

GENETIC STRUCTURE OF *Copaifera langsdorffii* Desf. NATURAL POPULATIONS

Dulcinéia de Carvalho¹; Alessandro Fabiano de Oliveira²

(Recebido: 9 de dezembro de 2003; aceito: 19 de maio de 2004)

ABSTRACT: Three natural populations of *Copaifera langsdorffii* Desf. were studied by isoenzyme electrophoresis to determine the genetic structure, levels of genetic variability, gene flow and effective size of the populations. The populations are located in the county of Lavras, in Southern Minas Gerais state and are represented by three types of vegetation (Savanna, Riparian forest and Semi-deciduous forest). Samples of leaf tissue were taken from 20 adult plants from each population and from 400 progenies for isoenzyme electrophoresis analysis. Five enzymatic systems were evaluated in adults (24 alleles average) and progenies (29 alleles total). The proportion of polymorphic loci (0,95) among populations was 72.73% to 87.50%. The average number of alleles per locus was 2.2 to 2.5 and the expected heterozygosity (\hat{H}_e) was 0.368 to 0.435. The estimated parameters revealed inbreeding in the adult populations ($\hat{F} = 0.130$) and excess of heterozygotes in the progenies ($\hat{F} = -0.033$). The genetic diversity index of the populations was 0.142 for the Savanna-Riparian forest, 0.162 for the Riparian Semi-deciduous forest, and 0.073 for the Savanna-Semi-deciduous forest. The gene flow, estimated by the number of migrants (\hat{N}_m), was 0.79 for Savanna-Semi-deciduous forest, 0.38 for Savanna-Riparian forest, 0.32 for Riparian-Semi-deciduous forest and 0.41 for all analyzed populations. The population from the Riparian forest has a relatively high genetic variability and is valuable for genetic conservation of *C. langsdorffii*.

Key words: Copaíba oil, Isoenzyme, Genetic conservation

ESTRUTURA GENÉTICA DE POPULAÇÕES DE *Copaifera langsdorffii* Desf.

RESUMO: Três populações naturais de *Copaifera langsdorffii* Desf., espécie arbórea comumente encontrada no Brasil, foram estudadas por meio da eletroforese de isoenzimas, visando determinar os níveis de variabilidade genética mantidos dentro e entre as populações, sua estrutura genética, a taxa de cruzamento, o fluxo gênico, o sistema reprodutivo e o tamanho efetivo das populações. As populações amostradas localizam-se no município de Lavras, sul de Minas Gerais, sendo que duas (cerrado e mata semidecidual) estão localizadas no campus da Universidade Federal de Lavras (UFLA) e a terceira (mata ciliar) em uma área de preservação permanente, entre os municípios de Lavras e Itumirim. Foram amostrados tecidos foliares de 20 indivíduos de cada população e analisaram-se 400 indivíduos jovens (progênies) procedentes de sementes coletadas de 20 matrizes na população da mata ciliar. Foram testados 21 sistemas enzimáticos e escolhidos os 5 melhores, revelando 35 alelos totais distribuídos em 12 locos. O polimorfismo (P) com limite de frequência igual ou inferior a 0,95 variou entre 72,73% a 87,50% entre as populações. O número médio de alelos por loco (A)

¹ Universidade Federal de Lavras – Departamento de Ciências Florestais, Caixa Postal 37, 37200-000 Lavras, MG, Brasil. Dulce@ufla.br

² IBAMA- Brasília, DF

variou entre 2,2 a 2,5 e a diversidade genética medida pela heteroziguidade média esperada (\hat{H}_e) variou entre 0,368 e 0,435. A estrutura genética revelou que há endogamia para o conjunto das populações adultas ($\hat{F} = 0,130$) e uma tendência de excesso de heterozigotos para as progênies ($\hat{F} = -0,033$). Os pares de populações cerrado-mata ciliar e mata semidecidual-mata ciliar apresentaram altos valores de \hat{q}_p (0,142 e 0,162), comparados ao par cerrado-mata semidecidual (0,073). O fluxo gênico medido pelo número de migrantes (\hat{N}_m) foi baixo, variando de 0,79 entre as populações cerrado-mata semidecidual, 0,38 entre cerrado-mata ciliar, 0,32 entre mata semidecidual-mata ciliar e 0,41 para o conjunto das três populações. A população localizada em Itumirim (mata ciliar) mostrou-se potencial para a conservação genética de *C. langsdorffii*.

Palavras chave: copaíba, isoenzimas, conservação genética

1 INTRODUCTION

Inadequate exploitation of natural resources has caused the extinction of many species in the different biomes of the planet, especially in those situated in the tropics. Habitat loss and fragmentation are among the most common causes of extinction. Therefore, there is an urgent need for population genetic studies of species which constitute these ecosystems to determine strategies for management and genetic conservation.

Many species which make up ecosystems are important for their economic and ecological values, but since the study of a large number of species is impractical, it is necessary to sample those that could serve as a model or that are representative of the bioma under study. Among these, *Copaifera langsdorffii*, commonly known as Copaíba oil tree, is suitable for population genetic studies because it is a widely distributed species, with great ecological plasticity, zoochoric and hydrochoric dispersion and bee pollination.

Use of isoenzyme analysis in genetic forest studies has been increasing recently. The main application of electrophoresis techniques has been in the investigation of genetic variation patterns (Lundkvist & Rudin, 1977) and crossing systems (Brown et al., 1975; Shaw & Brown, 1982). These

applications increase knowledge of the genetic structure in natural populations, increasing the efficiency of breeding and efforts for genetic conservation of endangered tree species.

This study was carried out to investigate the level and distribution of the genetic variability among and within the populations, through: i) quantification of the genetic variability; ii) characterization of the genetic structure of the population; iii) identification of the inter- and intra- population gene flow; and, iv) estimation of the effective size of the populations. Knowledge of these population characteristics is very important for the adoption of adequate management strategies for genetic conservation.

2 MATERIAL AND METHODS

Three *Copaifera langsdorffii* Desf. populations located in three distinct areas were sampled. The first location of study was a 5.83 ha forest fragment (Figure 1) located on the Campus of the Federal University of Lavras (UFLA), Lavras, MG, which is also known as "Matinha da UFLA". The vegetation is classified as Mountain Semi-deciduous Mesophytic forest (Semi-deciduous tropical forest) with emergent canopy (*Fme*) according to the classification system of the Fundação

Instituto Brasileiro de Geografia e Estatística – [Brazilian Foundation Institute of Geography and Statistics] (FIGBGE, 1993). This fragment is located at 21°14'40" SL and 44°57'50" WL, at an average altitude of 925m. The second study area is also located on the UFLA campus. It is a 3.2 ha (approximately) fragment of Savanna vegetation, known as "Cerradinho da UFLA". This fragment is located at 21°14'00" S and 44°59'00" W at an average altitude of 920 m. The third area is situated in an area subject to great human disturbance. It is a private property, located on the Capivari River bank, a tributary of "Rio Grande", between the counties of Lavras and Itumirim, in Minas Gerais state (at 21°18'40" SL and 44°53'30" WL at an altitude of approximately 918 m. It is a fragment of Riparian forest, with an area of almost 14.0 ha that extends along a stretch of the River Capivari.

The typical climate of the region is Cwb in the Köppen classification (Mesothermic with warm, mild summers and drought in the winter). The mean annual rainfall and temperature are 1,530 mm and 19.3°C, respectively. The monthly mean rainfall in the wet season (December to February) and dry season (May to June) are 254 mm and 25mm, respectively (Brasil, 1992).

The individuals were randomly sampled from practically the whole area of the three studied locations. Plant material was sampled by collecting leaves from 20 reproductively active individuals (considered adults) in each population and were taken to the forest Breeding and Molecular Biology laboratory (DCF – UFLA) and stored at -18°C. Seeds were also collected from 20 adult trees (not necessarily the same trees from which the leaves were collected) in each population for progeny analysis. Each plot contained 20 seedlings from each adult tree, totaling 400 seedlings. A total of 60 adult individuals from

the three populations and 400 seedlings from the Riparian forest population (Capivari) were analyzed.

Genetic characterization of the species was carried out using the most commonly used techniques of isoenzymatic electrophoresis, especially the methodology described by Alfenas et al. (1991) and Alfenas (1998). All laboratory work was carried out in the Forest Breeding and Molecular Biology Laboratory of the Department of Forest Science at the Federal University of Lavras (DCF – UFLA), Lavras MG. Discontinuous system vertical isoenzyme electrophoresis in polyacrilamide gel was used, with the concentration and separation gels at 4% and 12.5%, respectively. Each gel accommodated samples from 20 individuals. Gel development was based on the methodology described by Alfenas (1998).

The zymograms were interpreted after the laboratory work was finished and the genotype of each adult individual and the seedlings was obtained. Genetic parameters that characterized the populations effective size, among and within variability and gene flow were estimated. The allelic frequency at each segregating locus was also obtained from the interpretation of the zymograms. The following diversity indexes were estimated from these frequencies: a) observed heterozygosity; b) expected heterozygosity according to the Hardy-Weinberg Equilibrium (HWE); c) percentage of polymorphic loci; and, d) number of alleles per locus (Brown & Weir, 1983; Nei, 1987). The estimates of these parameters that measure the intrapopulation variability were obtained using the BIOSYS-1 software (Swofford & Selander, 1989). The structure genetic of populations was characterized by Cockerham coancestry (\hat{q}_P) obtained by allelic frequencies variances using the software GDA (Lewis & Zaykin, 2000).

The flow gene (\hat{N}_m) among populations was estimated according Crow & Aoki (1984) and the effective size of populations (\hat{N}_e) was estimated according Vencovsky (1992).

3 RESULTS AND DISCUSSION

The Alfa Esterase, Glutamate Oxaloacetate Transaminase, Phosphatase Acid, Malate Dehydrogenase, Peroxidase, Sorbitol Dehydrogenase and Dehydrogenase Glutamate enzymatic systems were chosen among the 21 tested to establish the protocol for the *C. langsdorffii* Desf. tree species. However, only the first five showed enough resolution to allow interpretation. The analysis involved 100 gels from the progenies and 15 from the adult individuals. Estimates of the genetic diversity indexes and of the genetic structure parameters of the *C. langsdorffii* Desf. populations were obtained from the frequencies of 35 alleles distributed in 12 aloenzymatic loci (Table 1).

The allelic frequencies varied in each adult population from the complete fixation to very low frequencies of alleles. For example, the Acp-4 and Got-1 loci in the Savanna population, Got-1 in the Semi-deciduous forest and Acp-1 and Acp-4 in the adults and progenies of the Riparian forest populations showed complete fixation, while allele 2 in the Acp-3 locus (0.05), allele 3 in the Mdh-2 locus (0.075) in the Savanna population, allele 3 in the Est-2 locus (0.033) in the Semi-deciduous forest and allele 3 in the Est-1 locus (0.029), allele 2 in Got-1 locus (0.024) in the adult individual population in the Riparian forest and allele 3 in the Est-1 (0.007), Est-2 (0.031), Acp-3 (0.025), Got-2 (0.026), Prx-1 (0.054) and Mdh-1 loci (0.024) and allele 2 in the Acp-1 (0.039), Got-1 (0.013) loci in the Riparian forest progeny showed very low

frequencies. Therefore, the data from the 11 aloenzymatic loci showed allelic frequency variation among and within populations, detecting differences between the adult populations. The variation in the frequencies of the most common alleles suggested high genetic divergence among the populations.

The comparison of the allelic frequencies of the adults and progenies of the Riparian forest population showed that few differences were present, suggesting the existence of equilibrium among generations. Oscillation in the allelic frequencies, allelic losses and fixation may be an indication of genetic drift or selection. Oscillations usually develop from random processes, in the absence of natural selection. In the studied locations, which characterize different environments with distinct variations, both genetic drift and natural selection may influence the allelic frequency oscillations in *C. langsdorffii* Desf. In the case of selection, the effect may be either direct, if the aloenzymes participate in metabolic routes that affect the characteristics under selection, or indirect, if the aloenzymatic loci are linked to the genes controlling those characteristics.

The means of the estimates obtained indicated that endogamy was present in the three adult population sets ($\hat{F} = 0.130$) and there was a tendency to an excessive number of heterozygotes in the progeny ($\hat{F} = -0.033$). The mean expectations of the within adult population (\hat{f}) parameters indicated low endogamy (-0.001), suggesting that, on average, the populations were not endogamic. Therefore, their genotypic frequencies are in HWE. However, the \hat{F} showed that the population set was not in HWE, was not panmictic and had, therefore, endogamy due to genetic drift and reproductive system effects (Table 2).

Table 1. Allelic frequencies and sample size (n) in 11 aloenzymatic loci in three natural *Copaifera langsdorffii* Desf. populations.**Tabela 1.** Frequências alélicas e tamanho amostral (n) em 11 locos aloenzimáticos em três populações naturais de *Copaifera langsdorffii* Desf.

Locus	Alleles	Populations			
		Savanna	Semi-deciduous	Riparian*	Progenies
α-Est-1	1	0,342	0,500	0,529	0,362
	2	0,579	0,400	0,441	0,631
	3	0,079	0,100	0,029	0,007
	n	19	10	17	290
α-Est-2	1	0,132	0,367	0,464	0,457
	2	0,447	0,467	0,286	0,512
	3	0,316	0,133	0,250	0,031
	4	0,105	0,033	0,000	0,000
n	19	15	28	383	
Acp-1	1	0,325	0,143	1,000	0,961
	2	0,675	0,857	0,000	0,039
	n	20	14	13	382
Acp-2	1	0,447	0,000	0,654	0,658
	2	0,500	0,000	0,346	0,342
	3	0,053	0,000	0,000	0,000
	n	19	00	13	360
Acp-3	1	0,950	0,000	0,769	0,676
	2	0,050	0,000	0,231	0,299
	3	-	-	-	0,025
	n	20	00	13	276
Acp-4	1	1,000	0,000	1,000	1,000
	2	-	-	-	-
	n	20	00	13	395
Got-1	1	1,000	1,000	0,976	0,987
	2	0,000	0,000	0,024	0,013
	n	20	16	21	356
Got-2	1	-	-	-	0,777
	2	-	-	-	0,196
	3	-	-	-	0,026
	n	-	-	-	191
Mdh-1	1	0,775	0,425	0,464	0,564
	2	0,225	0,575	0,304	0,412
	3	0,000	0,000	0,232	0,024
	n	20	20	28	373
Mdh-2	1	0,400	0,286	0,500	0,547
	2	0,525	0,714	0,500	0,453
	3	0,075	0,000	0,000	0,000
	n	20	14	13	320
Prx-1	1	0,188	0,722	0,485	0,478
	2	0,406	0,278	0,409	0,468
	3	0,219	0,000	0,106	0,054
	4	0,188	0,000	0,000	0,000
	n	16	09	33	231
Prx-2	1	0,175	0,175	0,577	0,556
	2	0,575	0,300	0,423	0,444
	3	0,250	0,525	0,000	0,000
	n	20	20	13	360
Alleles Total	35	28	19	24	29

* Oliveira et al. (2002)

Analysis of the progeny intrapopulation genetic structure showed that 16% of the genetic variability was detected among progenies and 84% among individuals within progenies. In half-sib, full-sib and S1 (one generation of self pollination) progenies, these values are expected to be at least 0.125, 0.25 and 0.50, respectively. Thus, the 0.163 value observed indicated that the progenies were not composed exclusively of half-sibs, but were also made up of full-sib and/or selfed individuals, obviously in a lower proportion than the half-sibs. This suggested that the crosses did not happen completely at random, and there may have been preferential mating (full-sib and/or selfing).

The genetic divergence among populations was high ($\hat{q}_p=0.131$), indicating that 13% and 87% of the genetic variability was detected among and within populations, respectively. This divergence was expected, given the differences observed among the allelic frequencies of the populations. Most of the tropical tree species have less than 5% genetic divergence among populations. However, the higher between population genetic divergence detected in this study (13%) is still in line with values observed for other species, which show greater proportions of the genetic variability within populations.

Ciampi (1999) working with AFLP (Amplified Fragment Length Polymorphism) and micro-satellites makers (SSR – Simple Sequence Repeats) found that most (about 90%) of the genetic variability in *C. langsdorffii* Desf. is obtained within population and only about 10% is detected among populations.

The pattern of organization of the *C. langsdorffii* Desf. genetic variation is similar to that commonly found in tropical tree species, which show high genetic variation within populations and low variation among populations.

However, there are cases where high levels of genetic diversity are found, such as in *Chorisia speciosa* (Souza, 1997), *Cryptocarya moschata* (Moraes, 1997) and *Syagrus romanzoffiana* (Negrão, 1999). Thus, *C. langsdorffii* Desf. showed an interpopulation genetic divergence value which is among the highest among tropical tree species. The high genetic divergence among populations estimated by fixation index (\hat{F}) and the low rate of within population endogamic crossing detected for the studied species are values in line with those expected from the action of random genetic drift (Futuyma, 1992). Random drift can lead each population to attain different allelic frequencies and, therefore, cause a large genetic divergence among populations. Consequently, the high divergence found among the populations in the present study indicates that random drift affected the genetic structure of the studied *C. langsdorffii* Desf. populations.

Souza (1997) suggested the association of high indices of \hat{F} with the action of genetic drift in *Chorisia speciosa*, studied in forest fragments in the interior of São Paulo state. The association among high interpopulational variation and low intrapopulational variation with genetic drift in tree species has already been proposed by various authors (Ge et al., 1998; Kitamura & Rahman, 1992; Moran & Hoper, 1983; Peters et al., 1990; Rocha & Lobo, 1996).

Divergence among the populations can be the result of the fragmentation process leading to genetic drift. However, considering the longevity of the life cycle of the species and the differences in the environments, this divergence may also be the result of selection.

The -0.234 \hat{f} value showed an excess of heterozygotes within progenies compared to the HWE expectation. However, the -0.033 \hat{f} value suggested that the progenies (and the population) did not show endogamy and were in HWE.

Table 2. Coefficient of coancestrality among three adult and one progeny populations of *Copaifera langsdorffii* Desf. Values within brackets correspond to the confidence intervals with 95% probability.**Tabela 2.** Coeficiente de coancestralidade entre três populações de indivíduos adultos e uma progênie de *Copaifera langsdorffii* Desf. Valores entre colchetes correspondem aos intervalos de confiança com 95% de probabilidade

	Locus	\hat{q}_P	\hat{q}_F	\hat{F}	\hat{f}
Adults	AES-1	-0,011	-	0,223	0,231
	AES-2	0,052	-	0,012	-0,042
	ACP-1	0,500	-	0,782	0,565
	ACP-2	0,020	-	0,210	0,194
	ACP-3	0,148	-	0,083	-0,076
	ACP-4	***	-	***	***
	GOT-1	-0,008	-	-0,003	0,005
	MDH-1	0,109	-	-0,279	0,020
	MDH-2	0,056	-	-0,056	-0,435
	PRX-1	0,087	-	0,105	-0,119
	PRX-2	0,195	-	0,059	-0,168
		Mean	0,131 [0,049 a 0,240]		0,130 [-0,034 a 0,330]
Progenies	AES-1	-	0,099	-0,002	-0,111
	AES-2	-	0,123	-0,114	-0,272
	ACP-1	-	0,649	0,531	-0,336
	ACP-2	-	0,359	0,193	-0,259
	ACP-3	-	0,251	0,214	-0,050
	ACP-4	-	-0,001	-0,000	0,001
	GOT-1	-	0,104	-0,007	-0,122
	GOT-2	-	0,080	-0,274	-0,385
	MDH-1	-	0,127	-0,053	-0,206
	PRX-1	-	0,183	0,129	-0,066
	PRX-2	-	0,097	-0,133	-0,254
		MDH-2		0,076	-0,364
	Mean		0,163 [0,111 a 0,230]	-0,033 [-0,156 a 0,085]	-0,234 [-0,325 a 0,147]

The negative mean value of endogamy ($\hat{f} = -0.001$) for the three population set is explained by the fact that the Savanna population had an excess of homozygotes (0.239), while the Riparian forest population showed an excess of heterozygotes (-0.203) (Table 3). The \hat{F} and \hat{f} results, with negative values for some loci, obtained for *C. langsdorffii* Desf., indicated high divergence in the majority of the adult and the progeny populations loci. Therefore, when considering the corresponding confidence intervals, it should be observed that these estimates were not statistically different from zero. These results suggested that, indeed, there was a defined population structure (a differentiation in relation to a reference ancestral population), but there was no significant endogamy when the populations were taken as a whole.

The detected population divergence, however, did not seem to have equal source in all the populations. This can be observed by the analysis of the genetic structure, taking the populations pairwise (Table 3). It can be seen that the population pairs Savanna-Riparian forests and the Semi-deciduous-Riparian forests presented high genetic divergence values ($\hat{q}_p = 0.142$ and 0.162 , respectively) compared to the Savanna-Semi-deciduous forests ($\hat{q}_p = 0.073$). This indicates that the Savanna and Semi-deciduous forest populations were more similar between each other and divergent from the Riparian forest population. This is coherent, as these two populations are found geographically closer (approximately at 1,400m). This low divergence and close geographical localization indicated that these populations

are linked by pollinator mediation and suggested that they have a common ancestral population. They may have formed a single population in the past, which gave rise to these distinct populations through fragmentation.

Partitioning the genetic structure analysis also indicated that the within populations excess of heterozygotes detected in the joint analysis of all populations occurred in consequence of the Riparian forest that showed a negative endogamy value (\hat{f}). This value was negative only when this population was compared. This suggests that the high genetic divergence among the three locations probably occurred due to a combined effect of genetic drift and selection, that is, the Riparian forest populations were under selection favoring heterozygotes and the Savanna and Semi-deciduous forests diverged by genetic drift. This hypothesis is based on the fact that drift generates endogamy and selection can favor homozygotes or heterozygotes, and in this case, selection favored the heterozygotes.

Population structure can be analyzed in the absence or presence of gene flow. If there is current gene flow, it should be between the Savanna-Semi-deciduous forest populations as they have low genetic divergence and, in an intermediate way, between the Savanna-Riparian forest and between the Semi-deciduous-Riparian forest populations, which showed greater divergence. Barriers would hinder the gene flow among these populations. In the absence of current gene flow, it can be considered that these populations may have been formed from a common ancestral population, but from different foundation events.

Table 3. Coancestry coefficients for the populations pairs. (\hat{f}): Within population endogamy coefficient. (\hat{F}): endogamy coefficient for the whole population set. (\hat{q}_p): genetic divergence among populations. Values in brackets indicate the confidence interval obtained from 10,000 bootstrapping.

Tabela 3. Coeficiente de coancestralidade tomando as populações duas a duas. \hat{f} : coeficiente de endogamia dentro das populações. \hat{F} : coeficiente de endogamia para o total das populações. \hat{q}_p : divergência genética entre populações. Valores entre parênteses indicam o intervalo de confiança obtido a partir de 10.000 bootstrapping.

Populations	\hat{f}	\hat{F}	\hat{q}_p
Savanna and Semi-Deciduous	0,173 (-0,049 a 0,353)	0,234 (0,046 a 0,398)	0,073 (0,014 a 0,132)
Savanna and Riparian	-0,015 (-0,132 a 0,107)	0,129 (-0,017 a 0,320)	0,142 (0,048 a 0,273)
Semi-Deciduous and Savanna	-0,119 (-0,280 a 0,078)	0,062 (-0,182 a 0,373)	0,162 (0,017 a 0,342)

The populations, in the present study, had average within population and small among population genetic variance. According to Loveless & Hamrick (1984), this is a coherent result since the typically self pollinating species have high within population genetic variance and low among population genetic variance. However, the divergence among populations is reduced proportionally to the amount of the gene flow (pollen and/or seeds).

When the percentage of polymorphic loci found in *C. langsdorffii* Desf. (Table 3) was compared with that of two tropical tree species, *Alseis blackiana* (89.3%) and *Psychotria horizontalis* (49.5%) (Hamrick & Loveless, 1986), values slightly lower than those of the first species and much higher than those of the second were observed. The analysis of the percentage of polymorphic loci showed that *C. langsdorffii* Desf. had high polymorphism, which makes these populations favorable for genetic conservation.

Regarding the number of alleles per locus, the mean values were 2.5 (Savanna), 2.4 (Semi-deciduous forest), 2.2 (Riparian Forest) and 2.5 (Riparian forest progenies), which did not statistically differ among each other, to judge from the standard deviation.

The expected heterozygosity mean of the Savanna, Semi-deciduous and Riparian forest populations of *C. langsdorffii* Desf. (Table 4) expressed a tendency to an excess of homozygotes in the first two cases and an excess of heterozygotes in the third case compared to the HWE expected values, which can be better visualized by the fixation index scores. These values were in line with those usually obtained in studies on tropical tree species and reinforce the fact that greater allele fixation is usually found in progenies, which reflects in lower polymorphism.

It was found that the heterozygosity observed and heterozygosity expected were high (>0.30) when compared with most tree species studies, both tropical and temperate,

showing a typical pattern of high density self-pollinating species (common species). The standard deviations of these estimates suggested that there were no among population differences for these parameters. Similarly, the average number of alleles per locus was similar among populations. However, the percentage of polymorphic loci indicated that the Savanna and Semi-deciduous forest populations had greater among loci variability, that is, they showed a higher number of polymorphic loci than the Riparian forest population.

When adults and progenies of the Riparian forest were compared, similar values for the diversity indexes were obtained, but heterozygosity tended to increase between the seedling and adult phases. This tendency has also been observed in most tropical tree species studies.

The tendency for heterozygosity to increase in adult individuals suggested that selective mechanisms favored the heterozygotes. This tendency has been observed in several studies, such as that of Hamrick et al. (1993) who studied three tropical tree species. Results of this type have also been found for other tropical tree species, such as *Cecropia obtusifolia* (Alvarez-Buylla & Garay, 1994), *Cedrela fissilis* (Gandara, 1996), *Chorisia speciosa* (Souza, 1997), *Genipa americana* (Sebbenn, 1997) and *Euterpe edulis* (Reis, 1996).

The standard deviation of alleles number (A) suggested that there were no differences among adults and progenies, but an increase in the number of alleles in the progenies was detected. This may have resulted from migration (gene flow) or from sampling problems in the adult phase, where the 20 individuals sampled may not have been representative of all the allelic variation present in the population.

The fixation index, or endogamy coefficient, suggested that the Savanna population is highly endogamic, that the Semi-deciduous forest population is under HWE and that there is selection favoring heterozygotes in the Riparian forest population.

The heterozygosity observed mean values of the *C. langsdorffii* adult populations were similar to those obtained by Hill et al. (1978) for three Riparian Amazon forest tree species, 0.100 to 0.425 for *Aeschynomene sensitiva* Sw. Var. *amazonica* Rudd., 0.184 to 0.364 for *Aeschynomene sensitiva* Sw. Var. *sensitiva* Rudd and 0.333 to 0.460 for *Mimosa pigra* L..

Ciampi (1999) analyzed eight SSR loci in *C. langsdorffii* Desf and observed genetic divergence values ($\hat{H}_e = 0.88$ and $\hat{H}_o = 0.89$) four to twelve-fold greater than those obtained with isoenzymes in different tree species (Sheely & Meagher, 1996; Loveless et al., 1998). Therefore, comparisons among diversity estimates obtained by isoenzymes and microsatellites become difficult to assess due to the fundamental differences in the nature of the sequences of the sampled DNA. Isoenzymes are coding sequences and microsatellites are non coding regions and are subject to high rates of mutation by their structure (Ciampi, 1999).

Hill et al. (1978) also observed that the populations located along river banks almost always show higher levels of heterozygosity than those located further away, and attributed this characteristic to hydrochoric dispersion. Since dispersion can also be hydrochoric in *C. langsdorffii* Desf. (Carvalho, 1994) and the Riparian forest population is on the banks of the River Capivari, its higher heterozygosity observed value compared to those of the other populations suggests that hydrochoric dispersion occurred. It would be interesting to complement this study by assessing other populations along the River Capivari to detect the extension of gene flow caused by water dispersion.

Table 4. Genetic diversity indexes of three natural *Copaifera langsdorffii* Desf. populations in the Lavras, MG region, based on 11 loci and five enzymatic systems (adult individuals) and 12 loci and five enzymatic systems (progenies of the Riparian forest population). * A locus is considered polymorphic if the most common allelic frequency does not exceed 0.95 probability; values within brackets correspond to the standard deviations and within square brackets to the confidence intervals.

Tabela 4. Índices de diversidade genética de três populações naturais de *Copaifera langsdorffii* Desf. na região de Lavras, MG, baseados em 11 locos e 5 sistemas enzimáticos (indivíduos adultos) e 12 locos e 5 sistemas enzimáticos (progênies da população da mata ciliar). * Um loco é considerado polimórfico se a frequência do alelo mais comum não exceder a 0,95 de probabilidade; os valores entre parênteses correspondem aos desvios padrões e entre colchetes aos intervalos de confiança.

Genetic Diversity Indexes	Populations			
	Savanna	Semi-Deciduous	Riparian	
	adults	adults	adults	Progenies
Heterozygosity mean observed (\hat{H}_o)	0,321 (0,064)	0,421 (0,110)	0,483 (0,103)	0,390 (0,069)
Heterozygosity mean expected (\hat{H}_e)	0,420 (0,081)	0,435 (0,078)	0,396 (0,077)	0,368 (0,060)
Percentage polymorphic loci* (0,95) (P)	81,8	87,5	72,7	75,0
Fixation index (f)	0,239 [0,087 a 0,374]	0,063 [-0,307 a 0,413]	-0,203 [-0,334 a -0,033]	-0,033 [-0,155 a 0,085]
Number of alleles (A)	2,5 (0,30)	2,4 (0,30)	2,2 (0,20)	2,5 (0,20)
Sample size (n)	19,4 (0,40)	14,8 (1,40)	18,6 (2,30)	326,4 (19,0)

The results showed that there were no significant differences for the heterozygosity observed (\hat{H}_o), heterozygosity expected (\hat{H}_e) and alleles number values among population and the \hat{H}_o and \hat{H}_e parameters showed that the species has high heterozygosity levels when compared with the average of forest species.

Cockerham's coancestry coefficient (\hat{q}_p) obtained from the analysis of the variance of the allelic frequencies, as suggested by Cockerham & Weir (1993), was used as a measure of genetic diversity among populations to estimate the gene flow (Table 5). The relationship between the genetic divergence among populations and the

migration intensity (\hat{N}_m) is given by the formula of Crow & Aoki (1984). The indirect estimates of the neighborhood size (\hat{N}_b) was made using a model of stepping-stones of population structure [Barton formula, (not published, citado por Slatkin & Barton, 1989)].

Table 5 shows that the mean number of migrants per generation (\hat{N}_m) was 0.41, indicating an apparent low gene flow, and the (N_b) value was about two individuals. According to Slatkin & Barton (1989), N_m and N_b determine whether genetic drift, alone, can produce substantial among location genetic variability. If N_m and N_b are greater than 1.0, the gene flow will be high enough to prevent differentiation due to genetic drift. Thus these values indicate that the gene flow was not enough to prevent population differentiation due to genetic drift, and can be

considered, however, relatively small when compared to the values obtained from other tropical tree species. They also indicated that the number of parents that randomly exchanged alleles was close to two individuals. A lower number of individuals in the same unit of area would imply in an increase in the among neighboring population divergence due to the internal increase of the endogamy levels, which alter the genetic structure of the populations in time (Reis, 1996). Furthermore, the effects of genetic drift would be pronounced, possibly leading to loss of the lower frequency alleles. Additionally, the heterozygosity levels would decrease, reducing the genetic diversity in the population. This reduction in diversity would have direct consequences in the population dynamics, implying a reduction in the dynamism of the species natural populations.

Table 5. Gene flow (\hat{N}_m) and indirect estimates of the neighborhood size (\hat{N}_b) obtained from genetic divergence estimates (\hat{q}_p) between pairs of three adult *Copaifera langsdorffii* Desf. populations and among the three populations. n = number of sampled individuals; np = number of populations.

Tabela 5. Fluxo gênico (\hat{N}_m), estimativas indiretas do tamanho da vizinhança (\hat{N}_b) obtidos a partir de estimativas de divergência genética (\hat{q}_p) de três populações de adultos de *Copaifera langsdorffii* Desf. tomadas duas a duas e das três populações juntas.

Populations	n	np	\hat{q}_p	\hat{N}_m	\hat{N}_b
Savanna and Semi-Deciduous	40	2	0,073	0,79	4,96
Savanna and Riparian	53	2	0,142	0,38	2,39
Semi-Deciduous and Riparian	53	2	0,162	0,32	2,01
Savanna, Semi-Deciduous and Riparian	73	3	0,131	0,41	2,58

The \hat{N}_m value calculated from the genetic divergence reflects the gene flow which occurred over a long period. It does not indicate whether there is any gene flow occurring in a given reproductive event, but estimates the levels of gene flow which must have occurred in the past to produce the observed genetic structure patterns. Thus the value of \hat{N}_m , instead of estimating the present gene flow, may indicate a narrower genetic base between the Savanna and Semi-deciduous forest populations than between the Savanna and Riparian forest or even between the Semi-deciduous and Riparian forests. It suggested that these populations were closely linked in the past or that they originated from a common ancestral. This is coherent with the hypothesis that the Savanna originated from a population previously linked to the Semi-deciduous forest, which in the past must have been a single population and was fragmented to give rise to these distinct units.

The effective population size \hat{N}_e is a very important parameter to know for germplasm preservation purposes, *in situ* conservation and seed collection, as it deals with sample genetic representativeness. According to Vencovsky (1987), the population effective size is very important in seed collecting activities of self pollinating plants since it is not sufficient to collect hundreds or even millions of seeds from only a single plant, because they will not be worth more than four seeds collected from four plants. Consequently, knowledge of effective size is essential to design the optimum collection strategies.

From the \hat{N}_e adult plant estimates (Table 6), it can be stated that the 20, 20 and 33 individuals (n) sampled in the Savanna, Semi-deciduous forest and Riparian forest populations represent, genetically, 16.1, 18.9

and 41.5 plants, respectively, from an ideal panmictic population.

For the Savanna population, the effective size was less than the number of sampled individuals. This difference is due to a lower genetic variation, which was a consequence of the endogamy found in the adult plants of this population. The effective size of the Semi-deciduous forest population was close to the number of individuals and, for the Riparian forest, the effective population size was greater than the number of individuals. This latter result is in line with the detected fixation indexes that indicated absence of endogamy in the adult plants of this population. The superiority resulted from the high heterozygosity levels detected in this generation, because the heterozygotes always carry two different alleles and may represent a greater number of individuals in the population.

The effective size estimated from the progenies was fairly low compared to the number of individuals (Table 6), which is naturally expected due to the relatedness among the progeny seedlings. This estimate showed that the 396 seedlings from the Riparian forest population corresponded to 56 individuals of an ideal panmictic population.

The $\hat{N}_{e/\text{parent}} = 2.80$ (55.93/20) value was considered of average magnitude. It resulted from the finite size of the offspring of the parent, which was 20 plants. The maximum value assumed by $\hat{N}_{e/\text{parent}}$ for infinite sample sizes, or in the order of hundreds for self-pollinating species, is $4F$ where F is the number of parents. Therefore, this value can reach the value of four and, in turn, can attain a maximum value of 80 (4×20). The population effective size was 55.93, which corresponded to 69.91% of the maximum. This value was considered good but revealed the presence of homozygous among the progenies,

which is in agreement with the fixation indexes of Wright.

C. langsdorffii Desf. seeds should ideally be collected in equal quantities per female parent, from at least 25 individuals randomly sampled in the populations. This strategy would guarantee the genetic representativeness of the seeds. The recuperation of degraded

areas with seeds collected following these principles would result in populations with high evolutionary potential due to the possibility of infinite new genotypic recombinations, which in turn will increase the adaptability of the species to the environment, to future environmental changes and to new areas.

Table 6. Effective size (\hat{N}_e) and number of individuals (n) of three natural *Copaifera langsdorffii* Desf. populations calculated from data on adult individuals and progenies of one of the populations.

Tabela 6. Tamanho efetivo (\hat{N}_e) e número de indivíduos (n) de três populações naturais de *Copaifera langsdorffii* Desf. a partir de dados de indivíduos adultos e de progênie de uma das populações.

Populations (Adults)	\hat{N}_e	$\hat{N}_{e/parent}$	n
Savanna	16,1	---	20
Semi- Deciduous	18,9	---	20
Riparian	41,5	---	33
Mean	25,5	---	24,3
Population (Progenies)	55,9	2,80	396

4 CONCLUSIONS

The study of natural *Copaifera langsdorffii* Desf. populations by isoenzymatic data allowed the following conclusions to be drawn:

The analyses of the within and among population allelic frequencies variability indicated that the three adult populations are significantly different and that there is a balance among the generations of the Riparian forest population.

The genetic structure of the populations was characterized by greater genetic variability within populations (87%) than among populations (13%).

The Savanna and Semi-deciduous forest populations are relatively more similar than the Riparian forest population.

The genetic diversity indices were higher than those of other tropical tree species, suggesting that the studied populations are amenable to *in situ* genetic conservation.

Most of the adult population loci are in HWE, suggesting that these populations were generated by random crosses.

The gene flow estimates indicated that there is greater similarity between the Savanna and Semi-deciduous forest populations than between the Savanna and Riparian forest populations or between the Semi-deciduous forest and Riparian forest populations. This

shows that they were either more linked previously or that they originated from a common ancestral population.

5 ACKNOWLEDGEMENTS

The authors are grateful to the CNPq (300542/89-5). We also wish to thank Sofia Caetano for her help with the english version of the manuscript.

6 BIBLIOGRAPHIC REFERENCES

- ALFENAS, A. C. **Eletroforese de isoenzimas e proteínas afins; fundamentos e aplicações em plantas e microrganismos**. Viçosa: Universidade Federal de Viçosa. 1998. 574 p.
- ALFENAS, A. C.; PETERS, I.; BRUNE, W.; PASSADOR, C. G. **Eletroforese de proteínas e isoenzimas de fungos e essências florestais**. Viçosa: Universidade Federal de Viçosa. 1991. 242 p.
- ALVAREZ-BUYLLA, E. R.; GARAY, A. A. Population genetic structure of *Cecropia obtusifolia*, a tropical pionner tree species. **Evolution**, Lawrence, v. 48, n. 2, Apr. p. 437-453, 1994.
- BRASIL. Ministério da Agricultura. Departamento Nacional de Metereologia. **Normas climatológicas (1961-1990)**, Brasília, 1992. 132 p.
- BROWN, A. H. D.; MATHESON, A. C.; ELDRIDGE, K. G. Estimation of the mating system for *Eucalyptus obliquia* L' Herit. using allozyme polymorphism. **Australian Journal of Botany**, Collingwood, v. 23, n. 6, p. 931-949, 1975.
- BROWN, A. H. D.; WEIR, B. S. Measuring genetic variability in plant populations. In: TANKSLEY, S. D.; ORTON, T. J. (Ed.). **Isozymes in plant genetics and breeding: part. A**. Amsterdam: Elsevier Science, 1983. p. 219-239.
- CARVALHO, P. E. R. **Espécies florestais brasileiras: recomendações silviculturais, potencialidades e uso de madeira**. Brasília: EMBRAPA-CNPq, 1994. 640 p.
- CIAMPI, A. Y. **Desenvolvimento e utilização de marcadores microsátélites, AFLP e sequenciamento de cpDNA, no estudo da estrutura genética e parentesco em populações de Copaíba (*Copaifera langsdorffii*) em matas de galeria no cerrado**. 1999. 109 p. Tese (Doutorado) - Universidade Estadual Paulista "Julio de Mesquita Filho"- Instituto de Bociências, Botucatu.
- COCKERHAM, C. C.; WEIR, B. S. Estimation of gene flow from F-statistics. **Evolution**, Lawrence, v. 47, n. 3, p. 855-863, June 1993.
- CROW, J. F.; AOKI, K. Group selection for polygenic behavioral trait: estimating the degree of population subdivision. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 81, n. 19, p. 6073-6077, 1984.
- FUNDAÇÃO INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Mapa de vegetação do Brasil**. Rio de Janeiro, 1993. (Mapa).
- FUTUYMA, D. J. **Biologia evolutiva**. 2. ed. Ribeirão Preto: Sociedade Brasileira de Genética/CNPq, 1992. 646 p.
- GANDARA, F. B. **Diversidade genética, taxa de cruzamento e estrutura espacial dos genótipos em uma população de *Cedrela fissilis* Vell. (Meliaceae)**. 1996. 69 p. Dissertação (Mestrado) – Universidade de Campinas, Campinas.
- GE, S.; HONG, D.; WANG, H.; LIU, Z.; ZHANG, C. Population genetic structure and conservation of an endangered conifer, *Cathaya argyrophylla* (Pinaceae). **International Journal Plant Science**, Chicago, v. 159, n. 2, p. 351-357, Mar. 1998.

- HAMRICK J. L.; MURAWSKI, D. A.; NASON, J. D. The influence of seed dispersal mechanisms on the genetic structure of tropical tree populations. **Vegetation**, Dordrecht, v. 107/108, p. 281-297, June 1993.
- HILL, R. J.; PRANCE, G. T.; MORI, S. A.; STEWARD, W. C.; SHIMABUKURU, D.; BERNARD, J. Estudo eletroforético da dinâmica de variação genética em três taxa ribeirinhas ao longo do rio Solimões, América do Sul. **Acta Amazônica**, Manaus, v. 8, n. 2, p. 183-199, abr./jun. 1978.
- KITAMURA, K.; RAHMAN, M. Y. B. A. Genetic diversity among natural populations of *Agathis borneensis*, a tropical rain forest conifer from Brunei Darussalam, Borneo, Southeast Asia. **Canadian Journal Botany**, Ottawa, v. 70, n. 10, p. 1945-1949, Oct. 1992.
- LEWIS, P. O.; ZAYKIN, D. **Genetic data analysis**: computer program for the analysis of allelic data. Version 1. 0 (d15). Free program distributed by the authors over the internet from the GDA. Disponível em: <<http://alleyn.eeb.uconn.edu/gda/2000>>. Acesso em: 2004.
- LOVELESS, M. D.; HAMRICK, J. L.; FOSTER, R. B. Population structure and mating system in *Tachigali versicolor*, a monocarpic neotropical tree. **Heredity**, Oxford, v. 81, n. 2, p. 134-143, Aug. 1998.
- LUNDKWIST, K.; RUDIN, D. Genetic variation in eleven populations of *Picea abies* as determined by isozyme analysis. **Hereditas**, Landskrona, v. 85, n. 1, p. 67-74, 1977.
- MORAES, M. L. T. **Variabilidade genética por isoenzimas e caracteres quantitativos em duas populações naturais de aroeira *Myracrodruon urundeuva* F. F. e M. F. Allemão (Anacardiaceae) Syn: *Astronium urundeuva* (Fr. Allemão) Engler**. 1997. 139 p. Tese (Doutorado) - Escola Superior de Agricultura "Luiz de Queiroz", Piracicaba.
- MORAN, G. F.; HOPPER, S. D. Genetic diversity and the insular population structure of the rare granite rock species, *Eucalyptus caesia* Benth. **Australian Journal Botany**, Collingwood, v. 31, n. 2, p. 161-172, 1983.
- NEGRÃO, M. O. **Estrutura genética de populações naturais de *Syagrus romanzoffiana* (Arecaceae) do continente (Serra do Mar, S. P.) e da Ilha dos Alcatrazes**. 1999. 78 p. Dissertação (Mestrado em Genética e Melhoramento de Plantas) - Escola Superior de Agricultura "Luiz de Queiroz", Piracicaba.
- NEI, M. **Molecular evolutionary genetics**. New York: Columbia University. 1987. 512 p.
- OLIVEIRA, A. F.; CARVALHO, D.; ROSADO, S. C. S. Taxa de cruzamento e sistema reprodutivo de uma população natural de *Copaifera langsdorffii* Desf, na região de Lavras, MG, a partir de isoenzimas. **Revista Brasileira de Botânica**, São Paulo, v. 25, n. 3, p. 331-338, jul./set. 2002.
- PETERS, G. B.; LONIE, J. S.; MORAN, G. F. The breeding system, genetic diversity and pollen sterility in *Eucalyptus pulverulenta*, a rare species with small disjunct populations. **Australian Journal of Botany**, Collingwood, v. 38, n. 6, p. 559-570, 1990.
- REIS, M. S. **Distribuição e dinâmica da variabilidade genética em populações naturais de palmitheiro (*Euterpe edulis* Martius)**. 1996. 203 p. Tese (Doutorado em Agronomia) - Escola Superior de Agricultura "Luiz de Queiroz", Piracicaba.
- ROCHA, O. J.; LOBO, J. A. Genetic variation and differentiation among five populations of the guanacaste tree (*Enterolobium cyclocarpum* Jacq.) **International Journal of Plant Science**, Chicago, v. 157, n. 2, p. 234-239, Mar. 1996.
- SEBBENN, A. M. **Estrutura genética de subpopulações de *Genipa americana* L. (Rubiaceae) a partir de isoenzimas**. 1997. 107 p. Dissertação (Mestrado em Ciências Florestais) -

Escola Superior de Agricultura "Luiz de Queiroz", Piracicaba.

SHAW, D. V.; BROWN, A. H. D. Optimum number of marker loci for estimating outcrossing in plant populations. **Theoretical and Applied Genetic** Berlin, v. 61, n. 4, p. 321-325, 1982.

SHEELY, D. L.; MEAGHER, T. Genetic diversity in Micronesian island populations of the tropical tree *Campnosperma brevipetiolata* (Anacardiaceae) **American Journal of Botany**, Columbus, v. 83, n. 12, p. 1571-1579, Dec. 1996.

SLATKIN, M.; BARTON, N. H. A comparison of three indirect methods for estimating average levels of gene flow. **Evolution**, Lawrence, v. 43, n. 7, p. 1349-1368, Nov. 1989.

SOUZA, L. M. F. I. **Estrutura genética de populações naturais de *Chorisia speciosa* St. Hil. (Bombacaceae) em fragmentos florestais na**

região de Bauru (SP). 1997. 76 p. Dissertação (Mestrado em Ciências Florestais) - Escola Superior de Agricultura "Luiz de Queiroz", Piracicaba.

SWOFFORD, D. L.; SELANDER R. B. **BIOSYS-1**. A computer program for the analysis of allelic variation in population genetics and biochemical systematics. Release 1. 7. Illinois: Natural History Survey, 1989. 43 p.

VENCOVSKY, R. Análise de variância de frequências alélicas. In: CONGRESSO LATINO AMERICANO DE GENÉTICA, 10., 1992, Rio de Janeiro. **Proceedings...** Rio de Janeiro, 1992. (Publicado na Revista Brasileira de Genética, Ribeirão Preto, v. 15, p. 53-60, 1992. Suplemento, 1.

VENCOVSKY, R. Tamanho efetivo populacional na coleta e preservação de germoplasma de espécies alógamas. **IPEF**, Piracicaba, v. 35, p. 79-84, abr. 1987.