GENETIC STRUCTURE OF NATURAL POPULATIONS OF CRYPTOCARYA MOSCHATA NEES (LAURACEAE) FROM SOUTHEASTERN BRAZILIAN ATLANTIC RAIN FOREST

Pedro Luís Rodrigues de Moraes1,3 & Maria Teresa Vitral de Carvalho Derbyshire2,4

Abstract

[Genetic structure of natural populations of Cryptocarya moschata Nees (Lauraceae) from southeastern Brazilian Atlantic rain forest]. This study was accomplished on a more comprehensive basis to evaluate previous questions that were raised from a preliminary article about the genetic structure of Cryptocarya moschata populations. Thus, through the analysis of 40 polymorphic allozyme loci, allele frequencies were estimated from 335 individuals of 11 natural populations of C. moschata from six hydrographic basins of São Paulo state and Serra da Estrela, Rio de Janeiro, Brazil. Estimates of Wright’s F statistics were done through the analysis of variance, presenting average values of $\hat{F} = 0.352$, $\hat{\theta}_P = 0.285$ and $\hat{f} = 0.097$. These results indicated that individuals within populations must be panmitic, and that the diversity among populations is fairly high, being superior to what would be expected for groups of plants having a full-sib family structure. From estimates of $\hat{\theta}_P$ obtained for populations taken two at a time, the model of isolation by distance was tested; data did not fit the model, showing that $\hat{\theta}_P$ did not increase by the respective increasing of the geographic distance. The estimated gene flow of 0.55 migrants per generation corroborated the pronounced populational differentiation, indicating that drift effects should be more important than the selection ones. The effective population sizes found from the sampled populations showed that there was an adequate genetic representativeness of the samples for those with relatively low values of $\hat{f}$ . Though, under a metapopulation context, the effective population size was 17.07 individuals, indicating that sampling performed for the species corresponded to 88.44% of the maximum effective size obtained from 11 populations with a $\hat{\theta}_P$ of 0.285, equivalent to only 5.09% individuals for the total sampled. Management and conservation strategies aimed at preserving high intrapopulation genetic variation in C. moschata would imply in the maintenance of populations with great number of individuals. Moreover, for the preservation of the species as a whole, the maintenance of many such populations would be mandatorily recommended, which denotes that the conservation of large areas of Atlantic rain forest should be necessary to hold its evolutionary dynamics.

Key words: allozymes, Lauraceae, genetic structure, Neotropics, Cryptocarya moschata, Atlantic rain forest, Brazil
Resumo

Estrutura genética de populações naturais de *Cryptocarya moschata* Nees (Lauraceae) na Mata Atlântica do sudeste brasileiro. Este estudo foi realizado, de forma mais abrangente, para avaliar questões prévias que foram levantadas a partir de um artigo preliminar sobre a estrutura genética de populações de *Cryptocarya moschata*. Assim, pela análise de 40 locos alozimicos polimórficos, estimaram-se as frequências alélicas para 335 indivíduos adultos de 11 populações naturais de *Cryptocarya moschata*, de seis bacias hidrográficas do estado de São Paulo e da Serra da Estrela, no Rio de Janeiro. Obtiveram-se estimativas das estatísticas $F$ de Wright, estimadas pelo método da análise da variância, que apresentaram valores de $\hat{F} = 0,352$, $\hat{\theta}_P = 0,285$ e $\hat{f} = 0,097$, indicando que os indivíduos dentro das populações devem ser panmíticos, que a diversidade entre as mesmas é bastante alta, superior à esperada para famílias com estruturação de irmãos-completos. Calculando $\hat{\theta}_P$ com as populações tomadas duas a duas, testou-se o modelo de isolamento pela distância, sendo que os dados não aderiram ao modelo, mostrando não haver aumento de $\hat{\theta}_P$ com o respectivo aumento da distância geográfica. O fluxo gênico estimado em 0,55 indivíduos por geração corrobora a pronunciada diferenciação populacional encontrada, indicando que efeitos de deriva devem ser mais importantes do que os de seleção. Os tamanhos efetivos populacionais estimados a partir das populações amostradas mostraram que houve uma adequada representatividade genética das amostras para aquelas com valores de $\hat{f}$ relativamente baixos. Contudo, sob um contexto metapopulacional, o tamanho efetivo populacional foi de 17,07 indivíduos, indicando que a amostragem realizada para a espécie correspondeu a 88,44% do tamanho efetivo máximo obtido a partir de 11 populações com um $\hat{\theta}_P$ de 0,285, correspondendo a apenas 5,09% dos indivíduos amostrados. As estratégias de manejo e conservação necessárias para a preservação da alta variabilidade genética intrapopulacional de *C. moschata* implicam na manutenção de populações com número grande de indivíduos. Além disso, para a conservação da espécie como um todo, a manutenção de muitas populações se faz necessária, o que denota que a conservação de grandes áreas de Mata Atlântica deva ser necessária para manter sua dinâmica evolutiva.

**Palavras-chave:** alozimas, Lauraceae, estrutura genética, Neotrópico, Cryptocarya moschata, Mata Atlântica, Brasil
1. Introduction

The allozyme variability in natural populations of plants has been intensively studied (Gottlieb 1977, Nevo 1978, Brown 1979, Hamrick et al. 1979, Hamrick & Godt 1990, Rocha & Lobo 1998, Sousa 2001). In tree species, the allozyme variability is related to longevity, ample geographic distribution, and the potential for a considerable gene flow, characteristics which, when combined, can contribute to an increase of the effective population size (Sousa 2001).

According to Rocha & Lobo 1998, the knowledge of populational structure is important so that one can avoid additional losses of genetic diversity in populations which are threatened by logging activities, deforestation and habitat fragmentation (Hall et al. 1994, Nason et al. 1997). It follows that the fragmentation of areas of forest, caused by agriculture and other anthropic activities, leads to the isolation of subpopulations and, consequently, to genetic drift and endogamy (Robinson 1998).

The information about the genetic variability in natural populations is fundamental to the progress of two current issues of major interest: speciation in tropical forests and conservation of genetic resources (Buckley et al. 1988). Nevertheless, the present knowledge of the Atlantic Rain Forest (sensu lato; see Joly et al. 1999) is still rather insignificant, mainly in what concerns the populational genetic structure of its species and suitable sampling procedures aiming at conservation and management.

On a previous work, Moraes et al. (1999) had studied 214 individuals of four Cryptocarya moschata Nees natural populations hailed from two regions (south and north coast) of São Paulo state, through 18 allozyme loci. The authors pointed out that the strategies of management and conservation, which are needed for the preservation of the high intrapopulational genetic variability found in C. moschata, would entail the maintenance of populations with a high number of individuals. Furthermore, for the preservation of the species as a whole, the maintenance of a large number of populations was suggested. For those adult populations of C. moschata, the divergences obtained through \( \hat{F}_{ST} \) estimates also suggested the existence of significant genetic drift and/or natural selection effects among populations. However, the lack of divergence between neighbouring populations of the northern coastal region, which had no physical isolation, indicated weaker genetic drift and/or natural selection effects when compared to southern populations, which presented potential physical isolation. This made reasonable to consider the stepping-stone as a possible model for the genetic variability organization within this species. Moreover, the coancestry coefficient for regions did not show any additional allele frequencies differences between individuals owed to the regions component, suggesting lower levels of gene flow than that found between neighbouring populations without isolation. An estimated gene flow of 0.9 migrants per generation corroborated the pronounced divergence found among populations. Additionally, assuming that the species was compounded by infinite populations, the effective population size was 10.05 individuals from the 214 sampled, which had indicated that the sampling procedure corresponded to 73.43% of the maximum effective size obtained from four populations with a \( \hat{F}_{ST} \) of 0.140, equivalent to only 5.00% individuals for the total sampled. These previous results showed that the high value of \( \hat{F}_{ST} \) and the low of \( \hat{Nm} \) had indicated a restricted gene flow among the populations, which were distanced about ten to hundreds of kilometers from each other.

Based mainly on the results obtained by Moraes et al. (1999), a more comprehensive sampling of the species in association with an expressive large number of allozyme loci have been used in the present study to verify if the slight tendency found before, of increase in populational diversity with distance, could be explained by the model of isolation by distance per se. Another issue that has been considered was to analyze how was the improvement in the genetical representativeness of the species, taken from \( \hat{Nm} \) under a better sampling strategy.

This study aims at the characterization of the genetic structure of Cryptocarya moschata Nees populations, by means of Wright’s F statistics estimates, through the variance analysis method. Estimates of indirect gene flow for the populational set were obtained, as well as intra and metapopulational effective sizes and neighbourhood sizes. Additionally, based on unbiased \( F_{ST} \) estimate, the isolation by distance model was assessed.

2. Materials and Methods

2.1. Cryptocarya moschata Nees

Cryptocarya moschata Nees belongs to genus Cryptocarya R. Brown, subtribe Cryptocaryineae Kosterm., tribe Cryptocaryae Meissn., subfamily Laurioidae Kosterm., according to classification by Kostermans (1957) for the family Lauraceae Lindl. Reproductive individuals of C. moschata are trees 15-30 m tall x 20-104 cm dbh (diameter at breast height; Moraes et al. 1999). The flowers are bisexual, trimerous, with two series of tepals (3 + 3), stamens 6 introrses and 3 extrorses, bilocular anthers, staminodes 3, ovary subsessile, stigma inconspicuous. The nucule-type fruit is 1.45-3.06 cm long (\( \bar{x} = 2.26 \pm 0.28 \) cm; N = 1,892), 1.29-2.55 cm wide (\( \bar{x} = 1.90 \pm 0.24 \) cm; N = 1,892), one-seeded. The diaspores (pericarp + seed) are 1.34-3.00 cm long (\( \bar{x} = 2.17 \pm 0.27 \) cm; N = 1,764), 1.16-1.92 cm wide (\( \bar{x} = 1.52 \pm 0.11 \) cm; N = 1,764; Moraes & Alves 1997, Moraes & Derbyshire, M. T. V. de C. - Biota Neotropica, v4 (n1) - BN00404012004
Alves 2002, Moraes et al. 2002; from fruit collections of PECB-S). Moraes & Monteiro (2002) presented the multilocus outcrossing rate estimate ($\hat{m} = 0.884 \pm 0.034$), which indicated a predominant outcrossing mating system for the analyzed population of PECB-S. Still, from Ritland’s “sibling-pair” model (correlated mating model), the results showed that there was variation in selfing rates among different trees, also indicating that when there was inbreeding, most seeds in the trees were likely full-sibs. The chromosome number is $2n = 24$ (Moraes & Gardingo 1996). This species is rich in alkaloids, styrylpyrones, and flavonoids (Cavalheiro & Yoshida 2000). Its bark, bitter and scented, is considered to be stomachical and helpful in fighting colic and diarrhea, in folk medicine (Vattimo 1957). The tea from its seeds is used against stomachache, and its crushed leaves against aches and colic; the fruit is carminative, and “muriquis” (woolly spider monkeys, Alouatta palliata) and cracids (Brachyteles arachnoides), are widely consumed by primate populations [brown howler monkeys (Alouatta fuscata), brown capuchins (Cebus apella), and “muriquí” (woolly spider monkeys, Brachyteles arachnoides)], and cracids (Pipile jacutinga and Penelope obscura); its wood is used in canoes manufacture (Moraes et al. 2002). C. moschata has a widespread geographic distribution, mainly in the Atlantic rain forest of Brazil (Coe-Texeira 1965, Kostermans 1937, 1938, Vattimo 1966), ranging from the state of Rio Grande do Sul to Pernambuco. Species identification was done from dried specimens housed at Herbarium ESA, Piracicaba - SP, whose vouchers are listed in table 1.

### 2.2. Plant collections and sampling

A foliar stem bearing mature leaves was collected from adult trees of Cryptocarya moschata, which was packed in plastic bags, and stored in stiff foam-boxes with ice for the transportation to the laboratory. In the laboratory, they were kept at $5 \pm ^\circ C$, and the third mature leaf from the stem apex was selected and extracted for electrophoresis, after two days on average. Table 1 and figure 1 present the locations where populations were sampled, with, at least, 20 individuals being collected, whenever possible.

### 2.3. Allozymes

Through allozyme markers revealed by starch gel electrophoresis, the genetic characterization of the analyzed populations was done according to procedures described by Alfenas et al. (1991).

The choice of allozyme systems for the genetic analysis of samples was preceded by a set of tests, which was carried out at the Laboratory of Plant Improvement (LAMP - CENA/USP), with the same procedures described by Moraes & Derbyshire (2002).

Thirty one enzyme systems were assayed. Among them, nineteen were selected for populational analyses due to good resolution of bands and polymorphism. From zymograms obtained, 40 putative loci were used in genetic analyses, such as: acid phosphatase (Acp1-1, Acp1-2, Acp2-3), alkaline phosphatase (Alph-1, Alph-2), cathalase (Cat-1 and Cat-2, method of Alfenas et al. 1991), cathalase (Cat-1, Cat-2 and Cat-3, method of Scandalias 1965), diaphorase (Dia-1, Dia-2, Dia-3), $\alpha$-esterase ($\alpha$-Est-1, $\alpha$-Est-2), $\beta$-esterase ($\beta$-Est-1, $\beta$-Est-2), glutamic dehydrogenase (Gdh), glutamate-oxaloacetate transaminase (Got), isocitrate dehydrogenase (Idh), malic dehydrogenase (Md1-1, Md1-2), nicotinamide adenine dinucleotide dehydrogenase (Nad1dh), nicotinamide adenine dinucleotide phosphate dehydrogenase (Nadphd1), 6-phosphogluconic dehydrogenase (6Pgd1h-1, 6Pgd1h-2, 6Pgd1h-3), phosphoglucose isomerase (Pgi-1, Pgi-2), polyphenoloxidase (Ppo-1, Ppo-2), peroxidase (Pox-1, Pox-2, Pox-3, Pox-5), shikimate dehydrogenase (Skdh-1, Skdh-2), and superoxide dismutase (Sod-1, Sod-2 and Sod-3). Different loci for the same enzyme were sequentially numbered, with the most anodally migrating locus being given the lowest number, and so on. Likewise, within each locus the most frequent migrating band was designated allele 1, and each successively slower band was numbered 2, 3, etc.

### 2.4. Data analyses

From the interpretation of zymograms and the following definition of genotypes for each individual analysed, the allelic frequencies of each population studied were calculated and also the main genetic variability measures: $A = \text{mean number of alleles per locus}$, $\theta = \text{effective number of alleles per locus}$, $P = \text{percent polymorphic loci (95\% criterion)}$, and $H = \text{gene diversity or expected heterozygosity (Nei 1987)}$.

Deviations from Hardy-Weinberg expectations (HWE) were described using fixation indexes ($\hat{F}_{is}$, Wright 1978) for each population, and their significance was evaluated by exact tests using a conventional Monte Carlo method, as proposed by Guo & Thompson (1992), performed by TFGPA 1.3 program (Miller 1997).

Nested analyses of molecular variance were done for a three-level sampling hierarchy, through the program GenAlEx v.5.04 (Peakall & Smouse 2001), to verify relatedness of pairs of alleles by virtue of being within individuals, between individuals within populations, or between populations within regions. The three degrees of relatedness were quantified by $F_{is}$ and $\theta_S$, respectively, with these parameters giving the measures of differentiation among individuals, populations, or regions, as presented by Weir (1996). Several possibilities for grouping the populations into regions were tested under different criteria, as by different hydrographic basins (Alto Paranapanema: popula-
Table 1: Sampling locations of individuals of Cryptocarya moschata Nees with their respective number of analyzed trees through allozymes and vouchers housed at Herbarium ESA. Location codes in parentheses.

<table>
<thead>
<tr>
<th>Location</th>
<th>Municipality</th>
<th>Nr. trees</th>
<th>Vouchers (Nr. ESA)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parque Estadual Carlos Botelho, Sede (PECB-S)</td>
<td>São Miguel Arcanjo</td>
<td>34</td>
<td>Moraes 31 (5722), 109 (5721), 198 (5736), 690 (7519), 758 (7837), 1010 (15977), 1018 (15970), 1214 (25900), 2217 (67775) and 2523 (76802)</td>
</tr>
<tr>
<td>Parque Estadual Carlos Botelho, Sete Barras (PECB-SB)</td>
<td>Sete Barras</td>
<td>30</td>
<td>Moraes 2340 (68797)</td>
</tr>
<tr>
<td>Parque Estadual Carlos Botelho, Rio Temível (PECB-T)</td>
<td>division of São Miguel Arcanjo and Sete Barras</td>
<td>28</td>
<td></td>
</tr>
<tr>
<td>Parque Estadual Carlos Botelho, Raizão (PECB-R)</td>
<td>Sete Barras</td>
<td>30</td>
<td>Moraes 2434 (71645) and 2435 (71644)</td>
</tr>
<tr>
<td>Parque Estadual Intervales, Núcleo Saibadela (PEI-NS)</td>
<td>Sete Barras</td>
<td>36</td>
<td>Moraes 2409 (69843), 2410 (69844), 2411 (69845), 2412 (69846) and 2414 (69847)</td>
</tr>
<tr>
<td>Parque Estadual da Cantareira (PEC)</td>
<td>Mairiporã</td>
<td>43</td>
<td>Moraes 2416 (70233)</td>
</tr>
<tr>
<td>Estação Experimental de Pariguera-Açu (EEPA)</td>
<td>Pariguera-Açu</td>
<td>34</td>
<td>Moraes 2141 (66229) and 2419 (70396)</td>
</tr>
<tr>
<td>Estação Ecológica Juréia-Itatins, Núcleo Arpoador (EEJI-NA)</td>
<td>Peruíbe</td>
<td>45</td>
<td>Moraes 2439 (71512), 2442 (71505), 2443 (71504), 2444 (71508), 2445 (71507) and 2448 (71513)</td>
</tr>
<tr>
<td>Serra da Estrela (SE)</td>
<td>Petrópolis</td>
<td>12</td>
<td>Moraes 2157 (67396), 2452 (71629) and 2453 (71630)</td>
</tr>
<tr>
<td>Parque Estadual da Serra do Mar, Núcleo Picinguaba (PESM-NP)</td>
<td>Ubatuba</td>
<td>21</td>
<td>Moraes 2471 (72463), 2473 (72461) and 2475 (72459)</td>
</tr>
<tr>
<td>Parque Estadual da Serra do Mar, Núcleo Santa Virgínia (PESM-NSV)</td>
<td>São Luiz do Paraitinga</td>
<td>22</td>
<td>Moraes 1241 (67403), 2478 (72456) and 2479 (72467)</td>
</tr>
</tbody>
</table>
Table 2: Estimates of genetic variability and heterozygosity averaged over 40 allozyme loci for 11 natural populations of Cryptocarya moschata Nees (Lauraceae). SD in parentheses.

<table>
<thead>
<tr>
<th>Population</th>
<th>Mean sample size/locus</th>
<th>Mean alleles/locus</th>
<th>Effective nr. alleles</th>
<th>% Loci polymorphic&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Mean heterozygosity&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Fixation index&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Observed</td>
<td>H-W exp.</td>
</tr>
<tr>
<td>PECB-S</td>
<td>64</td>
<td>2.40 (0.87)</td>
<td>1.73 (0.56)</td>
<td>87.80</td>
<td>0.343 (0.308)</td>
<td>0.369 (0.198)</td>
</tr>
<tr>
<td>PECB-SB</td>
<td>56</td>
<td>2.08 (0.67)</td>
<td>1.69 (0.56)</td>
<td>78.05</td>
<td>0.289 (0.345)</td>
<td>0.348 (0.226)</td>
</tr>
<tr>
<td>PECB-T</td>
<td>55</td>
<td>2.07 (0.57)</td>
<td>1.57 (0.46)</td>
<td>85.37</td>
<td>0.316 (0.329)</td>
<td>0.312 (0.203)</td>
</tr>
<tr>
<td>PECB-R</td>
<td>58</td>
<td>2.35 (0.62)</td>
<td>1.88 (0.54)</td>
<td>97.56</td>
<td>0.375 (0.282)</td>
<td>0.433 (0.173)</td>
</tr>
<tr>
<td>PEI- NS</td>
<td>70</td>
<td>2.25 (0.84)</td>
<td>1.72 (0.61)</td>
<td>80.49</td>
<td>0.307 (0.364)</td>
<td>0.350 (0.225)</td>
</tr>
<tr>
<td>PEC</td>
<td>84</td>
<td>2.42 (0.71)</td>
<td>1.83 (0.59)</td>
<td>95.12</td>
<td>0.374 (0.324)</td>
<td>0.404 (0.188)</td>
</tr>
<tr>
<td>EEPA</td>
<td>66</td>
<td>2.27 (0.71)</td>
<td>1.73 (0.54)</td>
<td>92.68</td>
<td>0.379 (0.349)</td>
<td>0.373 (0.196)</td>
</tr>
<tr>
<td>EEJI-NA</td>
<td>88</td>
<td>2.75 (0.77)</td>
<td>2.04 (0.54)</td>
<td>97.56</td>
<td>0.369 (0.286)</td>
<td>0.478 (0.152)</td>
</tr>
<tr>
<td>SE</td>
<td>22</td>
<td>2.47 (0.75)</td>
<td>2.01 (0.50)</td>
<td>92.68</td>
<td>0.437 (0.375)</td>
<td>0.492 (0.159)</td>
</tr>
<tr>
<td>PESM - NP</td>
<td>40</td>
<td>2.22 (0.80)</td>
<td>1.69 (0.57)</td>
<td>80.49</td>
<td>0.323 (0.328)</td>
<td>0.355 (0.212)</td>
</tr>
<tr>
<td>PESM - NSV</td>
<td>41</td>
<td>2.32 (0.93)</td>
<td>1.72 (0.64)</td>
<td>85.37</td>
<td>0.374 (0.368)</td>
<td>0.366 (0.245)</td>
</tr>
<tr>
<td>TOTAL</td>
<td>626</td>
<td>2.97 (0.88)</td>
<td>2.28 (0.63)</td>
<td>100.00</td>
<td>0.344 (0.262)</td>
<td>0.531 (0.127)</td>
</tr>
</tbody>
</table>

<sup>a</sup> A locus was considered polymorphic if the frequency of the most common allele did not exceed 0.95

<sup>b</sup> Unbiased estimate (Nei 1978); Hardy-Weinberg expectations

*** p < 0.001; ** p < 0.01; * p < 0.05
Table 3: Matrix of estimates of $\theta$ (above diagonal) and $N_e$ (below diagonal) for 11 populations of Cryptocarya moschata Nees

<table>
<thead>
<tr>
<th></th>
<th>PECB-S</th>
<th>PECB-SB</th>
<th>PECB-T</th>
<th>PECB-R</th>
<th>PEI-NS</th>
<th>PEC</th>
<th>EEPA</th>
<th>EEJI-NA</th>
<th>SE</th>
<th>PESM-NP</th>
<th>PESM-NSV</th>
</tr>
</thead>
<tbody>
<tr>
<td>PECB-S</td>
<td>0.328</td>
<td>0.316</td>
<td>0.258</td>
<td>0.298</td>
<td>0.297</td>
<td>0.322</td>
<td>0.254</td>
<td>0.276</td>
<td>0.388</td>
<td>0.354</td>
<td></td>
</tr>
<tr>
<td>PECB-SB</td>
<td>0.51</td>
<td>0.323</td>
<td>0.196</td>
<td>0.298</td>
<td>0.275</td>
<td>0.346</td>
<td>0.282</td>
<td>0.263</td>
<td>0.371</td>
<td>0.379</td>
<td></td>
</tr>
<tr>
<td>PECB-T</td>
<td>0.54</td>
<td>0.52</td>
<td>0.259</td>
<td>0.329</td>
<td>0.332</td>
<td>0.357</td>
<td>0.260</td>
<td>0.315</td>
<td>0.416</td>
<td>0.351</td>
<td></td>
</tr>
<tr>
<td>PECB-R</td>
<td>0.72</td>
<td>1.03</td>
<td>0.71</td>
<td>0.192</td>
<td>0.191</td>
<td>0.281</td>
<td>0.168</td>
<td>0.172</td>
<td>0.318</td>
<td>0.292</td>
<td></td>
</tr>
<tr>
<td>PEI-NS</td>
<td>0.59</td>
<td>0.59</td>
<td>0.51</td>
<td>1.05</td>
<td>0.188</td>
<td>0.283</td>
<td>0.264</td>
<td>0.247</td>
<td>0.395</td>
<td>0.323</td>
<td></td>
</tr>
<tr>
<td>PEC</td>
<td>0.59</td>
<td>0.66</td>
<td>0.50</td>
<td>1.06</td>
<td>1.08</td>
<td>0.267</td>
<td>0.223</td>
<td>0.234</td>
<td>0.384</td>
<td>0.314</td>
<td></td>
</tr>
<tr>
<td>EEPA</td>
<td>0.53</td>
<td>0.47</td>
<td>0.45</td>
<td>0.64</td>
<td>0.63</td>
<td>0.69</td>
<td>0.281</td>
<td>0.238</td>
<td>0.381</td>
<td>0.363</td>
<td></td>
</tr>
<tr>
<td>EEJI-NA</td>
<td>0.73</td>
<td>0.64</td>
<td>0.71</td>
<td>1.24</td>
<td>0.70</td>
<td>0.87</td>
<td>0.64</td>
<td>0.173</td>
<td>0.208</td>
<td>0.247</td>
<td></td>
</tr>
<tr>
<td>SE</td>
<td>0.66</td>
<td>0.70</td>
<td>0.54</td>
<td>1.20</td>
<td>0.76</td>
<td>0.82</td>
<td>0.80</td>
<td>1.19</td>
<td>0.262</td>
<td>0.221</td>
<td></td>
</tr>
<tr>
<td>PESM-NP</td>
<td>0.39</td>
<td>0.42</td>
<td>0.35</td>
<td>0.54</td>
<td>0.38</td>
<td>0.40</td>
<td>0.41</td>
<td>0.95</td>
<td>0.71</td>
<td>0.243</td>
<td></td>
</tr>
<tr>
<td>PESM-NSV</td>
<td>0.46</td>
<td>0.41</td>
<td>0.46</td>
<td>0.61</td>
<td>0.52</td>
<td>0.55</td>
<td>0.44</td>
<td>0.76</td>
<td>0.88</td>
<td>0.78</td>
<td></td>
</tr>
</tbody>
</table>

Table 4: Sample sizes with respective estimates of effective population sizes ($N_e$) for populations of Cryptocarya moschata Nees.

<table>
<thead>
<tr>
<th>População</th>
<th>N</th>
<th>$f = F_{IS}$</th>
<th>$N_e$ a</th>
</tr>
</thead>
<tbody>
<tr>
<td>PECB-S</td>
<td>34</td>
<td>0.070</td>
<td>31.78</td>
</tr>
<tr>
<td>PECB-SB</td>
<td>30</td>
<td>0.174</td>
<td>25.55</td>
</tr>
<tr>
<td>PECB-T</td>
<td>28</td>
<td>-0.013</td>
<td>28.37</td>
</tr>
<tr>
<td>PECB-R</td>
<td>30</td>
<td>0.136</td>
<td>26.41</td>
</tr>
<tr>
<td>PEI-NS</td>
<td>36</td>
<td>0.126</td>
<td>31.97</td>
</tr>
<tr>
<td>PEC</td>
<td>43</td>
<td>0.076</td>
<td>39.96</td>
</tr>
<tr>
<td>EEPA</td>
<td>34</td>
<td>-0.017</td>
<td>34.59</td>
</tr>
<tr>
<td>EEJI-NA</td>
<td>45</td>
<td>0.229</td>
<td>36.61</td>
</tr>
<tr>
<td>SE</td>
<td>12</td>
<td>0.119</td>
<td>10.72</td>
</tr>
<tr>
<td>PESM-NP</td>
<td>21</td>
<td>0.092</td>
<td>19.23</td>
</tr>
<tr>
<td>PESM-NSV</td>
<td>22</td>
<td>-0.024</td>
<td>22.54</td>
</tr>
<tr>
<td>TOTAL</td>
<td>335</td>
<td><strong>0.097</strong></td>
<td><strong>17.07</strong></td>
</tr>
</tbody>
</table>

a for a population without genetic structure: $\hat{N}_e = N/(1 + \hat{f})$; b for many populations in a model of infinite populations: $\hat{N}_e = 0.5/[1_{p}[(1 + C_{p})/P - (1/N)] + (1 + \hat{F})/2N]$ . (Vencovsky 1997).
Figure 1: Map of São Paulo and Rio de Janeiro states with the locations of sampled populations of Cryptocarya moschata Nees.
Figure 2: Mean estimates by allozyme loci from 11 natural populations of Cryptocarya moschata Nees. Confidence intervals obtained by numeric resampling: jackknife over samples; for the total, bootstrap over loci.

Figure 3: $\theta_p$ mean estimates by allozyme loci from 11 natural populations of Cryptocarya moschata Nees. Confidence intervals obtained by numeric resampling: jackknife over samples; for the total, bootstrap over loci.
Figure 4: $F$ mean estimates by allozyme loci from 11 natural populations of Cryptocarya moschata Nees. Confidence intervals obtained by numeric resampling: jackknife over samples; for the total, bootstrap over loci.

Figure 5: $\log(\hat{Nm})$ against $\log($distance in km$)$ for populations of Cryptocarya moschata Nees.
tions PECB-S and PECB-T; Ribeira do Iguape: PECB-R, PECB-SB, PEI-NS, and EEPA; Baixada Santista: EEJI-NA; Paraba do Sul: PESM-NSV; Litoral Norte: PESM-NP; Alto Tietê: PEC; and Serra da Estrela, RJ: SE), different altitudes (around 800 m: PECB-S and PECB-T; around 300 m: PECB-R; around 60-100 m: PECB-SB, PEI-NS, PESM-NP, EEPA, EEJI-NA, and SE; around 900-1000 m: PEC and PESM-NSV), geographic proximity, and by a combination of them.

Similarly, but with no hierarchy of populations within regions, for each allele, Wright’s $F$ statistics were calculated through Cockerham-Weir’s estimators (the overall inbreeding coefficient $F = Wright’s F_{IT}$; the coancestry $\hat{\theta}_p = Wright’s F_{ST}$; and $f = Wright’s F_{IS}$; Cockerham & Weir 1993), with permutation tests to verify their significance, done by FSTAT 2.9.3.2 program (Goudet 1995, 2002). Confidence intervals were calculated through resampling procedures like bootstrapping over loci, for their mean, and jackknifing over samples, for each locus.

For indirect estimation of $N_m$, the formula $N_m = 0.25(1 - F_{ST}/F_{ST})$ (Wright 1969) was applied, using the estimator $\hat{\theta}_p$ for $\hat{F}_{ST}$, as demonstrated by Cockerham & Weir (1993), for the fact it is less biased. This model assumes equilibrium between the homogenizing effects of gene flow and the disruptive effects of genetic drift and it also assumes that parameters such as population size and migration rate are uniform and constant over space and time (Sork et al. 1999). From $N_m$ estimate, the neighbourhood size ($N_b$) was calculated from the formula proposed by Slatkin & Barton (1989): $N_b = 2\pi N_m$ (Eguiarte et al. 1993). In this way, considering the density of reproductive individuals per area (d), the neighbourhood area is defined as $A = N_0 d$ (Eguiarte et al. 1993), which was calculated for the populations that presented available estimates of densities.

Information on densities of individuals, which were utilized to calculate neighbourhood areas, was obtained from phytosociological studies already concluded and developed in some of the sampling areas of C. moschata of the present study. So, for PECB-S population, the density found was that of 12 adult individuals.ha$^{-1}$ (A.C. Dias, unpublished data); for PEI-NS, it was of 15.4 adult ind.ha$^{-1}$ (R.J. Almeida-Scabia, unpublished data); 10 adult ind.ha$^{-1}$ for EEPA (N.M. Ivanauskas, unpublished data); 17.3 adult ind.ha$^{-1}$ for PEC (F.A.R.D.P. Arzolla, unpublished data); 7.5 adult ind.ha$^{-1}$ for PESM-NP (M. Sanchez, unpublished data); and 2 adult ind.ha$^{-1}$ for PESM-NSV (M. Tabarelli, unpublished data).

The model of isolation by distance was tested according to Slatkin (1993), making use of a Mantel’s test (Manel 1967, see Manly 1985, Sokal 1979, and Sokal & Rohlf 1995), applying the Pearson’s correlation coefficient ($r$). This test was performed by the following programs: GENEPOP 3.3 (Raymond & Rousset 1995); FSTAT 2.9.3.2 (Goudet 1995, 2002); and GENETIX 4.04 (Belkhir et al. 1996-2002), which calculate $\hat{\theta}_p$ and the number of migrants per generation ($Nm$) between pairs of populations.

The effective population size was obtained through variance components, according methodology presented by Vencovsky (1997) and carried out by Moraes et al. (1999) and Moraes & Derbyshire (2002).

3. Results and Discussion

The main indexes of gene diversity, as well as the observed and expected heterozygosities are shown on table 2. The C. moschata populations displayed high rates of gene diversity, similar to the highest ones found in other tropical tree species. The PECB-T, EEPA and PESM-NSV populations proved to be consistent with the HWE proportions, in spite of the fact that the negative values of the intrapopulation fixation index ($f = F_{IS}$) indicated an excessive number of heterozygotes, although at non-significant levels. The remaining populations presented significant to highly significant intrapopulational endogamy. The total fixation index to the 11 populations grouped was highly significant, with a deficit of heterozygotes. These results per se can only point to the existence of differences among the populations in what regards allelic frequencies and levels of diversity, hinting at a divergence among them. In the absence of inbreeding or population genetic subdivision ($F_{IT} = 0$), $f$ is equal to the probability that a randomly chosen gamete from individual $i$ and a randomly chosen gamete from individual $j$ carry alleles that are identical by descent. This probability is, by definition, the inbreeding coefficient of a hypothetical offspring of $i$ and $j$ so that $f$ has expected values of 0.25 and 0.125 for full and half-sibling family structure, respectively. From table 2, it can be observed that population EEJI-NA showed a full-sib pattern of genetic structure, whereas populations PECB-SB, PECB-R, PEI-NS, SE, and PESM-NP, and PECB-S and PEC have indicated kinship at approximately the level of half-sibs and first cousins, respectively. The estimates of gene and heterozygote variability to the C. moschata populations were slightly higher than those found among the analysed populations of C. aschersoniana by Moraes & Derbyshire (2002), utilizing the same set of allozymes used in the present study.

From the analyses of variance for three-level hierarchy, in all but one of the cases, results showed non significant contribution of the component “regions” ($\theta_R$) to the total variance found under different criteria for combining the populations, and, in the only exception, the variance component due to populations within regions was 3.9% ($p = 0.001$) of the total, when grouped in six regions, as follows: region 1: PECB-S and PECB-T; region 2: PECB-R, PECB-SB and PEI-NS; region 3: EEPA and EEJI-NA; region 4: SE; region 5: PESM-NP and PESM-NSV; and region 6: PEC.
addition, the values of $\hat{\theta}_P$ and $\hat{F}$ were 0.257 and 0.285, respectively. Thus, the smallness of the average fixation index for the 40 loci among the above regions indicated relatively weak differentiation among large areas and, consequently, the apparent genetic structure among regions was not given further consideration. Since the greatest environmental differences are related to the different regions, it is probable that selective differences were of little or no importance and that the fixation indexes were essentially measures of random drift due to endogamous mating in populations (P.L.R. Moraes, unpublished data), and probably accidental differences in the founding individuals among the populations which are largely smoothed out in the regions, as discussed by Wright (1978).

According to figure 2, in which the $f(F_{ST})$ mean estimates by allozyme loci are shown, based on 11 natural populations of *Cryptocarya moschata*, 22 loci (55%) presented significant positive deviations (heterozygote deficit, endogamy) in relation to the HWE proportions, while 11 (27.5%) presented significant negative deviations (excess of heterozygotes, systematic avoidance of consanguine mating within subdivisions) and 7 loci (17.5%) presented themselves in equilibrium (random matings). As pointed out by Weir (1996), values of $f$ can be nulls, as in the case of loci under HWE, positives, as in the case of loci under Wright’s genetic equilibrium (which results from the same conditions that promote the HWE, but the outcrossing rates can be smaller than 1), or even negatives, as in the case of loci under favourable selection to heterozygotes.

Nowadays, according to Coelho (2002), there are copious publications that present the genetic characterization of natural populations, with many of them using estimated values of the intrapopulation fixation index from data of different loci. Generally, the evaluation of those results has demonstrated that the values are highly variable among loci even when measured in the same group of individuals, suggesting that other factors, besides those described in Wright’s genetic equilibrium, must be operating. Several authors have proposed an explanation for this phenomenon. Charlesworth (1991), for instance, ascribed such variations to effects of genotypic disequilibria present among loci in populations of mixed mate system, with intermediate levels of endogamy ($f$ values between 0 and 1), joined to different selective forces in different genomic regions.

Likewise, in figure 3 the $\hat{\theta}_P$ ($F_{ST}$) mean estimates are presented for the 11 populations. It has been observed that all loci (40) proved to be significantly positive. This high divergence can be attributed to the Wahlund effect, as pointed out for *C. aschersoniana* by Moraes & Derbyshire (2002). The overall estimate of $\hat{\theta}_P$ (0.285) was typical of animal-pollinated, mixed selfing and outcrossing plant species (mean $F_{ST}$ = 0.22; Frankham et al. 2002), which is the case for *C. moschata* (Moraes & Monteiro 2002).

According to figure 4, in which the mean values of $F(F_{IT})$ to the *C. moschata* populations are displayed, it can be proved that 28 loci (70%) presented $F_{IT}$ values that were significantly positive, pointing to systematic subdivision, whether into demes ($F_{IT} = 0$ and $F_{ST} = F_{IT}$) or into inbred groups, while 6 (15%) presented significantly negative values, indicating little or no systematic subdivision and prevailing avoidance of consanguine mating (Wright 1969), and another 6 loci showed values not significantly different from zero. Among the various factors that can promote an $F$ estimate different from zero, endogamy, assortative matings, populational subdivision and selection are the most important according to Nei (1987). The analysis of population structure indicated that the subpopulations did not form a single panmictic unit with free gene flow throughout the geographic range of the sampled populations.

The estimate of the number of migrants per generation for the populational set resulted in $\hat{Nm}$ of 0.55, with a neighbourhood size ($\hat{N}_b$) of 3.46 individuals, from the formula to a stepping-stone model (Slatkin & Barton 1989). The values of $Nm$ and $N_b$ determine if the genetic drift per generation can produce substantial genetic variability among locations. If $Nm$ is greater than 1.0, the gene flow will be high enough to avoid a differentiation due to the drift, as the ones found for some pairs of populations (PECB-R and PECB-SB [1.03], PECB-R and PEI-NS [1.05], for example), on table 3. In this study, the mean number of migrants per generation ($\hat{Nm}$) was significantly inferior to 1.0, indicating that the gene flow is not enough to avoid a populational differentiation due to genetic drift, and can be considered relatively low when compared to the values obtained from other tropical Angiosperm trees. The number of parents that exchange genes randomly is approximately 3.5 individuals, whereas a lower number of individuals in the same area unit (surface) would entail an increase in the divergence between neighbourhoods. These estimated parameters were also smaller than those calculated by Moraes et al. (1999) from PECB-S, PECB-SB, PESM-NP and PESM-NSV populations, since the divergence among present populations was considerably higher.

Long-lived woody plants allow the study of both long-term temporal and spatial variation in population structure. Since the present study had no age-structured data, estimates of $Nm$ and $Nb$ were calculated here from populations with no discrete generations, assuming that they were at migration-drift equilibrium, an assumption of Wright’s Island model of population genetic structure (Wright 1969), in relation to species longevity, or life cycle, Hamrick et al.
(1979) demonstrated that perennial, longevoous tree species had showed the highest levels of isoenzymatic variability. According to the authors, longevity could guarantee that many generations would be represented in those populations, since each generation would be related to different selection pressures, from which a great variety of genotypes could be maintained. Although the spatial genetic structure would be a result of natural selection, Wright (1951) and others had showed that a genetic subdivision would be caused by a restricted gene flow and a small effective population size. From the Shifting Balance Theory (Wright 1932), a populational subdivision into small groups (demes), with low gene flow in their midst, would result in a fast differentiation due to drift. However, as pointed out by Sork et al. (1999), this approach for estimating indirectly the gene flow from \( F_{ST} \) provides little insight into the landscape-level processes that differentially may affect spatial patterns of genetic structure and rates of gene flow among individual populations, because \( F_{ST} \) is a summary statistic for a set of populations.

From the estimate values of \( N_{th} \), the neighbourhood areas were estimated for the populations with data on density of individuals, arriving at an average for the species, based on the afore mentioned areas. The area was 2,883 m\(^2\), for PECB-S, taking a density of 12 reproductive individuals.ha\(^{-1}\) into account. As for PEI-NS the area was 2,247 m\(^2\), with a density of 15.4 ind.ha\(^{-1}\). For EEPA the area was 3,460 m\(^2\), with a density of 10 ind.ha\(^{-1}\). For PEC the area was 2,000 m\(^2\), with a density of 17.3 ind.ha\(^{-1}\). For PESM-NP the area was 4,613 m\(^2\), with a density of 7.5 ind.ha\(^{-1}\) and, for PESM-NSV, the area was 17,300 m\(^2\), with a density of 2 ind.ha\(^{-1}\). Consequently, the area of average neighbourhood found was 5,417.17 m\(^2\), from \( \hat{\theta}_p \). It means that circa 6.4 reproductive individuals.ha\(^{-1}\) would be necessary to maintain the actual genetic structure for the populations of \( C. moschata \) investigated, in mean terms (corresponds to the calculation of the harmonic mean of density).

The values of \( \hat{\theta}_p \), calculated among the pairs of populations, can be found on table 3 and are all significantly different from zero. One can conclude that several \( \hat{\theta}_p \) values are discrepant in relation to what could be expected from the geographical distances among the pairs of populations. By the two-dimensional stepping-stone model (i.e., nearest neighbor dispersal), it might reasonably be expected that subpopulations which never exchange migrants should not exhibit much higher \( F_{ST} \) values than those that do (Rousset 1997), as what was found for the PECB-T and PECBR (\( \hat{\theta}_p = 0.259 \)) and PECB-T and PESM-NP (\( \hat{\theta}_p = 0.416 \)) populations, for example. Likewise, based on a classical isolation by distance model, where gene flow is considered solely in terms of gene exchange during mating, an increase of \( \hat{\theta}_p \) would be expected with the increase of the distance among the pairs (Raybould et al. 1996).

Unlike the island model, the isolation by distance model incorporates spatial information and thus permits tests of hypotheses about relationships between the effective migration rate and spatial patterns of interpopulation connectivity (Sork et al. 1999). In order to test the isolation by distance model, a nonparametric test of the two matrixes of distances was carried out. For the \( C. moschata \) populations, the value obtained for the Mantel coefficient (Z) among the logarithms matrixes of the geographic distances and the number of migrants per generation (Nm) was -46.60, with the probability (P = 0.6338) of error in rejecting the null hypothesis of independence, based on 10,000 permutations, thus indicating that there is no isolation by distance among the populations of \( C. moschata \) analysed. In the same way, based on the graph in logarithm scale for \( \hat{\theta}_N \) versus the geographic distance (figure 5), an approximately linear regression was expected for a population in equilibrium, with restricted dispersion. However, the regression coefficient was -0.025, with the Y intercept explaining no more than 0.75% of the log (\( \hat{\theta}_N \)) total variance, through regression, demonstrating once again that the pattern obtained does not point to a tendency of the species to present isolation by distance, being in equilibrium with migration and drift.

The \( \hat{\theta}_p \) values found for the pairs of populations indicated, in this way, that factors such as drift and founder effects might probably be related to the high divergence found among populations. In this sense, according to the tests of adequacy to the Infinite Allele Model and the Stepwise Mutation Model (P.L.R. Moraes & Derbyshire, unpublished data), it is observed that the PECB-S, PECB-T, PEC, EEPA, PESM-NP and PESM-NSV presented mutation-drift equilibrium, while the others presented significant deviations, indicating bottlenecks occurrence. There is another prominent difference between the \( C. moschata \) and \( C. aschersoniana \) populations (Moraes & Derbyshire 2002), which were all in mutation-drift disequilibrium and are mostly fragmented and disturbed. These results also showed that the hypothesis of isolation by distance formulated from the tendency found by Moraes et al. (1999) and Moraes et al. (2002), based on data for only four \( C. moschata \) populations and 18 allozyme loci, were not confirmed for the group herein analysed of 11 populations and 40 alloenzymatic loci. On the contrary, the results seemed to match a random model, which states that the level of similarity in gene frequencies among subpopulations would not be correlated to the distances that separate them, or to a metapopulation model. However, all the models are idealized, and the true model describing \( C. moschata \) population genetics is likely to be
different to any of the proposed models that can be found in literature. It is also worth to mention that these estimates of indirect gene flow from \( F_{ST} \) are historical and do not reflect contemporary variation in gene exchange among populations or current changes to dispersal processes (Sork et al. 1999).

On table 4, it is observed that, in an intrapopulational context, for the sampling of *C. moschata* individuals from PECB-S, 34 sampled individuals represented an \( N_e \) of 31.78 individuals. Likewise, the effective sizes found for the samplings of the remaining populations showed that there was an adequate genetic representativeness of the samples obtained, from the allele frequencies of the analysed loci, from those with low values of \( \hat{f} \). Therefore, from the formula

\[
\hat{N}_e = \frac{N}{1 + \hat{f}}
\]

one can also observe that endogamy reduces the effective population size, and the negligible value found for \( \hat{f} \) causes \( \hat{N}_e \) to be practically the same as the number of individuals sampled, as presented by Vencovsky (1997).

In a metapopulation context, taking the 11 populations sampled for the species (populations as sampling units), considering it consists of infinite populations, the effective population size that resulted was of 17.07 individuals, for the 335 sampled. For the limit, with \( \theta_P = 0 \) and \( N \) being sufficiently large, \( N_e \rightarrow (0.5 P)/\theta \), equaled 19.29 individuals. Thus the sampling carried out for the species corresponded to 88.44% of the maximum effective size obtained from the 11 populations with a \( \hat{\theta}_P \) of 0.285, representing only 5.09% of the individuals sampled. However, it is clear that the effective size obtained for the species was considerably small, given that the divergence among populations was relatively high, showing that the sampling of different populations is an important strategy, even when the divergence among them is relatively small, as pointed out by Vencovsky (1997) and analysed for the *C. moschata* populations by Moraes et al. (1999) and for the *C. aschersoniana* populations by Moraes & Derbyshire (2002).

In summary, the present study of genetic structure of natural populations of *Cryptocarya moschata* revealed a higher interpopulation genetic diversity than what was obtained previously, and it was suggested that this is because restricted gene flow which has promoted that the correlation among allelic frequencies from different individuals of the same population (\( \theta_P \)) was basically a measure of random drift due to endogamous mating in populations, and probably random differences in the founding individuals within and among populations (Wahlund effect). Moreover, both isolation by distance and stepping stone models were inadequate to explain the results, which suggested that a random model of gene flow, or a classical metapopulation model, could be more suitable. Overall, populations that presented relative high intrapopulational fixation indexes (\( \hat{f} \)) showed losses in genetic representativeness from 11% (SE) to almost 19% (EEJI-NA). Although there was an increase of 15% in genetic representativeness of the sampling procedure related to the maximum effective population size for the species as a whole, there was an improvement of only 0.09% in genetic representativeness for the total number of individuals sampled, compared to what was obtained from the previous work. Thus, it was demonstrated that the high genetic divergence found among *C. moschata* populations could be a current characteristic of the species evolutionary dynamics, which indicates that the in situ conservation of the species must encompass large areas of Atlantic rain forest. This scenario has been probably realized in the extensive reserves of the State Parks of São Paulo. However, since the smallest populations have also showed genetic diversities as high as the largest ones, their management in the future should be also important in order to avoid detrimental effects and losses of variability caused by endogamy and/or genetic drift.

4. Acknowledgements

To “Instituto Florestal de São Paulo”, “Fundação Florestal”, “Instituto Agronômico de Campinas” and to “Superintendência da IMBEL”, for permission of collection in the respective areas, in their custody. To Marco Aurélio Nalon for constructing the figure 1. To Fernando H. Romano and Mariana C. Daher for their laboratory assistance. A special word of thanks to three anonymous reviewers for their criticism, and to Angela M. Vasconcellos for her support with the manuscript translation into English.

5. References


