ORIGINAL ARTICLE

Early genetic consequences of defaunation in a large-seeded vertebrate-dispersed palm (*Syagrus romanzoffiana*)

MI Giombini[1,2], SP Bravo[1], YV Sica[1] and DS Tosto[2]

Plant populations are seriously threatened by anthropogenic habitat disturbance. In particular, defaunation may disrupt plant-disperser mutualisms, thus reducing levels of seed-mediated gene flow and genetic variation in animal-dispersed plants. This may ultimately limit their adaptive potential and ability to cope with environmental change. Tropical forest remnants are typically deprived of medium to large vertebrates upon which many large-seeded plants rely for accomplishing effective seed dispersal. Our main goal was to examine the potential early genetic consequences of the loss of large vertebrates for large-seeded vertebrate-dispersed plants. We compared the genetic variation in early-stage individuals of the large-seeded palm *Syagrus romanzoffiana* between continuous protected forest and nearby partially defaunated fragments in the Atlantic Forest of South America. Using nine microsatellites, we found lower allelic richness and stronger fine-scale spatial genetic structure in the disturbed area. In addition, the percentage of dispersed recruits around conspecific adults was lower, although not significantly, in the disturbed area (median values: 0.0 vs 14.4%). On the other hand, no evidence of increased inbreeding or reduced pollen-mediated gene flow (selfing rate and diversity of pollen donors) was found in the disturbed area. Our findings are strongly suggestive of some early genetic consequences resulting from the limitation in contemporary gene flow via seeds, but not pollen, in defaunated areas. Plant-disperser mutualisms involving medium–large frugivores, which are seriously threatened in tropical systems, should therefore be protected to warrant the maintenance of seed-mediated gene flow and genetic diversity in large-seeded plants.

Heredity advance online publication, 25 January 2017; doi:10.1038/hdy.2016.130

INTRODUCTION

Anthropogenic disturbance has severely affected most terrestrial ecosystems. Habitat loss and fragmentation has now become a common scenario in tropical and subtropical regions (Laurance, 2010), where plant populations are expected to suffer several genetic consequences (Ellstrand and Elam, 1993; Young *et al.*, 1996; Lowe *et al.*, 2005; Aguilar *et al.*, 2008). Furthermore, human-induced changes in animal abundance, composition and movement patterns may alter pollen- and seed-dispersal mutualisms (Dirzo *et al.*, 2014; Neuschulz *et al.*, 2016) that are critical to maintaining plant gene flow and genetic diversity (Sork and Smouse, 2006). Therefore, plants are not only threatened by fragmentation and demographic reduction, but also by the loss of dispersal and gene flow capacities that may ultimately limit their adaptive potential and ability to cope with the fast-paced environmental change (Jump and Peñuelas, 2005).

A population size reduction is expected to be followed by the loss of rare alleles and, subsequently, by a progressive loss of genetic variation through increased genetic drift (Loveless and Hamrick, 1984; Young *et al.*, 1996). Mating among relatives and self-fertilization may also become more frequent in smaller populations that consequently promotes inbreeding and inbreeding depression (Young *et al.*, 1996; Lowe *et al.*, 2005). Limited gene flow should increase the spatial genetic structure ('SGS') within continuous populations while promoting genetic differentiation among separate (sub)populations (Loveless and Hamrick, 1984; Lowe *et al.*, 2005).

Evidence for most of these predictions, however, is controversial as indicated by dissimilar, and often contrasting, results (Lowe *et al.*, 2005; Sork and Smouse, 2006; Kramer *et al.*, 2008). In part, such discrepancies may be because disturbance is a very recent phenomenon with respect to the longevity of trees, so that the surveyed adult individuals represented genetic variation in predisturbance times (Lowe *et al.*, 2005; Aguilar *et al.*, 2008). In addition, contrary to initial expectations, some fragmented populations may not suffer a decline in gene flow levels (Kramer *et al.*, 2008; see below). Thus, the detection of the genetic impact will depend on the age of sampled individuals and the time elapsed since the habitat disturbance began, as well as the sensitivity of genetic variation measures to show changes after disturbance. In general, genetic effects have been more easily detected in early-stage (progenies) versus adult individuals, and by changes in allelic richness and SGS versus gene diversity and inbreeding coefficients (Lowe *et al.*, 2005; Aguilar *et al.*, 2008; Wang *et al.*, 2011; Browne *et al.*, 2015; Carvalho *et al.*, 2015).

Elucidating the underlying mechanisms of those genetic consequences poses an additional challenge because multiple related factors may operate simultaneously. For instance, both forest fragmentation and logging, which commonly act in concert in anthropogenically disturbed habitats, may reduce the effective sizes and the genetic

[1]IEGEBA, Instituto de Ecología, Genética y Evolución de Buenos Aires, UBA-CONICET, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Ciudad Autónoma de Buenos Aires, Argentina and [2]Instituto de Biotecnología, Instituto Nacional de Tecnología Agropecuaria INTA-Castelar, Hurlingham, Buenos Aires, Argentina
Correspondence: Dr MI Giombini, Instituto de Biología Subtropical–nodo Iguazú (CONICET–UNaM), Moisés Bertoni 85, Puerto Iguazú, Misiones 3370, Argentina.
E-mail: mgiombi@yahoo.com

2

connectivity within and among plant populations (Ellstrand and Elam, 1993; Young et al., 1996; Lowe et al., 2005). The former may determine landscape barriers to gene flow among forest remnants and the latter could increase the degree of isolation among reproductive trees. The genetic effects may be difficult to disentangle because both processes are expected to erode genetic diversity and increase inbreeding and spatial genetic structure (Ellstrand and Elam, 1993; Young et al., 1996; Lowe et al., 2005).

Disturbed areas may also be subject to the restriction of contemporary gene flow levels through the loss, or movement limitation, of pollinators and seed dispersers (Sork and Smouse, 2006). For instance, a higher degree of spatial isolation of reproductive individuals and/or restriction in pollinator displacements might limit the ability of plants to exchange pollen. This may reduce the outcrossing rates and the number of pollen donors that fertilize each maternal tree, potentially increasing population inbreeding levels (Lowe et al., 2005; Sork and Smouse, 2006). In particular, several studies have shown that selfing rates tend to increase when adult tree density decreases, as is the case in disturbed areas (Ward et al., 2005). Despite detrimental effects, derived from changes in pollination and mating patterns, have been occasionally documented, there is also evidence that the amount of pollen-mediated gene flow is not necessarily reduced (reviewed in Lowe et al., 2005; Ward et al., 2005; Sork and Smouse, 2006; Aguilar et al., 2008; Kramer et al., 2008). Interestingly, the movement distances of pollinating insects (either native or exotic) are often extensive, and even higher, in fragmented populations, allowing the maintenance of genetic diversity and connectivity (see, for example, White et al., 2002; Dick et al., 2003; Hanson et al., 2008).

Seed dispersal and plant regeneration may be substantially affected by defaunation (Neuschulz et al., 2016). Recruitment patterns may be altered by the extirpation of both seed dispersers and seed predators (Kurten, 2013). The outcome will depend on which of these assemblages are more severely impacted, and the interplay with density-dependent processes that influence mortality rates, such as Janzen–Connell effects (Nathan and Muller-Landau, 2000; Kurten, 2013; Galetti et al., 2015). For instance, given that seed predation mostly occurs near maternal plants, a relative increase of recruitment around adult trees may result from the loss of dispersers and/or seed predators. Instead, overall recruitment levels would be reduced or enhanced, respectively.

So far, little is known about the impact of defaunation on seed-mediated gene flow and plant genetic structure. Frugivore-mediated dispersal shapes the primary distribution of maternal progenies and genotypes across the landscape (García and Grivet; 2011). Each dispersal agent may contribute differentially to the emerging spatial patterns of genetic variation (Hamrick et al., 1993; García and Grivet, 2011). Some of them may perform clumped dispersal of seeds containing highly structured gene pools (see, for example, Bialozyt et al., 2014; Karubian et al., 2015), whereas others promote an admixture of genotypes and a decrease of fine-scale SGS (see, for example, Karubian et al., 2010; Giombini et al., 2016). In general, the loss of dispersers should increase progeny aggregation around maternal plants because of lower seed removal (Hamrick et al., 1993; Kurten, 2013). In addition, an overall restriction of gene dispersal via seeds should increase the strength of SGS (Loveless and Hamrick, 1984; Hamrick et al., 1993; Gehring and Delph, 1999).

Tropical forest remnants are typically deprived of medium to large vertebrates (Peres, 2001; Stoner et al., 2007) upon which many large-seeded plants rely for accomplishing effective and long-distance seed dispersal (Kurten, 2013). Our main goal was to examine the potential early genetic consequences of the loss of large vertebrates for large-seeded vertebrate-dispersed plants. To this end, we compared the genetic variation in early-stage individuals of the large-seeded palm Syagrus romanzoffiana between continuous protected forest and nearby partially defaunated fragments in the Atlantic Forest of South America. This palm is mainly dispersed by the Amazonian tapir (Tapirus terrestris), which is the largest frugivore in the region and highly sensitive to habitat disturbance, and by other medium–large mammals and birds (Olmos et al., 1999; Guix and Ruiz, 2000; Giombini et al., 2009; Sica et al., 2014). Specifically, we addressed whether palm recruits in partially defaunated fragments present lower genetic diversity, higher inbreeding and stronger fine-scale SGS than in protected forest, where the main disperser is still abundant. We also evaluated whether recruitment around adult palms in the fragments was composed of a lower proportion of dispersed recruits (that is, recruits not assigned as the progeny of the focal adult). Finally, we assessed differences in mating patterns (that is, selfing rate and diversity of pollen donors) between areas.

MATERIALS AND METHODS
Study species and region
Syagrus romanzoffiana (Cham.) Glassman is widely distributed in the southeast of South America, and one of the most common palms in the Atlantic Forest. It is a monoecious species mainly pollinated by bees and beetles. Its mature infructescences produce 500–1000 single-seeded drupes of 2–3 cm, with an orange fleshy pericarp. The fruiting period extends most of the year, thus offering a keystone resource for numerous frugivores during periods of fruit scarcity (Keuroghlian and Eaton, 2008; Genini et al., 2009). Vertebrates that consume these fruits (Guix and Ruiz, 2000) are far more numerous than those capable of ingesting and effectively dispersing the seeds (Giombini et al., 2016). Besides tapirs, other potentially important dispersers are crab-eating foxes, coatis and large birds (for example, cracids and toucans). Some smaller frugivores (for example, crab-eating racoons, several rodents) may occasionally contribute to seed dispersal. In continuous protected areas, the density of early-stage recruits is typically highest in tapir latrines, and is higher in the vicinity of adult palms than in random sites (Giombini et al., 2009; Sica et al., 2014). Only few vertebrates predate S. romanzoffiana seeds, including some terrestrial rodents and, mainly, squirrels. The larva of the weevil Revena rubiginosa kills the seed embryo in early developmental, predispersal stages and is probably the main seed predator (Olmos et al., 1999; Mendes et al., 2016).

This study was conducted in the north of Misiones Province, northeastern Argentina (Supplementary Figure S1) that represents the southern limit of the Upper Paraná Atlantic Forest of South America. This ecoregion has been globally reduced to ∼8% of its original extension during the second half of the twentieth century, and currently presents a high degree of fragmentation (Di Bitetti et al., 2003). The study region contains both large areas of protected continuous forest and disturbed forest fragments that were originated by human settlement and local demographic growth since five decades ago. Sampling was conducted in the Iguazú National Park (henceforth 'protected area'), which is a 58 600 ha natural reserve, and five forest remnants ('disturbed area'). Both areas have been connected by continuous forest until fragmentation began (Supplementary Figure S1). The forest fragments range from 20 to 150 ha and originated <40 years ago. They are located in the surroundings (2.5–5.5 km distant) of the rural locality of Comandante Andresito that is 10 km east of Iguazú National Park eastern limit (Supplementary Figure S1). The protected area harbours a considerable richness and abundance of large vertebrates, including tapirs and the remaining potential dispersers of S. romanzoffiana. The disturbed area includes forest fragments that are partially defaunated (Iezzi et al., unpublished) while preserving similar vegetation structure and composition with respect to the protected area. In particular, the tapir is locally extinct. Coatis and large birds (that is, cracids and toucans) are still present with roughly similar abundances in the two areas, whereas foxes seem to be more abundant in the disturbed area. Agoutis and pacas, two medium–large rodents that probably play a minor role in effective dispersal of S. romanzoffiana, seem to be scarcer in the disturbed area (Iezzi et al., unpublished).

In the study region, *S. romanzoffiana* adults are normally found at densities of 10–20 individuals per ha, with a fairly random spatial distribution (Olmos *et al.*, 1999). Notably, disturbed areas, such as logged fragments and even crop fields, often retain similar adult densities. This is probably because this palm has a well-recognized cultural and ornamental value among native people. In addition, *S. romanzoffiana* is not considered a valuable timber species and local inhabitants usually avoid cutting its stem because it damages saws and chainsaws. Therefore, the spatial distribution of adults does not strictly resemble the pattern of landscape fragmentation but exhibits a higher degree of continuity. The similarity in densities and spatial distribution between protected and disturbed areas should ameliorate the effect of some potentially confounding factors on genetic variation (for example, pollination distances, mating patterns and number of nearby maternal seed sources; see Discussion).

Sample collection and sampling schemes
Leaf tissue was collected from early-stage individuals (hereafter 'recruits') and adult palms representing candidate maternal sources (see below). Recruits corresponded either to seedlings presumably <1 year old, or saplings (early juveniles) >1 year old, with wider (but still not pinnate) leaves.

Consequences of seed-dispersal limitation on genetic variation and spatial genetic structure. The first sampling scheme (Table 1) was carried out to detect the potential short-term consequences of seed-dispersal limitation, with a particular focus on tapir extinction, on genetic variation and fine-scale SGS of recruits. We sampled recruits from plots of 100 m × 10 m ('100 × 10 m plots') established in four forest fragments and four sites in the protected area (Table 1; Supplementary Figure S1) in order to estimate the genetic diversity, inbreeding levels and fine-scale SGS. In two sites (D4 and P1), extremely high-density vegetation precluded sampling from some portions of the plot, and hence the sampling was extended beyond 100 m (up to 120 m) in order to compensate for the unsampled portion (Supplementary Figure S2). All recruits within each plot were counted and mapped by defining an X–Y coordinate system and then a subset of 51–52 recruits were sampled at random, seeking to encompass the entire plot (Supplementary Figure S2). Within the protected area, recruits tend to be spatially associated with tapir latrines (Giombini *et al.*, 2009; Sica *et al.*, 2014). In addition to its prominent contribution to *S. romanzoffiana* recruitment, this large frugivore shows the most contrasting abundances between the protected area and forest fragments (that is, abundant and absent, respectively). Therefore, plots in the protected area were located so that they included tapir latrines. The number of plots and number of individuals per plot were chosen in accordance to the compromise between each other for permitting an affordable experimental effort and cost. Whereas sample sizes of 50 individuals may be somewhat borderline for detecting SGS (Epperson, 2005), sampling

from four sites is preferable than from a single site, as patterns of SGS often show considerable heterogeneity (Smouse *et al.*, 2008). In addition, the use of highly variable markers (that is, microsatellites) may also aid in estimating SGS accurately (Epperson, 2005). Because the local density of nearby adults (potential contributing mothers) may affect the strength of the spatial genetic structure of recruits (Hamrick *et al.*, 1993; Vekemans and Hardy, 2004), we also counted and mapped the adults within plots extended (symmetrically) to 140 m × 40 m (Supplementary Figure S2).

Comparison of parental sources of progenies recruited around adult palms. The second sampling design was conducted both to evaluate whether the disturbed area showed a stronger aggregation of progenies around maternal trees and whether mating patterns pointed to lower levels of pollen-mediated gene flow. In each area, eight focal adults and a subset of recruits established around them were sampled to estimate the proportion of dispersed recruits, the diversity of pollen donors and the selfing rate. At each focal site, a 3 m radius circular plot centred on the adult stem was delimited and 10–16 recruits (≈15 on average) were randomly sampled. Leaf tissue from the focal adult (candidate mother) was also collected. Focal adults in the protected area were located along a 4 km road section. In the disturbed area, focal adults were situated in four forest fragments, three of which were the same as those used for estimating genetic diversity and SGS (Supplementary Figure S1). In all cases, focal adults were separated by at least 350 m.

DNA extraction and genotyping
We extracted DNA from 5–35 mg leaf tissue with the NucleoSpin 96 Plant II kit (Macherey-Nagel, Düren, Germany) and genotyped all individuals using a set of nine, previously characterized microsatellites (http://tomato.bio.trinity.edu/manuscripts/12-6/mer-12-0111.pdf). The nine microsatellites, genotyping procedures and the estimation of genotyping errors and null alleles were described in detail in Giombini *et al.* (2016). Briefly, the frequency of null alleles was <0.05 for all loci and the allele typing error rate per locus was <0.01. Despite their very low occurrences, null alleles and genotyping errors were considered in maternity analyses as described by Giombini *et al.* (2016). The whole data set contained <0.5% of missing data. The genotypes of all individuals ($N = 674$), as well as geographic coordinates of sampling locations, were deposited at Dryad (doi:10.5061/dryad.s0vf6).

Data analyses
Genetic diversity. Allelic richness (that is, the number of different alleles) and gene diversity (that is, the expected heterozygosity under random mating) were estimated using SPAGeDi 1.4 (Hardy and Vekemans, 2002). Allelic richness was corrected to standardized (rarefacted) sample sizes (that is, equal number of individuals and genes) (Kalinowski, 2004). Gene diversity was calculated according to the unbiased estimator of Nei (1978). Both genetic diversity estimates were compared between the protected and disturbed area by the Wilcoxon matched-pairs (one-sided) test, using STATISTICA 6.0 (StatSoft, Inc., 2001). This test corresponds to a nonparametric randomized blocks design for paired comparisons (Sokal and Rohlf, 1995). For our purposes, the two areas (protected and disturbed area) were the treatments of interest, whereas loci represented the blocks. Thus, variation in genetic diversity levels among loci, which may arise from different mutation rates, potential selective constraints and so on, was controlled by pairing treatments by locus (Kalinowski, 2004).

Inbreeding. The inbreeding coefficient (F_{is}) was computed as the kinship coefficient (estimator defined in Loiselle *et al.*, 1995, see further details below) between homologous genes (alleles) within individuals (Hardy, 2003) using SPAGeDi 1.4. This coefficient estimates the probability of identity between two alleles drawn from the same individual with respect to the probability of identity among two alleles randomly sampled from the 'reference population', and provides a measure of the departure from random mating (Hardy, 2003; Vekemans and Hardy, 2004). The 'reference allele frequencies' (Hardy, 2003), which are required for this and other estimations (see below), were obtained by pooling the genotypes of 819 individuals (including those analysed herein) sampled across a broad area in the north of Misiones province. These reference allele frequencies are provided in Giombini *et al.* (2016). A one-sided test was

Table 1 Summary of sample collection of *Syagrus romanzoffiana* recruits in a disturbed and a protected area of the Atlantic Forest in the northeast of Argentina

Area	Site[a]	Fragment size (ha)	N	n	Number of nearby adults[b]
Disturbed	D1	90	129	52	6
	D2	150	110	52	4
	D3	25	75	52	4
	D4	150	188	52	7
Protected	P1	58 600	77	52	7
	P2	58 600	116	52	2
	P3	58 600	421	51	6
	P4	58 600	560	51	2

Abbreviations: *N*, total number of recruits found in the 100 × 10 m plots; *n*, number of genotyped recruits.
Four sites (D1–D4 and P1–P4) were sampled in each area.
[a]Geographic coordinates have been deposited at Dryad (doi:10.5061/dryad.s0vf6).
[b]The number of adults found within the extended 140 × 40 m plots (see Supplementary Figure S2).

applied to assess whether inbreeding coefficients were significantly higher than expected under random mating, after generating the null distribution based on 1000 permutations of alleles among individuals (see SPAGeDi 1.4 user's manual).

Fine-scale SGS. We characterized the fine-scale SGS of recruits within the 100×10 m plots according to Vekemans and Hardy (2004). This approach is based on the analysis of spatial genetic autocorrelation (that is, the relation between pairwise kinship coefficients and spatial distance among individuals) and was conducted with SPAGeDi 1.4 (Hardy and Vekemans, 2002). We used the kinship estimator (ρ_{ij}) proposed in Loiselle *et al.* (1995) that is generally the most accurate and does not suffer biases in the presence of low frequency alleles (Vekemans and Hardy, 2004; see SPAGeDi 1.4 user's manual). Kinship-distance curves were characterized by the slope of the regression (b) of pairwise kinship values on the natural logarithm of physical distance (Vekemans and Hardy, 2004). The overall strength of SGS was quantified by the Sp statistic, which is defined as $Sp = -b/(1-\rho_1)$, where ρ_1 is the average kinship between nearest-neighbour individuals (that is, individuals in the first distance class) (Vekemans and Hardy, 2004). Sp provides a single summary statistic that synthesizes the strength of SGS and is fairly insensitive to the sampling scheme. Moreover, as is the case for b, it does not depend on the set of distance intervals arbitrarily chosen for constructing autocorrelograms. This statistic is thus suitable for measuring the strength of SGS on fine spatial scales and for comparing values among populations, sites or samples under different conditions (Vekemans and Hardy, 2004; Wang *et al.*, 2011). The SGS patterns were visualized by constructing autocorrelograms with six distance intervals (I_1: 0–5 m, I_2: >5–10 m, I_3: >10–20 m, I_4: >20–40 m, I_5: >40–80 m and I_6: >80—maximum distance), so that the estimates of average kinship coefficients were supported by more than 100 pairwise comparisons for most distance intervals. Standard errors for the estimated b and ρ_1 values were obtained by a jackknife procedure over loci. Statistical significances were assessed after obtaining the 95% confidence limits of the null distribution by performing 1000 random permutations of the spatial locations of individuals within plots and evaluating whether the observed b or ρ_1 values were lower or higher than the 5% lower or upper limits of the null distribution (one-sided test), respectively. The P-values for the eight plots were then adjusted for multiple testing by the sequential Bonferroni procedure, declaring significance at an experiment-wise level of $\alpha = 0.05$. The statistical comparison of ρ_1, b and Sp between the disturbed and protected area was conducted by using per-locus estimates averaged across sites in each area and applying a one-sided Wilcoxon matched-pairs test, analogously to the comparison of genetic diversity. In addition, for descriptive purposes, average multilocus values of ρ_1, b and Sp were computed from the four sites in both areas.

We complemented the above-mentioned analysis by evaluating the heterogeneity in fine-scale SGS patterns within and among sampling sites and areas. For this purpose, we applied the method developed by Smouse *et al.* (2008), as implemented in GenAlEx 6.5 (Peakall and Smouse, 2012). This method is based on a genetic correlation coefficient (r) that is closely related to the kinship coefficient because it depends on the proportion of alleles shared between individuals. However, r does not include information on external allele frequencies of a reference population (Smouse *et al.*, 2008). Using the multipopulation testing approach (see Smouse *et al.* (2008), and GenAlEx user's guide), we specifically evaluated: (1) whether the combined patterns of genetic autocorrelation in each area indicated a nonrandom distribution of genotypes across space (by testing the null hypothesis that $r = 0$ for all distance intervals of the resulting, multisite correlogram by the ω_D statistic in each area), (2) whether the four autocorrelograms in each area were homogeneous (by means of the ω_E statistic in each area), (3) whether the overall patterns of genetic autocorrelation were significantly different between the disturbed and protected area (by comparing the resulting correlograms between areas by the ω_E statistic) and (4) which distance intervals showed more divergent r values between the disturbed and protected area (by comparing r values at each distance interval between the resulting correlograms by the t^2 statistic). Statistical significances were determined based on 1000 random permutations of individual spatial positions and a significance level of $\alpha = 0.01$.

Proportion of dispersed recruits around adults. The proportion of recruits not mothered by the focal adult (candidate mother) in the 3 m radius circular plots was estimated by three maternity exclusion methods: (1) full-pedigree maximum likelihood (hereafter 'ML') reconstruction (Wang and Santure, 2009), (2) ML estimation of pairwise relationships for recruit–candidate mother dyads (Wagner *et al.*, 2006) and (3) crude exclusion analysis based on at least one and two mismatching loci (Wang, 2010). They were implemented with COLONY v2.0.5.0 (Jones and Wang, 2010), ML-RELATE (Kalinowski *et al.*, 2006) and both MicroErrorAnalyzer (Wang, 2010) and WHICHPARENTS 1.0 (Eichert, 1999), respectively. These three alternative approaches were used to obtain a more robust comparison of the relative amounts of dispersed recruits, although the former generally outperforms the others, because it exploits more information from the data (Wang, 2012). The accommodation of genotyping errors and potential occurrence of null alleles, as well as further details of the implementation of the three methods, are explained in Giombini *et al.* (2016). We tested whether the relative amount of dispersed recruits, as inferred from full-pedigree ML reconstruction, was lower in the disturbed than in the protected area by the nonparametric (one-sided) Mann–Whitney test implemented in STATISTICA 6.0 (StatSoft, Inc., 2001).

Selfing rate and diversity of pollen donors

The selfing rate (s) and diversity of pollen donors were estimated based on the pedigree relationships and parental genotypes of the recruits sampled around adult palms, as inferred by full-pedigree reconstruction with COLONY (Wang *et al.*, 2012). Those recruits not assigned as the progeny of the focal adults (that is, dispersed recruits) were discarded in these estimations, because the inferred parental genotypes cannot be assigned unambiguously to paternal (pollen) and maternal (seed) sources. We thus obtained a data set analogous to those based on progeny arrays in which the maternal (focal adult) genotype is known and the inference of paternal genotypes is straightforward. We computed s as the proportion of recruits for which the focal adult was assigned as both the mother and the father (Wang *et al.*, 2012). Provided the method developed by Wang *et al.* (2012) is based on pedigree reconstruction without requiring *a priori* information on any parental genotype, we also obtained an alternative estimation of selfing rate (s') from the group of recruits sampled in the 100×10 m plots.

The diversity of pollen donors was quantified both in terms of the absolute number of different contributing fathers or 'paternal richness' (N_p) and the effective number of contributing fathers or 'paternal diversity' (N_{ep}) that is informative about the number and evenness of paternal contributions. Owing to the unequal number of nondispersed recruits among focal sites, N_p was estimated as a function of the number of nondispersed recruits using the rarefaction/extrapolation methods ("individual-based abundance data') implemented in EstimateS 9.0 (Colwell, 2013). The effective number of fathers was calculated as $N_{ep} = 1/PPI$, where PPI is the unbiased estimator of the probability of paternal identity (Smouse and Robledo-Arnuncio, 2005), computed with an analogous equation to that described in Scofield *et al.* (2011) for the probability of maternal identity (q^*_{gg}). We applied the nonparametric (two-sided) Mann–Whitney test for comparing s, N_p for standardized samples of 7 and 16 recruits ($N_p(7)$ and N_p (16), respectively), and N_{ep} between the protected and disturbed area using STATISTICA 6.0.

RESULTS
Genetic diversity and inbreeding
A total of 414 recruits sampled from the 100×10 m plots in the disturbed ($N = 208$) and protected ($N = 206$) area were genotyped (Table 1). Allelic richness was significantly higher in the protected area, as revealed by a higher number of alleles across all but two loci (Table 2, $T_{wilcoxon} = 2$, $P = 0.013$). EE54 was uninformative in this comparison as it only presented two alleles (Table 2) that occurred at least once in all plots. Gene diversity was identical in both areas (Table 2, $T_{wilcoxon} = 19$, $P = 0.339$).

Inbreeding coefficients were consistent with random mating patterns for all plots (Table 3) and showed a lower average value in the disturbed vs protected area ($F_{is} = 0.015$ vs $F_{is} = 0.038$, respectively).

Table 2 Summary statistics of genetic diversity of *Syagrus romanzoffiana* recruits sampled in a disturbed and a protected area of the Atlantic Forest in the northeast of Argentina

Area																				
Site									Locus											
	mBgCIR091		CNZO4		mBgCIR053		CAC2		BgO2-10		EE54		CNZ26		CNZ44		CNZ50		Multilocus average	
	A	H_e	A	H_e	A	H_e	A	H_e	A	H_e	A	H_e	A	H_e	A	H_e	A	H_e	A	H_e
Disturbed																				
D1	5.99	0.65	6.83	0.70	6.88	0.56	4.94	0.53	6.88	0.65	2.00	0.51	6.94	0.33	3.94	0.23	4.00	0.64	5.38	0.53
D2	6.88	0.68	6.96	0.64	6.94	0.63	6.00	0.59	6.00	0.76	2.00	0.50	6.99	0.34	3.00	0.41	4.00	0.45	5.42	0.56
D3	6.00	0.72	7.88	0.66	7.88	0.63	5.83	0.49	7.88	0.67	2.00	0.50	7.88	0.30	2.00	0.18	4.00	0.62	5.70	0.53
D4	6.00	0.65	8.77	0.68	6.94	0.60	5.88	0.44	7.00	0.70	2.00	0.50	7.00	0.48	3.88	0.19	4.00	0.34	5.72	0.51
Average	6.22	0.67	7.61	0.67	7.16	0.61	5.66	0.51	6.94	0.69	2.00	0.50	7.20	0.36	3.21	0.25	4.00	0.51	5.56	0.53
Protected																				
P1	5.00	0.67	7.88	0.75	5.94	0.60	6.96	0.68	7.83	0.74	2.00	0.49	7.96	0.40	2.94	0.09	4.00	0.42	5.61	0.54
P2	6.00	0.71	9.77	0.70	7.82	0.61	6.94	0.49	9.77	0.78	2.00	0.50	7.94	0.42	3.00	0.18	4.00	0.52	6.36	0.55
P3	4.96	0.65	5.92	0.56	7.92	0.72	4.92	0.15	6.92	0.73	2.00	0.46	9.92	0.64	3.00	0.20	4.96	0.56	5.61	0.52
P4	7.96	0.77	6.96	0.65	9.84	0.68	8.00	0.18	7.92	0.74	2.00	0.51	8.92	0.51	4.92	0.15	4.00	0.64	6.72	0.54
Average	5.98	0.70	7.63	0.66	7.88	0.65	6.70	0.37	8.11	0.75	2.00	0.49	8.69	0.49	3.47	0.16	4.24	0.53	6.08	0.53

Abbreviations: A, allelic richness rarefied to minimum sample sizes of 98 genes; H_e, gene diversity.
Four sites (D1–D4 and P1–P4) were sampled in each area. Shown are the values for each site and locus, as well as multisite and multilocus averages.

Table 3 Inbreeding levels and spatial genetic autocorrelation of *Syagrus romanzoffiana* recruits sampled in a disturbed and a protected area of the Atlantic Forest in the northeast of Argentina

Area	Inbreeding levels		Spatial genetic autocorrelation		
Site	F_{is}	P-value	ρ_1 (s.e.)	b (s.e.)	Sp
Disturbed					
D1	0.008	0.723	0.126[a] (0.031)	− 0.027[a] (0.008)	0.031
D2	0.006	0.829	0.070[a] (0.024)	− 0.008 (0.007)	0.008
D3	− 0.008	0.911	0.102[a] (0.024)	− 0.029[a] (0.009)	0.032
D4	0.053	0.124	0.009 (0.016)	− 0.006 (0.003)	0.006
Average	0.015		0.077	− 0.017	0.019
Protected					
P1	0.053	0.306	0.078[a] (0.018)	− 0.009 (0.003)	0.009
P2	0.055	0.062	0.025[a] (0.011)	− 0.008 (0.003)	0.008
P3	0.040	0.460	0.070[a] (0.029)	− 0.024[a] (0.007)	0.026
P4	0.005	0.626	0.018 (0.011)	0.003 (0.003)	− 0.003
Average	0.038		0.048	− 0.010	0.010

Abbreviations: b, slope of regression of pairwise kinship on the natural logarithm of distance among recruits; F_{is}, inbreeding coefficient; ρ_1, average kinship between nearest-neighbour recruits; Sp, Sp statistic.
Four sites (D1–D4 and P1–P4) were sampled in each area. See the main text for a further definition of the parameters.
[a]Significant values (P<0.05) after correction for multiple testing.

Fine-scale spatial genetic structure

Two disturbed and one protected sites showed a significant decay of kinship with distance among recruits (Table 3). Nearest-neighbour recruits (0–5 m distance interval) exhibited higher kinship values than expected under no spatial structuring in three out of four sites in both areas (Table 3 and Figure 1). A more negative average regression slope along with a higher average kinship among nearest neighbours resulted in an almost twofold average Sp statistic in the disturbed area ($Sp = 0.019$ vs $Sp = 0.010$; Table 3). Seven out of nine loci yielded higher ρ_1, more negative b and higher Sp values in the disturbed area.

These differences were significant or marginally significant ($T_{wilcoxon} = 7$, $P = 0.033$ for comparison of ρ_1, and $T_{wilcoxon} = 11$, $P = 0.087$ for comparison of both b and Sp).

As suggested by the visual inspection of the correlograms (Figure 1), the results of the heterogeneity analyses (Supplementary Table S1) indicated that: (1) the genotypes of the sampled recruits were nonrandomly distributed across space in both the disturbed ($\omega_D = 68.87$, $P = 0.001$) and protected areas ($\omega_D = 66.56$, $P = 0.001$), (2) the genetic autocorrelation patterns were quite heterogeneous among sites in both the disturbed ($\omega_E = 51.78$, $P = 0.001$) and protected areas ($\omega_E = 47.15$, $P = 0.001$), (3) the overall patterns of genetic autocorrelation differed significantly between areas ($\omega_E = 35.22$, $P = 0.001$) and (4) the larger differences in the overall patterns of genetic autocorrelation between areas occurred for the last distance intervals; the r value was higher in the disturbed area at the 40–80 m interval ($r = -0.014$ vs $r = -0.038$, $t^2 = 8.31$, $P = 0.006$) and lower at distances of >80 m ($r = -0.045$ vs $r = 0.027$, $t^2 = 11.77$, $P = 0.001$). The r value for the first distance interval was slightly but not significantly higher in the disturbed area ($r = 0.086$ vs $r = 0.067$, $t^2 = 1.88$, $P = 0.173$).

Proportion of dispersed recruits around adults

A total of 244 recruits sampled around 16 focal adults in the disturbed ($N = 128$) and protected ($N = 116$) areas were genotyped (Table 4). All maternity exclusion methods indicated a lower proportion of dispersed recruits around adults in the disturbed area (Figure 2). The median percentages (first–third quartiles) of dispersed recruits inferred by full-pedigree ML reconstruction, which should provide the most accurate estimates, were 0.0 (0.0–15.6) and 14.4 (0.0–39.3) in the disturbed and protected area, respectively (Figure 2). The difference, however, was not significant ($U_{Mann–Whitney} = 20.5$, $P = 0.117$).

Selfing rate and diversity of pollen donors

Average selfing rates estimated from the focal sites ($n = 8$, $s = 0.01 \pm 0.01$ s.e. and $s = 0.04 \pm 0.04$ s.e. for disturbed and protected

6
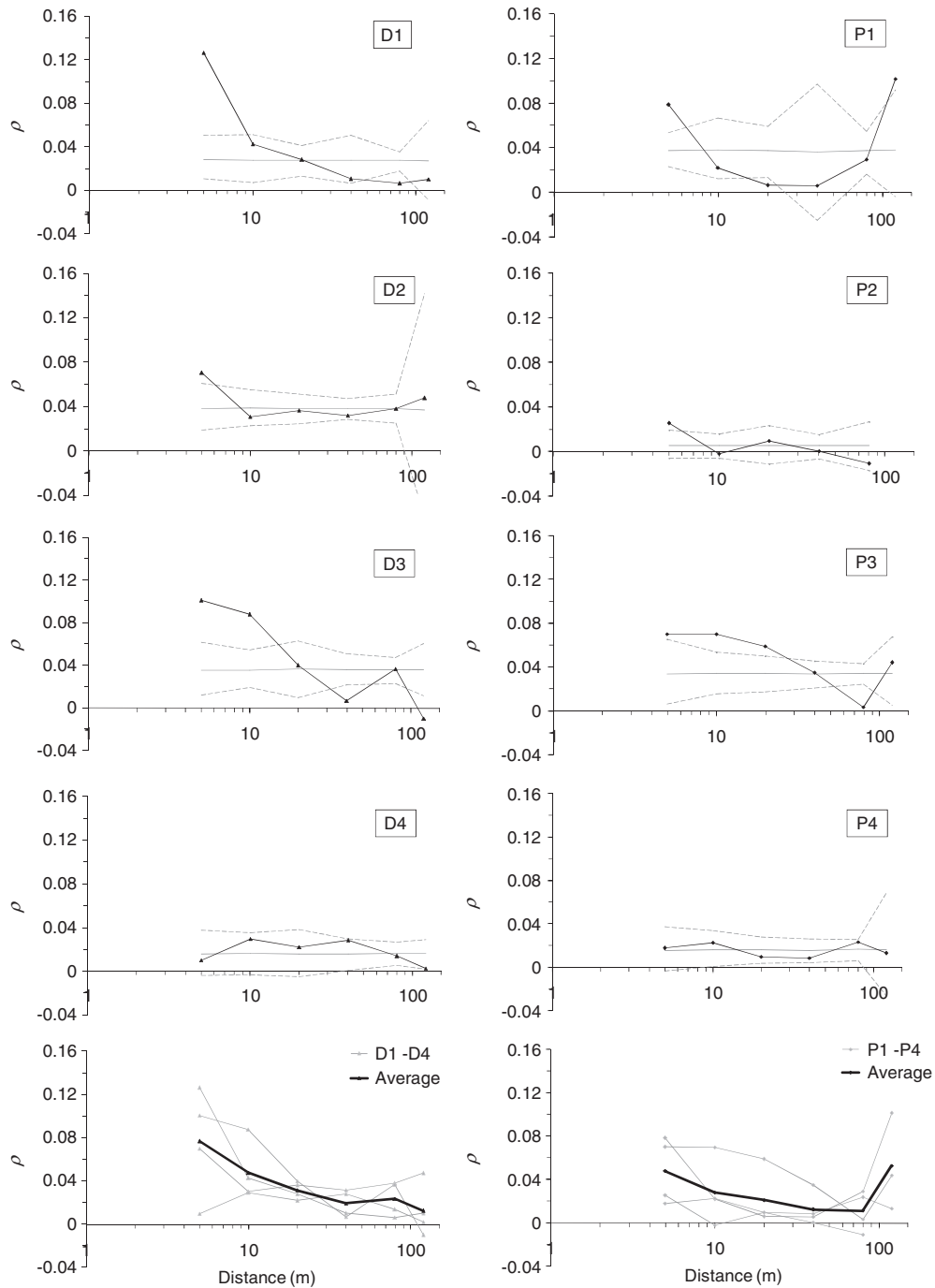
Figure 1 Spatial genetic autocorrelograms for *Syagrus romanzoffiana* recruits sampled in a disturbed and a protected area of the Atlantic Forest in the northeast of Argentina. The autocorrelograms for the four disturbed (D1–D4) and protected (P1–P4) sites are shown in left and right columns, respectively. All x axes represent distance in m (logarithmic scale). Average kinship values for each distance interval (black markers connected by lines) are indicated at the upper limit of the interval. Distance intervals are as follows: 0–5 m, >5–10 m, >10–20 m, >20–40 m, >40–80 m and >80–120 m. Also shown are the average expected values under random spatial distribution of genotypes (grey continuous line) and their 95% confidence limits (grey dotted lines). Single-site (grey lines and markers) and multisite average (black lines and markers) autocorrelograms are plotted together at the bottom panels.

area, respectively; Table 4) were different from those obtained from recruits sampled in the 100×10 m plots ($n = 4$, $s' = 0.12 \pm 0.01$ s.e. and $s' = 0.11 \pm 0.01$ s.e. for disturbed and protected area, respectively; Supplementary Table S2). In any case, and regarding our purposes, the estimated values did not differ significantly between areas, as indicated

by the Mann–Whitney test applied for the comparison of s ($U = 27$, $P = 0.955$; Table 4) or by the largely overlapping distributions of s' estimates (Supplementary Table S2).

The richness and diversity of pollen donors were also similar between areas, with not (or marginally) significantly lower values in

Table 4 Mating pattern parameters (selfing rate and diversity of pollen donors) in *Syagrus romanzoffiana* estimated from recruits sampled around focal adults in a disturbed and a protected area of the Atlantic Forest in the northeast of Argentina

Area	Focal site[a]	N	n	n'	s	$N_p(7)$	$N_p(16)$	N_{ep}
Disturbed	d1	64	16	16	0.00	3.32	4.00	3.36
	d2	51	16	13	0.08	2.98	4.35	2.47
	d3	51	16	16	0.00	2.76	3.00	2.88
	d4	72	16	13	0.00	2.69	3.00	5.64
	d5	77	16	14	0.00	3.86	5.10	4.38
	d6	62	16	16	0.00	2.37	3.00	1.75
	d7	69	16	16	0.00	3.18	4.00	2.88
	d8	47	16	16	0.00	3.42	4.00	4.03
	Average	61.6	16	15	0.01	3.07	3.81	2.80[b]
	Median[c]				0.00	3.08	4.00	3.12
	Quartiles				0.00–0.00	2.73–3.37	3.00–4.17	2.67–4.21
Protected	p1	14	14	10	0.00	3.77	4.00	5.10
	p2	14	14	7	0.29	4.00	4.17	5.44
	p3[d]	15	14	3	—	—	—	—
	p4	27	16	16	0.00	3.32	4.00	3.36
	p5	25	16	13	0.00	3.27	4.17	4.93
	p6	18	16	16	0.00	3.42	4.00	4.03
	p7	10	10	9	0.00	2.90	3.00	2.84
	p8	19	16	16	0.00	3.19	4.00	3.95
	Average	17.7	14.5	11.2	0.04	3.41	3.91	3.65[b]
	Median[c]				0.00	3.32	4.00	4.03
	Quartiles				0.00–0.00	3.19–3.77	4.00–4.17	3.36–5.10

Abbreviations: N, total number of recruits found in the 3 m radius plots around focal adults; n, number of genotyped recruits; n', number of recruits assigned as the progeny of the focal adult; N_{ep}, effective number of pollen donors; $N_p(x)$, number of different pollen donors estimated for rarefied/extrapolated samples of x recruits; s, selfing rate.
Eight focal sites (d1–d8 and p1–p8) were sampled in each area.
[a]Geographic coordinates have been deposited at Dryad (doi:10.5061/dryad.s0vf6).
[b]Calculated as $[PPI_M]^{-1}$ where PPI_M is the weighted mean of the probability of paternal identity across sites, according to Scofield et al. (2011).
[c]Medians plus first and third quartiles are only provided for the estimates that were subject to statistical comparison by the Mann–Whitney test.
[d]Estimates of mating patterns were not computed as n' was extremely low (n' = 3).

the disturbed vs protected area (Table 4; Mann–Whitney test for comparisons of N_p (7), N_p (16) and N_{ep}, respectively: $U = 16$, $P = 0.189$; $U = 24.5$, $P = 0.694$ and $U = 13$, $P = 0.094$).

DISCUSSION

In contemporary times, plants are probably suffering from various short-term genetic consequences because of anthropogenic disturbance (Aguilar et al., 2008). However, empirical evidence indicates that plant populations may respond in a complex or subtle way and hence the genetic impact may not be easily predicted (Lowe et al., 2005; Kramer et al., 2008). Elucidating the causes and consequences of the changes in the patterns of plant gene flow is crucial for defining conservation actions. Overall, our findings are strongly suggestive of some early genetic consequences resulting from the limitation in contemporary gene flow via seeds, but not pollen, in defaunated areas.

First, allelic richness, but not gene diversity, was significantly lower in the disturbed area. Such pattern may be expected from the rapid loss of low frequency (rare) alleles soon after population shrinkages (Young et al., 1996), as has been detected in some small or fragmented populations (see, for example, Lowe et al., 2005; Jump and Peñuelas, 2006). However, a reduction of allelic richness is unlikely to occur as a consequence of a genetic bottleneck for long-lived trees when < 50
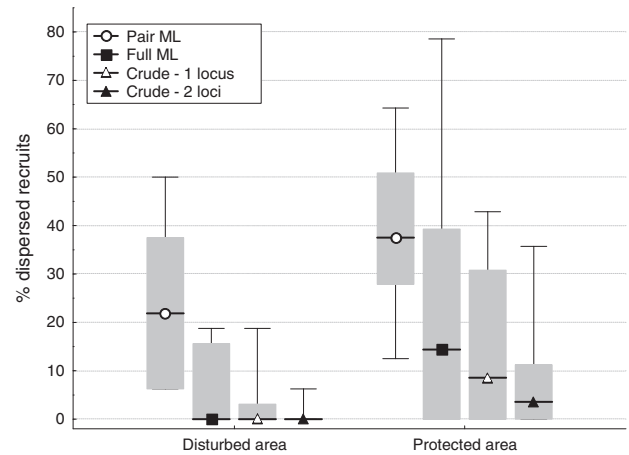
Figure 2 Percentages of *Syagrus romanzoffiana* dispersed recruits sampled around conspecific adults in a disturbed and a protected area of the Atlantic Forest in the northeast of Argentina. Shown are the values of median (black thick lines with markers), first and third quartiles (bottom and top of the boxes) and entire range (whiskers) estimated from the eight focal sites in each area. Values are presented for the four maternity exclusion methods. Crude-1 locus and Crude-2 loci, crude exclusion methods based on at least one and two mismatching loci, respectively (implemented with WHICHPARENTS); Full ML, full-pedigree maximum likelihood method (implemented with COLONY); Pair ML, pairwise maximum likelihood method (implemented with ML-RELATE).

years have elapsed since fragmentation (Aguilar et al., 2008; Wang et al., 2011; Browne et al., 2015), as it is the case for the present study. Moreover, as mentioned above, the fragmentation and potential decline of the effective population size in *S. romanzoffiana* was probably not very severe. Thus, although not surveyed, we do not expect that the parental populations (*S. romanzoffiana* adults) in the study areas showed different allelic richness to account for the difference observed in recruits. Therefore, in the context of our study, the difference in allelic richness may be interpreted more plausibly as a spatial restriction for the occurrence of certain alleles in defaunated sites, as a consequence of gene dispersal limitation, rather than the loss of those alleles by a population bottleneck.

Second, the disturbed area showed an overall trend toward a stronger fine-scale SGS. The differences of b and Sp among areas, however, were marginally significant and two disturbed sites (D2 and D4) showed no significant decrease of kinship with distance (Table 3). The moderate difference between areas is consistent with the fact that the disperser assemblage in the disturbed area is only partially depleted. Thus, the remaining dispersers probably contribute, in some degree, to the admixture of genotypes over the spatial scale studied here. In addition, SGS patterns were fairly heterogeneous, suggesting that several factors, beside seed dispersal levels, were also involved in shaping the SGS at each site. Such heterogeneity seems to be the norm in empirical studies evaluating SGS from different sites, populations or conditions (see, for example, Gehring and Delph, 1999; Smouse et al., 2008; De-Lucas et al., 2009; Zeng et al., 2012).

The differences in allelic richness and SGS between areas may be partly accounted for by the contribution of tapir-mediated dispersal, as sampling sites in the protected area included tapir latrines. This frugivore has been shown to promote the recruitment of unrelated neighbouring seedlings and increase the number of alleles and maternal genotypes as more defecations are accumulated (Giombini et al., 2016). Therefore, the observed patterns would be somewhat biased towards a genetic footprint left by tapirs. As our sampling

scheme does not allow us to tease apart the relative contribution of tapirs and other dispersers, we cannot determine the extent to which the former could explain the differences among the protected and disturbed area. Given that tapir latrines (and single defecations) containing plant recruits are quite widespread across well-protected areas (Fragoso and Huffman, 2000; Giombini et al., 2009; Barcelos et al., 2013), the effect on the emerging patterns of genetic variations might be significant. Addressing this issue would be valuable because tapirs are among the most sensitive vertebrates to habitat disturbance and their extirpation is becoming increasingly common in South America (Peres, 2001; Stoner et al., 2007; Cruz et al., 2014).

Other factors potentially influencing SGS patterns and allelic richness did not show significant differences between the disturbed and protected area. For instance, a higher local density of adult palms may reduce SGS because of a larger overlap of maternal seed shadows (Hamrick et al., 1993; Vekemans and Hardy, 2004). Nonetheless, the natural density of S. romanzoffiana adults is fairly uniform across the entire study region and is not expected to be currently reduced in the disturbed area. In fact, the number of nearby adults was not higher in the protected sites (4.25 vs 5.25 on average; Table 1) and we found no correlation between Sp values and the number of nearby adults ($R_{spearman} = 0.22$, $P = 0.600$).

An increase in inbreeding levels is a potential short-term consequence of habitat disturbance (see, for example, Lowe et al., 2005; Jump and Peñuelas, 2006; Aguilar et al., 2008) and may affect and be affected by SGS patterns (Loveless and Hamrick, 1984; Ellstrand and Elam, 1993). Inbreeding is largely determined by selfing rates and mating among closely related individuals (that is, biparental inbreeding). Remarkably, neither selfing rates nor overall inbreeding were higher in the disturbed area. In addition, the richness and diversity of pollen donors were also similar in both areas. The diversity of pollen donors, however, showed a marginally significant higher value in the protected area, and thus cannot be completely discarded as a factor contributing to the higher allelic richness and weaker SGS in that area. Overall, our results suggest that patterns of pollen-mediated gene flow are not severely affected in the disturbed area.

No previous assessment of S. romanzoffiana mating system has been reported so far. Measured selfing rates varied depending on the method used for the estimation. This is not surprising because each method involved a different type of sample. Whereas the recruits sampled from the focal sites represented the progeny of a single, known maternal source, those sampled from the 100×10 m plots most probably originated from multiple, unknown parents. Regardless of the estimation method, global averages were lower than 0.20 ($s = 0.02 \pm 0.08$ s.d., $n = 16$ and $s' = 0.11 \pm 0.02$ s.d., $n = 8$, respectively). Therefore, based on these rough estimates, and as previously suggested by Giombini et al. (2016), S. romanzoffiana may be considered a predominantly outcrossing species (Goodwillie et al., 2005). Such mating pattern seems to be pervasive among palms and other tree families in the Neotropics (Ward et al., 2005; Nazareno and Dos Reis, 2012; and references therein). However, selfing rates could have been underestimated by our procedure if some selfed progeny underwent embryo abortion and/or differential mortality at the seed or early seedling stages (Husband and Schemske, 1996).

Another suggestive signal of restricted seed dispersal was revealed by the comparison of the proportion of dispersed recruits around adults. A stronger spatial aggregation of seeds or seedlings around conspecific (putative maternal) adults in defaunated areas seems to be a widespread pattern for animal-dispersed plants (Kurten, 2013). The number of S. romanzoffiana recruits around adults was indeed larger in the disturbed area (Sica et al., 2014; see also Table 4). It is unlikely

that a decrease in seed predation is responsible for this pattern, although we cannot rule it out completely. Seed predators of S. romanzoffiana should be affected less drastically than dispersers, because the former involve smaller vertebrates (for example, squirrels and small rodents) that are more resilient to disturbance than the latter. Remarkably, vertebrate seed predation of another palm has been shown to increase in defaunated areas in the Atlantic Forest because of higher abundance or activity of small rodents, including squirrels (Galetti et al., 2015). Furthermore, a potential decrease in vertebrate seed predation is likely to be compensated, at least partially, by an increase in invertebrate seed predation (Kurten, 2013; Mendes et al., 2016). In particular, a negative correlation has been found between seed predation by squirrels and invertebrates in S. romanzoffiana, and this is as expected because the former prey not only on seeds but also on the invertebrate larvae (Mendes et al., 2016).

In addition to being more numerous, recruits around adults in the disturbed area consisted of a lower proportion of dispersed individuals (that is, those originated from a maternal source other than the focal adult). This latter result cannot be explained by a lower rate of seed predation per se. Instead, this probably reflects the lower abundance of dispersal agents. Few studies have so far documented this pattern using genetic analysis (Pacheco and Simonetti, 2000; Wang et al., 2007), although it might be a widespread signature in the so-called 'empty' or 'half-empty' forests (Kurten, 2013). However, we cannot be conclusive on this respect as the observed differences, although consistent across several maternity exclusion methods, were not significant. A limited statistical power might be, at least partially, responsible for such result.

Overall, our results pointed consistently to a disturbance on the same underlying process, that is, the restriction of gene dispersal via seeds. The contrasting genetic patterns revealed in this study may be signalling potentially deleterious effects in the short and long term that would result from the loss of genetic variation and an increase in SGS. However, because of selective and demographic effects occurring throughout the life cycle, patterns of genetic structure in early cohorts may or may not be imprinted in subsequent stages and generations (Loveless and Hamrick, 1984; García and Grivet, 2011; Bialozyt et al., 2014). Further research is therefore needed to examine the extent to which these incipient changes in genetic patterns are ultimately detrimental to the viability and adaptive potential of plant populations.

As already mentioned, many plant populations are capable of maintaining, and even increasing, levels of pollen-mediated gene flow in disturbed areas. Remarkably, such capability has been verified in both wind- and insect-pollinated species, likely induced by opening vegetation structure and increasing pollinator foraging distances, respectively (see Lowe et al., 2005; Kramer et al., 2008; and references therein). In contrast, an increment of seed-mediated gene flow in deforested landscapes has been only reported for wind-dispersed species (Bacles et al., 2006), whereas opposite trends have been revealed in animal-dispersed species (Wang et al., 2011; Browne et al., 2015; this study). For instance, a significant increment of the SGS in recruits, but not adults, occurred in recently (30–40 years) fragmented areas in the neotropical, vertebrate-dispersed palm Oenocarpus bataua (Browne et al., 2015). Likewise, Wang et al. (2011) detected an increase of SGS in cohorts of the tropical Asian tree Castanopsis sclerophylla recruited in post-fragmentation times (< 50 years ago). The results from both studies suggested that the stronger SGS in the fragmented area was more likely a consequence of a restriction in animal-mediated seed dispersal than in pollen flow. Based on such evidence along with our results, we may propose that an increment in SGS should be occurring quite frequently in post-disturbance generations of tropical trees dispersed by frugivorous

vertebrates sensitive to anthropogenic disturbance. Moreover, as our findings concern the consequences of partial (modest) defaunation levels, a stronger genetic impact should be expected for more severely defaunated areas. Although the protection of plant-pollinators mutualisms, and associated gene flow, is undoubtedly desirable to sustain genetically viable plant populations, preservation of plant-disperser mutualisms, largely threatened in tropical systems, should also be imperative.

DATA ARCHIVING

Data available from the Dryad Digital Repository: http://dx.doi.org/10.5061/dryad.s0vf6.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

Aguilar R, Quesada M, Ashworth L, Herrerias-Diego Y, Lobo J (2008). Genetic consequences of habitat fragmentation in plant populations: susceptible signals in plant traits and methodological approaches. *Mol Ecol* **17**: 5177–5188.

Bacles CFE, Lowe AJ, Ennos RA (2006). Effective seed dispersal across a fragmented landscape. *Science* **311**: 628.

Barcelos AR, Bobrowiec PED, Sanaiotti TM, Gribel R (2013). Seed germination from lowland tapir (*Tapirus terrestris*) fecal samples collected during the dry season in the northern Brazilian Amazon. *Integr Zool* **8**: 63–73.

Bialozyt R, Luettmann K, Michalczyk IM, Pinedo Saboya PP, Ziegenhagen B, Heymann EW (2014). Primate seed dispersal leaves spatial genetic imprint throughout subsequent life stages of the Neotropical tree *Parkia panurensis*. *Trees* **28**: 1569–1575.

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.

Carvalho CS, Ribeiro MC, Côrtes MC, Galetti M, Collevatti RG (2015). Contemporary and historic factors influence differently genetic differentiation and diversity in a tropical palm. *Heredity* **115**: 216–224.

Colwell RK (2013). *EstimateS: Statistical Estimation of Species Richness and Shared Species from Samples*. Version 9. User's Guide and application published at http://purl.oclc.org/estimates.

Cruz MP, Paviolo A, Bó R, Thompson J, Di Bitetti MS (2014). Daily activity patterns and habitat use of the lowland tapir (*Tapirus terrestris*) in the Atlantic Forest. *Mamm Biol* **79**: 376–383.

De-Lucas AI, González-Martínez SC, Vendramin GG, Hidalgo E, Heuertz M (2009). Spatial genetic structure in continuous and fragmented populations of *Pinus pinaster* Aiton. *Mol Ecol* **18**: 4564–4576.

Di Bitetti MS, Placci LG, Dietz LA (2003). *A Biodiversity Vision for the Upper Paraná Atlantic Forest Eco-region: Designing a Biodiversity Conservation Landscape and Setting Priorities for Conservation Action*. WWF: Washington, USA.

Dick CW, Etchelecu G, Austerlitz F (2003). 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.

Dirzo R, Young HS, Galetti M, Ceballos G (2014). Defaunation in the Anthropocene. *Science* **345**: 401–406.

Eichert W (1999). *WHICHPARENTS (version 1.0)*: a windows application for determining the most likely parents of offspring using multilocus genotype data. Computer software distributed from http://bml.ucdavis.edu/research/research-programs/ecology-evolution-conservation/salmon-research/salmon-genetics-software/.

Ellstrand NC, Elam DR (1993). Consequences of small population size: implications for plant conservation. *Annu Rev Ecol Syst* **24**: 217–242.

Epperson BK (2005). Estimating dispersal from short distance spatial autocorrelation. *Heredity* **95**: 7–15.

Fragoso JMV, Huffman JM (2000). Seed-dispersal and seedling recruitment patterns by the last Neotropical megafaunal element in Amazonia, the tapir. *J Trop Ecol* **16**: 369–385.

Galetti M, Bovendorp RS, Guevara R (2015). Defaunation of large mammals leads to an increase in seed predation in the Atlantic forests. *Glob Ecol Conserv* **3**: 824–830.

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

Gehring J, Delph L (1999). Fine-scale genetic structure and clinal variation in *Silene acaulis* despite high gene flow. *Heredity* **82**: 628–637.

Genini J, Galetti M, Morellato LPC (2009). Fruiting phenology of palms and trees in an Atlantic rainforest land-bridge island. *Flora Morphol Distrib Funct Ecol Plants* **204**: 131–145.

Giombini MI, Bravo SP, Martínez MF (2009). Seed dispersal of the palm *Syagrus romanzoffiana* by tapirs in the semi-deciduous Atlantic Forest of Argentina. *Biotropica* **41**: 408–413.

Giombini MI, Bravo SP, Tosto DS (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.

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

Guix JC, Ruiz X (2000). Plant-disperser-pest evolutionary triads: how widespread are they? *Orsis* **15**: 121–126.

Hamrick JL, Murawski DA, Nason JD (1993). The influence of seed dispersal mechanisms on the genetic structure of tropical tree populations. *Vegetatio* **107/108**: 281–297.

Hanson TR, Brunsfeld SJ, Finegan B, Waits LP (2008). Pollen dispersal and genetic structure of the tropical tree *Dipteryx panamensis* in a fragmented Costa Rican landscape. *Mol Ecol* **17**: 2060–2073.

Hardy OJ (2003). Estimation of pairwise relatedness between individuals and characterization of isolation-by-distance processes using dominant genetic markers. *Mol Ecol* **12**: 1577–1588.

Hardy OJ, Vekemans X (2002). SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Mol Ecol Notes* **22**: 618–620.

Husband BC, Schemske DW (1996). Evolution of the magnitude and timing of inbreeding depression in plants. *Evolution* **50**: 54–70.

Jones OR, Wang J (2010). COLONY: a program for parentage and sibship inference from multilocus genotype data. *Mol Ecol Resour* **10**: 551–555.

Jump AS, Peñuelas J (2005). Running to stand still: adaptation and the response of plants to rapid climate change. *Ecol Lett* **8**: 1010–1020.

Jump AS, Peñuelas J (2006). Genetic effects of chronic habitat fragmentation in a wind-pollinated tree. *Proc Natl Acad Sci USA* **103**: 8096–8100.

Kalinowski ST (2004). Counting alleles with rarefaction: private alleles and hierarchical sampling designs. *Conserv Genet* **5**: 539–543.

Kalinowski ST, Wagner AP, Taper ML (2006). ML-RELATE: a computer program for maximum likelihood estimation of relatedness and relationship. *Mol Ecol Notes* **6**: 576–579.

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 VL, Roorda T, Durães R, Smith TB (2010). Destination-based seed dispersal homogenizes genetic structure of a tropical palm. *Mol Ecol* **19**: 1745–1753.

Keuroghlian A, Eaton DP (2008). Fruit availability and peccary frugivory in an isolated Atlantic Forest fragment: effects on peccary ranging behavior and habitat use. *Biotropica* **40**: 62–70.

Kramer AT, Ison JL, Ashley MV, Howe HF (2008). The paradox of forest fragmentation genetics. *Conserv Biol* **22**: 878–885.

Kurten EL (2013). Cascading effects of contemporaneous defaunation on tropical forest communities. *Biol Conserv* **163**: 22–32.

Laurance WF (2010). Habitat destruction: death by a thousand cuts. In: Sodhi NS, Ehrlich PR (eds), *Conservation Biology for All*. Oxford University Press: Oxford, UK, pp 73–87.

Loiselle BA, Sork VL, Nason J, Graham C (1995). Spatial genetic structure of a tropical understorey shrub, *Psychotria officinalis* (Rubiaceae). *Am J Bot* **82**: 1420–1425.

Loveless MD, Hamrick JL (1984). Ecological determinants of genetic structure in plant populations. *Annu Rev Ecol Syst* **15**: 65–95.

Lowe AJ, Boshier D, Ward M, Bacles CFE, Navarro C (2005). Genetic resource impacts of habitat loss and degradation; reconciling empirical evidence and predicted theory for neotropical trees. *Heredity* **95**: 255–273.

Mendes CP, Ribeiro MC, Galetti M (2016). Patch size, shape and edge distance influence seed predation on a palm species in the Atlantic forest. *Ecography* **39**: 465–475.

Nathan R, Muller-Landau H (2000). Spatial patterns of seed dispersal, their determinants and consequences for recruitment. *Trends Ecol Evol* **15**: 278–285.

10

Nazareno AG, Dos Reis MS (2012). Linking phenology to mating system: exploring the reproductive biology of the threatened palm species *Butia eriospatha*. *J Hered* **103**: 842–852.

Nei M (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* **89**: 583–590.

Neuschulz EL, Mueller T, Schleuning M, Böhning-Gaese K (2016). Pollination and seed dispersal are the most threatened processes of plant regeneration. *Sci Rep* **6**: 29839.

Olmos F, Pardini R, Boulhosa RLP, Bürgi R, Morsello C (1999). Do tapirs steal food from palm seed predators or give them a lift? *Biotropica* **31**: 375–379.

Pacheco LF, Simonetti JA (2000). Genetic structure of a mimosoid tree deprived of its seed disperser, the spider monkey. *Conserv Biol* **14**: 1766–1775.

Peakall R, Smouse P (2012). GenAlEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research–an update. *Bioinformatics* **1**: 1–3.

Peres CA (2001). Synergistic effects of subsistence hunting and habitat fragmentation on Amazonian forest vertebrates. *Conserv Biol* **15**: 1490–1505.

Scofield DG, Alfaro VR, Sork VL, Grivet D, Martinez E, Papp J *et al.* (2011). Foraging patterns of acorn woodpeckers (*Melanerpes formicivorus*) on valley oak (*Quercus lobata* Née) in two California oak savanna-woodlands. *Oecologia* **166**: 187–196.

Sica YV, Bravo SP, Giombini MI (2014). Spatial patterns of pindó palm (*Syagrus romanzoffiana*) recruitment in Argentinian Atlantic Forest: the importance of tapir and effects of defaunation. *Biotropica* **46**: 696–703.

Smouse PE, Peakall R, Gonzales E (2008). A heterogeneity test for fine-scale genetic structure. *Mol Ecol* **17**: 3389–3400.

Smouse PE, Robledo-Arnuncio JJ (2005). Measuring the genetic structure of the pollen pool as the probability of paternal identity. *Heredity* **94**: 640–649.

Sokal RR, Rohlf FJ (1995). *Biometry*. 3rd edn. WH Freeman and Company: New York, NY, USA.

Sork VL, Smouse PE (2006). Genetic analysis of landscape connectivity in tree populations. *Landsc Ecol* **21**: 821–836.

StatSoft, Inc. (2001). STATISTICA (data analysis software system). Version 6. Available at: www.statsoft.com.

Stoner KE, Riba-Hernández P, Vulinec K, Lambert JE (2007). The role of mammals in creating and modifying seedshadows in tropical forests and some possible consequences of their elimination. *Biotropica* **39**: 316–327.

Vekemans X, Hardy OJ (2004). New insights from fine-scale spatial genetic structure analyses in plant populations. *Mol Ecol* **13**: 921–935.

Wagner AP, Creel S, Kalinowski ST (2006). Estimating relatedness and relationships using microsatellite loci with null alleles. *Heredity* **97**: 336–345.

Wang BC, Sork VL, Leong MT, Smith TB (2007). Hunting of mammals reduces seed removal and dispersal of the afrotropical tree *Antrocaryon klaineanum* (Anacardiaceae). *Biotropica* **39**: 340–347.

Wang J (2010). Effects of genotyping errors on parentage exclusion analysis. *Mol Ecol* **19**: 5061–5078.

Wang J (2012). Computationally efficient sibship and parentage assignment from multi-locus marker data. *Genetics* **191**: 183–194.

Wang J, El-Kassaby YA, Ritland K (2012). Estimating selfing rates from reconstructed pedigrees using multilocus genotype data. *Mol Ecol* **21**: 100–116.

Wang J, Santure AW (2009). Parentage and sibship inference from multilocus genotype data under polygamy. *Genetics* **181**: 1579–1594.

Wang R, Compton SG, Chen X-Y (2011). Fragmentation can increase spatial genetic structure without decreasing pollen-mediated gene flow in a wind-pollinated tree. *Mol Ecol* **20**: 4421–4432.

Ward M, Dick CW, Gribel R, Lowe AJ (2005). To self, or not to self... A review of outcrossing and pollen-mediated gene flow in neotropical trees. *Heredity* **95**: 246–254.

White GM, Boshier DH, Powell W (2002). Increased pollen flow counteracts fragmentation in a tropical dry forest: An example from *Swietenia humilis* Zuccarini. *Proc Natl Acad Sci USA* **99**: 2038–2042.

Young AG, Boyle T, Brown T (1996). The population genetic consequences of habitat fragmentation for plants. *Trends Ecol Evol* **11**: 413–418.

Zeng X, Michalski SG, Fischer M, Durka W (2012). Species diversity and population density affect genetic structure and gene dispersal in a subtropical understory shrub. *J Plant Ecol* **5**: 270–278.