystematic survey of the goldfield genus Lasthenia (Compositae: Helenieae). University of California Press Berkeley, California, USA.

PÉREZ, F., M.T.K. ARROYO, AND J.J. ARMESTO. 2009. Evolution of autonomous selfing accompanies increased specialization in the pollination system of Schizanthus (Solanaceae). *American journal of botany* 96: 1168–76.

RITLAND, C., AND K. RITLAND. 1989. Variation of sex allocation among 8 taxa of the Mimulus guttatus species complex (Scrophulariaceae). *American Journal of Botany* 76: 1731–1739.

SCHOEN, D., M. JOHNSTON, A. L'HEUREUX, AND J. MARSOLAIS. 1997. Evolutionary history of the mating system in Amsinckia (Boraginaceae). *Evolution* 51: 1090–1099.

SCHULTHEIS, L. 2001. Systematics of Downingia (Campanulaceae) based on molecular sequence data: implications for floral and chromosome evolution. *Systematic Botany* 26: 603–621.

THEISS, K.E., K.E. HOLSINGER, AND M.E.K. EVANS. 2010. Breeding system variation in 10 evening primroses (Oenothera sections Anogra and Kleinia; Onagraceae). *American journal of botany* 97: 1031–9.

TREMAYNE, M., AND A. RICHARDS. 2000. Seed weight and seed number affect subsequent fitness in outcrossing and selfing Primula species. *New Phytologist* 34: 149–161.

WELLER, S., W. WAGNER, AND A. SAKAI. 1995. A phylogenetic analysis of Schiedea and Alsinidendron (Caryophyllaceae□: Alsinoideae): Implications for the evolution of breeding systems. *Systematic Botany* 20: 315–337.

bioRxiv preprint doi: https://doi.org/10.1101/016261; this version posted March 9, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Appendix S4. Results of beta regression models analyzing the effects of divergence time, 'sister pair mating system', and their interaction on co-occurrence, estimated at four spatial scales (**A-D**).

	Coefficient (s.e.)	*Wald's z-value*	*P*
A. 0.05 decimal degree (~25 km^2)			
Intercept	-2.116 (0.227)	-9.316	<0.001
selfer-outcrosser	-0.049 (0.391)	-0.125	0.900
selfer-selfer	0.345 (0.482)	0.716	0.474
divergence time	-0.118 (0.179)	-0.662	0.508
selfer-outcrosser : divergence time	-0.042 (0.272)	-0.153	0.878
selfer-selfer : divergence time	0.297 (0.347)	0.855	0.392
B. 0.1 decimal degree (~100 km^2)			
Intercept	-1.820 (0.335)	-5.431	<0.001
selfer-outcrosser	0.159 (0.404)	0.393	0.695
selfer-selfer	0.266 (0.502)	0.530	0.596
divergence time	-0.155 (0.185)	-0.836	0.403
selfer-outcrosser : divergence time	-0.147 (0.280)	-0.524	0.600
selfer-selfer : divergence time	0.326 (0.364)	0.896	0.370
C. 0.5 decimal degree (~2500 km^2)			
Intercept	-0.475 (0.388)	-1.225	0.221
selfer-outcrosser	0.542 (0.508)	1.067	0.286
selfer-selfer	0.154 (0.632)	0.244	0.807
divergence time	-0.100 (0.228)	-0.437	0.662
selfer-outcrosser : divergence time	-0.360 (0.344)	-1.047	0.295
selfer-selfer : divergence time	0.014 (0.457)	0.031	0.975
D. 1 decimal degree (~10,000 km^2)			
Intercept	0.234 (0.394)	0.594	0.553
selfer-outcrosser	0.049 (0.514)	0.096	0.924
selfer-selfer	-0.010 (0.643)	-0.015	0.988
divergence time	-0.252 (0.236)	-1.069	0.285

bioRxiv preprint doi: https://doi.org/10.1101/016261; this version posted March 9, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

selfer-outcrosser : divergence time	-0.119 (0.352)	-0.337	0.736
selfer-selfer : divergence time	-0.002 (0.466)	-0.004	0.997

Note: The categorical coefficient estimates are log-odds ratios, measured as departures from the 'outcrosser-outcrosser' category.