

Rare genotype advantage promotes survival and genetic diversity of a tropical palm

Luke Browne^{1,2,3} and Jordan Karubian^{1,2}

¹Department of Ecology and Evolutionary Biology, Tulane University, 400 Lindy Boggs, New Orleans, LA 70118, USA; ²Foundation for the Conservation of the Tropical Andes, Quito, Ecuador; ³UCLA La Kretz Center for California Conservation Science, Institute of the Environment and Sustainability, University of California Los Angeles, Los Angeles, CA 90095, USA

Summary

Author for correspondence:

Luke Browne

Tel: +1 310 794 1431

Email: lukembrowne@gmail.com

Received: 5 December 2017

Accepted: 13 February 2018

New Phytologist (2018)

doi: 10.1111/nph.15107

Key words: forest ecology and evolution, genetic diversity, intraspecific variation, negative density dependence, negative frequency-dependent selection, plant recruitment, tropics.

- Negative density dependence, where survival decreases as density increases, is a well-established driver of species diversity at the community level, but the degree to which a similar process might act on the density or frequency of genotypes within a single plant species to maintain genetic diversity has not been well studied in natural systems.
- In this study, we determined the maternal genotype of naturally dispersed seeds of the palm *Oenocarpus bataua* within a tropical forest in northwest Ecuador, tracked the recruitment of each seed, and assessed the role of individual-level genotypic rarity on survival.
- We demonstrate that negative frequency-dependent selection within this species conferred a survival advantage to rare maternal genotypes and promoted population-level genetic diversity. The strength of the observed rare genotype survival advantage was comparable to the effect of conspecific density regardless of genotype.
- These findings corroborate an earlier, experimental study and implicate negative frequency-dependent selection of genotypes as an important, but currently underappreciated, determinant of plant recruitment and within-species genetic diversity. Incorporating intraspecific genetic variation into studies and theory of forest dynamics may improve our ability to understand and manage forests, and the processes that maintain their diversity.

Introduction

Understanding the processes that generate and maintain biodiversity at all levels – from genes, to species, to ecosystems – is a long-standing goal in the field of plant biology, made all the more urgent by anthropogenic threats to biodiversity. In tropical and temperate forests, patterns of species diversity are strongly influenced by the local density of members of the same species (i.e. conspecifics), as a result of well-documented impacts of competition or susceptibility to attack by pathogens and predators (Harms *et al.*, 2000; Volkov *et al.*, 2005; Wills *et al.*, 2006; Metz *et al.*, 2010; Bagchi *et al.*, 2014; Comita *et al.*, 2014; LaManna *et al.*, 2017). Much less is known about the degree to which within-species genetic diversity is influenced by similar processes, whereby the density and rarity of genotypes (i.e. different genetic variants within a single species) mediate local biotic interactions and survival, and in turn impact population-level patterns of genetic diversity. This represents a critical gap in knowledge because genetic diversity underlies adaptive potential and can also impact patterns of species diversity and overall ecosystem function (Whitham *et al.*, 2006; Johnson & Stinchcombe, 2007; Hughes *et al.*, 2008; Bailey *et al.*, 2009).

One of the most powerful processes that maintains biodiversity in natural systems is negative frequency- or density-dependent selection, which occurs when rarity provides a relative advantage

in survival or fitness (Ayala & Campbell, 1974; Chesson, 2000; HilleRisLambers *et al.*, 2012). At the species level, negative density dependence is commonly shown through reduced seed and seedling survival in areas with high densities or close proximity to conspecifics ('Janzen-Connell effects') (Janzen, 1970; Connell, 1971). This stabilizes the population sizes of rare species, prevents competitive exclusion by common species, and promotes overall degree of species diversity (Chesson, 2000). At the level of genotypes within a single species, negative frequency-dependent selection has long been theorized to serve as a potentially important ecological and evolutionary process that could maintain degrees of genetic diversity within populations by providing a relative advantage to rare genotypes (Haldane, 1949; Ayala & Campbell, 1974; Antonovics, 1976). Yet, although genotypic negative frequency-dependent selection has been shown to occur in agricultural and plant breeding contexts (Burdon, 1987; Barrett, 1988; Thompson & Burdon, 1992), our understanding of the impact of frequency-dependent selection of genotypes remains limited for plants in natural systems (Whitham *et al.*, 2006; Liu *et al.*, 2012; Thrall *et al.*, 2012; Browne & Karubian, 2016). Determining the degree to which genotypic rarity impacts plant survival and subsequent patterns of intraspecific genetic variation would improve our ability to understand and manage the processes driving patterns of recruitment and genetic diversity in forests.

In a previous study of the canopy palm *Oenocarpus bataua* in the Chocó rainforests of northwest Ecuador, we found that experimentally planted seedlings with rare genotypes exhibited a survival advantage relative to seedlings with common genotypes (Browne & Karubian, 2016), providing one of the first examples of negative frequency-dependent selection of seedling genotypes in a tropical forest. The seedlings used in the experiment were germinated in a shade house after collecting seeds directly from maternal trees and planted in equal densities across plots, which captured natural pollination dynamics, but by design omitted the influence of seed dispersal dynamics and natural variation in the density of nearby conspecifics. Thus, several significant questions remain unanswered – are the effects of genotypic frequency-dependent selection relevant for naturally dispersed and recruiting seeds and seedlings, or are they only detectable in experimental settings? Is the strength of genotypic frequency-dependent selection comparable to the strength of conspecific negative density dependence, one of the most pervasive and well-studied demographic processes in tropical forests? In addition, because in the previous study seedlings were planted as seedlings, we were not able to assess dynamics across the seed to seedling transition, a critical demographic transition in tropical forests (Harms *et al.*, 2000).

In the current study, we tracked the recruitment of naturally dispersed seeds of *O. bataua* across a 5 yr time period to improve our understanding of these issues in a natural context. We hypothesized that rare genotypes would show a survival advantage over common genotypes, which would lead to overall higher degrees of genetic diversity than in the absence of a rare genotype survival advantage. Specifically, we assessed the degree to which survival was associated with the genetic rarity of the seed's maternal genotype relative to the maternal genotypes of other dispersed seeds and adults in our study population; and the genetic relatedness of the seed's maternal genotype relative to the nearest adult conspecific where the seed was deposited. These complementary metrics provide information on the relative genetic rarity of a seed at the level of the population and at more local spatial scales. We next evaluated the relative importance of genetic rarity and genetic relatedness of the maternal genotype on seed and seedling survival compared with factors generally shown to be important for plant recruitment (Comita *et al.*, 2014): the density of nearby conspecific seeds and seedlings regardless of genetic identity; spatial distance to the nearest adult conspecific regardless of genetic identity; and seed dispersal distance, where seeds dispersed farther from adult trees are expected to experience higher survival (Howe & Smallwood, 1982). Finally, we investigated the consequences of a rare genotype survival advantage on population-level genetic diversity.

Materials and Methods

Study species and area

This study took place at Bilsa Biological Station (BBS; 00°21'33" N, 79°42'02" W; 300–750 m above sea level), a 3300 ha private forest reserve in Esmeraldas province, northwest Ecuador. The

focal study species *Oenocarpus bataua* is a hyperabundant, generalist canopy palm that grows to 20–40 m in height (Henderson *et al.*, 1995; ter Steege *et al.*, 2013). Adult *O. bataua* in BBS show low spatial genetic structure, with low genetic relatedness among neighbors (Ottewell *et al.*, 2012). *O. bataua* produces large infructescences with thousands of fruits, each containing a single large seed (39.7 ± 3.4 mm length, 22.7 ± 2.2 mm width, mean \pm SD, $n = 394$) covered by a thin, lipid-rich aril. Primary seed dispersers that remove fruit directly from the infructescence of *O. bataua* in BBS include the long-wattled umbrellabird (*Cephalopterus penduliger*), toucans (*Ramphastos spp.*), squirrels (*Sciurus spp.*), with occasional dispersal by small rodents, kinkajous (*Potos flavus*), and oilbirds (*Steatornis caripensis*) (J. Karubian & L. Browne, unpublished data). Relatively little is known about potential predators, pathogens, or herbivores of *O. bataua*, although preliminary data suggest that major seed predators are bark beetles (Scolytidae; L. Browne & J. Karubian, unpublished).

Sample collection and censuses

Adults Within a 235 ha study area in BBS, we mapped all *O. bataua* adults (Fig. 1). This includes adults sampled in the core 130 ha study parcel described in Ottewell *et al.* (2012), individuals that have transitioned to adults since 2012, and the inclusion of 45 new adults within a 200 m buffer area surrounding the original 130 ha study parcel sampled for this study (Fig. 1), for a total of 242 adults. Leaf, bark, or root tissue was taken from each adult for genetic analysis (see later).

Seeds In our core 130 ha study parcel, we placed a total of 390 1×1 m seed traps in groups of one to four throughout the study parcel (Fig. 1). Of the 390 traps, we placed 56 traps within an active umbrellabird lek in groups of one to two traps to cover the area where umbrellabirds were actively displaying (Karubian *et al.*, 2010; Ottewell *et al.*, 2018); there was no difference in survival among seeds arriving at lek vs nonlek traps (see also Karubian *et al.*, 2016), so we pooled these samples together. We placed 39 traps as singletons < 30 m from an adult *O. bataua* to increase sampling at short dispersal distances. The remaining 295 traps were placed in groups of four at random locations within the study parcel. From August 2011 to July 2016, we checked seed traps monthly. When an *O. bataua* seed was found in a trap, we carefully removed a small portion of the fibrous outer seed coat that contains only maternal DNA for genetic parentage analysis (see later). We were not able to obtain tissue containing both maternal and paternal genetic contributions without destructively sampling the seed to access the embryo, which would make it impossible to track survival across the seed-to-seedling transition. In light of this, we focus in this study on the maternal genotype of each seed, which has been shown to be a predictor of the genetic rarity of the diploid leaf seedling tissue (Browne & Karubian, 2016). Taking a genetic sample from the outer seed coat does not affect *O. bataua* germination rates (Supporting Information Table S1). After taking a genetic sample, we then placed the seed on the forest floor to mimic primary dispersal.

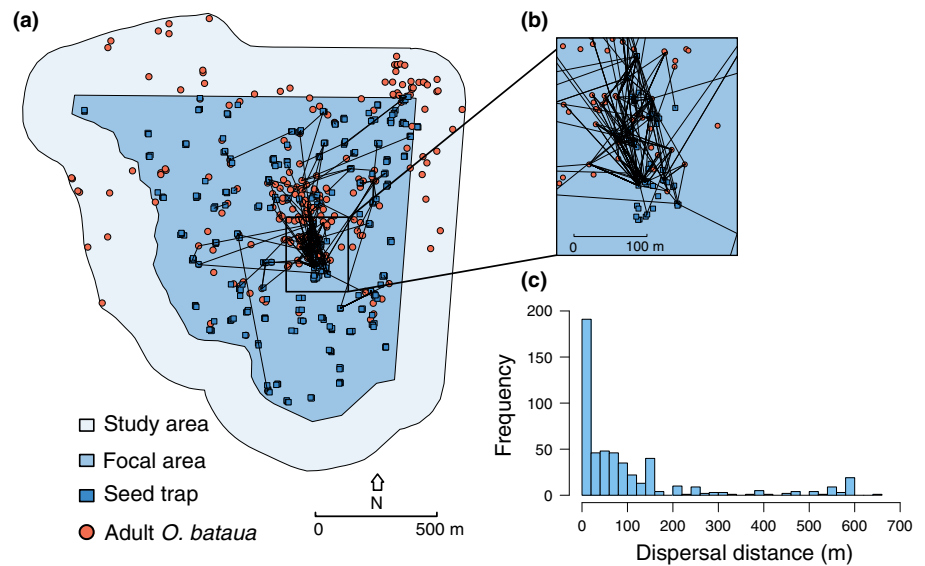


Fig. 1 Map of study area and seed dispersal events. (a) Study area (235 ha) located in Bilsa Biological Reserve, Esmeraldas Province, Ecuador, where all adult *Oenocarpus bataua* palm trees are mapped and genotyped (circles) and 135 ha focal study area with 390 seed traps (squares); (b) inset of area with high density of dispersal events; (c) frequency histogram of dispersal distances. The lines connect seed dispersal events from maternal tree to seed trap.

Seeds were placed 1–2 m away from the seed trap within an area of *c.* 2 m² adjacent to, but not beneath, the seed trap from which they were encountered to minimize microhabitat differences among individuals. The total number of seeds placed beside a given trap was determined by the number of seeds naturally dispersed into that trap. We attached each seed to a flag by wrapping metal wire once around the seed to prevent secondary dispersal, but despite this precaution, a portion of seeds (74/525, 14%) were presumably removed by secondary dispersal by either animals or rain, and their fate is unknown. We conducted censuses approximately yearly (October 2012, July 2013, June 2014, July 2015, July 2016), where we assessed whether each individual was alive or dead.

Genotyping. We extracted genomic DNA from seed, leaf, bark, and root tissue with a Qiagen DNeasy kit and screened a total of 10 microsatellite loci (Table S2) following the protocol described in Ottewell *et al.* (2012). In this study, our goal was to use an array of microsatellite loci that are assumed to be selectively neutral to sample the genome-wide genetic rarity of individuals, rather than identify specific loci under selection as in Browne & Karubian (2016). However, we acknowledge that in some cases, microsatellite loci may be under selection because of linkage with other loci under selection or because they have a functional role in gene expression (Koshi & King, 2006; Selkoe & Toonen, 2006; Putman & Carbone, 2014). We obtained multilocus genotypes for 237/242 adults in the study area, and 499/525 seeds, after culling samples with missing data at five or more loci. Samples that were missing data at five or more loci generally amplified poorly during PCR, producing unreliable allele calls, and were thus excluded from analyses. We checked for evidence of null alleles, linkage disequilibrium (LD), and deviation from Hardy–Weinberg equilibrium across unique maternal genotypes in the study and removed one locus from our dataset as a result of LD, leaving us with a suite of nine unlinked loci (Table S2). Locus diversity, probability of identity, and genotyping error rates are available in Table S2.

Maternal genotype matching and seed dispersal distance

Because the outer seed coat tissue of a seed contains the full diploid DNA of the mother tree (Godoy & Jordano, 2001; Jones *et al.*, 2005), we were able to assess whether dispersed seeds originated from one of the adults in our study area using direct genotype matching. We matched the genotypes of seeds to their maternal tree using the ‘identity match’ algorithm in CERVUS v.3.0.7 (Kalinowski *et al.*, 2007), with a minimum of five loci needed for a match and allowing for at most two loci mismatch, to account for possible genotyping error. We matched 97% (484/499) of seeds to at least one maternal tree in our study area. For seeds with matches to a single adult in the study area ($n = 481$), we calculated dispersal distance as the Euclidian distance between the adult source tree and the seed trap where the seed fell. For seeds that were matched to multiple maternal trees ($n = 3$), we assumed that the closest adult tree was the source tree, which would conservatively underestimate dispersal distances (Nathan & Muller-Landau, 2000). For seeds that were not matched with an adult in the study area ($n = 15$), we assumed these seeds originated outside the study parcel and calculated a minimum dispersal distance as the Euclidian distance between the seed trap where the seed fell and the nearest border of the study area, which again would conservatively underestimate dispersal distance (Jones *et al.*, 2005). Alternatively, these seeds could have originated from the five adults in our study parcel that were not genotyped. Excluding ‘immigrant’ seeds from analyses did not change qualitative results (Table S3).

Genetic rarity and genetic relatedness

Following the genetic rarity index developed by Browne & Karubian (2016), we estimated individual-level genetic rarity of the maternal genotype of each dispersed seed by averaging across the frequencies of alleles within each seed’s maternal genotype relative to a reference population. We calculated the reference allele frequencies based on the maternal genotypes of all seeds

genotyped in our study ($n=499$), which assesses the genetic rarity of a seed's maternal genotype in comparison to other maternal genotypes with which a seed is likely to interact directly or indirectly. We chose to use the cohort of dispersed seeds as our reference population rather than the overall adult population because only a subset of adults produced fruit during the study period, and we do not expect the genotype of the adults that did not fruit to play a strong role in the survival of dispersed seeds in this study. However, we do provide results of the survival analysis with genetic rarity calculated based on the genotypes of all unique maternal genotypes in our study area for reference. We averaged across the relative frequencies (in relation to the reference population) of each allele within each seed genotype, to calculate the genetic rarity index of individual i following the equation:

$$\text{Genetic rarity index}_i = \frac{\sum_{l=1}^L (p_{a_scaled} + p_{b_scaled})_l}{2L}$$

where p_{a_scaled} and p_{b_scaled} are the scaled relative frequencies of alleles a and b at locus l in the reference population, and L is the total number of loci (Browne & Karubian, 2016). Allele frequencies at each locus were scaled to a mean of zero and unit variance to ensure all loci were equally weighted in the calculation of genetic rarity. In the case of homozygotes, $p_{a_scaled} = p_{b_scaled}$. Note that because we are genotyping the maternal seed tissue, the genetic rarity index in this study represents the relative genetic rarity of the maternal tree of each seed, which has been shown to be a significant predictor of the genetic rarity of the biparentally inherited genotype (Browne & Karubian, 2016).

We estimated the genetic relatedness of the maternal genotype of each seed and its nearest adult conspecific in the study area. We calculated the kinship coefficient of Loiselle *et al.* (1995) using all unique maternal genotypes sampled in this study to generate the reference allele frequencies in the program SPAGED1 v.1.5 (Hardy & Vekemans, 2002). We also calculated genetic relatedness using the estimators of Queller & Goodnight (1989) and Ritland (1996), along with the kinship estimator of Lynch & Ritland (1999). Model results were robust to the choice of relatedness estimator (Table S4), and the results using F_{ij} of Loiselle *et al.* (1995) are presented in the main text.

Survival model

We modeled survival (coded binary as 0 = dead or 1 = alive) using a binomial regression and a complementary log-log link function (Paine *et al.*, 2012; Zhu *et al.*, 2015) with the following predictor variables: genetic rarity index of the seed's maternal genotype, genetic relatedness of the seed's maternal genotype to nearest adult conspecific, density of conspecifics in the study alive in the plot within a 30 d range at each census point, distance to nearest adult conspecific (m), seed dispersal distance (m), and age in yr. We investigated 30, 60, 90 and 180 d ranges for estimating conspecific density and found the 30 d range to be the best overall predictor of survival based on Akaike's information criterion (AIC) values (30 d threshold

AIC = 1052.9; 60 d threshold AIC = 1058.4; 90 d threshold AIC: 1058.1; 180 d threshold AIC: 1059.7). Because we were interested in how the effects on survival of genetic rarity of the seed's maternal genotype, genetic relatedness of the seed's maternal genotype to the nearest adult conspecific, conspecific density, distance to adult conspecific, and seed dispersal distance potentially change depending on the age of an individual, we also assessed interactions between these variables and age. We treated plot (i.e. the seed trap where the seed fell) and year of census as random effects to account for spatial and temporal variation in survival. We also included an individual-level random effect to account for repeated measures of the same individual across multiple censuses. The model followed the formula:

$$\begin{aligned} p_{ik} &= \ln(-\ln(1 - \hat{y}_{ik})) \\ &= \alpha + \alpha_{\text{trap}_p} + \alpha_{\text{year}_y} + \alpha_{\text{ind}_i} + \beta_1 \times \text{RarityIndex}_i + \beta_2 \\ &\quad \times \text{RelatednessNN}_i + \beta_3 \times \text{Density}_{ik-1} + \beta_4 \times \text{DistNN}_i \\ &\quad + \beta_5 \times \text{DispDist}_i + \beta_6 \times \text{Age}_{ik-1} + \beta_7 \times \text{RarityIndex}_i \\ &\quad \times \text{Age}_{ik-1} + \beta_8 \times \text{RelatednessNN}_i \times \text{Age}_{ik-1} + \beta_9 \\ &\quad \times \text{Density}_{ik-1} \times \text{Age}_{ik-1} + \beta_{10} \times \text{DistNN}_i \times \text{Age}_{ik-1} \\ &\quad + \beta_{11} \times \text{DispDist}_i \times \text{Age}_{ik-1} \end{aligned}$$

where p_{ik} is the predicted survival probability of individual i at the k th census, α is the global intercept, α_{trap_p} is the random effect term of the p th seed trap, α_{year_y} is the random effect term of the y th year, α_{ind_i} is the random effect term of individual i , β_1 is the effect of genetic rarity, β_2 is the effect of genetic relatedness to the nearest adult conspecific, β_3 is the effect of density of conspecific seeds or seedlings at the time of the previous census $k-1$, β_4 is the effect of distance to nearest adult conspecific, β_5 is the effect seed dispersal distance, and β_6 is the effect of age at the time of the previous census $k-1$. β_7 – β_{11} are the interaction terms that estimate how the slopes of genetic rarity, genetic relatedness to nearest conspecific adult, conspecific density, and distance to nearest adult conspecific vary with age. Because census intervals were not equally spaced across individuals, we included a time offset of log-transformed years to produce survival estimates in terms of yearly survival probability (Egli & Schmid, 2001; Paine *et al.*, 2012; Zhu *et al.*, 2015). For the main survival model, we excluded seedlings that had an unknown survival status, that were missing data at five or more loci, and that had not been observed for at least 30 d (*O. batava* germination from planting in our study area takes $c.$ 1 month), leading to a total sample size of 416 seedlings observed a total of 762 times in 87 plots across 6 yr. We mean-centered and standardized covariates to unit variance to facilitate comparison of coefficient estimates and aid in model convergence (Schielzeth, 2010; Bolker *et al.*, 2013). The mean and SD of each covariate is available in Table S5 to facilitate back-transformation. There was no evidence of strong collinearity among predictor variables: the variance inflation factor (VIF) was <1.3 for all parameters, below the threshold of VIF = 3 suggested by Zuur *et al.* (2010). We fitted the model to our data using the R package 'lme4' (Bates *et al.*, 2015).

Comparing genetic diversity

To test whether a rare genotype survival advantage led to increased genetic diversity, we calculated population-level gene diversity (Nei, 1987) (H_s) of the maternal genotype of seedlings that were alive at the time of last census or dead at the time of last census and compared these estimates with the expected degrees of gene diversity if survival was random with respect to genotype. To generate the null distribution of gene diversity, we randomly shuffled whether a seedling was alive or dead at time of last census 1000 times. If observed gene diversity did not overlap with the 95% confidence interval (CI) of the null distribution, we rejected the null hypothesis that seedling survival was random with respect to genotype. We calculated gene diversity with the R package 'HIERFSTAT' (Goudet, 2005; R Core Team, 2015).

Simulation model

We parameterized a simulation model based on our study population of *O. batavia* to test the effects of the impacts of genetic rarity, genetic relatedness to the nearest conspecific, conspecific density, distance to nearest adult conspecific, and age on population-level gene diversity (Nei, 1987) (H_s). Simulations began in a landscape populated by all adults genotyped in our 235 ha study area ($n = 237$ total), with their spatial locations and genotypes mirroring their spatial locations and genotypes in our study area (Fig. 1). The simulated landscape also included an array of 390 seed traps (1×1 m) that mirror exactly the locations of seeds traps in our study area (Fig. 1). To simulate seed rain, we calculated the probability of a seed arriving in a given seed trap from a given adult based on the Euclidian distance of the seed trap to the adult. Dispersal probabilities were modeled from a log-normal distribution with mean = 3.72 and SD = 1.45 in the R package 'FITDISTRPLUS' (Delignette-Muller & Dutang, 2015), which provided the best fit to our observed dispersal distances (lognormal AIC = 5784) compared with either Weibull (AIC = 5844), Gamma (AIC = 5870) or exponential (AIC = 5949) distributions.

Using the predicted dispersal probabilities, we estimated the number of seeds arriving in a trap from a given adult with a binomial distribution, with the number of trials equal to the number of dispersed seeds for each adult ($n = 250$), and the dispersal probability. We estimated a total of 250 dispersed seeds from each adult roughly based on direct observation of fruiting *O. batavia* in our study area and the average duration of an infructescence (J. Karubian & L. Browne, unpublished), though varying the number of dispersed seeds per adult did not qualitatively change our results. Using the spatial coordinates of both simulated seed traps and adults, we then calculated the distance to the nearest adult conspecific. Because adults in the simulated landscape had equivalent genotypes to our study population, we were able to calculate the genetic relatedness using F_{ij} (Loiselle *et al.*, 1995) of each simulated seed to the nearest adult conspecific and also the corresponding genetic rarity index of each seed's maternal genotype.

Once seed rain was simulated across the landscape, we estimated the cumulative survival probability of each seed to 5 yr

based on mean parameter estimates of our survival model. Simulations were performed in four scenarios: a scenario that included the effects of age on survival only ('age only'); a scenario that included age effects along with the effects of conspecific density and distance to nearest adult conspecific ('JC' for Janzen–Connell effects); a scenario that included age effects along with the effects of genetic rarity and genetic relatedness to the nearest adult conspecific ('NFDS' for negative frequency-dependent selection); a scenario that includes all estimated effects – age effects, JC effects, and negative frequency-dependent selection effects ('NFDS + JC'). Note that the effects of distance to the nearest adult conspecific and genetic relatedness to the nearest adult conspecific varied in the simulation with age, as there was a significant age interaction in our survival model. We did not include any effect of seed dispersal distance on survival because it was not a significant predictor of survival in our observed data. Each scenario was replicated 1000 times to capture variability across runs.

Data availability

Genotype and survival data from the individuals included in this study have been made publicly available at <https://doi.org/10.6084/m9.figshare.4987184>.

Results

Using an array of 390 seed traps in a 235 ha study area (Fig. 1), we recovered 499 naturally dispersed *O. batavia* seeds from 2011 to 2016. We found strong evidence for a rare genotype survival advantage that promotes genetic diversity for naturally dispersed and recruited seedlings of *O. batavia*. Consistent with predictions of genotypic negative frequency-dependent selection, individuals with rare maternal genotypes were more likely to survive than individuals with common maternal genotypes (Fig. 2a; Table 1). There was no significant interaction between the impact of genetic rarity of the seed or seedling's maternal genotype and age (Fig. 2a; Table 1). At local spatial scales, the genetic relatedness of the maternal genotype of a seed or seedling to the nearest adult conspecific also had a strong impact on survival probabilities, which varied with age (Fig. 2b; Table 1). At older ages (> 2 yr), individuals that had a more genetically similar maternal genotype to the nearest adult conspecific had lower survival probability than individuals that were less genetically similar (Fig. 2b). However, at younger ages (< 2 yr), yearly survival was higher if a seed or seedling had a more genetically similar maternal genotype to the nearest adult (Fig. 2b).

Survival was reduced in areas of high densities of conspecific seeds and seedlings, and close to adult conspecifics (Fig. 2c,d; Table 1). The negative effect on survival of being close to an adult strengthened with seedling age, while the negative effect of conspecific density remained similar across ages (Fig. 2c,d). The relative importance of the effects of genetic rarity and relatedness to the nearest adult of the maternal genotype was similar to that of density and distance to the nearest adult conspecific, as shown by the similar magnitude of the standardized regression coefficients (Table 1). By contrast, we found no significant relationship between seed dispersal distance and survival (Fig. 2e; Table 1).

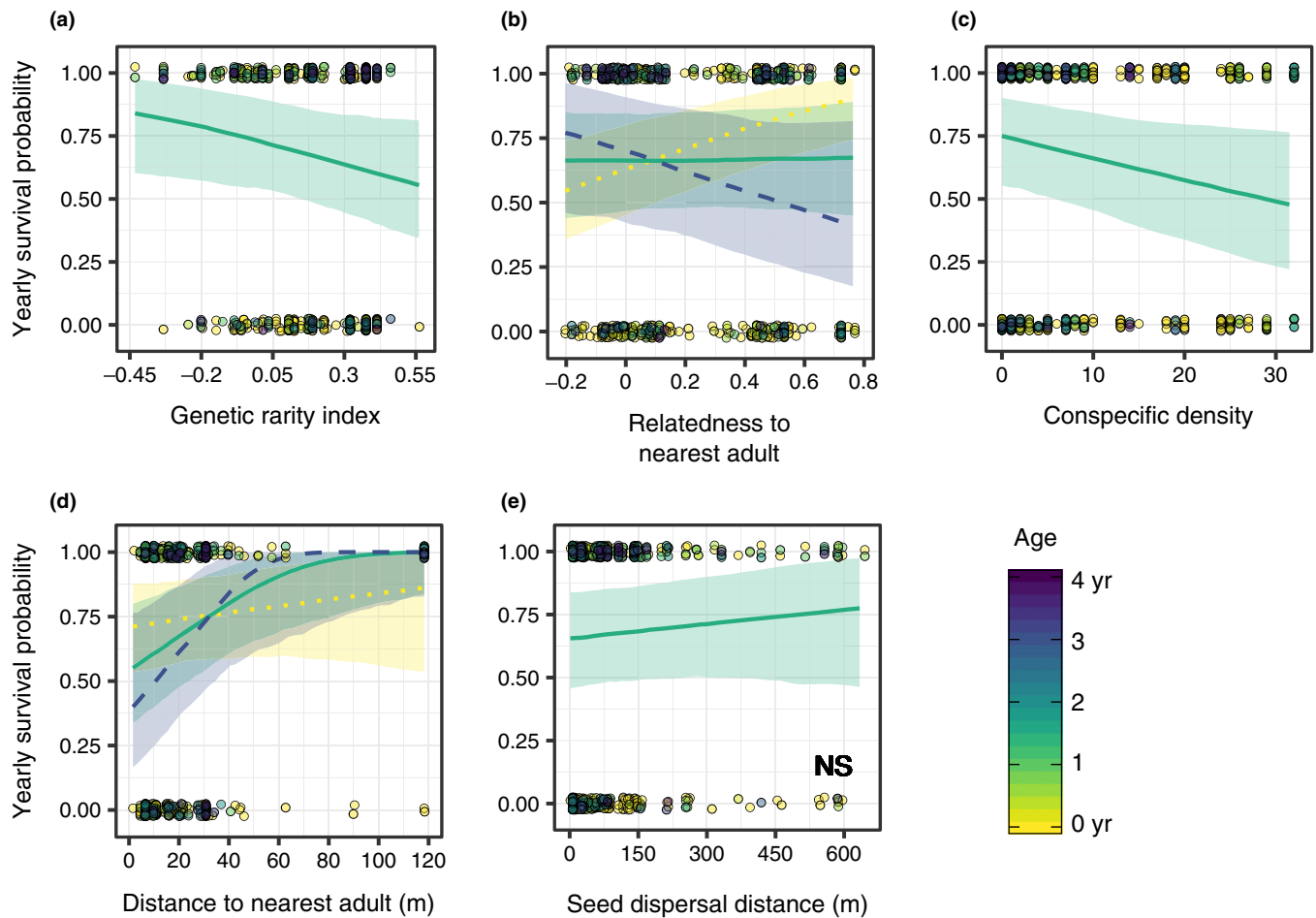


Fig. 2 Survival probability of *Oenocarpus bataua*. Relationship between yearly survival probability and: (a) genetic rarity of a seed's maternal genotype relative to other dispersed seeds – lower values of the genetic rarity index indicate genotypes that are relatively rare in the population; (b) genetic relatedness (F_{ij}) of a seed's maternal genotype to nearest adult conspecific; (c) density of conspecific seedlings; (d) distance to nearest adult conspecific (m); (e) seed dispersal distance (m). We show interactions of survival at ages 0 yr (yellow dotted line), 1.5 yr (green solid line), and 3 yr (purple dashed line) in (b) and (d), and survival probabilities at 1.5 yr (green solid line) in (a), (c) and (e). Circles show observed data jittered along the y-axis and colored by age. Lines show mean parameter estimates, and shaded areas show approximate 95% confidence intervals. 'NS' indicates a nonsignificant statistical relationship between seed dispersal distance and survival probability. Model parameter estimates are available in Table 1.

We next assessed the degree to which a rare genotype survival advantage leads to increased genetic diversity both in our study population and in simulated populations. As predicted by genotypic negative frequency-dependent selection, gene diversity of maternal genotypes was higher among seedlings that survived and lower among seedlings that died than would be expected if survival was random with respect to genotype (Fig. 3a). Simulations confirmed that genetic diversity of maternal genotypes was higher in scenarios where survival was driven by genetic rarity and genetic relatedness to the nearest adult than when only age or Janzen-Connell effects were present (Fig. 3b).

Discussion

Using the maternal genotypes of naturally recruiting seedlings of the palm *O. bataua*, we found that a rare genotype survival advantage led to increased population-level genetic diversity, as predicted under negative frequency-dependent selection. These findings

provide empirical evidence that genotypic negative frequency-dependent selection may be an important, but currently underappreciated, driver of tropical plant survival that can help to explain patterns of nonrandom mortality among recruiting seedlings and overall patterns of genetic diversity at the population level.

Among *O. bataua* seeds and seedlings, the strength of the rare genotype survival advantage of maternal genotypes we documented is comparable in magnitude to more commonly studied effects of density and proximity to conspecifics, the so-called 'Janzen-Connell' effects (Harms *et al.*, 2000; Comita *et al.*, 2014). Interestingly, our results indicate that seed dispersal distance *per se*, which is often used as a proxy for conspecific density and proximity, has little direct impact on survival in this system. Instead, the context of the location into which a seed is dispersed (i.e. the density, proximity, and genetic identity of conspecific neighbors) took priority over advantages of being dispersed long distances. These findings point to the potential importance of frugivore foraging ecology and other behaviors that shape seed

Table 1 Regression coefficients of survival model for *Oenocarpus bataua*

Parameter	Genetic rarity calculated relative to other seeds				Genetic rarity calculated relative to adults			
	Estimate	SE	Z	P	Estimate	SE	Z	P
Intercept	0.178	0.251	0.709	0.478	0.176	0.245	0.716	0.474
Age	-0.125	0.108	-1.155	0.248	-0.126	0.108	-1.161	0.245
Genetic rarity	-0.209	0.068	-3.075	0.002	-0.037	0.069	-0.533	0.594
Genetic rarity × age	0.094	0.058	1.632	0.103	-0.061	0.062	-0.974	0.330
Relatedness to nearest adult conspecific	0.142	0.085	1.673	0.094	0.062	0.089	0.694	0.488
Relatedness to nearest adult conspecific × age	-0.209	0.077	-2.693	0.007	-0.192	0.079	-2.423	0.015
Conspecific density	-0.235	0.099	-2.384	0.017	-0.270	0.096	-2.802	0.005
Conspecific density × age	-0.003	0.084	-0.038	0.970	0.025	0.081	0.303	0.762
Distance to nearest adult conspecific	0.238	0.091	2.618	0.009	0.229	0.096	2.395	0.017
Distance to nearest adult conspecific × age	0.196	0.095	2.068	0.039	0.197	0.097	2.021	0.043
Dispersal distance	0.077	0.066	1.163	0.245	0.086	0.068	1.279	0.201
Dispersal distance × age	-0.009	0.065	-0.142	0.887	0.007	0.069	0.098	0.922

The estimated effects on survival of genetic rarity of the maternal genotype, genetic relatedness of the maternal genotype to the nearest adult conspecific, density of conspecifics, spatial distance to nearest adult conspecific, and seed dispersal distance. Interactions with age were also assessed for each parameter. Shown are separate models where genetic rarity of a seed's maternal genotype was calculated in reference to other sampled seeds or in reference to adults in the study population. The table shows standardized regression coefficients of survival model (binomial regression), with the parameter estimate, standard error (SE), Z-value, and P-values. Parameters with $P < 0.05$ are shown in bold. The mean and SD of covariates are provided in Supporting Information Table S5 to allow back-transformation to original units.

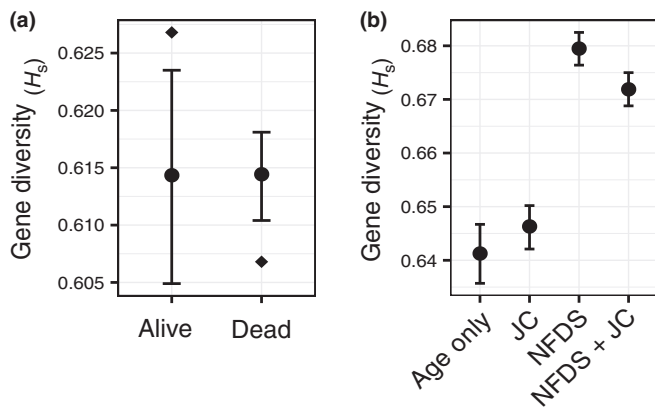


Fig. 3 Patterns of genetic diversity. (a) Comparison of gene diversity (H_s) of the maternal genotypes of recruited vs dead individuals in this study. Diamonds show observed values of gene diversity, and circles and bars show the mean and 95% range of values from a null model simulated 1000 times where survival was random with respect to genotype. (b) Genetic diversity of maternal genotypes of simulated 5-yr-old cohorts of *Oenocarpus bataua* where survival is dependent on age only, Janzen–Connell effects and age (JC), genetic rarity and relatedness and age (NFDS), or all effects combined (NFDS + JC). Circles show mean simulated values and lines show the 95% range of simulated values across 1000 repetitions.

deposition patterns and resulting degrees of clustering and genetic similarity among seeds away from source trees (García & Grivet, 2011; Scofield *et al.*, 2012). Moreover, the results of this study show that the genotypic identity of recruiting plants and nearby conspecifics, even if it is only the maternal genotype, can provide useful information to predict survival probabilities. Taken together, these results suggest that to achieve a complete understanding of the processes driving forest regeneration, it is appropriate to consider individual-level genetic variation in forest dynamics, along with the more commonly studied JC effects.

This represents a change in perspective in the field, which until recently has generally ignored such individual-level genetic variation (see later).

We were not able to determine the underlying mechanisms driving the observed rare genotype survival advantage but consider it likely that coevolution with soil-based microbes, which are key agents involved in plant growth and survival, may be involved. In agricultural and crop systems, it is common for some genotypes within a species to be more susceptible to highly specialized microbial pathogens than to others (Thompson & Burdon, 1992; Burdon *et al.*, 2006), with genotypes and their specialized pathogens engaging in a frequency-dependent coevolutionary arms race (Clay & Kover, 1996). In natural systems, soilborne microbial pathogens have been implicated in driving negative density dependence at the species level (Bell *et al.*, 2006; Bagchi *et al.*, 2014; Bever *et al.*, 2015) and among phylogenetically closely related species (Liu *et al.*, 2012; Gilbert & Parker, 2016), but few studies have shown the role of microbes in causing genotypic frequency or density-dependent selection within a species (Liu *et al.*, 2015). The occurrence and overall influence of genotype-specific microbial pathogens in natural systems remain unclear. However, we cannot rule out competition among genotypes as another mechanism that would produce this same pattern, and thus further experiments are needed to determine the exact mechanism involved in producing a rare genotype survival advantage.

Interestingly, we found that the impact of genetic relatedness of a seed or seedling's maternal genotype to the nearest conspecific adult on survival was dependent on seedling age, in that a 'home-field advantage' experienced at young ages reversed over time. This pattern could be caused by beneficial microbes passed from the mother plant to nearby offspring promoting survival at young ages, and shared pathogens reducing survival probability at older ages (Van der Putten *et al.*, 2013). To the extent that

these patterns are driven by microbial interactions, these results imply that different genotypes of *O. bataua* may have associations with coevolved microbial communities, which are then persistent across generations (Whitham *et al.*, 2006, 2008). Studies in common gardens of *Populus angustifolia* from different populations have shown that soil microbial community composition varies predictably with plant genotype (Schweitzer *et al.*, 2008), but our results suggest that this effect may also occur on much finer spatial scales within a single interbreeding population. However, we currently lack evidence that microbes are responsible for the observed effects, which at a minimum would require linking microbial community composition across space and time with differential survival of plant genotypes.

A limitation of this study is that we are lacking information on the biparentally inherited genotype of individuals, as a result of sequencing only the maternal tissue of each seed. As negative frequency-dependent selection and the genetic diversity generated through recombination are thought to play an important role in the evolution of sexual reproduction in plants (Antonovics & Ellstrand, 1984; Kelley, 1994), the lack of information about pollen donors represents a gap in knowledge of our understanding of how negative frequency-dependent selection of genotypes in this system relates to broader evolutionary patterns. In addition, for many species, including *O. bataua* (Karubian *et al.*, unpublished), pollen donors provide more allelic diversity to recruiting seedlings than seed donors (Sork *et al.*, 2015), which suggests that the movement of pollen both within and between populations may be an important avenue for the introduction and maintenance of rare alleles which may provide a survival advantage. We chose to sequence only maternal tissue because obtaining biparentally inherited tissue from a seed usually requires destructive sampling of the embryo, which would have prevented the examination of the relationship between genotype and survival in this study. Consequently, our estimates of the strength of the rare genotype survival advantage may be conservative, as they only include half the information contained in biparentally inherited tissue. Results from previous work, however, indicate that the maternal genotype of a seed provides useful information about the effects of genetic rarity on survival; our earlier experimental study found that genetic rarity of a seedling can be predicted by the genetic rarity of the maternal tree and that maternal genotypic rarity can predict seedling survival (Browne & Karubian, 2016).

Our results add to a growing appreciation of the importance of intraspecific genetic variation in the ecological dynamics of forests. For example, Marden *et al.* (2017) showed that populations of six tropical tree species with lower diversity in resistance genes had increased susceptibility to potentially pathogenic fungi and showed increased strength of conspecific negative density dependence. In subtropical China, soil microbiota were implicated as providing a survival advantage to seedlings of two species that were relatively genetically distinct compared with the population where the soil was collected (Liu *et al.*, 2015). And in our previous study on experimentally planted seedlings of *O. bataua*, individuals with rare genotypes showed a survival advantage over those with common genotypes (Browne & Karubian, 2016). In the current study, we show that the rare genotype survival

advantage also occurs among naturally dispersed and recruiting seeds and seedlings, a critical step in understanding the overall importance of genotypic negative frequency-dependent selection in natural forest regeneration processes by showing that the rare genotype survival advantage occurs in natural, nonexperimental conditions. Further research on a greater number of species at various life stages in both tropical and temperate forests is warranted to determine the ubiquity of genotypic negative frequency-dependent selection in driving plant recruitment in natural systems. There is also a need to extend our understanding of the degree to which a rare genotype advantage is driven by microbial interactions and the broader effects of a rare genotype advantage within and across trophic levels, given the potential impacts of genotypic diversity on species coexistence, species diversity, and ecosystem-level processes (Crutsinger *et al.*, 2006; Bolnick *et al.*, 2011; Ehlers *et al.*, 2016). Overall, incorporating intraspecific genetic variation and genetic rarity with existing theory of forest ecology and evolution is expected to improve our ability to understand, manage, and conserve these habitats in the face of continued threats.

Acknowledgements

We are grateful for the support of FCAT (Fundación para la Conservación de los Andes Tropicales), the Jatun Sacha Foundation, and the Ecuadorian Ministry of the Environment. We especially thank D. Cabrera for his enormous contribution to this project. We also thank J. Cabrera, F. Castillo, G. DiRenzo, M. Gonzalez, N. Gonzalez, K. Harms, J. Olivo, K. Ottewell, T. Smith, V. Sork, C. Taylor, S. Van Bael, and the helpful comments of two anonymous reviewers. This project was supported by the Conservation, Food & Health Foundation; the Disney Conservation Fund; the National Science Foundation (EAGER #1548548, DDIG #1501514 and a Graduate Research Fellowship to L.B.); the National Geographic Society; Tulane University; and the United States Fish & Wildlife Service (NMBCA #5605). All research was conducted with the approval of the Ecuadorian Ministry of the Environment (Permit 010-2015-IC-FLO-FAU-DPE-MA).

Author contributions

J.K. designed and supervised the project. L.B. performed laboratory work, assisted with field work, and analyzed the data. Both authors wrote the manuscript.

References

- Antonovics J. 1976. The input from population genetics: the new ecological genetics. *Systematic Botany* 5: 86–104.
- Antonovics J, Ellstrand NC. 1984. Experimental studies of the evolutionary significance of sexual reproduction. I. A test of the density-dependent selection hypothesis. *Evolution* 38: 103–115.
- Ayala FJ, Campbell CA. 1974. Frequency-dependent selection. *Annual Review of Ecology and Systematics* 5: 115–138.
- Bagchi R, Gallery RE, Gripenberg S, Gurr SJ, Narayan L, Addis CE, Freckleton RP, Lewis OT. 2014. Pathogens and insect herbivores drive rainforest plant diversity and composition. *Nature* 506: 85–88.

- Bailey JK, Schweitzer JA, Ubeda F, Koricheva J, LeRoy CJ, Madritch MD, Rehill BJ, Bangert RK, Fischer DG, Allan GJ *et al.* 2009. From genes to ecosystems: a synthesis of the effects of plant genetic factors across levels of organization. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences* 364: 1607–1616.
- Barrett JA. 1988. Frequency-dependent selection in plant-fungal interactions. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences* 319: 473–483.
- Bates D, Mächler M, Bolker BM, Walker SC. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67: 1–48.
- Bell T, Freckleton RP, Lewis OT. 2006. Plant pathogens drive density-dependent seedling mortality in a tropical tree. *Ecology Letters* 9: 569–574.
- Bever JD, Mangan S, Alexander HM. 2015. Maintenance of plant species diversity by pathogens. *Annual Review of Ecology, Evolution, and Systematics* 46: 305–325.
- Bolker BM, Gardner B, Maunder M, Berg CW, Brooks M, Comita L, Crone E, Cubaynes S, Davies T, de Valpine P *et al.* 2013. Strategies for fitting nonlinear ecological models in R, AD Model Builder, and BUGS. *Methods in Ecology and Evolution* 4: 501–512.
- Bolnick DI, Amarasekare P, Araújo MS, Bürger R, Levine JM, Novak M, Rudolf VHW, Schreiber SJ, Urban MC, Vasseur DA. 2011. Why intraspecific trait variation matters in community ecology. *Trends in Ecology and Evolution* 26: 183–192.
- Browne L, Karubian J. 2016. Frequency-dependent selection for rare genotypes promotes genetic diversity of a tropical palm. *Ecology Letters* 19: 1439–1447.
- Burdon JJ. 1987. *Diseases and plant population biology*. Cambridge, UK: Cambridge University Press.
- Burdon JJ, Thrall PH, Ericson AL. 2006. The current and future dynamics of disease in plant communities. *Annual Review of Phytopathology* 44: 19–39.
- Chesson P. 2000. Mechanisms of maintenance of species diversity. *Annual Review of Ecology and Systematics* 31: 343–366.
- Clay K, Kover PX. 1996. The Red Queen Hypothesis and plant/pathogen interactions. *Annual Review in Phytopathology* 34: 29–50.
- Comita LS, Queenborough SA, Murphy SJ, Eck JL, Xu K, Krishnadass M, Beckman N, Zhu Y. 2014. Testing predictions of the Janzen–Connell hypothesis: a meta-analysis of experimental evidence for distance- and density-dependent seed and seedling survival. *Journal of Ecology* 102: 845–856.
- Connell J. 1971. On the role of natural enemies in preventing competitive exclusion in some marine animals and in forest trees. In: Den Boer PJ, Gradwell GR, eds. *Dynamics of populations*. Wageningen, the Netherlands: Centre for Agricultural Publishing and Documentation, 298–312.
- Crutsinger GM, Collins MD, Fordyce JA, Gompert Z, Nice CC, Sanders NJ. 2006. Plant genotypic diversity predicts community structure and governs an ecosystem process. *Science* 313: 966–968.
- Delignette-Muller ML, Dutang C. 2015. fitdistrplus: an R package for fitting distributions. *Journal of Statistical Software* 64: 1–34.
- Egli P, Schmid B. 2001. The analysis of complex leaf survival data. *Basic and Applied Ecology* 2: 223–231.
- Ehlers BK, Damgard CF, Laroche F. 2016. Intraspecific genetic variation and species co-existence in plant communities. *Biology Letters* 12: 20150853.
- García C, Grivet D. 2011. Molecular insights into seed dispersal mutualisms driving plant population recruitment. *Acta Oecologica* 37: 632–640.
- Gilbert GS, Parker IM. 2016. The evolutionary ecology of plant disease: a phylogenetic perspective. *Annual Review of Phytopathology* 54: 549–578.
- Godoy JA, Jordano P. 2001. Seed dispersal by animals: exact identification of source trees with endocarp DNA microsatellites. *Molecular Ecology* 10: 2275–2283.
- Goudet J. 2005. HIERFSTAT, a package for R to compute and test hierarchical F-statistics. *Molecular Ecology Notes* 5: 184–186.
- Haldane JBS. 1949. Disease and evolution. *La Ricerca Scientifica Supplemento* 19: 1–11.
- Hardy O, Vekemans X. 2002. SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Notes* 2: 618–620.
- Harms KE, Wright SJ, Calderón O, Hernández A, Herre EA. 2000. Pervasive density-dependent recruitment enhances seedling diversity in a tropical forest. *Nature* 404: 493–495.
- Henderson A, Galeano G, Bernal R. 1995. *Field guide to the palms of America*. Princeton, NJ, USA: Princeton University Press.
- HilleRisLambers J, Adler PB, Harpole WS, Levine JM, Mayfield MM. 2012. Rethinking community assembly through the lens of coexistence theory. *Annual Review of Ecology, Evolution, and Systematics* 43: 227–248.
- Howe HF, Smallwood J. 1982. Ecology of seed dispersal. *Annual Review of Ecology and Systematics* 13: 201–228.
- Hughes AR, Inouye BD, Johnson MTJ, Underwood N, Vellend M. 2008. Ecological consequences of genetic diversity. *Ecology Letters* 11: 609–623.
- Janzen D. 1970. Herbivores and the number of tree species in tropical forests. *American Naturalist* 104: 501–528.
- Johnson MTJ, Stinchcombe JR. 2007. An emerging synthesis between community ecology and evolutionary biology. *Trends in Ecology and Evolution* 22: 250–257.
- Jones FA, Chen J, Weng G-J, Hubbell SP. 2005. A genetic evaluation of seed dispersal in the neotropical tree *Jacaranda copaia* (Bignoniaceae). *American Naturalist* 166: 543–555.
- Kalinowski ST, Taper ML, Marshall TC. 2007. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Molecular Ecology* 16: 1099–1106.
- Karubian J, Browne L, Cabrera D, Olivo J. 2016. Relative influence of relatedness, conspecific density and microhabitat on seedling survival and growth of an animal-dispersed Neotropical palm, *Oenocarpus bataua*. *Botanical Journal of the Linnean Society* 182: 425–438.
- Karubian J, Sork VL, Roorda T, Durães R, Smith TB. 2010. Destination-based seed dispersal homogenizes genetic structure of a tropical palm. *Molecular Ecology* 19: 1745–1753.
- Kelley SE. 1994. Viral pathogens and the advantage of sex in the perennial grass *Anthoxanthum odoratum*. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences* 346: 295–302.
- Koshi Y, King D. 2006. Simple sequence repeats as advantageous molecular markers – mutations in evolution. *Trends in Genetics* 22: 252–259.
- LaManna JA, Mangan SA, Alonso A, Bourg NA, Brockelman WY, Bunyavechewin S, Chang L, Chiang J, Chuyong GB, Clay K *et al.* 2017. Plant diversity increases with the strength of negative density dependence at the global scale. *Science* 356: 1389–1392.
- Liu X, Etienne RS, Liang M, Wang Y, Yu S. 2015. Experimental evidence for an intraspecific Janzen–Connell effect mediated by soil biota. *Ecology* 96: 662–671.
- Liu X, Liang M, Etienne RS, Wang Y, Staehelin C, Yu S. 2012. Experimental evidence for a phylogenetic Janzen–Connell effect in a subtropical forest. *Ecology Letters* 15: 111–118.
- Loiselle BA, Sork VL, Nason J, Graham C. 1995. Spatial genetic structure of a tropical understory shrub, *Psychotria officinalis* Rubiaceae. *American Journal of Botany* 82: 1420–1425.
- Lynch M, Ritland K. 1999. Estimation of pairwise relatedness with molecular markers. *Genetics* 152: 1753–1766.
- Marden JH, Mangan SA, Peterson MP, Wafula E, Fescemyer HW, Der JP, Comita LS. 2017. Ecological genomics of tropical trees: how local population size and allelic diversity of resistance genes relate to immune responses, co-susceptibility to pathogens, and negative density dependence. *Molecular Ecology* 26: 2498–2513.
- Metz MR, Sousa WP, Valencia R. 2010. Widespread density-dependent seedling mortality promotes species coexistence in a highly diverse Amazonian rain forest. *Ecology* 91: 3675–3685.
- Nathan R, Muller-Landau H. 2000. Spatial patterns of seed dispersal, their determinants and consequences for recruitment. *Trends in Ecology & Evolution* 15: 278–285.
- Nei M. 1987. *Molecular evolutionary genetics*. New York, NY, USA: Columbia University Press.
- Ottewell K, Browne L, Cabrera D, Olivo J, Karubian J. 2018. Genetic diversity of dispersed seeds is highly variable among leks of the long-wattled umbrellabird. *Acta Oecologica* 86: 31–37.
- Ottewell K, Grey E, Castillo F, Karubian J. 2012. The pollen dispersal kernel and mating system of an insect-pollinated tropical palm, *Oenocarpus bataua*. *Heredity* 109: 332–339.

- Paine CET, Norden N, Chave J, Forget P-M, Fortunel C, Dexter KG, Baraloto C. 2012. Phylogenetic density dependence and environmental filtering predict seedling mortality in a tropical forest. *Ecology Letters* 15: 34–41.
- Putman AJ, Carbone I. 2014. Challenges in analysis and interpretation of microsatellite data for population genetic studies. *Ecology and Evolution* 4: 4399–4428.
- Queller D, Goodnight K. 1989. Estimating relatedness using genetic markers. *Evolution* 43: 258–275.
- R Core Team. 2015. *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Ritland K. 1996. Estimators for pairwise relatedness and individual inbreeding coefficients. *Genetical Research* 67: 175–185.
- Schielzeth H. 2010. Simple means to improve the interpretability of regression coefficients. *Methods in Ecology and Evolution* 1: 103–113.
- Schweitzer JA, Bailey JK, Fischer DG, LeRoy CJ, Lonsdorf EV, Whitham TG, Hart SC, Lonsdorf EV, Whitham TG, Hart SC. 2008. Plant–soil–microorganism interactions: heritable relationship between plant genotype and associated soil microorganisms. *Ecology* 89: 773–781.
- Scofield DG, Smouse PE, Karubian J, Sork VL. 2012. Use of alpha, beta, and gamma diversity measures to characterize seed dispersal by animals. *American Naturalist* 180: 719–732.
- Selkoe KA, Toonen RJ. 2006. Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology Letters* 9: 615–629.
- Sork VL, Smouse PE, Grivet D, Scofield DG. 2015. Impact of asymmetric male and female gamete dispersal on allelic diversity and spatial genetic structure in valley oak (*Quercus lobata* Née). *Evolutionary Ecology* 29: 927–945.
- ter Steege H, Pitman NCA, Sabatier D, Baraloto C, Salomão RP, Guevara JE, Phillips OL, Castillo CV, Magnusson WE, Molino J-F *et al.* 2013. Hyperdominance in the Amazonian tree flora. *Science* 342: 325–342.
- Thompson JN, Burdon JJ. 1992. Gene-for-gene coevolution between plants and parasites. *Nature* 355: 242–244.
- Thrall PH, Laine AL, Ravensdale M, Nemri A, Dodds PN, Barrett LG, Burdon JJ. 2012. Rapid genetic change underpins antagonistic coevolution in a natural host–pathogen metapopulation. *Ecology Letters* 15: 425–435.
- Van der Putten WH, Bardgett RD, Bever JD, Bezemer TM, Casper BB, Fukami T, Kardol P, Klironomos JN, Kulmatiski A, Schweitzer JA *et al.* 2013. Plant–soil feedbacks: the past, the present and future challenges. *Journal of Ecology* 101: 265–276.
- Volkov I, Banavar JR, He F, Hubbell SP, Maritan A. 2005. Density dependence explains tree species abundance and diversity in tropical forests. *Nature* 438: 658–661.
- Whitham TG, Bailey JK, Schweitzer JA, Shuster SM, Bangert RK, LeRoy CJ, Lonsdorf EV, Allan GJ, DiFazio SP, Potts BM *et al.* 2006. A framework for community and ecosystem genetics: from genes to ecosystems. *Nature Reviews Genetics* 7: 510–523.
- Whitham TG, DiFazio SP, Schweitzer JA, Shuster SM, Allan GJ, Bailey JK, Woolbright SA. 2008. Extending genomics to natural communities and ecosystems. *Science* 320: 492–495.
- Wills C, Harms KE, Condit R, King D, Thompson J, He F, Muller-Landau HC, Ashton P, Losos E, Comita L *et al.* 2006. Nonrandom processes maintain diversity in tropical forests. *Science* 311: 527–531.
- Zhu Y, Comita LS, Hubbell SP, Ma K. 2015. Conspecific and phylogenetic density-dependent survival differs across life stages in a tropical forest. *Journal of Ecology* 103: 957–966.
- Zuur AF, Ieno EN, Elphick CS. 2010. A protocol for data exploration to avoid common statistical problems. *Methods in Ecology and Evolution* 1: 3–14.

Supporting Information

Additional Supporting Information may be found online in the Supporting Information tab for this article:

Table S1 Removal of genetic tissue sample for genetic analysis does not significantly impact germination of *Oenocarpus bataua* seeds

Table S2 Genetic diversity for unique *Oenocarpus bataua* maternal genotypes ($n = 252$) for an array of 10 microsatellite markers

Table S3 Regression coefficients of survival model of *Oenocarpus bataua* with ‘immigrant’ seeds excluded

Table S4 Choice of genetic relatedness estimator did not affect survival model results

Table S5 Covariate summary statistics

Please note: Wiley Blackwell are not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.