Journal of Heredity 2003:94(5):399–406
DOI: 10.1093/jhered/esg087

Genetic Structure, Mating System, and Long-Distance Gene Flow in Heart of Palm (*Euterpe edulis* Mart.)

F. A. Gaiotto, D. Grattapaglia, and R. Vencovsky

From the Laboratório de Genética de Plantas, EMBRAPA–Recursos Genéticos e Biotecnologia, C.P. 02372, Brasília, DF 70770-900, Brazil (Gaiotto and Grattapaglia); Departamento de Genética, ESALQ/USP, Av. Pádua Dias, 11 C.P. 83, Piracicaba, SP 13400-970 Brazil (Gaiotto and Vencovsky); and Laboratório de Biotecnologia Genômica, Universidade Católica de Brasília, SGAN 916 Módulo B Brasília, DF 70790-160, Brazil (Grattapaglia). F. A. Gaiotto is currently at the Departamento de Ciências Biológicas, Universidade Estadual de Santa Cruz—UESC, Rodovia Ilhéus-Itabuna, km 16 Salobrinho, Ilhéus/BA CEP 45650-000.

Address correspondence to D. Grattapaglia at one of the addresses above, or e-mail: dario@cenargen.embrapa.br.

Abstract

We report a detailed analysis of the population genetic structure, mating system, and gene flow of heart of palm (*Euterpe edulis* Mart.–*Arecaceae*) in central Brazil. This palm is considered a keystone species because it supplies fruits for birds and rodents all year and is intensively harvested for culinary purposes. Two populations of this palm tree were examined, using 18 microsatellite loci. The species displays a predominantly outcrossed mating system ($t_m = 0.94$), with a probability of full sibship greater than 70% within open-pollinated families. The following estimates of interpopulation genetic variation were calculated and found significant: $F_{IT} = 0.17$, $F_{IS} = 0.12$, $F_{ST} = 0.06$, and $R_{ST} = 0.07$. This low but significant level of interpopulation genetic variation indicates high levels of gene flow. Two adult trees were identified as likely seed parents ($P > 99.9\%$) of juveniles located at a distance of 22 km. Gene flow over such distances has not been reported before for tropical tree species. The establishment and management of *in situ* genetic reserves or *ex situ* conservation and breeding populations for *E. edulis* should contemplate the collection of several hundreds open-pollinated maternal families from relatively few distant populations to maximize the genetic sampling of a larger number of pollen parents.

The heart of the palm is the edible apical meristem of wild palms, considered an exotic gourmet commodity, and therefore has a high monetary value. *Euterpe edulis* Mart.–*Arecaceae*, yields the most desirable heart, and it is therefore commonly denominated *heart of palm*. This species is found throughout the Atlantic Forest, as well as in gallery forests of Cerrado vegetation from southern to northeastern Brazil. This palm requires 8 to 10 years to reach harvestable maturity. Harvesting of the heart from this single-stemmed palm requires its destruction. The high commercial value of the heart of palm and poor incentives for an adequate management of the existing resources have resulted in a serious threat to the existing populations in some areas of the Brazilian Atlantic Forest (Galetti and Fernandez 1998; Orlande et al. 1996).

The inflorescence of *E. edulis* is formed by a composite of floral triads, usually consisting of two male and one female flower each, with strong protandry. This flower structure prevents autogamy. Although the main pollinator of the palm is a short-flying-distance bee (*Trigona spinipes* Fabricius), it is also subject to wind pollination (Reis et al. 1993). Flower and fruit production of *E. edulis* can last more than four months. Because of the high annual fruit yield of *E. edulis*, it is an important food source for mammals and birds and therefore considered a "keystone species" (Terborgh 1986).

Population dynamics of plants can significantly affect evolutionary factors such as selection and genetic drift, and consequently have important impacts on the genetic structure of populations. However, the magnitude of effects of drift and selection on the patterns of genetic variation will depend on the reproductive ability of the organism and the rates of gene flow. The relationship between gene dispersal and levels of genetic diversity within populations can be determined by examining the effect of plant-breeding-system and pollen-dispersal mechanisms on levels of genetic diversity (Hamrick and Nason 1996). There are different ways to study the movement of genes within and between populations. One of the most direct methods for studying

gene flow is performing population-specific paternity tests using highly informative microsatellite markers. Microsatellites, also called simple sequence repeats (SSR), have been increasingly used for genetic structure analysis, gene flow determinations, and paternity studies of tropical forest populations (Chase et al. 1996a,b; Collevatti et al. 2001; Dayanandan et al. 1999). Such investigations have yielded data that have improved understanding of the population dynamics of a number of species. However, essentially all studies of mating systems of Neotropical trees have been performed by examining tropical forest species in Central America, particularly Panamá, Costa Rica and Mexico, and Asian tropics (e.g., James et al. 1998; Nason et al. 1998). Few studies have been published that focus on Brazilian tree species (e.g., Collevatti et al. 2001; O'Malley et al. 1988). Moreover, for *Euterpe* species, and in fact for all species of the large group of palm trees, important components of tropical forest ecosystems, detailed studies of gene flow, genetic structure, and paternity have not yet been reported.

We are interested in understanding the population genetics and mating system of economically important tropical trees in order to generate useful information for conservation, domestication, and prebreeding programs. In this study we employed highly informative microsatellite markers to investigate the preferential mating system, the distribution of genetic variation, and the extent of gene flow between two populations of *E. edulis* in gallery forests of Cerrado vegetation in central Brazil.

Materials and Methods

Plant Material

Two natural populations of *E. edulis* were studied, located in Brasilia National Park (PNB) (15°45′48″S, 47°58′80″W), and in the Ecological Reserve of the Brazilian Institute of Geography and Statistics (IBGE; 15°56′26″S, 47°53′37″W). Both populations are found in gallery forests of Cerrado vegetation in central Brazil. The distance between these populations is 22 km.

For the mating system study, open-pollinated maternal half-sib seeds were collected from 13 randomly selected adult trees from IBGE and 12 adults in PNB. Seeds were sown in sterile soil in 2 L pots, and plants were grown in a greenhouse under controlled temperature (20–35°C) and natural photoperiod. Progeny arrays of 12 individuals per each open-pollinated family were used to determine outcrossing rates.

To study genetic structure and gene flow, leaf material from adults and juvenile palm trees was sampled randomly along a transect in both populations. All juvenile individuals were approximately 1 year old, for all existing seedlings from older crops were marked with red flags tied to a metal wire in the previous year as part of a parallel ongoing experiment. A total of 111 adults and 231 juveniles in PNB population and 33 adults and 208 juveniles in IBGE population were sampled. Precise map position was recorded for all adult individuals along the transects (Figure 1).

Microsatellite Marker Analysis

Genomic DNA extraction from fresh leaf tissue followed standard CTAB procedure (Doyle and Doyle 1987). Microsatellite markers used in this study and basic PCR conditions were described previously (Gaiotto et al. 2001). Six microsatellite loci were used in the mating system study, and up to 18 microsatellite loci were used for genetic structure analysis and paternity determinations. Marker loci were PCR amplified individually or in duplex and triplex reactions and analyzed in multiplex gel loading assays of up to six loci, based on the allele size range, to optimize laboratory time and costs. Multiplex PCR reactions could be performed with the following loci: duplex of loci *EE8* and *EE23*; duplex of loci *EE2* and *EE32*; triplex of loci *EE43*, *EE45*, and *EE52*; and duplex of loci *EE48* and *EE54*. The PCR cocktail (13 μl) contained 7.5 ng of genomic DNA, 250 μM dNTPs (Gibco-BRL), 0.75 μM $MgCl_2$ (Gibco-BRL, Bethesda, MD), 1 × PCR buffer (Gibco-BRL), (10mM Tris-HCl, 50mM KCl, 1.5mM $MgCl_2$ pH8.3), 2.5 μg/μl BSA (New England Biolabs), 0.2 μM of each primer (Operon Technologies, Alameda, CA), one of them dye-labeled (6-FAM, HEX, or TET dyes), and 1 U of *Taq* DNA polymerase (Gibco-BRL). Amplifications were performed with a MJ Research PTC-100 thermal controller, using the following protocol: 96°C for 2 min; 30 cycles of 94°C for 1 min, the primer-specific annealing temperature for 1 min (see Gaiotto et al. 2001), 72°C for 1 min, and ending with 72°C for 7 min. Loci were loaded in the following loading multiplex systems: (1) Hexaplex I composed of loci *EE15*, *EE2*, and *EE32* (labeled with 6-FAM); *EE3*, *EE8* (labeled with HEX) and *EE23* (labeled with TET); (2) Hexaplex II composed of loci *EE43*, *EE52* (labeled with TET), *EE45*, *EE48* (labeled with 6-FAM), *EE54*, and *EE9* (labeled with HEX); Pentaplex composed of loci *EE41*, *EE59* (labeled with TET), *EE47*, *EE63* (labeled with 6-FAM), and *EE25* (labeled with HEX); and a single locus *EE5* labeled with 6-FAM. Within each loading multiplex the PCR products of each locus were pooled in equal quantities, and 2 μl of the mixture were combined with 5.2 μl of 1:5 loading buffer (25 mM EDTA, 50 mg/ml blue dextran): deionized formamide and 0.6 μl of an internal fluorescently labeled DNA standard (Brondani and Grattapaglia 2001). Electrophoresis and fragment sizing was carried out with an automated DNA sequencer (Applied Biosystems 377-XL), with virtual filter C, using the software GeneScan 2.1 and Genotyper 2.5.x (Applied Biosystems, Foster City, CA).

Data Analysis

Genetic diversity measures and Wright *F* statistics were estimated under a random model (Weir 1996; i.e., populations sampled are considered to represent the species and have a common evolutionary history). Allelic frequencies, number of alleles per locus (*A*), observed heterozygosity (H_o), expected heterozigosity (H_e), and *F* statistics (F_{IS}, F_{ST}, and F_{IT}) were estimated with the software GDA 1.0 (Lewis and Zaykin 2000).

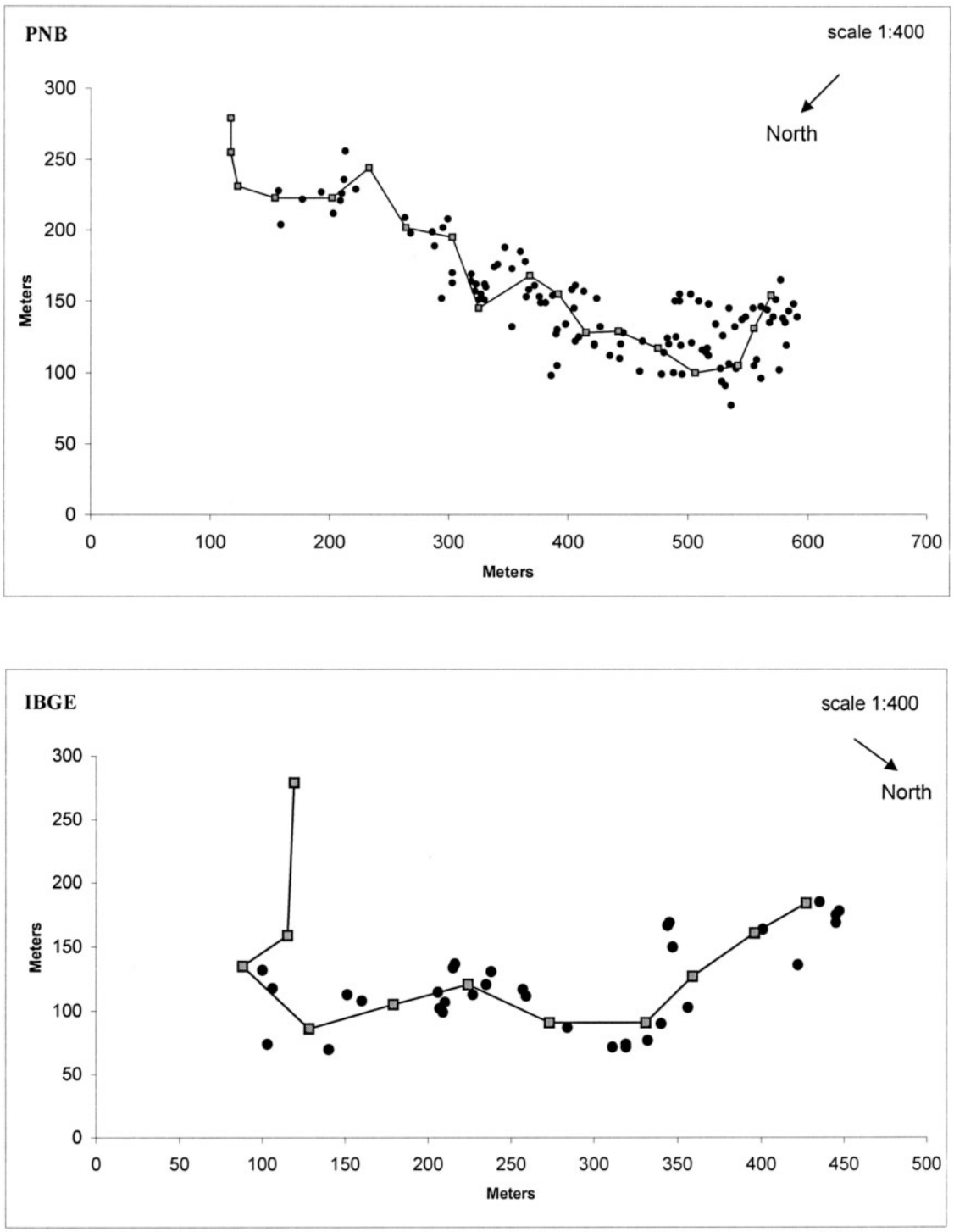

Figure 1. Map showing the spatial distribution of the sampled adult trees of *E. edulis* (black dots) along the transects (grey lines) in PNB and IBGE populations.

Because the mutation process at microsatellites does not conform to an infinite allele model, we estimated the genetic differentiation R_{ST}, an analogue of F_{ST} developed specifically for microsatellite data, following a stepwise mutation model (Slatkin 1995). Estimates of R_{ST} and gene flow (N_m) were calculated using the software R_{ST} Calc 2.2 (Goodman 1997).

When populations are under Wright's equilibrium (Wright 1965), outcrossing rate is a function of the within-population inbreeding measure (F_{IS}). Apparent outcrossing

Table 1. Estimates of genetic variation for adults, juveniles, and greenhouse seedlings for the two populations (PNB and IBGE)

Population	N	L	#A	H_o	H_e	F_{is}	L	U	t_a	t_m
PNB										
Adults	111	18	11.2	0.70	0.75	0.06	−.02	.14	0.88	
Juveniles	231	18	11.3	0.66	0.73	0.09	.04	.15	0.83	
Seedlings	144	06	9.5	0.73	0.76	0.04	.02	.05	0.93	0.90 (0.04)
IBGE										
Adults	33	18	8.5	0.66	0.74	0.11	−.02	.24	0.80	
Juveniles	208	18	11.6	0.62	0.74	0.16	.09	.23	0.73	
Seedlings	156	06	10.5	0.79	0.79	0.01	−.03	.05	0.99	0.98 (0.02)
Mean	10.8	0.69	0.75	—	0.86	0.94 (0.03)				

N is the number of individuals sampled; L is the number of microsatellite loci typed; #A is the mean number of alleles; H_o is the observed heterozygosity; H_e is the expected heterozigosity; F_{is} is the within-population inbreeding; L and U are, respectively, the lower and upper 95% confidence interval limits on the estimates of F_{is}; t_a is the apparent outcrossing rate; and t_m is the multilocus outcrossing rate with its standard deviation.

rate (t_a) was estimated by using the estimated inbreeding coefficient (F_{IS}), using the following estimator as described by Weir (1996):

$$t_a = \frac{(1 - \hat{F}_{IS})}{(1 + \hat{F}_{IS})} \quad (1)$$

The multilocus outcrossing rate (t_m) was estimated under the mixed mating model described by Ritland and Jain (1981), with the software MLTR 2.3 (Ritland 1990). The procedure estimates the multilocus outcrossing rate (t_m), the mean single-locus outcrossing rate (t_s), the difference between estimates ($t_m - t_s$) that represents the outcrossing rate between related individuals, the correlation of outcrossing rate within progeny arrays (r_t), and the correlation of outcrossed paternity within progeny arrays (or the probability that a randomly chosen pair of progeny from the same array are full sibs; r_p). Standard errors were estimated based on 10,000 bootstraps between individuals within progeny arrays.

Paternity determinations were carried out with standard paternity testing methods (Weir 1996). Initially, exclusion analysis was done by comparing genotypes of all adult individuals in turn with each one of the juvenile individuals at all typed loci. When the adult and the juvenile did not share any allele, the adult was excluded from the group of palm trees that could be a parent of the juvenile. When the adult could not be excluded (i.e., the adult and the juvenile shared one or two alleles), specific calculations were carried out to estimate the paternity index (PI) for each locus. Using the product rule, a multilocus combined paternity index (CPI) was calculated for loci that were determined to be independent by linkage disequilibrium tests carried out in a previous study (Gaiotto et al. 2001). From the CPI, and with a specific prior probability of paternity (P), an estimate of the posterior probability of paternity (PP) was obtained, according to Weir (1996), as

$$\mathrm{PP} = \frac{\mathrm{CPI} \cdot P}{[\mathrm{CPI} \cdot P + (1 - P)]} \quad (2)$$

We used three different prior probabilities (P) to arrive at the estimates of PP, including: (1) a probability of .5, which is the prior probability suggested in the literature and typically used in human paternity testing, which assumes equal probabilities of paternity and nonpaternity before the test (Gjertson and Morris 1995; Smouse and Chakraborty 1986); (2) a probability of .007, which is the probability that any adult palm tree sampled is the parent (1/144); and (3) a probability of .0001, which is an almost zero prior probability of any adult individual's being the parent, the most conservative approach. Likelihood ratios for paternity calculations were obtained, based on a database of allele frequencies estimated from all 144 adult individuals, using a conservative minimum allele frequency of 5/2 n (i.e., 0.017), to compensate for sampling bias, as suggested by Budowle and Monson (1993).

Results

Genetic Variation

Both populations examined showed high levels of expected heterozygosity (H_e), varying from 0.73 to 0.79. For the 18 microsatellite loci analyzed, observed heterozygosities (H_o) varied from 0.62 to 0.79 (Table 1). Mean number of alleles per locus was also high, varying from 8.5 to 11.5. Estimates of the inbreeding coefficient F_{IS} for all developmental stages were low and not significant for the adult stage in both populations and for the seedling population from the IBGE population (Table 1).

Genetic Structure (F Statistics and R_{ST}) and Gene Flow

The inbreeding coefficient (F_{IS}) was estimated, with consideration of all the field individuals, adults and juveniles, of both populations (Table 2). The estimated R_{ST} was 0.07 and F_{ST} was 0.06. Both estimates, although low, were significantly different from zero (Table 3). Despite the significance of F_{ST} and R_{ST} values, N_m based on the estimated R_{ST} was 3.4 (Table 2), indicating that there is a relatively large amount of migration even between geographically distant populations. Analyses between populations, using juveniles and adults, showed an F_{IT} value of 0.17. All the estimates of F statistics and R_{ST} were statistically significant using a confidence interval of 95%.

Table 2. Wright's F statistics, R_{ST}, and relative number of migrants per generation (N_m), based on adults and juveniles from the two populations of *E. edulis* (PNB and IBGE) and the 95% confidence intervals on the estimates

	F_{IS}	F_{ST}	F_{IT}	R_{ST}	N_m
Estimate	0.12	0.06	0.17	0.07	3.37
Upper bound	0.18	0.09	0.24	0.09	3.98
Lower bound	0.06	0.04	0.11	0.06	2.65

F_{IS} = within-population inbreeding coefficient; F_{ST} = fixation index, genetic differentiation due to random genetic drift; F_{IT} = overall inbreeding coefficient; R_{ST} = genetic differentiation under the stepwise mutation model (Slatkin 1995).

Table 3. Estimates of the mating system parameters for the two populations (PNB and IBGE)

	PNB	IBGE
t_m	0.90 (0.04)	0.98 (0.02)
t_s	0.94 (0.04)	0.94 (0.04)
$t_m - t_s$	−0.04 (0.03)	0.04 (0.04)
r_p	0.71 (0.12)	0.73 (0.07)
r_t	0.23 (0.04)	0.54 (0.06)

t_m is the multilocus outcrossing rate; t_s is the mean single locus outcrossing rate; r_p is the correlation of outcrossed paternity within progeny arrays; and r_t is the correlation of outcrossing rate within progeny arrays; standard errors based on 10,000 bootstraps are in parentheses.

Mating System

Results on the outcrossing rate (apparent and multilocus) are presented in Table 3. The apparent outcrossing rate (t_a), calculated using the estimated F_{IS}, allowed for the estimation of the outcrossing rate for different developmental stages in the populations (i.e., adults, juveniles from the field, and planted seedlings in the greenhouse). In general, the estimates of average t_a were slightly lower than the average t_m. It should be noted, however, that the mixed model estimate was obtained for only planted seedlings for which open-pollinated half-sib progeny arrays were available. The apparent outcrossing rate for this class of plants in both populations (0.93 and 0.99 for PNB and IBGE, respectively) was very much in agreement with the multilocus outcrossing rate (0.90 and 0.98 for PNB and IBGE, respectively). Under the mixed mating model, the average single locus outcrossing rates (t_s) were also estimated (Table 3), and the values were similar for both populations. The difference between t_m and t_s represents the outcrossing rate between related individuals, which contributes to the increase of inbreeding within populations. These differences were small, (−0.04 for PNB, and 0.04 for IBGE), indicating that the occurrence of biparental inbreeding (i.e., cross-fertilization events between close relatives) is not a common occurrence.

The probabilities of a randomly chosen pair of individuals from the same open-pollinated half-sib family being full sibs (r_p) were 0.73 (± 0.07) for IBGE and 0.71 (± 0.12) for the PNB population, suggesting full sibship in the same half-sib family. Correlations between outcrossing rates within families (r_t) were 0.54 (± 0.06) and 0.23 (± 0.037) for IBGE and PNB populations, respectively (Table 3), indicating that self-fertilization had occurred in some of the studied families.

Paternity Analysis

For 153 out of 231 (66.2%) juveniles sampled in the PNB population, one parent was determined, and for 54 of them both parents (23.4%) were identified, according to the paternity exclusion analysis. In the IBGE population, out of 208 juvenile sampled, 109 (52.4%) had one parent determined, and for 29 both parents (13.9%) were determined. Surprisingly, however, the paternity or maternity of five juvenile individuals sampled in the IBGE population was attributed to adult individuals located in the PNB, with a probability greater than 99.9% (Table 4). Varying the prior probability from .5 to a conservative .0001, the posterior probability of paternity or maternity remained greater than 93% for two of the juvenile individuals (Table 4).

Variation in reproductive success among adult individuals was observed in both populations. Two adult individuals showed high probabilities to be the parents, most likely maternal, of 9.5% and 19.0%, respectively, of the 231 juveniles sampled in PNB. A similar trend was seen in the IBGE population, where one adult was determined to be the parent of 21.6% of the 208 juveniles sampled in this population.

Discussion

The 18 microsatellite loci used for genetic analyses in this work detected high levels of allelic variation in both *E. edulis* populations, confirming the usefulness of these markers for refined study of population structure, mating system, and paternity in species of *Euterpe* (Gaiotto et al. 2001). This is one of the first reports to use relatively large batteries of microsatellite markers to estimate mating system parameters and long-distance gene flow events in a tropical tree. Although an increasing number of gene flow and paternity analyses have been carried out with this class of markers (e.g., Chase et al. 1996a,b; James et al. 1998), only recently has

Table 4. Probability of paternity or maternity with which adult individuals from PNB population were declared parents of juvenile individuals from the IBGE population, 22 km away

Juvenile individual	Putative parent	Probability of paternity/maternity		
		$P = .5$	$P = .007$	$P = .0001$
IBGE 27	PNB71	99%, 95%	93%, 33%	16%, 57%
IBGE 36	PNB4	99%, 89%	86%, 95%	8%, 64%
IBGE 121	PNB13	100%, 00%	99%, 95%	96%, 86%
IBGE 124	PNB19	100%, 00%	99%, 90%	93%, 38%
IBGE 146	PNB13	99%, 90%	87%, 34%	8%, 91%
IBGE 169	PNB92	99%, 57%	61%, 88%	2%, 25%

Probabilities of paternity/maternity were estimated with increasingly conservative prior probabilities (P) of paternity.

a detailed understanding of mating systems in tropical trees been pursued with large batteries of microsatellites, allowing the precise detection of outcrossing events even between close relatives (Collevatti et al. 2001).

Studies evaluating the rate of inbreeding in forest tree populations have demonstrated that, during the early stages of development, values of F_{IS} are typically positive and significant. As trees grow older this estimate tends to become not significantly different from zero or even display negative values due to selection's operating against inbred individuals after seedling establishment. This trend has often been attributed to selection against homozygous individuals (e.g. Eguiarte et al. 1992; Gaiotto et al. 1997). For both *E. edulis* populations this was also the trend observed (Table 1). This behavior can be explained by the mating patterns detected in this study. *E. edulis* displays a predominantly outcrossing mating system with a low but significant amount of selfing, around 7–10%, however with no detectable biparental inbreeding. Furthermore, an interesting characteristic that is contributing to the significant within-population inbreeding is the high probability of full sibship among individuals within open-pollinated progeny (>70%).

The close agreement between the estimates of apparent and multilocus outcrossing rates is a clear indication that seedling populations are under Wright equilibrium. Similar results were also obtained using isozymes for other populations of *E. edulis* sampled throughout the Atlantic Forest in southern Brazil (Reis et al. 1997). Our results indirectly corroborate the adequacy of microsatellite markers for mating system studies when compared to the more traditionally employed isozymes.

Low or negative values for the difference between t_m and t_s (Table 2) suggest that breeding among related individuals is not a common occurrence in undisturbed *E. edulis* populations. On the other hand, the high r_p values in both populations support the hypothesis that the open-pollinated families are most likely composed of groups of full sibs. This observation might be explained by two alternative or complementary hypotheses as pointed out earlier by Ritland (1989): (1) the visitation behavior of the pollinator bees is such that multiple depositions of pollen loads, each load consisting of pollen from a single source, are very frequent, and (2) some adult individuals within the sampled areas display a much greater reproductive success when compared to others in the same population. The results of the detailed paternity analysis carried out in conjunction with the mating system study herein strongly support the second hypothesis, although the first one cannot be disregarded. A particular pair of adult individuals was found to be the parents of a large proportion of juveniles in both populations. For most of these juveniles, one parent, most likely the female parent, was near the juvenile plant analyzed, whereas the other, most likely the pollen parent, was distant from it.

Two measures of interpopulation genetic differentiation were used in this study (F_{ST} and R_{ST}). Both estimates, although relatively low, were numerically very similar and significantly different from zero, suggesting that little differentiation exists among these populations located in the same geographical region and that a significant rate of gene flow is occurring among them. F_{ST} is based on probabilities of identity by descent using the infinite allele model, whereas R_{ST} is estimated under the stepwise mutation model. In the stepwise mutation model alleles can be identical in state because of mutation, which acts to increase or decrease the size of the microsatellite repeats and, by consequence, the eletrophoretic mobility of an allele by one repeat unit (Rousset 1996). According to Slatkin (1995), this mutation model is more appropriate for the analysis of loci with a high mutation rate, such as microsatellites. In estimation of R_{ST}, the estimation procedure accounts for potential genotyping inaccuracies due to stutter bands or peaks that are typical of dinucleotide microsatellite loci. Inaccurate allele sizing, especially across gels, can occur, even when automatic sequencers are used for sizing microsatellite products (Haberl and Tautz 1999). In contrast to the estimate of F_{ST}, the impact of errors in allele size declaration is minimized during the estimation of R_{ST}. Such inaccuracies were avoided in this study by consistently running a number of control samples in every gel to ensure the accuracy of size estimation across gels. This procedure was a likely reason for R_{ST} and F_{ST} value similarities, suggesting that allele sizing conducted herein was accurate.

In the paternity analysis, one parent could be unequivocally determined for more than 50% of the juveniles in both populations. As individuals were sampled over a distance of 150–240 m, corresponding to the average distance flown by pollinators of *E. edulis* (Reis et al. 1993), these results indicate that gene flow within populations is occurring in agreement with the general expectations. Furthermore, the high degree of confidence in the determination of parentage corroborates earlier reports that microsatellite markers are useful for detailed paternity and maternity investigation in natural populations (Chase et al. 1996b).

As in others' studies (Kaufman et al. 1998; Schnabel et al. 1998), we observed that some adult individuals displayed significantly greater reproductive success. We believe that these individuals are most likely the female parents, because of the small spatial distance between them and their probable offspring. When both parents were determined, one parent was close to the juvenile plant and the other was distant. These observations suggest that pollen comes from long distances, whereas most seed germinates near the mother plant for *E. edulis*. It has been reported for the tree species *Gleditsia triacanthos* L. that, when nearby neighbors are more closely related than more distant neighbors, selection could act against pollen of nearest neighbors to reduce inbred progenies (Schnabel and Hamrick 1995). Thus, distant or nonrelated neighbors might display an increased reproductive success by pollen flow.

The detailed paternity investigation carried out herein detected an unexpected long-distance gene flow in a tropical tree species. The number of microsatellite loci used had a very high power of paternity exclusion (>99.999%; i.e., the power with which the loci together exclude an erroneously alleged tree of being the parent of an offspring; Gaiotto et al. 2001).

In a previous study using isozymes, Nason et al. (1998), detected a pollen dispersal of 5.8–14.2 km in species of the *Moraceae* family in Panama. In our study we detected gene flow at a distance of 22 km, basing this distance on high likelihood ratios for parentage. Because both maternal and paternal genotypes were unknown, it could not be determined whether the long-distance gene flow is due to seed or pollen dispersion. The adult PNB population plants determined to be probable parents of the juveniles in the IBGE population had a high probability of paternity or maternity even when a prior probability of .0001 was used (Table 4). Although the most likely pollinator for *E. edulis* species is a little bee (*Trigona spinipes* Fabricius) that has a short flight (50–100 m), a hypothesis has been proposed that pollen of *Euterpe edulis* may also be carried by the wind (Reis et al. 1993). Furthermore, alternative pollen dispersors might exist with much longer flight, such as birds. As many as 22 bird species have been recently described as potential seed dispersers of *E. edulis*, including species of Toucans (*Ramphastidae*), Trogons (*Trogonidae*), and Cracids (*Cracidae*) (Pizo and Simao 2001). In fact, paternity studies in angiosperms have revealed that individuals from outside the population frequently father significant portions of the seed crop of a target population (Devlin and Ellstrand 1990).

We could not determine the second parent for these two juveniles involved in this long-distance gene flow event. Although pollen movement is a possibility, because of the geography of the region we believe that most likely the distant parent is in fact the seed parent. Still, this is possibly the most distant gene flow event so far detected for a tropical tree. It is possible that such long-distance gene flow events are more frequent than expected in most tropical tree populations, and that only now do they begin to be documented, because more powerful genetic tools are available for population studies.

In conclusion, we have described the patterns of distribution of genetic variation and gene flow for heart of palm in Cerrado of central Brazil. Three main features were observed that are relevant to the conservation and prebreeding of heart of palm: (1) a high probability of full sibship within open-pollinated families, (2) a low but significant level of interpopulation genetic variation, and (3) the occurrence of gene flow at very long distances. Thus, although most of the genetic variation is found within populations, only a few pollen parents might be represented within a particular open-pollinated family. The establishment and long-term management of in situ genetic reserves could therefore concentrate efforts in a relatively small number of populations. However, for ex-situ conservation or the establishment of a breeding program, ideally open-pollinated seeds from several hundreds of trees should be collected so as to maximize genetic sampling from a larger number of pollen parents and therefore increase the genetic base of the population.

Acknowledgments

We thank Marcelo Bramucci, Scott J. Westenberger, and Kleyne Cristina for their invaluable help in the field. We also acknowledge the support given by the Brazilian National Research Council CNPq through PADCT grant #6250059.97-4 to D.G., a Doctoral Fellowship to F.A.G., and a Research Fellowship to D.G. and R.V.

References

Brondani RPV and Grattapaglia D, 2001. A simple and cost effective method to synthesize a fluorescent labeled internal DNA standard for fragment sizing in an automatic sequencer. Biotechniques 31:802–810.

Budowle B and Monson KL, 1993. The forensic significance of various reference population databases for estimating the rarity of variable number of tandem repeat (VNTR) loci profiles. In: DNA fingerprinting: state of the science (Pena SDJ, Chakraborty R, Epplen JT, and Jeffreys AJ, eds). Berlin: Springer Verlag; 177–191.

Chase MR, Kesseli R, and Bawa KS, 1996a. Microsatellite markers for population and conservation genetics of tropical trees. Am J Bot 83:51–57.

Chase MR, Moller C, Kesseli R, and Bawa KS, 1996b. Distant gene flow in tropical trees. Nature 383:398–399.

Collevatti RG, Grattapaglia D, and Hay JD, 2001. High resolution microsatellite based analysis of the mating system allows the detection of significant biparental inbreeding in *Caryocar brasiliense*, an endangered tropical tree species. Heredity 86:60–67.

Dayanandan S, Dole J, Bawa K, and Kesseli R, 1999. Population structure delineated with microsatellite markers in fragmented populations of a tropical tree, *Carapa guianensis* (Meliaceae). Mol Ecol 8:1585–1592.

Devlin B and Ellstrand NC, 1990. The development and application of a refined method for estimating gene flow from angiosperm paternity analysis. Evolution 44:248–259.

Doyle JJ and Doyle JL, 1987. Isolation of plant DNA from fresh tissue. Focus 12:13–15.

Eguiarte LE, Perez-Nasser N, and Piñero D, 1992. Genetic structure, outcrossing rate and heterosis in *Astrocaryum mexicanum* (tropical palm): implications for evolution and conservation. Heredity 69:217–228.

Gaiotto FA, Bramucci M, and Grattapaglia D, 1997. Estimation of outcrossing rate in a breeding population of *Eucalyptus urophylla* with dominant RAPD and AFLP markers. Theor Appl Genet 95:842–849.

Gaiotto FA, Brondani RPV, and Grattapaglia D, 2001. Microsatellite markers for Heart of Palm—*Euterpe edulis* and *E. oleracea* Mart. (Arecaceae). Mol Ecol Notes 1/2:86–88.

Galetti M and Fernandez C, 1998. Palm heart harvesting in the Brazilian Atlantic Forest: changes in industry structure and the illegal trade. J Appl Ecol 35:294–301.

Gjertson DW and Morris JW, 1995. Assessing probability of paternity and the product rule in DNA systems. Genetica 96:89–98.

Goodman SJ, 1997. R_{ST} Calc: a collection of computer programs for calculating estimates of genetic differentiation from microsatellite data and determining their significance. Mol Ecol 6:881–885.

Haberl M and Tautz D, 1999. Comparative allele sizing can produce inaccurate allele size differences for microsatellites. Mol Ecol 8:1347–1350.

Hamrick JL and Nason JD, 1996. Consequences of dispersal in plants. In: Population dynamics in ecological space and time (Rodhes OEJ, Chesser RK, and Smith MH, eds). Chicago, IL: University of Chicago Press; 203–236.

James T, Vege S, Aldrich P, and Hamrick JL, 1998. Mating system of three tropical dry forest tree species. Biotropica 30:587–594.

Kaufman SR, Smouse PE, and Alvarez-Buylla ER, 1998. Pollen-mediated gene flow and differential male reproductive success in a tropical pioneer tree, *Cecropia obtusifolia* Bertol. (Moraceae): a paternity analysis. Heredity 81:164–173.

Lewis PO and Zaykin D, 2000. Genetic data analysis, version 1.0 (d15): computer program for the analysis of allelic data. Free program distributed

by the authors over the Internet from the GDA home page at http://alleyn.eeb.uconn.edu/gda/. Visited on March 6, 2003.

Nason JD, Herre EA, and Hamrick JL, 1998. The breeding structure of a tropical keystone plant resource. Nature 391:685–687.

O'Malley DM, Buckley DP, Prance GT, and Bawa KS, 1988. Genetics of Brazil nut (*Bertholletia excelsa* Humb. & Bonpl.: Lecythidaceae), 2. mating system. Theor Appl Genet 76:929–932.

Orlande T, Laarman J, and Mortimer J, 1996. Palmito sustainability and economics in Brazil's Atlantic coastal forest. For Ecol Manag 80:257–265.

Pizo MA, and Simão I, 2001. Seed deposition patterns and the survival of seeds and seedlings of the palm *Euterpe edulis*. Acta Oecologica 22:229–233.

Reis MS, Guerra MP, and Nodari RO, 1997. Management of natural populations and maintenance of genetic diversity of *Euterpe edulis*, the heart-of-palm tree. In: Recent advances in biotechnology for tree conservation and management. Florianópolis, Brazil: Sociedade Brasileira de Silvicultura SBS; 145–156.

Reis MS, Guimarães E, and Oliveira GP, 1993. Estudos preliminares da biologia reprodutiva do palmiteiro (*Euterpe edulis*) em mata residual do estado de São Paulo. In: 7° Congresso Florestal Brasileiro, Curitiba Anais; 358–360.

Ritland K, 1989. Correlated matings in the partial selfer *Mimulus guttatus*. Evolution 43:848–859.

Ritland K, 1990. A series of FORTRAN computer programs for estimating plant mating systems. J Hered 81:235–237.

Ritland K and Jain SK, 1981. A model for estimation of outcrossing rate and gene frequencies using *n* independent loci. Heredity 47:35–52.

Rousset F, 1996. Equilibrium values of measure of population subdivision for stepwise mutation process. Genetics 142:1357–1362.

Schnabel A, and Hamrick JL, 1995. Understanding the population genetic structure of *Gleditsia triacanthos* L.: the scale and pattern of pollen gene flow. Evolution 49:921–931.

Schnabel A, Nason JD, and Hamrick JL, 1998. Understanding the population genetic structure of *Gleditsia triacanthos* L.: seed dispersal and variation in female reproductive success. Mol Ecol 7:819–832.

Slatkin M, 1995. A measure of population subdivision based on microsatellite allele frequencies. Genetics 139:457–462.

Smouse PE and Chakraborty R, 1986. The use of restriction fragment length polymorphism in paternity analysis. Am J Hum Genet 38:918–939.

Terborgh J, 1986. Keystone plant in a tropical forest. In: Conservation biology: the science of scarcity and diversity (Soulé ME, ed). Sunderland, MA: Sinauer.

Weir BS, 1996. Genetic data analysis II. Sunderland, MA: Sinauer.

Wright S, 1965. The interpretation of population structure by *F*-statistics with special regard to systems of mating. Evolution 19:395–420.

Received July 29, 2002
Accepted May 30, 2003

Corresponding Editor: Irwin L. Goldman