BMC Genetics

Open Access

Spatial genetic structure, genetic diversity and pollen dispersal in a harvested population of *Astrocaryum aculeatum* in the Brazilian Amazon

Santiago Linorio Ferreyra Ramos[1*], Gabriel Dequigiovanni[1], Alexandre Magno Sebbenn[2], Maria Teresa Gomes Lopes[3], Paulo Yoshio Kageyama[4], Jeferson Luis Vasconcelos de Macêdo[5], Matias Kirst[6] and Elizabeth Ann Veasey[1*]

Abstract

Background: *Astrocaryum aculeatum* is a palm tree species native to the tropical regions of South America, exploited commercially by local farmers for the pulp extracted from its fruits. The objective of this research was to compare the genetic diversity between adult plants and seedlings from open-pollinated seeds, quantify the pollen flow and dispersal, the spatial genetic structure, and the effective size of a population that has been continuously harvested for its fruits. The study was carried out in a natural population of *A. aculeatum* distributed over approximately 8 ha in the State of Amazonas (Brazil), separated by 400 m from the closest neighboring population. In total, 112 potential pollen donors, 12 mother plants and 120 offspring were mapped and genotyped.

Results: Genetic diversity was high for parents and the offspring. The fixation indexes for adults ($F = -0.035$) and offspring ($F = -0.060$) were negative and not significant. A significant spatial genetic structure was detected for the adult plants (up to the distance of 45 m) indicating short-distance seed dispersal. Paternity analysis detected 9.2 % of pollen immigration and the average distance of pollination within the population was 81 m. The average effective pollination neighborhood area between plants was 1.51 ha.

Conclusions: Our results indicate that substantial introduction of new alleles has occurred in the population through pollen immigration, contributing to the maintenance of genetic diversity. Conservation efforts aimed at maintaining the gene pool of the current population or establishing new populations should utilize offspring from mother plants selected to be spaced by at least 50 m to prevent collecting seeds from relatives.

Keywords: Coancestry coefficient, Effective size, Gene flow, Microsatellites, Paternity analysis, Population genetics, Genetic structure, Tucumã of amazonas

* Correspondence: slfr03@hotmail.com; eaveasey@usp.br
[1]Escola Superior de Agricultura "Luiz de Queiroz"/Universidade de São Paulo (ESALQ/USP), Departamento de Genêtica, C.P. 9, CEP 13418-900 Av. Pádua Dias, 11, Piracicaba, SP, Brazil
Full list of author information is available at the end of the article

Ramos *et al. BMC Genetics* (2016) 17:63

Page 2 of 10

Background

Palm trees are considered one of the most useful groups of plants, especially for rural communities to which they provide building material, fabric, fuel, food, ornamental and medicinal plants [1]. The palm tree *Astrocaryum aculeatum* (Arecaceae) is found in the tropical forest of the Brazilian Amazon region. It is a monoecious species, morphologically showing an inflorescence that has the two floral structures (male and female) and that presents protogyny. Popularly known as tucumã of Amazonas, it is of significant economical importance for the populations of the Amazon region [2]. The pulp from the fresh fruit is directly consumed and also used to extract oil for the cosmetic industry and human consumption. The seed endocarp is used to make crafts [3, 4]. The species is adapted to non-flooded areas and is more frequently found in deforested areas or in areas that underwent anthropic action [3]. The distribution of *A. aculeatum* is restricted to the Western and Central Brazilian Amazon region, more specifically in the states of Acre, Mato Grosso, Rondônia, Roraima, part of Pará [4] and Amazonas. Amazonas is probably its most important center of genetic diversity [5]. Commercial harvest of *A. aculeatum* fruit is almost entirely dependent on extractivism. Consequently, highly irregular amounts of fruits of variable quality are commonly harvested. The growing market demand and the high price of the fruit has motivated farmers in the state of Amazonas to increasingly exploit commercially these plants, but the genetic consequences on natural populations has been unknown. Highly informative co-dominant microsatellite molecular markers [6] can be applied to measure the impact of current harvesting methods on natural populations of *A. aculeatum*. In particular, these markers can be used to measure the degree of genetic variation within and between populations, supplying important answers in population genetics and in ecological and evolutionary approaches [7], as well as estimating genetic diversity, endogamy, spatial population structure, mating system, and gene flow [8].

Gene flow in trees comprises both pollen and seeds dispersal [9, 10]. The male gametes are pollen-dispersed and the embryos carrying contributions from both parents are seed-dispersed [11]. Pollen and seed dispersal patterns significantly influence the genetic structure and the effective population size [12]. High levels of gene flow allow the maintenance of the genetic cohesion between populations whereas low levels result in genetic differentiation through genetic drift and local selection [13]. Studies based on genetic markers have shown that pollen can be dispersed over long distances in animal pollinated tree species [14, 15], preventing possible effects of genetic drift, such as loss of alleles, reduction in heterozygosity, increase in inbreeding and decrease in effective population size within populations [16].

Here we present a population genetics study of a harvested natural population of *A. aculeatum* established in the Brazilian Amazon region. The objectives of this study were to quantify the genetic diversity in adult individuals and open-pollinated seeds, quantify the pollen flow and dispersal, the spatial genetic structure, and the effective population size of this population. Estimating the effective number of pollen donors and the effective size of a natural population defines the minimum distance required between seed trees to avoid genetic relatedness in seed collections. This information is critical for conservation and tree breeding programs [14]. The associated parentage analysis methods [17] determined through categorical likelihood paternity analysis [18] is also important for allowing the genealogical reconstruction of the relatedness between individuals within and among progeny [14]. In summary, estimating parameters of genetic diversity, reproduction system, gene flow, pollen dispersal, and spatial genetic structure in harvested populations of *A. aculeatum* are fundamental for the establishment of adequate strategies for the use of this genetic resource.

Results

Genetic diversity

A set of 12 microsatellite loci was used in this study, of which two (Aac01 and Aac13) were monomorphic in the population and excluded from further analysis. A total of 81 alleles were identified in the ten polymorphic loci, distributed among all 244 sampled plants (adults + offspring). In adults, the total number of alleles per locus varied from 3 to 19, with a mean of 7.7. In offspring, the total number of allele was lower, varying from 2 to 14, with a mean of 6.4 (Table 1). Adults had also more private alleles (17) than offspring (4), suggesting low pollen immigration from other areas or recent population establishment. The observed and expected heterozygosity of the adults ($H_o = 0.566$; $H_e = 0.547$) were marginally lower than the offspring ($H_o = 0.594$; $H_e = 0.560$). The mean fixation index (F) was negative and not significantly different from zero for adults (-0.035) and offspring (-0.060), indicating absence of inbreeding.

Spatial genetic structure

The spatial genetic structure was significant up to 45 m (Fig. 1), suggesting that near neighbor adult plants are relatives. Above this distance the θ_{ij} values were not significantly different from zero or significantly lower than zero. The slope of the b_k regression of the pairwise coancestry coefficient over the logarithm of spatial distance scale (0 – 351 m) was significantly negative ($b_k = -0.014$), showing a seed dispersal pattern of isolation by distance. The S_p statistic was 0.0139.

Ramos *et al. BMC Genetics* (2016) 17:63

Page 3 of 10

Table 1 Genetic diversity indexes[a] for adults and offspring in an *Astrocaryum aculeatum* population

Sample	n	k	A_p	A	H_e	H_o	F	P_{2p}
Adults	122.7	77	17	7.7 ns	0.547 ns	0.566 ns	−0.035 ns	0.996080
Offspring	118	64	4	6.4 ns	0.560 ns	0.594 ns	−0.060 ns	-

[a]n sample size, k total number of alleles over loci, A_p number of private alleles in each generation, A is the average number of alleles per loci, H_e and H_o expected and observed heterozygosities, respectively, F fixation index, P_{2p} theoretical combined non exclusion probability for second parent. [b]*ns* non significant by the Student's t- test with p (probability). A, $p = 0.543$; H_e, $p = 0.885$; H_o, $p = 0.809$; F, $p = 0.894$

Effective population size

The group coancestry coefficient (Θ) for adults was 0.023, suggesting that the expected rate of inbreeding by mating among relatives is very low (<3 %). The estimated effective population size indicates that the 124 adult plants correspond to 22 (N_e) unrelated and non-inbred individuals.

Pollen flow, *dispersal patterns, and dispersal kernel estimation*

The combined probability of exclusion of the second parent (Table 1) was high ($P_{2p} = 0.99608$) in the paternity analyses. This indicates that pollen flow may have been overestimated because the higher number of pollen donors has been attributed for seeds using the delta statistic, and some true fathers may not have been assigned due to restrictions of this statistic. Among the sample of 120 offspring, 109 (90.8 %) were identified as being sired by an individual from within the sampled population. The other 11 offspring were probably sired by pollen donors from other populations, suggesting a 9.2 % of pollen immigration rate (Table 2). The 109 offspring were apparently generated by 56.4 % of the adult plants

(70 of the 124 reproductive trees). No selfed offspring were detected, indicating a zero selfing rate.

The pollen dispersal distance (δ) ranged from 3 to 194 m, with a mean of 81 m (standard deviation of 49 m) and a median of 70 m (Fig. 2a). In the study area, the estimate of the correlation coefficient among the number of offspring fertilized by pollen of male genitors (adult plants genotyped in this study) and the distance between mother plants was high and significantly different than zero ($R^2 = 0.74$, $p < 0.05$). This suggests that the distance between plants had an impact on the mating probability. A Kolmogorov-Smirnov test was not significant (D = 0.064, $p = 0.82$; Fig. 2a), indicating that the spatial distance between trees explains the observed pattern of pollen dispersal. The mean effective pollination neighborhood area was 1.51 ha and the average effective pollination ratio of pollen dispersal was 69 m.

Pollen dispersal kernel for *A. aculeatum* was a slightly fat-tailed dispersal (Fig. 2b), indicating a high probability of long-distance dispersal of pollen (immigration). We also found absence of selfing (s = 0.0), pollen immigration rate of 0.098, mean distance of
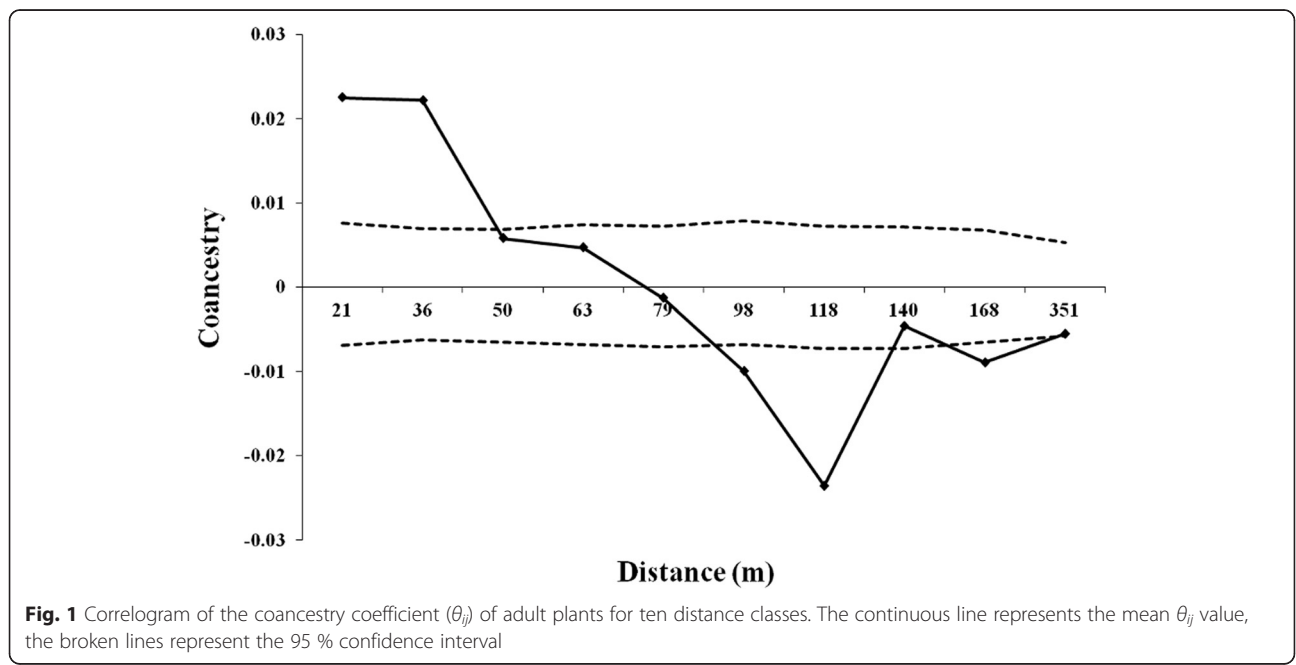

Fig. 1 Correlogram of the coancestry coefficient (θ_{ij}) of adult plants for ten distance classes. The continuous line represents the mean θ_{ij} value, the broken lines represent the 95 % confidence interval

Ramos *et al. BMC Genetics* (2016) 17:63

Page 4 of 10

Table 2 Pollen flow and dispersal in an *Astrocaryum aculeatum* population

Mother plants	M 04	M 05	M 06	M 07	M 15	M 37	M 40	M 41	M 45	M 47	M 48	M 49	Total
Np	15	2	7	14	10	8	8	15	13	2	14	12	120
Nm	1	-	3	-	2	1	1	-	-	-	1	2	11
m_p (%)	6.7	0	42.8	0	20	12.5	12.5	0	0	0	7.1	16.7	9.2

Np number of offspring genotyped per mother plants, *Nm* number of offspring with pollen donors located outside of the population, m_p percent of pollen flow

pollen dispersal of 90 m, dispersal kernel scale (a = 20.0739 m) and a slightly fat-tailed shape (b = 0.9).
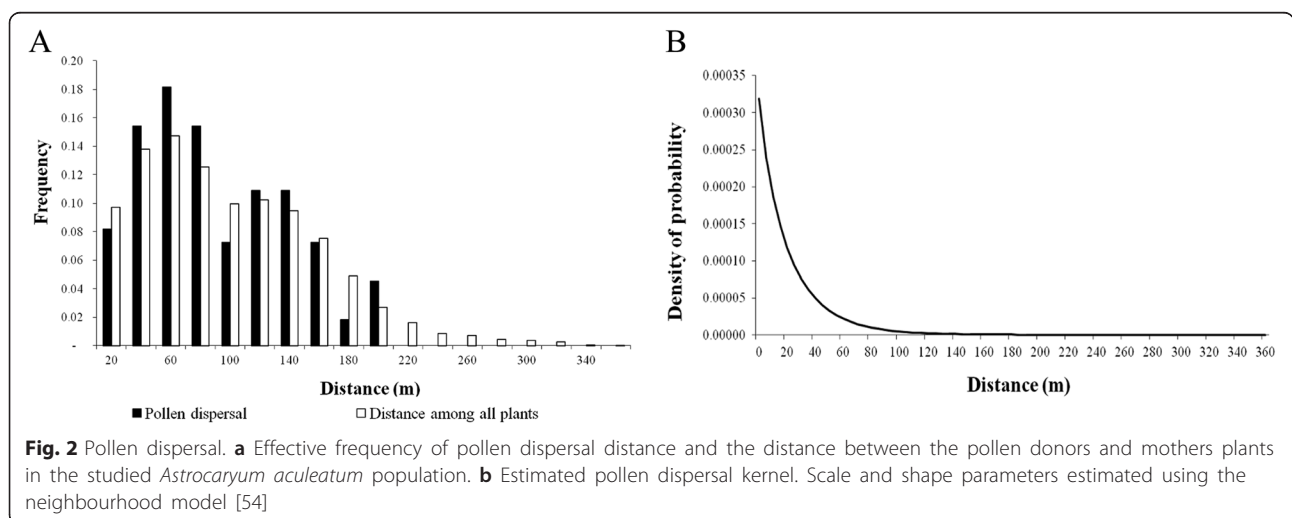
Discussion

Genetic diversity

This is the first population genetic study in *A. aculeatum*, aimed at understanding gene flow in a natural population exploited through extractivism. The results show that this harvest practice has not had a negative effect on the genetic diversity of *A. aculeatum* offspring. By examining the mean number of alleles per locus (*A*), observed (*H$_o$*) and expected (*H$_e$*) heterozygosities, it was concluded that the levels of genetic diversity are similar between adults and offspring.

The great majority of the private alleles were found in the adults, suggesting that genetic drift may have occurred during the reproductive events, since not all of the alleles observed in the adults were transmitted to the offspring. Genetic drift is the random change in allele frequency that occurs because gametes transmitted from one generation to the next carry only a sample of the alleles present in the parental generation, and is more significant in small populations. In this species, female flower anthesis takes place in the evening and flowers remain viable for 24 h. Male flowers start their anthesis after female flowers and remain viable for only six hours [19]. *A. aculeatum* is a monoecious species with protogyny, without overlapping of sexual phases

[19] and, consequently, expected to be predominantly outcrossing [20]. This was confirmed by the paternity analysis, which did not identify any instance of self-fertilization.

Spatial genetic structure

The spatial genetic structure detected among the adults suggests a seed dispersal pattern of isolation by distance. Thus, near neighbor plants are probably relatives. Spatial genetic structure results from to the seed and pollen dispersal near the mother plant. The seeds from *A. aculeatum* are primarily dispersed by gravity, usually concentrated near the canopy projection of up to 3.5 m [4]. Secondary dispersal by rodents, such as *Dasyprocta azarae* and *Myoprocta* sp., places the seeds close to the mother plants [21]. This may explain the observed spatial genetic structure. The results also showed relatively low levels of coancestry in the two first distance classes (0 – 21 and 21 – 36 m), these being expected values between first degree cousins ($\theta_{ij} = 0.0625$). Comparing the S_p-statistic with other species, the observed value (0.0139) is similar to those identified in populations with a high density of individuals, such as *Dicorynia guianensis* ($S_p = 0.026$), *Vouacapoua americana* ($S_p = 0.012$) [22], *Sanicula odorata* ($S_p = 0.0181$), and *Silene acaulis* ($S_p = 0.0144$) [23], showing that density is one of the most important factors determining the spatial genetic structure.

Fig. 2 Pollen dispersal. **a** Effective frequency of pollen dispersal distance and the distance between the pollen donors and mothers plants in the studied *Astrocaryum aculeatum* population. **b** Estimated pollen dispersal kernel. Scale and shape parameters estimated using the neighbourhood model [54]

Ramos *et al. BMC Genetics* (2016) 17:63

Page 5 of 10

Effective population size

The adult coancestry group ($\Theta = 0.023$) suggests that under random mating a low level of inbreeding is expected (< 3 %) [24], in agreement with the estimated inbreeding in both adults and offspring. Due to the low Θ, the effective population size ($N_e = 22$) was low [25], showing a high proportion of related individuals within the population.

The cause of the high proportion of relatedness within the population is likely due to short distance seed dispersal, as shown by the analysis of the spatial genetic structure. The pollen immigration rate (9.2 %), however, may in the future increase genetic diversity and effective size and so offset the negative effects of genetic drift.

Pollen flow and dispersal patterns

Pollen flow from outside of the population was moderate (9.2 %). This gene flow was similar to the ones observed in other tropical tree species of wind and animal pollinated populations, which are isolated by a distance superior to 1 km. For instance, a pollen immigration rate of 10 % was observed in a population of *Araucaria angustifolia* distributed in an area of 5.4 ha and isolated from other populations by more than 1.7 km [13]. In a 4.8 ha forest fragment isolated from the nearest individual by a distance of 1.2 km, pollen migration in *Copaifera langsdorffii* was estimated as 5 to 8 % [14, 24]. In other studies with non isolated populations or localized at distances lower than 1 km, the immigration rates were higher: 49 % in *Symphonia globulifera* [26], 61.3 % in *Theobroma cacao* [27] and > 38 % in *Hymenaea courbaril* [28]. Pollen immigration increases the genetic diversity and effective population size, due to the introduction of new alleles [24]. Thus, the result suggests that pollen immigration contributed moderately to the preservation of the genetic diversity in the *A. aculeatum* population.

Pollen dispersal reached long distances within the population (194 m), considering that the maximum distance between two plants was 211 m. However, pollen dispersal followed an isolation by distance pattern (Fig. 2a), shown by the high correlation ($R^2 = 0.74$) between the number of seeds fertilized by pollen donors located close to the mother plants. There are no previous studies on the pollinators of *A. aculeatum*. However, floral morphology and phenology are similar to the congener *Astrocaryum vulgare* [29], suggesting that *A. aculeatum* has similar pollinators. *A. vulgare* is pollinated by bees of the species *Trigona sp.* and *Apis mellifera* (Hymenoptera) and Coleoptera belonging to the families Nitidulidae (*Mystrops* sp.) and Curculionidae (*Terires minusculus*) [29]. We observed the species *Trigona* sp. and *A. mellifera* visiting flower buds in pre-anthesis and anthesis of male flowers (floral damaging these parts in search of pollen).

In addition, Coleoptera (*Mystrops* sp. and *Terires minusculus*) insects were observed both in male and female flowers of *A. aculeatum*. These insects have potential for long-distance pollen dispersal [30]. This could explain the observed high effective pollination area (1.51 ha).

The mean pollen dispersal distance of *A. aculeatum* (81 m) by insects pollinators was similar to that detected in high-density populations (> 5 tree/ha) [15]. In general, this distance is lower than 100 m. Pollen dispersal distance for high-density populations rarely exceeds 300 m in forests [15]. For example, in *Astrocaryum mexicanum* the average varies between 13 to 23 m [31], while in *Oenocarpus bataua*, a species that occurs in low-density populations, the estimated distances ranges from 113 to 1263 m [32]. In tree species from other families, pollination dispersal distances range from 28 m in *Theobroma cacao* [27], 65 m in *Himatathus drasticus* [33] and 66 to 94 m in *Copaifera langsdorffii* [14, 24, 34]. Our results showed that a higher frequency of short pollen dispersal distance relative to long pollen dispersal distance suggests that *A. aculeatum* is primarily an animal-pollinated species [15].

The exponential power distribution of inter-mate distances within our study plot had a similar behavior of a slightly leptokurtic shape (b = 0.9), indicating some long-distance mating events. Similarly, the dispersal kernel inferred for *A. aculeatum* using the spatially-explicit Neighborhood model showed a distribution with high probability of dispersal at low-distances. However, a very rapid decline in dispersal probability with increasing distance was observed. Similar long-distance dispersal events at pollination and overall pollen movement have been observed in recent studies (for example, with *Araucaria angustifolia* [13] and *Phoenix canariensis* [35]), suggesting a pattern in pollination events that may not be uncommon, especially amongst tropical species, because it showed a non-leptokurtic pattern [32].

The results observed in *A. aculeatum* would indicate that the genetic data used for the underlying model of dispersal, which typically assumes a long tail of dispersal [32] are in agreement to the model. The shape of the tail of the dispersal kernel (that is, whether thin- or fat-tailed) impacts the ultimate distribution of genetic diversity within and between populations. Most studies report that pollen dispersal kernels are fat-tailed in tree species [35–38]. This shows that the slightly fat-tailed dispersal distribution found in the sampled *A. aculeatum* is related to a few long distance dispersal events, due to the relatively low pollen immigration detected (9.8 %).

Implications for conservation, cultivation of trees and seed harvesting

Our results have important implications for seed harvesting strategies for *ex situ* conservation and commercial

Ramos *et al. BMC Genetics* (2016) 17:63

Page 6 of 10

reforestation of *A. aculeatum*. The presence of spatial genetic structure within the population indicates that seeds should be harvested from mother plants located at distances larger than 45 m from each other. This practice will limit the probability of collecting seeds from related plants, which would reduce the effective size of collected progeny. The inclusion of offspring from different and non-related mother plants in germplasm banks increases the effective size of the retained population [13]. However, considering that spatial genetic structure occurs in the studied population and 50 % of the pollen is dispersed below 70 m (median pollen dispersal distance), we can expect some levels of inbreeding in the sampled progeny arrays, originated from mating among relatives [13]. Some inbreeding may be eliminated in nurseries, by excluding seedlings with low growth and poorly developed. Therefore, the estimation of genetic parameters in progeny tests with seeds from this population must be adjusted to accommodate inbreeding.

Conclusions

The studied population presents high levels of genetic diversity in both adults and open-pollinated seed, due to pollen flow. However, the effective size of the adult population is low due to the presence of spatial genetic structure caused by short seed dispersal distances. Also, pollen dispersal follows a model of isolation by distance. Consequently, most of the mating occurred among neighboring plants and open-pollinated seeds probably present some levels of inbreeding. It will be critical to repeat this study in other populations and/or the next reproductive events of the mother plants to ascertain the observed results.

Methods

Study area

This study was undertaken in a natural population of *A. aculeatum*, located in the lot number 12 of the Manaus Agrarian Project, in a rural property named "Natajuba" (latitude -02°53'27.9"S and longitude -60°06'08.2"W) in the state of Amazonas, Brazil (Fig. 3). The region is characterized by a tropical forest climate (Af type), according to the Koppen-Geiger world map of climate classification [39]. The population is bordered to the East by the hydrographic basin of the Tarumã-Açu river, and to the West by a native forest that is part of the legal reserve of the "Natajuba" property and other properties within the Manaus Agrarian Project. North and South boarders are delimited by the rivers "Cuieiras" and "Jacaré", respectively. Each side of the streams has a riparian forest of width larger than 200 m. Two other natural populations of *A. aculeatum* are located 400 m from the northern and more than 1,000 m from the southern boarders of the population studied here.
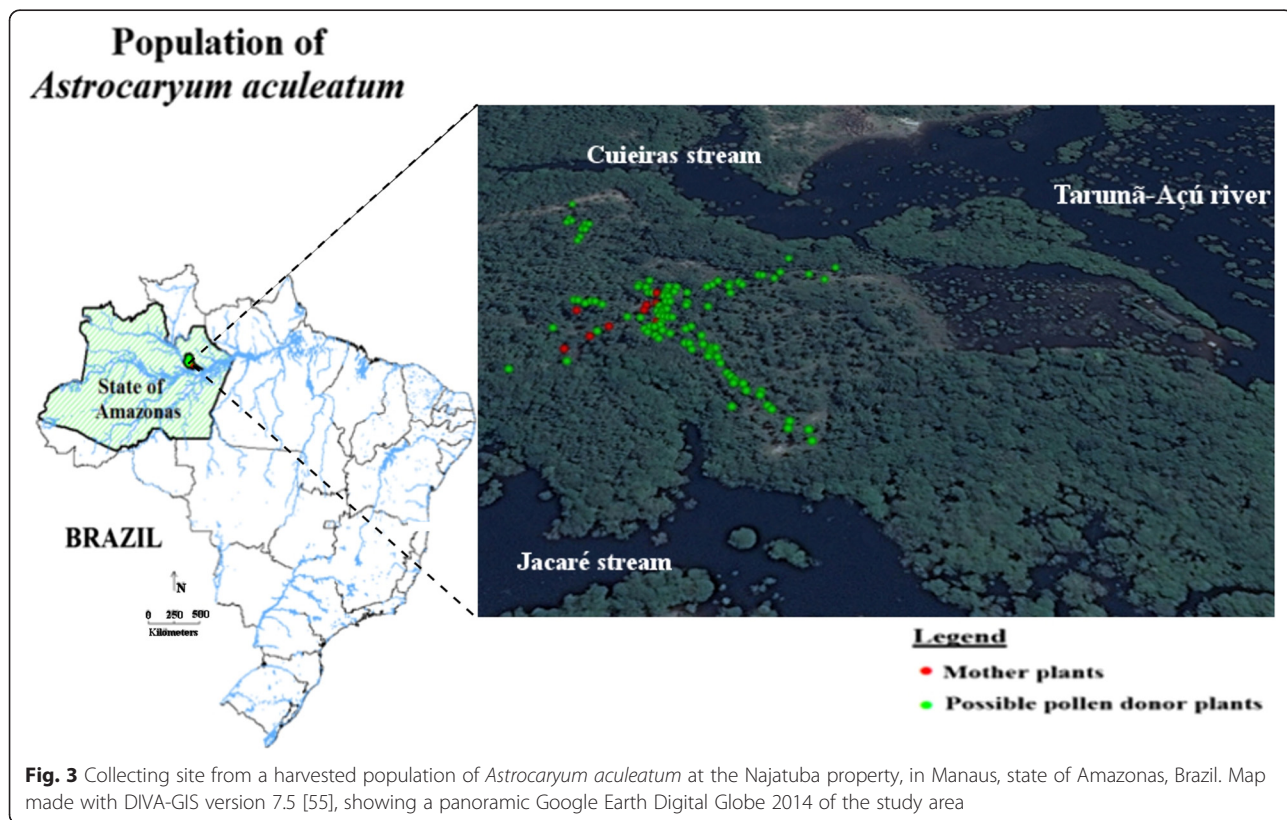

Fig. 3 Collecting site from a harvested population of *Astrocaryum aculeatum* at the Najatuba property, in Manaus, state of Amazonas, Brazil. Map made with DIVA-GIS version 7.5 [55], showing a panoramic Google Earth Digital Globe 2014 of the study area

Ramos *et al. BMC Genetics* (2016) 17:63

Page 7 of 10

Additional *A. aculeatum* populations are located at distances higher than 3,000 m. The population used in this study has been exploited continuously since 1996 to provide *A. aculeatum* fruits to the Manaus market. This population is also part of the *in situ* conservation of the superior germplams program from Embrapa Western Amazonia, within the project "Research, development, and innovation in oil producing palm plants and the economical use of by-products and residues", PROPALMA (Embrapa-Propalma).

Sampling

In March 2011, 12 mother plants of *A. aculeatum* with maturing fruits generated by open-pollination were identified in the population area [4]. Twenty-five fruits were collected from each mother plant and placed within properly identified polyethylene bags and taken to the seed laboratory of the Western Amazonia Embrapa, in Manaus (Brazil). The pulp of each fruit was removed and the seeds obtained were dried at a temperature of 30 °C up to the point where the moisture content was 14.5 %, allowing the separation of the seed tegument by mechanical breakage [3]. Soaking and germination processes were immediately carried out [3]. The offspring germination phase was conducted during three months in the greenhouse. A total of 120 offspring from 12 mother plants of *A. aculeatum* was obtained. The number of offspring obtained per plant varied from 2 to 15. The distance between the mother plants in the area ranged from 6.2 to 83.2 m, with a mean of 37.7 m and a median of 37.1 m. Based on the geographic distribution of the mother plants, the center point between them was identified. From this central point four transects were established to conduct sampling of adult plants that are potential parents (pollen donors) of the offspring obtained from the mother plants. The end of each transect was established when there were no more plants within a distance of 100 m. Each transect was 10 m wide. These transects were extended to the Northeast, Southeast, Northwest and Southwest directions. In the four transects, 112 possible pollen donor plants were sampled. The distance between possible pollen donors varied from 0.5 to 351.4 m, with a mean of 89.6 m and a median of 78.9 m. The geographic position of each adult plant (pollen donor and mother plant) was determined by using a global positioning system (GPSmap 60CSx - GARMIN).

Leaf samples from 120 offspring, 12 mother plants and 112 potentials parent plants (pollen donors), were collected and stored in silica gel at -20 °C in the Laboratory of Molecular Biology at the National Institute of Amazonian Research (LTBM-INPA). Total DNA was extracted according to the cationic detergent protocol CTABLE 2× (Cationic Hexadecyl Trimethyl Ammonium Bromide) [40] and quantification was performed according to Ramos et al. [20].

This research was supported by the norms of Resolution 21 - 31 August 2006 - of CGEN (*Conselho de Gestão Do Patrimônio Genético - Ministério do Meio Ambiente*) (http://www.mma.gov.br/estruturas/sbf_dpg/_arquivos/res21cons.pdf). Material collection was registered in SISBIO (*Sistema de Autorização e Informação em Biodiversidade/Instituto Chico Mendes de Conservação da Biodiversidade - ICMBio/Ministério do Meio Ambiente - MMA*), voucher number 39950-4. A sample of the taxon *Astrocaryum aculeatum* G. Meyer (tucumã-do-Amazonas) was deposited in the INPA herbarium under No. 246369.

Microsatellites amplification

In this study, 12 microsatellite loci developed for *A. aculeatum* (Aac01, Aac02, Aac03, Aac04, Aac06, Aac07, Aac09, Aac10, Aac11, Aac12, Aac13 and Aac14) were used [41]. These microsatellites were amplified by polymerase chain reaction (PCR) using the Veriti Thermal Cycler (Applied Biosystems) in a total reaction volume of 10 μL, containing 10 ng genomic DNA, 1× buffer (10× standard *Taq* reaction buffer), 210 μM of each dNTP, 1.5 mM $MgCl_2$, 0.16 μM of forward and M13 labeled primers (FAM or NED dyes) [42], 0.32 μM of reverse primers, 1.05 U *Taq* DNA polymerase (Invitrogen), and 3.49 μL of ultra pure water. The amplifications *via* PCR occurred in two phases, the first being specific for the primers and the second to connect the M13. The first stage began by stabilizing the temperature at 68 °C for 2 min and at 92 °C for 30 s, followed by 30 cycles (30 s at 92 °C for denaturation process, 35 s at the primer-specific annealing temperature {Table 1 of [41]}, and 30 s at 68 °C {72 °C for Aac07 and Aac11} for extension); the second step consisted of 15 cycles (30 s at 92 °C, 30 s at 53 °C, 30 s at 72 °C) and a final extension at 72 °C for 15 min followed by a period of 15 min at 68 °C [20, 41].

Amplification products were checked by electrophoresis on 1.5 % agarose gels stained with GelRed (Biotium) in 1× TBE buffer (pH 8.0). Amplified products of the PCR were submitted to an automatic DNA analyzer by capillary electrophoresis in the ABI 3130XL Genetic Analyzer (Applied Biosystems). The ET-550 ROX size standard (GE Healthcare) was used to determine the size of the alleles. Amplified fragments were observed and analyzed with the GENEMAPPER v4.0 software (Applied Biosystems).

Statistical analysis

Analysis of genetic diversity and fixation index

Genetic diversity was determined to compare adults and offspring, using the indexes total number of alleles over loci (*k*), average number of alleles per locus (*A*), number

Ramos *et al. BMC Genetics* (2016) 17:63

Page 8 of 10

of private alleles in each generation (A_p), and the observed (H_o) and expected (H_e) heterozygosities. These indexes were estimated using the GDA program [43]. Inbreeding was estimated using the fixation index (F). To test whether the F values were statistically different from zero, 1,000 Monte Carlo permutations of alleles among individuals, associated to a *Bonferroni* correction (95 %, $\alpha = 0.05$), were obtained using SPAGeDi 1.3 [44]. To investigate if the mean values of A, H_o, H_e and F were significantly different between adults and offspring, the Student *t-test* was used, with a prior verification of the homogeneity of variances of the two groups, using a Fisher's *F-test*. These analyzes were performed using the *var.test* and *t.test* functions of R package from the R project [45].

Analysis of the spatial genetic structure

The intrapopulation spatial genetic structure was studied using the mean coancestry coefficient (θ_{ij}) between pairs of adult plants, calculated according to Loiselle et al. [46] and using the SPAGeDI program. To visualize the spatial genetic structure, values of θ_{xy} were plotted against ten distance classes with the same number of pairwise individuals. In order to verify whether the spatial genetic structure had a significant deviation from a random structure, the CI of 95 % was calculated for each θ_{ij} observed value and each distance class, using 10,000 Monte Carlo permutations of individuals among different distance classes. To compare the spatial genetic structure with other studies we estimated the S_p statistic: $Sp = -b_k/(1 - \theta_1)$ [23], where θ_1 is the average coancestry coefficient calculated in the first distance class (0 to 21 m), and b_k is the slope of the regression curve in relation to the logarithm of the spatial distance (up to 361 m). To test the intensity of SGS, the spatial position of the individual was permuted 10,000 times to obtain the distribution frequency of b_k where the null hypothesis states that θ_1 and $\ln d_{xy}$ are not correlated (d_{xy} is the spatial distance between individuals x and y). These analyses were run using SPAGeDI 1.3 program.

Analysis of the group coancestry and population effective size

The group coancestry (Θ) [47] was estimated for the adult plants from pairwise coancestry coefficient between all pairs of individuals (θ_{ij}), using the estimator described in Loiselle et al. [46], implemented in the SPAGeDI program: $\Theta = [0.5n(1 + F_p) + \sum_{i=1}^{n}\sum_{j \neq i}^{n}\theta_{ij}]/n^2$, where n is the number of sampled individuals, F_p is the inbreeding coefficient of the population, estimated from the fixation index (negative value are assumed as zero). The effective population size (N_e was calculated following Cockerham [48] from the variance of gene frequencies due to genetic drift ($\sigma_p^2 = [(n - 1)/n]\Theta + (1 + F)/2n]p(1 - p)$, where n is the

sample size, p is frequency for a given neutral allele and F is the average inbreeding coefficient. In an idealized population under random mating, σ_p^2 value for a group of n individuals is $\sigma_p^2 = p(1 - p)/2n$ and as in a idealized population there is not related and inbred individuals, the term n can be substituted by N_e: $\sigma_p^2 = p(1 - p)/2N_e$. Thus, we can equate both σ_p^2 expression and derive the variance effective population size, $N_e = \frac{0.5}{\Theta\left(\frac{n-1}{n}\right)+\frac{1+F}{2n}}$.

Analysis of the pollen flow, dispersal patterns and dispersal kernel estimation

For the paternity analysis, the CERVUS 3.0.3 program [49] was used, based on a categorical maximum likelihood method. The offspring paternity was determined by the Δ estimated statistic, calculated using simulations, considering 10,000 repetitions (simulated for the offspring), zero error rate at the loci (0.00) and all the 124 reproductive palm trees (112 adults + 12 mother plants) as pollen candidates for the offspring (60 % of sampled pollen donors collected in the study area). We adopted the confidence levels of 80 % as suggested by Marshall et al. [49] for the paternity assigned. Self-fertilization was also considered as a possibility and was estimated. The pollen immigration rate (m_p) within the area was estimated as the number of offspring for which no father was assigned in the sampled area, divided by the total number of sampled offspring. The pollen dispersal distance for each progeny was calculated as the distance between the seed trees and the putative pollen donors by the Euclidian distance between two points. To verify whether the reproduction patterns were due to the distance between plants, the frequency of pollen dispersal curve was compared with the spatial distance among all plants using the Kolmogorov-Smirnov test [50]. The effective pollination neighboring area (A_{ep}) was calculated assuming a circular area around a central seed tree, by $A_{ep} = 2\pi\sigma_p^2$ [51], where σ_p^2 is the axial variance of the pollen dispersal.

The combined probability of exclusion of second parent, P_{2p} [52], was estimated with the NM+ program [53]. We also estimated pollen flow, selfing and pollen dispersal distance assuming an exponential power dispersal kernel [37], implemented in the NM+ program [53]. This program is based on neighborhood model [54]. In this model, the pollen dispersal distance and patterns are not derived from individual paternity assignments, as in the case of Cervus, but indirectly from a spatial explicit mating model. The model considers that paternity of an offspring may result from: i) self-fertilization with probability s; ii) migrant pollen from outside the plot, with probability m_p, or iii) outcrossing with a male located within the plot, with probability 1-s-m_p [54]. The NM+ was matched with initial settings using categorical

Ramos *et al. BMC Genetics* (2016) 17:63

Page 9 of 10

paternity assignment for our study plot. The neighborhood parameter was set to 'infinite' to include all sampled adults in our study plot as the neighborhood size [53]. Pollen dispersal was modeled using the exponential-power family parameter [37, 53] with estimates given of the scale (*a*) and shape (*b*) parameters from which the average distance of pollen dispersal (δ) is estimated.

Ethics
Not applicable.

Consent to publish
Not applicable.

Availability of supporting data
The data sets supporting the results of this article are included within the article.

Abbreviations
σ_p^2: Variance of gene frequencies due to genetic drift; *A*: Average number of alleles per locus; *a*: Dispersal kernel scale; A_{ep}: Effective pollination neighboring area; A_p: Number of private alleles; *b*: Shape parameters; b_k: Slope of the regression curve in relation to the logarithm of the spatial distance; C_{gf}: Cryptic gene flow; D: Kolmogorov-Smirnov test; d_{xy}: Spatial distance between individuals *x* and *y*; *F*: Fixation indexes; H_e: Expected heterozygosity; H_o: Observed heterozygosity; m_p: Pollen immigration rate; *n*: Sample size; N_e: Effective population size; n_p: Number of candidate parents; *p*: Frequency for a given neutral allele; P_{2p}: The combined probability of exclusion of the second parent; R^2: Correlation coefficient among the number of offspring fertilized by pollen of male genitor and the distance between mother plants; *s*: Selfing rate; SGS: Spatial genetic structure; S_p: Statistic which measure the extension of spatial genetic structure in the fist distance class; δ: Pollen dispersal distance; Θ: The group coancestry coefficient; θ_1: Average coancestry coefficient calculated in the first distance class; θ_{ij}: Mean coancestry coefficient.

Competing interests
The authors declare that they have no competing interests.

Authors' contribution
SLFR contributed in all steps, planning the study design, sampling the populations, laboratory analyses, conducting the data analysis and writing of the manuscript. GD contributed in data analysis and writing of the manuscript. AMS contributed in data analysis and writing of the manuscript. MTGL contributed in the laboratory analysis and writing of the manuscript. PYK contributed conceiving the main idea for the study design and writing of the manuscript. JLVM contributed in the sampling of the populations, farmers contacts and writing of the manuscript. MK contributed in data interpretation and writing of the manuscript. EAV contributed in discussing the main ideas and writing of the manuscript. All authors read and approved the final manuscript.

Authors' information
For correspondence concerning the species please contact with SFLR, JLVM and MTGL. For data analysis please contact mainly AMS. For all other correspondence please contact EAV and SLFR.

Author details
[1]Escola Superior de Agricultura "Luiz de Queiroz"/Universidade de São Paulo (ESALQ/USP), Departamento de Genética, C.P. 9, CEP 13418-900 Av. Pádua Dias, 11, Piracicaba, SP, Brazil. [2]Instituto Florestal de São Paulo, Seção de Melhoramento e Conservação Genética Florestal, C.P. 1322, CEP 01059-970 São Paulo, SP, Brazil. [3]Faculdade de Ciências Agrárias, Universidade Federal do Amazonas (UFAM), CEP 60077-000 Manaus, AM, Brazil. [4]ESALQ/USP, Departamento de Ciências Florestais, Av. Pádua Dias, 11, C.P. 9, CEP 13418-900 Piracicaba, SP, Brazil. [5]Embrapa Amazônia Ocidental, C.P. 319, CEP 69048-660 Manaus, AM, Brazil. [6]University of Florida, School of Forest Resources & Conservation, PO Box 110410, Gainesville, FL 32611-0410, USA.

References
1. Sosnowska J, Balslev H. American palm ethnomedicine: a meta-analysis. J Ethnobiol Ethnomed. 2009;5:43–53.
2. Kahn F, Granville J-JD. Palms and forest management in Amazonia. In: Kahn F, Granville J-JD, editors. Palms in forest ecosystems of Amazonia. New York: Springer New York LLC; 1992. p. 155–67.
3. Ramos SLF, Macêdo JLV, Martins CC, Lopes R, Lopes MTG. Tratamentos pré-germinativos e procedência de sementes do tucumã-do-Amazonas para a produção de mudas. Rev Bras Frutic. 2011;33:962–9.
4. Lopes MTG, Macêdo JLVD, Lopes R, Leeuwen JV, Ramos SLF, Bernardes LG. Domestication and breeding of the Tucum Palm. In: Borém A, Lopes MTG, Clement CR, Noda H, editors. Domestication and breeding: Amazonian species. Viçosa: Universidade Federal de Viçosa; 2012. p. 421–36.
5. Roncal J, Kahn F, Millan B, Couvreur TLP, Pintaud J-C. Cenozoic colonization and diversification patterns of tropical American palms: evidence from *Astrocaryum* (Arecaceae). Bot J Linn Soc. 2013;171:120–39.
6. Kalia RK, Rai MK, Kalia S, Singh R, Dhawan AK. Microsatellite markers: an overview of the recent progress in plants. Euphytica. 2011;177:309–34.
7. Davey JW, Hohenlohe PA, Etter PD, Boone JQ, Catchen JM, Blaxter ML. Genome-wide genetic marker discovery and genotyping using next-generation sequencing. Nat Rev Genet. 2011;12:499–510.
8. Ouborg NJ, Pertoldi C, Loeschcke V, Bijlsma R, Hedrick PW. Conservation genetics in transition to conservation genomics. Trends Genet. 2010;26:177–87.
9. Godoy JA, Jordano P. Seed dispersal by animal: exact identification of source trees with endocarp DNA microsatellites. Mol Ecol. 2001;10:2275–83.
10. Smouse PE, Sork VL. Measuring pollen flow in forest trees: an exposition of alternative approaches. For Ecol Manage. 2004;197:21–38.
11. Hamrick JL, Murawski DA, Nason JD. The influence of seed dispersal mechanisms on the genetic structure of tropical tree populations. Vegetatio. 1993;107-108:281–97.
12. Dow BD, Ashley MV. High levels of gene flow in bur oak revealed by paternity analysis using microsatellites. J Hered. 1998;89:62–70.
13. Bittencourt JM, Sebbenn AM. Patterns of pollen and seed dispersal in a small fragmented population of a wind pollinated *Araucaria angustifolia* in southern Brazil. Heredity. 2007;99:580–91.
14. Manoel RO, Alves PF, Dourado CL, Gaino APSC, Freitas MLM, Moraes MLT, et al. Contemporary pollen flow, mating patterns and effective population size inferred from paternity analysis in a small fragmented population of the Neotropical tree *Copaifera langsdorffii* Desf. (Leguminosae-Caesalpinioideae). Conserv Genet. 2012;13:613–23.
15. Degen B, Sebbenn AM. Genetics and Tropical Forests. In: Köhl M, Pancel L, editors. Tropical Forestry Handbook. Berlin Heidelberg: Springer; 2014. p. 1–30.
16. Burczyk J, DiFazio SP, Adams WT. Gene flow in forest trees: how far do genes really travel? For Genet. 2004;11:1–14.
17. Ashley MV. Plant parentage, pollination, and dispersal: how DNA microsatellite have altered the landscape. CRC Crit Rev Plant Sci. 2010;29:148–61.
18. Meagher TR. Analysis of paternity within a natural population of *Chamaelirium luteum*.1. Identification of most-likely male parents. Am Nat. 1986;128:199–215.

Ramos *et al. BMC Genetics* (2016) 17:63

Page 10 of 10

19. Bacelar-Lima CG, Mendonça MS, Barbosa TCTS. Morfologia floral de uma população de tucumã *Astrocaryum aculeatum* G. Mey (Arecaceae) na Amazônia central. Acta Amazon. 2006;36:407–12.

20. Ramos SLF, Lopes MTG, Lopes R, Cunha RNV, Macêdo JLV, Contim LAS, et al. Determination of the mating system of Tucumã palm using microsatellite markers. Crop Breed Appl Biotechnol. 2011;11:181–5.

21. Jorge MLSP, Howe HF. Can forest fragmentation disrupt a conditional mutualism? A case from central Amazon. Oecologia. 2009;161:709–18.

22. Hardy OJ, Maggia L, Bandou E, Breyne P, Caron H, Chevallier MH, et al. Fine-Scale genetic structure and gene dispersal inferences in 10 Neotropical tree species. Mol Ecol. 2006;15:559–71.

23. Vekemans X, Hardy OJ. New insights from fine-scale spatial genetic structure analysis in plant populations. Mol Ecol. 2004;13:921–35.

24. Sebbenn AM, Carvalho ACM, Freitas MLM, Moraes SMB, Gaino APSC, Silva JM, et al. Low levels of realized seed and pollen gene flow and strong spatial genetic structure in a small, isolated and fragmented population of the tropical tree *Copaifera langsdorffii* Desf. Heredity. 2011;106:134–45.

25. Ribeiro FE, Baudouin L, Lebrun P, Chaves LJ, Brondani C, Zucchi MI, et al. Population structures of Brazilian tall coconut (*Cocos nucifera* L.) by microsatellite markers. Genet Mol Biol. 2010;33:696–702.

26. Carneiro F, Degen B, Kanashiro M, Lacerda AE, Sebbenn AM. High levels of pollen dispersal detected through paternity analysis from a continuous *Symphonia globulifera* population in the Brazilian Amazon. For Ecol Manage. 2009;258:1260–6.

27. Silva CRS, Albuquerque PSB, Ervedosa FR, Mota JWS, Figueira A, Sebbenn AM. Understanding the genetic diversity, spatial genetic structure and mating system at the hierarchical levels of fruits and individuals of a continuous *Theobroma cacao* population from the Brazilian Amazon. Heredity. 2011;106:973–85.

28. Carneiro FS, Lacerda AEB, Lemes MR, Gribel R, Kanashiro M, Wadt LHO, et al. Effects of selective logging on the mating system and pollen dispersal of *Hymenaea courbaril* L. (Leguminosae) in the Eastern Brazilian Amazon as revealed by microsatellite analysis. For Ecol Manage. 2011;262:1758–65.

29. Oliveira MSPD, Couturier G, Beserra P. Biologia da polinização da palmeira tucumã (*Astrocaryum vulgare* Mart.) em Belém, Pará, Brasil. Acta Bot Brasilica. 2003;17:343–53.

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

31. Eguiarte LE, Burquez A, Rodriguez J, Martinez-Ramos M, Sarukhan J, Pinero D. Direct and indirect estimates of neighborhood and effective population size in a tropical palm. Astrocaryum mexicanum. Evolution. 1993;47:75–87.

32. Ottewell K, Grey E, Castillo F, Karubian J. The pollen dispersal kernel and mating system of an insect-pollinated tropical palm. Oenocarpus bataua. Heredity. 2012;109:332–9.

33. Baldauf C, Ciampi-Guillardi M, Aguirra TJ, Correa CE, Santos FAM, Souza AP, et al. Genetic diversity, spatial genetic structure and realised seed and pollen dispersal of *Himatanthus drasticus* (Apocynaceae) in the Brazilian savanna. Conserv Genet. 2014;15:1073–83.

34. Tarazi R, Sebbenn AM, Kageyama PY, Vencovsky R. Long-distance dispersal in a fire- and livestock-protected savanna. Ecol Evoll. 2013;3:1003–15.

35. Saro I, Robledo-Arnuncio JJ, González-Pérez MA, Sosa PA. Patterns of pollen dispersal in a small population of the Canarian endemic palm (Phoenix canariensis). Heredity. 2014;113:3215–23.

36. Klein EK, Lavigne C, Gouyon P-H. Mixing of propagules from discrete sources at long distance: comparing a dispersal tail to an exponential. BMC Ecol. 2006;6:3.

37. Austerlitz F, Dick CW, Dutech C, Klein EK, Oddou-Muratorio S, Smouse PE, et al. Using genetic markers to estimate the pollen dispersal curve. Mol Ecol. 2004;13:937–54.

38. Matter P, Kettle CJ, Ghazoul J, Hahn T, Pluess AR. Evaluating contemporary pollen dispersal in two common grassland species *Ranunculus bulbosus* L. (Ranunculaceae) and *Trifolium montanum* L. (Fabaceae) using an experimental approach. Plant Biol. 2013;15:583–92.

39. Kottek M, Grieser J, Beck C, Rudolf B, Rubel F. World Map of the Köppen-Geiger climate classification updated. Meteorol Z. 2006;15:259–63.

40. Doyle JJ, Doyle JL. Isolation of plant DNA from fresh tissue. Focus. 1990;12:13–5.

41. Ramos SLF, Macedo JLV, Lopes MTG, Batista JS, Formiga KM, Da Silva PP, et al. Microsatellite loci for tucumã of Amazonas (*Astrocaryum aculeatum*) and amplification in other Arecaceae. Am J Bot. 2012;99:e508–10.

42. Schuelke M. An economic method for the fluorescent labeling of PCR fragments. Nat Biotechnol. 2000;18:233–4.

43. Lewis PO, Zaykin D: Genetic data analysis: computer programe for the analyse of allelics data (software) version 1.0 (d16c). http://phylogeny.uconn.edu/software/. Accessed 04 Oct 2011.

44. Hardy OJ, Vekemans X. SPAGeDI: a versatile computer program to analyze spatial genetic structure at the individual or population levels. Mol Ecol Notes. 2002;2:618–20.

45. R Core Team. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2013. https://www.r-project.org. Accessed 30 Dec 2013.

46. Loiselle BA, Sork VL, Nason J, Graham C. Spatial genetic structure of a tropical understory shrub, *Psychotria officinalis* (Rubiaceae). Am J Bot. 1995;82:1420–5.

47. Lindgren D, Mullin TJ. Relatedness and status number in seed orchard crops. Can J For Res. 1998;28:276–83.

48. Cockerham CC. Variance of gene frequencies. Evolution. 1969;23:72–84.

49. Marshall TC, Slate J, Kruuk LEB, Pemberton JM. Statistical confidence for likelihood-based paternity inference in natural populations. Mol Ecol. 1998;7:639–55.

50. Sokal RR, Rohlf FJ. Biometry: Principles and Practices of Statistics in Biological Research. 3rd ed. New York: W.H. Freeman; 1995.

51. Levin DA. The paternity pool plants. Am Nat. 1988;132:309–17.

52. Dow BD, Ashley MV. Microsatellite analysis of seed dispersal and parentage of sampling in bur oak. Quercus macrocarpa. Mol Ecol. 1996;5:615–27.

53. Chybicki IJ, Burczyk J. NM+: software implementing parentage-based models for estimating gene dispersal and mating patterns in plants. Mol Ecol. 2010;10:1071–5.

54. Burczyk J, Adams WT, Moran JF, Griffin AR. Complex patterns of mating revealed in a *Eucalyptus regnans* seed orchard using allozyme markers and the neighbourhood model. Mol Ecol. 2002;11:2379–91.

55. Hijmans RJ, Guarino L, Mathur P. DIVA-GIS a geographic information system for the analysis of biodiversity data: Manual, Version 7.5. International Potato Center: Lima; 2012. http://www.diva-gis.org/download.