

OPEN

# Janzen-Connell effects shape gene flow patterns and realized fitness in the tropical dioecious tree *Spondias purpurea* (ANACARDIACEAE)

E. Jacob Cristóbal-Pérez<sup>1,2</sup>, Eric J. Fuchs<sup>2,4</sup>, Ulises Olivares-Pinto<sup>2,3</sup> & Mauricio Quesada<sup>1,2\*</sup>

Pollination and seed dispersal patterns determine gene flow within plant populations. In tropical forests, a high proportion of trees are dioecious, insect pollinated and dispersed by vertebrates. Dispersal vectors and density dependent factors may modulate realized gene flow and influence the magnitude of Fine Scale Genetic Structure (FSGS), affecting individual fitness. *Spondias purpurea* is a vertebrate-dispersed, insect-pollinated dioecious tropical tree. We assessed the influence of sex ratio, effective and realized gene flow on genetic diversity, FSGS and individual fitness within a 30 ha plot in the tropical dry forest reserve of Chamela-Cuixmala, Mexico. All individuals within the plot were tagged, geo-referenced and sampled for genetic analysis. We measured dbh and monitored sex expression during two reproductive seasons for all individuals. We collected seeds directly from maternal trees for effective pollen dispersal analysis, and analyzed established seedlings to assess realized pollen and seed dispersal. Nine microsatellite loci were used to describe genetic diversity parameters, FSGS and gene flow patterns among different size classes. A total of 354 individuals were located and classified into three size classes based on their dbh (<10, 10–20, and >20 cm). Population sex ratios were male biased and diametric size distributions differed among sexes, these differences may be the result of precocious male reproduction at early stages. Autocorrelation analyses indicate low FSGS ( $F_j < 0.07$ ) across all size classes. Long realized pollen and seed dispersal and differences among effective and realized gene flow were detected. In our study site low FSGS is associated with high gene flow levels. Effective and realized gene flow indicate a population recruitment curve indicating Janzen-Connell effects and suggesting fitness advantages for long-distance pollen and seed dispersal events.

Dispersal is a central ecological mechanism that involves the movement of genes or individuals from one location to another. In plants, gene movement occurs through pollen flow (paternal gamete dispersal) and seed dispersal, which is facilitated by animals such as pollinators and seed dispersers or abiotic agents like wind and water<sup>1,2</sup>. Gene flow patterns are important for population and conservation genetics since they directly influence genetic diversity within populations and determine the genetic structure among populations<sup>3</sup>. Within plant populations, limited gene flow could result in fine spatial genetic structure (FSGS), that is, the non-random spatial distribution of genotypes<sup>4</sup>, resulting in neighborhoods of related individuals at a restricted spatial scale<sup>4,5</sup>. Other factors besides gene flow may generate FSGS, including colonization history<sup>6,7</sup>, demographic changes<sup>8–10</sup> and the mating system and sexual expression of plants<sup>3,11</sup>.

Spatial distribution patterns in plants are determined by seed dispersal patterns and post-dispersal processes such as seed predation and seedling survival. For example, in zoochorous plants<sup>12</sup> short-range dispersers forage near fruiting trees, resulting in higher seed deposition near maternal plants and FSGS. Further, seeds may accumulate in roosting or sleeping sites which would again result in aggregated dispersal and consequently resulting in FSGS<sup>12</sup>. According to the Janzen-Connell hypothesis, the establishment and survivorship of seedlings near

<sup>1</sup>Instituto de Investigaciones en Ecosistemas y Sustentabilidad, Universidad Nacional Autónoma de México, Morelia, Michoacán, México. <sup>2</sup>Laboratorio Nacional de Análisis y Síntesis Ecológica, Escuela Nacional de Estudios Superiores Unidad Morelia, Universidad Nacional Autónoma de México, Morelia, Michoacán, México. <sup>3</sup>Escuela Nacional de Estudios Superiores Unidad Juriquilla, Universidad Nacional Autónoma de México, Santiago de Querétaro, Querétaro, México. <sup>4</sup>Escuela de Biología, Universidad de Costa Rica, San José, 11501-2060, Costa Rica. \*email: [mquesada@cieco.unam.mx](mailto:mquesada@cieco.unam.mx)

maternal or conspecific plants could be affected by negative density-dependent processes such as intra-specific competition, host-specific predation, herbivory or pathogens<sup>13,14</sup>. Although density-dependence is commonly accepted as a factor responsible for high tropical diversity<sup>15</sup>, its influence on FSGS patterns is poorly understood. Proximate density dependent factors may increase mortality of seedlings near conspecifics or at high densities<sup>13,14</sup>, reducing FSGS. Dioecious trees are relatively common in tropical forests (ca. 27% trees)<sup>16–18</sup>. Dioecy is a reproductive system in which the male and female reproductive functions exist in separate individuals. A high proportion of dioecious species share reproductive traits such as small insect pollinated flowers and fleshy fruits dispersed by vertebrates<sup>18,19</sup>. Although dioecious trees are frequent in tropical forests, gene flow patterns<sup>20,21</sup> and FSGS have been rarely studied<sup>11,22</sup>. In dioecious trees the magnitude of FSGS will depend, on the one hand, on the movement ability and behaviour of pollinators and seed dispersers and, on the other hand, on demographic traits associated with the reproductive system, such as sexual ratio and spatial distribution of adults. Theory of density-dependent animal pollination assumes that tree species occurring at low densities receive pollen from fewer individuals than trees in denser populations, and from longer pollen flow distances<sup>23–26</sup>. Theory predicts a 1:1 sex ratio maintained by negative frequency-dependent selection<sup>27</sup>, however intersex differences in growth, reproduction and survival may result in biases in sex ratio<sup>28</sup>. Specifically, males are expected to have higher growth rates, survivorship and reproduction than females, due to the higher energy costs associated with gamete and fruit production in females; resulting in male-biased sex ratios<sup>29</sup>. For dioecious plants, the spatial distribution of male and female trees as well as the population sex ratios could affect seed disperser behaviour playing an important role in determining FSGS patterns.

Fitness may be defined as an individual's genetic contribution to future generations<sup>30</sup>. The number of surviving offspring of a given individual in the population, determined through paternity analyses, is a good proxy of plant fitness<sup>31,32</sup>. In plants, individual fitness will depend on their ability to produce viable seeds (i.e. fecundity) and the success of seeds to disperse and establish (i.e. survival). Therefore, factors affecting individual plant fitness as pollen gene flow distances, seed or seedling predation near maternal trees may affect FSGS formation. Due to the difficulty to estimate overall fitness in natural conditions, fitness analyses frequently rely on traits associated with one component of individual fitness, generally associated to fecundity (e.g. fruit production, seed production)<sup>33</sup>, ignoring the transition between seed formation and seedling emergence<sup>34</sup>. An approach to analyze fitness and fill the gap in this demographic transition is to estimate effective pollen flow and realized pollen and seed dispersal. Effective pollen dispersal is the pollen flow that occurs before seed dispersal, while realized pollen and seed dispersal refer to pollen and seed flow measured on established seedlings and juveniles<sup>35</sup>. This approach allows the evaluation of individual fitness through an estimate of the number of offspring by each parental plant, and secondly to analyse the contribution of gene flow patterns on FSGS.

In this study, we analyze fine-scale genetic structure (FSGS) in relation to gene flow patterns and fitness on the dioecious, insect pollinated and vertebrate seed dispersed tree, *Spondias purpurea* L. We analyse the effect of demography, gene flow and sex ratio on effective gene flow patterns. Specifically, we evaluate the following objectives: (1) Determine the relationship between population sex ratios, size classes and FSGS; (2) Determine the relationship between gene flow patterns via seed and pollen dispersal, with respect to FSGS, and (3) Determine the relationship of effective and realized pollen dispersal, as a proxy of plant fitness in relation to FSGS.

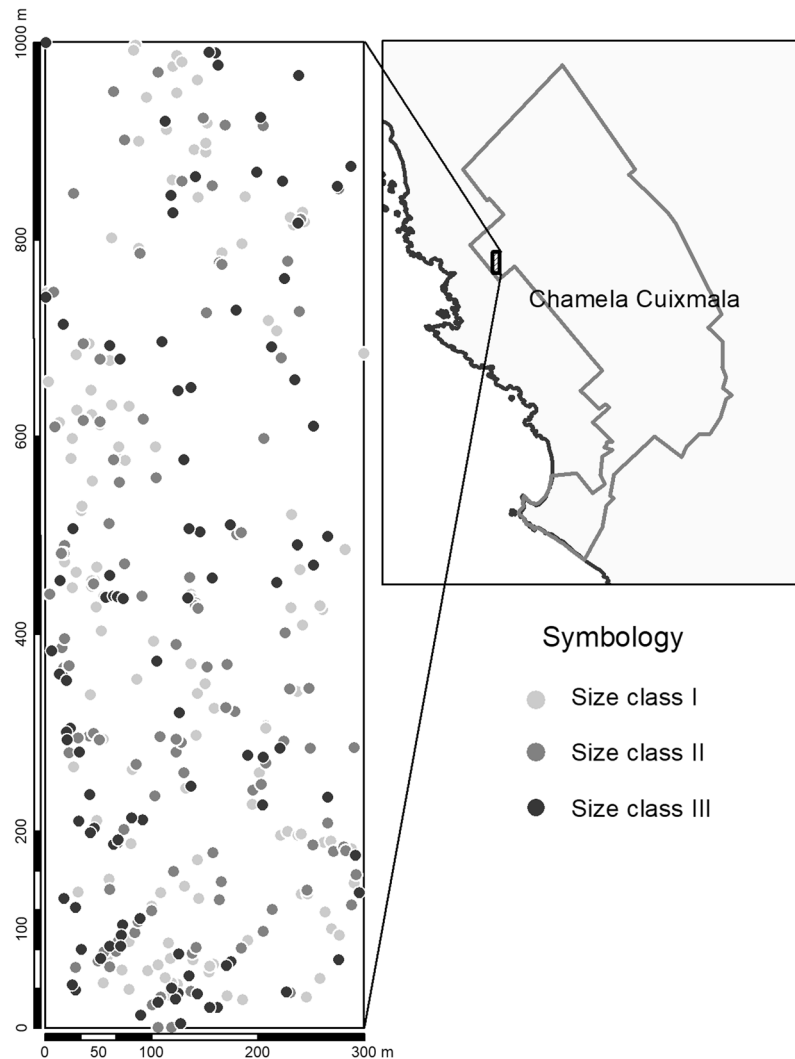
## Results

**Sexual ratios.** A total of 354 *S. purpurea* individuals were mapped within the 30 ha plot, 144 individuals were included in size class I, 113 in size class II and 97 individuals in size class III (Fig. 1). During two reproductive seasons we observed 62 size class I and 32 size class II individuals that did not flower, these individuals were included as non-reproductive individuals (Fig. 2). We observed significant differences in sex ratios across size classes ( $G_{\text{heterogeneity}} = 22.809$ , d.f. = 2,  $p = 0.0011$ ). A male-biased sex ratio was observed for the pooled population and for size classes I and II ( $p < 0.05$ ; Fig. 2). For size class III, the sex ratio did not significantly deviate from unity ( $p > 0.05$ ) (Fig. 2). We observed significant differences in diameter size distribution between sexes (Fig. 2,  $D = 0.33$ ,  $p < 0.001$ ).

**Genetic diversity analysis.** All 354 mapped individuals were genetically analysed. All nine loci were polymorphic and had similar levels of genetic diversity across size classes (Table 1). For the entire population, the average number of alleles per locus was  $N_a = 6.51$  ( $\pm 0.231$ ), the allelic richness was  $A_r = 6.07$  ( $\pm 0.582$ ), the observed and expected heterozygosities were  $H_o = 0.430$  ( $\pm 0.09$ ) and  $H_e = 0.502$  ( $\pm 0.006$ ) respectively.

**Fine scale genetic structure and spatial analysis.** Our results indicated relatively low levels of FSGS across different size classes. Individuals of size class I had significant but low pairwise kinship coefficient for the 20 and 40 m distance classes,  $F_{ij} = 0.0531$  and  $F_{ij} = 0.069$  respectively (Fig. 3A). Individuals of size class II had significant FSGS for the 20 m distance class ( $F_{ij} = 0.077$ ) (Fig. 3B), similarly size class III individuals showed significant family structure for the 20 m distance class ( $F_{ij} = 0.05$ ) (Fig. 3C). Although significant for the first distance classes, kinship coefficients were low ( $F_{ij} \leq 0.077$ ) in this *S. purpurea* population. The *Sp* statistic confirms these results, with all regression slopes being statistically significant (Table 2). The between-size class FSGS analysis, showed that size class I individuals exhibited low genetic relatedness with nearby size class III individuals ( $F_{ij} = 0.043$  and  $F_{ij} = 0.030$  for 20 and 40 m respectively), than expected for parent-progeny pairs ( $F_{ij} \approx 0.5$ ) (Fig. 4). Results of spatial interaction analyses indicated spatial association between individuals of size class I and III (Fig. 5).

**Parentage analysis.** The exclusion probability estimated over all individuals and over nine loci was 99.99%. Out of 127 genotyped seeds for effective pollen dispersal estimates, 68 were assigned to pollen donors within the plot, 58 of which with 95% confidence. We were able to identify six pollen donors within the plot that sired between 5 and 12 seeds (mean = 9.6 seeds by pollen donor). Effective pollen dispersal distances ranged between 27.9 m



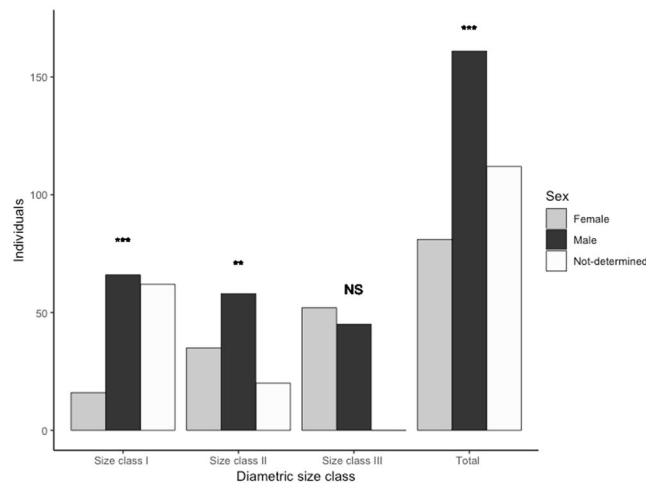
**Figure 1.** Spatial distribution *Spondias purpurea* individuals in a 30 ha plot located in Chamela, Jalisco Mexico. Size class I (DBH < 10 cm), Size class II (10 cm < DBH < 20 cm), Size class III (DBH > 20 cm). This map was created in Quantum GIS v.3.4 (Quantum GIS Development Team (2018). Quantum GIS Geographic Information System. Open Source Geospatial Foundation Project. <http://qgis.osgeo.org>).

and 828.2 m with an average of 306.78 m ( $\pm 28.57$ ) and a median of 263.29 m. Approximately 30% of all effective pollination events assigned through paternity analysis, occurred between trees that were less than 150 m apart, 70% occurred between trees >200 and 600 m apart and 12% occurred between trees 600 to 850 m apart (Fig. 6A).

We determined realized pollen dispersal distances on 144 individuals in size class I, 31 were assigned to a parent-pair within the plot, 29 of which were assigned with >95% confidence. We were able to identify 11 pollen donors within the plot that sired between 1 and 5 established seedlings by each donor (mean = 2.8 seedling by pollen donor). The realized pollen dispersal distance ranged from 96.6 m to 854.6 m with an average of 331.6 m ( $\pm 31.14$ ) and a median of 293 m. For realized pollen dispersal, only 6.4% of the events occurred between trees <150 m apart, 93% occurred between trees >150 m apart and 12% occurred between trees 600 to 854 m apart (Fig. 6B).

The same 144 individuals of the previous analyses were used to estimate seed dispersal distances, 77 of these seedlings were assigned to a mother (13 maternal plants) within the plot (67 with >95% confidence). Seed dispersal distances ranged between 7.2 m and 833.2 m with an average of 305 m ( $\pm 20.6$ ) and a median of 255.9 m, 19.4% of all seed dispersal events occurred at distances <150 m apart from assigned maternal trees, 80.5% occurred between 150 m and 600, and 12.9% occurred between 600 and 833 m apart from the assigned maternal trees (Fig. 6C).

We did not find any significant evidence of a relationship between the number of established progeny produced by an individual and its DBH ( $r^2 < 0.44$ ,  $p > 0.144$ ), nor between the number of progeny established and their distance from any given maternal tree ( $r^2 = 0.09$ ,  $p = 0.99$ ).



**Figure 2.** Number of individuals in each size class for each sex and for not-determined individuals over two reproductive seasons of *Spondias purpurea* within a 30 ha plot in Chamela Jalisco, Mexico. Notes: \*\*\*indicates  $p < 0.001$ , \*\* $p < 0.05$ , NS:  $p > 0.05$  (G-test goodness of fit).

Size class	N	$N_a$	$A_r$	$H_o$	$H_e$	F
Size class I	144	6.333 (0.273)	6.02 (1.05)	0.41822 (0.077)	0.49947 (0.078)	0.116 (-0.0692–0.3636)
Size class II	113	6.444 (0.297)	6.08 (1.01)	0.43545 (0.071)	0.49751 (0.075)	0.072 (-0.1256–0.3179)
Size class III	97	6.778 (0.037)	6.12 (1.08)	0.43775 (0.082)	0.51023 (0.078)	0.099 (-0.1151–0.3545)

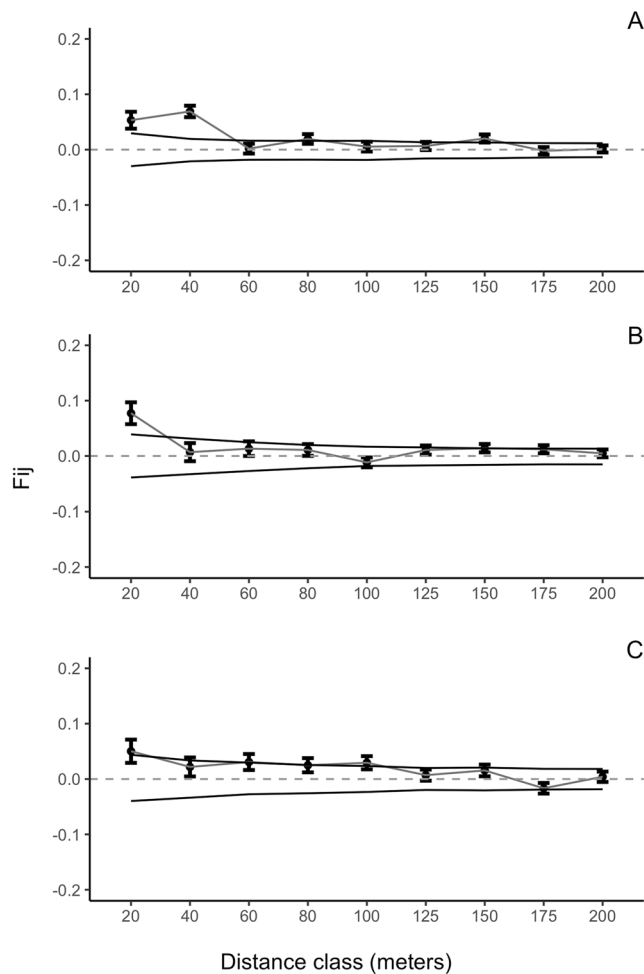
**Table 1.** Genetic diversity estimates for each size class within 30 ha plot of *Spondias purpurea*. N: Number of individuals;  $N_a$ : mean allele number per locus ( $\pm$ SD);  $A_r$ : Allelic richness ( $\pm$ SD);  $H_o$ : observed heterozygosity ( $\pm$ SD);  $H_e$ : expected heterozygosity ( $\pm$ SD); F: fixation index (Bootstrap estimates of confidence intervals).

## Discussion

Sex ratios in the studied *S. purpurea* population were significantly male-biased, but departures from 1:1 ratio varied across sizes classes. While smaller size classes (DBH < 20 cm) were mostly male-biased, the largest size classes (>20 cm) did not deviate from a 1:1 sex ratio. In the two sampling periods of this study, we did not find differences in flowering frequency between male and female *Spondias* trees across size classes. Differences in size distributions between males and females are unlikely due to greater growth rates in males, but instead by an overrepresentation of males in the smaller size classes. The 1:1 sex ratio in the larger size classes probably indicates that many of the non-flowering individuals in the smaller size classes are likely females. Thus, *S. purpurea* females reach age at first reproduction at a minimum threshold size determined by a minimum budget of resources necessary for flower and fruit production. Flower production in males start at an early stage in individuals that are 2 cm in diameter and less than 1 m in height; while in females, flowering starts when they grow taller than 5.2 cm in diameter and >1 m in height. This result suggests precocious male reproduction in *S. purpurea*, as also found in other tropical species<sup>36–41</sup>. Female trees likely require a minimum branch size to bear fruit (i.e. >1 cm) which is related to the overall development of the plant.

We observed intermediate levels of genetic diversity in all three size classes of *S. purpurea* as expected for tropical trees<sup>42–44</sup> and for dioecious species with obligate outcrossing<sup>45,46</sup>. Our data showed significant but low fine-scale structure up to 40 meters for the smallest size class and up to 20 meters for size classes II and III.  $Sp$  values were comparable to estimates for animal dispersed tropical trees ( $Sp \in [0.01 \dots 0.03]$ ) and showed no evidence of FSGS across all size classes, which indicates that genotypes were largely randomly distributed in space. These results contradict expectations of high and significant structure in the smaller size classes of animal dispersed trees due to limited seed dispersal<sup>20,45,47,48</sup>, and a reduction in FSGS in larger diametric size classes due to post-dispersal demographic thinning (e.g. density-dependent predation and competition)<sup>20,45,48–50</sup>. Our expectations were based on previous findings for other vertebrate dispersed tropical dioecious tree species such as *Pouteria reticulata* (Sapotaceae), *Simarouba amara* (Simaroubaceae), *Ficus cyrtophylla* (Moraceae), and *Protium spruceanum* (Burseraceae); where significant FSGS was detected in sapling and adult size classes at short distances, and FSGS declined in larger diametric classes<sup>20,45,51,52</sup>. This pattern has also been shown for dioecious species such as *F. hispida*, *F. exasperata*, *F. pumila*, *Ceratiola ericoides*, *Eurya emarginata*, *Myracrodruon urundeuva*<sup>46,53–56</sup>; however this is not the pattern that we found for *S. purpurea* at Chamela, our data shows low genetic structure and no differences among size classes.

The unexpected low FSGS in *S. purpurea* is likely a result of long-distance seed dispersal and early density-dependent factors shaping FSGS patterns. In the study area two bird species (*Ortalis poliocephala* and

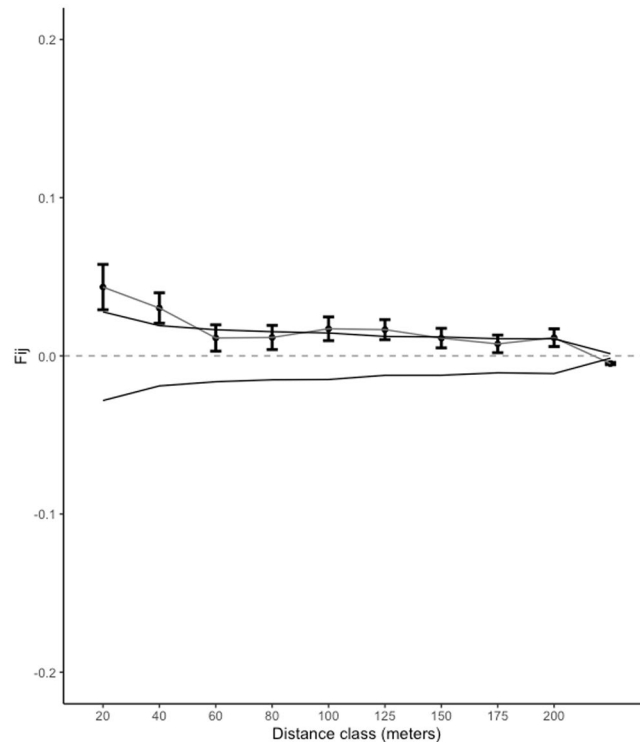


**Figure 3.** Autocorrelograms for estimated pairwise kinship correlation  $F_{ij}$  for three diametric size classes: **(A)** size class I (DBH < 10 cm), **(B)** size class II (10 < DBH ≤ 20 cm), **(C)** size class III (DBH > 20 cm). Each point represents calculated  $r_j$  values for each distance class. The solid lines represent upper and lower 95% confidence envelopes for the null hypothesis of no correlation ( $r_j = 0$ ).

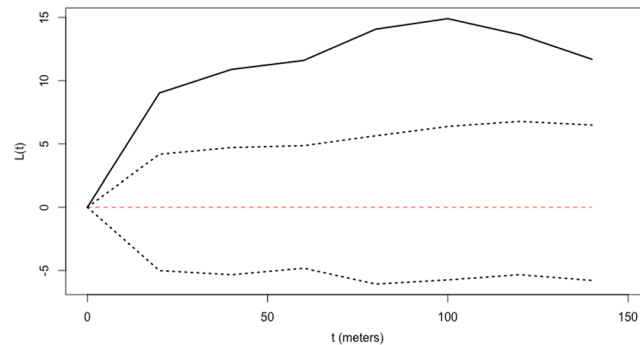
Size class	$b_{\log}$	$F_{(1)}$	Sp
Size class I	-0.0121 <sup>NS</sup>	0.0531	0.0128
Size class II	-0.01 <sup>NS</sup>	0.0771	0.0108
Size class III	-0.01 <sup>NS</sup>	0.0502	0.0105

**Table 2.** Slope of the regression ( $b_{\log}$ ) between kinship coefficient and the log of the geographic distance between individuals, the average kinship between individuals separated by less than 20 meters ( $F_{(1)}$ ) and the Sp statistic calculated for each size class within the plot. NS: Not significant.

*Icterus pustulatus*), eight mammal species (*Odocoileus virginianus*, *Pecari tajacu*, *Canis latrans*, *Nasua narica*, *Urocyon cinereoargenteus*, *Didelphis virginiana*, *Sciurus colliciae* and *Liomys pictus*) and one reptile (*Ctenosaura pectinata*) have been documented to feed on *S. purpurea* seeds<sup>57</sup>. Collared peccaries (*P. tajacu*) predate seeds<sup>57</sup>, however the rest are all potential dispersers that are capable of long-distance seed dispersal. Only three species consumed and left seeds underneath tree crowns (i.e. *D. virginiana*, *S. colliciae* and *I. pustulatus*), while the other species moved fruits away from maternal trees<sup>57</sup>. Intact seeds were observed in feces of *C. latrans*, *N. narica*, *U. cinereoargenteus* and *C. pectinata*; regurgitated seeds were also observed on rest spots and roosting sites of *O. virginianus* and *O. poliocephala*, respectively<sup>57</sup>. Seed dispersal studies assume that most seed dispersal follows a leptokurtic distribution, in which most seeds are dispersed nearby maternal trees<sup>58</sup>. However, seed dispersers may collect seeds on one tree and move them to other conspecifics in their feeding routes<sup>59–62</sup>. For most animal dispersed plant species, ingested seeds are normally dispersed away from maternal trees<sup>62–64</sup>. This seed dispersal behaviour has two consequences on gene flow patterns associated with seed dispersal. First, it promotes gene movement over long distances. In our parentage analysis, we were able to assign maternity to 53% of offspring within the plot, 85% of these assignments occurred at distances over 150 m away from the maternal tree. We



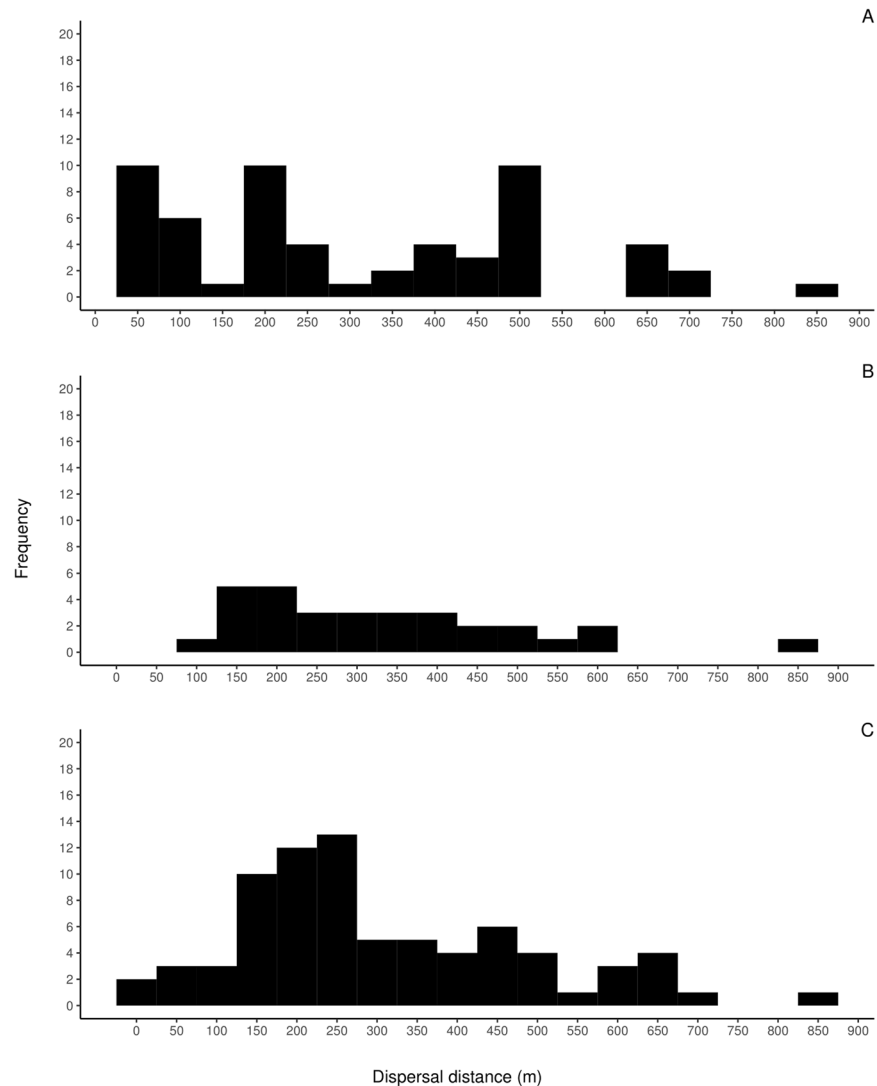
**Figure 4.** Kinship values ( $F_{ij}$ ) of size class I - size class III pairs. Each point represents calculated  $r_j$  values for each distance class. The solid lines represent upper and lower 95% confidence envelopes for the null hypothesis of no correlation ( $r_j = 0$ ).



**Figure 5.** Results of Ripley's bivariate  $L(t)$  analyses for size class I and size class III individuals in a 30 ha plot located in Chamela, Jalisco, Mexico. Solid lines represent the  $L(t)$  values. Dotted lines represent 95% confidence intervals. Spatial association between size classes is suggested if values of  $L(t) > 0$ . Conversely,  $L(t) < 0$  indicates repulsion between size classes. If  $L(t)$  values fall within the confidence envelope, the spatial relationship between size classes is deemed random.

deduced that offspring not assigned to maternal trees within the plot may be the result of seed dispersal events by mothers located outside our plot. Our results are comparable to paternity assignments obtained for other species using microsatellite markers<sup>10,65,66</sup>. Seed dispersal distance is an important result, as most studies that analyze seed dispersal rarely estimate seed dispersal over such long distances<sup>9,20,67,68</sup>. Second, dispersers may consume seeds from multiple trees before depositing them in resting places, creating a mixture of different genotypes<sup>64</sup>; both factors should reduce FSGS. Similar to seed dispersal, effective and realized pollen flow occurred more frequently at distances greater than 150 meters (70% and 93% respectively), suggesting that for *S. purpurea* insects also provide long-distance gene dispersal services, comparable to other tropical trees<sup>69</sup>.

In addition to long-dispersal gene flow, at shorter distances other factors may operate to limit FSGS and male and female fitness. Spatial association results between the smallest size class and the largest size class indicated that there is recruitment near conspecific trees, however only if the genetic relatedness between these two size classes is low (Figs. 4 and 5), suggesting a parentage recruitment effect. Furthermore, based on pre-dispersed seeds (effective pollen dispersal) paternity analysis show that 27% of pollen donors are located within 100 meters



**Figure 6.** Frequency distribution of dispersal event distances: (A) Effective pollen dispersal, (B) Realized pollen dispersal and (C) Realized seed dispersal. Pollen dispersal is based on the results of a paternity assignment on seeds ( $n = 127$ ) collected from 10 *Spondias purpurea* trees. Effective pollen and seed dispersal is based on parent-pair assignment on *Spondias purpurea* seedlings in the diametric size class I ( $n = 144$ ) within of 30 ha plot in Chamela.

of maternal trees (Fig. 6A), meanwhile parent-pair analyses of established seedlings (realized gene flow) showed that less than 7% of all seedlings had a parent within the same distance range (100 m) (Fig. 6B,C). In other words, seeds sired by distant male trees have a higher likelihood of establishing than seeds sired by nearby males; in the same way seeds dispersed away from maternal trees have a greater probability of germination and survivorship in comparison with seed dispersed at shorter distances. This indicates that the probability of surviving and establishing in the population is a function of distance to their parental trees (both male and female) congruent with the Janzen-Connell (J-C) hypothesis<sup>13,14</sup>. According to J-C, recruitment is limited in nearby neighborhoods of paternal plants due to asymmetric competition and predator and pathogen transferal from related nearby individuals. Thus, dispersal away from parental genotypes should increase the survival probability of seedlings, thereby in accordance with our paternity data, lower FSGS could be indicative of strong J-C effects. In our study site, Mendez-Toribio and colleagues<sup>70</sup> compared recruited seedling richness and density underneath female and male *S. purpurea* trees. They found that seedling density and richness was higher underneath female trees compared to male trees; furthermore density of zoochorous species was greater underneath the canopy of *S. purpurea* females suggesting a directional dispersal bias towards them, mediated by seed dispersers<sup>70</sup>. Interestingly no seedlings of *S. purpurea* were recorded directly underneath conspecific trees possibly due to strong negative density dependent effects (e.g. seed predation, interspecific competition)<sup>70-72</sup> reinforcing the idea that J-C effects strongly shape *S. purpurea* seedling recruitment.

To our knowledge only two studies have previously shown J-C effects using genetic data, in *Pinus halepensis* (Pinaceae), Steinitz and colleagues<sup>73</sup> used parentage analysis to estimate differences among seed dispersal kernels



and effective dispersal kernels demonstrating higher seedling survival with increasing distance from maternal plants. Similar results were found by Berens and colleagues<sup>74</sup> in the vertebrate seed dispersed tropical tree *Prunus africana* (Rosaceae); by comparing dispersal distances among different seedling stages, they found that dispersal distances were higher in older seedlings.

Our paternity analyses also showed evidence of other factors that affected individual plant fitness in *S. purpurea* including pollen and seed dispersal distances. Previous studies have analysed individual plant fitness associating only fecundity (seed production) with tree size indicating that in several species, the size of the plant is positively related to greater fecundity<sup>75</sup>; however, information on the success of these seeds to survive and establish in the population is scarce<sup>75</sup>. Genetic markers are a useful tool to trace the number of offspring recruited in the population sired or dispersed by a particular parental tree, data that can be used as an individual fitness measure<sup>75</sup>. Our results indicated that there is not a relationship between the number of established progeny and tree size (DBH); larger individuals could have the potential to produce more offspring due to higher fecundity<sup>75</sup>, however, other factors can also influence progeny success (e.g., timing of germination, seed predation, dispersal of offspring). In addition, the number of progeny assigned is not a linear function of the distance from its maternal tree. Our realized seed dispersal data suggests a population recruitment curve (PRC) between 150 and 250 meters away from the maternal tree (Fig. 6C). In accordance with the original proposal of Janzen<sup>13</sup> PRC is the result of the interaction between seed dispersal and seed or seedling predation<sup>13</sup>. In addition, our results show that the effective pollen flow distances are also comparable to the realized seed dispersal PRC. The proposal by Janzen<sup>13</sup> did not originally considered pollen flow in the model; interestingly, our results show that Density Dependent (DD) factors also regulate effective pollen dispersal distances (Fig. 6B). Therefore, DD factors may regulate the parental structure (both maternal and paternal) of the established progeny, also regulating the spatial distribution of genetic diversity within populations.

By comparing effective and realized pollen dispersal results, we can observe that seeds sired by males separated by at least 150 m from female trees have a greater likelihood of establishing in the population than seeds sired by nearby male trees. As shown for several high density insect pollinated tropical trees, effective pollination with nearby neighbours also occurred in *S. purpurea*<sup>76,77</sup>. However, our realized pollen flow data showed that 85% of established seedlings are sired by pollen donors that are >150 m away from maternal trees, indicating that long distance sires have higher male fitness. Similarly, our realized seed dispersal data revealed that seedlings are dispersed more frequently at distances over 150 m, indicating that long dispersal distances from maternal trees also had a strong effect on individual seedling fitness. It has been proposed that in order to coexist with hermaphrodites, dioecious species could have compensatory fitness mechanisms<sup>58,78</sup> for example in some tropical forests females of dioecious species produce more seeds than hermaphroditic species<sup>79</sup>. However, such greater seed production can result in greater seed deposition nearby maternal female trees due to limited seed dispersal; consequently, more aggregated seedlings should be expected and this aggregation could lead to a reduction in fitness via decreased seedling performance<sup>15,80,81</sup>. Under this scenario dioecious plants may allocate more resources to increase seed dispersal to enhance fitness as we show here for *S. purpurea*.

In summary, we evaluated the population sex ratio and the effects of gene dispersal through effective and realized seed and pollen dispersal on FSGS across size classes of the dioecious tree *S. purpurea*. Our results indicated that the male biased sex ratio of this species is due to precocious male reproduction; and the low levels of FSGS among size classes is likely determined by a combination of factors related with long dispersal gene flow and J-C effects during early recruitment stages. J-C density dependent factors regulate realized gene flow patterns and fitness in *S. purpurea*; both long-distance pollen donors and long-distance seed dispersers have higher male and female fitness, respectively, all of which result in low FSGS. Therefore, we propose that low levels of FSGS in tropical trees may be indicative of density dependent factors regulating seedling populations of tropical trees.

## Materials and Methods

**Study area and study species.** This study was conducted in the Chamela-Cuixmala Biosphere Reserve at Chamela Biological Station, UNAM (ChaBS) (19°30'N, 105°03'W) in the central Pacific coast of Jalisco, Mexico. This reserve protects 13,142 hectares of Tropical Dry Forest (TDF), a highly endangered ecosystem. Permission for sampling was obtained from ChaBS director.

*Spondias purpurea* L. (ANACARDIACEAE) is a small (3–10 m), deciduous and dioecious tree native to the tropical dry forests of Mexico and Central America<sup>82,83</sup>. Reproduction occurs between December and May. This tree depends on insects for cross pollination, primarily stingless bees and wasps<sup>84</sup>. Female trees produce small, bright red, juicy and sweet-acidic fruits that are eaten by various mammals and birds<sup>57</sup>. These fruits represent an important source of water and nutrients in the dry season when high temperatures, water and food scarcity are commonplace<sup>57,70</sup>.

**Experimental design and sample collection.** We set a 30 ha plot (1000 × 300 m) where all *S. purpurea* individuals were tagged and geo-referenced. Plants were grouped into size classes using diameter at breast height (DBH), except for trees less than 2 m tall, for which we used the diameter at the base of the stem. We defined three size classes: class I (DBH < 10 cm), class II (10 cm < DBH < 20 cm), class III (DBH > 20 cm). Seedlings were included in size class I. Sexual expression was monitored for all individuals in two consecutive flowering seasons (2015, 2016). We collected flowers from different sections of the canopy to accurately determine the gender of all individuals.

To estimate pollen dispersal distances we collected twenty fruits from ten randomly chosen female trees within the plot. Fruits were collected directly from the canopy of each individual to guarantee that we sampled progeny from a single maternal tree.



**Sex ratios.** Sex ratio was expressed as the proportion of males in the population: males/ (females + males)<sup>85</sup>. For each sex we determined the frequency of individuals in each diametric size class. To determine deviations from 1:1 sex ratio we performed a goodness of fit G-test for each size class and for pooled size classes. To evaluate differences in diameter size distribution between sexes, we used a Kolmogorov-Smirnov two sample test. Statistical analyses were performed using the *RVAideMemoire*<sup>86</sup> and *stats* libraries implemented in the R computing environment<sup>87</sup>.

**DNA extraction and microsatellite amplification.** We collected fresh leaf tissue from all individuals within the 30 ha plot and stored them at  $-20^{\circ}\text{C}$  until DNA extraction. For pollen dispersal analyses we dissected fruits and sampled embryos for DNA extraction. DNA from leaves and embryos was extracted using a modified Cetyltrimethylammonium Bromide (CTAB) protocol<sup>88</sup>. Nine microsatellites previously developed for *S. purpurea*<sup>89</sup> were amplified via multiplex PCR using QIAGEN Multiplex kit (QIAGEN, Hilden, Germany) in 12  $\mu\text{L}$  reaction volumes. The first multiplex mix used primers SPUR44, SPUR40, SPUR28 and SPUR41, the second mix included the primers SPUR35, SPUR29 and SPUR33 and the third set contained the primers SPUR42 and SPUR39; in all cases primers were at a 0.2  $\mu\text{M}$  concentration. The PCR amplification profile included an initial activation step of  $95^{\circ}\text{C}$  for 15 min, followed by a touchdown PCR consisting of 32 cycles with denaturation at  $95^{\circ}\text{C}$  for 90 s; annealing for 60 s with temperature decreasing  $1^{\circ}\text{C}$  every two cycles from  $64^{\circ}\text{C}$  (12 cycles), then 10 cycles at  $58^{\circ}\text{C}$  and 10 cycles at  $57^{\circ}\text{C}$ ; elongation at  $72^{\circ}\text{C}$  for 60 s; and a final extension at  $72^{\circ}\text{C}$  for 5 min. Fragments were analysed on an automatic ABI-PRISM 3100-Avant sequencer (APPLIED BIOSYSTEMS, Carlsbad, California, USA) using GeneScan LIZ 600 (APPLIED BIOSYSTEMS, Carlsbad, California, USA) to determine fragment sizes. Alleles were scored manually using *GeneMarker Software* version 2.6.4 (SOFTGENETICS). To reduce genotyping error, all samples were independently genotyped by two different researchers to reach consensus in the final data set. The primer SPUR57 was excluded from the analyses because in several individuals it produced multiple peaks that could not be genotyped properly. Erroneous genotypes may critically bias paternity assignments and eliminating SPUR57 did not significantly reduce the exclusion probability of the remaining 9 markers.

**Genetic diversity analysis.** Genetic diversity was quantified for each size class using the following parameters: allele number averaged across loci ( $N_a$ ), Allelic richness ( $A_r$ ), observed ( $H_o$ ) and expected heterozygosities ( $H_e$ ) and fixation indexes ( $F$ ). These parameters were calculated using the software *ARLEQUIN*<sup>90</sup>. Confidence intervals for  $F$  values were obtained by bootstrapping over loci using the library *hierfstat*<sup>91</sup> implemented in R<sup>87</sup>.

**Fine scale genetic structure and spatial analysis.** Fine-scale genetic structure was estimated for all size classes using the software *SPAGeDi*<sup>92</sup>. We estimated the mean and the confidence intervals of pairwise kinship coefficients between all pairs of individuals within each size class. For each analysis, we selected eight distance categories; the first five categories are separated by 20 m intervals up to 100 m, with the remaining categories separated by 25 m intervals up to 200 m. All categories included more than 100 pairwise comparisons. To test the null hypothesis of a random distribution of genotypes in space, estimated  $F_{ij}$  values were compared to the 95% confidence intervals generated by 10,000 random permutations of individuals in space.  $F_{ij}$  values were regressed onto  $\ln(d_{ij})$ , where  $d_{ij}$  is the spatial distance between all pairs of  $ij$ -individuals. To quantify the FSGS intensity and to compare it with other studies the  $S_p$  statistic was calculated according to Vekemans and Hardy<sup>3</sup>. The significance of the slope was assessed by comparing the estimated value with 10 000 random shifts of individuals among locations. Standard errors (SE) for all slopes were determined by jackknifing across loci. In order to analyze parentage and conspecific effects on seedling recruitment, we tested the spatial interaction (i.e., association or repulsion) and inter-size class genetic structure between the smallest size class (size class I) and the largest size class (size class III). To evaluate relatedness patterns among plants between size class I and III we performed a “between-size class” FSGS analysis; in which relatedness is calculated for pairs of individuals including an individual from size class I and an individual from size class III, yielding the degree of relatedness of each “seedling” with their closest adults. The distances classes analysed and parameters used for to test the null hypothesis of a random distribution of genotypes in space were the same used in the FSGS within size classes. The spatial interaction (association or repulsion) between individuals of size class I and III was evaluated using Ripley’s second order  $K_{12}(t)$  function.  $K_{12}(t)$  estimates the expected number of type 1 (size class I individuals in this study) within a distance ( $t$ ) of a randomly chosen type 2 point (size class III). To reduce scale dependencies and stabilize variances, we used the  $L(t)$  square root transformation of  $K(t)$ <sup>21,93</sup>. Spatial interaction was evaluated for  $t$  distances between 0 and 150 m and an isotropic edge correction was also applied. We chose ‘random labeling’ as the null hypothesis to construct confidence envelopes<sup>94</sup>. Confidence envelopes were constructed from 999 permutations. Spatial association between size class I and size class III individuals is suggested if values of  $L(t) > 0$ . Conversely,  $L(t) < 0$  indicates repulsion between size class I and size class III individuals. If  $L(t)$  values fall within the confidence envelope, the spatial relationship is deemed random. Spatial analysis was conducted using the *spatstat* library<sup>95</sup> implemented in the R computing environment<sup>87</sup>.

**Parentage analysis.** To determine gene flow levels via pollen and seeds we performed two separate parentage analysis. We first evaluated effective pollen dispersal via paternity analysis on seed embryos as offspring genotypes. Males that reproduced in the same year when seeds were collected were considered potential fathers in the analysis. The effective pollen dispersal distance was calculated as the distance between the father assigned to the embryo and the maternal tree. In a second analysis we evaluated realized pollen and seed dispersal distances. To this end we used the genotypes of individuals in the diametric class I as offspring, while reproductive individuals (i.e. identified as male or females by flower observation) in diametric classes II and III were used as potential parents. These analyses resulted in two dispersal distances: (i) the realized pollen dispersal distance estimated as the distance between the mother and father assigned to an established seedling when both parents

occur within the plot<sup>20</sup>; (ii) and the effective seed dispersal distance calculated as the distance between seedlings and their assigned mother. Parentage analyses were conducted using Colony V2<sup>96</sup>. We used the parameter of a polygamous outcrossing dioecious species, and selected the full likelihood method with a large run. We assigned paternity using “strict” (>80%) confidence levels. We included a genotyping error probability of 0.0923 based on the number of mismatches observed between the genotypes of maternal trees and their offspring (calculated from the data set of embryos with known maternal genotypes), estimated using Cervus 3.0<sup>97</sup>.

## Data availability

The datasets generated and analysed during the current study are available from the corresponding author upon request.

Received: 26 October 2019; Accepted: 3 February 2020;

Published online: 12 March 2020

## References

- Ghazoul, J. Pollen and seed dispersal among dispersed plants. *Biol. Rev.* **80**, 413 (2005).
- Jordano, P. What is long-distance dispersal? And a taxonomy of dispersal events. *J. Ecol.* **105**, 75–84 (2017).
- Vekemans, X. & Hardy, O. J. New insights from fine-scale spatial genetic structure analyses in plant populations. *Mol. Ecol.* **13**, 921–935 (2004).
- Wright, S. Isolation by distance. *Genetics* **28**, 114–138 (1943).
- Bohonak, A. J. Dispersal, Gene Flow, and Population Structure. *Q. Rev. Biol.* **74**, 21–45 (1999).
- Kalisz, S., Nason, J. D., Hanzawa, F. M. & Tonsor, S. J. Spatial population genetic structure in *Trillium grandiflorum*: the roles of dispersal, mating, history and selection. *Evolution* **55**, 1560–1568 (2001).
- Pardini, E. A. & Hamrick, J. L. Inferring recruitment history from spatial genetic structure within populations of the colonizing tree *Albizia julibrissin* (Fabaceae). *Mol. Ecol.* **17**, 2865–2879 (2008).
- Jacquemyn, H., Brys, R., Vandepitte, K., Honnay, O. & Roldán-Ruiz, I. Fine-scale genetic structure of life history stages in the food-deceptive orchid *Orchis purpurea*. *Mol. Ecol.* **15**, 2801–2808 (2006).
- Jones, F. A. & Hubbell, S. P. Demographic spatial genetic structure of the Neotropical tree, *Jacaranda copaia*. *Mol. Ecol.* **15**, 3205–3217 (2006).
- Choo, J., Juenger, T. E. & Simpson, B. B. Consequences of frugivore-mediated seed dispersal for the spatial and genetic structures of a neotropical palm. *Mol. Ecol.* **21**, 1019–1031 (2012).
- Nazareno, A. G., Alzate-Marin, A. L. & Pereira, R. A. S. Dioecy, more than monoecy, affects plant spatial genetic structure: the case study of *Ficus*. *Ecol. Evol.* **3**(10), 3495–3508 (2013).
- Gelmi-Candusso, T. A., Heymann, E. W. & Heer, K. Effects of zoochory on the spatial genetic structure of plant populations. *Mol. Ecol.* **26**, 5896–5910 (2017).
- Janzen, D. H. Herbivores and the Number of Tree Species in Tropical Forests. *Am. Nat.* **104**, 501–528 (1970).
- Connell, J. H. On the role of natural enemies in preventing competitive exclusion in some marine animals and in rain forest trees in Dy namics of Populations. (eds. Den Boer, P. J. & Gradwell, G. R.) 298–312 (Centre for Agricultural Publishing and Documentation, 1971).
- Comita, L. S. *et al.* Testing predictions of the Janzen–Connell hypothesis: a meta-analysis of experimental evidence for distance- and density-dependent seed and seedling survival. *J. Ecol.* **102**, 845–856 (2014).
- Bawa, K. S., Opler, P. A., Rica, D. C. & Rica, C. Dioecism in Tropical Forest Trees. *Evolution* **29**, 167–179 (1975).
- Bawa, K. S., Perry, D. R. & Beach, J. H. Reproductive Biology of Tropical Lowland Rain Forest Trees. I. Sexual Systems and Incompatibility Mechanisms. *Am. J. Bot.* **72**, 331 (1985).
- Ibarra-Manriquez, G. & Oyama, K. Ecological Correlates of Reproductive Traits of Mexican Rain Forest Trees. *Am. J. Bot.* **79**, 383 (1992).
- Renner, S. S. & Ricklefs, R. E. Dioecy and its correlates in the flowering plants. *Am. J. Bot.* **82**, 596–606 (1995).
- Zhou, H. P. & Chen, J. Spatial genetic structure in an understory dioecious fig species: the roles of seed rain, seed and pollen-mediated gene flow, and local selection. *J. Ecol.* **98**, 1168–1177 (2010).
- Riba-Hernández, P., Segura, J. L. & Muñoz-Valverde, J. Female fruit production depends on female flower production and crown size rather than male density in a continuous population of a tropical dioecious tree (*Virolo surinamensis*). *Am. J. Bot.* **103**, 1990–1999 (2016).
- Degen, B. & Sebbenn, A. M. Genetics and Tropical Forests in *Tropical Forestry Handbook* (eds. Köhl, M. & Pancel, L) (Springer, 2014).
- Murawski, D. & Hamrick, J. The effect of the density of flowering individuals on the mating systems of nine tropical tree species. *Heredity* **67**, 167–174 (1991).
- Bianchi, F. J. J. A. & Cunningham, S. A. Unravelling the role of mate density and sex ratio in competition for pollen. *Oikos* **121**, 219–227 (2012).
- Castilla, A. R. *et al.* Adding landscape genetics and individual traits to the ecosystem function paradigm reveals the importance of species functional breadth. *Proc Natl Acad Sci USA* **114**, 12761–12766 (2017).
- Grant, E. L. *et al.* Short distance pollen dispersal and low genetic diversity in a subcanopy tropical rainforest tree, *Fontainea picrosperma* (Euphorbiaceae). *Heredity* **123**, 503–516 (2019).
- Barrett, S. C. H., Yakimowski, S. B., Field, D. L. & Pickup, M. Ecological genetics of sex ratios in plant populations. *Philos. T. R. Soc. B* **365**, 2549–57 (2010).
- Lloyd, D. G. & Webb, C. J. Secondary sex characters in plants. *Bot. Rev.* **43**, 177–216 (1977).
- Obeso, J. R., Alvarez-Santullano, M. & Retuerto, R. Sex ratios, size distributions, and sexual dimorphism in the dioecious tree *Ilex aquifolium* (Aquifoliaceae). *Am. J. Bot.* **85**, 1602–1608 (1998).
- Wilson, A. J. & Nussey, D. H. What is individual quality? An evolutionary perspective. *Trends Ecol. Evol.* **25**(4), 207–214 (2010).
- Meagher, T. R. & Thompson, E. Analysis of Parentage for Naturally Established Seedlings of *Chamaelirium luteum* (Liliaceae). *Ecology* **68**, 803–812 (1987).
- Primack, R. B. & Kang, H. Measuring fitness and natural selection in wild plant populations. *Annu. Rev. Ecol. Syst.* **20**, 367–396 (1987).
- Kingsolver, J. G., Diamond, S. E., Siepielski, A. M. & Stephanie, M. C. Synthetic analyses of phenotypic selection in natural populations: lessons, limitations and future directions. *Evol. Ecol.* **26**, 1101 (2012).
- Campbell, D. R., Brody, A. K., Price, M. V., Waser, N. M. & Aldridge, G. Is Plant Fitness Proportional to Seed Set? An Experiment and a Spatial Model. *Am. Nat.* **190**(6), 818–827 (2017).
- Sant’Anna, C. *et al.* Realized pollen and seed dispersal within a continuous population of the dioecious coniferous Brazilian pine [*Araucaria angustifolia* (Bertol.) Kuntze]. *Conserv. Genet.* **14**(3), 601–613 (2013).
- Opler, P. A. & Bawa, K. S. Sex ratios in tropical forest trees. *Evolution* **32**, 812–821 (1978).

37. Bullock, S. H. & Bawa, K. S. Sexual Dimorphism and the Annual Flowering Pattern in *Jacaratia Dolichaula* (D. Smith) Woodson (Caricaceae) in a Costa Rican Rain Forest. *Ecology* **62**, 1494–1504 (1981).
38. Allen, G. A. & Antos, J. A. Sex Ratio Variation in the Dioecious Shrub *Oemeleria cerasiformis*. *Am. Nat.* **141**, 537–553 (1993).
39. Nicotra, A. B. Sex ratio variation and spatial distribution of *Siparuna grandiflora*, a tropical dioecious shrub. *Oecologia* **115**, 102–113 (1998).
40. Thomas, S. C., LaFrankie, J. V., Thomas, A. S. C. & LaFrankie, J. V. Sex, size and interyear variation in flowering among dioecious trees of the malayan rain forest. *Ecology* **74**, 1529–1537 (1993).
41. García, M. & Antor, R. Sex Ratio and Sexual Dimorphism in the Dioecious *Borderea pyrenaica* (Dioscoreaceae). *Oecologia* **101**(1), 59–67 (1995).
42. Hamrick, J. L., Godt, M. J. W. & Sherman-Broyles, S. L. Factors influencing levels of genetic diversity in woody plant species. *New Forest* **6**, 95–124 (1992).
43. Hamrick, J. L. Response of forest trees to global environmental changes. *Forest Ecol. Manag.* **197**, 323–335 (2004).
44. Dick, C. W., Hardy, O. J., Jones, F. A. & Petit, R. J. Spatial Scales of Pollen and Seed-Mediated Gene Flow in Tropical Rain Forest Trees. *Trop. Plant Biol.* **1**, 20–33 (2008).
45. Hardesty, B. D., Dick, C. W., Kremer, A., Hubbell, S. & Bermingham, E. Spatial genetic structure of *Simarouba amara* Aubl. (Simaroubaceae), a dioecious, animal-dispersed Neotropical tree, on Barro Colorado Island, Panama. *Heredity* **95**, 290–297 (2005).
46. Dev, S. A., Kjellberg, F., Hossaert-McKey, M. & Borges, R. M. Fine-scale Population Genetic Structure of Two Dioecious Indian Keystone Species, *Ficus hispida* and *Ficus exasperata* (Moraceae). *Biotropica* **43**, 309–316 (2011).
47. Loiselle, B. A., Sork, V. L., Nason, J. & Graham C. Spatial genetic structure of a tropical understory shrub, *Psychotria officinalis* (Rubiaceae). *Am. J. Bot.* **82**, 1420–1425 (1995).
48. Chung, M. Y., Epperson, B. K. & Chung, M. G. Genetic structure of age classes in *Camellia japonica* (Theaceae). *Evolution* **57**, 62–73 (2003).
49. Epperson, B. K. & Alvarez-Buylla, E. R. Limited seed dispersal and genetic structure in life stages of *cecropia obtusifolia*. *Evolution* **51**, 275–282 (1997).
50. Kitamura, K., Nakanishi, A., Lian, C. & Goto, S. Distinctions in Fine-Scale Spatial Genetic Structure Between Growth Stages of *Picea jezoensis* Carr. *Front. Genet.* **9**, 490 (2018).
51. Schroeder, J. W., Trana, H. T. & Dick, C. W. Fine scale spatial genetic structure in *Pouteria reticulata* (Engl.) Eyma (Sapotaceae), a dioecious, vertebrate dispersed tropical rain forest tree species. *Glob. Ecol. Conserv.* **1**, 43–49 (2014).
52. Vieira, F. A., Fajardo, C. G., Souza, A. M., Reis, C. A. F. & Carvalho, D. Fine-scale genetic dynamics of a dominant neotropical tree in the threatened Brazilian Atlantic Rainforest. *Tree Genet. Genomes* **8**, 1191–1201 (2012).
53. Chung, M. G. & Epperson, B. K. Clonal and spatial genetic structure in *Eurya emarginata* (Theaceae). *Heredity* **84**, 170–177 (2000).
54. Trapnell, D. W., Schmidt, J. P. & Hamrick, J. L. Spatial genetic structure of the Southeastern North American endemic, *Ceratiola ericoides* (Empetraceae). *J. Hered.* **99**, 604–609 (2008).
55. Wang, R. *et al.* Spatial genetic structure and restricted gene flow in a functionally dioecious fig *Ficus pumila* L. *pumila* (Moraceae). *Popul. Ecol.* **51**, 307–315 (2009).
56. Gaino, A. P. S. C. *et al.* Understanding the effects of isolation on seed and pollen flow, spatial genetic structure and effective population size of the dioecious tropical tree *Myracrodruon urundeuva*. *Conserv. Genet.* **11**, 1631–1643 (2010).
57. Mandujano, S., Gallina, S. & Bullock, S. H. Frugivory and dispersal of *Spondias purpurea* (Anacardiaceae) in a tropical deciduous forest in México. *Rev. Biol. Trop.* **42**, 107–114 (1994).
58. Heilbuth, J. C., Ilves, K. L. & Otto, S. P. The consequences of dioecy for seed dispersal: modeling the seed-shadow handicap. *Evolution* **55**, 880–888 (2001).
59. Hardesty, B. D., Hubbell, S. P. & Bermingham, E. Genetic evidence of frequent long-distance recruitment in a vertebrate-dispersed tree. *Ecol. Lett.* **9**, 516–525 (2006).
60. Wang, B. C., Sork, V. L., Leong, M. T. & Smith, T. B. Hunting of Mammals Reduces Seed Removal and Dispersal of the Afrotropical Tree *Antrocaryon klaineum* (Anacardiaceae). *Biotropica* **39**(3), 340–347 (2007).
61. Sezen, U. U., Chazdon, R. L. & Holsinger, K. E. Proximity is not a proxy for parentage in an animal-dispersed Neotropical canopy palm. *Proc. R. Soc. B.* **276**, 2037–2044 (2009).
62. Solís-Hernández, W. & Fuchs, E. J. Effective gene flow patterns across a fragmented landscape in southern Costa Rica for *Symphonia globulifera* (Clusiaceae): a species with mobile seed and pollen dispersers. *Rev. Biol. Trop.* **67**(2), 95–111 (2019).
63. Howe, H. F. Bird Activity and Seed Dispersal of a Tropical Wet Forest Tree. *Ecology* **58**, 539–550 (1977).
64. Fuchs, E. J. & Hamrick, J. L. Spatial genetic structure within size classes of the endangered tropical tree *Guaiacum sanctum* (Zygophyllaceae). *Am. J. Bot.* **97**, 1200–1207 (2010).
65. Dow, B. D. & Ashley, M. V. Microsatellite analysis of seed dispersal and parentage of saplings in bur oak, *Quercus macrocarpa*. *Mol. Ecol.* **5**, 615–627 (1996).
66. Streiff, R. *et al.* Pollen dispersal inferred from paternity analysis in a mixed oak stand of *Quercus robur* L. and *Q. petraea* (Matt.) Liebl. *Mol. Ecol.* **8**, 831–841 (1999).
67. Nathan, R. & Muller-Landau, H. C. Spatial patterns of seed dispersal, their determinants and consequences for recruitment. *Trends Ecol. Evol.* **15**(7), 278–285 (2000).
68. Godoy, J. A. & Jordano, P. Seed dispersal by animals: exact identification of source trees with endocarp DNA microsatellites. *Mol. Ecol.* **10**, 2275–2283 (2001).
69. Dick, C. W., Etchelecu, G. & Austerlitz, F. Pollen dispersal of tropical trees (*Dinizia excelsa*: Fabaceae) by native insects and African honeybees in pristine and fragmented Amazonian rainforest. *Mol. Ecol.* **12**, 753–764 (2003).
70. Méndez-Toribio, M., Pierrro, A. M. G.-D., Quesada, M. & Benítez-Malvido, J. Regeneration beneath a dioecious tree species in a Mexican tropical dry forest. *J. Trop. Ecol.* **30**, 265–268 (2014).
71. Sugiyama, A., Comita, L., Masaki, T., Condit, R. & Hubbell, S. Resolving the paradox of clumped seed dispersal: positive density and distance dependence in a bat-dispersed species. *Ecology* **99**(1), 2583–2591 (2019).
72. Zhu, Y. *et al.* Density-dependent survival varies with species life-history strategy in a tropical forest. *Ecol. Lett.* **21**(4), 506–515 (2018).
73. Steinitz, O., Troupin, D., Vendramin, G. G. & Nathan, R. Genetic evidence for a Janzen–Connell recruitment pattern in reproductive offspring of *Pinus halepensis* trees. *Mol. Ecol.* **20**, 4152–4164 (2011).
74. Berens, D. G. *et al.* Changes of effective gene dispersal distances by pollen and seeds across successive life stages in a tropical tree. *Oikos* **122**, 1616–1625 (2013).
75. Younginger, B. S., Sirová, D., Cruzan, M. B. & Ballhorn, D. J. Is biomass a reliable estimate of plant fitness? *Appl. Plant Sci.* **5**, 1600094, <https://doi.org/10.3732/apps.1600094> (2017).
76. Stacy, E. A. *et al.* Pollen Dispersal in Low-Density Populations of Three Neotropical Tree Species. *Am. Nat.* **148**(2), 275–298 (1996).
77. Ward, M., Dick, C. W., Gribel, R. & Lowe, A. J. To self, or not to self... A review of outcrossing and pollen-mediated gene flow in neotropical trees. *Heredity* **95**, 246–254 (2005).
78. Bawa, K. S. Evolution of dioecy in flowering plants. *Ann. Rev. Ecol. Syst.* **11**, 15–39 (1980).
79. Bruijning, M. *et al.* Surviving in a Cosexual World: A Cost-Benefit Analysis of Dioecy in Tropical Trees. *Am. Nat.* **189**(3), 297–314 (2017).

80. Harms, K. E., Wright, J. S., Calderón, O., Hernández, A. & Herre, E. A. Pervasive density-dependent recruitment enhances seedling diversity in a tropical forest. *Nature* **404**, 493–495 (2000).
81. Lebrija-Trejos, E., Reich, P. B., Hernández, A. & Wright, S. J. Species with greater seed mass are more tolerant of conspecific neighbours: a key driver of early survival and future abundances in a tropical forest. *Ecol. Lett.* **19**, 1071–1080 (2016).
82. Miller, A. J. & Knouft, J. H. GIS-based characterization of the geographic distributions of wild and cultivated populations of the Mesoamerican fruit tree *Spondias purpurea* (Anacardiaceae). *Am. J. Bot.* **93**, 1757–1767 (2006).
83. Miller, A. & Schaal, B. Domestication of a Mesoamerican cultivated fruit tree, *Spondias purpurea*. *Proc. Natl. Acad. Sci. USA* **102**, 12801–12806 (2005).
84. Mitchell, J. D. & Daly, D. C. A revision of *Spondias* L. (Anacardiaceae) in the Neotropics. *PhytoKeys* **55**, 1–92 (2015).
85. Wilson, K. & Hardy, I. C. W. Statistical analysis of sex ratios: an introduction. (ed. Hardy, I. C. W.) *Sex Ratios*. 48–92 (Cambridge University Press, 2002).
86. Hervé, M. RVAideMemoire: testing and plotting procedures for biostatistics. R package version 0.9-69-3, <https://CRAN.R-project.org/package=RVAideMemoire> Computer software (2018).
87. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria (2018).
88. Doyle, J. J. & Doyle, J. L. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem. Bull.* **19**, 11–15 (1987).
89. Cristóbal-Pérez, E. J., Fuchs, E. J., Harvey, N. & Quesada, M. Isolation and characterization of microsatellites loci in *Spondias purpurea* (Anacardiaceae) and cross amplification in congeneric species. *Mol. Biol. Rep.* **46**(5), 5581–5585 (2019).
90. Excoffier, L. & Lischer, H. E. L. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* **10**, 564–567 (2010).
91. Goudet, J. H. a package for *r* to compute and test hierarchical F-statistics. *Mol. Ecol. Res.* **5**, 184–186 (2005).
92. Hardy, O. J. & Vekemans, X. Spagedi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Mol. Ecol. Res.* **2**, 618–620 (2002).
93. Besag, J. Contribution to the discussion of Dr. Ripley's paper. *J. Roy. Stat. Soc. B* **39**, 193–195 (1977).
94. Goreaud, F. & Pélissier, R. Avoiding misinterpretation of biotic interactions with the intertype K12-function: population independence vs. random labeling hypotheses. *J. Veg. Sci.* **14**, 681–692 (2003).
95. Baddeley, A. & Turner, R. Spatstat: an R package for analyzing spatial point patterns. *J. Stat. Softw.* **12**, 1–42 (2005).
96. Jones, O. R. & Wang, J. COLONY: a program for parentage and sibship inference from multilocus genotype data. *Mol. Ecol. Res.* **10**, 551–555 (2010).
97. Kalinowski, S. T., Taper, M. L. & Marshall, T. C. Revising how the computer program Cervus accommodates genotyping error increases success in paternity assignment. *Mol. Ecol.* **16**, 1099–1106 (2007).

## Acknowledgements

This study was performed in partial fulfillment of the requirements of the Ph. D. degree of E. Jacob Cristóbal-Pérez at the Graduate Program in Biological Sciences, UNAM. We thank G. Sanchez-Montoya, M. Aguilar-Aguilar, R. García, S. Novais, K. Sanchez-Gomez and Z. Carrillo for help in the Lab and field. This work was supported by grants from Universidad Nacional Autónoma de México (PAPIIT # IA207618, IV200418, IA105920), SAGARPA-CONACYT 291333, CONACYT-UNAM-UAGro to LANASE (2015-LN250996, 2016-LN271449, 2017-LN280505, 2018-LN293701, 2019-LN299033), Programa Iberoamericana de Ciencia y Tecnología para el Desarrollo RED CYTED-SEPODI (417RT0527), and graduate fellowship (CONACYT) to E.J.C.P. We thank the Chamela-Cuixmala Biosphere Reserve, Chamela Biological Station UNAM for providing permission for sampling, physical structure and logistical support for the study.

## Author contributions

E.J.C.P., E.J.F. and M.Q. designed the study. E.J.C.P. and M.Q. collected the samples. E.J.C.P. performed lab work. E.J.C.P., U.O.P. and E.J.F. performed statistical and genetic analyzes. E.J.C.P. drafted the manuscript. All authors completed and approved the final manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to M.Q.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020