Global biogeography of mating system variation in seed plants

David A. Moeller[1]*, Ryan D. Briscoe Runquist[1]*, Annika M. Moe[1]*, Monica A. Geber[2], Carol

Goodwillie[3], Pierre-Olivier Cheptou[4], Christopher G. Eckert[5], Elizabeth Elle[6], Mark O. Johnston[7],

Susan Kalisz[8], Richard H. Ree[9], Risa D. Sargent[10], Mario Vallejo-Marin[11], and Alice A. Winn[12]

* D.A.M., R.D.B.R, and A.M.M. contributed equally to this work.
[1] Department of Plant Biology, University of Minnesota, St. Paul, MN 55108, USA
[2] Department of Ecology and Evolutionary Biology, Cornell University, Ithaca, NY 14853, USA
[3] Department of Biology, East Carolina University, Greenville, NC 27858, USA
[4] Centre d'Ecologie Fonctionnelle et Evolutive (CNRS), Montpelier Cedex 05, France
[5] Department of Biology, Queen's University, Kingston, Ontario K7L 3N6, Canada
[6] Department of Biological Sciences, Simon Fraser University, Burnaby, BC V5A 1S6, Canada
[7] Department of Biology, Dalhousie University, Halifax, NS B3H 4R2, Canada
[8] Department of Ecology and Evolutionary Biology, University of Tennessee, Knoxville, TN 37996,
   USA
[9] Botany Department, Field Museum of Natural History, Chicago, IL 60605, USA
[10] Department of Biology, University of Ottawa, Ottawa, Ontario K1N 6N5, Canada
[11] Biological and Environmental Sciences, University of Stirling, Stirling, FK9 4LA. Scotland,
   United Kingdom
[12] Department of Biological Science, Florida State University, Tallahassee, FL 32306, USA

Running header: Biogeography of plant mating systems

Keywords: biotic interactions, breeding system, floral evolution, latitudinal gradient, life history, outcrossing, plant-pollinator interaction, pollination, self-fertilization, sexual system

Type of article: Letter; No. of words: ; No. of references: 66; No. of figures: 4

Correspondence: David A. Moeller; Dept. of Plant Biology, U. of Minnesota, St. Paul, MN 55108.
email: moeller@umn.edu; phone: (612) 624-1037; fax: (612) 625-1738.

Statement of Authorship: All authors contributed to the design of the study, DM, RBR, AM, MG, and CG collected the data, RBR, AM, and DM conducted the analyses, DM wrote the first draft of the manuscript, and all authors contributed substantially to revisions.

Data accessibility statement: Should the manuscript be accepted, all data supporting the results will be archived in Dryad and the data DOI will be included at the end of the article.

ABSTRACT

Latitudinal gradients in biotic interactions have often been suggested as causes of global patterns of biodiversity and phenotypic variation. Plant biologists have long speculated that primarily outcrossing mating systems are more common at low than high latitudes owing to a greater predictability of plant-pollinator interactions in the tropics; however, these ideas have not previously been tested. Here, we present the first global biogeographic analysis of plant mating systems based on 624 published studies from 492 taxa. We found a weak decline in outcrossing rate toward higher latitudes and among some biomes, but no biogeographic patterns in the frequency of self-incompatibility. Incorporating life history and growth form into biogeographic analyses reduced or eliminated the importance of latitude and biome in predicting outcrossing or self-incompatibility. Our results suggest that biogeographic patterns in mating system are more likely a reflection of the frequency of life forms across latitudes rather than the strength of plant-pollinator interactions.

INTRODUCTION

Geographic variation in the strength of biotic interactions is often considered central to explanations for latitudinal diversity gradients and the evolution of traits that mediate interactions (Mittelbach et al. 2007; Schemske et al. 2009). Stronger biotic interactions in the tropics may occur due to greater productivity and climate stability (Dobzhansky 1950; Paine 1966, Pianka 1966). Those stronger biotic interactions, in turn, may have consequences for the evolution of enemy defenses and traits that mediate mutualistic interactions (e.g. Pennings et al. 2001; Dyer and Coley 2002; Moles et al. 2007). There is evidence from some studies of stronger interactions and greater specialization at low latitudes (reviewed in Coley and Barone 1996; Schemske et al. 2009) but not others (e.g. Hille Ris Lambers et al. 2002; Moles et al. 2011; Ollerton 2012; Schleuning et al. 2012). Despite intensive efforts, there remains no consensus about the ubiquity of latitudinal biotic interactions gradients and their evolutionary consequences.

Plant-pollinator interactions have been intensively studied because of their role in floral diversification (Kay and Sargent 2009). The strength and specialization of plant-pollinator interactions may be particularly consequential for effecting outcrossing and the evolution of mating systems. Mating system variation is common within and among seed plant lineages (Stebbins 1974; Schemske & Lande 1985; Barrett & Eckert 1990) including not only alternative extremes of predominant outcrossing or selfing but also "mixed mating" systems (Vogler & Kalisz 2001; reviewed in Goodwillie et al. 2005). It has been increasingly appreciated that plant-pollinator interactions contribute substantially to the selection of mating systems by affecting pollen limitation of reproduction (Morgan & Wilson 2005), geitonogamous selfing (Lloyd 1992; Harder & Barrett 1995), and gamete discounting (Harder & Wilson 1998; Porcher & Lande 2005; Johnston et al. 2009). Studies of geographic variation in mating system have been particularly revealing of the contribution of plant-pollinator interactions to adaptive differentiation in mating system (e.g. Barrett et al. 1989; Moeller 2006). Although there is now evidence that selfers often occur at higher latitudes than their sister outcrossers (Grossenbacher et al. 2015),

3

and that geographic variation in mating system is related to the predictability of pollination environments (e.g. Moeller 2006; Alonso et al. 2007), it remains unclear as to whether these observations for single species or small groups of taxa scale up to global biogeographic patterns.

Self-fertilization may be adaptive when it provides reproductive assurance in poor pollination environments (Darwin 1876; Lloyd 1992). This reproductive assurance hypothesis is often invoked because of the observation that selfing taxa are commonly associated with ephemeral and extreme environments where mate and pollinator availability are likely to be low (e.g. Baker's Law: Baker 1955, Evans et al. 2011; Cheptou 2012). Experimental studies within individual systems provide support for the hypothesis that selfing is favored by selection in chronically poor pollination environments (e.g. Fishman & Wyatt 1999; Moeller & Geber 2005; Eckert et al. 2006), particularly when pollinator availability fluctuates temporally (Kalisz et al. 2004), and in portions of geographic ranges where pollinators are scarce (Barrett et al. 1989; Moeller 2006). It has been suggested that the frequency of different plant mating systems varies among floras globally (reviewed in Lloyd 1980), but the extent to which such patterns exist and are driven by the biogeography of plant-pollinator interactions versus other factors remains unknown.

Previous studies propose the existence of a latitudinal gradient in plant mating systems where outcrossers are more frequent at low compared to high latitudes (reviewed in Lloyd 1980). For example, some authors have suggested that the frequency of autonomous selfing increases with latitude within temperate regions (Kerner 1896; Kevan 1972). Others have noted that mechanisms enforcing outcrossing are especially prevalent at low latitudes (Ashton 1969; Bawa 1974). The possibility of a latitudinal gradient in mating system has been attributed to two potential mechanisms. First, animal pollinators may be less diverse and unreliable at high latitudes such that selfing ensures reproduction (hereafter the "pollination environment hypothesis": Hagerup 1951; Baker 1966; see also Ollerton et al. 2006). This idea is consistent with general predictions that biotic interactions are stronger and more predictable at low compared to high latitudes (Dobzhansky 1950; MacArthur 1969; reviewed in Schemske et al.

2009). Support for this hypothesis is mixed: comparative studies suggest that plants have a greater diversity of pollination systems in the tropics (Ollerton et al. 2006); however, plant-pollinator interactions are not more specialized there (Ollerton & Cranmer 2002; Schleuning *et al.* 2012) and pollen limitation of reproduction is potentially greater in regions of high plant species diversity (Vamosi *et al.* 2006). A latitudinal gradient in biotic interactions could expose plant populations to uncertain pollination environments and strong pollen limitation more often at high than low latitudes. Overall, global patterns of mating system variation and their relation to the diversity and predictability of plant-pollinator interactions remain poorly understood.

Second, it has been suggested that the frequency of outcrossing may parallel geographic variation in life history and growth form (hereafter the "life form hypothesis": Grant 1975; Levin 1975; reviewed in Lloyd 1980). Here, outcrossing may be more common in the tropics because those regions are dominated by a higher diversity of long-lived trees and other woody plants, whereas selfing may be more common in temperate environments that tend to contain more short-lived species. However, at the highest latitudes long-lived taxa also dominate ecosystems, which may suggest that latitudinal patterns are not linear. Previous analyses have shown a strong association between outcrossing and life history and growth form, even when accounting for phylogeny (Barrett et al. 1996). This close association has suggested that selection for higher selfing rates may be strongest in short-lived herbaceous plants where there is a premium on reproducing quickly because the number of future opportunities to reproduce is limited. By contrast, selection for selfing may be weaker in long-lived and/or woody plants where there are multiple opportunities for future reproduction (Morgan et al. 1997) and where inbreeding depression is stronger (Morgan 2001; Scofield & Schulz 2006). If the potential for alleles that promote selfing to increase in populations is greater for annuals than long-lived perennials, broad-scale patterns of mating system variation may be more strongly attributed to geographic variation in life form than pollinator predictability. These two hypotheses are not mutually exclusive but represent different mechanisms and make contrasting predictions.

The objective of this study was to examine the global biogeography of mating system variation. The proportion of seeds produced through outcrossing (t_m) has been estimated in many lineages of seed plants, typically using molecular markers, and this information has been synthesized in several previous analyses (Schemske & Lande 1985; Barrett & Eckert 1990; Barrett et al. 1996; Vogler & Kalisz 2001; Goodwillie et al. 2005). Here, we updated the existing database and collected new information on biogeography and life form. We tested whether mating system variation is associated with latitude, biome, life history, and growth form, while accounting for phylogeny. The finding of a latitudinal gradient in outcrossing rates, where outcrossing tended to decline with latitude, motivated us to evaluate alternative hypotheses for the observed pattern. If biogeography explains significant variation in outcrossing even after accounting for variation in life form, then this is consistent with the pollination environment hypothesis. Alternatively, a finding that life form is the only predictor of variation in outcrossing, with biogeography failing to explain significant residual variation, would more strongly support the life form hypothesis.

METHODS

Data collection

We added published estimates of outcrossing rates for seed plants to an existing dataset that has accumulated through several decades and has been periodically analyzed (Schemske & Lande 1985; Barrett & Eckert 1990; Barrett et al. 1996; Vogler & Kalisz 2001; Goodwillie et al. 2005). To obtain new outcrossing rates, we searched databases (WEB OF SCIENCE and GOOGLE SCHOLAR) for "t_m", "outcrossing rate", and "selfing rate" through 2012. Only studies reporting outcrossing rates of natural plant populations were included. We treated subspecies as independent taxa. In total, our dataset included 624 studies of 492 taxa, with 197 studies and 147 taxa added since the last published analysis (Goodwillie et al. 2005). We recorded the outcrossing rate (t_m= the proportion of seeds outcrossed as opposed to self-fertilized) and whether the taxon exhibits self-incompatibility. Species were only designated as self-

incompatible when a genetic mechanism of self-incompatibility was reported in the publication or in other primary literature (often discerned from crossing experiments). Species were unclassified if there was not an explicit mention as to whether a taxon was self-incompatible (SI) or self-compatible (SC). We were able to classify 331 taxa (67.3%) as SI (84, 25%) or SC (248, 75). As with previous versions of the dataset, we averaged estimates of outcrossing rates within each study when multiple populations and years were included.

For each study, we recorded the latitude of the study population when reported or we used GOOGLE EARTH to find the approximate latitude of the site(s) based on the published localities. Where multiple populations were studied, the latitudes were averaged. We used the absolute value of latitude (i.e. degrees away from the equator). We classified each study site(s) into one of eight biomes (Fig. 2; Table S4) based upon a description of the biome reported in the paper and by consulting published maps of biomes (e.g. Olson et al. 2001). Biome variation, as we have classified it, describes a gradient in net primary productivity (Whittaker 1975). We classified latitude for 550 studies (88.1%) and biome for 471 studies (75.5%). Unclassified publications are those where it was not possible to determine a specific location for the study, and therefore we were unable to determine a latitude or biome with confidence. Categories of biome were represented as follows: desert and arid scrub (36 studies, 7.6%), temperate grassland (28, 5.9%), Mediterranean/chapparal (101, 21.4%), boreal forest and taiga (35, 7.4%), tropical savanna (19, 4.0%), temperate deciduous forest (132, 28.0%), tropical seasonal forest (37, 7.9%), and tropical rainforest (83, 17.6%).

We recorded life history and growth form from each publication, when reported, or via searches of the primary literature and regional floras. We classified life history for 619 studies (99.2%) and growth form for all studies. Life history was designated as annual (160 studies, 25.8%), biennial (9, 1.5%), semelparous perennial (9, 1.5%), iteroparous perennial (438, 70.5%), or varies (5, 0.8%), indicating that individuals may differ in life history. Growth form was designated as herb (302 studies, 48.4%), vine (16, 2.6%), shrub (82, 13.1%), tree (223, 35.7%), or varies (1, 0.2%), indicating that individuals may differ in growth form. Theoretical studies

support this categorization of life history and growth form (Morgan 2001; Scofield & Schultz 2006).

Statistical Analyses

For 97 taxa in the dataset, there were multiple published estimates of outcrossing rates (hereafter "outcrossing" or "t_m"). Averaging across those estimates can be problematic when (1) mating systems are highly diverged among populations examined in different studies and (2) life form and biogeographical variables are not consistent among populations examined in different studies (e.g. annual populations examined in one study, perennial populations in another). Both of these issues were present, although not common, in our dataset. Therefore, we conducted each analysis by randomly selecting one published study from each taxon in the dataset. For each analysis, we repeated this procedure 1000 times and generated a distribution of p-values for each independent variable.

To test for associations between outcrossing rate and biogeographic, growth form, and life history variables we first conducted a series of ordinary least squares (OLS) regressions that incorporated each potential predictor variable separately (latitude, biome, life history, and growth form), in pairwise combination, in triplet combination, and all four combined. We also explored whether a possible latitudinal relationship was curvilinear by incorporating a quadratic term into these models. We conducted the series of analyses with outcrossing rate and self-incompatibility (SI) as alternative response variables. For outcrossing, we conducted analyses with outcrossing logit transformed and untransformed; the results were qualitatively the same so we only report analyses of untransformed data. For SI, we used a binomial model with a probit link function (a logit link provided very similar results). For both t_m and SI, we performed separate analyses on data from (1) all seed plants, (2) only angiosperms (i.e. excluding gymnosperms), and (3) only animal-pollinated taxa. Variance inflation factors (VIF) indicated minimal collinearity among predictors and therefore we used all predictors in our analyses. Diagnostic plots of the residuals confirmed that there were no strong deviations from

homoscedasticity or high leverage points. Analyses were performed for each of the 1000 datasets generated by resampling; we report the mean and range of p-values for each variable. We considered variables to be significant when the 95% confidence interval of the distribution of p-values fell below 0.05. All analyses were conducted in R (R Core Team 2014).

Because life history and growth form proved to be the strongest predictors of variation in t_m, and because latitudinal patterns were not significant when these factors were included in models, we conducted analyses examining latitudinal and biome variation within life form categories. We had sufficient sample sizes to examine two life history categories (annuals, iteroparous perennials) and two growth form categories (herbs and trees). In all cases, we conducted univariate analyses taking latitude and biome separately, along with bivariate analyses including both predictors.

Finally, we also tested whether the frequency of life histories and growth forms varied with latitude using multinomial logistic regressions in the nnet package of R (version 7.3-8; Venables & Ripley 2002). To examine potential curvilinearity to these relationships across latitudes, we constructed models that included latitude as both linear and quadratic terms.

Analyses incorporating phylogeny

We conducted a second set of analyses that accounted for evolutionary history using phylogenetic comparative methods. First, we used PHYLOMATIC v.3 (R20120829.new tree; http://phylodiversity.net/phylomatic/; Webb & Donoghue 2005) to generate a phylogeny for all taxa in the dataset. Second, we quantified phylogenetic signal in outcrossing rate for 10 randomly selected datasets using Blomberg's K (Blomberg et al. 2003), which was conducted using the picante package v. 1.6-2 in R (Kembel et al. 2010). Values of K near zero indicate a lack of phylogenetic signal, with no correlation between phylogenetic relatedness and trait value. Values approaching one indicate trait evolution along a phylogeny consistent with Brownian motion. We tested for significant differences from zero by comparing the observed K to those calculated after randomly permuting taxa across the tips of the tree 9,999 times. Third, we

examined the relation between outcrossing and each potential predictor while accounting for phylogeny. We tested each variable separately and all combined. For the continuous predictor (latitude), we used a phylogenetic least squares (PGLS) regression. For the categorical predictors (biome, life history, growth form), we performed phylogenetic independent contrasts using the brunch algorithm, which allows for categorical variables to be considered nominal (Burt 1989). Both sets of analyses were conducted using the caper package v. 0.5.2 in R (Orme et al. 2013). For the combined analysis, using all four predictors, we conducted a PGLS regression. This approach requires categorical variables to be ordinal. Biome was ordered according to net ecosystem productivity (from low to high: desert and arid shrubland, temperate grassland, chaparral and shrubland, taiga, savanna, temperate deciduous forest, tropical seasonal forest, tropical rainforest; Reichle 1970; Whittaker 1975), life history was ordered by frequency of reproduction and then lifespan (annual, biennial, semelparous perennial, iteroparous perennial), and growth form was ordered by secondary growth and size (herbaceous, shrub, vine, tree). Because the resampling procedure used for OLS analyses revealed consistent outcomes among the 1000 resampled datasets, we conducted PGLS analyses on a randomly selected subset of 10 datasets.

We also examined latitudinal variation in outcrossing within the four families best represented in the dataset: Fabaceae (70 spp.), Pinaceae (42 spp.), Asteraceae (32 spp.), and Myrtaceae (30 spp). In each case, we conducted univariate analyses of latitude and multivariate analyses that included life history and growth form, if possible. The Pinaceae in our dataset were all iteroparous trees and the Myrtaceae were all perennials and only exhibited growth form variation (shrub vs. tree). It was not possible to include biome as a factor as there were many biome categories and too little replication of each within families.

Finally, we examined the relationship between SI and each predictor while accounting for phylogeny using phylogenetic logistic regressions (Ives & Garland 2010; Ho & Ane 2014). These analyses were conducted using the phyloglm function in the R package phylolm v. 2.0 (Ho

& Ane 2014). As with analyses of outcrossing rate, we conducted separate univariate analyses for each variable along with a multivariate analysis including all four independent variables.

RESULTS

Biogeographic and ecological correlates of outcrossing

A complete reanalysis of the distribution of outcrossing rates based upon our updated dataset is presented in Appendix S1. Those analyses provided three key insights: (1) the distribution was best fit by a Gaussian mixture model with four modes, and differed significantly from the first report of the database in 1985 (2) there was a significant increase in outcrossing (t_m) over the 57 years of study, and (3) t_m was greater in self-incompatible than self-compatible taxa (for more detail please see Appendix S1).

For all seed plants, ordinary least squares (OLS) regressions indicated that all four predictors were significantly correlated with outcrossing rate when each predictor was considered alone (Appendix S2, S3). Latitude and biome explained little variance in t_m (mean R^2 = 0.01 and 0.05, respectively; Appendix S2) whereas life history and growth form explained substantial variance in t_m (mean R^2 = 0.25 and 0.24, respectively). The 1000 resampled data sets consistently exhibited these patterns. Outcrossing rate declined significantly with increasing latitude (Fig. 1). We found that t_m was consistently lower in chapparal/Mediterranean than tropical rainforest biomes and in some resampled datasets t_m was also lower in chapparal/Mediterranean than taiga/boreal forest and savanna biomes (Fig. 2, Appendix S4). The association between life history, growth form, and t_m closely paralleled previous findings. Annuals had lower t_m than semelparous perennials or iteroparous perennials (Fig. 3, Appendix S4). Herbs had consistently lower t_m than trees; other contrasts were significantly different in some datasets but not others (Fig. 3, Appendix S4).

When all variables were considered together, only life history and growth form predicted t_m across seed plants (Appendices S2, S3). In fact, latitude and biome were never significant when models included either life history or growth form (Appendices S2, S3). This

was also true if we allowed for a curvilinear relationship (quadratic term) between t_m and latitude (results not shown). Analyses including only angiosperms or only animal-pollinated taxa were very similar to those presented here for all taxa (Appendices S2, S3, S4).

Given the importance of life history and growth form as predictors of outcrossing, we conducted separate analyses to ask whether latitudinal gradients or biome variation in outcrossing rate were apparent when life form categories were examined separately (e.g. latitudinal variation within annuals). Sample sizes were sufficient to examine annuals, iteroparous perennials, herbs, and trees. In no case did we detect a significant latitudinal gradient either when latitude was analyzed alone or in combination with biome (Appendix S5).

As expected, t_m was significantly greater for SI compared to SC taxa in our dataset (means: SI = 0.86, SC = 0.60; t = 7.4, P < 0.0001). Binomial models did not detect a significant association between the probability of self-incompatibility and latitude or biome from univariate analyses (Appendices S6, S7). As with analyses of outcrossing, we found a significant association between SI and both life history and growth form. The probability of SI was highest for iteroparous perennials and trees and lowest for annuals and herbs (Fig. 3, Appendix S8). Analyses incorporating all four predictor variables revealed similar patterns. Analyses including all taxa or just angiosperms revealed that life history was the only significant predictor, whereas analyses of just animal-pollinated species revealed that both life history and growth form were significant predictors (Appendix S8).

Life history and growth form varied significantly with latitude (Appendix S9). For life history, annuals increased in frequency relative to iteroparous perennials with latitude; however, this relationship was significantly curvilinear with annuals reaching peak frequency in the temperate zone (ca. 35-45 degrees; Appendix S9). For growth form, herbaceous plants increased in frequency relative to trees and shrubs with latitude. There was also significant curvilinearity to growth form variation primarily because shrubs were most common in subtropical to temperate zones (ca. 15-35 degrees).

Accounting for phylogeny in biogeographic analyses

We detected significant phylogenetic signal for outcrossing rate and latitude (mean K = 0.12 for both); although K values were not large, they differed significantly from zero (Appendix S10). When excluding gymnosperms, both variables still showed significant signal (mean K = 0.14 and 0.17, respectively; Appendix S11).

When accounting for phylogeny, we found associations between outcrossing and both life history and growth form; however, neither latitude nor biome was significant (Appendices S2, S4). Similarly, when all four factors were incorporated into a combined analysis, only life history and growth form explained significant variation in outcrossing (Appendix S2). These results were consistent across the 10 resampled datasets.

We also examined biogeographic patterns within the four clades represented by the most taxa (Fig. 4). Univariate analyses revealed a significant negative correlation of outcrossing with latitude in the Fabaceae (Fig. 5; R^2 = 0.18, $P < 0.001$) and Myrtaceae (R^2 = 0.21, P = 0.013) but not the Asteraceae (R^2 = -0.003, P = 0.373) or Pinaceae (R^2 = 0.05, P = 0.110). All Pinaceae in our dataset had outcrossing rates above 0.5, limiting the potential to detect any latitudinal pattern. For the Fabaceae and Asteraceae, we were able to conduct analyses including life history and growth form and in both cases found that life history was the strongest (and significant) predictor (Fabaceae: LH, $P < 0.001$; GF, P = 0.281; Asteraceae: LH, P = 0.004; GF, P = 0.864) and that latitude was not significant in either case (Fabaceae: P = 0.520; Asteraceae: P = 0.057). For the Myrtaceae, there was only variation in growth form (shrub vs. tree) and analyses indicated that both growth form and latitude were significant predictors (GF, P = 0.014; Latitude, P = 0.010). For the Pinaceae, all representatives were iteroparous perennial trees so no additional analyses were possible.

Life history and growth form predicted SI in univariate analyses; however as with t_m, latitude and biome were not significant predictors (Appendices S5, S6). Similarly, iteroparous perennials and trees were more likely to be SI than annuals and herbs. In the multivariate

analysis, only life history predicted SI. The results were consistent across the 10 resampled datasets.

Latitudinal gradients in biotic interactions have been of broad interest to ecologists because of their potential importance in driving patterns of biodiversity and phenotypic variation. In plants, there has long been speculation about latitudinal gradients in plant-pollinator interactions, and their consequences for mating system evolution. However, no previous study has attempted to synthesize across studies to test for global biogeographic patterns. Using a database of outcrossing estimates (t_m) from seed plants published over 57 years, we tested for biogeographic patterns and the potential underlying mechanisms generating those patterns. We detected a decline in t_m with increasing latitude and heterogeneity in t_m among biomes (e.g. tropical vs. Mediterranean). By contrast, the probability of SI did not vary with latitude and was only weakly and inconsistently associated with biome. The presence of a latitudinal gradient in t_m but not SI may occur because genetic SI is only one way in which plants can enforce outcrossing; floral mechanisms in SC taxa (dichogamy and herkogamy) can also minimize selfing. Our results are consistent with speculation in the literature that mating systems vary latitudinally (reviewed in Lloyd 1980) but suggest that the pattern is not particularly strong and the causative interpretation complex.

Mating systems are highly labile features of plants (Stebbins 1974) and our quantitative analyses of phylogenetic signal are consistent with this. Both t_m and SI had low but significant values of K (indicating little phylogenetic conservation). This signal across very wide phylogenetic breadth of seed plants may represent variation among clades in the probability that selfing taxa arise from outcrossers and/or persist over evolutionary time (Goldberg et al. 2010). Alternatively, signal may arise in our data set because taxa are so sparsely and unevenly distributed across the seed plant phylogeny. Some families were sampled out of proportion to the number of species in those families. For example, the dataset contains 70 species of

14

Fabaceae (albeit a large family) whereas many others contain single species (32 of 91 families). Although phylogenetic signal was weak, accounting for it in biogeographic analyses affected the significance of latitude and biome as predictors of t_m. We found significant phylogenetic signal for latitude, suggesting that some clades are non-randomly distributed across latitude. If mating system is conserved to some extent in those clades, the association between mating system and latitude could be explained in part by phylogenetic patterns. For example, gymnosperms are more diverse at mid compared to low latitudes (peak diversity between ca. 30-40° N and S of equator: Farjon & Filer 2013) unlike angiosperms, which increase in diversity continuously up to the equator (Kreft & Jetz 2007). Since gymnosperms are typically highly outcrossing, we expected that this geographic pattern could contribute to the overall latitudinal pattern that we observed. However, when gymnosperms were excluded from the analysis the biogeographic patterns remained qualitatively similar. Clades other than gymnosperms must also be influencing the latitudinal gradient observed in OLS analyses.

We tested the prediction that biomes with high productivity differ in mating system because of the reliability of plant-pollinator interactions (Ollerton et al. 2006) but found only weak support. In univariate analyses, outcrossing was higher among taxa from tropical rainforests than chapparal/Mediterannean habitats. This pattern was also apparent in multivariate analyses of SI that accounted for phylogeny. However, tropical rainforest taxa exhibited a similar distribution of t_m compared to those from boreal forest/taiga and desert/shrubland. Despite some significant contrasts among individual biomes, t_m was highly variable within and mean t_m was rather consistent across biomes.

Pollinator vs. life form hypothesis

The two hypotheses that we explored are not necessarily mutually exclusive but lead to alternative predictions about the importance of life form variation, alone, in explaining patterns of mating system variation. It is possible that both biogeography (latitude and biome) and life form (life history and growth form) jointly determine biogeographic patterns of t_m. However, in

our analyses only the latter predicted outcrossing. Life form variables remained significant predictors in multivariate analyses that accounted for phylogeny. We found very similar results for SI. Taken together, these findings support the premise that latitudinal variation in outcrossing is due in large part to the biogeographical variation in plant life history and growth form evident in our dataset (Appendix S9) and previous biogeographic surveys (Grant 1975; Levin 1975; reviewed in Lloyd 1980). It is important to note that while there is a significant linear component to the latitudinal gradient in life history and growth form (greater frequency of iteroparous perennials and trees in the tropics), the relationship is also significantly curvilinear because the highest frequency of annuals and shrubs occurs at mid-latitudes (Appendix S9). Interestingly, although there is some suggestion of curvilinearity to the latitudinal pattern of outcrossing, a quadratic term was not significant in our analyses.

Although biotic interactions have been hypothesized to be stronger and more specialized at low latitudes (Schemske et al. 2009), recent analyses of plant-pollinator interaction networks have indicated that these interactions may be less specialized in the tropics than originally thought (Olesen & Jordano 2002; Ollerton & Cranmer 2002; Schleuning et al. 2012). There is also evidence that the diversity of pollination systems is greater in the tropics (Ollerton et al. 2006); however, it remains unclear whether this broad functional variation in pollination systems contributes in any meaningful way to the predictability of pollinator service and the evolution of mating system strategies. Finally, a synthetic analysis has suggested that pollen limitation of reproduction is strongest in areas where plant species richness is highest, potentially owing to greater competition for pollinators in those regions (Vamosi et al. 2006). This pattern is especially pronounced for self-incompatible species and trees. An alternative explanation is that the density and abundance of mates is reduced in highly diverse areas (high alpha diversity is associated with low density of compatible mates), which could similarly result in pollen limitation. Greater plant competition via pollinators, mate limitation due to low densities, and lower specificity of plant-pollinator relations in the tropics may all contribute to the shallow latitudinal gradient in mating systems.

Our work complements a recent comparative analysis which suggested that selfers occur at somewhat higher latitudes and have larger geographic ranges than their sister outcrossers (Grossenbacher et al. 2015). It is also consistent with the finding that asexuals occur at higher latitudes than sexuals in two sections of the genus *Oenothera* (Johnson et al. 2010). The complementary results of these studies suggest that the divergence of selfers from outcrossers (and asexuals from sexuals) may occur more frequently at species' northern range margins; however the causes of these patterns remain poorly understood. For individual systems, the bias toward selfers at higher latitudes could reflect selection for selfing that occurred during range expansion, such as during postglacial migration. Populations at the leading edge of an expanding range may suffer from a shortage of both compatible mates and effective pollinators (Hargreaves & Eckert 2013). This sort of phenomenon could influence comparative analyses of mating system variation as most taxa studied occur in temperate and boreal regions that were strongly affected by glacial cycles.

Considerations and Conclusions

Estimates of outcrossing (t_m) in natural plant populations have been generated for > 57 years by many researchers for a variety of reasons and hence we still lack a truly random sample of taxa across biogeography or life forms. In previous datasets, sampling biases against species with extreme outcrossing values (both high and low t_m) were identified due to the exclusion of highly selfing species, species with SI, and dioecious species (Igic & Kohn 2006; Goodwillie et al. 2010). In the current version of the dataset, we found a considerably higher representation of highly outcrossing, SI, and mixed mating taxa compared to previous versions (Appendix S1). For example, the fraction of taxa that are SI in the current dataset is nearly twice as great as reported previously (25 vs. 13%; Igic and Kohn 2006); nevertheless, the current dataset is not free from bias. The sparse sampling of taxa across the seed plant phylogeny also makes it challenging to resolve phylogenetic signal and its potential contribution to biogeographic patterns. Nonetheless, the major patterns observed in our study are not likely to

be affected by lingering biases in the dataset given that those biases would have to occur differently across latitudes or with respect to life form to have an important influence.

Overall, our work has suggested that there is a latitudinal gradient in mating system (as quantified by outcrossing) that is influenced to some extent by phylogenetic signal. Our analyses indicate that mating system variation is more strongly related to life form variation than geography, which casts some doubt on the idea that pollination environments are the primary driver of the broader latitudinal gradient. Instead they suggest that for shorter-lived organisms, there is a premium on reproducing, which may often select for selfing due to its reproductive assurance value (e.g. Herlihy & Eckert 2002; Kalisz et al. 2004; Moeller & Geber 2005; Eckert et al. 2006). Resolving the mechanisms driving the latitudinal gradient described here will require detailed examination of mating system variation among populations within individual species along with plant-pollinator interactions (Moeller 2006; Hargreaves et al. 2015). To date, few studies have tackled this problem, but it is likely that species with broad latitudinal distributions spanning biomes may be the most profitable to examine first.

REFERENCES

Alonso, C., Mutikainen, P., & Herrera, C.M. (2007). Ecological context of breeding system variation: sex, size, and pollination in a (predominantly) gynodioecious shrub. *Ann Bot.,* **100**, 1547—1556.

Ashton, P.S. (1969). Speciation among tropical forest trees. *Biol. J. Linn. Soc.,* **1**, 155—196.

Baker, H.G. (1955). Self-compatibility and establishment after 'long distance' dispersal. *Evolution,* **9**, 347–348.

Baker, H.G. (1966) The evolution, functioning, and breakdown of heteromorphic incompatibility systems. *Evolution*, **20**, 349—368.

Barrett, S.C.H. (2003). Mating strategies in flowering plants: the outcrossing-selfing paradigm and beyond. *Phil. Trans. R. Soc. Lond. B*, **358**, 991—1004.

Barrett, S.C.H. & Harder, L.D. (1996). Ecology and evolution of plant mating. *Trends in Ecology & Evolution*, **11**, 73-79.

Barrett, S.C.H. & Eckert, C.G. (1990). Variation and evolution of mating systems in seed plants. In: Biological approaches and evolutionary trends in plants (ed. Kawano, S). Academic Press Limited.

Barrett, S. C. H., M. T. Morgan, & B. C. Husband (1989). The dissolution of a complex genetic polymorphism: the evolution of self-fertilization in tristylous *Eichornia paniculata* (Pontederiaceae). *Evolution* **43**, 1398–1416.

Barrett, S.C.H., Harder, L.D. & Worley, A.C. (1996). The comparative biology of pollination and mating in flowering plants. *Phil. Trans. R. Soc. Lond. B*, **351**, 1271—1280.

Bawa, K.S. (1974). Breeding systems of tree species of a lowland tropical community. *Evolution*, **28**, 85—92.

Blomberg, S.P., Garland, T., & Ives, A.R. (2003). Testing for phylogenetic signal in comparative data: Behavioral traits are more labile. *Evolution,* **57**, 717—745.

Burt, A. (1989). Comparative methods using phylogenetically independent contrasts. *Oxford Surveys in Evolutionary Biology*, **6**, 33–53.

Cheptou, P.-O. (2012). Clarifying Baker's law. *Ann. Bot.*, **109**, 633—641.

Darwin, C.R. (1876). The effects of cross and self fertilisation in the vegetable kingdom. John Murray, London.

Dobzhanksy, T. (1950). Evolution in the tropics. *Am. Sci.,* **38**, 209–21.

Eckert, C.G., Samis, K.E., & Dart, S. (2006). Reproductive assurance and the evolution of uniparental reproduction in flowering plants. In: Ecology and evolution of flowers (ed. Harder, L.D. and Barrett, S.C.H.). Oxford University Press.

Evans, M. E., Hearn, D. J., Theiss, K. E., Cranston, K., Holsinger, K. E., & Donoghue, M. J. (2011). Extreme environments select for reproductive assurance: evidence from evening primroses (Oenothera). *New Phytologist*, **191**, 555-563.

Farjon, A., & Filer, D. (2013). *An atlas of the world's conifers: an analysis of their distribution, biogeography, diversity and conservation status*. Brill.

Fishman, L., & Wyatt, R. (1999). Pollinator-mediated competition, reproductive character displacement, and the evolution of selfing in *Arenaria uniflora* (Caryophyllaceae). *Evolution*, **53**, 1723-1733.

Goldberg, E.E. & Igic, B. (2012). Tempo and mode in plant breeding system evolution. *Evolution*, **66**, 3701-3709.

Goodwillie, C., Kalisz, S., & Eckert, C.G. (2005). The evolutionary enigma of mixed mating systems in plants: Occurrence, theoretical explanations, and empirical evidence. *Annu. Rev. Ecol. Evol. Syst.* **36**, 47–79.

Goodwillie, C., Sargent, R.D., Eckert, C.G., Elle, E., Geber, M.A., Johnston, *et al.* (2010). Correlated evolution of mating system and floral display traits in flowering plants and its implications for the distribution of mating system variation. *New Phytologist*, **185**, 311-321.

Grant, V. (1975). *Genetics of Flowering Plants*. New York, Columbia University Press.

Grossenbacher, D., Briscoe Runquist, R., Goldberg, E. E., & Brandvain, Y. (2015). Geographic range size is predicted by plant mating system. *Ecology Letters*, 18, 706-713.

Hagerup, O. (1951). Poillination in the Faroes—in spite of rain and poverty in insects. *D. Kgl. danske Vidensk. Selsk. Biol. Medd.* 18, 1–47.

Hargreaves, A.L., & Eckert, C.G. (2013). Evolution of dispersal and mating systems along geographic gradients: implications for shifting ranges. *Funct. Ecol*. **28**, 5–21.

Hargreaves, A.L., Weiner, J.L. & Eckert, C.G. (2015). High-elevation range limit of an annual herb is neither caused nor reinforced by declining pollinator service. *J. Ecol*. **103**, 572–584.

Harder, L.D., & Barrett, S.C.H. (1995). Mating cost of large floral displays in hermaphrodite plants. *Nature*, 373, 512-515.

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

Ho, L.S.T., & Ane, C. (2014). A linear-time algorithm for Gaussian and non-Gaussian trait evolution models. *Syst. Biol.*, **63**, 397—408.

Igic, B., & Kohn, J.R. (2006). The distribution of plant mating systems: study bias against obligately outcrossing species. *Evolution*, **60**, 1098—1103.

Ives, A.R., & Garland, T. (2010). Phylogenetic logistic regression for binary dependent variables. *Syst. Biol.*, **59**, 9—26.

Johnson, M.T.J., Smith, S., & Rausher, M. (2010). The effect of plant sex on range distributions and allocation to reproduction. *New Phytologist,* **186**: 769–779.

Johnston, M.O., Porcher, E., Cheptou, P.-O., Eckert, C.G., Elle, E., Geber, M.A. *et al*. 2009. Correlations among fertility components can maintain mixed mating in plants. *American Naturalist,* **173**: 1-11.

Kalisz S., Vogler D.W., & Hanley, K.M. (2004). Context-dependant autonomous self-fertilisation yields reproductive assurance and mixed mating. *Nature,* **430**, 884–886.

Kembel, S.W., Cowan, P.D., Helmus, M.R., Cornwell, W.K., Morlon, H., Ackerly, D.D., *et al.* (2010). Picante: R tools for integrating phylogenies and ecology. *Bioinformatics*, **26**, 1463-1464.

Kerner, A. (1896). *The Natural History of Plants*. Vol. 2 London, Gresham Publishing Co.

Kevan, P.G. (1972). Insect pollination of high arctic flowers. *J. Ecol.*, **60**, 831—847.

Kreft, H., & Jetz, W. (2007). Global patterns and determinants of vascular plant diversity. *Proceedings of the National Academy of Sciences*, **10**4, 5925-5930.

Lande, R., & Schemske, D.W. (1985). The evolution of self-fertilization and inbreeding

   depression in plants. 1. Genetic models. *Evolution,* **39**, 24–40.

Levin, D.A. (1975). Pest pressure and recombination systems in plants. *Am. Nat.*, **109**, 437—451.

Lloyd, D.G. (1980). Demographic factors and mating patterns in angiosperms. In: Demography

   and evolution in plant populations (ed. Solbrig, OT). University of California Press Berkeley,

   CA, U.S.A., pp. 67-88.

Lloyd, D. G. (1992). Self– and cross–fertilization in plants. II. The selection of self–fertilization.

   Int. J. Plant Sci., **153**, 370–380.

MacArthur, R.H. (1969). Patterns of communities in the tropics. *Biol*. *J. Linn. Soc.,* **1**, 19–30.

Mittelbach, G.G, Schemske, D.W., Cornell, H.V., Allen, A.P., Brown, J.M., Bush, M.B., *et al.*

   (2007). Evolution and the latitudinal diversity gradient: speciation, extinction and

   biogeography. *Ecol. Lett.,* **10**, 315–31

Moeller, D.A. (2006). Geographic structure of pollinator communities, reproductive assurance,

   and the evolution of self–pollination. *Ecology,* **87**, 1510–1522.

Moeller, D.A., & Geber, M.A. (2005). Ecological context of the evolution of self–pollination in

   Clarkia xantiana: population size, plant communities, and reproductive assurance.

   Evolution, **59**, 786–799.

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., Schoen, D. J., & Bataillon, T. M. (1997). The evolution of self‐fertilization in

   perennials. *American Naturalist*, **150**, 618-638.

Ollerton, J., & Cranmer, L. (2002). Latitudinal trends in plant-pollinator interactions: are

   tropical plants more specialised? *Oikos*, **98**, 340–350.

Ollerton, J., Johnson, S.D., & Hingston, A.B. (2006) Geographical variation in diversity and

   specificity of pollination systems. Plant-Pollinator Interactions: from Specialization to

Generalization (Eds N.M. Waser & J. Ollerton), pp. 283—308. University of Chicago Press, Chicago,USA.

Olesen, J. M. & Jordano, P. (2002). Geographic patterns in plant-pollinator mutualistic networks. *Ecology*, **83**, 2416–2424.

Orme, D., Freckleton, R., Thomas, G., Petzoldt, T., Fritz, S., Isaac, N., *et al.* 2013. Package 'Caper': Comparative Analyses of Phylogenetics and Evolution in R. https://cran.r-project.org/web/packages/caper/index.html.

Porcher, E., & R. Lande. 2005. The evolution of self-fertilization and inbreeding depression under pollen discounting and pollen limitation. J. Evol. Biol. 18:497–508.

R Core Team. 2014. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Reichle, D.E. (1970). Analysis of temperate forest ecosystems. Springer-Verlag, Berlin.

Schemske, D.W., & Lande, R. (1985). The evolution of self-fertilization and inbreeding depression in plants. 2. Empirical observations. *Evolution,* **39**, 41–52.

Schemske, D.W., Mittelbach, G.G., Cornell, H.V., Sobel, J.M. & Roy, K. (2009). Is there a latitudinal gradient in the importance of biotic interactions? *Annu. Rev. Ecol. Evol. Syst.*, **40**, 245–269.

Schleuning, M., Fründ, J., Klein, A.-M., Abrahamczyk, S., Alarcón, R., Albrecht, M., *et al*. (2012). Specialization of mutualistic interaction networks decreases towards tropical latitudes. *Curr. Biol.*, **22**, 1925–1931.

Scofield, D.G., & Schulz, S.T. (2006). Mitosis, stature and evolution of plant mating systems: low-Φ and high-Φ plants. *Proc R. Soc. Lond. B.* **273**, 275—282.

Stebbins, G.L. (1974). Flowering plants: Evolution above the species level. Harvard University Press, Cambridge, MA.

Vamosi, J.C., Knight, T.M., Steets, J.A., Mazer, S.J., Burd, M., & Ashman, T.-L. (2006). Pollination decays in biodiversity hotspots. *Proc. Natl. Acad. Sci. USA* **103**, 956–61.

Venables, W. N. & Ripley, B.D. (2002). Modern applied statistics with S. Fourth Edition. Springer, New York

Vogler, D.W. & Kalisz, S. (2001). Sex among the flowers: the distribution of plant mating systems. *Evolution*, **55**, 202-204.

Webb, C.O. & Donoghue, M.J. (2005). Phylomatic: tree assembly for applied phylogenetics. *Mol. Ecol. Notes*, **5**, 181-183.

Whittaker, R.H. (1975). Communities and ecosystems. New York. Macmillan.

Figure 1. The upper two panels show the relationship between the probability of self-incompatibility and latitude and the relationship between outcrossing rate (t_m) and latitude. P-values are shown for univariate OLS analyses. The lower panel shows histograms of outcrossing rate at 10 degree intervals of latitude. For all panels, latitude is expressed as the absolute value of latitude, measuring the distance from the equator in both hemispheres.

Figure 2. Box plots of outcrossing rates for each of the eight biome categories used in our analyses. Biomes are arranged ordinally from left to right according to net ecosystem primary productivity. The number of species represented by each biome is shown above the box plots. For each boxplot, the middle line is the median and the top and bottom edges represent the 25th and 75th percentiles, respectively. The whiskers extend to 1.5 interquartile ranges from the median or to the maximum or minimum points, whichever is closer. The p-value is shown from the univariate OLS analysis.

Figure 3. The upper panels show box plots of outcrossing in relation to life history and growth form. The lower panels show means and standard errors for the probability of self-incompatibility for life history and growth form categories estimated from logistic regression models. The number of species represented by each category is shown above each plot. P-values are shown for the univariate OLS analyses.

Figure 4. Latitudinal variation in outcrossing rate for four seed plant families represented by the largest number of taxa. The predictor is the absolute latitude, i.e. the degrees from the equator. Slopes are shown for the two families for which the relation was significant. P-values are shown for the univariate OLS regressions.