
**Rudimar Conte, Mauricio Sedrez dos Reis, Adelar Mantovani, and Roland Vencovsky**

Monsanto do Brasil, Uberlandia Research Station, Rodovia Uberlandia/Araxa, BR 452, Km 149, 38405-232, Uberlandia, MG—Brazil, (Conte); Universidade Federal de Santa Catarina, Departamento de Fitotecnia, Rodovia Admar Gonzaga, 1346, CP 476, 88040-900, Florianópolis, SC, Brazil (Sedrez dos Reis); Universidade do Estado de Santa Catarina, Departamento de Engenharia Florestal, Av. Luiz de Camões, 2090, 88520-000, Lages, SC, Brazil (Mantovani); and Universidade de São Paulo, Escola Superior de Agricultura “Luiz de Queiroz”, Departamento de Genética, Av. Padua Dias, 11, CP 83, 13400-970, Piracicaba, SP, Brazil (Vencovsky).

Address correspondence to Rudimar Conte at the address above, or e-mail: rudimar.conte@monsanto.com.

**Abstract**

A comparative study between microsatellite and allozyme markers was conducted on the genetic structure and mating system in natural populations of *Euterpe edulis* Mart. Three cohorts, including seedlings, saplings, and adults, were examined in 4 populations using 10 allozyme loci and 10 microsatellite loci. As expected, microsatellite markers had a much higher degree of polymorphism than allozymes, but estimates of multilocus outcrossing rate \( \hat{t}_m \approx 1.00 \), as well as estimates of genetic structure \( F_{IS}, GST \), were similar for the 2 sets of markers. Estimates of \( R_{ST} \) for microsatellites, were higher than those of \( GST \), but results of both statistics revealed a close agreement for the genetic structure of the species. This study provides support for the important conclusion that allozymes are still useful and reliable markers to estimate population genetic parameters. Effects of sample size on estimates from hypervariable loci are also discussed in this paper.

In the last 4 decades, the main genetic markers used in population studies in most species have been allozymes (Hedrick 1999). Although the number of molecular markers (DNA markers) has increased significantly in recent years, allozymes still continue to be widely used in population genetic studies (Sebbenn et al. 2000; Mariot et al. 2002; Lumaret et al. 2004; Cuevas et al. 2006; Haag et al. 2006; Mantovani et al. 2006). An important aspect of these markers is related to their neutrality. Although most studies using allozyme markers assume selective neutrality, some studies have pointed out the existence of adaptive value associated with protein polymorphism (Eanes 1999; Conte et al. 2003; Borrell et al. 2004; Araki et al. 2005; Volis et al. 2005). This fact has generated some controversy about the usefulness of these markers in population genetics because natural selection may influence the maintenance and population organization of genetic polymorphism. However, there is evidence that only a few enzymes are actually under selection and that the rest must be neutral or nearly neutral (Eanes 1999; Thelen and Allendorf 2001; Volis et al. 2005).

The recent development of techniques for the examination of hypervariable regions of the genome has significantly raised the number of applications of genetic markers (Haig 1998; Zane et al. 2002). Currently, microsatellites or simple sequence repeats (SSRs) are a popular tool in population and conservation genetics (Collevatti et al. 2001; Gao et al. 2002; Gaiotto et al. 2003; Dutech et al. 2004). Because most microsatellites are located in noncoding regions and no direct evidence exists for a general role of SSRs in eukaryotic genomes, they are more likely to be neutral than functional markers such as allozymes (Queller et al. 1993; Oliveira et al. 2006). However, evidence for nonneutrality at particular microsatellite loci is accumulating (Larsson et al. 2007). The use of microsatellites has made a great contribution toward the improvement of genetic parameter estimates, especially in populations possessing insufficient diversity to be assessed with other techniques (Oliveira et al. 2006).

The most important difference between the 2 categories of markers is their level of polymorphism. Microsatellites are much more polymorphic than allozymes in natural
populations, with more alleles, and heterozygosities are often greater than 0.5 (Estoup et al. 1998). Mutation rates at microsatellite loci have been estimated to be in the order of $5 \times 10^{-5}$ to $5 \times 10^{-3}$ (Amos et al. 1996), that is 2–4 orders of magnitude higher than at allozyme loci (Estoup et al. 2002). Comparative analyses of microsatellite and allozyme markers in natural populations have revealed some divergences between the 2 markers, and some advantages have been pointed out for microsatellites (Estoup et al. 1998; Streiff et al. 1998; Gao et al. 2002). In particular, the higher level of variation at microsatellite loci suggests that they should be more powerful than allozymes to detect changes in effective population size, genetic structure, and outcrossing rates. However, according to Hedrick (1999), data from highly variable loci must be carefully evaluated because the information given may be quite different than that obtained from less variable markers.

Heart of palm (*Euterpe edulis* Mart.; Arecaceae) is a native shade-tolerant palm of the Brazilian Atlantic forest which is economically valuable because it is the main species used for palm heart extraction. Genetic studies using microsatellite and allozyme markers in natural populations have demonstrated high values of outcrossing rates and gene flow for the species together with high levels of genetic variability within populations and low interpopulation genetic divergence (Reis et al. 2000c; Conte et al. 2003; Gaiotto et al. 2003). In natural populations, the species presents a high density and a J-shaped age structure (Reis et al. 1996, 2000a). It is a monoeccious species with strongly protandrous flowers pollinated by insects (Mantovani and Morelato 2000). Currently, natural populations are intensely fragmented and reduced in its area of occurrence (Reis et al. 2000b). Although a sustainable management system has been proposed for the species (Reis et al. 2000a), the intensive exploitation occurring since the 1960’s has resulted in the elimination of several populations and serious alterations to remaining populations (Galetti and Fernandez 1998).

In this study, we combined analyses of microsatellite and allozyme loci in order to compare the genetic structure and mating system of *E. edulis* Mart. populations. Our working hypothesis was based on the assumption that, as the SSR loci are highly multiallelic, they improve resolution over allozymes for the estimation of population genetic parameters.

**Material and Methods**

**Populations and Sampling**

Four natural populations of *E. edulis* were surveyed in the districts of São Pedro de Alcântara and Ibirama, Santa Catarina, Brazil (Figure 1). The vegetation of the study sites is defined as “Evergreen Atlantic Tropical Forest.” In spite of the current fragmentation and degradation of the populations of *E. edulis*, many remnant fragments exist that are well conserved, as is the case of the populations here investigated. Also, the fragments still retain connectivity due to their mutual proximity, such that disturbed populations will benefit from the influence of the undisturbed populations in the surrounding area.

To study variability and genetic structure, leaf material from 3 age cohorts was sampled randomly in the 4 populations as follows: 1) Seedlings—insertion height of the youngest leaf less than 10 cm, corresponding to the individuals of the last reproductive period; 2) Saplings—insertion height from 30 cm to 1 m, absence of exposed stem and the presence of 4–5 mature leaves; and 3) Adults—plants usually higher than 10 m of height, presence of open inflorescence and/or abscission scars on the stem from previous events of reproduction. A total of 50 individuals were sampled per cohort in each population. For the mating system study, open-pollinated seeds were collected from 15 randomly selected adult trees in São Pedro (Population 4). Seeds were germinated and grown in a greenhouse under controlled temperature (20–35 °C) and natural photoperiod. Progeny arrays of 13 individuals per open-pollinated family were used to determine outcrossing rates.

**Genetic Analysis**

In total, 10 enzymatic systems were used: $\alpha$-esterase ($\alpha$-EST, EC 3.1.1.1), shikimic dehydrogenase (*SKDH*, EC 1.1.1.25), peroxidase (*PRX*, EC 1.11.1.7), 6-phosphogluconate dehydrogenase (*PGDH*, EC 1.1.1.44), nicotinamide adenine dinucleotide dehydrogenase (*NADH*DH, EC 1.6.59.3), malate dehydrogenase (*MDH*, EC 1.1.1.37), phosphoglucoisomerase (*PGI*, EC 5.3.1.9), isocitrate dehydrogenase (*IDH*, EC 1.1.1.42), phosphoglucomutase (*PGM*, EC 5.4.2.2), and glucose-6-phosphate dehydrogenase (*G6PDH*, EC 1.1.1.49). Among the 10 systems employed, 16 zones with enzymatic activity were observed, revealing 10 presumed polymorphic loci. Each zone that showed a behavior apparently independent of the others and with Mendelian
segregation was considered an allozyme locus, taking into account the possible quaternary structure of each enzyme in question (Kephart 1990). Only the 10 polymorphic loci were used in the genetic analyses.

A group of 10 SSR loci, previously developed and optimized for *E. edulis* (Gaiotto et al. 2001), were used to genotype all sampled individuals. The total reaction volume used in polymerase chain reaction was 13 μl, containing 7.5 ng of genomic DNA; 50 mM KCl, 20 mM Tris–HCl pH 8.4; 10% dimethyl sulfoxide; 1.5 mM MgCl₂; 250 μM of each deoxynucleoside triphosphates; 0.3 μM of each primer (forward and reverse); 1.0 unit of Taq DNA polymerase (Ludwig Biotecnologia Porto Alegre, Brazil); and ultrapure water to complete the final volume. Amplifications were performed using a PTC-100 thermal controller (MJ Research, Waltham, MA) with the following conditions: 94 °C for 1 min, 30 cycles of 94 °C for 1 min, the primer-specific annealing temperature (Gaiotto et al. 2001) for 1 min, 72 °C for 1 min, ending with 72 °C for 7 min. The amplified fragments were separated in 4% polyacrylamide gel in a run with 1 × Tris/Borate/EDTA at 60 W for 1 h, using 10-bp ladder size standard, and stained with silver nitrate (Creste et al. 2001).

### Data Analysis

Molecular genetic variability was described in terms of average number of alleles per locus (\(A\)), observed heterozygosity (\(H_o\)), Nei’s (1978) gene diversity (\(H_e\)), and fixation index (\(f\)) using the GDA program (Lewis and Zaykin 2001). The fixation index (\(f\) for \(F_{ST}\)) was estimated for each population by analysis of variance (ANOVA), and the significance was tested by bootstrapping over loci using 10 000 replicates in order to estimate interlocus variation. Genotypic frequencies of each locus were subjected to a goodness-of-fit test (Fisher’s exact test) to expected Hardy–Weinberg proportions, as defined by Weir (1996), using the GDA program (Lewis and Zaykin 2001). Such tests were performed by the conventional Monte Carlo method using 10 batches with 1000 permutations per batch.

Molecular genetic structure among populations was investigated by the analysis of gene diversity in subdivided populations (Nei 1973). Parameters \(H_o\), \(H_e\), \(D_{ST}\), and \(G_{ST}\) were estimated using the FSTAT program (Goudet 2002), and the null hypothesis was tested by bootstrapping over loci using 10 000 replicates. As most mutations within microsatellites involve the addition or subtraction of a small number of repeat units, the mutation process is not in line with expectations under an infinite allele model with low mutation rates. Therefore, an analogue of the \(F_{ST}\) statistics, namely the \(R_{ST}\) parameter (Slatkin 1995), developed for microsatellite data, was also used to quantify genetic structure. ANOVA of allele size followed Goodman (1997) using the FSTAT program (Goudet 2002).

The mating system of *E. edulis* was analyzed under the mixed mating model described by Ritland and Jain (1981) using the MLTR program (Ritland 1997). This procedure estimates the multilocus outcrossing rate (\(t_{m}\)), the average single-locus outcrossing rate (\(t_{a}\)), the outcrossing rate between related individuals (\(t_{m} - t_{s} \neq 0\)), the average single locus inbreeding coefficient of maternal parents (\(f\)), the correlation of outcrossing rate within progeny arrays (\(f_{p}\)), the correlation of outcrossed paternity within progeny arrays (or the probability that randomly chosen progeny pairs from the same array are full sibs) (\(f_{p}\)), and the pollen (\(p\)) and ovule (\(o\)) gene population frequencies. Standard errors (SEs) were estimated by bootstrapping over progenies using 10 000 replicates.

### Results

#### Number of Alleles at Allozyme and Microsatellite Loci

From the allozyme analysis, a total of 41 alleles at 16 loci were identified in the 3 cohorts of the 4 populations. From these loci, 6 were monomorphic (\(\alpha\)-Est-1, \(\alpha\)-Est-2, Pgdh-2, Nadhdh-1, Mdh-2, and Idh-1) and 10 were polymorphic (Skdh-1, Prx-2, Prx-3, Prx-4, Prx-5, Pgdh-2, Mdh-1, Pgi-2, Pgm-1, and G6pdh-1). The locus Prx-5 showed a very low level of polymorphism, with the most frequent allele close to fixation in all cohorts studied. However, some loci showed high levels of polymorphism (Prx-3, Prx-4, Pgm-1, with 4 alleles each, and Pgdh-2, with 5 alleles). All 10 microsatellite loci displayed high levels of polymorphism and a total of 161 alleles were described. The most variable loci (\(EE48\) and \(EE47\)) displayed 24 and 26 alleles, respectively, whereas the least variable (\(EE5\)) displayed 13 alleles.

#### Genetic Variability

The estimates of genetic variation within populations for allozymes and microsatellite markers are presented in Table 1. An average of 52 individuals was analyzed per cohort in each population. As expected, the values of \(A\), \(H_o\), and \(H_e\) were much higher for microsatellites than allozymes. For allozymes, the following diversity estimates were obtained among the cohorts: \(A\): 3.05–3.15; \(H_o\): 0.416–0.431; \(H_e\): 0.378–0.403. For microsatellites, the estimates were as follows: \(A\): 14.12–14.72; \(H_o\): 0.781–0.785; \(H_e\): 0.678–0.709.

Although considerable variation was observed among populations, both markers displayed similar trend within populations (Table 1). In general, both markers displayed lower \(f\) values for populations 3 and 4. For allozymes, the \(H_o\) values of the adult individuals were higher than the values of seedlings and saplings, mainly for the loci G6pdh-1, Pgdh-2, Prx-2, and Prx-3. Permutation analysis (Goudet 2002) revealed that the difference between the \(H_o\) values of the adults and seedlings sampled in Population 4 was significantly different from 0 (\(P < 0.05\)).

Departures from Hardy–Weinberg equilibrium were observed in a greater extent for microsatellites than allozymes, using Fisher’s exact test. Whereas allozymes presented an average of 70% of the loci at equilibrium for most populations, the average for microsatellites was 50%; some populations, however, presented values which were even lower for seedlings and saplings (37.5%). On the other hand, as Fisher’s exact test was performed for several loci, the number of loci where the null hypothesis was rejected...
Table 1. Estimates of genetic parameters for seedlings, saplings, and adults in 4 populations of *Euterpe edulis* based on 10 polymorphic allozymes and 10 microsatellite loci

<table>
<thead>
<tr>
<th>Cohorts/populations</th>
<th>Allozymes</th>
<th>Microsatellites</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\hat{A}$</td>
<td>$\hat{H}_e$</td>
</tr>
<tr>
<td>Seedlings</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Population 1</td>
<td>3.10</td>
<td>0.428</td>
</tr>
<tr>
<td>Population 2</td>
<td>3.00</td>
<td>0.441</td>
</tr>
<tr>
<td>Population 3</td>
<td>3.00</td>
<td>0.377</td>
</tr>
<tr>
<td>Population 4</td>
<td>3.10</td>
<td>0.419</td>
</tr>
<tr>
<td>Average</td>
<td>3.05</td>
<td>0.416</td>
</tr>
<tr>
<td>Saplings</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Population 1</td>
<td>3.20</td>
<td>0.437</td>
</tr>
<tr>
<td>Population 2</td>
<td>3.10</td>
<td>0.430</td>
</tr>
<tr>
<td>Population 3</td>
<td>3.00</td>
<td>0.401</td>
</tr>
<tr>
<td>Population 4</td>
<td>3.30</td>
<td>0.454</td>
</tr>
<tr>
<td>Average</td>
<td>3.15</td>
<td>0.431</td>
</tr>
<tr>
<td>Adults</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Population 1</td>
<td>3.00</td>
<td>0.434</td>
</tr>
<tr>
<td>Population 2</td>
<td>3.10</td>
<td>0.402</td>
</tr>
<tr>
<td>Population 3</td>
<td>3.00</td>
<td>0.410</td>
</tr>
<tr>
<td>Population 4</td>
<td>3.40</td>
<td>0.452</td>
</tr>
<tr>
<td>Average</td>
<td>3.12</td>
<td>0.424</td>
</tr>
</tbody>
</table>

$\hat{A}$, average number of alleles per locus; $\hat{H}_e$, expected heterozygosity; $\hat{H}_o$, observed heterozygosity; $\hat{f}$ = $F_{IS}$, fixation index;

*significant (confidence interval 95%: 10 000 bootstraps).

decreased when the Bonferroni correction was applied (Weir 1996).

**Genetic Structure**

The analysis of gene diversity in subdivided populations (Nei 1973) showed a high congruence between the 2 markers, both of which revealed low molecular genetic divergence among populations. On average, the estimated values of $G_{ST}$ for allozyme loci were low (0.023, 0.017, and 0.011 for seedlings, saplings, and adults, respectively), but significantly different from 0. Similarly, for microsatellites estimates of $G_{ST}$ were also low (0.024, 0.021, and 0.028, respectively) but significantly different from 0. In the pairwise comparisons, the estimated values of $G_{ST}$ varied from 0.013 to 0.038 for microsatellites and were significantly different than 0 in all comparisons performed with this marker. For allozymes, more variation of $G_{ST}$ was observed among the different comparisons (0.008–0.041), and the majority were not significantly different from 0, especially for populations of the same region.

The estimates of $R_{ST}$ over all populations ranged from 0.033 to 0.048 among the 3 cohorts. In the pairwise comparisons, these estimates ranged from 0.004 to 0.091 and mostly were significantly different from 0. Despite the fact that the estimates of $R_{ST}$ were higher than those of $G_{ST}$, results of both statistics revealed a close agreement, indicating that the divergence was mostly among regions.

**Mating system**

SEs of pollen and ovule gene frequencies were low for both markers (data not shown), indicating that the sample size used was adequate for the estimates of mating system parameters. Most differences in allele frequencies between pollen and ovules were not significant on the basis of SEs, indicating homogeneity of the pollen pool by each maternal tree. A great congruence between the 2 markers was observed in the estimates of mating system parameters (Table 2). The multilocus outcrossing rate ($\hat{t}$) was equal to 0.996 (0.006) for microsatellites and 1.000 (0.000) for allozymes, revealing preferentially allogamic behavior of *E. edulis*. The single locus outcrossing rate ($\hat{t}_s$) was high (0.987 for allozymes and 0.929 for microsatellites), but significantly different from the multilocus estimates, especially for microsatellites, suggesting that crosses among related individuals had occurred in the population. The fixation index for maternal genotypes ($\hat{f}$) was null in both markers, indicating the absence of inbreeding. The probability that a randomly chosen pair of progenies from the same array were full sibs ($\hat{F}_{IS}$) was equal to 0.109 (0.000) for allozymes and 0.103 (0.012) for microsatellites, suggesting that the effective number of pollen parents ($N_{ep} = 1/\hat{f}_s$) is somewhat less than 10 individuals. The normalized variation of outcrossing rate among progeny arrays ($\hat{F}_s$) was low and nonsignificant for both markers, which is in agreement with the floral biology of this strongly protandrous species.

**Discussion**

The 10 allozyme and 10 SSR loci used for genetic analysis in this study detected high levels of molecular genetic variation, confirming the results obtained in previous studies with *E. edulis* populations with allozymes (Reis et al. 2000c;
Table 2. Estimates of the mating system parameters for *Euterpe edulis*

<table>
<thead>
<tr>
<th></th>
<th>Allozymes</th>
<th>Microsatellites</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_m$</td>
<td>1.000 (0.000)</td>
<td>0.996 (0.006)</td>
</tr>
<tr>
<td>$t_s$</td>
<td>0.987 (0.001)</td>
<td>0.929 (0.018)</td>
</tr>
<tr>
<td>$t_m - t_s$</td>
<td>0.013 (0.001)</td>
<td>0.067 (0.017)</td>
</tr>
<tr>
<td>$f$</td>
<td>0.002 (0.006)</td>
<td>0.008 (0.009)</td>
</tr>
<tr>
<td>$r_s$</td>
<td>0.109 (0.000)</td>
<td>0.105 (0.012)</td>
</tr>
<tr>
<td>$r_f$</td>
<td>0.008 (0.009)</td>
<td>0.027 (0.034)</td>
</tr>
</tbody>
</table>

$t_m$, multilocus outcrossing rate; $t_s$, average single locus outcrossing rate; $f$, average single locus inbreeding coefficient of maternal parents; $r_s$, correlation of outcrossing rate within progeny arrays; $r_f$, correlation of outcrossing rate within progeny arrays; and SEs based on 10 000 bootstraps are in parentheses.

The results from both markers revealed that most of the molecular genetic variability is distributed within populations, confirming the results obtained in other studies with this species (Reis et al. 2000c; Conte et al. 2003; Gaiotto et al. 2003). The small amount of population genetic differentiation ($G_{ST}, R_{ST}$) is also consistent with the high rate of outcrossing of this species and the localized distribution of the study populations. Moreover, our results are in agreement with other studies on tropical species, where species with large geographic ranges, outcrossing breeding systems, and high rates of gene flow, have more genetic variability within populations and consequently low divergence among populations (Hamrick and Godt 1990).

Hedrick (1999) mentions that because microsatellite loci often have very high heterozygosity within populations, the magnitude of differentiation may be quite small. However, similar magnitude of genetic divergence was also observed using allozyme markers, which demonstrates the consistency of the genetic information obtained with these 2 markers. Similarly, an examination of the estimates of molecular genetic variability, such as the number of alleles per locus and the gene diversity, revealed that all 4 populations exhibited similar levels of variability for the 3 categories studied.

The multilocus outcrossing rate ($t_m$) in the present study ($t_m = 1.000$ for allozymes and $t_m = 0.996$ for microsatellites) confirmed the results obtained in previous studies of this species ranging from 0.940 to 1.040 for allozymes (Reis et al. 2000c) and from 0.910 to 0.980 for microsatellites (Gaiotto et al. 2003). These results are all consistent with the conclusion that *E. edulis* is an obligately outcrossing species. The estimates of $t_m$ are also congruent with the floral biology and pollination ecology of the species (Mantovani and Morelatto 2000).

Despite the high rate of outcrossing found in this species, nonrandom matings have been observed by the differences between the multilocus ($t_m$) and single locus ($t_s$) estimates for both markers, indicating that crosses among related individuals had occurred in the population. The occurrence of nonrandom matings was also observed in the estimates of correlation of outcrossed paternity within progeny arrays ($r_f$), revealing a substantial number of full sibs originating from a limited pool of pollen, a possible result of limited pollen dispersal and/or nonsynchronous flowering of all individuals. The larger SEs observed for microsatellites is the result of less accurate estimates typical for highly multiallelic loci.

In summary, SSR loci showed a much higher degree of absolute polymorphism than allozymes, but relative measures of the mating system and genetic structure were quite similar for the 2 sets of markers. Both markers displayed few pronounced genetic structure of *E. edulis*, with the exception of a small increase in the inbreeding levels among seedlings and saplings of the populations 1 and 2 (private areas).

In spite of limited number of estimates provided in this study, the present results give support for the important prediction that allozymes are still useful and reliable markers to estimate population genetic parameters. Regardless of the
limitation of sample size for the sampling of all possible genotypes, microsatellites also confirmed the high content of genetic information for studies of population genetic parameters.

Funding
Coordenação de Aperfeiçoamento de Pessoal de Ensino Superior and Conselho Nacional de Desenvolvimento Científico e Tecnológico.

Acknowledgments
We thank the researchers of the Núcleo de Pesquisas em Florestas Tropicais/Universidade Federal de Santa Catarina for their assistance in collecting and organizing the samples and also the Laboratório de Fisiologia do Desenvolvimento e Genética Vegetal/Universidade Federal de Santa Catarina for the infrastructure provided in support of this work.

References


Received August 7, 2007
Accepted May 22, 2008

Corresponding Editor: Jim Hamrick