



# Global patterns of population genetic differentiation in seed plants

Diana Gamba | Nathan Muchhala

Biology Department, University of Missouri,  
Saint Louis, MO, USA

**Correspondence**

Diana Gamba, One University Boulevard,  
223R Research Hall, Saint Louis, MO 63121,  
USA.

Email: dlgtk5@mail.umsl.edu

**Funding information**

Whitney R. Harris World Ecology Center,  
University of Missouri-St. Louis

**Abstract**

Evaluating the factors that drive patterns of population differentiation in plants is critical for understanding several biological processes such as local adaptation and incipient speciation. Previous studies have given conflicting results regarding the significance of pollination mode, seed dispersal mode, mating system, growth form and latitudinal region in shaping patterns of genetic structure, as estimated by  $F_{ST}$  values, and no study to date has tested their relative importance together across a broad scale. Here, we assembled a 337-species data set for seed plants from publications with data on  $F_{ST}$  from nuclear markers and species traits, including variables pertaining to the sampling scheme of each study. We used species traits, while accounting for sampling variables, to perform phylogenetic multiple regressions. Results demonstrated that  $F_{ST}$  values were higher for tropical, mixed-mating, non-woody species pollinated by small insects, indicating greater population differentiation, and lower for temperate, outcrossing trees pollinated by wind. Among the factors we tested, latitudinal region explained the largest portion of variance, followed by pollination mode, mating system and growth form, while seed dispersal mode did not significantly relate to  $F_{ST}$ . Our analyses provide the most robust and comprehensive evaluation to date of the main ecological factors predicted to drive population differentiation in seed plants, with important implications for understanding the basis of their genetic divergence. Our study supports previous findings showing greater population differentiation in tropical regions and is the first that we are aware of to robustly demonstrate greater population differentiation in species pollinated by small insects.

**KEYWORDS**

$F_{ST}$ , latitudinal region, life history traits, pollination mode, population differentiation, population genetic structure

## 1 | INTRODUCTION

Understanding the factors that drive patterns of genetic variation among plant populations is central in biology because genetic diversity is the raw material on which evolution acts. Quantifying population differentiation, which is most frequently done using the fixation index  $F_{ST}$  (Wright, 1951; see Holsinger & Weir, 2009; Meirmans &

Hedrick, 2011 for a review of  $F_{ST}$  and related metrics), is important for understanding the first stages of allopatric speciation (Harvey, Singhal, & Rabosky, 2019; Templeton, 1981), as well as the basis of local adaptation (Leimu & Fischer, 2008; Linhart & Grant, 1996), and provides critical information for conservation genetics (Ellstrand, 1992; Ellstrand & Elam, 1993; Kramer & Havens, 2009). Life history traits are expected to influence population genetic

structure in seed plants (Duminil et al., 2007; Hamrick & Godt, 1996; Loveless & Hamrick, 1984). However, previous studies have given conflicting results as to the importance of specific traits, such as pollination mode, seed dispersal mode, mating system, and growth form (e.g. Duminil et al., 2007; Hamrick & Godt, 1996), and only one study has compared patterns of  $F_{ST}$  variation between latitudinal regions (Dick, Hardy, Jones, & Petit, 2008). Furthermore, previous studies analysed each trait separately, and thus, little is known about the relative importance of these factors. Below, we discuss prior evidence for each of these factors in turn and then detail our approach to test them all together in a single analysis that also accounts for phylogenetic relatedness.

Pollination mode is predicted to affect population genetic structure, because pollen dispersal is critical to moving alleles between plant populations. Previous reviews have lumped different pollination mutualists together as animal pollination and compared them to wind pollination (Duminil et al., 2007; Hamrick, Godt, & Sherman-Broyles, 1992), suggesting that wind tends to reduce genetic structure. Although the idea has not been tested on a broad scale, it has long been thought that different types of animal pollinators should also lead to differences in population genetic structure due to differences in their movement patterns and pollen carry-over capacity (Castellanos, Wilson, & Thomson, 2003; Loveless & Hamrick, 1984). In fact, direct measures of pollen dispersal reveal that volant vertebrates and large bees transport pollen farther than small insects (Dick et al., 2008). Given these results, we predict that small insects restrict gene flow among plant populations and increase  $F_{ST}$ , compared to large insects, vertebrates or wind.

Seed dispersal mode is also expected to influence plant population genetic structure because, like pollination mode, it directly affects the movement of alleles and thus gene flow among populations. Strong evidence suggests that limited dispersal increases fine-scale spatial genetic structure in plants (Gelmi-Candusso, Heymann, & Heer, 2017) and in other organisms (Aguillon et al., 2017), which in consequence might scale up and lead to greater population genetic structure (Hamrick & Trapnell, 2011). In fact, reviews of the allozyme literature suggest that seed dispersal by wind and ectozoochory (i.e. externally attached to animals) results in lower  $F_{ST}$  than dispersal by gravity and endozoochory (i.e. ingested by animals) due to greater gene flow among populations from long distance dispersal events (Hamrick & Godt, 1996; Hamrick, Murawski, & Nason, 1993). However, Duminil et al. (2007) found that dispersal mode was not a significant predictor of  $F_{ST}$ . The lack of consistency among studies encourages further work with larger sample sizes to fully understand the role of seed dispersal mechanisms on population genetic structure.

Unlike pollination and seed dispersal modes, the effect of mating system on plant population genetic structure has been well-established in previous broad-scale studies (Duminil et al., 2007; Loveless & Hamrick, 1984), which suggest that it is the most important predictor of  $F_{ST}$  variation. Mating system affects inbreeding, which lowers within-population variation, inflating between-population  $F_{ST}$  values (Charlesworth, 2003). Duminil, Hardy, and Petit (2009) found that

the outcrossing rate and the inbreeding coefficient, which measures biparental inbreeding and selfing, are both significant predictors of  $F_{ST}$  in seed plants. Both selfing and inbreeding increase inbreeding depression and induce purging of deleterious alleles, reducing effective population size and increasing genetic drift, which can ultimately lead to fixation of different alleles in different populations (Angeloni, Ouborg, & Leimu, 2011; Wright, Ness, Foxe, & Barrett, 2008). In contrast, outcrossing increases gene flow within populations, potentially intensifying pollen-mediated gene flow among populations, which counteracts genetic drift and thus decreases population genetic structure (Duminil et al., 2009; Ellstrand, 2014).

Growth form is also an important predictor of population genetic structure. Broad-scale analyses (Duminil et al., 2009; Hamrick et al., 1992) have found strong associations between growth form and  $F_{ST}$ , with woody plants tending to have lower  $F_{ST}$  than herbaceous plants. The mechanism that causes this association is unclear, however, and may actually be driven by correlations between growth form and other factors. For example, Duminil et al. (2009) found that growth form only affects  $F_{ST}$  indirectly, through its influence on outcrossing rate ( $t_m$ ) and inbreeding coefficient ( $F_{IS}$ ); woody growth form is associated with greater  $t_m$  and lower  $F_{IS}$ . However, Hamrick and Godt (1996) reviewed the allozyme literature for over 300 species and found that when considering outcrossing plants, woody plants show lower levels of  $F_{ST}$  than herbs, which suggests that growth form directly affects gene flow among populations, decreasing population genetic structure. This could be because in trees greater geographic distance is presumably required for genetic differences to be detected among populations than in herbs, given that trees are larger than herbs. Thus, when considered at similar geographic scales, we predict that herbs have populations with greater genetic differentiation than trees.

Finally, the latitudinal region in which a plant occurs could also affect its population genetic structure due to differences among regions in spatial and climatic landscapes. In general, geographic heterogeneity and seasonal asynchrony over short distances are considerably higher in the tropics than in the temperate zones (Esquerré, Brennan, Catullo, Torres-Pérez, & Keogh, 2019; Ricklefs, 1977; Stein, Gerstner, & Krefth, 2014), which may act to disrupt mating among conspecific subpopulations and thus limit gene flow (Martin, Bonier, Moore, & Tewksbury, 2009; Quintero, González-Caro, Zalamea, & Cadena, 2014). For example, the “asynchrony of seasons hypothesis” (Martin et al., 2009) observes that seasons in temperate zones are determined primarily by temperature regimes, which are relatively constant over large geographic distances, while seasons in the tropics are primarily determined by rainfall, which can vary greatly over short distances. This leads to a geographic mosaic of seasonality in the tropics, that is high climatic asynchrony between nearby sites. Because organisms typically time their reproductive cycles to seasons, this hypothesis proposes that such climatic asynchrony leads to reproductive asynchrony, which in turn disrupts gene flow among subpopulations and promotes population genetic divergence in tropical regions. Additionally, genetic drift could have a more prominent role in the tropics than in the temperate zones, due to the fact that

most species in the tropics occur at low population densities and thus should have lower effective population sizes than in temperate zones (Dick et al., 2008; ter Steege et al., 2013). In fact, although their sample size was limited and phylogenetic autocorrelation was not accounted for, Dick et al. (2008) found that tropical trees have on average higher  $F_{ST}$  values than temperate trees. Given all of the above effects, we predict that  $F_{ST}$  is higher in the tropics than in the temperate zones.

Previous studies have not included all of the aforementioned factors together when modelling patterns of population genetic structure in seed plants (Duminil et al., 2007; Hamrick & Godt, 1996; Hamrick et al., 1992; Loveless & Hamrick, 1984; Nybom & Bartish, 2000). Furthermore, the most thorough study of  $F_{ST}$  in seed plants was over a decade ago (Duminil et al., 2007) and thus could not take advantage of the wealth of population genetic studies published since then. Here, we reviewed publications to assemble a 337-species database of seed plants with the goal of evaluating the factors predicted to best explain variation in plant population genetic structure. We focused on studies that used nuclear markers because their genetic structure should reflect both pollen and seed movement (due to biparental inheritance), unlike chloroplast markers, which only reflect seed movement in angiosperms (due to maternal inheritance) (McCauley, 1994), and pollen movement in gymnosperms (due to paternal inheritance) (Neale & Sederoff, 1988). We examined five ecological factors, including pollination mode, seed dispersal mode, mating system, growth form and latitudinal region, while controlling for phylogenetic autocorrelation. We also accounted for variables pertaining to the sampling scheme that have been shown to affect  $F_{ST}$  values for plants (Nybom & Bartish, 2000) and other systems (Blasco-Costa & Poulin, 2013; Pascual, Rives, Schunter, & Macpherson, 2017; Riginos, Douglas, Jin, Shanahan, & Tremblay, 2011), namely genotyping technique, distance between populations, and sample size. Using phylogenetic multiple regressions, we asked: (Q1) What set of life history traits promote population divergence in seed plants? (Q2) Do patterns of variation in  $F_{ST}$  differ between latitudinal regions? (Q3) What is the relative importance of these factors in explaining variation in  $F_{ST}$ ?

## 2 | MATERIALS AND METHODS

### 2.1 | Data collection

We constructed an  $F_{ST}$  data set through a systematic search in Google Scholar (key words: “plant” AND—the following words, each in a separate search—“genetic structure”, “population differentiation”, “population genetics”, “genetic diversity”, “population gene flow”) for articles published up until June 2018. The search yielded thousands of studies, which we reduced to 356 peer-reviewed publications on seed plants that determined population genetic structure ( $F_{ST}$ ) based on nuclear markers. When multiple studies reported  $F_{ST}$  values for the same species, we recorded the  $F_{ST}$  from the study with the largest geographic range, as this may better represent the

genetic diversity found in the species (Cavers et al., 2005). By this criterion, we compiled a data set that included 337 unique species. We extracted information for the predictor variables directly from the publications and infrequently complemented this, where necessary, with information from peer-reviewed literature on the studied species (see Appendix S1 and Table S1). Predictor variables were included in multiple regressions to explain variation in  $F_{ST}$  values (see section  $F_{ST}$  models). We also included three factors that pertained to the sampling scheme of each study and that can potentially affect  $F_{ST}$  (Nybom, 2004; Nybom & Bartish, 2000): genetic marker used, maximum distance between populations, and mean sample size per population. We used them to construct a null model to be compared against models with our factors of interest. Factors of interest consisted of five categorical variables with 2–4 levels: mating system (outcrossing, mixed-mating), growth form (non-woody, shrub, tree), pollination mode (large insects, small insects, vertebrates, wind), seed dispersal mode (animal, gravity, wind) and latitudinal region (tropics, subtropics, temperate). Below, we explain the  $F_{ST}$  estimates and all eight factors used in this study in greater detail.

### 2.2 | $F_{ST}$ estimates

We collected  $F_{ST}$  and  $F_{ST}$  analogs as measures of genetic differentiation (Holsinger & Weir, 2009; see Meirmans & Hedrick, 2011 for a review of  $F_{ST}$  and its analogs) which we collectively refer to  $F_{ST}$  throughout this paper. Assuming an island model of migration–drift equilibrium, Wright (1951) developed a theoretical framework for studying the gene frequency variation among subpopulations through the fixation indices, that is  $F$ -statistics (further developed in Wright, 1965, 1978). In this model,  $F_{ST}$  is the degree of gene differentiation among subpopulations for genes that have only two alleles. Nei (1973) expanded the model for polymorphic genes and proposed  $G_{ST}$  as a measure of the gene diversity partitioned among subpopulations, relative to the total gene diversity of the population. Subsequently, Weir and Cockerham (1984) proposed a standard measure (i.e. that accounts for genetic sampling) of genetic structure  $\theta$  based on Wright (1951). The statistic  $\theta$  is estimated per and across loci and represents the correlation of genes, or coancestry, among individuals in a given population. Excoffier, Smouse, and Quattro (1992) proposed analysis of molecular variance (AMOVA) and corresponding statistic  $\phi_{ST}$ ; the proportion of genetic diversity partitioned among populations, also accounting for genetic sampling. Finally, Hedrick (2005) proposed a standardized measure of population differentiation,  $G'_{ST}$ , which accounts for the level of heterozygosity of the marker used for genotyping individuals ( $G'_{ST} = G_{ST} / G_{STmax}$ ).

The most common statistic in our data set was AMOVA  $\phi_{ST}$  (116 studies), followed by  $\theta$  (111 studies). When  $\theta$  was reported per loci, we took the mean across loci as the global  $F_{ST}$  for that species. Some studies reported both  $\theta$  and  $\phi_{ST}$ , in which case we used  $\phi_{ST}$  as it likely better represents genetic structure among populations (Hey & Pinho, 2012). The statistics  $\theta$  and  $\phi_{ST}$  were, however, frequently

almost equivalent. Another common measure was  $G_{ST}$  (49 studies); when reported for multiple pairs of populations, we used the mean across all pairs. Other studies (49) reported  $F_{ST}$  measures using Wright's  $F_{ST}$  (1965; 1978). A few studies (12) reported  $G'_{ST}$ . It was not possible to back-transform  $G'_{ST}$  to  $G_{ST}$  because these studies did not report the maximum possible  $G_{ST}$  in their data (Hahn, Michalski, Fischer, & Durka, 2016). Even though  $G'_{ST}$  potentially yields a higher value than  $G_{ST}$  (or  $\theta$  and  $\phi_{ST}$ ) based on the same data (Hedrick, 2005; Meirmans & Hedrick, 2011), we opted to include  $G'_{ST}$  values, reasoning that any trend of variation in population genetic structure due to the variables here tested should still be present.

### 2.3 | Molecular markers

$F_{ST}$  values can be strongly affected by the genotyping technique implemented (Meirmans & Hedrick, 2011; Nybom, 2004; Nybom & Bartish, 2000), and thus, we included this factor in our null model. In our database, the majority of studies used nuclear microsatellites (140 species), followed by allozymes (114 species). Fewer studies used dominantly inherited markers, including Amplified Fragment Length Polymorphism (60 species), Random Amplification of Polymorphic DNA (16 species), and Inter-Simple Sequence Repeat (seven species).

### 2.4 | Distance between populations

Greater distance between populations should correspond to greater genetic differentiation based on an isolation by distance model (Wright, 1943). Thus, we also included in our null model the maximum distance between populations used in each study. We calculated this based on the coordinates of the two most distant populations. When this was not available, we used the scale bar of maps showing sampled populations. Distance varied from 0.01 to 9,900 km (mean =  $703 \pm 1,077$  SD).

### 2.5 | Mean sample size per population

The maximum value that  $F_{ST}$  can take decreases when the within-population expected heterozygosity increases. Thus, a general concern is that large sample sizes are required because small samples could under- or overestimate within-population heterozygosity, inflating or underestimating  $F_{ST}$  values, respectively (Holsinger & Weir, 2009; Kalinowski, 2005; Willing, Dreyer, & van Oosterhout, 2012). We accounted for this potential bias by including the mean sample size per population in our null model. Across the studies, this sample size ranged from three to 285 individuals per population, with an overall mean of  $40.12 (\pm 44.9$  SD).

### 2.6 | Pollination mode

Species were coded as pollinated by wind, small insects, large insects or vertebrates. Small insect pollinators included small Hymenoptera (i.e. *Trigona* and *Melipona* bees and wasps), Diptera (i.e. hoverflies and

gnats), Coleoptera (i.e. small curculionids), Hemiptera (i.e. Anthocoridae and Miridae) and Thysanoptera (i.e. thrips). Large insects included large bees (i.e. honeybees, bumblebees, carpenter bees, euglossine bees) and Lepidoptera (i.e. hawk moths and yucca moths, monarch butterflies). We included honeybees in the large insect category based on evidence showing that honeybees have flying and pollen carry-over capacity similar to bumblebees (Cresswell, Bassom, Bell, Collins, & Kelly, 1995; Escaravage & Wagner, 2004). Vertebrates included bats, hummingbirds and other nectarivorous birds such as honeyeaters and sunbirds. Some instances of vertebrate pollination were more generalized, with visitors including a combination of bats, birds, rodents and/or marsupials.

### 2.7 | Seed dispersal mode

Species were coded as dispersed by wind, animals or gravity. Plants adapted to wind dispersal presented fruits or seeds that were particularly light and/or winged. For those plants adapted to animal dispersal, exploratory analyses showed that different types of animal dispersal (i.e. different types of zoochory) were not significantly different (results not shown). Thus, we kept the animal dispersal category broad, including plants with fruits or seeds dispersed by endo-, ecto- or syn-zoochory (i.e. ingested, externally attached or buried by seed-cachers, respectively). Plants with no adaptations for vector-mediated seed dispersal were coded as gravity dispersed. Based on the information reported in publications with  $F_{ST}$  and trait data, we did not find evidence of secondary movement of fruits or seeds by biotic agents. In some instances, however, water may play a secondary role in dispersing seeds that fall under mother plants, as in the mangrove species *Avicennia* spp. and *Rhizophora* spp., and for *Beta vulgaris* L., *Casuarina cunninghamiana* Miq., *Cocos nucifera* L., and *Primula nutans* Georgi, as well as for many forest trees after floods or inhabiting riparian sites (Levine & Murrell, 2003; Nilsson, Brown, Jansson, & Merritt, 2010).

### 2.8 | Mating system

We coded species as selfing, mixed-mating or outcrossing, as identified by the authors in each study. Selfing species included strictly autogamous species. They were rare ( $N = 7$ ) and not included in the final 337-species data set, due to their low sample size. Mixed-mating species included those that undergo both outcrossing and selfing to some extent, through either autogamy or geitonogamy (Goodwillie, Kalisz, & Eckert, 2005). Outcrossing species included plants that are self-incompatible, unisexual (i.e. monoecious or dioecious) or dichogamous hermaphrodites; that is, either having the male reproductive organs come to maturity before the female organs (protandry), or vice versa (protogyny).

### 2.9 | Growth form

Species were coded as trees, shrubs or non-woody plants. Trees included woody plants >10 m tall, typically with a single trunk coming from the base. Shrubs included upright woody plants

<10 m tall, typically with one or several trunks coming from the base. We also included in the shrub category hemi-parasites and hemi-epiphytes. Non-woody plants included herbs, epiphytes and non-woody climbers. Growth form of species was often linked to habitat in that many non-woody plants and shrubs occurred in the forest understory, while many trees occurred in the subcanopy and canopy. However, non-woody plants, shrubs and trees also occurred in open habitats like prairies. We did not include habitat as an additional predictor in our models due to its high collinearity with growth form.

## 2.10 | Latitudinal region

We recorded the geographic location of each study to create an additional categorical variable for latitudinal region. Species were coded as tropical, subtropical or temperate. Tropical regions included sites between the tropics of Cancer and Capricorn (23.5° north and south of the equator, respectively), which are characterized by relatively low variation in daylight and temperature throughout the year, but with large environmental heterogeneity over short distances. Subtropical regions included latitudes from 23.5° to 35° (north and south). These regions have climates similar to the tropics, but with more seasonal fluctuations. Temperate regions included latitudes >35° north and south. These zones are characterized by a wide range of temperatures throughout the year, and by clearly marked seasonal changes.

## 2.11 | Analytical framework

Analyses were performed in R (R Core Team, 2018). Prior to model testing, we performed transformations of continuous data to improve normality of model residuals (details in Appendix S2).  $F_{ST}$  was transformed using Tukey's ladder of powers transformation (Tukey, 1970) with the function `transformTukey` from the R package `rcompanion` (Mangiafico, 2016). Transformed  $F_{ST}$  resulted in  $F_{ST}^{0.275}$  (Shapiro–Wilk statistic = 0.27,  $p = .7$ ). Continuous predictors were transformed using their natural logarithm. We also estimated correlations (Plackett, 1983) and evaluated multicollinearity issues (Acock & Stavig, 1979; Fox & Monette, 1992) among predictor variables (Appendix S3). The multicollinearity tests indicated that all predictors could be included together in a multiple regression (Table S2 and Table S3).

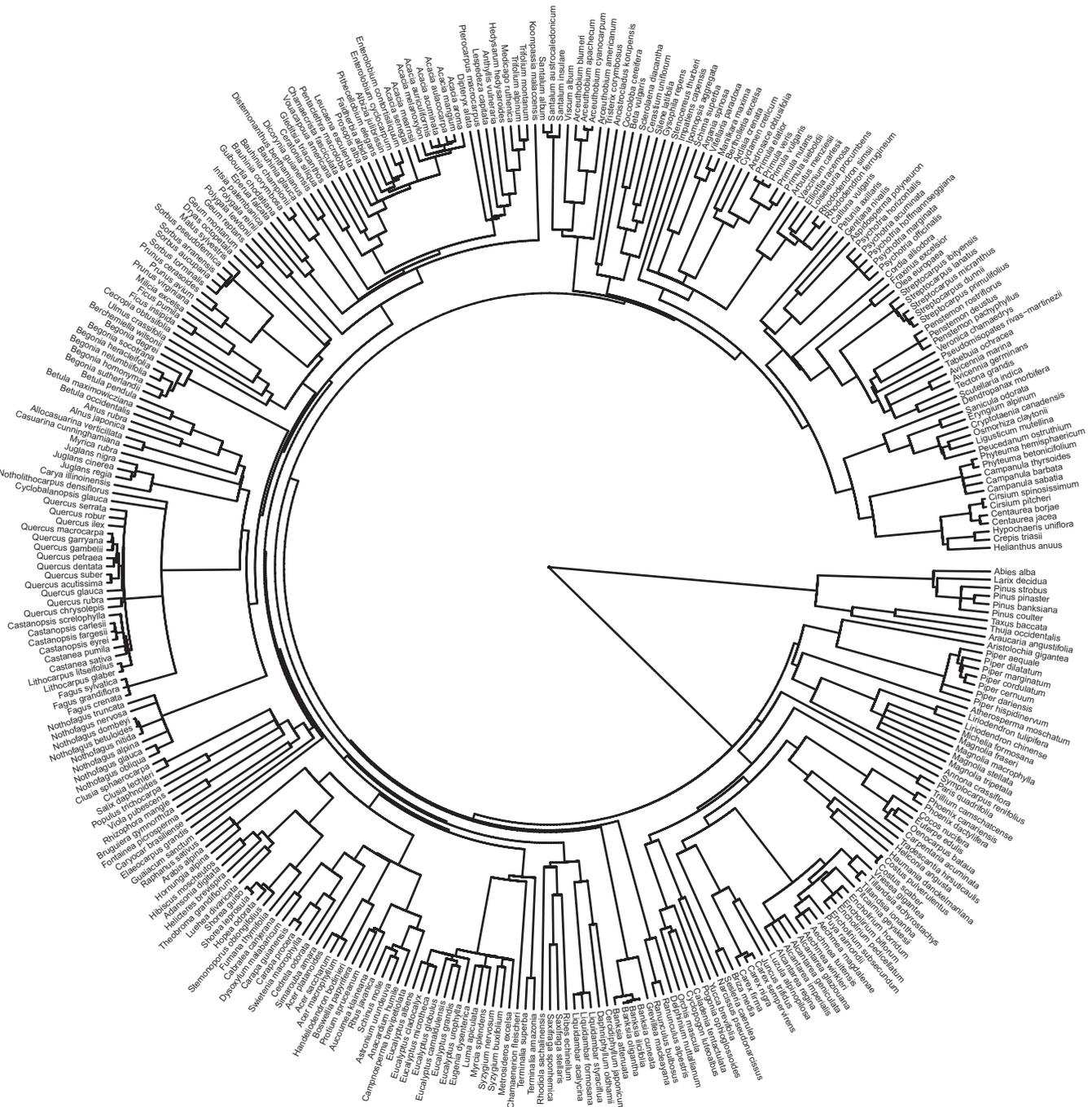
In order to perform models that correct for phylogenetic signal (Freckleton, Harvey, & Pagel, 2002), a species-level phylogeny (Figure 1) was produced with the R package `v.PHYLOMAKER` (Jin & Qian, 2019). This package prunes a custom list of species from the latest and most complete mega-tree of vascular plants (Smith & Brown, 2018) (see Appendix S4 for details). We then assessed phylogenetic signal in categorical predictors with Abouheif's (1999) method (Jombart, Balloux, & Dray, 2010; Pavoine, Ollier, Pontier, & Chessel, 2008), and in  $F_{ST}$  values with Pagel's (1999)

$\lambda$  (Molina-Venegas & Rodríguez, 2017; Revell, 2012) (Appendix S5). With both methods, a value significantly different from zero (and closer to 1) indicates phylogenetic autocorrelation in species traits. We found that closely related species tend to be more similar than expected by chance in their mating system, growth form, pollination mode, seed dispersal mode, latitudinal region and  $F_{ST}$ . The highest observed Moran's  $I$  was that of growth form, followed by pollination mode, latitudinal region, seed dispersal mode and lastly mating system (Fig. S1).  $F_{ST}$  values were also phylogenetically autocorrelated (Pagel's  $\lambda = 0.52$ ,  $p < .001$  and Pagel's  $\lambda = 0.53$ ,  $p < .001$  for raw and transformed  $F_{ST}$  values, respectively). Thus, in order to correct for the potential phylogenetic autocorrelation of residuals in  $F_{ST}$  models (explained below), we implemented phylogenetically informed multiple regressions (Revell, 2010; Symonds & Blomberg, 2014) with the function "phylolm" from the R package `phylolm` (Ho & Ané, 2014). For the fit of models, the likelihood of the parameters was calculated with a Brownian motion model of evolution (Ho & Ané, 2014) (Appendix S6). For the categorical predictors with more than two levels, we chose reference levels based on exploratory analyses with phylogenetic ANOVA and post hoc tests (Garland, Dickerman, Janis, & Jones, 1993; Revell, 2012). We selected the level which mean was most different from that of other levels (Tables S4 and S5). Reference levels were as follows: trees for growth form, small insects for pollination mode, gravity for dispersal mode, and temperate for latitudinal region.

## 2.12 | $F_{ST}$ models

We began our phylogenetic multiple regressions analyses of factors affecting genetic structure by constructing a null model with the sampling-scheme variables. We sequentially added the life history traits to this null model, checking whether each addition improved model fit of a multiple regression based on Akaike information criterion (AIC) scores (Akaike, 1974). Mating system and growth form were added together as there is ample evidence of their effect on  $F_{ST}$  (Duminil et al., 2007; Hamrick & Godt, 1996). We then added pollination mode and seed dispersal mode, to check whether either, or both together, improved the previous model. After finding the best model explaining  $F_{ST}$  with life history traits (Q1), we compared this model to one that included latitudinal region as an additional factor (Q2). We assessed the variance explained by each model with the R package `RR2` and the function "R2.PRED" (Ives, 2018; Ives & Li, 2018). We further evaluated the best-fit model through a backward stepwise model selection with the function "phylolm" in the `phylolm` package. The functions "phylolm" and "phylolm" were congruent in finding the same best model.

We then evaluated the importance of each variable in this best-fit model (Q3). We used the R package `RR2` and the function "R2.LIK" to obtain the unique contribution of each factor in terms of the amount of  $F_{ST}$  variance explained by comparing the best-fit model with a reduced model not including the factor of interest.



**FIGURE 1** Phylogeny of studied species included in this study, representing the major clades of seed plants: gymnosperms, magnoliids, monocots and eudicots

### 3 | RESULTS

#### 3.1 | Taxonomic scope and phylogeny

The 337 species were distributed in 210 genera, representing 96 families in 34 orders. The majority of species (268) belonged to the Eudicots, followed by 43 Monocots, 17 Magnoliids and nine Gymnosperms. The families Fabaceae (mostly *Acacia*; eight species) and Fagaceae (mostly *Quercus*; 13 species) were particularly

well represented, with 37 and 26 species, respectively (Table S1). The resulting phylogeny had 337 tips and 311 nodes (Figure 1). In other words, 92% of the phylogeny was resolved, and only 26 tips (8%) belonged to polytomies. These polytomies correspond to clades for which phylogenetic information remains scarce or unclear (Stevens, 2001): *Begonia* (Begoniaceae), *Alcantarea* and *Encholirium* (Bromeliaceae), *Streptocarpus* (Gesneriaceae), *Arceuthobium* (Santalaceae), *Magnolia* (Magnoliaceae), *Piper* (Piperaceae), *Psychotria* (Rubiaceae), *Acacia* (Fabaceae) and *Sorbus* (Rosaceae).

### 3.2 | Life history traits that promote population divergence in seed plants (Q1)

Among phylogenetic multiple regressions with the four life history traits (models 1–4, Table 1), model 4 was the best fit, indicating that mating system, growth form, pollination mode and seed dispersal mode all influence  $F_{ST}$  (AIC = -482.3). However, the performance of model 4 was almost indistinguishable from that of model 3 ( $\Delta AIC = 2.2$ ), which only differed in the lack of the factor seed dispersal mode. Further evidence for the relative unimportance of seed dispersal mode can be seen in the fact that adding seed dispersal mode to model 1 (which only has mating system and growth form) results in much less improvement of fit (models 2 vs. 1,  $\Delta AIC = 2.5$ ) than adding pollination mode (models 3 vs. 1,  $\Delta AIC = 16.6$ ).

### 3.3 | Differences among latitudinal regions (Q2)

Adding the factor latitudinal region to models with the four life history traits notably increased fit to the data (models 5–7, AIC = -488.6 to -503.9, Table 1), indicating that latitudinal region is a significant predictor of  $F_{ST}$ . This is particularly evident when comparing the best-fit models for each instance (models 4 vs. 6,  $\Delta AIC = 21.6$ ). Model performance was indistinguishable for models 6 versus 7 ( $\Delta AIC = 1$ ), which only differed in the addition of seed dispersal mode. Finally, in models 5 and 7 the factor seed dispersal mode was no longer a significant predictor of  $F_{ST}$  (Table 1 and 2). Below, we focus on results from model 7, as it is the most inclusive model of the factors we tested with the best fit to the data.

**TABLE 1** Phylogenetic multiple regressions explaining variation in  $F_{ST}$

MODEL	Variables <sup>a</sup>	R <sup>2</sup>	AIC	$\lambda$ fit
Null model	Genetic marker Mean sample size <sup>b</sup> Distance <sup>c</sup>	.36	-437	0.57
Model 1	Null model  	.41	-463.5	0.48
Model 2	Null model   	.42	-466	0.46
Model 3	Null model   	.43	-480.1	0.37
Model 4	Null model    	.44	-482.3	0.35
Model 5	Null model    	.42	-488.6	<0.001
Model 6	Null model    	.45	-503.9	<0.001
Model 7	Null model     	.46	-502.9	<0.001

Note: In each model, only the main effect of factors is considered, that is no interactions. AIC and  $\lambda$  fit (scaling parameter to correct for phylogeny) were estimated using maximum likelihood. Underlined variables indicate that at least one of their terms was a significant factor in the corresponding model. (Thick underline:  $p \leq .005$ , thin underline:  $.005 < p < .05$ ).

<sup>a</sup>Yellow circle: mating system, green circle: growth form, brown circle: seed dispersal mode, red circle: pollination mode, blue circle: latitudinal region.

<sup>b</sup>Mean sample size: natural logarithm of the mean sample size of individuals per population.

<sup>c</sup>Distance: natural logarithm of the maximum distance between populations.

**TABLE 2** Details of model 7, the most inclusive phylogenetic model with factors of interest

Variable	N	Estimate	SE	T value	p value
Intercept		0.59	0.04	14.1	<.001
Mating system					
<b>Mixed-mating</b>	80				
Outcrossing	257	-0.07	0.01	-4.7	<.001
Growth form					
<b>Tree</b>	163				
Non-woody	121	0.09	0.02	5.3	<.001
Shrub	53	0.06	0.02	3	.003
Pollination mode					
<b>Small insects</b>	176				
Large insects	48	-0.06	0.02	-3.4	.001
Vertebrates	44	-0.05	0.02	-2.6	.01
Wind	69	-0.05	0.02	-3	.003
Seed dispersal mode					
<b>Gravity</b>	82				
Animals	147	-0.003	0.02	-0.2	.8
Wind	108	-0.02	0.02	-1.4	.1
Latitudinal region					
<b>Temperate</b>	134				
Subtropical	78	0.07	0.02	4.5	<.001
Tropical	125	0.09	0.02	5.4	<.001

Note: Variables in bold indicate the reference level for each categorical factor. *N* indicates the sample size of each group without phylogenetic correction. Significant *p* values are in bold.

### 3.4 | General pattern of $F_{ST}$ variation

Figure 2 shows how the levels of each factor affect population differentiation as measured by  $F_{ST}$  values (after transformation). The marginal (i.e. predicted) effect of each factor on  $F_{ST}^{0.275}$  is depicted while holding constant the other independent variables in model 7. For mating system, outcrossers tend to have lower population differentiation than mixed-mating plants (Figure 2a). Trees tend to have significantly lower population differentiation relative to non-woody plants and shrubs, while the latter two growth forms did not differ between each other (Figure 2b). Pollination by small insects leads to significantly greater differentiation compared to large insect, vertebrate and wind pollination, while the latter three pollination modes did not differ between each other (Figure 2c). Temperate zones have significantly lower  $F_{ST}$  values than tropics and subtropics, and the latter two regions did not differ from each other (Figure 2e). Finally, seed dispersal mode was not a significant predictor of population genetic differentiation.  $F_{ST}$  values associated with gravity dispersal were highly variable, and although gravity dispersal results in higher  $F_{ST}$  values compared to wind dispersal, this difference was not significant. Animal dispersal did not differ from other dispersal modes either (Figure 2d).

### 3.5 | Most important factor for explaining $F_{ST}$ (Q3)

Of all of the factors that we analysed, latitudinal region explained the highest per cent variation (7%), higher than the life history traits in model 7 (0.9%–6%, Figure 2f). Of the life history traits, mating system and pollination mode had the highest independent contribution to the variation in  $F_{ST}$  values (6% each), followed by growth form (4%), while the contribution of dispersal mode was very low (0.9%) and not statistically significant (Figure 2f).

### 3.6 | Influence of variables in the null model

Variables in the null model were significant predictors of  $F_{ST}$  in all multiple regressions (Table 1) and in model 7 (see Table S6 for results pertaining to the variables in the null model). Distance had the highest independent contribution (8%), compared to genetic marker and mean sample size (4% each). In general,  $F_{ST}$  values become larger when the geographic scale of studies increases. In contrast,  $F_{ST}$  values decrease with larger mean sample sizes of individuals per population. There were significant differences between codominant markers (microsatellites and allozymes) and dominant markers (AFLP and RAPD); the former result in lower  $F_{ST}$  values, while the latter result in higher  $F_{ST}$  values. ISSR markers did not differ from others.

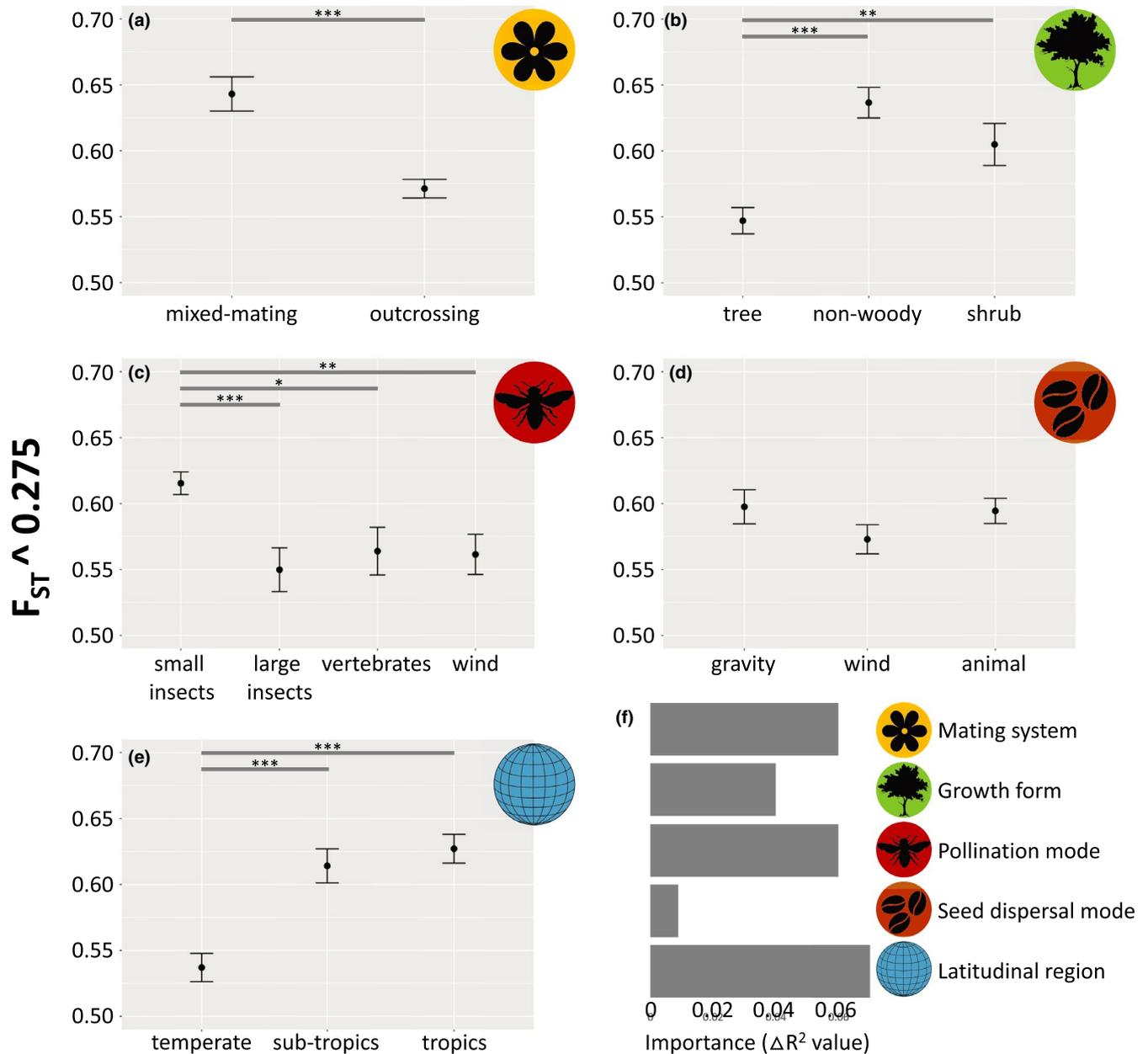
### 3.7 | Further examination of results

Including marker type in our null model helps to account for possible differences between these in the resulting  $F_{ST}$  values. However, as a further check of the robustness of our results, we also ran the same set of parameters as model 7 for each genetic marker (Table S7a–c), each marker type (i.e. codominant vs. dominant; Table S7d,e) and each  $F_{ST}$  analog separately (Table S7f–i). We found that qualitative patterns remained unchanged relative to model 7, with the same traits remaining significant for codominant markers alone, although these differences were not always significant for dominant markers and  $F_{ST}$  analogs, likely due to decreased statistical power given lower sample sizes.

Finally, we note that only subtropical (11 studies) and tropical species (one study) used the statistic  $G'_{ST}$  (Hedrick, 2005), raising the possibility that this may be driving the significant differences we found among latitudinal regions (given that  $G'_{ST}$  usually takes larger values than other  $F_{ST}$  analogs). We thus re-analysed the data after removing these studies and again find identical results to those in model 7 (Table S7j). Together, these additional analyses demonstrate that results are robust to differences in marker type and  $F_{ST}$  analog used in the reviewed studies.

## 4 | DISCUSSION

Here, we provide the most robust and comprehensive evaluation to date of factors driving population genetic differentiation in seed



**FIGURE 2** Marginal effects of factors on predicted  $F_{ST}$  values (after transformation) in model 7: (a) mating system, (b) growth form, (c) pollination mode, (d) seed dispersal mode and (e) latitudinal region. Black dots are predicted  $F_{ST}^{0.275}$  means, and surrounding bars correspond to  $\pm 1$  SD. Significant differences between groups are depicted by the upper horizontal grey lines (\*\* $p < .01$ , \*\*\* $p < .001$ , \* $p < .05$ ). (f) Relative importance of each factor ( $\Delta R^2$  value); the change in  $R^2$  in model 7 relative to a separate model with the individual factor removed

plants. We largely found support for our hypothesis of factors that significantly influence  $F_{ST}$  and several intriguing patterns emerge from our analyses. Overall, we found higher  $F_{ST}$  for tropical, mixed-mating, non-woody species pollinated by small insects, and lower  $F_{ST}$  for temperate, outcrossing, trees pollinated by wind. Latitudinal region was the most important predictor for  $F_{ST}$  relative to the others tested. Mating system and pollination mode had equal contributions for explaining  $F_{ST}$ . Growth form was also a key factor influencing  $F_{ST}$ , while seed dispersal mode was not important in our most inclusive model (Table 2, Figure 2).

#### 4.1 | Influence of latitudinal region on $F_{ST}$

Population differentiation was higher in the tropics and subtropics than in temperate regions (Figure 2e). This result supports the idea that patterns of local diversity, such as the partitioning of genetic diversity among plant populations, cannot be explained in isolation from the geographic and historic processes of each region (Ricklefs, 1987, 2004, 2006). Some factors that may contribute include regional differences in seasonality and geography, which have more generally been hypothesized to contribute to the

latitudinal diversity gradient (i.e. increased species richness closer to the equator) (Mittelbach et al., 2007; Rolland, Condamine, Jiguet, & Morlon, 2014; Schemske, Mittelbach, Cornell, Sobel, & Roy, 2009). Below, we discuss some of these ideas, all of which allude to gene flow being more restricted in tropical regions.

One compelling explanation for the regional differences in  $F_{ST}$  is based on the observation that the tropics can have highly asynchronous rainfall patterns over small spatial scales (Martin et al., 2009). Given that most plants time their flowering to seasons (Crimmins, Crimmins, & Bertelsen, 2011; Gaudinier & Blackman, 2019) and that seasons are largely determined by rainfall in the tropics, small-scale differences in rainfall potentially disrupt gene flow and cause high population differentiation over short distances compared to the temperate zones. In contrast, seasonality in the temperate zones is largely determined by temperature regimes, which remain relatively constant over large distances. Because temperature is the main seasonal cue for flowering time in the temperate zones (i.e. most species flower during the spring), conspecific populations likely flower at similar times throughout their range, promoting gene flow and decreasing population differentiation. This is the aforementioned “asynchrony of seasons hypothesis” (Martin et al., 2009), and our analyses support the prediction of higher population differentiation in the tropics. We note that the tropics and subtropics did not differ in  $F_{ST}$  and that these regions have comparable climatic patterns (Sitnikov, 2009), and thus, the ASH may extend to subtropical regions.

Other factors which can lead to gene flow being more restricted in the tropics include the heterogeneous orogeny and rich fluvial systems characteristic of these latitudes. Such geographic differences have also been hypothesized to contribute to the latitudinal diversity gradient (e.g. Smith et al., 2014; Wallace, 1854). For example, large Amazonian rivers represent barriers to dispersal for many animals, and to some extent for plants (e.g. Rio Negro; Nazareno, Dick, & Lohman, 2017, 2019), promoting isolation between populations and thus genetic differentiation (although this might not be the case for large trees; Dexter et al., 2017). Furthermore, if extant tropical species (like those in our data set) have been in place for a longer time since divergence (or have expanded their ranges more anciently) than temperate counterparts, tropical populations might have had more time to build up genetic differentiation (Kisel & Barraclough, 2010; Smith et al., 2014). This argument becomes particularly compelling in combination with the fact that temperature does not vary as extremely through the year in the tropics. Given this, different subpopulations would be expected to evolve narrower physiological niches that adapt them to particular altitudinal zones, and a similarly sized mountain would impose a greater barrier to dispersal, and thus to gene flow among subpopulations, in tropical than in temperate regions (Ghalambor, 2006; Janzen, 1967).

Thus, overall, our results are in line with hypotheses that suggest greater species diversity in the tropics is due to higher speciation rates rather than lower extinction rates. While the specific mechanisms differ, including those mentioned above and others (see Mittelbach et al., 2007), these hypotheses all posit greater

population-level differentiation that then scales up to faster speciation rates via allopatric or parapatric speciation. On the other hand, greater plant diversity in the tropics could be unrelated with higher speciation rates. For instance, Igea and Tanentzap (2020) found that angiosperm speciation rates are actually higher at higher latitudes, which might imply that our result of higher population differentiation in the tropics does not relate to higher speciation rates. Direct tests on the influence of population differentiation on speciation rates are thus necessary in order to establish whether population differentiation is a rate-limiting step of the speciation process (Harvey et al., 2019). Such tests are scarce and have only focused on vertebrates, finding a positive association in New World birds (Harvey et al., 2017), and no association in Australian lizards (Singhal et al., 2018). We encourage similar tests in seed plants at a global scale. Nevertheless, our result supports two previous studies that document such a pattern of greater population differentiation in the tropics in animals (Martin & McKay, 2004) and in plants (Eo, Wares, & Carrol, 2008).

## 4.2 | Influence of pollination mode on $F_{ST}$

We found that pollination mode plays a key role in population differentiation. Our result is novel compared to the previous most recent review on the subject, which did not find a difference between wind versus animal pollination (Duminil et al., 2007). Here, we subdivide animal pollination into small insects, large insects and vertebrates and find that species pollinated by small insects have significantly higher  $F_{ST}$  than those with other pollination modes (Figure 2c). This pattern is likely due to reduced gene flow among plant populations. In fact, small insects have a lower pollen carry-over capacity than bumblebees and vertebrates (Dick et al., 2008; Rhodes, Fant, & Skogen, 2017), and studies of pollinator movement show that euglossine bees, hawkmoths and bats can all travel long distances, even across fragmented habitats (Brunet, Larson-Rabin, & Stewart, 2012; Finger, Kaiser-Bunbury, Kettle, Valentin, & Ghazoul, 2014; Janzen, 1971; López-Urbe, Oi, & Del Lama, 2008; McCulloch et al., 2013; Skogen, Overson, Hilpman, & Fant, 2019). Our results show that wind, large insects and vertebrates have similar and statistically indistinguishable homogenizing effects on plant  $F_{ST}$ . The similar effect of such pollination modes on plant  $F_{ST}$  is likely why Duminil et al. (2007) found that the broad category of “animal pollination” was not significantly different from wind pollination. Taken together, our findings suggest that plants pollinated by small insects might be more sensitive to habitat fragmentation; the inability of these pollinators to connect distant fragments may decrease genetic diversity within populations, and along with it the ability to adapt in response to anthropogenic change.

One important caveat is that the limited information on pollination systems for many species necessitated a relatively coarse-grained division of pollination mode into broad taxonomic groups. This approach overlooks potential behavioural differences within these groups. For instance, within the vertebrate pollination category,

territorial hummingbirds likely move pollen much shorter distances than trap-lining hummingbirds (Betts, Hadley, & Kress, 2015; Ohashi & Thomson, 2009), and bats may carry pollen more efficiently (Muchhala & Thomson, 2010) and to longer distances than hummingbirds (Lemke, 1984, 1985; Tello-Ramos, Hurly, & Healy, 2015).

### 4.3 | Influence of mating system on $F_{ST}$

Our results provide additional support for the idea that mating system is a strong predictor of  $F_{ST}$  (Figure 2a), even in the presence of covariates, in line with Duminil et al. (2007). Mating system associates with  $F_{ST}$  because any amount of inbreeding (through mixed-mating) increases homozygosity within a subpopulation and reduces its effective population size, leading to increased population structure due to genetic drift. In contrast, outcrossing maintains genetic cohesion within and among subpopulations, decreasing genetic drift and reducing population structure (Charlesworth, 2003; Duminil et al., 2009). Because populations of mixed-mating species are often highly differentiated, they will likely have populations with unique genetic diversity. Accordingly, conservation efforts for them should maximize the number of populations protected to maximize genetic diversity to increase their chances to adapt to environmental change (Ellstrand & Elam, 1993).

### 4.4 | Influence of growth form on $F_{ST}$

We found that trees have populations with significantly lower  $F_{ST}$  than both shrubs and non-woody plants (Figure 2b). Even though most trees are outcrossing in our data set, our results show that growth form contributes to the variation in  $F_{ST}$  independently from mating system. Duminil et al. (2007) found a similar pattern of  $F_{ST}$  variation in respect to growth form, but this result was no longer significant after accounting for phylogenetic signal. We suspect our larger data set, with its broader phylogenetic sampling (96 families and 210 genera in our study vs. 45 families and 86 genera in Duminil et al., 2007), allowed us to uncover the significant effect of growth form on  $F_{ST}$ . Growth form was the least important significant factor in model 7 (Figure 2f), which supports the idea that a large and diverse sample size is required to uncover its independent effect. The inherent difference in scale between growth forms may contribute to the difference in  $F_{ST}$  values, in that a given geographic distance between subpopulations may restrict gene flow much more for an herb than for a tree. In fact, neighbourhood size, that is the spatial extent of closely related individuals, is larger in trees than shrubs and herbs (Vekemans & Hardy, 2004). Furthermore, trees usually have greater longevity than shrubs and non-woody plants (Duminil et al., 2009), which may increase the chances of gene flow between tree subpopulations, more than for other growth forms. Finally, the fact that growth form and habitat are tightly linked may also contribute; many non-woody plants and shrubs in our data set occur in the

forest understory, while many trees reach the canopy. Givnish (2010) and Theim, Shirk, and Givnish (2014) hypothesized that the understory imposes more limits to gene flow than the canopy because of the sedentary lifestyle of animal mutualists in the understory.

### 4.5 | Seed dispersal and $F_{ST}$

Our results did not support the hypothesis that gravity-mediated seed dispersal increases population differentiation compared to wind or animal dispersal (Givnish, 2010) (Figure 2d). This is in line with previous findings suggesting that the genetic structure of nuclear markers is largely driven by pollen flow (Petit et al., 2005; Skogen et al., 2019; Sork, Nason, Campbell, & Fernandez, 1999) and that the effect of seed dispersal is only detectable in the population genetic structure of chloroplast genes (Duminil et al., 2007). However, we note that gravity dispersal resulted in highly variable  $F_{ST}$  values, potentially due to unrecorded secondary seed vectors.  $F_{ST}$  values for animal dispersal were also variable, which suggests that different animals (i.e. different taxonomic groups rather than different types of zoochory) could have different effects on population differentiation. Thus overall, as with vertebrate pollination, we suspect that more fine-scaled classifications of dispersers may improve our understanding of their effects on plant population genetic structure. Testing this idea, however, requires more detailed data on animal dispersers, which was often not available in the publications we reviewed, or was difficult to characterize. For example, in our study many species have a mix of seed dispersers, including small to large mammals and birds (like most *Arecaceae*, *Fabaceae*, *Fagaceae*, *Myrtaceae*, *Sapotaceae*, among others), making it difficult to assign plants to a disperser-specific taxonomic affiliation or foraging behavioural trait.

### 4.6 | Considerations on model inference

Phylogenetic multiple regressions allowed us to evaluate the unique effect of each predictor on  $F_{ST}$  while correcting for phylogenetic autocorrelation, which had not been accomplished in previous broad-scale studies. Additionally, we note that after adding the factor latitudinal region, the scaling parameter that corrects for phylogenetic autocorrelation ( $\lambda$  fit in Table 1) became insignificant. This suggests that latitudinal region decreases the phylogenetic autocorrelation in the residuals modelled by our phylogenetic regressions (Freckleton, 2009). We suspect that region captured important phylogenetic information in  $F_{ST}$  and species traits; within each regional species pool, lineages share strong biogeographic and phylogenetic affinities. Put another way, we think that regional affiliation is the most important underlying factor influencing  $F_{ST}$  values at a global scale, and when not included, phylogenetic signal becomes a proxy for latitudinal region due to the tendency for closely related species to occur in similar regions.

## 4.7 | Future directions

Understanding how plant population genetic structure is affected by life history traits can greatly improve management strategies for populations facing increasingly fragmented habitats due to human-accelerated global change. Our study reveals that gene flow is generally more limited in non-woody species pollinated by small insects, making them more susceptible to isolation and loss of genetic diversity. Thus, in order to preserve the largest amount of genetic diversity for species with such traits, conservation efforts should seek to maintain numerous subpopulations spanning a wide geographic extent. Future broad-scale studies of  $F_{ST}$  variation could provide even greater insights for conservation by including population densities (Murawski & Hamrick, 1991; Sork et al., 1999), effects of habitat fragmentation (Aguilar, Quesada, Ashworth, Herrerias-Diego, & Lobo, 2008; Skogen et al., 2019), and the landscape context of populations (Sork et al., 1999).

Another avenue for future research involves linking patterns of genetic variation at different scales. Little is known about how factors that affect genetic patterns over fine spatial scales (i.e. within subpopulations) extend to genetic patterns over larger spatial scales (i.e. among subpopulations). Intuitively, species with greater fine-scale genetic structure (Loiselle, Sork, Nason, & Graham, 1995) should also have greater population genetic structure, but this has rarely been tested. For example, a recent review found greater fine-scale genetic structure in species with short-distance dispersers, than those dispersed by birds (Gelmi-Candusso et al., 2017), but it is unclear whether this difference would extend over larger distances. Overall, we expect that more comprehensive studies of ecological interactions in combination with increasing amounts of genetic data collected at various spatial scales will continue to improve our understanding of the factors that influence population genetic structure in seed plants.

### ACKNOWLEDGEMENTS

We thank the researchers whose published data we analysed in this paper. Thanks to Robert Ricklefs, Christine Edwards and Carmen Ulloa for advice in this study. Isabel Loza, Justin Baldwin and Sebastián Tello provided valuable help with statistical analyses. Many thanks to Justin Zweck, Justin Baldwin, Krissa Skogen and to members of the Muchhala Lab at the University of Missouri–Saint Louis for constructive discussions on a previous version of this manuscript. We also thank two anonymous reviewers whose comments helped improve this manuscript. This research was supported by the Whitney Harris World Ecology Center at the University of Missouri at Saint Louis.

### AUTHOR CONTRIBUTIONS

D.G. and N.M. planned and designed the research. D.G. collected and analysed the data. D.G. wrote the first draft of the manuscript. D.G. and N.M. contributed equally to substantial revisions of the manuscript.

### DATA AVAILABILITY STATEMENT

The data sets generated and analysed during the current study, along with the R script supporting the results, are available in the Dryad repository (<https://doi.org/10.5061/dryad.d2547d819>). No new data were used in this research because analyses were based on a literature review of published studies (Gamba & Muchhala, 2020).

### ORCID

Diana Gamba  <https://orcid.org/0000-0002-0421-6437>

Nathan Muchhala  <https://orcid.org/0000-0002-4423-5130>

### REFERENCES

- Abouheif, E. (1999). A method for testing the assumption of phylogenetic independence in comparative data. *Evolutionary Ecology Research*, 1(8), 895–909.
- Acocck, A. C., & Stavig, G. R. (1979). A measure of association for non-parametric statistics. *Social Forces*, 57(4), 1381–1386. <https://doi.org/10.1093/sf/57.4.1381>
- Aguilar, R., Quesada, M., Ashworth, L., Herrerias-Diego, Y., & Lobo, J. (2008). Genetic consequences of habitat fragmentation in plant populations: Susceptible signals in plant traits and methodological approaches. *Molecular Ecology*, 17(24), 5177–5188. <https://doi.org/10.1111/j.1365-294X.2008.03971.x>
- Aguillon, S. M., Fitzpatrick, J. W., Bowman, R., Schoech, S. J., Clark, A. G., Coop, G., & Chen, N. (2017). Deconstructing isolation-by-distance: The genomic consequences of limited dispersal. *PLoS Genetics*, 13(8), e1006911. <https://doi.org/10.1371/journal.pgen.1006911>
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6), 716–723. <https://doi.org/10.1109/TAC.1974.1100705>
- Angeloni, F., Ouborg, N. J., & Leimu, R. (2011). Meta-analysis on the association of population size and life history with inbreeding depression in plants. *Biological Conservation*, 144(1), 35–43. <https://doi.org/10.1016/j.biocon.2010.08.016>
- Betts, M. G., Hadley, A. S., & Kress, W. J. (2015). Pollinator recognition by a keystone tropical plant. *Proceedings of the National Academy of Sciences of the United States of America*, 112(11), 3433–3438. <https://doi.org/10.1073/pnas.1419522112>
- Blasco-Costa, I., & Poulin, R. (2013). Host traits explain the genetic structure of parasites: A meta-analysis. *Parasitology*, 140(10), 1316–1322. <https://doi.org/10.1017/S0031182013000784>
- Brunet, J., Larson-Rabin, Z., & Stewart, C. M. (2012). The distribution of genetic diversity within and among populations of the rocky mountain columbine: The impact of gene flow, pollinators, and mating system. *International Journal of Plant Sciences*, 173(5), 484–494. <https://doi.org/10.1086/665263>
- Castellanos, M. C., Wilson, P., & Thomson, J. D. (2003). Pollen transfer by hummingbirds and bumblebees, and the divergence of pollination modes in *Penstemon*. *Evolution*, 57(12), 2742–2752. <https://doi.org/10.1111/j.0014-3820.2003.tb01516.x>
- Cavers, S., Degen, B., Caron, H., Lemes, M. R., Margis, R., Salgueiro, F., & Lowe, A. J. (2005). Optimal sampling strategy for estimation of spatial genetic structure in tree populations. *Heredity*, 95(4), 281–289. <https://doi.org/10.1038/sj.hdy.6800709>
- Charlesworth, D. (2003). Effects of inbreeding on the genetic diversity of populations. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 358(1434), 1051–1070. <https://doi.org/10.1098/rstb.2003.1296>
- Cresswell, J. E., Bassom, A. P., Bell, S. A., Collins, S. J., & Kelly, T. B. (1995). Predicted pollen dispersal by honey-bees and three

- species of bumble-bees foraging on oil-seed rape: A comparison of three models. *Functional Ecology*, 9, 829–841. <https://doi.org/10.2307/2389980>
- Crimmins, T. M., Crimmins, M. A., & Bertelsen, C. D. (2011). Onset of summer flowering in a 'Sky Island' is driven by monsoon moisture. *New Phytologist*, 191(2), 468–479. <https://doi.org/10.1111/j.1469-8137.2011.03705.x>
- Dexter, K. G., Lavin, M., Torke, B. M., Twyford, A. D., Kursar, T. A., Coley, P. D., ... Pennington, R. T. (2017). Dispersal assembly of rain forest tree communities across the Amazon basin. *Proceedings of the National Academy of Sciences of the United States of America*, 114(10), 2645–2650. <https://doi.org/10.1073/pnas.1613655114>
- Dick, C. W., Hardy, O. J., Jones, F. A., & Petit, R. J. (2008). Spatial scales of pollen and seed-mediated gene flow in tropical rain forest trees. *Tropical Plant Biology*, 1(1), 20–33. <https://doi.org/10.1007/s12042-007-9006-6>
- Duminil, J., Fineschi, S., Hampe, A., Jordano, P., Salvini, D., Vendramin, G. G., & Petit, R. J. (2007). Can population genetic structure be predicted from life-history traits? *The American Naturalist*, 169(5), 662–672. <https://doi.org/10.1086/513490>
- Duminil, J., Hardy, O. J., & Petit, R. J. (2009). Plant traits correlated with generation time directly affect inbreeding depression and mating system and indirectly genetic structure. *BMC Evolutionary Biology*, 9(1), 177. <https://doi.org/10.1186/1471-2148-9-177>
- Ellstrand, N. C. (1992). Gene flow by pollen: Implications for plant conservation genetics. *Oikos*, 63(1), 77–86. <https://doi.org/10.2307/3545517>
- Ellstrand, N. C. (2014). Is gene flow the most important evolutionary force in plants? *American Journal of Botany*, 101(5), 737–753. <https://doi.org/10.3732/ajb.1400024>
- Ellstrand, N. C., & Elam, D. R. (1993). Population genetic consequences of small population size: Implications for plant conservation. *Annual Review of Ecology and Systematics*, 24(1), 217–242. <https://doi.org/10.1146/annurev.es.24.110193.001245>
- Eo, S. H., Wares, J. P., & Carroll, J. P. (2008). Population divergence in plant species reflects latitudinal biodiversity gradients. *Biology Letters*, 4(4), 382–384. <https://doi.org/10.1098/rsbl.2008.0109>
- Escaravage, N., & Wagner, J. (2004). Pollination effectiveness and pollen dispersal in a *Rhododendron ferrugineum* (Ericaceae) population. *Plant Biology*, 6(05), 606–615. <https://doi.org/10.1055/s-2004-821143>
- Esquerré, D., Brennan, I. G., Catullo, R. A., Torres-Pérez, F., & Keogh, J. S. (2019). How mountains shape biodiversity: The role of the Andes in biogeography, diversification, and reproductive biology in South America's most species-rich lizard radiation (Squamata: Liolaemidae). *Evolution*, 73(2), 214–230. <https://doi.org/10.1111/evo.13657>
- Excoffier, L., Smouse, P. E., & Quattro, J. M. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics*, 131(2), 479–491.
- Finger, A., Kaiser-Bunbury, C. N., Kettle, C. J., Valentin, T., & Ghazoul, J. (2014). Genetic connectivity of the moth pollinated tree *Glionnetia sericea* in a highly fragmented habitat. *PLoS One*, 9(10), e111111. <https://doi.org/10.1371/journal.pone.0111111>
- Fox, J., & Monette, G. (1992). Generalized collinearity diagnostics. *Journal of the American Statistical Association*, 87(417), 178–183. <https://doi.org/10.1080/01621459.1992.10475190>
- Freckleton, R. P. (2009). The seven deadly sins of comparative analysis. *Journal of Evolutionary Biology*, 22(7), 1367–1375. <https://doi.org/10.1111/j.1420-9101.2009.01757.x>
- Freckleton, R. P., Harvey, P. H., & Pagel, M. (2002). Phylogenetic analysis and comparative data: A test and review of evidence. *The American Naturalist*, 160(6), 712–726. <https://doi.org/10.1086/343873>
- Gamba, D., & Muchhala, N. (2020). Global patterns of population genetic differentiation in seed plants. *Dryad*: <https://doi.org/10.5061/dryad.d2547d819>
- Garland, T., Dickerman, A. W., Janis, C. M., & Jones, J. A. (1993). Phylogenetic analysis of covariance by computer simulation. *Systematic Biology*, 42(3), 265–292. <https://doi.org/10.1093/sysbio/42.3.265>
- Gaudinier, A., & Blackman, B. K. (2019). Evolutionary processes from the perspective of flowering time diversity. *New Phytologist*, 225(5), 1883–1898. <https://doi.org/10.1111/nph.16205>
- Gelmi-Candusso, T. A., Heymann, E. W., & Heer, K. (2017). Effects of zoochory on the spatial genetic structure of plant populations. *Molecular Ecology*, 26(21), 5896–5910. <https://doi.org/10.1111/mec.14351>
- Ghalambor, C. K. (2006). Are mountain passes higher in the tropics? Janzen's hypothesis revisited. *Integrative and Comparative Biology*, 46(1), 5–17. <https://doi.org/10.1093/icb/46.1>
- Givnish, T. J. (2010). Ecology of plant speciation. *Taxon*, 59(5), 1326–1366. <https://doi.org/10.1002/tax.595003>
- Goodwillie, C., Kalisz, S., & Eckert, C. G. (2005). The evolutionary enigma of mixed mating systems in plants: Occurrence, theoretical explanations, and empirical evidence. *Annual Review of Ecology, Evolution, and Systematics*, 36(1), 47–79. <https://doi.org/10.1146/annurev.ecolsys.36.091704.175539>
- Hahn, C. Z., Michalski, S. G., Fischer, M., & Durka, W. (2016). Genetic diversity and differentiation follow secondary succession in a multi-species study on woody plants from subtropical China. *Journal of Plant Ecology*, 10(1), 213–221. <https://doi.org/10.1093/jpe/rtw054>
- Hamrick, J. L., & Godt, M. J. W. (1996). Effects of life history traits on genetic diversity in plant species. *Philosophical Transactions: Biological Sciences*, 351(1345), 1291–1298. <https://doi.org/10.1098/rstb.1996.0112>
- Hamrick, J. L., Godt, M. J. W., & Sherman-Broyles, S. L. (1992). Factors influencing levels of genetic diversity in woody plant species. In W. T. Adams, S. H. Strauss, D. L. Copes & A. R. Griffin (Eds.), *Population genetics of forest trees. forestry sciences*, Vol. 42 (pp. 95–124). Berlin, Germany: Springer. [https://doi.org/10.1007/978-94-011-2815-5\\_7](https://doi.org/10.1007/978-94-011-2815-5_7)
- Hamrick, J. L., Murawski, D. A., & Nason, J. D. (1993). The influence of seed dispersal mechanisms on the genetic structure of tropical tree populations. *Vegetatio*, 107(1), 281–297. <https://doi.org/10.1007/BF00052230>
- Hamrick, J. L., & Trapnell, D. W. (2011). Using population genetic analyses to understand seed dispersal patterns. *Acta Oecologica*, 37(6), 641–649. <https://doi.org/10.1016/j.actao.2011.05.008>
- Harvey, M. G., Seeholzer, G. F., Smith, B. T., Rabosky, D. L., Cuervo, A. M., & Brumfield, R. T. (2017). Positive association between population genetic differentiation and speciation rates in New World birds. *Proceedings of the National Academy of Sciences of the United States of America*, 114(24), 6328–6333. <https://doi.org/10.1073/pnas.1617397114>
- Harvey, M. G., Singhal, S., & Rabosky, D. L. (2019). Beyond reproductive isolation: Demographic controls on the speciation process. *Annual Review of Ecology, Evolution, and Systematics*, 50(1), 75–95. <https://doi.org/10.1146/annurev-ecolsys-110218-024701>
- Hedrick, P. W. (2005). A standardized genetic differentiation measure. *Evolution*, 59(8), 1633–1638. <https://doi.org/10.1111/j.0014-3820.2005.tb01814.x>
- Hey, J., & Pinho, C. (2012). Population genetics and objectivity in species diagnosis: Population genetics and species diagnosis. *Evolution*, 66(5), 1413–1429. <https://doi.org/10.1111/j.1558-5646.2011.01542.x>
- Ho, T. L., & Ané, C. (2014). A linear-time algorithm for gaussian and non-gaussian trait evolution models. *Systematic Biology*, 63(3), 397–408. <https://doi.org/10.1093/sysbio/syu005>
- Holsinger, K. E., & Weir, B. S. (2009). Genetics in geographically structured populations: Defining, estimating and interpreting  $F_{ST}$ . *Nature Reviews Genetics*, 10(9), 639–650. <https://doi.org/10.1038/nrg2611>
- Igea, J., & Tanentzap, A. J. (2020). Angiosperm speciation cools down in the tropics. *Ecology Letters*, 23(4), 692–700. <https://doi.org/10.1111/ele.13476>

- Ives, A. R. (2018). R2s for correlated data: Phylogenetic models, LMMs, and GLMMs. *Systematic Biology*, 68(2), 234–251. <https://doi.org/10.1093/sysbio/syy060>
- Ives, A. R., & Li, D. (2018). rr2: An R package to calculate R2s for regression models. *J. Open Source Software*, 3(30), 1028. <https://doi.org/10.21105/joss.01028>
- Janzen, D. H. (1967). Why mountain passes are higher in the tropics. *The American Naturalist*, 101(919), 233–249. <https://doi.org/10.1086/282487>
- Janzen, D. H. (1971). Euglossine bees as long-distance pollinators of tropical plants. *Science, New Series*, 171(3967), 203–205. <https://doi.org/10.1126/science.171.3967.203>
- Jin, Y., & Qian, H. (2019). V.PhyloMaker: An R package that can generate very large phylogenies for vascular plants. *Ecography*, 42(8), 1353–1359. <https://doi.org/10.1111/ecog.04434>
- Jombart, T., Balloux, F., & Dray, S. (2010). Adephylo: New tools for investigating the phylogenetic signal in biological traits. *Bioinformatics*, 26(15), 1907–1909. <https://doi.org/10.1093/bioinformatics/btq292>
- Kalinowski, S. T. (2005). Do polymorphic loci require large sample sizes to estimate genetic distances? *Heredity*, 94(1), 33–36. <https://doi.org/10.1038/sj.hdy.6800548>
- Kisel, Y., & Barraclough, T. G. (2010). Speciation has a spatial scale that depends on levels of gene flow. *The American Naturalist*, 175(3), 316–334. <https://doi.org/10.1086/650369>
- Kramer, A. T., & Havens, K. (2009). Plant conservation genetics in a changing world. *Special Issue: Plant Science Research in Botanic Gardens*, 14(11), 599–607. <https://doi.org/10.1016/j.tplants.2009.08.005>
- Leimu, R., & Fischer, M. (2008). A meta-analysis of local adaptation in plants. *PLoS One*, 3(12), e4010. <https://doi.org/10.1371/journal.pone.0004010>
- Lemke, T. O. (1984). Foraging ecology of the long-nosed bat, *Glossophaga soricina*, with respect to resource availability. *Ecology*, 65(2), 538–548. <https://doi.org/10.2307/1941416>
- Lemke, T. O. (1985). Pollen carrying by the nectar-feeding bat *Glossophaga soricina* in a suburban environment. *Biotropica*, 17(2), 107–111. <https://doi.org/10.2307/2388502>
- Levine, J. M., & Murrell, D. J. (2003). The community-level consequences of seed dispersal patterns. *Annual Review of Ecology, Evolution, and Systematics*, 34(1), 549–574. <https://doi.org/10.1146/annurev.ecolsys.34.011802.132400>
- Linhart, Y. B., & Grant, M. C. (1996). Evolutionary significance of local genetic differentiation in plants. *Annual Review of Ecology and Systematics*, 27, 237–277. <https://doi.org/10.1146/annurev.ecolsys.27.1.237>
- Loiselle, B. A., Sork, V. L., Nason, J., & Graham, C. (1995). Spatial genetic structure of a tropical understory shrub, *Psychotria officinalis* (Rubiaceae). *American Journal of Botany*, 82(11), 1420–1425. <https://doi.org/10.1002/j.1537-2197.1995.tb12679.x>
- López-Urbe, M. M., Oi, C. A., & Del Lama, M. A. (2008). Nectar-foraging behavior of Euglossine bees (Hymenoptera: Apidae) in urban areas. *Apidologie*, 39(4), 410–418. <https://doi.org/10.1051/apido:2008023>
- Loveless, M. D., & Hamrick, J. L. (1984). Ecological determinants of genetic structure in plant populations. *Annual Review of Ecology and Systematics*, 15, 65–95. <https://doi.org/10.1146/annurev.ecolsys.15.110184.000433>
- Mangiafico, S. S. (2016). *Summary and Analysis of Extension Program Evaluation in R, version 1.18.1*. [rcompanion.org/handbook/](http://rcompanion.org/handbook/)
- Martin, P. R., Bonier, F., Moore, I., & Tewksbury, J. (2009). Latitudinal variation in the asynchrony of seasons: Implications for higher rates of population differentiation and speciation in the tropics. *Ideas in Ecology and Evolution*, 2, 9–17. <https://doi.org/10.4033/iee.2009.2.3.n>
- Martin, P. R., & McKay, J. K. (2004). Latitudinal variation in genetic divergence of populations and the potential for future speciation. *Evolution*, 58(5), 938–945. <https://doi.org/10.1111/j.0014-3820.2004.tb00428.x>
- McCauley, D. E. (1994). Contrasting the distribution of chloroplast DNA and allozyme polymorphism among local populations of *Silene alba*: Implications for studies of gene flow in plants. *Proceedings of the National Academy of Sciences of the United States of America*, 91(17), 8127–8131. <https://doi.org/10.1073/pnas.91.17.8127>
- McCulloch, E. S., Tello, J. S., Whitehead, A., Rolón-Mendoza, C. M., Maldonado-Rodríguez, M. C., & Stevens, R. D. (2013). Fragmentation of Atlantic Forest has not affected gene flow of a widespread seed-dispersing bat. *Molecular Ecology*, 22(18), 4619–4633. <https://doi.org/10.1111/mec.12418>
- Meirmans, P. G., & Hedrick, P. W. (2011). Assessing population structure:  $F_{ST}$  and related measures. *Molecular Ecology Resources*, 11(1), 5–18. <https://doi.org/10.1111/j.1755-0998.2010.02927.x>
- Mittelbach, G. G., Schemske, D. W., Cornell, H. V., Allen, A. P., Brown, J. M., Bush, M. B., ... Turelli, M. (2007). Evolution and the latitudinal diversity gradient: Speciation, extinction and biogeography. *Ecology Letters*, 10(4), 315–331. <https://doi.org/10.1111/j.1461-0248.2007.01020.x>
- Molina-Venegas, R., & Rodríguez, M. Á. (2017). Revisiting phylogenetic signal; strong or negligible impacts of polytomies and branch length information? *BMC Evolutionary Biology*, 17(1), 53. <https://doi.org/10.1186/s12862-017-0898-y>
- Muchhala, N., & Thomson, J. D. (2010). Fur versus feathers: Pollen delivery by bats and hummingbirds and consequences for pollen production. *The American Naturalist*, 175(6), 717–726. <https://doi.org/10.1086/652473>
- Murawski, D. A., & Hamrick, J. L. (1991). The effect of the density of flowering individuals on the mating systems of nine tropical tree species. *Heredity*, 67(2), 167–174. <https://doi.org/10.1038/hdy.1991.76>
- Nazareno, A. G., Dick, C. W., & Lohmann, L. G. (2017). Wide but not impermeable: Testing the riverine barrier hypothesis for an Amazonian plant species. *Molecular Ecology*, 26(14), 3636–3648. <https://doi.org/10.1111/mec.14142>
- Nazareno, A. G., Dick, C. W., & Lohmann, L. G. (2019). Tangled banks: A landscape genomic evaluation of Wallace's Riverine barrier hypothesis for three Amazon plant species. *Molecular Ecology*, 28(5), 980–997. <https://doi.org/10.1111/mec.14948>
- Neale, D. B., & Sederoff, R. R. (1988). Inheritance and evolution of conifer organelle genomes. In J. W. Hanover, D. E. Keathley, C. M. Wilson & G. Kuny (Eds.), *Genetic manipulation of woody plants* (pp. 251–264). Boston, MA: Springer.
- Nei, M. (1973). Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences*, 70(12), 3321. <https://doi.org/10.1073/pnas.70.12.3321>
- Nilsson, C., Brown, R. L., Jansson, R., & Merritt, D. M. (2010). The role of hydrochory in structuring riparian and wetland vegetation. *Biological Reviews*, 85(4), 837–858. <https://doi.org/10.1111/j.1469-185X.2010.00129.x>
- Nybom, H. (2004). Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. *Molecular Ecology*, 13(5), 1143–1155. <https://doi.org/10.1111/j.1365-294X.2004.02141.x>
- Nybom, H., & Bartish, I. V. (2000). Effects of life history traits and sampling strategies on genetic diversity estimates obtained with RAPD markers in plants. *Perspectives in Plant Ecology, Evolution and Systematics*, 3(2), 93–114. <https://doi.org/10.1078/1433-8319-00006>
- Ohashi, K., & Thomson, J. D. (2009). Trapline foraging by pollinators: Its ontogeny, economics and possible consequences for plants. *Annals of Botany*, 103(9), 1365–1378. <https://doi.org/10.1093/aob/mcp088>
- Pagel, M. (1999). Inferring the historical patterns of biological evolution. *Nature*, 401(6756), 877–884. <https://doi.org/10.1038/44766>
- Pascual, M., Rives, B., Schunter, C., & Macpherson, E. (2017). Impact of life history traits on gene flow: A multispecies systematic review across oceanographic barriers in the Mediterranean Sea. *PLoS One*, 12(5), e0176419. <https://doi.org/10.1371/journal.pone.0176419>

- Pavoine, S., Ollier, S., Pontier, D., & Chessel, D. (2008). Testing for phylogenetic signal in phenotypic traits: New matrices of phylogenetic proximities. *Theoretical Population Biology*, 73(1), 79–91. <https://doi.org/10.1016/j.tpb.2007.10.001>
- Petit, R. J., Duminil, J., Fineschi, S., Hampe, A., Salvini, D., & Vendramin, G. G. (2005). Invited review: Comparative organization of chloroplast, mitochondrial and nuclear diversity in plant populations. *Molecular Ecology*, 14(3), 689–701. <https://doi.org/10.1111/j.1365-294X.2004.02410.x>
- Plackett, R. L. (1983). Karl Pearson and the Chi-Squared Test. *International Statistical Review / Revue Internationale de Statistique*, 51(1), 59–72. <https://doi.org/10.2307/1402731>
- Quintero, I., González-Caro, S., Zalamea, P.-C., & Cadena, C. D. (2014). Asynchrony of seasons: Genetic differentiation associated with geographic variation in climatic seasonality and reproductive phenology. *The American Naturalist*, 184(3), 352–363. <https://doi.org/10.1086/677261>
- R Core Team (2018). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>
- Revell, L. J. (2010). Phylogenetic signal and linear regression on species data. *Methods in Ecology and Evolution*, 1(4), 319–329. <https://doi.org/10.1111/j.2041-210X.2010.00044.x>
- Revell, L. J. (2012). phytools: An R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution*, 3(2), 217–223. <https://doi.org/10.1111/j.2041-210X.2011.00169.x>
- Rhodes, M. K., Fant, J. B., & Skogen, K. A. (2017). Pollinator identity and spatial isolation influence multiple paternity in an annual plant. *Molecular Ecology*, 26(16), 4296–4308. <https://doi.org/10.1111/mec.14115>
- Ricklefs, R. E. (1977). Environmental heterogeneity and plant species diversity: A hypothesis. *The American Naturalist*, 111(978), 376–381. <https://doi.org/10.1086/283169>
- Ricklefs, R. E. (1987). Community diversity: Relative roles of local and regional processes. *Science*, 235(4785), 167–171. <https://doi.org/10.1126/science.235.4785.167>
- Ricklefs, R. E. (2004). A comprehensive framework for global patterns in biodiversity. *Ecology Letters*, 7(1), 1–15. <https://doi.org/10.1046/j.1461-0248.2003.00554.x>
- Ricklefs, R. E. (2006). Evolutionary diversification and the origin of the diversity–environment relationship. *Ecology*, 87(sp7), S3–S13. [https://doi.org/10.1890/0012-9658\(2006\)87\[3:EDATO\]2.0.CO;2](https://doi.org/10.1890/0012-9658(2006)87[3:EDATO]2.0.CO;2)
- Riginos, C., Douglas, K. E., Jin, Y., Shanahan, D. F., & Tremblay, E. A. (2011). Effects of geography and life history traits on genetic differentiation in benthic marine fishes. *Ecography*, 34(4), 566–575. <https://doi.org/10.1111/j.1600-0587.2010.06511.x>
- Rolland, J., Condamine, F. L., Jiguet, F., & Morlon, H. (2014). Faster speciation and reduced extinction in the tropics contribute to the mammalian latitudinal diversity gradient. *PLoS Biology*, 12(1), e1001775. <https://doi.org/10.1371/journal.pbio.1001775>
- 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? *Annual Review of Ecology, Evolution, and Systematics*, 40(1), 245–269. <https://doi.org/10.1146/annurev.ecolsys.39.110707.173430>
- Singhal, S., Huang, H., Grundler, M. R., Marchán-Rivadeneira, M. R., Holmes, I., Title, P. O., ... Rabosky, D. L. (2018). Does population structure predict the rate of speciation? A comparative test across Australia's most diverse vertebrate radiation. *The American Naturalist*, 192(4), 432–447. <https://doi.org/10.1086/699515>
- Sitnikov, I. G. (2009). Principal weather systems in subtropical and tropical zones. In G. V. Gruzaayloa (Ed.), *Environmental structure and function: Climate system* (pp. 120–139). Oxford, UK: EOLSS.
- Skogen, K. A., Overson, R. P., Hilpman, E. T., & Fant, J. B. (2019). Hawkmoth pollination facilitates long-distance pollen dispersal and reduces isolation across a gradient of land-use change. *Annals of the Missouri Botanical Garden*, 104(3), 495–511. <https://doi.org/10.3417/2019475>
- Smith, B. T., McCormack, J. E., Cuervo, A. M., Hickerson, M. J., Aleixo, A., Cadena, C. D., ... Brumfield, R. T. (2014). The drivers of tropical speciation. *Nature*, 515(7527), 406. <https://doi.org/10.1038/nature13687>
- Smith, S. A., & Brown, J. W. (2018). Constructing a broadly inclusive seed plant phylogeny. *American Journal of Botany*, 105(3), 302–314. <https://doi.org/10.1002/ajb2.1019>
- Sork, V. L., Nason, J., Campbell, D. R., & Fernandez, J. F. (1999). Landscape approaches to historical and contemporary gene flow in plants. *Trends in Ecology and Evolution*, 14(6), 219–224. [https://doi.org/10.1016/S0169-5347\(98\)01585-7](https://doi.org/10.1016/S0169-5347(98)01585-7)
- Stein, A., Gerstner, K., & Kreft, H. (2014). Environmental heterogeneity as a universal driver of species richness across taxa, biomes and spatial scales. *Ecology Letters*, 17(7), 866–880. <https://doi.org/10.1111/ele.12277>
- Stevens, P. F. (2001). *Angiosperm Phylogeny Website*. Version 14, July 2017 [and more or less continuously updated since]. Retrieved from <http://www.mobot.org/MOBOT/research/APweb/>
- Symonds, M. R. E., & Blomberg, S. P. (2014). A primer on phylogenetic generalized least squares. In L. Z. Garamszegi (Ed.), *Modern phylogenetic comparative methods and their application in evolutionary biology* (pp. 105–130). Berlin, Heidelberg: Springer. [https://doi.org/10.1007/978-3-662-43550-2\\_5](https://doi.org/10.1007/978-3-662-43550-2_5)
- Tello-Ramos, M. C., Hurly, T. A., & Healy, S. D. (2015). Traplining in hummingbirds: Flying short-distance sequences among several locations. *Behavioral Ecology*, 26(3), 812–819. <https://doi.org/10.1093/beheco/arv014>
- Templeton, A. R. (1981). Mechanisms of speciation - A population genetic approach. *Annual Review of Ecology and Systematics*, 12(1), 23–48. <https://doi.org/10.1146/annurev.es.12.110181.000323>
- ter Steege, H., Pitman, N. C. A., Sabatier, D., Baraloto, C., Salomao, R. P., Guevara, J. E., ... Silman, M. R. (2013). Hyperdominance in the Amazonian tree flora. *Science*, 342(6156), 1243092. <https://doi.org/10.1126/science.1243092>
- Theim, T. J., Shirk, R. Y., & Givnish, T. J. (2014). Spatial genetic structure in four understory *Psychotria* species (Rubiaceae) and implications for tropical forest diversity. *American Journal of Botany*, 101(7), 1189–1199. <https://doi.org/10.3732/ajb.1300460>
- Tukey, J. W. (1970). *Exploratory data analysis: Limited*, preliminary ed. Boston, MA: Addison-Wesley Publishing Company.
- Vekemans, X., & Hardy, O. J. (2004). New insights from fine-scale spatial genetic structure analyses in plant populations. *Molecular Ecology*, 13(4), 921–935. <https://doi.org/10.1046/j.1365-294X.2004.02076.x>
- Wallace, A. R. (1854). On the monkeys of the Amazon. *Annals and Magazine of Natural History*, 14(84), 451–454. <https://doi.org/10.1080/037454809494374>
- Weir, B. S., & Cockerham, C. C. (1984). Estimating F-Statistics for the analysis of population structure. *Evolution*, 38(6), 1358. <https://doi.org/10.2307/2408641>
- Willing, E.-M., Dreyer, C., & van Oosterhout, C. (2012). Estimates of genetic differentiation measured by  $F_{ST}$  do not necessarily require large sample sizes when using many SNP markers. *PLoS One*, 7(8), e42649. <https://doi.org/10.1371/journal.pone.0042649>
- Wright, S. (1943). Isolation by distance. *Genetics*, 28(2), 114–138.
- Wright, S. (1951). The genetical structure of populations. *Annals of Eugenics*, 15(1), 323–354. <https://doi.org/10.1111/j.1469-1809.1949.tb02451.x>
- Wright, S. (1965). The interpretation of population structure by F-statistics with special regard to systems of mating. *Evolution*, 19(3), 395–420. <https://doi.org/10.1111/j.1558-5646.1965.tb01731.x>

- Wright, S. (1978). *Evolution and the genetics of populations: a treatise in four volumes: Vol. 4: Variability within and among natural populations*. Chicago, IL: University of Chicago Press.
- Wright, S. I., Ness, R. W., Foxe, J. P., & Barrett, S. C. H. (2008). Genomic consequences of outcrossing and selfing in plants. *International Journal of Plant Sciences*, 169(1), 105–118. <https://doi.org/10.1086/523366>

**How to cite this article:** Gamba D, Muchhala N. Global patterns of population genetic differentiation in seed plants. *Mol Ecol*. 2020;29:3413–3428. <https://doi.org/10.1111/mec.15575>

#### SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.