Received: 12 March 2021 | Accepted: 27 August 2021

DOI: 10.1002/ajb2.1766

RESEARCH ARTICLE

American Journal of Botany

Selfing rate variation within species is unrelated to life-history traits or geographic range position

Carly J. Prior | **Jeremiah W. Busch**

School of Biological Sciences, Washington State University, Pullman, WA 99164, USA

Correspondence
Carly J. Prior, School of Biological Sciences, Washington State University, Pullman, WA 99164, USA.
Email: carly.prior@wsu.edu

Abstract

Premise: In plants, populations and species vary widely along the continuum from outcrossing to selfing. Life-history traits and ecological circumstances influence among-species variation in selfing rates, but their general role in explaining intraspecific variation is unknown. Using a database of plant species, we test whether life-history traits, geographic range position, or abundance predict selfing rate variation among populations.

Methods: We identified species where selfing rates were estimated in at least three populations at known locations. Two key life-history traits (generation time and growth form) were used to predict within-species selfing rate variation. Populations sampled within a species' native range were assessed for proximity to the nearest edge and abundance. Finally, we conducted linear and segmented regressions to determine functional relationships between selfing rate and geographic range position within species.

Results: Selfing rates for woody species varied less than for herbs, which is explained by the lower average selfing rate of woody species. Relationships between selfing and peripherality or abundance significantly varied among species in their direction and magnitude. However, there was no general pattern of increased selfing toward range edges. A power analysis shows that tests of this hypothesis require studying many (i.e., 40+) populations.

Conclusions: Intraspecific variation in plant mating systems is often substantial yet remains difficult to explain. Beyond sampling more populations, future tests of biogeographic hypotheses will benefit from phylogeographic information concerning specific range edges, the study of traits influencing mating system (e.g., herkogamy), and measures of abundance at local scales (e.g., population density).

KEYWORDS

abundant center hypothesis, Baker's rule, biogeography, geographic range, life-history traits, mating system, self-fertilization

Plant mating systems are highly variable, with species ranging from primarily self-fertilizing (selfing) to obligately outcrossing, with many exhibiting some mixture of selfing and outcrossing (Schemske and Lande, 1985; Vogler and Kalisz, 2001). Biologists quantify mating systems by tabulating the number of selfed offspring (i.e., the selfing rate; Ritland, 1990, 2002). Mating systems profoundly affect the genetic structure and persistence of populations (Schoen and Brown, 1991; Hamrick and Godt, 1996; Siol et al., 2007; Crawford et al., 2008; Barrett and Harder, 2017; Noël et al., 2017), in addition to their niche breadth (Randle et al., 2009; Grossenbacher et al., 2015; Grant and Kalisz, 2020; but see Lowry and Lester, 2006). Because of these impacts, there has been keen interest in explaining variation in the amount of self-fertilization expressed by plants (Holsinger, 2001). In studying variation among species, selfing rates are not randomly associated with life-history traits (Barrett and Eckert, 1990; Barrett et al., 1996; Hamrick and Godt, 1996). In addition, more highly selfing species tend to be associated with particular geographic contexts, since they tend to colonize peripheral, isolated, or island environments (Baker, 1967; Grossenbacher et al., 2017). Additionally, selfing species have been shown to have greater niche breadth (Grant and Kalisz, 2020) and larger geographic ranges than outcrossing sister taxa (Randle et al., 2009; Grossenbacher et al., 2015; but see Lowry and Lester, 2006). While among-species variation in mating systems has been well studied, variability in the amount of selfing within species is not well understood (Jain, 1976; Whitehead et al., 2018).

Life-history traits connect variation in selfing rate among and within species

The primary cost of selfing is the reduction in fitness of self-fertilized progeny (i.e., inbreeding depression; Lande and Schemske, 1985; Charlesworth and Willis, 2009), while the primary benefit of selfing is the production of seed when cross-pollen is limited (i.e., reproductive assurance; Jain, 1976; Solbrig, 1980; Schoen et al., 1996; Eckert et al., 2006; Busch and Delph, 2012). There is a well-supported association between mean selfing rate and life-history traits, with woody and perennial species having lower selfing rates (Barrett and Eckert, 1990; Barrett et al., 1996; Hamrick and Godt, 1996; Moeller et al. 2017). Given that a balance between the costs and benefits of selfing influences the selfing rate at the population level (Herlihy and Eckert, 2002; Layman et al., 2017), the life-history traits of species should set bounds on the amount of selfing rate variation that is expressed among populations of a species. For one, woody species accumulate more somatic mutations than herbs because they undergo more mitotic cell divisions during individual growth (Klekowski and Godfrey, 1989; Scofield and Schultz, 2006; Duminil et al., 2009), making the evolution of higher selfing rates increasingly difficult (Morgan et al., 1997; Morgan, 2001). In a similar vein, perennial species experience diminished selection for reproductive assurance because the life-time fitness of individuals reflects several reproductive episodes, constraining the evolution of high selfing rates (Morgan et al., 1997; Pannell and Barrett, 1998). In other words, populations with high selfing rates should rarely be observed because of evolutionary constraints in species where the costs of self-fertilization are generally high. Such an argument based on evolutionary constraints predicts that woody species and perennial species will, on average, express less within-species variation in selfing rates than otherwise identical herbaceous or annual species.

The geography of selfing rate variation within species

Selfing rate variation within species is also expected to reflect local environmental conditions and therefore have a geographic component. Neither inbreeding depression nor pollen limitation are expected to be constant when populations inhabit heterogenous environments across a species' range (Baker, 1955, 1965; Fishman and Wyatt, 1999; Moeller and Geber, 2005; Haddadchi and Fatemi, 2015; Moeller et al., 2017). For one, small or recently colonized populations are expected to have a combination of lower inbreeding depression (Charlesworth et al., 1992; Kirkpatrick and Jarne, 2000; Pujol et al., 2009), a dearth of mates (Jacquemyn et al., 2002; Knight, 2003), and potentially fewer pollinators (Mustajärvi et al., 2001; Knight et al., 2005; Morgan et al., 2005). Moreover, a finding of populations with high selfing rates in newly colonized locations or environments with few mates or pollinators does not mean plant traits have evolved to secure self-pollination. Selfing rates can also naturally increase whenever outcross pollen is hard to find (Baker, 1955; Karron et al., 1995; Kalisz et al., 2004; Ghazoul, 2005; Coates et al., 2007; Eckert et al., 2009; Yin et al., 2016). Range-edge environments may be associated with more selfing (Moeller, 2006; Barrett, 2010; Moeller et al., 2012; Hargreaves and Eckert, 2014), though such an expectation depends on whether abundance generally declines near range boundaries (Brown, 1984; Sagarin and Gaines, 2002; Eckert et al., 2008; Dallas et al., 2017; Pironon et al., 2017). Of studies in single species, several support the hypothesis of increased selfing on species' range edges (Rick et al., 1977; Moeller and Geber, 2005; Darling et al., 2008; Dart et al., 2012; Griffin and Willi, 2014; but see Herlihy and Eckert, 2005). However, many of these studies infer mating systems indirectly (e.g., petal size) or were based on a priori knowledge of geographic variation in a species' key floral traits (Hargreaves and Eckert, 2014). Whether increased selfing rates on species' range edges is a common biogeographic pattern is currently unknown.

In addition, the expected shape of relationships between selfing rates and proximity to species' range edges is unclear. Gradual declines in selfing rates toward the range edge may be expected in species where pollinators or mates gradually decline in abundance in environments that are increasingly close to range boundaries. Conversely, abrupt shifts in mating system may occur very near range edges if environmental quality suddenly deteriorates (Santini et al., 2019). While linear, monotonic relationships represent a reasonable starting place when investigating biogeographic patterns (Gaston and Blackburn, 2000; Freckleton et al., 2003; Meiri et al., 2007; Pironon et al., 2017), more complex relationships may emerge from empirical studies (Yancovitch Shalom et al., 2020). To better understand the relationships between the geography of selfing rate variation and its underlying ecological predictors, analyses would benefit from confronting data with additional models that incorporate the potential for nonlinear relationships across a species' range (Hilborn and Mangel, 1997). Specifically, abundance surfaces are likely to fluctuate across a species' range, potentially complicating simple relationships between selfing rates and a population's range position or abundance (Yancovitch Shalom et al., 2020). Accommodating for multiple relationships between selfing and geographic range position provides an opportunity to determine where such alterations occur and whether they are more likely near species' range edges.

Since the development of methods to estimate selfing rates (Ritland and Jain, 1981; Ritland, 2002), there has been an impressive number of studies on selfing rate variation among populations and its underlying ecological causes (Schoen, 1982; Barrett and Husband, 1990; Dart et al., 2012; Whitehead et al., 2018). If there are any general predictors of selfing rate variation within species, revealing them would require studying many species across the tree of life of seed plants. These types of studies could assess whether selfing rates are predicted by geographic range position or local abundance. Here, we assembled a diverse, multispecies data set, quantified the relative proximity of populations to range edges, and interpolated abundance patterns. These data on geographic range position serve as predictors of variation among populations in the proportion of progeny produced by self-fertilization (i.e., the selfing rate). With this multispecies data set, we asked the following: (1) Is the degree of variation in selfing rate within species predicted by life-history traits? (2) Is there a relationship between selfing rates of populations and either their geographic location or abundance? (3) What is the general shape of the relationship between selfing rate and geographic location?

MATERIALS AND METHODS

Assembling a data set of within-species variation in selfing rate

To assess whether selfing is more common near species' range edges, a database of studies that measured population-level selfing rates and provided locality information was assembled. Most studies reported outcrossing rates (t_m), which were transformed into selfing rates ($1 - t_m$). Only multilocus outcrossing rates were used to acquire estimates of the realized mating system within populations. We initially searched published databases that gathered empirical estimates from studies of plant mating systems (Goodwillie et al., 2005; Moeller et al., 2017; Whitehead et al., 2018). To be included in our database, each study needed to report selfing rates for at least three populations (e.g., Whitehead et al., 2018) and provide locality information for each population. Initially, other databases that focus on mating system studies were searched for papers that fit these qualifications (Goodwillie et al., 2005; Moeller et al., 2017; Whitehead et al., 2018). To capture as many species as possible, an additional Web of Science search was conducted, using these criteria, for studies published between September 2016 and April 2019 that cite seminal papers in the calculation of outcrossing rates (e.g., Ritland and Jain, 1981; Ritland, 1990, 2002).

Cultivated or experimentally manipulated populations were excluded from this data set. Species with known hybridization (e.g., *Camissoniopsis chieranthifolia*; López-Villalobos and Eckert, 2019) were also excluded from this study as it complicates the interpretation of within-species selfing rate variation. When outcrossing rates were reported graphically, the R package digitize 0.0.4 (Poisot, 2011) was used to extract the values. If a study reported outcrossing rates for the same population over different years or if separate studies reported outcrossing rates for the same populations, an average was taken. If studies reported outcrossing rates for different seed or flower morphs within a population, an average was calculated for each population across categories. The platform GEOLocate (www.geo-locate.org) was used to obtain geographic coordinates for populations when only location names were provided. Additionally, if only a county name was provided as the populations' locality, the center of the county was used ($N = 41$). The assumption of county center for these populations should not introduce bias into the resulting data set as it is random in both spatial dimensions. From these searches, a database was developed containing population-level multilocus outcrossing rates and locality data for 97 vascular plants; hereafter, these data will be referred to as selfing rate data. All species names were checked with the Taxonomic Name Resolution Service (TNRS v4.1; Boyle et al., 2013).

Variation in life-history traits and intraspecific selfing rates

To determine whether life-history traits influence within-species selfing rate variation, generation time and growth form data were collected from the citations reporting selfing rates, a large biogeographic database (Moeller et al., 2017), and the TRY plant traits database (Kattge et al., 2020). Two species were dropped due to lack of information on their generation times and growth forms ($N = 97$). Generation time was collapsed into two categories: annual and perennial. Biennial species were categorized as annuals as they only undergo a single reproductive event. Growth form was also collapsed into two categories as well: herbaceous and woody (Scofield and Schultz, 2006).

Two analyses were performed to determine whether selfing rate variation within species is elevated in annual or herbaceous species as compared to perennial or woody species. In the first analysis, each of the 97 species' selfing rates, the mean, standard deviation (SD), and coefficient of variation (CV) were calculated. The CV was used because the SD and mean values were found to be correlated. Mann–Whitney *U* tests were used to compare each statistic between the groupings of generation time (annual vs. perennial) and growth form (herbaceous vs. woody).

Comparisons of selfing rate variation are complicated by the fact that species with intermediate selfing rates are expected to exhibit greater variation among populations. In the second analysis, binomial expectations were used to control for the increased variance of intermediate values of selfing rates, which are bounded on the 0–1 interval (Whitehead et al., 2018). The sample variance (σ^2) was corrected by the binomial variance expected given the species' mean selfing rate (p): $\sigma^2/p(1 - p)$. This average-corrected variance was calculated for each of the four life-history categories, and observed differences were determined for generation time and growth form. Two thousand bootstrapped replicates were sampled for each life-history category, and bootstrapped differences were calculated for generation time and growth form. The 95% confidence interval of the bootstrapped differences was compared to the observed differences to determine significance.

The geographic ranges of species in the data set

Each species' unique geographic range was estimated with occurrence data. Occurrence data were downloaded using the BIEN R package version 1.2.3 (Botanical Information and Ecology Network R package; Maitner et al., 2018). BIEN R searches were completed using TNRS-approved names. Subspecies names are not accepted within the search query and were therefore dropped. BIEN R occurrences were then annotated with information on native status and political boundaries for subsequent filtering. When an occurrence record did not provide location names, a database of global administrative areas (GADM, Hijmans et al., 2018; Weiss et al., 2018) was used to determine location names based on geographic coordinates to facilitate later native status filtering through a custom R script. These data will hereafter be referred to as occurrence data.

Data from BIEN R primarily originate from the Global Biodiversity Information Facility (GBIF), which may lead to records for cultivated or invasive instances. As our aim is to assess whether native plant ranges display more selfing toward range edges, non-native occurrence data were filtered out. Native ranges were ascertained via USDA-ARS Germplasm Resources Information Network (GRIN) and the Euro+Med Plantbase (Grossenbacher et al., 2017) using only those locations marked as "native". Species that were not present or did not have native range information in either GRIN or Euro+Med databases were removed (e.g., *Eichhornia paniculata*). Native range information from both sources was then translated into Biodiversity Information Standards (TDWG; http://www.twdg.org/standards/109) codes using the R package bazr (Walker, 2020). These codes provide a standard reference of whether each geographic location is contained within the set of native regions for each species (Grossenbacher et al., 2017). Native status for occurrence data was determined through comparison of native locations and occurrence records for each species (Figure 1A). Custom scripts were written to identify all native occurrences within the data set obtained from BIEN R. All subsequent analyses were performed on occurrence data that passed this filtration step.

Native occurrence data were then used to create a convex hull for each species range using the function convexhull.xy from the R package Spatstat (Baddeley et al., 2015). These convex hulls represent the shape and extent of the native range of each species (Figure 1B). Convex hulls were also used to determine whether populations in the selfing rate data set were within their respective native ranges. Any populations outside of the native convex hull plus a 25-km buffer zone were removed from subsequent analyses as were any species that dropped below

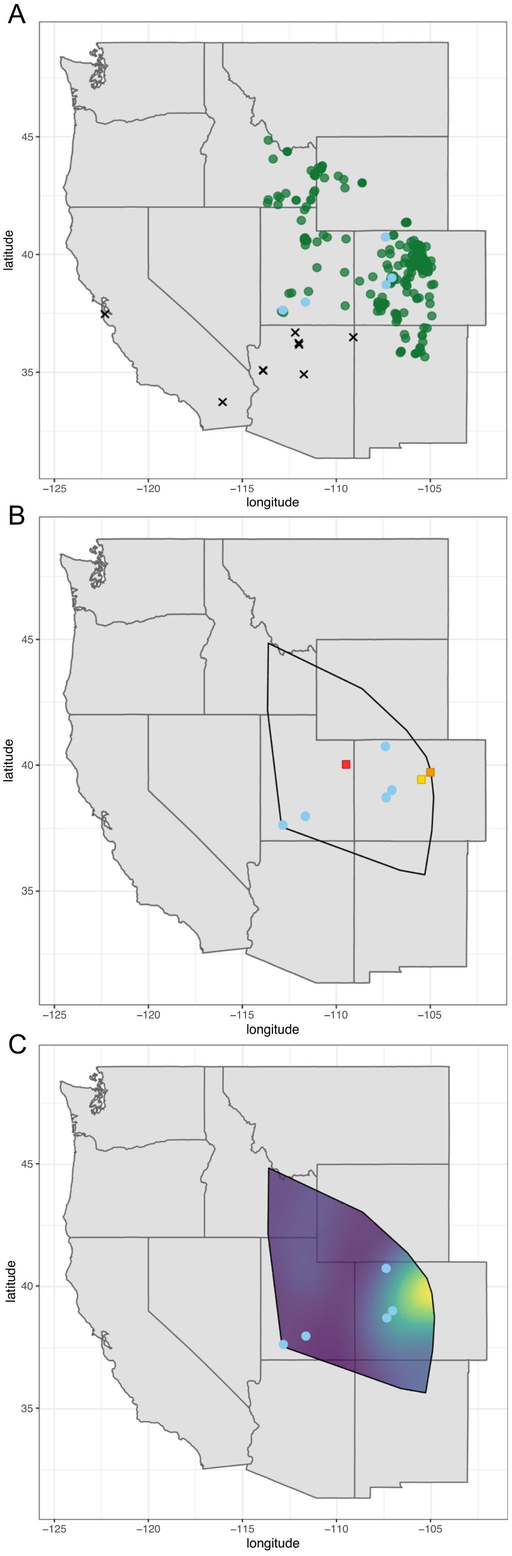

FIGURE 1 Maps depicting key steps for determining native ranges, range centers, and interpolation of abundances using *Aquilegia coerulea* as an example. In all maps, light blue points represent the populations with selfing rate data. (A) Green points indicate native occurrences of *A. coerulea* from BIEN R. Black x's indicate non-native occurrences obtained from BIEN R. (B) The black polygon represents the convex hull which is based on native occurrence data. Squares indicate the different range centers: red is the geometric center, orange is the predicted abundance center, and yellow is the friction abundance center. (C) Predicted abundance surface. Lightest colors indicate areas with the highest abundance

the initial three-population threshold. After these filtration steps, 56 species remained to be analyzed.

Predictors of selfing rate variation: peripherality and abundance

To characterize the geographic location of each population in the selfing rate data set, the center of each species geographic range was determined. Three range centers were evaluated as a means to investigate the effect of center location on the relationship of geographic location and selfing rates (Figure 1B). First, the geometric range center was defined using the centroid function from the R package Geosphere (Hijmans, 2019) for each species' convex hull. Second, densities were interpolated across the geographic range based on occurrence records, yielding a predicted abundance surface (Figure 1C). Abundance interpolation was performed with the density.ppp function from Spatstat (Baddeley et al., 2015). The predicted abundance range center was then defined as the point of highest abundance within the range. Finally, a friction abundance center was interpolated to reduce the potential sampling bias in the occurrence data set. Friction abundance was interpolated by weighing each occurrence point based on a friction surface that reflects global human accessibility (Weiss et al., 2018). Hull peripherality (P_H) was calculated using the geometric center of a species' range. The remaining two peripherality indexes are based on interpolated abundance surfaces. Abundance peripherality (P_A) was calculated with the highest predicted abundance point as the range center. Friction peripherality (P_F) was calculated with the highest predicted abundance point when accounting for accessibility.

Both the predicted and friction abundance for each population in the selfing rate data set were saved for later analysis. The distHaversine function from Geosphere (Hijmans, 2019) was used to calculate the distance from populations and each of the three range centers. Distance to nearest range edge was also calculated for each population using the dist2Line function from Geosphere (Hijmans, 2019). Both distance to range center and distance to nearest edge were measured in meters.

To create a simple measure of a population's proximity to the nearest geographic range edge, we calculated a peripherality index. For each of the three range centers described above, peripherality scores were calculated as follows:

$$\text{Peripherality} = \frac{D_c}{D_c + D_e},$$

where D_c is the distance to the range center and D_e is the distance to the nearest range edge. Peripherality is bounded between zero and one, with high values indicating populations close to a range edge and low values indicating populations close to the range center. Because we employed three methods to estimate range centers, there are three corresponding peripherality indices to analyze.

Statistical analyses

A single, generalized linear model was used to test the hypothesis that selfing rates within species increase toward species' geographic range edges. Selfing rates were Z-transformed within each species so that all data points are expressed in terms of standard normal deviates ($m = 0$, $s = 1$). This standardization ensures that our analyses only consider within-species variation in selfing rates, since there is no mean difference in selfing rate among species. This approach also treats each species as independently sampled, which seems reasonable since data points represent standard normal deviations within each reproductively isolated lineage. Distance toward geographic range edges was determined based on the three peripherality indexes and two measures of abundance. Five distinct models were analyzed using each standardized predictor and the interaction between species and the predictor variable. In these models, species is used as part of the interaction term and explains no variation as a main effect. We also examined correlations among predictor variables and report results. Only species with ≥5 populations in the selfing rate data set were used in these analyses to reduce the influence of sampling on the ability to detect relationships. This filter reduced the number of species to 28. Analysis of variance was used to determine which, if any, effects were significant.

Functional relationships between geographic location and selfing rates

While linear models are useful in exploring predictors of selfing rate variation, they may fail to distinguish changes in relationships across the geographic range for each species. To determine the shape of the relationship between selfing rates and peripherality across each species' range, we compared linear and segmented regressions. Segmented regressions allow for insights into where changes in this relationship occur within the species' range. Segmented regressions were performed with the R package segmented (Muggeo, 2008). For these analyses, the number of species was limited to those that had the most variation in selfing rate and peripherality (top three quantiles of SD; $SD > 0.083$ for selfing rates, $SD > 0.031$ for peripherality). This filter reduced the number of species to 15. The final analysis presented here has 14 species as one species' segmented regression failed to converge. We performed linear and segmented regressions for each species with peripherality as the dependent variable. P-values cannot be calculated for segmented regressions, so these were not reported. The goal with these analyses was to compare the relative fits of linear and segmented regression models, as such AIC and BIC scores were calculated to determine which model best describes the relationship between within-species selfing rates and peripherality.

Statistical power of testing the hypothesis that selfing rates are associated with geography

Empirical studies of selfing rates are limited by the number of populations that are studied (i.e., sample size) within a species. We explored the influence of sample size on the statistical power to reject the null hypothesis that there is no association between selfing rates and geographic location. We considered cases where the true underlying correlation between selfing rates and distance to range edge varied from ($r = 0$, 0.1, 0.2, and 0.4). The distribution of these variables was assumed to follow the bivariate standard normal distribution ($m = 0$, s = 1). We simulated 1000 independent data sets, where populations were randomly sampled ($N = 5$, 10, 20, or 40). These samples span the range of populations, which are typically studied empirically. The mean estimated correlation coefficient and its 95% confidence interval was calculated, as was the average power of the test, taken from the fraction of data sets that reject the null hypothesis of $r = 0.0$.

RESULTS

Variation in life-history traits and intraspecific selfing rates

Of the 97 sampled species, 74 were perennial, and 50 were herbaceous. Annual species had significantly higher mean selfing rates than perennial species (annual = 0.460, perennial = 0.282; $P = 0.002$). Neither the SD nor the CV was significantly different for generation time (SD: $P = 0.147$; CV: $P = 0.959$; Figure 2A). Herbaceous species had significantly higher mean selfing rates than woody species (herbaceous = 0.408, woody = 0.234; $P < 0.001$). Standard deviations were also significantly elevated in herbs compared to woody species ($P = 0.009$; Figure 2B). Coefficients of variation were not significantly different between growth forms ($P = 0.938$; Figure 2B). When selfing rate variation within species was corrected by the level expected under the binomial distribution, the observed difference between annual and perennial selfing rate variation (−0.011) was not significant (95% CI of the bootstrapped differences −0.038–0.021). The observed difference between herbaceous and woody species' selfing rate variation (0.014) was also found to be not significant (95% CI of the bootstrapped differences −0.012–0.042).

Peripherality, abundance, and within-species selfing rates

In the three peripherality models, 27 species were analyzed, while 22 species were analyzed in the two abundance models. Five quantitative predictors were evaluated to determine which, if any, influenced selfing rates within species. All of these predictors were found to be highly correlated with each other, except hull peripherality, which was not significantly related to either of the two abundance measures (Figure 3). Hull peripherality is expected to be less strongly correlated with other variables, since it is based on the geometric center of a species' range, which is unlikely to have a relationship with abundance. Abundance peripherality and friction peripherality are negatively correlated with both abundance measures as expected since abundance and friction peripherality are based on predicted and friction abundance, respectively (Figure 3). On the basis of these correlations, we report only the results of the linear models for abundance peripherality and predicted abundance in the main text for ease of interpretation (for all models, see Appendix S1).

Within species, selfing rate was negatively related to both abundance peripherality (P_A) and predicted abundance, but neither were statistically significant (P_A: slope = −0.036, SE = 0.244, $P = 0.884$; predicted abundance: slope = −0.018, SE = 0.247, $P = 0.942$; Table 1, Figure 4). The interaction between abundance peripherality and species on selfing rate was marginally significant (SS = 38.937, $P = 0.057$; Table 1), while the interaction between predicted abundance and species was not (SS = 22.504, $P = 0.439$; Table 1). Selfing rate was also not significantly associated with abundance peripherality of northern range edges specifically (Appendix S1). A minority of the species (6 of 14) violate the assumptions for linear regressions, but these violations likely have little influence on our conclusions, given the apparently weak effects of predictors in the models.

Functional relationships between geographic location and selfing rates

The 14 species in this analysis had between 5 and 58 ($\bar{x} = 17.5$, SD = 16.161) populations with selfing rate data. Eight of the 14 species displayed a positive slope in the linear regression, which is the expected direction for the relationship between selfing rate and peripherality (Table 2). Eight of the 14 species displayed a breakpoint greater than 0.5, and the average breakpoint was 0.621, indicating that shifts in the relationship between selfing and peripherality occur more often toward species' range edges (Table 2). The linear model only explained a large amount of selfing rate variation in *Aquilegia coerulea* ($R^2 = 0.688$; Table 2). The segmented regression explained a large amount of selfing rate variation for both *Mimulus* guttatus ($R^2 = 0.834$; Table 2) and *Daviesia mimosoides* ($R^2 = 0.956$; Table 2). All other species had very little variance explained by either the linear or the segmented regressions (Table 2). Only four of

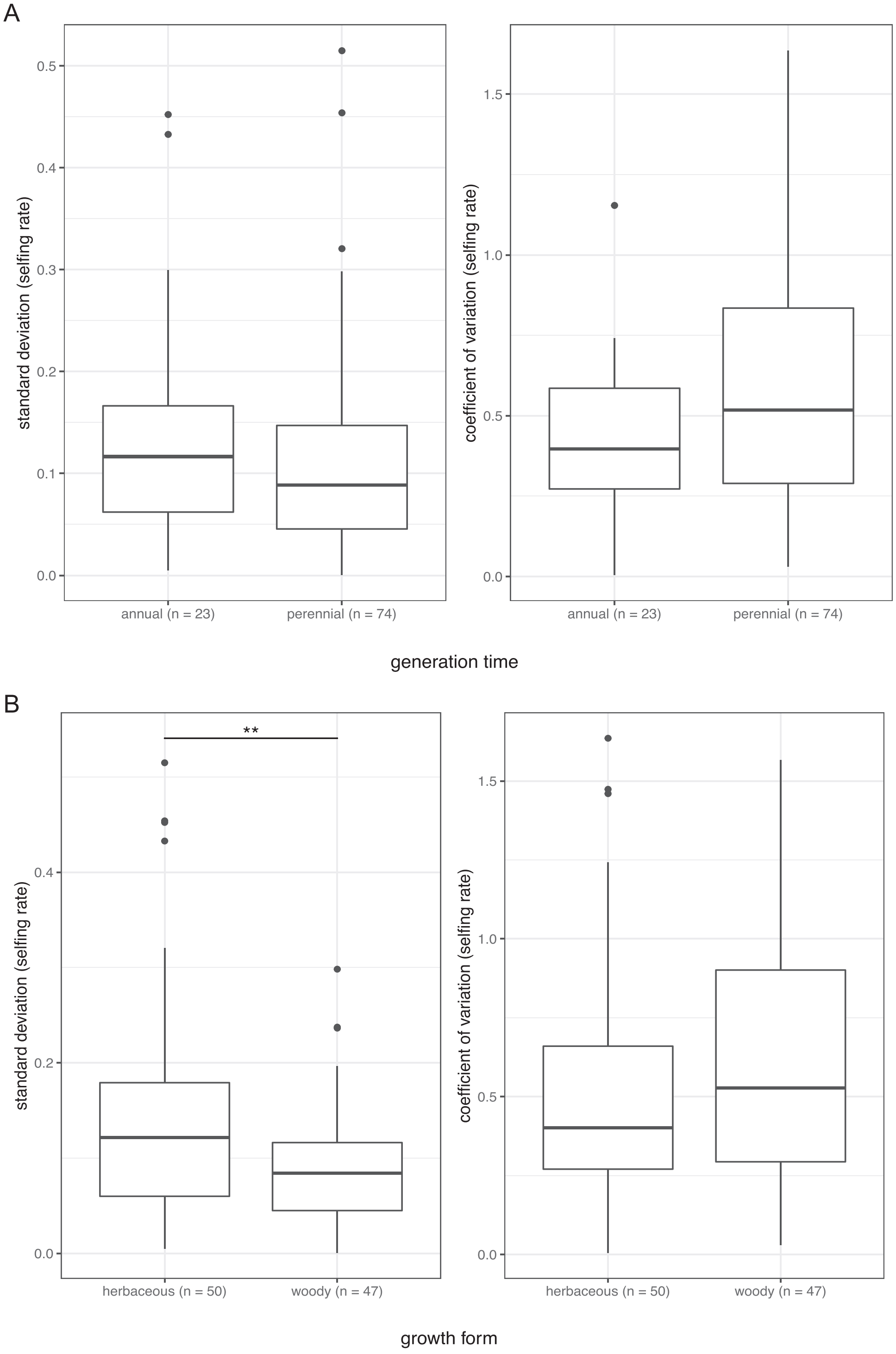

FIGURE 2 Boxplots showing differences in within-species selfing rate variation between two key life-history traits. (A) Standard deviation and coefficient of variation comparisons for within-species selfing rates between annual and perennial species. (B) Standard deviation and coefficient of variation for within-species selfing rates compared between herbaceous and woody species. Asterisks above plots indicate $P < 0.01$

the 14 species displayed a better fit with the segmented regression than the linear regression (Table 3). Of these four species, three displayed sharp changes in the direction of the relationship between selfing and peripherality (Figure 5). Only one species (*Impatiens capensis*) displayed a negative first slope and a positive second slope (Table 2, Figure 5).

Statistical power of testing the hypothesis that selfing rates are associated with geography

If there are weak or moderate correlations between selfing rates and distance to range edge ($r = 0.1$ or 0.2), studies with 20 or fewer populations will reject the null hypothesis ($r = 0$)

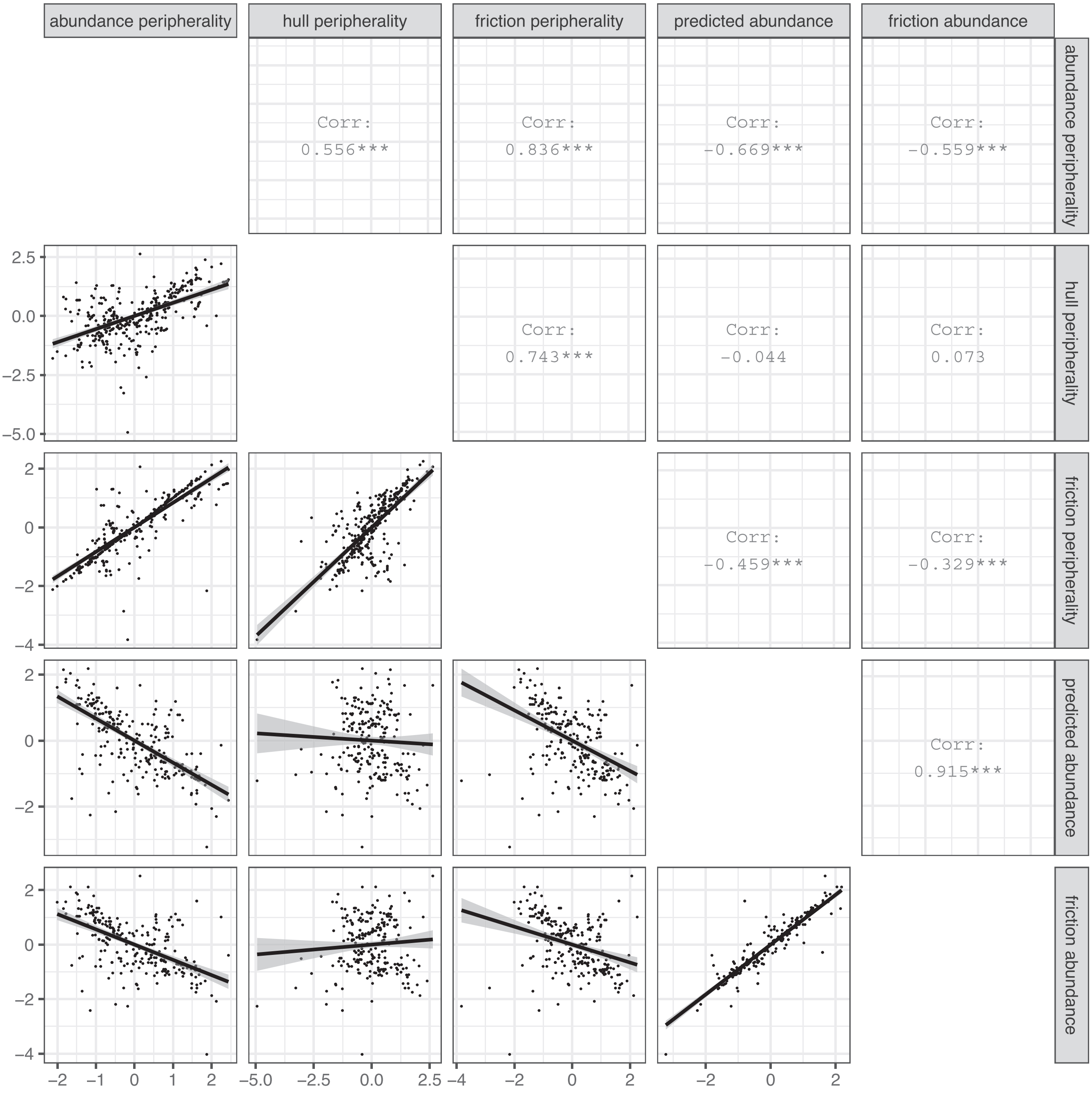

FIGURE 3 Relationships among predictors of within-species selfing rate variation. Correlation coefficients are reported between the three measures of peripherality ($N = 27$ species) and two measures of abundance ($N = 22$ species). Cells above the diagonal report correlation coefficients; those below display the data and predicted regression line. Asterisks denote significant correlations (***$P < 0.001$)

at most 13% of the time (Table 4). The power to reject the null hypothesis improves marginally with 40 sampled populations if there is a weak correlation ($r = 0.1$), though the power nearly doubles to 24.3% if there is a moderate correlation ($r = 0.2$). If there is a relatively strong correlation between selfing rates and distance to range edge ($r = 0.4$), these tests have more than twice the power over the above cases, but power still falls below 50% in empirical data sets with 20 or fewer sampled populations. In cases with a strong correlation and 40 sampled populations, power to reject the null hypothesis approaches 75% (Table 4).

DISCUSSION

In this study, variation in selfing rates within species is not well explained by life-history traits, abundance patterns, or the proximity of populations to range edges. Here, these findings

TABLE 1 Analyses of variance for two models explaining within-species variation in selfing rate. (A) Abundance peripherality and its interaction with species and (B) predicted abundance and its interaction with species. The abundance peripherality model includes 28 species with five or more populations with measured selfing rates. The predicted abundance model includes 23 species, five species were lost due to lack of variation. Significance of predictors was assessed with sums of squares (SS). The main effect of species is zero in all models as selfing rates, peripherality and abundance measures were *Z*-transformed, as such species is not reported. Bolded *P*-values indicate marginal significance

Effects	SS	df	*P*
(A) Abundance peripherality	0.019	1	0.888
Abundance peripherality × Species	38.937	27	**0.057**
(B) Predicted abundance	0.005	1	0.945
Predicted abundance × Species	22.504	22	0.439

are placed in a broader context. In addition, we make suggestions for future empirical work to better explain within-species variation in selfing rates.

Variation in life-history traits and selfing rates within species

While we found a significant difference in the standard deviation of selfing rates between herbaceous and woody species (Figure 2B), the significance does not persist when average-corrected differences are considered using coefficients of variation (Figure 2B) or binomial corrections for each species. Life-history traits therefore fail to predict the magnitude of selfing rate variation within species, though we found that with annuals and herbaceous species exhibited higher mean selfing rates than perennial and woody species, respectively. This

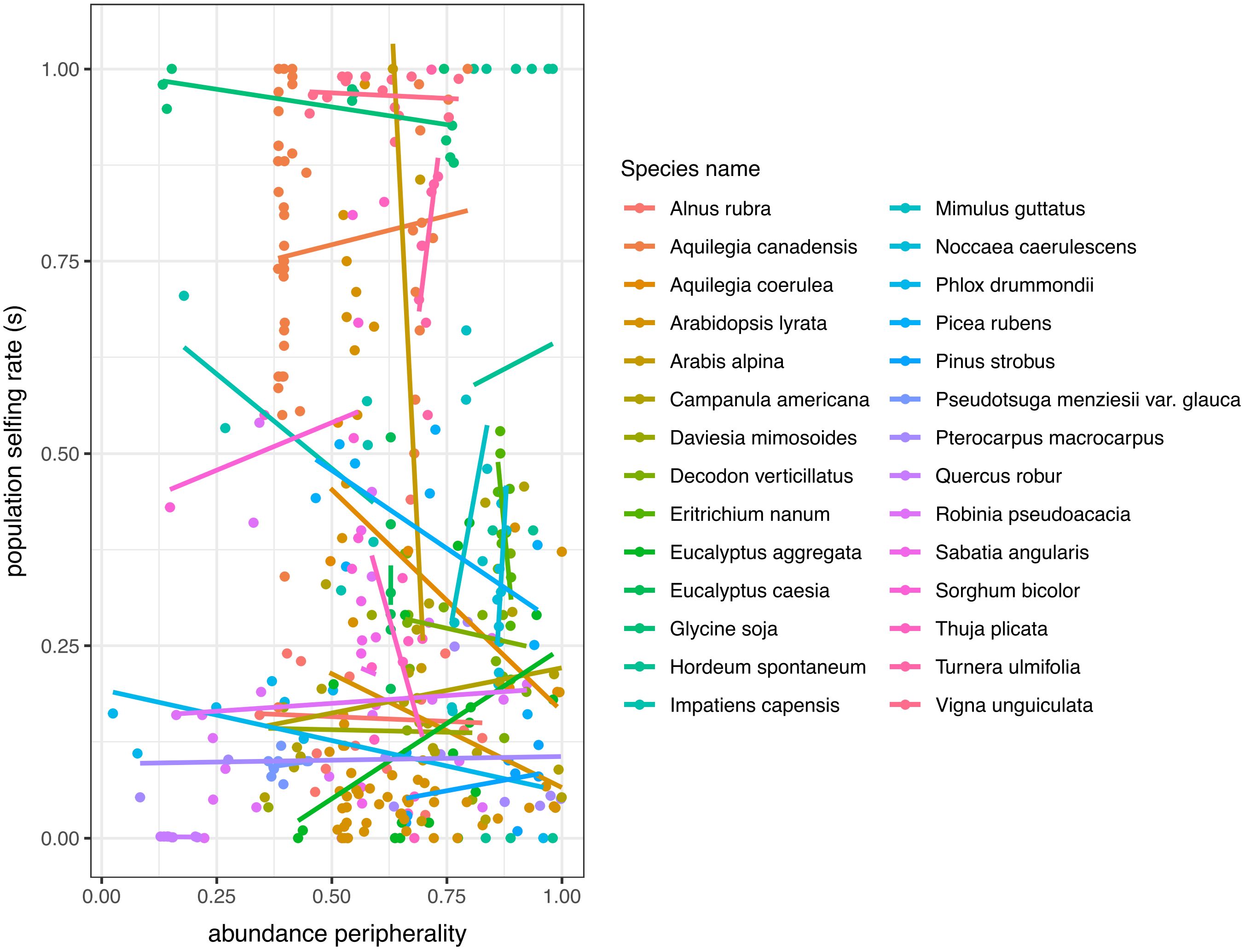

FIGURE 4 Relationship between raw abundance peripherality and population selfing rates for 27 species. Overall within species relationships are depicted by least-squares regression (solid lines)

TABLE 2 Comparison of relative fits for linear and segmented regressions for 14 species. Both regressions describe the relationship between selfing rate and abundance peripherality. Species are ordered from most positive slope to most negative slope in the linear regression. Slope and adjusted R^2 are reported for each linear regression. Segmented regressions produce two slopes which change in direction or magnitude at the breakpoint and an adjusted R^2

		Linear regression		Segmented regression			
Species	***N***	**Slope**	**Adj. R^2**	**Slope 1**	**Slope 2**	**Breakpoint**	**Adj. R^2**
Mimulus guttatus	7	3.309	0.141	21.516	−4.293	0.784	0.834
Eucalyptus aggregata	17	0.392	0.136	0.159	0.529	0.638	0.020
Hordeum spontaneum	11	0.314	−0.109	−16.944	2.254	0.842	−0.275
Sorghum bicolor	7	0.247	−0.132	0.585	−1.141	0.483	−0.848
Aquilegia canadensis	47	0.151	−0.010	−0.131	2.715	0.679	0.007
Campanula americana	23	0.118	−0.006	1.587	−0.080	0.487	0.024
Robinia pseudoacacia	23	0.041	−0.043	1.141	−0.138	0.329	−0.043
Pterocarpus macrocarpus	11	0.010	−0.110	0.186	−0.729	0.795	0.214
Daviesia mimosoides	5	−0.013	−0.333	3.873	−1.095	0.460	0.956
Alnus rubra	16	−0.025	−0.070	−0.597	0.074	0.468	−0.212
Arabidopsis lyrata	58	−0.296	0.014	−1.016	0.658	0.752	0.030
Picea rubens	9	−0.406	0.331	0.167	−1.086	0.738	0.377
Impatiens capensis	6	−0.491	0.282	−1.916	2.115	0.447	0.325
Aquilegia coerulea	5	−0.576	0.688	−0.646	10.967	0.789	0.110

TABLE 3 Comparison of relative fits for linear and segmented regressions for 14 species. The order of the species is maintained from Table 2. Bolded values indicate which model provides the best fit for each species

		Linear regression		Segmented regression	
Species	***N***	**AIC**	**BIC**	**AIC**	**BIC**
Mimulus guttatus	7	−0.531	−0.694	**−11.593**	**−11.864**
Eucalyptus aggregata	17	**−17.262**	**−14.762**	−13.553	−9.387
Hordeum spontaneum	11	**19.066**	**20.260**	21.834	23.824
Sorghum bicolor	7	**−0.973**	**−1.135**	2.882	2.611
Aquilegia canadensis	47	**−21.622**	**−16.072**	−20.556	−11.306
Campanula americana	23	**−29.479**	**−26.072**	−28.473	−22.796
Robinia pseudoacacia	23	**−17.881**	**−14.474**	−16.180	−10.502
Pterocarpus macrocarpus	11	−18.229	−17.036	**−20.795**	**−18.806**
Daviesia mimosoides	5	−3.334	−4.931	**−22.334**	**−24.287**
Alnus rubra	16	**−22.258**	**−19.940**	−18.734	−14.871
Arabidopsis lyrata	58	**6.258**	**12.440**	7.223	17.525
Picea rubens	9	**−11.909**	**−11.317**	−11.573	−10.587
Impatiens capensis	6	−5.320	−5.945	**−5.854**	**−6.895**
Aquilegia coerulea	5	**−7.471**	**−8.642**	−3.714	−5.666

finding supports other studies where herbaceous annuals have the highest mean selfing rates, followed by herbaceous perennials, and then woody perennials (Barrett and Eckert, 1990; Barrett et al., 1996; Hamrick and Godt, 1996). While there are both annual and perennial herbaceous plants in our data set, all woody species are perennial. Therefore, woody species should pay the compound cost of increased inbreeding depression but should also reap fewer potential benefits from reproductive assurance. This hypothetical compounding of costs may explain, in part, why a difference in standard deviation was seen in the growth form comparison but not in the generation time comparison (Figure 2A, B).

Relationships between selfing rates, range edges, and abundance

Within species, we found no evidence to support a general trend of increased selfing rates in range-edge populations (Figure 4). Regardless of how range centers were defined,

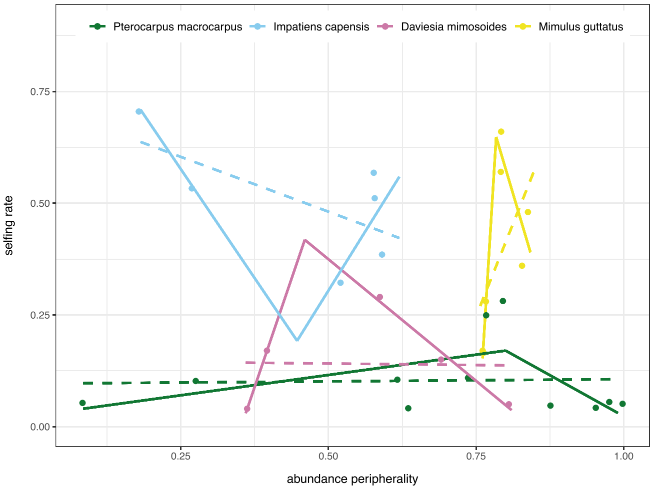

FIGURE 5 Plots of the four species for which the relationship between selfing rates and abundance peripherality were better predicted by segmented regression than linear regression. Linear regressions are represented by dashed lines and segmented regressions are represented by solid lines.

the generalized linear models only suggested variability among species in relationships between selfing rates and proximity to a geographic range edge (Table 1; Appendix S1). We also found no evidence that the absolute distance to the nearest edge (D_e) reflected a consistent relationship with selfing rates within species (Appendix S1). There has been some supposition that temperate and boreal species should display this biogeographic pattern more clearly toward northern edges because of recent colonization of high latitude environments following glacial recession (Moeller et al., 2017), a pattern that is also not supported within species in this study (Appendix S1). In perhaps the most thorough empirical challenge of this hypothesis, northern, range-edge populations did not exhibit higher selfing rates or reduced density in comparison to central populations in *Aquilegia canadensis* (Herlihy and Eckert, 2005). Even in a broad meta-analysis of selfing rate variation among species, Moeller et al. (2017) found a weak relationship with latitude. This weak association arose primarily because outcrossing rates are higher in perennial species, which are more common at higher latitudes. Thus, it appears there is an emerging consensus that selfing rates are not generally higher near the edges of species' geographic ranges (but see Griffin and Willi, 2014).

Selfing rates are expected to increase in populations that are small, patchy, or receive poor pollinator service (Baker, 1955, 1965; Fishman and Wyatt, 1999; Moeller and Geber, 2005; Morgan et al., 2005), which are often associated with proximity to geographic range edges (Abeli et al., 2014; Hargreaves and Eckert, 2014; Hargreaves et al., 2014). Our segmented regressions indicate that, on average, the relationship between selfing rates and abundance peripherality tends to change as populations near the range edge (average breakpoint = 0.621; Table 2), though segmented regressions rarely provided a better fit than linear models. This finding provides a small degree of evidence for breakpoints in the relationship between selfing rates and geography tending to occur toward range edges. However, our data do not support the hypothesis that selfing generally increases toward range edges, as only 50% of our analyzed species display a more positive slope after the breakpoint (Table 2).

TABLE 4 The power of testing the null hypothesis that selfing rates are not associated with geographic range position. Selfing rates were assumed to be positively correlated ($r = 0.1$, 0.2, or 0.4) with distance to the nearest range edge. We then randomly sampled populations from this distribution ($N = 5$, 10, 20, 40) to generate 1000 experimental data sets. Power reflects the fraction of 1000 different experiments that reject the null hypothesis of $r = 0.0$. Both selfing rates and distance to the nearest range edge were simulated from the bivariate standard normal distribution

No. of populations	True correlation	Estimated correlation	Lower 95% CI	Upper 95% CI	Power to reject H0
5	0.0	−0.016	−0.794	0.789	0.052
10	0.0	−0.003	−0.587	0.583	0.047
20	0.0	0.001	−0.421	0.422	0.056
40	0.0	−0.002	−0.306	0.302	0.053
5	0.1	0.079	−0.760	0.823	0.050
10	0.1	0.085	−0.520	0.638	0.061
20	0.1	0.103	−0.334	0.505	0.075
40	0.1	0.088	−0.222	0.381	0.082
5	0.2	0.200	−0.691	0.861	0.069
10	0.2	0.183	−0.439	0.691	0.096
20	0.2	0.192	−0.250	0.571	0.129
40	0.2	0.199	−0.112	0.475	0.243
5	0.4	0.350	−0.585	0.896	0.111
10	0.4	0.369	−0.269	0.791	0.192
20	0.4	0.390	−0.038	0.700	0.453
40	0.4	0.393	0.101	0.624	0.730

While this study found no evidence for an association of selfing rates with proximity to range edges, in general, there are well-known examples where traits facilitating self-fertilization are common near range boundaries, for example, in wild *Solanum* (Rick et al., 1977; Broz et al., 2017). In *Eichhornia paniculata*, heterostylous morphs predominate in the native range, with multiple, disjunct populations of homostylous morphs with higher selfing rates in more northerly locales (Barrett et al., 2009). Similar processes explain the association of self-compatible genotypes near range edges of *Leavenworthia alabamica* (Busch et al., 2011) and individuals with relatively low herkogamy near a range edge of *Clarkia xantiana* (Moeller and Geber, 2005; Pettengill and Moeller, 2012). In these cases, selfing traits become associated with a particular range edge, primarily because expansion at that edge provided permissive conditions for the evolution of self-fertilization (Koski et al., 2019b; Encinas-Viso et al., 2020). Additional historical processes, such as spatially restricted hybridization with a congener, explain the association of self-compatible, more highly selfing populations of *Camissoniopsis cheiranthifolia* with a specific range edge (Dart et al., 2012; López-Villalobos and Eckert, 2019). In all of these instances, knowledge of a species' phylogeographic history is essential to explain why self-fertilization sometimes predominates near specific edges of a species range, where recent declines in mates or pollinators may have occurred (Griffin and Willi, 2014).

In this study, abundance measures failed to reveal higher selfing in low abundance populations, regardless of geographic location. Relationships between density and selfing rates have been well documented, further supporting the idea that selfing rates increase whenever mates are scarce (Barrett and Eckert, 1990; Karron et al., 1995; Brunet and Sweet, 2006; Opedal, 2018; Koski et al., 2019a). Since relations between density and abundance relationships are generally positive (Gaston et al., 2000), abundances should also exhibit a positive relationship with selfing rates, albeit to a weaker extent. Relationships between density and abundance are not universally recovered. Studies use various abundance measures, some of which capture temporal fluctuations in abundance better than others (Gaston et al., 2000). Population-level density is commonly identified through classic quadrat techniques where reproductive individuals are counted across each population (Herlihy and Eckert, 2005). Density measures therefore capture information about each population in the year in which they were measured. In contrast, abundances reflect occurrence data for species, which naturally accumulate over long periods of time and are sensitive to sampling bias (Boakes et al., 2010; Beck et al., 2014).

A few remedies have been proposed to ensure that abundances accurately reflect the density of individuals. Occurrence data are commonly informed by spatial variation in climatic variables, since the latter provide insight on the realized niche of species (Peterson et al., 2011; Yañez-Arenas et al., 2014; Carrascal et al., 2015). More recently, Yin and He (2014) proposed a method that takes spatial autocorrelation into account when assessing abundance from occurrence data. Our method also accounts for spatial autocorrelation while interpolating many years of species occurrence data across a species' native range. In doing so, our method permits abundance to be estimated over relatively large scales, but this approach is insensitive to temporal fluctuations that potentially influence selfing rates. While aggregating occurrences over time generates more data on abundance, combining data from many years further separates abundance from population-level demographic changes (Gaston et al., 1999, 2000; De Bona et al., 2019). Given that selfing rates fluctuate reliably with density, future studies on the geography of selfing rate variation should measure density at the time in which parents mate and produce offspring, as has been done reliably in the past (Karron et al., 1995; Brunet and Sweet, 2006; Opedal, 2018).

CONCLUSIONS

Although plant mating systems have been studied extensively, there is still a need to explain variation in selfing rates within species (Whitehead et al., 2018). When exploring intraspecific variation in species' mating systems across their geographic ranges, species were required to have at least five populations with selfing rate estimates for inclusion in this study. With this threshold and native filtering of occurrence data, 73% of species were lost from the original data set (71 out of 97 species) for all geographic analyses. Given that selfing rates are rarely estimated in more than 10 populations within species, it is perhaps not surprising we failed to detect the hypothesized pattern of increased selfing toward species range edges (Table 4). Even in cases with a strong underlying association between selfing and geography, empirical tests will lack sufficient power unless many populations are studied. However, even in species with the most sampled populations (*Arabidopsis lyrata*, $N = 58$; *Aquilegia canadensis*, $N = 47$; *Robina pseudoacacia*, $N = 23$; *Campanula americana*, $N = 23$) no clear biogeographic pattern in selfing rates arises (Table 2). It seems likely either that selfing is unrelated to range edges, or at most is only weakly associated with general proximity to range edges.

Even if large samples of populations ($N = 40+$) are studied empirically in the future, additional hurdles remain. Selfing rates, like most traits expressed at the population level, are often not evenly estimated across species' geographic ranges (Sagarin and Gaines, 2002; Herlihy and Eckert, 2005), though some studies have approximated this sampling scheme (Koski et al., 2019a). Sampling populations in a few locations across the geographic range should represent a considerable constraint, especially for widespread species (Santini et al., 2019). Moreover, if most populations are sampled at similar distances from geographic range limits, measures such as abundance peripherality will lack utility in predicting patterns of mating systems. Uneven sampling effort across a geographic range has emerged as a common hurdle in other studies attempting to detect biogeographic patterns using opportunistically collected data (Santini et al., 2019). In addition to these issues regarding the spatial aggregation of samples, temporal scales must also be considered. Given the potential for large among-year fluctuations in selfing rates within natural populations (Eckert et al., 2009), traits that predispose plants to self-fertilize (e.g., herkogamy, autonomy) yet do not fluctuate year-to-year are better for investigating spatial patterns in the evolution and ecology of mating systems (Rick et al., 1977; Opedal, 2018; Koski et al., 2019b).

ACKNOWLEDGMENTS

The authors thank O. Cornejo, E. Hilpman, N. Layman, J. Rainbow, and S. Tisinai for critiquing the ideas presented here. J. Brunner suggested the power analysis, and anonymous reviewers also provided helpful suggestions for improvement. This study was supported by funds from NSF 1457037 and NSF 1911313.

AUTHOR CONTRIBUTIONS

C.J.P. conceived the project, gathered data, conducted analyses, and wrote the manuscript. J.W.B. conceived the project, interpreted data, and helped write the manuscript.

DATA AVAILABILITY STATEMENT

All data and scripts are available in the Dryad repository at https://doi.org/10.5061/dryad.3j9kd51jw (Busch and Prior, 2021).

ORCID

Carly J. Prior http://orcid.org/0000-0002-2194-7864

REFERENCES

Abeli, T., R. Gentili, A. Mondoni, S. Orsenigo, S., and G. Rossi. 2014. Effects of marginality on plant population performance. *Journal of Biogeography* 41: 239–249.

Baddeley, A., E. Rubak, and R. Turner. 2015. Spatial point patterns: methodology and applications with R. Chapman and Hall, London, UK/CRC Press, Boca Raton, FL, USA.

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

Baker, H. G. 1965. Characteristics and modes of origin of weeds. *In* G. L. Stebbins and H. G. Baker [eds.], The genetics of colonizing species, 147–172. Academic Press, NY, NY, USA.

Baker, H. G. 1967. Support for Baker's law-as a rule. *Evolution* 21: 853–856.

Barrett, S. C. H. 2010. Understanding plant reproductive diversity. *Philosophical Transactions of the Royal Society, B, Biological Sciences* 365: 99–109.

Barrett, S. C. H., and C. G. Eckert. 1990. Variation and evolution of mating systems in seed plants. *In* S. Kawano [ed.], Biological approaches and evolutionary trends in plants, 229–254. Academic Press, San Diego, CA, USA.

Barrett, S. C. H., and L. D. Harder. 2017. The ecology of mating and its evolutionary consequences in seed plants. *Annual Review of Ecology, Evolution, and Systematics* 48: 135–157.

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 of London, B, Biological Sciences* 351: 1271–1280.

Barrett, S. C. H., and B. C. Husband. 1990. Variation in outcrossing rates in *Eichhornia paniculata*: the role of demographic and reproductive factors. *Plant Species Biology* 5: 41–55.

Barrett, S. C. H., R. W. Ness, and M. Vallejo-Marín. 2009. Evolutionary pathways to self-fertilization in a tristylous plant species. *New Phytologist* 183: 546–556.

Beck, J., M. Böller, A. Erhardt, and W. Schwanghart. 2014. Spatial bias in the GBIF database and its effect on modeling species' geographic distributions. *Ecological Informatics* 19: 10–15.

Boakes, E. H., P. J. McGowan, R. A. Fuller, C.-q. Ding, N. E. Clark, K. O'Connor, and G. M. Mace. 2010. Distorted views of biodiversity: spatial and temporal bias in species occurrence data. *PLoS Biology* 8: e1000385.

Boyle, B., N. Hopkins, Z. Lu, J. A. R. Garay, D. Mozzherin, T. Rees, N. Matasci, et al. 2013. The taxonomic name resolution service: an online tool for automated standardization of plant names. *BMC Bioinformatics* 14: 16.

Brown, J. H. 1984. On the relationship between abundance and distribution of species. *American Naturalist* 124: 255–279.

Broz, A. K., A. M. Randle, S. A. Sianta, A. Tovar-Mendez, and P. A. Bedinger. 2017. Mating system transitions in *Solanum habrochaites* impact interactions between populations and species. *New Phytologist* 213: 440–454.

Brunet, J., and H. R. Sweet. 2006. Impact of insect pollinator group and floral display size on outcrossing rate. *Evolution* 60: 234–246.

Busch, J. W., and L. F. Delph. 2012. The relative importance of reproductive assurance and automatic selection as hypotheses for the evolution of self-fertilization. *Annals of Botany* 109: 553–562.

Busch, J. W., S. Joly, and D. J. Schoen. 2011. Demographic signatures accompanying the evolution of selfing in *Leavenworthia alabamica*. *Molecular Biology and Evolution* 28: 1717–1729.

Busch, J., and C. Prior. 2021. Data from: selfing rate variation within species is unrelated to life-history traits or geographic range position. *Dryad Digital Repository*. https://doi.org/10.5061/dryad.3j9kd51jw

Carrascal, L. M., P. Aragón, D. Palomino, and J. M. Lobo. 2015. Predicting regional densities from bird occurrence data: validation and effects of species traits in a Macaronesian island. *Diversity and Distributions* 21: 1284–1294.

Charlesworth, D., and J. H. Willis. 2009. The genetics of inbreeding depression. *Nature Reviews Genetics* 10: 783–796.

Charlesworth, D., M. T. Morgan, and B. Charlesworth. 1992. The effect of linkage and population size on inbreeding depression due to mutational load. *Genetical Research* 59: 49–61.

Coates, D. J., J. F. Sampson, and C. J. Yates. 2007. Plant mating systems and assessing population persistence in fragmented landscapes. *Australian Journal of Botany* 55: 239–249.

Crawford, D. J., J. K. Archibald, D. Stoermer, M. E. Mort, J. K. Kelly, and A. Santos-Guerra. 2008. A test of Baker's law: breeding systems and the radiation of *Tolpis* (Asteraceae) in the Canary Islands. *International Journal of Plant Sciences* 169: 782–791.

Dallas, T., R. R. Decker, and A. Hastings. 2017. Species are not most abundant in the centre of their geographic range or climatic niche. *Ecology Letters* 20: 1526–1533.

Darling, E., K. E. Samis, and C. G. Eckert. 2008. Increased seed dispersal potential towards geographic range limits in a Pacific coast dune plant. *New Phytologist* 178: 424–435.

Dart, S. R., K. E. Samis, E. Austen, and C. G. Eckert. 2012. Broad geographic covariation between floral traits and the mating system in *Camissoniopsis cheiranthifolia* (Onagraceae): multiple stable mixed mating systems across the species' range? *Annals of Botany* 109: 599–611.

De Bona, S., M. Bruneaux, A. E. Lee, D. N. Reznick, P. Bentzen, and A. López-Sepulcre. 2019. Spatio-temporal dynamics of density-dependent dispersal during a population colonisation. *Ecology Letters* 22: 634–644.

Duminil, J., O. J. Hardy, and R. J. Petit. 2009. Plant traits correlated with generation time directly affect inbreeding depression and mating system and indirectly genetic structure. *BMC Evolutionary Biology* 9: 177.

Eckert, C. G., B. Ozimec, C. R. Herlihy, C. A. Griffin, and M. B. Routley. 2009. Floral morphology mediates temporal variation in the mating system of a self-compatible plant. *Ecology* 90: 1540–1548.

Eckert, C. G., K. E. Samis, and S. Dart. 2006. Reproductive assurance and the evolution of uniparental reproduction in flowering plants. *In* Barrett, S. C. H., and L. D. Harder [eds.], The ecology and evolution of flowers, 183–203. Oxford University Press, Oxford, UK.

Eckert, C. G., K. E. Samis, and S. C. Lougheed. 2008. Genetic variation across species' geographical ranges: the central–marginal hypothesis and beyond. *Molecular Ecology* 17: 1170–1188.

Encinas-Viso, F., A. G. Young, and J. R. Pannell. 2020. The loss of self-incompatibility in a range expansion. *Journal of Evolutionary Biology* 33: 1235–1244.

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.

Freckleton, R. P., P. H. Harvey, and M. Pagel. 2003. Bergmann's rule and body size in mammals. *American Naturalist* 161: 821–825.

Gaston, K. J., and T. M. Blackburn. 2000. Pattern and process in macroecology. Blackwell, Malden, MA, USA.

Gaston, K. J., T. M. Blackburn, J. J. Greenwood, R. D. Gregory, R. M. Quinn, and J. H. Lawton. 2000. Abundance–occupancy relationships. *Journal of Applied Ecology* 37: 39–59.

Gaston, K. J., R. D. Gregory, and T. M. Blackburn. 1999. Intraspecific relationships between abundance and occupancy among species of Paridae and Sylviidae in Britian. *Ecoscience* 6: 131-142.

Ghazoul, J. 2005. Buzziness as usual? Questioning the global pollination crisis. *Trends in Ecology and Evolution* 20: 367–373.

Goodwillie, C., S. Kalisz, and C. G. Eckert. 2005. The evolutionary enigma of mixed mating systems in plants: occurrence, theoretical explanations, and empirical evidence. *Annual Review of Ecology, Evolution, and Systematics* 36: 47–79.

Grant, A. G., and S. Kalisz. 2020. Do selfing species have greater niche breadth? Support from ecological niche modeling. *Evolution* 74: 73–88.

Griffin, P. C., and Y. Willi. 2014. Evolutionary shifts to self-fertilisation restricted to geographic range margins in North American *Arabidopsis lyrata*. *Ecology Letters* 17: 484-490.

Grossenbacher, D. L., Y. Brandvain, J. R. Auld, M. Burd, P. O. Cheptou, J. K. Conner, A. Grant, et al. 2017. Self-compatibility is over-represented on islands. *New Phytologist* 215: 469–478.

Grossenbacher, D., R. Briscoe Runquist, E. E. Goldberg, and Y. Brandvain. 2015. Geographic range size is predicted by plant mating system. *Ecology Letters* 18: 706–713.

Haddadchi, A., and M. Fatemi. 2015. Self-compatibility and floral traits adapted for self-pollination allow homostylous *Nymphoides geminata* (Menyanthaceae) to persist in marginal habitats. *Plant Systematics and Evolution* 301: 239–250.

Hamrick, J. L., and M. J. W. Godt. 1996. Effects of life history traits on genetic diversity in plant species. *Philosophical Transactions of the Royal Society of London, B, Biological Sciences* 351: 1291–1298.

Hargreaves, A. L., and C. G. Eckert. 2014. Evolution of dispersal and mating systems along geographic gradients: implications for shifting ranges. *Functional Ecology* 28: 5–21.

Hargreaves, A. L., K. E. Samis, and C. G. Eckert. 2014. Are species' range limits simply niche limits writ large? A review of transplant experiments beyond the range. *American Naturalist* 183: 157–173.

Herlihy, C. R., and C. G. Eckert. 2002. Genetic cost of reproductive assurance in a self-fertilizing plant. *Nature* 416: 320–323.

Herlihy, C. R., and C. G. Eckert. 2005. Evolution of self-fertilization at geographical range margins? A comparison of demographic, floral, and mating system variables in central vs. peripheral populations of *Aquilegia canadensis* (Ranunculaceae). *American Journal of Botany* 92: 744–751.

Hijmans, R. J. 2019. geosphere: spherical trigonometry. R package version 1.5-10. Website: https://CRAN.R-project.org/package=geosphere

Hijmans, R., N. Garcia, and J. Wieczorek. 2018. Global administrative areas database (GADM) version 3.6. Website: gadm.org. UN, NY, NY, USA.

Hilborn, R., and M. Mangel. 1997. The ecological detective: confronting models with data. Princeton University Press, Princeton, NJ, USA.

Holsinger, K. 2001. Reproductive systems and evolution in vascular plants. *Proceedings of the National Academy of Sciences, USA* 97: 7037–7042.

Jacquemyn, H., R. Brys, and M. Hermy. 2002. Patch occupancy, population size and reproductive success of a forest herb (*Primula elatior*) in a fragmented landscape. *Oecologia* 130: 617–625.

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

Kalisz, S., D. W. Vogler, and K. M. Hanley. 2004. Context-dependent autonomous self-fertilization yields reproductive assurance and mixed mating. *Nature* 430: 884–887.

Karron, J. D., N. N. Thumser, R. Tucker, and A. J. Hessenauer. 1995. The influence of population density on outcrossing rates in *Mimulus ringens*. *Heredity* 75: 175–180.

Kattge, J., G. Bönisch, S. Díaz, S. Lavorel, I. C. Prentice, P. Leadley, S. Tauthenhan, et al. 2020. TRY plant trait database–enhanced coverage and open access. *Global Change Biology* 26: 119–188.

Kirkpatrick, M., and P. Jarne. 2000. The effect of a bottleneck on inbreeding depression and the genetic load. *American Naturalist* 155: 154–167.

Klekowski, E. J., and P. J. Godfrey. 1989. Ageing and mutation in plants. *Nature* 340: 389–391.

Knight, T. M. 2003. Floral density, pollen limitation, and reproductive success in *Trillium grandiflorum*. *Oecologia* 137: 557–563.

Knight, T. M., J. A. Steets, J. C. Vamosi, S. J. Mazer, M. Burd, D. R. Campbell, M. R. Dudash, et al. 2005. Pollen limitation of plant reproduction: pattern and process. *Annual Review of Ecology, Evolution, and Systematics* 36: 467–497.

Koski, M. H., L. F. Galloway, and J. W. Busch. 2019a. Pollen limitation and autonomous selfing ability interact to shape variation in outcrossing rate across a species range. *American Journal of Botany* 106: 1240–1247.

Koski, M. H., N. C. Layman, C. J. Prior, J. W. Busch, and L. F. Galloway. 2019b. Selfing ability and drift load evolve with range expansion. *Evolution Letters* 3: 500–512.

Lande, R., and D. W. Schemske. 1985. The evolution of self-fertilization and inbreeding depression in plants. I. Genetic models. *Evolution* 39: 24–40.

Layman, N. C., M. T. R. Fernando, C. R. Herlihy, and J. W. Busch. 2017. Costs of selfing prevent the spread of a self-compatibility mutation that causes reproductive assurance. *Evolution* 71: 884–897.

López-Villalobos, A., and C. G. Eckert. 2019. The contribution of hybridization to range-wide population genetic structure in a Pacific coastal dune plant. *American Journal of Botany* 106: 1575–1588.

Lowry, E., and S. E. Lester. 2006. The biogeography of plant reproduction: potential determinants of species' range sizes. *Journal of Biogeography* 33: 1975–1982.

Maitner, B. S., B. Boyle, N. Casler, R. Condit, J. Donoghue, S. M. Durán, D. Guaderrama, et al. 2018. The BIEN R package: a tool to access the Botanical Information and Ecology Network (BIEN) database. *Methods in Ecology and Evolution* 9: 373–379.

Meiri, S., N. Cooper, and A. Purvis. 2007. The island rule: made to be broken? *Proceedings of the Royal Society of London, B, Biological Sciences* 275: 141–148.

Moeller, D. A. 2006. Geographic structure of pollinator communities, reproductive assurance, and the evolution of self-pollination. *Ecology* 87: 1510–1522.

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

Moeller, D. A., M. A. Geber, V. M. Eckhart, and P. Tiffin. 2012. Reduced pollinator service and elevated pollen limitation at the geographic range limit of an annual plant. *Ecology* 93: 1036–1048.

Moeller, D. A., R. D. Briscoe Runquist, A. M. Moe, M. A. Geber, C. Goodwillie, P. O. Cheptou, C. G. Eckert, et al. 2017. Global biogeography of mating system variation in seed plants. *Ecology Letters* 20: 375–384.

Morgan, M. T. 2001. Consequences of life history for inbreeding depression and mating system evolution in plants. *Proceedings of the Royal Society of London, B, Biological Sciences* 268: 1817–1824.

Morgan, M. T., D. J. Schoen, and T. M. Bataillon. 1997. The evolution of self-fertilization in perennials. *American Naturalist* 150: 618–638.

Morgan, M. T., W. G. Wilson, and T. M. Knight. 2005. Plant population dynamics, pollinator foraging, and the selection of self-fertilization. *American Naturalist* 166: 169–183.

Muggeo, V. M. 2008. Segmented: an R package to fit regression models with broken-line relationships. *R News* 8: 20–25.

Mustajärvi, K., P. Siikamäki, S. Rytkönen, and A. Lammi. 2001. Consequences of plant population size and density for plant–pollinator interactions and plant performance. *Journal of Ecology* 89: 80–87.

Noël, E., P. Jarne, S. Glémin, A. MacKenzie, A. Segard, V. Sarda, and P. David. 2017. Experimental evidence for the negative effects of self-fertilization on the adaptive potential of populations. *Current Biology* 27: 237–242.

Opedal, Ø. H. 2018. Herkogamy, a principal functional trait of plant reproductive biology. *International Journal of Plant Sciences* 179: 677–687.

Pannell, J. R., and S. C. H. Barrett. 1998. Baker's law revisited: reproductive assurance in a metapopulation. *Evolution* 52: 657–668.

Peterson, A. T., J. Soberón, R. G. Pearson, R. P. Anderson, E. Martínez-Meyer, M. Nakamura, and M. B. Araújo. 2011. Ecological niches and geographic distributions. Princeton University Press, Princeton, NJ, USA.

Pettengill, J. B., and D. A. Moeller. 2012. Phylogeography of speciation: allopatric divergence and secondary contact between outcrossing and selfing *Clarkia*. *Molecular Ecology* 21: 4578–4592.

Pironon, S., G. Papuga, J. Villellas, A. L. Angert, M. B. García, and J. D. Thompson. 2017. Geographic variation in genetic and demographic performance: new insights from an old biogeographical paradigm. *Biological Reviews* 92: 1877–1909.

Poisot, T. 2011. The digitize package: extracting numerical data from scatterplots. *R Journal* 3: 25–26.

Pujol, B., S. R. Zhou, J. S. Vilas, and J. R. Pannell. 2009. Reduced inbreeding depression after species range expansion. *Proceedings of the National Academy of Sciences, USA* 106: 15379–15383.

Randle, A. M., J. B. Slyder, and S. Kalisz. 2009. Can differences in autonomous selfing ability explain differences in range size among sister-taxa pairs of *Collinsia* (Plantaginaceae)? An extension of Baker's Law. *New Phytologist* 183: 618–629.

Rick, C. M., J. F. Fobes, and M. Holle. 1977. Genetic variation in *Lycopersicon pimpinellifolium*: evidence of evolutionary change in mating systems. *Plant Systematics and Evolution* 127: 139–170.

Ritland, K. 1990. A series of FORTRAN computer programs for estimating plant mating systems. *Journal of Heredity* 81: 236–237.

Ritland, K. 2002. Extensions of models for the estimation of mating systems using *n* independent loci. *Heredity* 88: 221–228.

Ritland, K., and S. Jain. 1981. A model for the estimation of outcrossing rate and gene frequencies using *n* independent loci. *Heredity* 47: 35–52.

Sagarin, R. D., and S. D. Gaines. 2002. The 'abundant centre' distribution: To what extent is it a biogeographical rule? *Ecology Letters* 5: 137–147.

Santini, L., S. Pironon, L. Maiorano, and W. Thuiller. 2019. Addressing common pitfalls does not provide more support to geographical and ecological abundant-centre hypotheses. *Ecography* 42: 696–705.

Schemske, D. W., and R. Lande. 1985. The evolution of self-fertilization and inbreeding depression in plants. II. Empirical observations. *Evolution* 39: 41–52.

Schoen, D. J. 1982. The breeding system of *Gilia achilleifolia*: variation in floral characteristics and outcrossing rate. *Evolution* 36: 352–360.

Schoen, D. J., and A. H. D. Brown. 1991. Intraspecific variation in population gene diversity and effective population size correlates with the mating system in plants. *Proceedings of the National Academy of Sciences, USA* 88: 4494–4497.

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

Scofield, D. G., and S. T. Schultz. 2006. Mitosis, stature and evolution of plant mating systems: low-Φ and high-Φ plants. *Proceedings of the Royal Society, B, Biological Sciences* 273: 275–282.

Siol, M., I. Bonnin, I. Olivieri, J. M. Prosperi, and J. Ronfort. 2007. Effective population size associated with self-fertilization: lessons from temporal changes in allele frequencies in the selfing annual *Medicago truncatula*. *Journal of Evolutionary Biology* 20: 2349–2360.

Solbrig, O. T. [ed.]. 1980. Demography and evolution in plant populations. University of California Press, Berkeley, CA, USA.

Vogler, D. W., and S. Kalisz. 2001. Sex among the flowers: the distribution of plant mating systems. *Evolution* 55: 202–204.

Walker, B. 2020. bazr: Baz's useful functions. R package version 0.0.0.9000. Website: https://github.com/barnabywalker/bazr

Weiss, D. J., A. Nelson, H. S. Gibson, W. Temperley, S. Peedell, A. Lieber, M. Hancher, et al. 2018. A global map of travel time to cities to assess inequalities in accessibility in 2015. *Nature* 553: 333-336.

Whitehead, M. R., R. Lanfear, R. J. Mitchell, and J. D. Karron. 2018. Plant mating systems often vary widely among populations. *Frontiers in Ecology and Evolution* 6: 38.

Yancovitch Shalom, H., I. Granot, S. A. Blowes, A. Friedlander, C. Mellin, C. E. L. Ferreira, J. E. Arias-González, et al. 2020. A closer examination of the 'abundant centre' hypothesis for reef fishes. *Journal of Biogeography* 47: 2194–2209.

Yañez-Arenas, C., R. Guevara, E. Martínez-Meyer, S. Mandujano, and J. M. Lobo. 2014. Predicting species' abundances from occurrence data: effects of sample size and bias. *Ecological Modelling* 294: 36–41.

Yin, D., and F. He. 2014. A simple method for estimating species abundance from occurrence maps. *Methods in Ecology and Evolution* 5: 336–343.

Yin, G., S. C. H. Barrett, Y. B. Luo, and W. N. Bai. 2016. Seasonal variation in the mating system of a selfing annual with large floral displays. *Annals of Botany* 117: 391–400.

SUPPORTING INFORMATION

Additional supporting information may be found online in the supporting information section at the end of the article.

Appendix S1. Analyses of variance for two models explaining within-species variation in selfing rate. (A) Abundance peripherality and its interaction with species, (B) hull peripherality and its interaction with species, (C) friction peripherality and its interaction with species, (D) predicted abundance and its interaction with species, (E) predicted abundance and its interaction with species, (F) friction abundance and its interaction with species, and (G) distance to the northern most edge and its interaction with species. The peripherality models and the distance to northernmost-edge model includes 28 species with five or more populations with measured selfing rates. The abundance models include 23 species; five species were lost due to lack of variation. Significance of predictors was assessed with sums of squares (SS). Bolded *P*-values indicate marginal significance.

How to cite this article: Prior, C. J., and J. W. Busch. 2021. Selfing rate variation within species is unrelated to life-history traits or geographic range position. *American Journal of Botany* 108(11): 2294–2308. https://doi.org/10.1002/ajb2.1766