

Genetic diversity of *Annona crassiflora* (Annonaceae) in northern Minas Gerais State

L.G. Cota¹, F.A. Vieira², A.F. Melo Júnior¹, M.M. Brandão³,
K.N.O. Santana¹, M.L. Guedes¹ and D.A. Oliveira¹

¹Laboratório de Bioprospecção e Recursos Genéticos,
Departamento de Biologia Geral, Universidade Estadual de Montes Claros,
Montes Claros, MG, Brasil

²Unidade Acadêmica Especializada em Ciências Agrárias,
Universidade Federal do Rio Grande do Norte, Macaíba, RN, Brasil

³Programa de Pós-Graduação da Universidade Federal de Lavras,
Lavras, MG, Brasil

Corresponding author: D.A. Oliveira
E-mail: dario.oliveira@unimontes.br

Genet. Mol. Res. 10 (3): 2172-2180 (2011)

Received December 2, 2010

Accepted August 12, 2011

Published September 23, 2011

DOI <http://dx.doi.org/10.4238/vol10-3gmr1188>

ABSTRACT. Genetic diversity analyses of tropical tree species are relevant to landscape management, plant genetic resource inventory, and biological conservation of threatened species. *Annona crassiflora* is an endangered fruit tree native to the Cerrado biome that is threatened by reduction of natural populations and fruit extraction. We examined the intra- and interpopulational genetic diversity of this species in the northern region of Minas Gerais State. Seventy-two individuals from four natural populations were genotyped using RAPD markers. We found moderate genetic diversity among populations, with Shannon's I index varying between 0.31 and 0.44, and Nei's genetic diversity (H_E) for the population set equal to 0.31. AMOVA indicated a greater genetic

variation within (77.38%) rather than among populations (22.62%), tending towards isolation by distance (Mantel's $r = 0.914$; $P = 0.089$). Nei's genetic identity estimates among populations revealed a hierarchical pattern of genetic similarity of form [(CA1, CA2), MC], [(GM)], corroborating the high genetic differentiation between spatially isolated populations.

Key words: *Annona crassiflora*; Northern of Minas Gerais; Genetic diversity

INTRODUCTION

The Cerrado biome is responsible for a large part of the megadiversity found in Brazil (Mittermeier et al., 2005), standing out as the second largest biome of the country and covering nearly 23% of the national territory (Ribeiro and Walter, 1998). There are numerous plant species with high economic potential (Caramori et al., 2004), mainly because of the diversity of native fruit trees with high nutritional value, medicinal properties, and their potential for use by the lumber industry (Ribeiro and Rodrigues, 2006).

Among the typical fruit trees of Cerrado formations is *Annona crassiflora* Mart. (Annonaceae), popularly known as "marolo", "araticum" or "cabeça de negro". In Brazil, *A. crassiflora* is distributed in the States of MG, SP, MS, GO, BA, TO, MA, PI, PA, and the Federal District, and occurs in the Cerrado, "Cerradão", and "campos rupestres" (Lorenzi, 1998; Silva Junior, 2005). It is a deciduous, heliophytic, xerophytic, hermaphrodite tree reaching up to 8 m high (Lorenzi, 1998). The fruits are appreciated *in natura* or in delicacies, since the pulp is sweet and very aromatic; the leaves have anti-diarrhea and antimicrobial properties, and the seeds are used as insecticides (Silva Junior, 2005).

The extractivism of "araticum" is lucrative because of its multiple uses, but uncontrolled exploitation may affect perpetuation of the species (Ribeiro and Pasqual, 2005). Therefore, *A. crassiflora* has become a target species for research regarding conservation-oriented genetic characterization. Knowledge of the structure of native populations is one way of guiding future actions by defining adequate monitoring parameters.

One of the tools for determining population structure and conserving biodiversity is the use of conservation genetics, which seeks a better understanding of population dynamics (Santos and Redondo, 2003). Conservation programs are already using data generated by molecular techniques that evaluate the genetic structure and diversity of natural populations. Using molecular markers, it is also possible to quantify the variability level within populations and to evaluate their distribution among original populations (Souza et al., 2007; Vieira and Carvalho, 2008). RAPD (random amplified polymorphic DNA) is an established technique commonly used to provide information regarding DNA-level variability for many applications in genetic analyses (Ferreira and Grattapaglia, 1995). The wide use of these markers in conservation programs can be credited to their usability for any kind of organism and to their rapid results (Lopes et al., 2002).

Thus, this study had the objective of characterizing the genetic diversity of four natural *A. crassiflora* populations in northern of Minas Gerais State through RAPD markers, generating information for management and conservation strategies.

MATERIAL AND METHODS

Study and sampling sites

A total of 72 individuals were randomly selected from four natural *A. crassiflora* Mart. populations in northern Minas Gerais State (Table 1). The geographical distances of the populations varied between 9.7 km (CA1 and CA2) and 195 km (GM and CA2). The populations were found in diversified habitats: an agroforestry reserve (GM), an ecological reserve (MC) and a private property (CA1 and CA2). All the individuals in the populations were marked and identified according to their geographical coordinates. Young leaves were collected and taken to the Bioprospection and Genetic Resources Laboratory of the State University of Montes Claros (UNIMONTES), where they were kept in a freezer at -20°C until DNA extraction.

Table 1. Location of the *Annona crassiflora* Mart. populations studied and the number (N) of individual samples.

Population	Location	N	Coordinates	Altitude (m)
GM	Grão Mogol, MG	21	16°22'47"S 43°00'29"W	800
MC	Montes Claros, MG	21	16°54'92"S 43°55'88"W	1003
CA1	Campo Azul, MG	15	16°26'23"S 44°50'01"W	686
CA2	Campo Azul, MG	15	16°26'27"S 44°44'33"W	637

Genomic DNA extraction and amplification

DNA extraction was performed according to the protocol proposed by Doyle and Doyle (1990) and modified by Faleiro et al. (2002). The extracted DNA was quantified in a spectrophotometer measuring absorbance at 260 and 280 nm. The integrity of the DNA was determined on a 1.2% agarose gel electrophoresis and photographed under ultraviolet light. Next, the DNA samples representative of each population were submitted to the polymerase chain reaction (PCR)-RAPD technique. The amplification reactions were carried out in a final volume of 25 µL, with 17 primers tested. Each reaction consisted of 10 mM Tris-HCl, 50 mM KCl, 2 mM MgCl₂, 100 µM of each deoxynucleotide (dATP, dTTP, dGTP, and dCTP), 0.4 µM primer, 1 U Taq DNA Polymerase, and 50 ng template DNA. The reactions were amplified in an Applied Biosystems thermal cycler model Veriti 96 Well. The molecular separation of the amplified products was performed by horizontal electrophoresis on a 1.2% agarose gel for 3 h at 120 V, after which the gels were visualized and photographed using photodocumentation equipment through UV Vilber Loumart transillumination.

Data analysis

After analyzing the gel photographs (genotyping), a phenotypic matrix composed of zeros and ones was generated, with 0 indicating the absence of a band and 1 indicating the presence. In order to determine with precision the genetic similarity estimates

between individuals, a bootstrap analysis (10,000 permutations) was performed to verify whether the number of generated polymorphic markers was sufficient. Genetic similarity was estimated for each pair of individuals using the GENES program (Cruz, 2001). The parameters sum of the squared deviations in relation to the resamplings and stress value (S), which indicate the adjustment between the original and the sampling matrix, were made available during the analyses. The number of polymorphic fragments is considered to be precise when stress does not reach the value of 0.05 (Kruskal, 1964).

The estimates of intrapopulation genetic diversity and of the population set were obtained using the POPGENE 1.32 program (Yeh et al., 1997), with P as the percentage of polymorphic loci, I as Shannon's index (Lewontin, 1972), and H_E as Nei (1978) genetic diversity. A locus was considered to be polymorphic for the population when at least two individuals were different from the others in terms of presence or absence of bands. The H_E and I indexes were compared among the four populations through the analysis of variance (ANOVA, F -test) at a probability of 95% and complemented with Bonferroni's *a priori* t -test (Rice, 1989). Nei (1978) genetic identity matrix among populations obtained in the POPGENE program was used for the construction of the genetic similarity dendrogram using the UPGMA method (unweighted pair-group method with arithmetic averages), adopting the SAHN routine (sequential agglomerative, hierarchical and nested clustering) in the NTSYS version 2.11x program (Rohlf, 2000). The consistency of the clusters (dendrogram nodes) was verified using the TFPGA 1.3 program (tools for population genetic analysis) (Miller, 1997), by means of 1000 bootstraps.

The genetic structure of the populations was inferred through an analysis of molecular variance (AMOVA) obtained with the aid of the ARLEQUIN version 3.11 program (Excoffier et al., 2002). In order to evaluate the spatial genetic variation patterns of the populations, Pearson's correlation coefficient (r) was estimated between the genetic distance matrixes (calculated from Jaccard's similarity index) and the geographical distance matrix of the populations. The isolation by distance hypothesis was tested among matrixes by Mantel's Z -statistic, with the option of using the Monte Carlo test (1000 permutations) of the PC-ORD 4.10 program (McCune and Mefford, 1997).

RESULTS

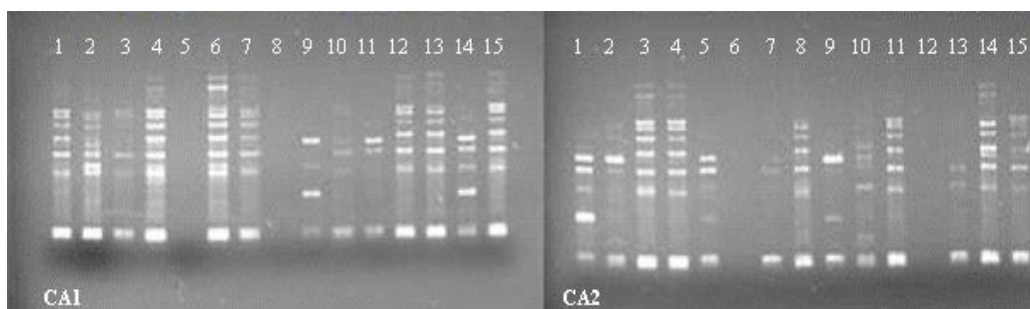
Based on the amplification patterns, 15 RAPD primers were selected from a total of 17 primers tested. These primers produced unequivocal fragments that allowed the interpretation of the data. Of the 140 loci generated from the primers, 123 were polymorphic (Table 2). Each utilized primer revealed from 4 to 13 polymorphic loci, with an average of nine polymorphic loci per locus. Figure 1 shows the amplification profile obtained with primer P04 in *A. crassiflora* individuals belonging to the CA1 and CA2 populations.

The values of the genetic diversity indexes (H_E and I) estimated for the four *A. crassiflora* populations and compared by ANOVA presented significant differences between the GM \times CA1 and between the GM \times CA2 populations (Table 3). No significant differences were observed comparing the other value pairs.

The distribution of variability among and within the natural *A. crassiflora* populations obtained by AMOVA revealed that the highest percentage of genetic variation occurred within populations (77.38%).

Table 2. Selected primers with their corresponding nucleotide sequences, total number of fragments and number of polymorphic loci.

Primer	Nucleotide sequence	Total number of fragments	Number of polymorphic fragments
P01	5'-GTGACATGCC-3'	6	4
P02	5'-AGATGCAGCC-3'	11	11
P03	5'-ATGGCTCAGC-3'	13	12
P04	5'-CAGGCCCTTC-3'	9	8
P05	5'-CTCTTGGGCT-3'	13	11
P06	5'-ACCACCCGCT-3'	13	12
P07	5'-ACCCACACAC-3'	13	12
P08	5'-AGCGTGTCTG-3'	5	4
P09	5'-TCACGTCCAC-3'	5	4
P10	5'-GGCTCATGTG-3'	10	9
P11	5'-GTCAGGGCAA-3'	15	13
P12	5'-GGTGCACGTT-3'	6	5
P15	5'-GCTTGTGAAC-3'	5	5
P16	5'-GGGACCGTGT-3'	8	7
P17	5'-GGGTCTCGGT-3'	8	6
Total		140	123

**Figure 1.** Profile of the fragments amplified by primer P04 and detected on agarose gel under UV light for *Annona crassiflora* individuals belonging to the CA1 and CA2 populations.**Table 3.** Genetic diversity indexes for four natural *Annona crassiflora* Mart. populations from northern Minas Gerais.

Populations	$H_e \pm SD$	$I \pm SD$	P
GM	0.21 \pm 0.203	0.31 \pm 0.290	59.31
MC	0.25 \pm 0.203	0.37 \pm 0.285	71.05
CA1	0.28 \pm 0.199	0.41 \pm 0.275	76.55
CA2	0.30 \pm 0.190	0.44 \pm 0.261	81.38
F_{ANOVA}	6.418*	6.706*	-
For the population set	0.31 \pm 0.168	0.47 \pm 0.224	91.72

*Significant at 95% probability. H_e = Nei's genetic diversity; I = Shannon's index; P = percentage of polymorphic loci.

The GM population was more genetically distant than the others, with a mean genetic distance of 0.139. The CA1 and CA2 populations presented a smaller genetic distance among themselves. The dendrogram constructed by the UPGMA method based on Nei's genetic identity matrix (Nei, 1978) revealed a hierarchical pattern of genetic similarity of form [(CA1, CA2), MC], [(GM)] (Figure 2).

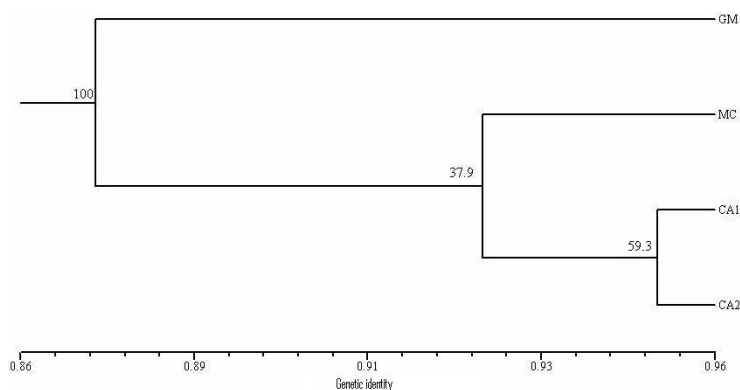


Figure 2. Dendrogram constructed from Nei's genetic identity matrix (Nei, 1978) using the UPGMA method for four *Annona crassiflora* Mart. populations. The consistency values of the nodes (%) are indicated.

DISCUSSION

According to Ferreira and Grattapaglia (1995), RAPD markers are sensitive and may generate different quantities of amplified fragments depending on the quality and quantity of the DNA used, as well as the amplification conditions. The number of fragments observed in this study was satisfactory for interpretation and conversion into molecular data for the populations studied, since according to Kruskal (1964), stress values under 0.05 are a good sign of precision.

The resampling method (bootstrap) performed to verify whether the number of loci was enough to determine with precision genetic similarity estimates among genotypes indicated that starting from 105 bands, the stress value (S) was 0.049 and the correlation level between original and sample matrixes was 0.969, suggesting a sampling sufficiency with the 140 bands analyzed.

The percentage of polymorphic loci was lower for the GM population (59.31%), varying between 71.05 and 81.38% in the other *A. crassiflora* populations. Similar results were found for *Schizolobium parahyba* populations, with polymorphism percentages varying from 65.6 to 87.5% (Freire et al., 2007). Torezan et al. (2005) found polymorphism rates of 41 and 83% among adult *Aspidosperma polyneuron* (Apocynaceae) populations using RAPD data. Zimback et al. (2004) suggested that the natural *Trichilia pallida* populations are not genetically isolated because of the high polymorphism index found (90.3 to 97.3%). Analogously, it is arguable that the GM population sampled in the present study tends toward genetic distancing owing to the lower polymorphic loci rate presented.

The estimates of Nei's genetic diversity (H_E) indicated that the CA1 and CA2 populations presented the highest diversity values: 0.28 and 0.30. An H_E of 0.31 was observed for the *A. crassiflora* population set. These results were in conformity with other studies that used RAPD marker data. For example, Freire et al. (2007) evaluated natural *Schizolobium parahyba* (Leguminosae) populations and found diversity indexes between 0.25 and 0.37; furthermore, in a study of the genetic variability of *Dimorphandra mollis* populations, Oliveira et al. (2008) obtained an H_E varying between 0.17 and 0.28.

The correlation between the genetic and geographical distances of the populations indicated a tendency toward isolation by distance (Mantel's $r = 0.914$; $P = 0.089$).

The populations studied presented similar genetic diversity indexes. The I values evidence a moderate genetic diversity within *A. crassiflora* populations when the results are compared to those obtained by Moura (2005), who used RAPD markers and observed a high genetic diversity in *Eremanthus erythropappus* populations, with values of Shannon's index between 0.498 and 0.539. This index varies between zero and one, with values closer to one indicating greater genetic diversity (Estopa et al., 2006). While studying natural populations of *Plathymenia reticulata*, Lacerda et al. (2001) verified that populations with more elevated values of Shannon's index presented higher molecular variance rates and vice-versa, evidencing a clear correspondence with AMOVA.

Variability distribution in the populations analyzed in this study was in conformity with that expected for tree species, which generally present greater genetic variation within populations (Zimback et al., 2004). However, the results obtained in the present study evidenced the existence of an elevated variation between populations (22.62%). It is probable that natural or artificial barriers are hindering allelic flow performed by pollinators and dispersers. The pollination of *A. crassiflora* is performed by coleopterans, and seed dispersion is carried out by gravity and animals. According to Loveless and Hamrick (1984), entomophily may limit the movement of pollen to less foraged sites, thus enhancing differentiation among populations. The same authors also reported that seed dispersion performed by gravity or animals that ingest the seeds results in intermediary levels of genetic variation among populations. It is also important to emphasize that the genetic differentiation among populations may not be a result of current gene flow, but rather of gene flow that occurred in the past when the forests were continuous (Moraes et al., 2005). Therefore, it is assumed that the divergence found among the *A. crassiflora* populations is probably a reflection of the isolation between populations that occurred because of the fragmentation processes and reduction of population sizes.

Analyzing the dendrogram, a tendency toward grouping is observed for geographically closer populations. In fact, the geographically closer CA1 and CA2 populations presented higher genetic identity, grouping with a similarity of 95.5%. The spatial proximity between these populations probably offers higher allelic flow, thus allowing smaller genetic differentiation. Nevertheless, consistency analysis of the groupings revealed that the group formed by the MC, CA1 and CA2 populations presented low consistency (37.9%). Wendt et al. (2007) emphasize that when a low consistency is obtained, it is necessary to raise the number of markers used with the goal of securing higher precision of the results.

Therefore, the general results obtained show that the levels of genetic diversity and polymorphism for the species are moderate, which shows the potential for genetic conservation of the populations. The variation rates indicate isolation by distance, which emphasizes the importance of maintaining the levels of genetic variation within and among *Annona crassiflora* Mart. populations.

ACKNOWLEDGMENTS

We thank the Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) for the Research and Technological Development Incentive (BIPDT) scholarships, the Capacitation Program of Human Resources (PCRH) scholarships, the

Regional Scientific and Technological Development (BDCTR) scholarship, and the Scientific Initiation scholarship (PROBIC) granted to the authors; the State University of Montes Claros (UNIMONTES) for logistics support, and the Alternative Agriculture Center of Northern Minas Gerais - CAA for support and aid in selecting the study areas.

REFERENCES

- Caramori SS, Lima CS and Fernandes KF (2004). Biochemical characterization of selected plant species from Brazilian Savannas. *Braz. Arch. Biol. Technol.* 47: 253-259.
- Cruz CD (2001). Programa GENES: Versão Windows. Universidade Federal de Viçosa, Viçosa.
- Doyle JJ and Doyle JL (1990). Isolation of plant DNA from fresh tissue. *Focus* 12: 13-15.
- Estopa RA, Souza AM, Moura MCO, Botrel MCG, et al. (2006). Diversidade genética em populações naturais de candeia (*Eremanthus erythropappus* (DC.) MacLeish). *Sci. Forestalis* 70: 97-106.
- Excoffier L, Laval G and Schneider S (2002). Arlequin: a Software for Population Data Analysis. Version 3.1. University of Geneva, Geneva.
- Faleiro FG, Araújo IS, Bahia RCS, Santos RF, et al. (2002). Otimização da extração e amplificação de DNA de *Theobroma cacao* L. visando obtenção de marcadores RAPD. *Agrotropica* 14: 31-34.
- Ferreira ME and Grattapaglia D (1995). Introdução ao Uso de Marcadores RAPD e RFLP em Análise Genética. Embrapa - Cenargen, Brasília.
- Freire JM, Piña-Rodrigues FCM, Lima ER, Sodrê SRC, et al. (2007). Estrutura genética de populações de *Schizolobium parahyba* (Vell.) Blake (guapuruvu) por meio de marcadores RAPD. *Sci. Forestalis* 74: 27-35.
- Kruskal JB (1964). Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika* 29: 1-27.
- Lacerda DR, Acedo MD, Filho JP and Lovato MB (2001). Genetic diversity and structure of natural populations of *Plathymenia reticulata* (Mimosoideae), a tropical tree from the Brazilian Cerrado. *Mol. Ecol.* 10: 1143-1152.
- Lewontin RC (1972). The Apportionment of Human Diversity. In: *Evolutionary Biology* (Dobzhansky T, Hecht MK and Steere WC, eds.). Appleton-Century-Crofts, New York, 381-398.
- Lopes R, Lopes MTG, Figueira AVO, Camargo LEA, et al. (2002). Marcadores moleculares dominantes (RAPD e AFLP). *Biotecnol. Cien. Desenvol.* 29: 56-60.
- Lorenzi H (1998). Árvores Brasileiras/Manual de Identificação e Cultivo de Plantas Arbóreas Nativas do Brasil. 2ª ed. Editora Plantarum, Nova Odessa.
- Loveless MD and Hamrick JL (1984). Ecological determinants of genetic structure in plant populations. *Ann. Rev. Ecol. Syst.* 15: 65-95.
- McCune B and Mefford MJ (1997). PC-ORD: Multivariate Analysis of Ecological Data. Version 3.0. Gleneden, Beach.
- Miller MP (1997). Tools for Population Genetic Analysis. Version 1.3. Northern Arizona University, Flagstaff.
- Mittermeier RA, Fonseca GAB, Rylands AB and Brandon K (2005). Uma breve história da conservação da biodiversidade no Brasil. *Megadiversidade* 1: 14-21.
- Moraes MLT, Kageyama PY and Sebbenn AM (2005). Diversidade e estrutura genética espacial em duas populações de *Myracrodruon urundeuva* Fr. All. Sob diferentes condições antrópicas. *Rev. Arvore* 2: 281-289.
- Moura MCO (2005). Distribuição da Variabilidade Genética em Populações Naturais de *Eremanthus erythropappus* (DC.) MacLeish por Isoenzimas e RAPD. Doctoral thesis, Universidade Federal de Lavras, Lavras.
- Nei M (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 89: 583-590.
- Oliveira DA, de Paula MFB, Pimenta MAS, Braga RF, et al. (2008). Variabilidade genética de populações de fava d'anta (*Dimorphandra mollis*) da região norte do Estado de Minas Gerais. *Rev. Arvore* 32: 355-363.
- Ribeiro JF and Walter BMT (1998). O Bioma Cerrado. In: *Cerrado: Ambiente e Flora* (Sano SM and Almeida SP, eds.). Embrapa - CPAC, Planaltina.
- Ribeiro MNO and Pasqual M (2005). Tecnologia da Produção do Marolo, Boletim. Editora UFLA, Lavras.
- Ribeiro RA and Rodrigues FM (2006). Genética da conservação em espécies vegetais do cerrado. *Rev. Cienc. Med. Biol.* 5: 253-260.
- Rice WR (1989). Analyzing tables of statistical tests. *Evolution* 43: 223-225.
- Rohlf FJ (2000). Numerical Taxonomy and Multivariate Analysis System. Version 2.11. Applied Biostatistics, New York.
- Santos FR and Redondo AF (2003). A Biotecnologia Aplicada à Conservação de Espécies Silvestres. In: *Biotecnologia de A a Z* (Borem A, Santos FR and Almeida MR, eds.). Editora Folha de Viçosa, Viçosa, 229.
- Silva Junior MC (2005). 100 Árvores do Cerrado, Guia de Campo. Rede de Sementes do Cerrado, Brasília.

- Souza AM, Carvalho D, Vieira FA, Nascimento LH, et al. (2007). Estrutura genética e espacial de populações naturais de *Calophyllum brasiliense* Camb. em mata de galeria. *Cerne* 13: 239-247.
- Torezan JMD, Souza RF, Ruas PM, Ruas CF, et al. (2005). Genetic variability of pre and post-fragmentation cohorts of *Aspidosperma polyneuron* Muell. Arg. (Apocynaceae). *Braz. Arch. Biol. Tech. Curitiba* 48: 171-180.
- Vieira FA and Carvalho D (2008). Genetic structure of an insect-pollinated and bird-dispersed tropical tree in vegetation fragments and corridors: implications for conservation. *Biodivers. Conserv.* 17: 2305-2321.
- Wendt SN, Sousa VA, Quoirin M, Sebbenn AM, et al. (2007). Caracterização genética de procedências e progênies de *Ilex paraguariensis* St. Hil. utilizando marcadores RAPD. *Sci. Forestalis* 73: 47-53.
- Yeh FC, Yang RC, Boyle TBJ, Ye ZH, et al (1997). POPGENE, the User-Friendly Shareware for Population Genetic Analysis Molecular Biology and Biotechnology Centre. University of Alberta, Edmonton.
- Zimback L, Mori ES, Kageyama PY, Veiga RFA, et al. (2004). Estrutura genética de populações de *Trichilia pallida* Swartz (Meliaceae) por marcadores RAPD. *Sci. Forestalis* 65: 114-119.