ORIGINAL ARTICLE

WILEY **MOLECULAR ECOLOGY**

The relative contributions of seed and pollen dispersal to gene flow and genetic diversity in seedlings of a tropical palm

Luke Browne[1,2,3,4] (iD) | Kym Ottewell[5] (iD) | Victoria L. Sork[3,4] (iD) | Jordan Karubian[1,2] (iD)

[1]400 Lindy Boggs, Department of Ecology and Evolutionary Biology, Tulane University, New Orleans, Louisiana

[2]Fundación para la Conservación de los Andes Tropicales, Quito, Ecuador

[3]Institute of the Environment and Sustainability, University of California, Los Angeles, California

[4]Department of Ecology and Evolutionary Biology, University of California, Los Angeles, California

[5]Department of Biodiversity, Conservation and Attractions, Kensington, WA, Australia

Correspondence
Luke Browne, UCLA La Kretz Center for California Conservation Science, Institute of the Environment and Sustainability, University of California, Los Angeles, CA 90095-7239.
Email: lukembrowne@gmail.com

Abstract

Seed and pollen dispersal shape patterns of gene flow and genetic diversity in plants. Pollen is generally thought to travel longer distances than seeds, but seeds determine the ultimate location of gametes. Resolving how interactions between these two dispersal processes shape microevolutionary processes is a long-standing research priority. We unambiguously isolated the separate and combined contributions of these two dispersal processes in seedlings of the animal-dispersed palm *Oenocarpus bataua* to address two questions. First, what is the spatial extent of pollen versus seed movement in a system characterized by long-distance seed dispersal? Second, how does seed dispersal mediate seedling genetic diversity? Despite evidence of frequent long-distance seed dispersal, we found that pollen moves much further than seeds. Nonetheless, seed dispersal ultimately mediates genetic diversity and fine-scale spatial genetic structure. Compared to undispersed seedlings, seedlings dispersed by vertebrates were characterized by higher female gametic and diploid seedling diversity and weaker fine-scale spatial genetic structure for female gametes, male gametes and diploid seedlings. Interestingly, the diversity of maternal seed sources at seed deposition sites (N_{em}) was associated with higher effective number of pollen sources (N_{ep}), higher effective number of parents (N_e) and weaker spatial genetic structure, whereas seed dispersal distance had little impact on these or other parameters we measured. These findings highlight the importance maternal seed source diversity (N_{em}) at frugivore seed deposition sites in driving emergent patterns of fine-scale genetic diversity and structure.

KEYWORDS
alpha, beta and gamma allelic diversity, Arecaceae, frugivore foraging ecology, pollen flow, seed dispersal, spatial genetic structure

1 | INTRODUCTION

Among plants, gene movement occurs through both pollen and seed dispersal, and the combined effects of these two dispersal processes shape patterns of genetic diversity and structure. These two dispersal processes are not independent, in that seed dispersal moves both the male and female gametic genomes and therefore determines the final location of male gametes (Crawford 1984; Hamilton, 1999; Isagi, Saito, Kawaguchi, Tateno, & Watanabe, 2007; Grivet, Robledo-

Arnuncio, Smouse, & Sork, 2009; Sork, Smouse, Grivet, & Scofield, 2015). In other words, seed dispersal determines the degree to which "primary" pollen dispersal (i.e., the movement of male gametes from sire to maternal tree) differs from patterns of "realized" paternal gene dispersal (i.e., the movement of male gametes from sire to established seedling, as mediated by seed dispersal). This in turn has important consequences for the size of local genetic neighbourhoods and associated opportunities for natural selection and genetic drift in plant populations (Wright, 1943, 1946). As such, resolving the ways

in which seed dispersal shapes the movement of male gametes and subsequent genetic parameters is a long-standing goal of plant evolutionary ecologists (Ennos, 1994; García & Grivet, 2011; Grivet et al., 2009; Hamilton & Miller, 2002; Heuertz, Vekemans, Hausman, Palada, & Hardy, 2003; Jordano, 2010; Levin, 1981; Sork et al., 2015).

The spatial extent of pollen movement is thought to exceed that of seeds in many plant species, and in some cases, differences appear to be extreme (e.g., Bittencourt & Sebbenn, 2007; Chybicki & Burczyk, 2010; Ennos, 1994; Hanson, Brunsfeld, Finegan, & Waits, 2008; Nakanishi, Tomaru, Yoshimaru, Manabe, & Yamamoto, 2009). For this reason, it is often thought that pollen dispersal contributes more to promoting gene flow and weakening genetic structure than does seed dispersal (Hamrick, 2010; Sork & Smouse, 2006). However, few direct tests of this expectation exist, with most studies instead relying on indirect estimates, different sets of molecular markers for the two dispersal modalities, or untested assumptions about pollen and/or seed movement (e.g., the closest adult is assumed to be the seed source; but see Grivet et al., 2009; Sork et al., 2015). Moreover, few studies have examined plant species in which there are a priori reasons to expect that seed movement may be equivalent to, or greater than, pollen movement (e.g., based on long-scale movements by seed dispersal agents; Holbrook, Smith, & Hardesty, 2002; Nathan, 2006; Abedi-Lartey, Dechmann, Wikelski, Scharf, & Fahr, 2016). Direct measures of seed and pollen movement among the same set of individual recruits in a system where long-distance seed dispersal is common are required to determine whether, at least in some cases, movement of seeds may exceed that of pollen.

The spatial extent of seed movement relative to pollen is likely to influence the degree to which seed dispersal directly influences seedling genetic diversity. Holding other factors such as directionality of dispersal processes and spatial configuration of source trees constant, if seeds move farther than pollen, then seed dispersal should increase realized paternal gene dispersal relative to primary pollen dispersal, effectively promoting gene flow and reducing spatial genetic structure in male gametes. In contrast, if the spatial extent of pollen movement greatly exceeds that of seed movement, then seed dispersal should create a genetic bottleneck that would serve to reduce effective gene flow and increase spatial genetic structure among male gametes (Grivet et al., 2009; Sork et al., 2015).

In addition to the distance seeds move, seed dispersal also determines the effective number of maternal source trees represented at a given seed deposition site, denoted as N_{em} (Grivet, Smouse, & Sork, 2005). Like seed dispersal distance, N_{em} varies considerably both within and among animal-dispersed plant systems, with some patches of seedlings originating from multiple maternal source trees and others originating from a single source tree (e.g., Giombini, Bravo, & Tosto, 2016; Jordano, García, Godoy, & García-Castaño, 2007; Karubian, Ottewell, Link, & Di Fiore, 2015; Ottewell, Browne, Cabrera, Olivo, & Karubian, 2018; Scofield, Smouse, Karubian, & Sork, 2012). Because seeds move both male and female gametes, the influence of N_{em} on subsequent patterns of genetic diversity extends beyond female gametes to influence both male and diploid gametic diversity

(Grivet et al., 2009). For example, an increase in N_{em} can also increase N_{ep}, the effective number of pollen sources represented in a patch of seedlings, and by extension N_e, the effective number of parents represented in a patch of seedlings. In contrast, more homogeneous seed pools characterized by low N_{em} may restrict N_{ep} and N_e, thereby dampening overall diploid genetic diversity. The relationship between seed dispersal distance and N_{em} is not well understood. For some systems, long-distance seed dispersal is associated with low diversity of maternal source trees at deposition sites, whereas in other systems, shorter distance seed movements may be associated with high diversity of maternal source trees (García, Jordano, Arroyo, & Godoy, 2009; Sork et al., 2015). Moreover, the relative influence of dispersal distance versus N_{em} on realized paternal gene flow and seedling genetic diversity is currently unknown.

Contemporary technical and analytical approaches now provide a powerful platform from which to investigate how interactions between pollen flow and seed dispersal impact patterns of genetic diversity in seedlings. The fundamental advance underlying these approaches is the use of genetic markers to unambiguously partition male versus female gametic contributions to diploid seedling tissue, which is possible when maternal seed tissue remains attached to the established seedling (Garcia, Jordano, & Godoy, 2007; Grivet et al., 2009; Robledo-Arnuncio, Grivet, Smouse, & Sork, 2012). In such cases, and when the locations of adult trees are known, we can directly measure the spatial extent of seed dispersal. primary pollen dispersal and realized paternal gene dispersal (Garcia et al., 2007). We can also estimate the effective number of maternal and paternal source trees represented within patches of dispersed seedlings (i.e., N_{em} and N_{ep}, respectively; Grivet et al., 2009; Sork et al., 2015). Moreover, the specific contributions of male versus female alleles to diploid diversity can be calculated using α (alpha), β (beta) and γ (gamma) diversity indices that are analogous to more familiar indices of species diversity (Scofield et al., 2012; Sork et al., 2015). Using this information, the spatial distribution of genetic variation of male and female gametes can also be separately assessed with spatial autocorrelation analyses (Heywood, 1991; Vekemans & Hardy, 2004). Sork et al. (2015) used a subset of these tools in the valley oak *Quercus lobata* to show that pollen dispersal by wind promoted overall diversity and reduced spatial structure, whereas seed dispersal by animals reduced overall diversity and markedly increased spatial genetic structure.

In this study, we leverage these approaches to investigate how seed and pollen dispersal interact to shape patterns of genetic diversity among seedlings of the tropical palm *Oenocarpus bataua*. We had two main goals: (a) to characterize the spatial extent of pollen versus seed movement in this species; and (b) to evaluate how seed dispersal shapes male gametic and seedling genetic diversity, with particular reference to the relative impacts of seed movement versus N_{em}. *Oenocarpus bataua* is pollinated by small flying insects, especially beetles (Núñez-Avellaneda & Rojas-Robles, 2008), whereas large-bodied avian frugivores characterized by long-distance seed dispersal dominate primary seed dispersal (J. Karubian, unpublished data). Importantly, we can distinguish three different types of seed

deposition sites, each characterized by a distinctive seed dispersal mechanism (*Lek plots* within leks of the long-wattled umbrellabird *Cephalopterus penduliger*, *Away plots* outside leks, and *Beneath plots* located beneath adult *O. bataua* trees; see *Methods* section for more details). Seed dispersal distances into these plot types have not been previously measured, but it is known that Lek plots are characterized by high N_{em} values and Away plots by intermediate N_{em} values (Karubian, Sork, Roorda, Duraes, & Smith, 2010; Scofield et al., 2012). The existence of these three distinct types of seed dispersal plots provides a useful context to evaluate how differences in seed dispersal impact the movement of male gametes and resulting patterns of diploid genetic diversity and structure.

Our overarching hypotheses were that seed movement would be extensive in this system; that seed movement (or the lack thereof) would shape male gametic and seedling genetic diversity; and that the impact of N_{em} on male gametic and seedling genetic diversity would be stronger than that of seed dispersal distance. We used a hierarchical approach to explore these issues. First, we estimated seed dispersal, primary pollen dispersal and realized paternal gene dispersal distances separately for each of our three plot types. Because the dominant seed dispersal agents for *O. bataua* are all capable of frequent long-distance dispersal, we predicted that seed dispersal distances for Lek and Away plots would be much greater than in Beneath plots and, similar to that of pollen dispersal distances. Here, we also took advantage of an opportunity to test the expectation that realized paternal gene dispersal would exceed primary pollen dispersal in both Lek and Away plots (i.e., seed dispersal would increase the effective distance male gametes are moved), whereas differences would be minimal in Beneath plots. Next, we evaluated how N_{em} varies among plot types. On the basis of previous work (e.g., Karubian et al., 2010; Scofield et al., 2012) and because umbrellabirds bring seeds into Lek plots year-round, whereas Away plots are putatively used ephemerally for activities such as roosting, we expected N_{em} to be higher in Lek plots than Away plots; Beneath plots, where most seeds are expected to be gravity-dispersed, were expected to present very low N_{em} values. Finally, we evaluated the importance of seed dispersal (i.e., Lek and Away plots vs. Beneath plots), and of seed dispersal distance versus N_{em} (i.e., Lek vs. Away plots), on three indices of genetic diversity: (a) the effective number of pollen sources N_{ep} and parents N_e per plot; (b) female and male gametic diversity and seedling diversity; and (c) fine-scale spatial genetic structure. Our expectation was that seed dispersal would have an important impact on all the indices we examined but that among seeds that have been dispersed, the impact of N_{em} would be greater than that of seed dispersal distance.

2 | METHODS

2.1 | Study system

We conducted field research in Bilsa Biological Station (BBS; 79°45′ W, 0°22′ N; 330–730 m elevation; 3.4 m rain/year), a 3500-ha reserve of humid Chocó rainforest surrounded by an increasingly agricultural landscape in the Mache-Chindul Reserve, northwest Ecuador. *Oenocarpus bataua* is a widespread canopy palm in Latin America (ter Steege et al., 2013) and is common at BBS. The species is monoecious, meaning that the same individual can potentially serve as both a seed source and a pollen source, but is also almost entirely outcrossing, making it unlikely that the same adult would be the seed and pollen source of a given seedling (Ottewell, Grey, Castillo, & Karubian, 2012). Pollination services are provided by small insects, particularly beetles (Núñez-Avellaneda & Rojas-Robles, 2008). Adults at BBS exhibit very weak spatial genetic structure, likely as a consequence of extensive pollen movement and outcrossing (Ottewell et al., 2012) and seed dispersal (Karubian, Duraes, Storey, & Smith, 2012). At BBS, *O. bataua* has flowers and ripe fruit available in all months of the year, but the proportion of flowering and fruiting adults in the population is usually relatively low, punctuated by population-wide peaks in fruit production at approximately 30-month intervals (J. Karubian unpublished data; see also Rojas-Robles & Stiles, 2009). Age at first reproduction is unknown for our population but likely to be on the order of 40 years or more.

Oenocarpus bataua produces large-seeded, lipid-rich fruits in single infructescences of up to 2,000 fruits (Goulding & Smith, 2007). Ripe fruits, available for 4–8 weeks on a given infructescence, are an important food source for frugivorous birds and mammals (Goulding & Smith, 2007). At BBS, three large-bodied avian frugivores provide most primary seed dispersal: the long-wattled umbrellabird and two toucan species (*Ramphastos swainsonii* and *R. brevis*) (J. Karubian, unpublished data). We have also observed the oilbird (*Steatornis caripensis*) removing fruit, but have not quantified removal rates by this or any other nocturnal frugivores. Secondary dispersal of already-dispersed seeds is relatively rare in this system (L. Browne and J. Karubian, unpublished data). All four avian frugivores swallow *O. bataua* fruits whole and regurgitate intact, viable seeds, and all are known to be capable of long-distance dispersal (Holbrook, 2011; Holland, Wikelski, Kümmeth, & Bosque, 2009; Karubian et al., 2012).

This system allows us to distinguish three distinct seed dispersal plot types, each characterized by relatively high densities of *O. bataua* seedlings. *Lek plots* are generated by male long-wattled umbrellabirds that deposit seeds year-round beneath display perches in a lek, often transporting seeds long distances (Karubian et al., 2012). *Away plots* are located away from *O. bataua* adults, but not in leks, with most seeds likely deposited by toucans, other avian frugivores and potentially some arboreal mammals that use these areas for sleeping, resting, nesting or food processing. *Beneath plots* are located directly beneath the canopies of *O. bataua* adults, and most seeds are likely gravity-dispersed from the nearest adult *O. bataua* to these sites.
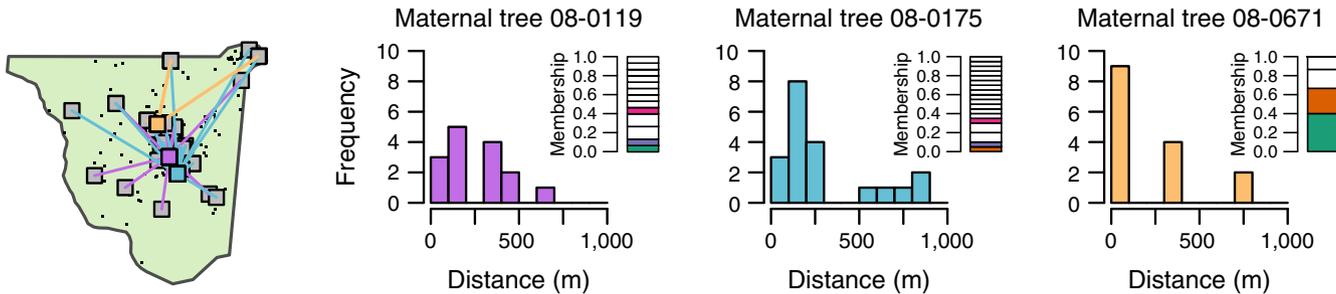
2.2 | Sampling design

Sample collection took place in a 130-ha study area within BBS in which all *O. bataua* adults (i.e., reproductive individuals as inferred from the presence of active or previous reproductive structures; $n = 185$ individuals, 1.5 per ha) are marked and genotyped (Figure 1a). To sample seeds and seedlings, we established sample
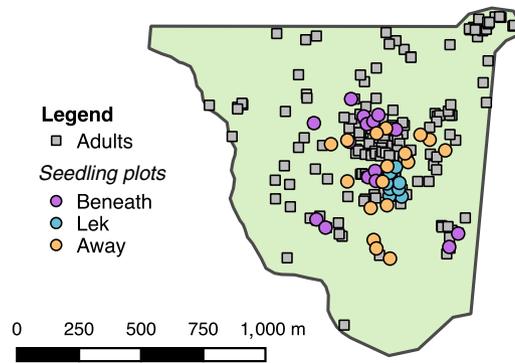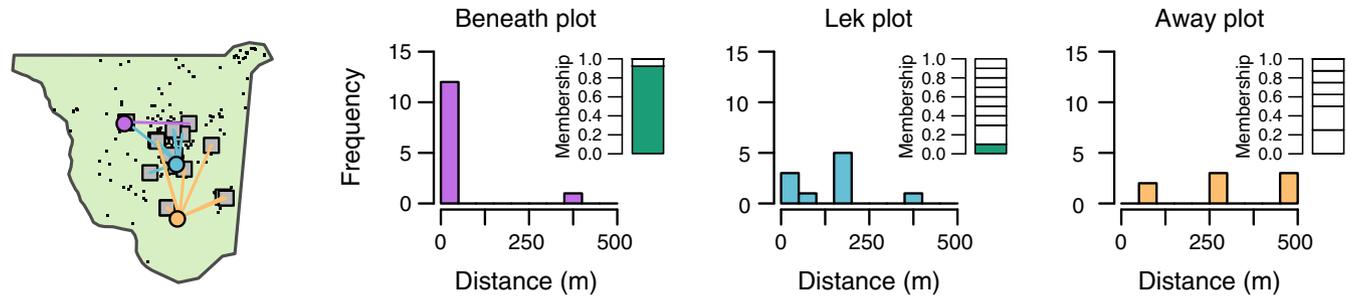
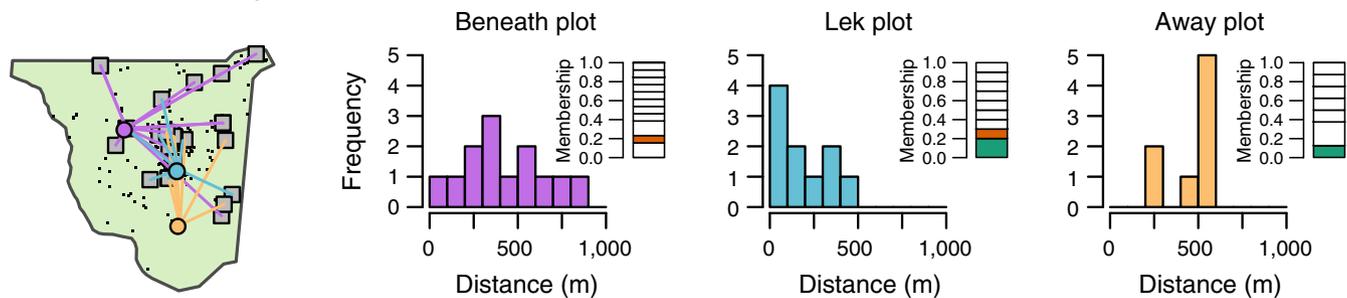**FIGURE 1** Map of the study area and representative patterns for primary pollen dispersal, seed dispersal and realized paternal gene dispersal for *Oenocarpus bataua* in northwest Ecuador. Panel (a) depicts adult *O. bataua* trees (squares) in 130-ha study area and the study plots (circles) where Beneath, Lek and Away seedlings were sampled. Panel (b) illustrates primary pollen movement for three arbitrarily selected maternal trees included in the study, along with the frequency histograms of pollen dispersal distances and membership plots depicting sires per source tree. Lines in the miniature map connect maternal trees to pollen donors, and each line in the membership plot represents a different pollen donor, with coloured lines representing pollen donors shared among maternal trees. Panels (c) and (d) provide examples of seed movement and realized paternal gene movement, respectively, into one representative Beneath plot, one Lek plot and one Away plot. Frequency histograms of movement distances and membership plots showing number of maternal source trees per plot (for seed movement) and number of sire trees per plot (for realized paternal gene movement) are provided

collection plots in areas corresponding to our three categories of deposition sites: Beneath (n = 15 plots), Lek (n = 8) and Away (n = 17). Lek plots were situated >10 m from an *O. bataua* adult to ensure all seeds collected had been dispersed into the plot and located within an umbrellabird lek, with most seedlings likely dispersed by displaying males (Karubian & Durães, 2014; Karubian et al., 2010). The lek is located in centrally within the 130-ha study plot and consists of 8–12 males umbrellabirds, each of which holds a fixed territory of ~ 25 m², the entire lek is ~ 1 ha in area. Away plots (also referred to as "Background" plots in Karubian et al., 2010 and "Random" plots in Scofield et al., 2012) were also situated >10 m from an *O. bataua* adult but were outside the umbrellabird lek; most of these seedlings were likely dispersed by toucans, umbrellabirds and oilbirds that used branches on trees above these locations for resting, sleeping or food processing, although other sources of dispersal are possible. Beneath plots were located directly beneath the canopies of *O. bataua* adults, and most seedlings were dispersed by gravity, although some proportion of seeds from other adult trees could have been dispersed to these sites. Each circular sampling plot had a 10 m diameter and contained >8 *O. bataua* established seedlings with seeds still attached. In all plot types, samples of leaf tissue and seed pericarp tissue were sampled nondestructively and stored in individually labelled envelopes in dry conditions until laboratory analyses were conducted.

2.3 | Molecular analyses and gametic extraction

We extracted DNA from leaf and seed tissue samples and amplified 11 microsatellite loci using PCR methods described in Ottewell et al. (2012). We extracted DNA from 175 Beneath seedlings (15 plots), 139 Lek seedlings (8 plots) and 194 Away seedlings (17 plots) and genotyped these individuals for both leaf and seed tissue (N = 508 total individuals from 40 total plots). To determine the separate haploid male and female gametic genotypes, we used a modified TwoGener gametic extraction (Smouse, Dyer, Westfall, & Sork, 2001; Sork et al., 2015), using the diploid leaf and seed tissue of each seedling. In cases where the male and female gametic contribution could not be assigned definitively (e.g., where both leaf and seed tissue are heterozygous at the same alleles) representing ~ 11% of cases, we assigned a 50% probability of the ambiguous alleles coming from either the male or female tree. Genotypes of adults in the 130-ha study plot (N = 185) were used from previously published studies (Browne, Ottewell, & Karubian, 2015; Ottewell et al., 2012). Genotype and location data for adults and seedlings are publicly available from Figshare (https://doi.org/10.6084/m9.figshare.6434144). The R code used for gametic extraction is available on Github (https://github.com/lukembrowne/gametic-extraction; https://doi. org/10.5281/zenodo.1274538).

2.4 | Statistical analyses

2.4.1 | Parentage assignment

To assess seed dispersal, we matched genotypes of (maternal) pericarp tissue from seeds to those of potential maternal source trees in

CERVUS v. 3.0.3 (Marshall, Slate, Kruuk, & Pemberton, 1998). When a "match" (defined as an exact match or a mismatch of up to two loci) between seed pericarp and potential source tree was observed, the seed was considered to have originated from that source tree. The probability of identity (p_{ID}, *i.e.*, the probability that two randomly drawn unrelated individuals from the population will share the same genotype) and the probability of parentage exclusion (p_{EX}) when one parent genotype is known for each locus and combined across loci are available in Supporting Information Table S1. Of our total sample size of 508 seedlings, 24 were excluded because we could not amplify microsatellite loci for pericarp tissue, yielding 484 seedlings included in the maternal pericarp matching analysis.

To assess pollen dispersal, we used the program CERVUS v. 3.0.3 (Marshall et al., 1998) to assign paternity to both those offspring that we were able to assign to a maternal source tree (i.e., using the direct matching methods of seed pericarp tissue described in the previous paragraph) and offspring who were not assigned a maternal source tree. For those seedlings that had a maternal match, we included this information as a known maternal tree in the CERVUS analysis. As *O. bataua* is monoecious, it is possible for an individual to be both a seed donor and to contribute paternity to the offspring of other maternal trees. We conducted parentage analysis on all 508 seedlings, although 20 seedlings were excluded for not amplifying at leaf tissue. Critical values (Δ) for parentage assignment were estimated using a simulation with the following conditions: 10,000 offspring genotypes, 185 candidate fathers, assuming 98% of candidate fathers were sampled and with a genotyping error rate of 0.03 based on previous studies from the same system (Ottewell et al., 2012). We assigned paternity at an 80% confidence threshold, although qualitative results did not change using more stringent thresholds (e.g., 95%). Overall, we obtained both a maternal and paternal match for most seedlings (n = 145 in Beneath plots, n = 105 in Lek plots, n = 141 in Away plots), with no seedlings having only a maternal match, but no paternal match, and 10, 8 and 14 seedlings having only a paternal match, but no maternal match, in Beneath, Lek and Away plots, respectively.

2.4.2 | Observed dispersal distances

We calculated "observed" dispersal distances for all seedlings for which we could assign a maternal or paternal source tree, respectively, located within the study plot using the methods described above. We calculated, observed seed dispersal distance as the Euclidean distance between the maternal source tree as determined by maternal genotyping matching and the spot where the established seedling was encountered. For pollen, we calculated two measures of observed dispersal distance: (a) primary dispersal distance, from the paternal source tree to the maternal tree where fertilization occurred, and (b) realized paternal gene dispersal, from the paternal source tree to the spot where the established seedling was encountered. In other words, realized paternal gene dispersal distance is a measure of how far male gametes move from the paternal source tree to the location where the individual seedling establishes. We

were only able to calculate primary pollen dispersal distances for the subset of seedlings that had maternal + paternal matches. To test whether observed pollen and seed dispersal distances among plot types differed from each other, we used a nonparametric Kruskal–Wallis rank-sum test and Dunn's test for post hoc analysis using a Bonferroni correction for multiple comparisons in R v 3.3.1 (R Core Team 2016). We excluded seedlings that did not have a maternal or paternal match to an adult in our study parcel when estimating observed dispersal distances.

2.4.3 | Dispersal kernels and immigration rates

We estimated pollen and seed dispersal kernels, i.e., the probability of pollen or seed travelling from its origin (0,0) to fertilization or deposition at position (x, y), respectively, using the NEIGHBORHOOD model with the software NMπ (Chybicki, 2017), which is a reimplementation of NM+ (Chybicki & Burczyk, 2010). To model pollen dispersal, the NEIGHBORHOOD model uses a maximum-likelihood fractional paternity assignment approach where for each seedling, paternity may result from (a) a paternal tree located outside the study plot due to pollen immigration with probability m_p or (b) by a paternal tree located within the study plot (with probability ($1-m_p$)). Offspring may be the result of self-pollination with probability s. Similarly, for seed dispersal, seeds either arise from a maternal tree located within the study parcel with probability m_s or may result from seed immigration from a maternal tree outside the study parcel with probability $1-m_s$. We used the two-parameter Weibull dispersal kernel (Chybicki & Burczyk, 2010) to model primary pollen and seed dispersal kernels. The Weibull dispersal kernel has been shown to be a good fit for pollen dispersal in *O. bataua* (Ottewell et al., 2012) and for both pollen and seed dispersal kernels in other species (Bullock et al., 2017). We estimated the scale a and shape b parameters of the Weibull dispersal kernel, where $b < 1$ indicates a fat-tailed distribution and $b > 1$ indicates a thin-tailed distribution with fewer long-distance dispersal events, whereas $b = 2$ approximates a normal distribution (Chybicki & Burczyk, 2010). We also estimated the frequency of self-fertilization (s), the frequency of seed immigration (m_s) and the frequency of pollen immigration (m_p). Genotyping error rates were set at 0.03 across loci. We set the NEIGHBORHOOD parameter to include all sampled adults in our study plot as the neighbourhood size (Chybicki, 2017; Chybicki & Burczyk, 2010). Kernel analyses produced by the NEIGHBORHOOD model are designed to avoid dependence on the spatial configuration of source trees relative to the sampling plots, whereas estimates of observed dispersal distances generated from direct parentage (above) are not; for this reason we prioritize results from NMπ over that of parentage assignment in cases where results from the two methods differ (see also *Discussion*, below).

2.4.4 | Parental correlations and effective parental sizes

To assess shared paternity and maternity among seedlings and the number of effective parents contributing to different patches, we

ran a parental structure analysis (PSA; Robledo-Arnuncio et al., 2012; see also Grivet et al., 2009). The PSA estimates the correlation of paternity within patches—Q_w^p, the probability that two seedlings drawn at random from the same patch share the same father; Q_w^m, correlation of maternity within patches—the probability that two seedlings drawn at random from the same patch share the same mother; and Q_w^{mp}, cross-parental correlation within patches—the probability that two seedlings drawn at random from the same patch have a cross-parental match. We also estimated parental correlations among patches: Q_b^p, correlation of paternity among patches—the probability that two seedlings drawn at random from two different patches share the same father; Q_b^m, correlation of maternity among patches—the probability that two seedlings drawn at random from two different patches share the same mother; Q_b^{mp}, cross-parental correlation among patches—the probability that two seedlings drawn at random from two different patches have a cross-parental match. We used the parental correlation estimates to estimate the effective number of fathers ($N_{ep} = 1/Q_w^p$), effective number of mothers ($N_{em} = 1/Q_w^m$) and effective number of parents ($N_e = 4/(Q_w^p + Q_w^m + 2Q_w^{mp})$) per patch. We used jackknifing across individuals to estimate approximate standard errors of parental correlations and effective parental sizes, and PSA was run separately for each plot type (e.g., Beneath, Away and Lek). To test for differences in within-patch paternal and maternal correlations across plot types, we used a nonparametric Kruskal–Wallis rank-sum test and Dunn's test for post hoc analysis with a Bonferroni correction for multiple comparisons R in R v 3.3.1 (R Core Team 2016). Two Beneath plots were excluded from these analyses for having <6 seedlings.

2.4.5 | Alpha, beta and gamma allelic diversity

To estimate the relative contributions of male gametes from pollen and female gametes from seeds to patterns of allelic diversity, we used the genetic counterparts of the species diversity estimates of α, β, δ and γ diversity (Scofield et al., 2012; Sork et al., 2015) for male gametes, female gametes and diploid leaf tissue of seedlings in each plot type. Within-patch α diversity estimates the effective number of alleles per locus per sampled patch and is calculated as the reciprocal of r_{gg}, the unbiased estimate of the probability of drawing identical alleles for a locus in a sampled patch (Grivet et al., 2005; Sork et al., 2015). Between-patch β diversity estimates the effective number of genetically nonoverlapping patches and provides information on the degree of allelic turnover across sampled patches, with higher values indicating more turnover between patches. Lek patches were on average closer to each other than were Away or Beneath plots (next section), which could influence patterns of β diversity, so we reran our β diversity analyses including those plots with similar pairwise distances to ensure that qualitative patterns were robust to this sampling area effect. We also measured allelic divergence across patches in terms of δ (delta) (Sork et al., 2015), where δ = 1 represents no overlap of alleles across patches and δ = 0, represents complete sharing of alleles across patches. Cumulative allelic diversity accumulated across plots was estimated with γ, which is the

effective number of alleles per locus across the entire group of seedlings without regard to patch. We did not conduct statistical comparisons of β, δ and γ across plot types because there is currently no method to statistically compare β and δ across plot types, and because of differences in the total sampling area may bias estimates of γ (above). We averaged diversity estimates across 10 loci (Ob19 was excluded from analysis due to not amplifying at a subset of seedlings). We also present scaled diversity metrics (between 0 and 1) from Sork et al. (2015) to allow comparison to other studies that may differ in overall levels of genetic diversity. To test for differences in within-patch α diversity of male gametes, female gametes and diploid seedlings across and within plot types, we used a nonparametric Kruskal–Wallis rank-sum test and Dunn's test for post hoc analysis with a Bonferroni correction for multiple comparisons in R v 3.3.1 (R Core Team 2016). Two Beneath plots were excluded from these analyses for having <6 seedlings.

2.4.6 | Fine-scale spatial genetic structure

To assess fine-scale spatial patterns of genetic variation among seedlings, we estimated fine-scale spatial genetic structure with the kinship coefficient F_{ij} of Loiselle, Sork, Nason, and Graham (1995) using the program SPAGeDi v1.5 (Hardy & Vekemans, 2002). We estimated the strength of spatial genetic structure using the Sp statistic of Vekemans and Hardy (2004), which is calculated as $-b_{Flog}/(1-F_1)$, where b_{Flog} is the mean slope of the regression coefficient of F_{ij} on a log distance scale and F_1 is the mean estimate of the kinship coefficient of the first distance class. We restricted the maximum distance used when estimating b_{Flog} to 150 m, which is approximately the maximum pairwise distance between Lek plots; the maximum pairwise distance for Away plots (547 m) and Beneath plots were greater (740 m). Average pairwise distances between Lek plots (mean ± standard deviation; 59 ± 26 m) were also smaller than in Beneath (317 ± 189 m) or Away (269 ± 132 m) plots. Higher values of the Sp statistic indicated stronger patterns of spatial genetic structure, and the Sp statistic is robust to the choice of distance intervals (Vekemans & Hardy, 2004). To visualize how kinship changes with spatial distance, we plotted F_{ij} across four distance intervals: 0–10 m, 10–50 m, 50–100 m and 100–150 m, which were chosen to include a minimum of 100 pairwise comparisons, >50% of individuals and a coefficient of variation of participation ≤1.0 for each plot type following the recommendations of Hardy and Vekemans (2002). We calculated reference allele frequencies used in estimating F_{ij} separately for each plot type. We assessed statistical significance at each distance interval based on the 95% confidence interval of the null distribution of permuting individuals among locations n = 9,999 times. Following Hardy et al. (2004) and Nakanishi et al. (2009), we converted the haploid male and female gametic genotypes to diploid homozygous genotypes prior to analysis. To account for ambiguous cases where both leaf and seed tissue were heterozygous at the same allele (see above), the maternal and paternal haploid genotypes were converted to the corresponding heterozygous genotypes. We tested for statistical differences in Sp across male gametes, female gametes and diploid seedlings, and leaf genotypes with an ANOVA using locus as the blocking factor. To compare female gametes to male gametes within each plot type, we used a Wilcoxon signed-rank test with locus as the pairing factor.

3 | RESULTS

3.1 | Pollen, realized paternal gene and seed dispersal distances

Kernel estimates from NMπ and direct parentage assignment both indicate long-distance pollen dispersal in the O. bataua system (Tables 1 and 2). Comparing across plot types, mean pollen dispersal distances and shapes of the pollen dispersal kernels estimated in NMπ were similar across plot types, with overlapping 95% confidence intervals (Table 1, Figure 2). Under direct parentage assignment in CERVUS, primary pollen dispersal distances differed across plot types (Kruskal–Wallis chi-squared = 6.95, df = 2, p = 0.03; Table 2), being greater for Away seedlings than Beneath seedlings (Dunn's z = −2.44, p = 0.022), with no statistical differences in other pairwise comparisons (p > 0.68 for both tests). Patterns of realized paternal gene dispersal distances (i.e., distance from the sire to the site of seedling establishment) across plot types were similar to those of primary pollen dispersal, with Away seedlings being greater than Beneath seedlings (Dunn's z = −3.94, p = 0.0001), but no statistically significant differences among other pairwise comparisons (p > 0.08; Table 2). In terms of magnitude, realized paternal gene dispersal distances were similar to primary pollen dispersal distances across all plot types (Table 2). Pollen immigration rates were similar among Lek and Away seedlings, but significantly higher in Lek seedlings compared to Beneath seedlings (Table 1).

In contrast to pollen movement, seed movement was spatially restricted and varied among plot types. Mean and maximum modelled (Table 1, Figure 2) and observed (Table 2) dispersal distances for seeds in each plot type were all lower than corresponding values for primary pollen dispersal and realized paternal gene movement. The shape parameter (b) in the modelled dispersal kernels for Lek and Away seedlings was <1 for seed and >1 for pollen, indicating a fatter-tailed dispersal kernel for seeds compared to pollen among these plot types (Table 1, Figure 2). We found significant differences in seed dispersal distances among plot types (Kruskal–Wallis chi-squared = 233.08, df = 2, p = <0.001), where both observed and modelled seed dispersal distances were greater in Lek (Dunn's z = −11.702, p < 0.001) and Away (Dunn's z = −14.20, p < 0.001) seedlings than in Beneath seedlings; there were no differences between Away and Lek plots (Dunn's z = −1.736, p = 0.124; Tables 1 and 2). Seed immigration rates estimated by dispersal kernel modelling were near 0, and generally lower than pollen immigration rates (Table 1).

TABLE 1 Modelled pollen and seed dispersal kernel estimates from NMπ for the palm *Oenocarpus bataua*, in northwest Ecuador for Beneath, Lek, and Away plot types

Plot type	ds	bs	dp	bp	s	ms	mp
Beneath	10.2^a (8.3–13.2)	1.41^a (1.08–1.75)	568.8^a (364.9–1289.5)	1.37^a (1.18–1.69)	0.01^a (0.00–0.03)	0.00*	0.07^a (0.00–0.14)
Lek	97.7^b (57.8–315.0)	0.73^b (0.40–1.05)	417.2^a (222.5–3352.0)	1.69^a (1.30–2.07)	0.10^a (0.03–0.17)	0.01 (0.00–0.15)	0.34^b (0.17–0.51)
Away	107.0^b (72.0–208.0)	0.75^{ab} (0.39–1.12)	552.4^a (297.8–3812.6)	1.52^a (1.09–1.94)	0.04^a (0.00–0.08)	0.00*	0.22^{ab} (0.10–0.33)

The table includes *ds* (mean distance of seed dispersal in metres), *bs* (shape parameter of seed dispersal kernel), *dp* (mean distance of pollen dispersal in metres) and *bp* (shape parameter of pollen dispersal kernel) and approximate 95% CI (confidence intervals) in parentheses. Also shown are *s* (frequency of self-fertilization), *ms* (frequency of seed immigration) and *mp* (frequency of pollen immigration) ± standard error (SE). Superscript letters indicate where the CI of parameter estimates comparing Beneath, Lek and Away seedlings do not overlap. Both pollen and seed dispersal kernels were fit to a Weibull distribution. The CI of seed immigration rates of Beneath and Away seedlings could not be estimated (*).

TABLE 2 Summary of observed pollen, realized paternal gene and seed dispersal distances for the palm *Oenocarpus bataua* in northwest Ecuador from direct parentage assignment for Beneath, Lek, and Away plots

	Min.	Median	Mean	Max.
Primary pollen				
Beneath^a	32.8	214.7	254.8	1106.7
Lek^{ab}	0.0	213.7	263.3	844.8
Away^b	0.0	275.0	313.3	1061.1
Realized paternal gene				
Beneath^a	0.0	162.0	218.1	1109.5
Lek^{ab}	23.0	204.1	254.8	828.9
Away^b	13.9	257.4	296.3	1039.2
Seed				
Beneath^a	0.0	0.0	20.8	656.2
Lek^b	10.3	47.2	76.4	536.7
Away^b	10.0	64.8	107.3	601.2

The minimum (Min.), median, mean and maximum (Max.) distances in metres are shown for all seedlings that could be directly linked to parental trees located within our study area. Primary pollen refers to the distance pollen moved between sire and maternal tree; Realized paternal gene refers to the distance male gametes moved between sire and established seedling; Seed refers to the distance seeds moved between maternal tree and established seeding. Shown are values for each of three plot types. Superscript letters indicate statistically significant differences among plot types.

3.2 | Parental correlations and effective number of parents

The probability of two seedlings within a patch sharing a father (i.e., paternal correlations, Q_w^p) was lowest in Lek plots, intermediate in Away plots and highest in Beneath plots (Table 3, Supporting Information Table S2), although these differences were not statistically significant (Kruskal–Wallis chi-squared = 5.59, *df* = 2, *p* = 0.06). The probability of two seedlings sharing the same mother (i.e., maternal correlations, Q_w^m) differed among plot types (Kruskal–Wallis chi-squared = 22.6, *df* = 2, *p* = <0.001; Table 3). More specifically, Beneath plots had significantly higher maternal correlations than did Lek (Dunn's *z* = 4.16, *p* < 0.001) and Away (Dunn's *z* = 3.97, *p* = <0.001) plots; maternal correlations were lower in Lek plots than Away plots, but this difference was not statistically significant (Dunn's *z* = −0.94, *p* = 0.516). Between-patch parental correlations were generally low for both sexes across all plot types (Table 3).

Consistent with parental correlation results, the effective number of fathers N_{ep} exceeded the effective number of mothers N_{em} within each plot type, although errors around these estimates were high (Table 3). Among plot types, both N_{ep} and N_{em} were higher in Lek and Away plots than in Beneath plots, with a disproportionately high ratio of $N_{ep}:N_{em}$ in Lek plots relative to other plot types (Table 3). Overall effective population size N_e was greatest in Lek plots, followed by Away plots and then Beneath plots (Table 3).
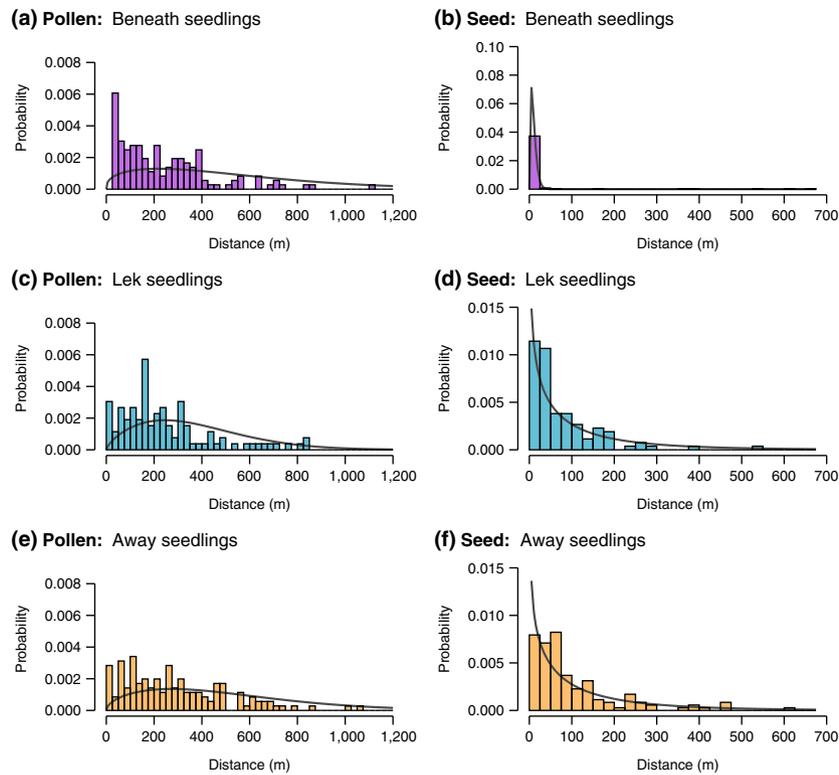
FIGURE 2 Estimated pollen and seed dispersal kernels (solid black lines) fit to a Weibull distribution for (a and b) Beneath, (c and d) Lek and (e and f) Away *Oenocarpus bataua* seedlings in northwest Ecuador, along with the frequency of observed dispersal distances (purple, blue, and yellow histograms) derived via parentage analysis + matching seed genotypes to maternal trees

3.3 | Allelic diversity

Allelic alpha diversity (α) was higher for male gametes than for female gametes within each plot type (Table 4, Supporting

TABLE 3 Mean estimates ± standard error (*SE*) of within-patch paternal (Q_w^p), maternal (Q_w^m) and cross-parental (Q_w^{mp}) correlations; between-patch paternal (Q_b^p), maternal (Q_b^m) and cross-parental (Q_b^{mp}) correlations; and effective number of fathers (N_{ep}), mothers (N_{em}) and total parents (N_e) of *Oenocarpus bataua* seedlings in Beneath, Lek and Away plots

	Beneath	Lek	Away
Within-patch			
Q_w^p	0.099 ± 0.008[a]	0.004 ± 0.005[a]	0.048 ± 0.006[a]
Q_w^m	0.704 ± 0.014[a]	0.108 ± 0.003[b]	0.207 ± 0.005[b]
Q_w^{mp}	0.051 ± 0.008	−0.002 ± 0.004	0.016 ± 0.009
Between-patch			
Q_b^p	−0.006 ± 0.001	0.001 ± 0.001	−0.002 ± 0.001
Q_b^m	−0.051 ± 0.001	−0.014 ± 0.001	−0.011 ± 0.001
Q_b^{mp}	−0.024 ± 0.002	−0.012 ± 0.002	−0.004 ± 0.001
Effective number of parents			
N_{ep}	10.06 ± 0.86	227.78 ± 216.52	20.65 ± 2.90
N_{em}	1.42 ± 0.03	9.30 ± 0.25	4.84 ± 0.11
N_e	4.42 ± 0.09	37.05 ± 4.53	13.90 ± 1.30

Significant differences between within-patch paternal and maternal correlations among plot types are indicated by superscript letters.

Information Table S2; $p \leq 0.006$ for all). Comparing among plot types, male gametic α diversity was similar among Beneath, Away and Lek plots (Table 4). In contrast, female gametic α diversity was significantly lower in Beneath plots than in Lek and Away plots, which were similar to each other (Table 4). Allelic α diversity of diploid leaf tissue was also lower in Beneath plots than in Lek or Away plots, but Lek plots did not differ from Away plots (Table 4). The scaled version of α diversity (α') followed similar patterns (Table 4).

Across all plot types, there was less divergence (δ, delta) and lower β and β' diversity among male gametes than among female gametes, consistent with more extensive movement of pollen across the study area (Table 4). Overall, Beneath plots had higher divergence (i.e., less overlap), especially in female gametes, than did Away or Lek plots. Gamma diversity (γ), a measure of accumulated diversity summed across individual plots, and its scaled counterpart (γ') were higher for male gametes than female gametes for all plot types (Table 4). For δ, β and γ diversity and their scaled counterparts, values for diploid leaf tissue were intermediate between those of male and female gametes.

3.4 | Fine-scale spatial genetic structure

For all plot types, the strength of spatial genetic structure (Sp) of male gametes was weaker than that of female gametes (Supporting Information Table S3; $p \leq 0.006$ for all). For all but the

TABLE 4 Allelic diversity in male gametes, female gametes and diploid leaf tissue seedlings of *Oenocarpus bataua* in Beneath, Lek, and Away plot types in northwest Ecuador

	Male gametes			Female gametes			Diploid seedlings		
	Beneath	Lek	Away	Beneath	Lek	Away	Beneath	Lek	Away
Within patches									
α	2.688[a]	2.738[a]	2.884[a]	1.602[a]	2.109[b]	2.017[b]	2.279[a]	2.417[b]	2.475[b]
α′	0.673	0.697	0.697	0.403	0.577	0.538	0.579	0.612	0.614
Among patches									
β	1.088	0.996	1.012	1.475	1.053	1.097	1.189	1.028	1.048
β′	0.090	−0.005	0.013	0.358	0.056	0.098	0.176	0.03	0.051
δ	0.087	−0.006	0.011	0.382	0.095	0.120	0.211	0.045	0.064
Total across patches									
γ	2.925	2.727	2.918	2.363	2.221	2.213	2.710	2.484	2.594
γ′	0.662	0.639	0.661	0.580	0.554	0.551	0.633	0.600	0.616

Shown are within-patch α diversity and its scaled counterpart (α′), among-patch β diversity and its scaled counterpart (β′), divergence (δ), and across-patch γ diversity and its scaled counterpart (γ′). Significant differences between within-patch α diversity among plot types are indicated by superscript letters.
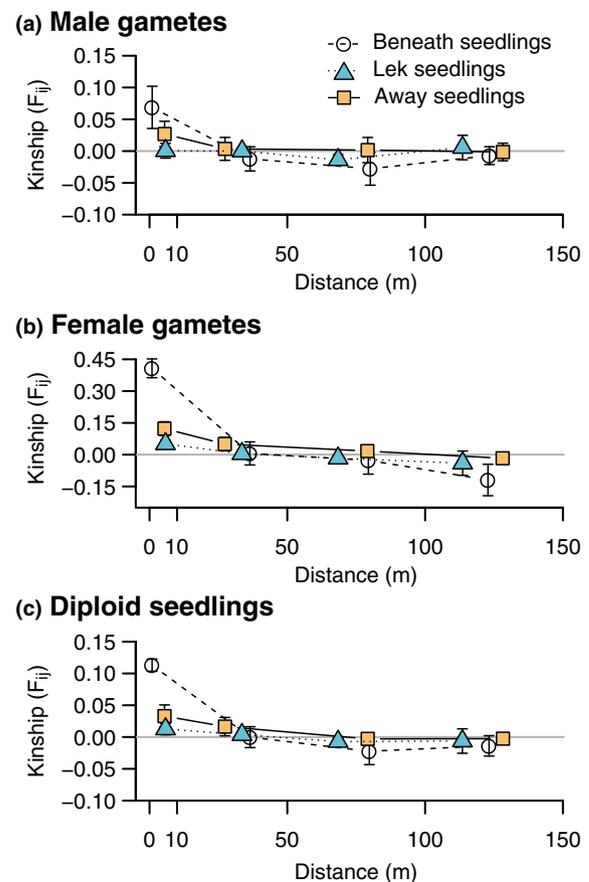
male gametes of Lek and Away seedlings, genetic kinship (F_{ij}) peaked at short distance intervals (0–10 m) and declined significantly with distance, showing local patterns of isolation by distance (Figure 3, Supporting Information Table S3). For male gametes, female gametes and diploid seedlings, spatial genetic structure was significantly stronger in Beneath seedlings than Lek ($p < 0.001$ for all comparisons) or Away ($p < 0.001$ for all comparisons) seedlings. Lek seedlings showed weaker spatial genetic structure in these categories relative to Away plots, but the difference was not significant ($p > 0.05$ for all; Figure 3, Supporting Information Table S3).

Female gametic correlograms exhibited two distinctive patterns that correspond to patches with dispersed versus undispersed seedlings. For the Beneath plots, female gametic correlations declined sharply with increasing distance, with a particularly strong drop off from the first distance class, which relates to a single patch and where F_{ij} values of 0.40 are nearly equivalent to that of full-siblings, to the second distance class, representing a neighbouring patch, for which F_{ij} values are close to 0. In contrast, for Away and Lek patches we documented much lower, but still nonzero correlation (F_{ij} = 0.12, 0.05, respectively) within the first distance classes that dropped off slowly as one moves out to the second and third distance classes. Male gametes yielded essentially flat correlograms for all three plot types, albeit with higher relatedness in the first distance class, which relates to a single patch, particularly for Beneath plots. Diploid seedlings exhibited an intermediate pattern.

4 | DISCUSSION

In this study, we addressed long-standing questions about the relative contributions of seed and pollen dispersal to gene flow in plant populations, and the way in which these two dispersal mechanisms

interact to shape patterns of genetic diversity and structure. We accomplished this using methods recently developed by Grivet et al. (2009), Robledo-Arnuncio et al. (2012) and Sork et al. (2015) that

FIGURE 3 Fine-scale spatial genetic structure in (a) male gametes, (b) female gametes and (c) diploid seedlings of *Oenocarpus bataua* seedlings in northwest ± Ecuador in Beneath, Lek, and Away plots. The kinship coefficient F_{ij} ± 2 standard error is plotted against distance intervals

allowed us to unambiguously partition the genetic variation contributed by male and female gametes in established seedlings of the animal-dispersed palm *Oenocarpus bataua*. Our findings highlight that although pollen moves considerably further than seeds in this system, frugivore-mediated seed dispersal, and in particular the diversity of maternal seed sources (N_{em}) that frugivores generate at seed deposition sites, plays a key role in driving patterns of fine-scale genetic diversity and structure among seedlings.

4.1 | Pollen versus seed dispersal

By directly comparing seed versus pollen movement for the same set of seedlings, we found that *O. bataua* pollen moves much greater distances than do seeds. This came as a surprise, because pollen in this system is moved by small insects, many of which are <10 mm in body length (Núñez-Avellaneda & Rojas-Robles, 2008), whereas seeds are moved by large-bodied bird species touted for their long-distance seed dispersal capabilities. Several lines of evidence supported our a priori assumption of long-distance seed dispersal: Radio-tracking studies indicate that umbrellabird males engage in a high incidence of long-distance dispersal into Lek plots (Karubian et al., 2012), and the toucans and oilbirds whose roosting and food processing behaviours putatively generate Away plots are also known for long-distance seed movement (Holbrook, 2011; Holland et al., 2009). In this study, we observed fat-tailed seed dispersal kernels ($b < 1$) in both Lek and Away plot types where these birds deposit seeds. Yet despite this, primary pollen dispersal distance was more than four times that of seed dispersal for all plots types as estimated by both kernel (mean dispersal distances dp versus ds; Table 1) and direct parentage (median dispersal distances; Table 2) approaches. Moreover, the presence versus absence of vertebrate seed dispersal (i.e., comparing Away and Lek plots vs. Beneath plots) appeared to have little direct impact on realized paternal gene dispersal distances, or on most metrics of male gametic diversity (e.g., α and β diversity), consistent with pollen movement greatly eclipsing that of seeds in this system. Thus, even for a system in which we expected the spatial extent of seed movement to rival that of pollen, and for which we could unambiguously distinguish and measure movement of pollen versus seeds, our measurements of pollen movement were much greater, supporting the general hypothesis that pollen movement is likely to exceed that of seeds for most plant species (Sork & Smouse, 2006; Hamrick, 2010; but see Monthe, Hardy, Doucet, Loo, & Duminil, 2017).

The geometry of the study design we used may have limited our ability to accurately detect long-distance seed and pollen dispersal in this system. Although our 130-ha study site is relatively large, it may not capture the full spatial scale at which relevant avian frugivores or insect pollinators move. This could lead to underestimates of long-distance dispersal for both pollen and seed. In particular, we consider it possible that a bimodal distribution of seed dispersal distances related to frugivore dispersal processes may exist (e.g., Jordano et al., 2007), but may be difficult to detect due to the relative rarity of long-distance dispersal events. Although anecdotal data

from GPS tracking devices indicate umbrellabirds are capable of dispersing seeds several kilometers (Karubian & Durães, 2014), we observed low rates of seed immigration and a maximum seed dispersal distance of 656 m in this study, similar to a 3% rate of seed immigration and maximum seed dispersal of 645 m in a separate study with naturally dispersed *O. bataua* seeds (Browne & Karubian, 2018). There were also differences among our three plot types in their spatial arrangement relative to source trees and to the boundaries of our study area that may have led us to underestimate long-distance dispersal events (both pollen and seed) to Lek plots, due to their lower average distance to source trees and longer distance from the study boundary. More generally, we note that dispersal estimates from the NEIGHBORHOOD model and parentage assignment approaches were highly congruent; in the few cases where they differed (e.g., primary pollen dispersal among plots types), we prioritize results from the NEIGHBORHOOD model because, unlike direct parentage, it is designed to deal with inherent biases related to the probability of detecting a dispersal event as the distance from a source tree increases (Burczyk, Adams, Birkes, & Chybicki, 2006; Burczyk, Adams, Moran, & Griffin, 2002). Additional GPS tracking of frugivores and expansion of the area where adult trees are exhaustively sampled may be required to fully capture and characterize long-distance dispersal in this system.

Despite the strong evidence for greater pollen movement in this system, we found robust support for the idea that seed dispersal ultimately mediates genetic diversity and genetic structure of female gametes, male gametes and diploid seedlings at fine spatial scales. Relative to seedlings that were mainly gravity-dispersed into Beneath plots, seedlings dispersed by frugivores into Away and Lek plots were characterized by: (a) higher effective numbers of maternal source trees (N_{em}), paternal source trees (N_{ep}) and overall parents (N_e); (b) higher female gametic and diploid seedling diversity; and (c) reduced fine-scale spatial genetic structure for female gametes, male gametes and diploid seedlings. We first discuss the relative influence of dispersal distance versus N_{em} on seedling genetic diversity, before moving on to patterns of gametic diversity and fine-scale spatial genetic structure.

4.2 | Effective number of maternal (N_{em}) and paternal (N_{ep}) source trees

A comparison of Lek versus Away plots provides indirect insights into how two key attributes of frugivore-mediated seed dispersal—seed dispersal distance and N_{em}—might contribute to patterns of genetic diversity. The key point here is that N_{em} at Lek plots was double that of Away plots, whereas seed dispersal distances did not differ. Parenthetically, we note that a similar lack of concordance between dispersal distance and N_{em} has also been reported from other systems (e.g., Garcia et al., 2007; Sork et al., 2015) and may also exist for pollen dispersal distance and pollen source diversity N_{ep} (Ashley, 2010). In the *O. bataua* system, the lack of concordance between dispersal distance and N_{em} provides an opportunity to separate how each of these two attributes of dispersal may be

correlated with subsequent patterns of genetic diversity, with the caveat that our study is correlational and has not established causality (i.e., we cannot rule out the possibility that both N_{em} and the metrics of genetic diversity we evaluated both covary in response to an unmeasured ecological variable).

Higher N_{em} values at Lek plots were associated with higher N_{ep} and N_e values, lower β diversity and reduced spatial genetic structure for male gametes, female gametes and diploid seedlings. This is consistent with the idea that among dispersed seeds, N_{em} makes a direct and substantial contribution to local patterns of genetic diversity. Under some simple assumptions (e.g., a dispersal kernel model with independence of dispersal events), N_{em} varies as a function of the seed dispersal kernel and the effective density of mothers (cf. comparable expressions for N_{ep} in the TwoGener approach; Austerlitz & Smouse, 2001). When N_{em} is observed to vary independently of dispersal distance, as in the current study, variation in the density of maternal source trees, or some departure to assumptions of a simple dispersal model, or both are likely to be responsible. The density of *O. bataua* maternal source trees was stable during the study period, but the distribution of concurrently fruiting individuals varied month to month, due to asynchronous fruit production within and among individuals (J. Karubian, unpublished data). Moreover, distinctive foraging ecologies of *O. bataua's* dispersal agents likely led to departures from the assumptions of random, independent dispersal events of a simple dispersal model. It therefore stands to reason that the lack of concordance between N_{em} and dispersal distance in Lek versus Away plots may be associated with some interaction between the distribution of fruiting trees and disperser behaviour. Because umbrellabird leks are active year-round and at least some *O. bataua* adults are flowering and fruiting in all months of the year, leks receive a steady stream of dispersed seeds, allowing N_{em} to steadily accumulate over time. In contrast, Away plots putatively correspond to discrete dispersal episodes associated with ephemeral use of an area by other avian frugivores for activities such as roosting, resting, food processing or nesting, which may restrict N_{em} in Away plots relative to Lek plots. Additional work on this putative mechanistic relationship is needed before any firm conclusions can be drawn.

The effective number of paternal source trees per plot N_{ep} follows a qualitatively similar pattern to N_{em} (i.e., Beneath < Away < Lek). The relatively small number of samples we analysed per plot yields considerable uncertainty around the ratio of $N_{ep}:N_{em}$, especially for higher values, but qualitative trends may be of interest biologically. Most striking is the disproportionate increase in N_{ep} relative to N_{em} observed in Lek plots, driven by the fact that N_{ep} is considerably higher than in Away plots or Beneath plots. The disproportionate increase in N_{ep} in Lek plots is also likely responsible for the fact that effective number of parents N_e in Lek plots is almost three times that of Away plots, and an order of magnitude higher than Beneath plots. This qualitative trend is consistent with the idea that seeds dispersed into Lek plots represent a higher diversity of pollen sources per effective maternal seed source, relative to both Away and Beneath plots, although we urge caution in interpreting these results because error associated with these estimates is

particularly large in Lek plots. Our working hypothesis is that the steady accumulation of seeds into leks by umbrellabirds increases representation in these sites of pollen from adult trees that flower relatively infrequently, as well as distinctive combinations of pollen donors corresponding to different fruiting events by the same maternal tree. Both of these scenarios would inflate $N_{ep}:N_{em}$ in leks. In contrast, Away plots may represent relatively short-lived "snapshots" that sample available pollen sources over much shorter time intervals, potentially leading to lower $N_{ep}:N_{em}$. Here, too, a lack of details on these putative mechanisms hinders our ability to draw firm conclusions: continued research on how differences in frugivore foraging ecology interact with tree phenology to drive patterns of fine-scale genetic diversity in animal-dispersed plants is a priority in this and other systems (García & Grivet, 2011; Karubian & Durães 2009, 2014; Schupp, Milleron, & Russo, 2002; Scofield et al., 2012).

4.3 | Allelic diversity & spatial genetic structure

Away and Lek plots had higher α diversity (the effective number of alleles per locus per plot) for all gamete types than did Beneath plots, although this difference was significant only for female gametes and diploid seedling genotypes. In contrast, male gametic α diversity did not vary between plot types, presumably because the spatial scale of pollen movement in this system is so extensive that it masks any effect of seed dispersal on α diversity of male gametes. The lack of strong patterns for alleles relative to estimates of effective parental sizes is consistent with other studies (e.g., Grivet et al., 2009; Sork et al., 2015), suggesting that the effective number of pollen and seed donors can be more strongly asymmetric than genetic contributions of male and female gametes to the resulting seedling progeny.

For measures of between-plot overlap, between-patch β diversity values among female gametes were substantially higher for Beneath plots than for either Away or Lek plots, consistent with highly restricted seed movement leading to very little overlap between Beneath plots (the divergence estimator δ gave qualitatively similar results throughout). Also, there was more overlap between Lek plots than Away plots, even when controlling for the fact that Lek plots are more spatially proximate to each other than are Away plots (i.e., we obtained qualitatively similar results when only including those Away plots with similar pairwise distances in our analyses). This pattern, which corroborates that of Karubian et al. (2012), is likely driven by different male umbrellabirds from the same lek foraging from the same *O. bataua* adults surrounding the lek. Turning to pollen movement, β diversity values were close to 1 for male gametes, indicating very low between-plot diversity, coupled with the fact that we observe very little differentiation between plot types, are again consistent with widespread pollen movement in this system.

The examination of γ diversity, which provides an effective way of summarizing accumulated genetic diversity across plots, showed that seed dispersal distance did not impact the overall genetic contribution of female gametes on a sitewide scale. Differences between plot types were modest, and in fact, γ diversity of diploid seedlings was highest in Beneath plots, likely because of the low overlap in

female gametes, such that each new Beneath plot introduced a new suite of gametes to the cumulative measure of γ diversity. This finding also has potentially important implications for the ultimate impact of seed dispersal differences on genetic diversity, as it indicates no bottlenecks exist at the scale of the study plot despite strong indications of structure at final spatial scales. For male gametes, within-patch α diversity is essentially equivalent to cumulative, across-site γ diversity, reinforcing that pollen dispersal is so extensive in this system that it has almost no differential effect on subsequent allelic diversity for dispersed (i.e., Away and Lek) versus undispersed (Beneath) seedlings.

The analysis of fine-scale genetic structure offers a spatial view of allelic diversity across the landscape and, in the context of the current study, allows insights into how seed dispersal may be impacting these patterns. The slow decline in fine-scale structure observed in both female gametic and diploid seedling correlograms in Lek and Away plots is consistent with relatively extensive seed dispersal and seed mixing, an inference also supported by the relatively low levels of β diversity observed among these plots. In contrast, the sharp drop off from the first to the second distance class observed in Beneath plots (and the associated increase in β diversity among Beneath plots) is presumably driven by the lack of seed dispersal and seed sharing between neighbouring plots. This same qualitative pattern of much higher spatial genetic structure among undispersed versus dispersed seedlings parallels patterns found in other systems, particularly oaks (Hampe, El Masri, & Petit, 2010; Sork et al., 2015). Consistent with higher N_{em} in Lek plots relative to Away plots, we observed weaker structure among Lek seedlings than Away seedlings at all distance classes, although this difference was not statistically significant for individual distance classes. Altogether, the spatial autocorrelation analyses support the notion that differences in seed dispersal contribute to fine-scale spatial genetic structure of seedlings; however, a major unanswered question is the degree to which these differences among plot types persist across life stages and impact the genetic structure of adults.

5 | CONCLUSIONS

This study provides a compelling argument for the importance of both pollen and seed dispersal in maintaining effective gene flow and genetic diversity. In our study system, pollen moves considerably further than do seeds and appears to be responsible for most long-distance gene flow. However, a lack of seed dispersal creates strong spatial genetic structure among seedlings despite robust pollen movement, supporting the perspective that seed dispersal is a key determinant of fine-scale genetic structure among seedlings (Hamrick, 2010; Sork & Smouse, 2006). Future research might extend this work on seedlings to investigate the ultimate consequences of these dispersal differences for genetic structure of adults; it is possible that effects may be limited due to and a survival advantage for rare genotypes (Browne & Karubian, 2016, 2018) in this study population.

Importantly, documenting the spatial extent of pollen and seed movement alone was not adequate to characterize dispersal in our study system; considering the diversity of maternal seed sources N_{em} and paternal pollen sources N_{ep} substantively improved our ability to assess the specific ways in which seed dispersal shapes genetic diversity and fine-scale genetic structure. Frugivore seed dispersal agents ultimately determine the diversity of seed sources, and our analyses also highlight how differing foraging ecologies may lead to differing genetic outcomes among dispersed seedlings (e.g., Lek vs. Away plots). Future work might explore in greater detail the proportion of seeds dispersed by different frugivores (e.g., Jordano et al., 2007) and how their various deposition patterns interact to shape observed patterns of diversity among seedlings. This might also have important conservation implications, as some key frugivores (e.g., umbrellabirds) in the *O. bataua* system are directly threatened by hunting, habitat degradation and habitat loss whereas others (e.g., toucans) appear to be more resilient (Walter et al., 2017). For example, in a companion paper we show evidence for a stronger impact of forest loss and fragmentation on seed dispersal than pollen flow for *O. bataua*, consistent with shifts in the relative contribution of different foraging ecologies to the dispersal and genetic structure of *O. bataua* (Browne & Karubian, 2018). In closing, we suggest future studies would benefit from considering both dispersal distance and effective number of sources at deposition sites (as opposed to dispersal distance alone) when attempting to characterize how seed dispersal may impact genetic diversity and fine-scale spatial genetic structure.

DATA ACCESSIBILITY

Sampling locations and microsatellite for *Oenocarpus bataua* adult and seedlings genotyped in this study are publicly available and archived on Figshare (https://doi.org/10.6084/m9.figshare.6434144). The R code used for gametic extraction is available on Github (https://github.com/lukembrowne/gametic-extraction; https://doi.org/10.5281/zenodo.1274538).

AUTHOR CONTRIBUTION

L.B. contributed to research design, sample collection, laboratory analysis, data analysis, and writing the manuscript. K.O. contributed to sample collection, laboratory analysis, interpretation, and writing the manuscript. V.L.S. contributed to research design, analysis, interpretation, and writing the manuscript. J.K. designed and led the research and contributed to analysis, interpretation, and writing the manuscript.

ORCID

Luke Browne http://orcid.org/0000-0002-2239-3007

Kym Ottewell http://orcid.org/0000-0002-4273-3224

Victoria L. Sork http://orcid.org/0000-0003-2677-1420

Jordan Karubian http://orcid.org/0000-0001-8201-9992

REFERENCES

Abedi-Lartey, M., Dechmann, D. K., Wikelski, M., Scharf, A. K., & Fahr, J. (2016). Long-distance seed dispersal by straw-coloured fruit bats varies by season and landscape. *Global Ecology and Conservation*, 7, 12–24.

Ashley, M. V. (2010). Plant parentage, pollination and dispersal: How DNA microsatellites have altered the landscape. *Critical Reviews in Plant Science*, 29, 148–161.

Austerlitz, F., & Smouse, P. E. (2001). Two-generation analysis of pollen flow across a landscape. III. Impact of adult population structure. *Genetical Research*, 78, 271–280.

Bittencourt, J. V. M., & Sebbenn, A. M. (2007). Patterns of pollen and seed dispersal in a small, fragmented population of the wind-pollinated tree *Araucaria angustifolia* in southern Brazil. *Heredity*, 99, 580–591.

Browne, L., & Karubian, J. (2018). Habitat loss and fragmentation reduce effective gene flow by disrupting seed dispersal in a neotropical palm. *Molecular Ecology*, https://doi.org/10.1111/mec.14765

Browne, L., & Karubian, J. (2016). Frequency dependent selection for rare genotypes promotes genetic diversity of a tropical palm. *Ecology Letters*, 19, 1439–1447.

Browne, L., & Karubian, J. (2018). Rare genotype advantage promotes survival and genetic diversity of a tropical palm. *New Phytologist*, 218, 1658–1667.

Browne, L., Ottewell, K., & Karubian, J. (2015). Short-term genetic consequences of habitat loss and fragmentation for the neotropical palm *Oenocarpus bataua*. *Heredity*, 115, 389–395.

Bullock, J. M., Mallada González, L., Tamme, R., Götzenberger, L., White, S. M., Pärtel, M., & Hooftman, D. A. (2017). A synthesis of empirical plant dispersal kernels. *Journal of Ecology*, 105, 6–19.

Burczyk, J., Adams, W. T., Birkes, D. S., & Chybicki, I. J. (2006). Using genetic markers to directly estimate gene flow and reproductive success parameters in plants on the basis of naturally regenerated seedlings. *Genetics*, 173, 363–372.

Burczyk, J., Adams, W. T., Moran, G. F., & Griffin, A. R. (2002). Complex patterns of mating revealed in a *Eucalyptus regnans* seed orchard using allozyme markers and the neighbourhood model. *Molecular Ecology*, 11, 2379–2391.

Chybicki, I. J. (2017). NMπ—improved re-implementation of NM+, a software for estimating gene dispersal and mating patterns. *Molecular Ecology Resources*, 18(1), 159–168.

Chybicki, I. J., & Burczyk, J. (2010). NM+: Software implementing parentage-based models for estimating gene dispersal and mating patterns in plants. *Molecular ecology resources*, 10, 1071–1075.

Crawford, T. J. (1984). The estimation on neighbourhood parameters for plant populations. *Heredity*, 52(2), 273–283. https://doi.org/10.1038/hdy.1984.29

Ennos, R. (1994). Estimating the relative rates of pollen and seed migration among plant populations. *Heredity*, 72, 250–259.

García, C., & Grivet, D. (2011). Molecular insights into seed dispersal mutualisms driving plant population recruitment. *Acta Oecologica*, 37, 632–640.

García, C., Jordano, P., Arroyo, J. M., & Godoy, J. A. (2009). Maternal genetic correlations in the seed rain: Effects of frugivore activity in heterogeneous landscapes. *Journal of Ecology*, 97, 1424–1435.

Garcia, C., Jordano, P., & Godoy, J. A. (2007). Contemporary pollen and seed dispersal in a *Prunus mahaleb* population: Patterns in distance and direction. *Molecular Ecology*, 16, 1947–1955.

Giombini, M. I., Bravo, S. P., & Tosto, D. S. (2016). The key role of the largest extant Neotropical frugivore (*Tapirus terrestris*) in promoting admixture of plant genotypes across the landscape. *Biotropica*, 48, 499–508.

Goulding, M., & Smith, N. (2007). *Palms: Sentinels for Amazon conservation*. Saint Louis: Missouri Botanical Garden Press.

Grivet, D., Robledo-Arnuncio, J. J., Smouse, P. E., & Sork, V. L. (2009). Relative contribution of contemporary pollen and seed dispersal to the effective parental size of seedling population of California valley oak (*Quercus lobata*, Née). *Molecular Ecology*, 18, 3967–3979.

Grivet, D., Smouse, P. E., & Sork, V. L. (2005). A novel approach to an old problem: Tracking dispersed seeds. *Molecular Ecology*, 14, 3585–3595.

Hamilton, M. B. (1999). Tropical tree gene flow and seed dispersal. *Nature*, 401, 8–9.

Hamilton, M. B., & Miller, J. R. (2002). Comparing relative rates of pollen and seed gene flow in the island model using nuclear and organelle measures of population structure. *Genetics*, 162, 1897–1909.

Hampe, A., El Masri, L., & Petit, R. J. (2010). Origin of spatial genetic structure in an expanding oak population. *Molecular Ecology*, 19(3), 459–471.

Hamrick, J. L. (2010). Pollen and seed movement in disturbed tropical landscapes. In J. A. DeWoody, C. H. Michler, et al. (Eds.), *Molecular approaches in natural resource conservation and management* (pp. 190–211). New York: Cambridge University Press.

Hanson, T. R., Brunsfeld, S. J., Finegan, B., & Waits, L. P. (2008). Pollen dispersal and genetic structure of the tropical tree *Dipteryx panamensis* in a fragmented Costa Rican landscape. *Molecular Ecology*, 17, 2060–2073.

Hardy, O. J., González-Martínez, S. C., Colas, B., Fréville, H., Mignot, A., & Olivieri, I. (2004). Fine-scale genetic structure and gene dispersal in Centaurea corymbosa (Asteraceae). II. Correlated paternity within and among sibships. *Genetics*, 168(3), 1601–1614.

Hardy, O. J., & 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.

Heuertz, M., Vekemans, X., Hausman, J. F., Palada, M., & Hardy, O. J. (2003). Estimating seed vs. pollen dispersal from spatial genetic structure in the common ash. *Molecular Ecology*, 12, 2483–2495.

Heywood, J. S. (1991). Spatial analysis of genetic variation in plant populations. *Annual Review of Ecology and Systematics*, 22(1), 335–355.

Holbrook, K. M. (2011). Home range and movement patterns of toucans: Implications for seed dispersal. *Biotropica*, 43, 357–364.

Holbrook, K. M., Smith, T. B., & Hardesty, B. D. (2002). Implications of long-distance movements of frugivorous rain forest hornbills. *Ecography*, 25(6), 745–749.

Holland, R. A., Wikelski, M., Kümmeth, F., & Bosque, C. (2009). The secret life of oilbirds: New insights into the movement ecology of a unique avian frugivore. *PLoS ONE*, 4, e8264.

Isagi, Y., Saito, D., Kawaguchi, H., Tateno, R., & Watanabe, S. (2007). Effective pollen dispersal is enhanced by the genetic structure of an *Aesculus turbinata* population. *Journal of Ecology*, 95, 983–990.

Jordano, P. (2010). Pollen, seeds and genes: The movement ecology of plants. *Heredity*, 105, 329–330. https://doi.org/10.1038/hdy.2010.28

Jordano, P., García, C., Godoy, J. A., & García-Castaño, J. L. (2007). Differential contribution of frugivores to complex seed dispersal patterns. *Proceedings of the National Academy of Sciences*, 104, 3278–3282.

Karubian, J., & Durães, R. (2009). Effects of seed disperser social behavior on patterns of seed movement and deposition. *Oecologia Brasiliensis*, 13(1), 45–57.

Karubian, J., & Durães, R. (2014). Impacts of mating behavior on plant-animal seed dispersal mutualisms: A case study from a Neotropical lek-breeding bird. In R. Macedo, & G. Machado (Eds.), *Sexual selection: insights from the neotropics* (pp. 365–390). London: Elsevier Press.

Karubian, J., Duraes, R., Storey, J. L., & Smith, T. B. (2012). Mating behavior drives seed dispersal by the long-wattled Umbrellabird *Cephalopterus penduliger*. *Biotropica*, 44, 689–698.

Karubian, J., Ottewell, K., Link, A., & Di Fiore, A. (2015). Genetic consequences of seed dispersal to sleeping trees by white-bellied spider monkeys. *Acta Oecologica*, 68, 50–58.

Karubian, J., Sork, V. L., Roorda, T., Duraes, R., & Smith, T. B. (2010). Destination-based seed dispersal homogenizes genetic structure of a tropical palm. *Molecular Ecology*, 19(8), 1745–1753.

Levin, D. A. (1981). Dispersal versus gene flow in plants. *Annals of the Missouri Botanical Garden*, 68, 233–253.

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, 1420–1425.

Marshall, T. C., Slate, J. B. K. E., Kruuk, L. E. B., & Pemberton, J. M. (1998). Statistical confidence for likelihood-based paternity inference in natural populations. *Molecular Ecology*, 7, 639–655.

Monthe, F. K., Hardy, O. J., Doucet, J.-L., Loo, J., & Duminil, J. (2017). Extensive seed and pollen dispersal and assortative mating in the rain forest tree *Entandrophragma cylindricum* (Meliaceae) inferred from indirect and direct analyses. *Molecular Ecology*, 26, 5279–5291.

Nakanishi, A., Tomaru, N., Yoshimaru, H., Manabe, T., & Yamamoto, S. (2009). Effects of seed-and pollen-mediated gene dispersal on genetic structure among *Quercus salicina* saplings. *Heredity*, 102, 182–189.

Nathan, R. (2006). Long-distance dispersal of plants. *Science*, 313, 86–788.

Núñez-Avellaneda, L., & Rojas-Robles, R. (2008). Reproductive biology and pollination ecology of the milpesos palm *Oenocarpus bataua* in the Colombian Andes. *Caldasia*, 30, 101–125.

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.

R Core Team (2016) R: A language and environment for statistical computing.

Robledo-Arnuncio, J. J., Grivet, D., Smouse, P. E., & Sork, V. L. (2012). PSA: Software for parental structure analysis of seed or seedling patches. *Molecular Ecology Resources*, 12, 1180–1189.

Rojas-Robles, R., & Stiles, F. G. (2009). Analysis of a supra-annual cycle: Reproductive phenology of the palm *Oenocarpus bataua* in a forest of the Colombian Andes. *Journal of Tropical Ecology*, 25, 41–51.

Schupp, E. W., Milleron, T., & Russo, S. E. (2002). Dissemination limitation and the origin and maintenance of species-rich tropical forests. In D. J. Levey, W. R. Silva, & M. Galetti (Eds.), *Seed dispersal and frugivory: Ecology, evolution and conservation* (pp. 19–33). New York: CABI Publishing.

Scofield, D. G., Smouse, P. E., Karubian, J., & Sork, V. L. (2012). Use of alpha, beta, and gamma diversity measures to characterize seed dispersal by animals. *The American Naturalist*, 180, 719–732.

Smouse, P., Dyer, R., Westfall, R., & Sork, V. L. (2001). Two-generation analysis of pollen flow across a landscape. I. male gamete heterogeneity among females. *Evolution*, 55, 260–271.

Sork, V. L., & Smouse, P. E. (2006). Genetic analysis of landscape connectivity in tree populations. *Landscape Ecology*, 21, 821–836.

Sork, V. L., Smouse, P. E., Grivet, D., & Scofield, D. G. (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, N. C., Sabatier, D., Baraloto, C., Salomão, R. P., Guevara, J. E., … Monteagudo, A. (2013). Hyperdominance in the Amazonian tree flora. *Science*, 342, 325–342.

Vekemans, X., & Hardy, O. J. (2004). New insights from fine-scale spatial genetic structure analyses in plant populations. *Molecular Ecology*, 13, 921–935.

Walter, S. T., Browne, L., Freile, J., Olivo, J., Gonzalez, M., & Karubian, J. (2017). Surrounding tree cover predicts diversity of large-bodied frugivorous birds in forest fragments. *Biotropica*, 49, 838–847.

Wright, S. (1943). Isolation by distance. *Genetics*, 28, 114–138.

Wright, S. (1946). Isolation by distance under diverse systems of mating. *Genetics*, 31, 39–59.

SUPPORTING INFORMATION

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