GMR

Diversity and genetic structure of the native Brazil nut tree (*Bertholletia excelsa* Bonpl.) population

J.C. Cabral[1], A.B. Baldoni[2], H. Tonini[2], V.C.R. Azevedo[3], L.D. Giustina[4], A.V. Tiago[1] and A.A.B. Rossi[1]

[1]Laboratório de Genética Vegetal e Biologia Molecular,
Faculdade de Ciências Biológicas e Agrárias,
Centro de Tecnologia da Amazônia Meridional,
Universidade do Estado de Mato Grosso, Campus de Alta Floresta,
Alta Floresta, MT, Brasil
[2]Embrapa Agrossilvipastoril, Sinop, MT, Brasil
[3]Embrapa Recursos Genéticos e Bioctecnologia, Brasília, DF, Brasil
[4]Programa de Pós-Graduação em Biotecnologia e Biodiversidade da Rede em Associação Ampla, Pró-Centro Oeste Universidade Federal de Mato Grosso, Cuiabá, MT, Brasil

Corresponding author: A.A.B. Rossi
E-mail: anabanrossi@unemat.br

Genet. Mol. Res. 16 (3): gmr16039702
Received April 17, 2017
Accepted June 13, 2017
Published July 6, 2017
DOI http://dx.doi.org/10.4238/gmr16039702

ABSTRACT. The present study was carried out in a native Brazil nut tree population (*Bertholletia excelsa* Bonpl., Lecythidaceae) to assess its genetic diversity and structure. Ten microsatellite markers were used to genotype 198 adult trees (*B. excelsa*). The population presented high genetic diversity and inbreeding absence rates. The empirical Bayesian method showed three distinct groups in the structure of this population. Molecular analysis of variance showed 98% variability within groups, and 2% between groups. The genetic divergence (F_{ST}) indicated little

difference between groups; thus, suggesting efficient gene flow between the analyzed *B. excelsa* adult trees.

Key words: *Bertholletia excelsa*; Genetic diversity; Amazonia

INTRODUCTION

The Brazil nut tree (*Bertholletia excelsa* Bonpl.) is native to the Amazon and presents large and irregular distribution in "terra firma" (not flooded) forests throughout the Amazon Basin (Mori and Prance, 1990). In the Brazilian Amazon, Brazil nut trees are found in Pará, Rondônia, Amazonas, and Acre States - Brazilian Amazon, as well as in Amapá, Roraima, Maranhão, and Mato Grosso States (Müller et al., 1995).

Brazil nut tree seeds have high nutritional value and play a relevant role in the Amazonian international trade because it is one of the main extractive products in regions' exportation agenda (Salomão, 2009). The harvest often takes place in natural populations and may represent up to 93% of the fruit picking from exploited trees (Zuidema and Boot René, 2002). The intense nut harvesting over decades is a threat to the species' natural regeneration. However, the ecological sustainability of Brazil nut harvesting was analyzed in several studies (Wadt et al., 2008; Scoles and Gribel, 2011; Ribeiro et al., 2014; Scoles and Gribel, 2015) according to which the current regeneration of the species is sufficient for its persistence in the study sites; nonetheless, genetic diversity conservation, besides maintaining tree population, is fundamental. Effective species conservation strategies must take the genetic, ecological, and population data into account since the genetic variability maintenance is a key to this process (Yeh et al., 1996). Knowing genetic diversity within and between populations helps to understand their history because the current genetic diversity levels can influence the success of future populations (Erickson et al., 2004). The awareness of population genetic structures and diversity levels is an essential instrument for breeding programs (Erickson et al., 2004).

There are several techniques available to detect genetic diversity at DNA sequence level. Molecular markers can be used in studies focused on the genetic diversity between individuals, and within and between populations (Souza et al., 2008). Microsatellite markers, also known as SSRs, are one of the most available polymorphic nowadays. They are codominant expressions of high replicability and present frequent and random distribution, which allows wide genome coverage (Caixeta et al., 2016).

Diversity and genetic structure studies using microsatellite markers to assess *B. excelsa* (Sujii et al., 2015) were successful but remain limited if one considers the species' great socio-economic importance. Accordingly, the aim of the present study was to assess the genetic diversity and population structure of a native *B. excelsa* population by using microsatellite markers.

MATERIAL AND METHODS

Study site and plant material

The study was carried out in Itaúba County, Mato Grosso State (MT) that is an area presenting a large extension of original Brazil nut tree formations. The experimental area is located approximately 30 km from Itaúba County, at Santo Ângelo Farm legal reserve

belonging to the Dal Pai Group. An 18.8 hectare (400 x 470 m) sub-area was installed in the site; 198 adult Brazil nut trees were sampled within it. A vascular cambium disk was extracted from each tree, approximately 1.30 m up in the stem and transferred to 2.0 mL specimen transport medium (STM; Digene Corporation) (300 µL 2% CTAB, 700 µL absolute ethanol). The material was stored at -20°C.

DNA extraction

The total genomic DNA extraction followed the CTAB method proposed by Doyle and Doyle (1987), with modifications by Machado et al. (2002). The DNA verification was performed in 1% agarose gel electrophoresis and stained with ethidium bromide at a concentration of 0.6 ng/µL. The DNA quantification was performed in small volume spectrophotometer (NanoDrop 2000 system, Thermo Scientific) and the DNA samples were diluted to a concentration of 3 ng/µL. The laboratory analyzes were performed in the Molecular Biology Laboratory of Embrapa Agrossilvipastoril, Sinop, MT, and at Embrapa Recursos Genéticos e Biotecnologia, Brasília, DF.

Genotyping

Ten microsatellite markers (primers) selected for *B. excelsa* were used in genotyping; four of them were developed by Sujii et al. (2013) and identified as "Bet"; six were developed by Reis et al. (2009) and identified as "Bex" (Table 1).

Table 1. Amplified microsatellite loci of individuals in the studied *Bertholletia excelsa* population.

Loci	Reason	Classification	Sequence primer (5'-3')
Bet12	$(TC)_7$	Simple perfect	F: ATAAGGACCGCCCATCATC R: ATAGCGAGAGCAACCTTTGAAC
Bet14	$(AG)_{15}$	Simple perfect	F: GTGTACTTCTCTGGTTGGGGC R: CCCGAGTTCATTACCCAAACT
Bet15	$(AG)_{18}$ $(AGA)_{13}$	Perfect composite	F: ACTGCCATCACCAGCATGTAG R: GTCCCTTGTGGTCTCTCACAAT
Bet16	$(CT)_{17}$ $(CCCT)_3$	Perfect composite	F: TTGATCTTCGCAAGGTCGGT R: ACTTCCTCAATCCATCGAGT
Bex02	$(CT)_8$ $(CA)_9$ $(CT)_5$ $(CA)_2$	Perfect composite	F: GCCATGTTCTCTACAGTCTC R: AGTCGGACATCCTTCGTGCT
Bex09	$(CT)_{32}$	Simple perfect	F: TATTCCATGGTCCTCCGT R: AGTCAATCATCTTCAAGAGT
Bex22	$(CT)_{38}$	Simple perfect	F: GCATTCTCTCATTTTCGCTTG R: CCCTAGCAATCGTCGTCTTC
Bex27	$(GA)_{20}$	Simple perfect	F: ACTGTTCTGATCCGCCATGT R: TTTCGACCGTTCAAATACG
Bex33	$(CT)_{37}$	Simple perfect	F: CAAGTCTCTGACTCATCGCCTA R: ACCAGGTTCAGCAGACGTTC
Bex37	$(CT)_{19}$	Simple perfect	F: TGCATGCTATGTTTCATTGCT R: CACGCAACCTCACAGTCTTG

The amplification reactions were performed in a total volume of 12 µL containing 3 µL DNA (3 ng/µL), 1.2 µL 10X PCR buffer (10 mM Tris-HCl, pH 8.3, 50 mM KCl), 1.2 µL dNTP (2.5 mM), 1.92 µL BSA (2.5 mg/mL), 0.24 µL Taq DNA Polymerase (5 U/µL), 1.5 µL of each primer (1 µM), and 1.44 µL ultrapure water. They were conducted in a thermal cycler (Applied Biosystems) under the following conditions: an initial denaturation cycle, at 94°C, for 5 min,

which was followed by 30 cycles of a 1-min denaturation, at 94°C, 1-min annealing, at 64°C, and an extension at 72°C, for 30 s, and a 7-min final extension at 72°C (Sujii et al., 2013). Primers were labeled using different fluorochromes to enable analyzing the PCR products and the duplex and multiplex systems in the ABI 3100 DNA analyzer (Applied Biosystems). The amplified fragments were detected in the GeneMapper software (v.4., Applied Biosystems).

Data analysis

The genetic diversity of the population was characterized by the mean number of alleles per locus (N_A), allelic frequency, observed heterozygosity (H_O) and expected heterozygosity (H_E) through the Hardy-Weinberg equilibrium. The Cervus 3.0.3 software was used to estimate and find the fixation index (Kalinowski et al., 2007).

The Bayesian analysis was performed in the Structure software (Pritchard et al., 2000). The number of K-groups was adjusted for 1 to 5 variations, with 20 independent interactions, using 250,000 burn-ins and 750,000 Markov chain Monte Carlo simulations, in each interaction. The genetic structure was analyzed in the Structure Harvester software (Earl and VonHoldt, 2012), according to the criterion by Evanno et al. (2005). Analysis of molecular variance (AMOVA) applied to the groups set the Structure was calculated in the GenAlEx 6.5 software (Peakall and Smouse, 2012).

RESULTS AND DISCUSSION

The sample of 10 microsatellite loci presented 38 alleles, in total, which ranged from 2 (Bet16 and Bex02) to 7 (Bex33), and 3.8 alleles per locus, on average (Table 2).

Reis et al. (2009) found 10.4 alleles per *B. excelsa* locus, on average, a number that is larger than the one identified in the present study. On the other hand, Sujii et al. (2015) studied the genetic diversity structure of five *B. excelsa* populations and found a range from 3.36 to 6.18 alleles per locus, on average, thus corroborating the herein found results.

Table 2. Estimated genetic diversity between *Bertholletia excelsa* individuals through 10 primer microsatellites.

Loci	Amplitude amplification (bp)	N	N_A	H_O	H_E	PIC	*F*
Bet12	112-118	3	189	0.646	0.582	0.498	-0.077
Bet14	105-121	5	190	0.432	0.617	0.540	0.165
Bet15	188-200	3	185	0.481	0.475	0.369	-0.008
Bet16	128-130	2	188	0.947	0.501	0.375	-0.308
Bex02	108-112	2	187	0.171	0.157	0.144	-0.037
Bex09	124-138	4	193	0.622	0.654	0.599	0.028
Bex22	128-150	3	193	0.591	0.616	0.535	0.014
Bex27	117-137	4	181	0.365	0.368	0.340	-0.015
Bex33	213-247	7	192	0.432	0.467	0.423	0.036
Bex37	185-207	5	173	0.434	0.482	0.424	0.041
Average	112-118	3.8	187.7	0.512	0.491	0.424	-0.016

N: number of individuals; N_A: number of alleles; H_O: observed heterozygosity; H_E: expected heterozygosity; PIC: polymorphism information content, *F:* fixation index.

Polymorphism information content (PIC) ranged from 0.144 (Bex02) to 0.599 (Bex09); mean PIC = 0.424 (Table 2). According to Botstein et al. (1980), PIC values below 0.25 are less informative, those between 0.25 and 0.50 are moderately informative, and the values above 0.50 are highly informative. Thus, the Bex02 locus is not informative or

recommended for further studies about the species. The loci Bet14, Bex09, and Bex22 are highly informative because they presented values higher than 0.50, and the remaining loci were moderately informative (Table 2).

The mean H_O and H_E (0.512 and 0.491, respectively) indicate population genetic diversity. Sujii et al. (2015) studied the genetic diversity of *B. excelsa* and observed higher H_O than H_E. The high heterozygosity rate presented by Brazil nut populations may be related to the breeding system. Cavalcante et al. (2012) noticed *B. excelsa* self-incompatibility and the absence of fruit generation when the plant receives pollen from its flowers. Self-incompatibility helps to maintain genetic diversity within populations since it favors cross-breeding. Other studies involving arboreal species have indicated the possible selection for the heterozygote of self-incompatible species (Sebbenn et al., 2000; Ribas and Kageyama, 2004).

Table 2 shows mean fixation index = -0.016, thus indicating the high heterozygosity and the low inbreeding of the herein studied *B. excelsa* population. According to Kageyama et al. (2003), the fixation index is one of the most important population genetic parameters, since it measures the balance between homozygotes and heterozygotes.

Based on the Bayesian analysis performed in the STRUCTURE software and on the ΔK method described by Evanno et al. (2005), we identified three groups of structures in the *B. excelsa* population, as well as overlaps between them. The ΔK value was higher than that of the remaining K in K = 3 (Figure 1A). Group I comprised 73 individuals; group II, 68; and group III, 57 (Figure 1B).

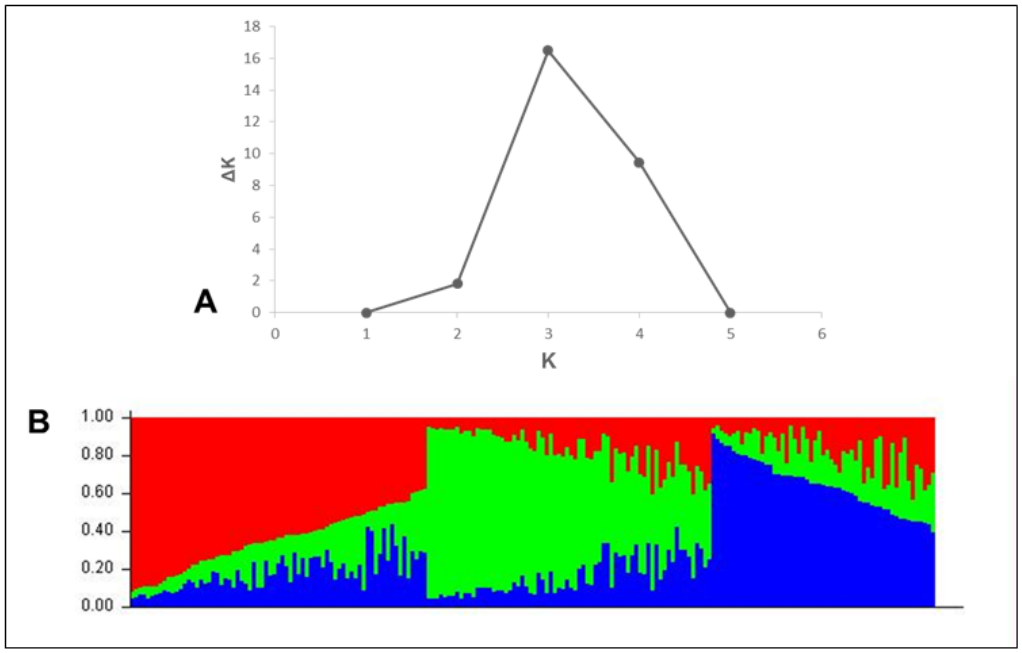

Figure 1. A. ΔK graphic representation of each K value (group) based on the 10 SSR locus information found in the Structure Harvester software. **B.** Representation of the 198 individuals from the native *Bertholletia excelsa* population. They are represented by the vertical color bars (red, green, and blue), depending on the genetic group that they belong.

AMOVA of the groups formed in the Structure software showed 98% variability within groups and 2% variability between groups (Table 3).

Such genetic structure pattern meets the studies in the literature: tropical tree species reproducing through allogamy maintain greater genetic variability proportion within populations than between populations (Zucchi et al., 2005).

Costa et al. (2015) studied the genetic structure of a *Nectandra megapotamica* population and found greater diversity within groups (75.11%) than among groups (24.89%). Soares (2016) observed the same effect when he studied the genetic diversity of a native *Hancornia speciosa* population; the greatest genetic variation occurred within groups (83.39%), rather than between groups (16.61%).

Table 3. Analysis of molecular variance (AMOVA) of native *Bertholletia excelsa* populations divided into groups, according to the molecular data from 10 SSR loci found in the "Structure" software.

Variation	d.f.	SS	CV	TV (%)	P value
Between groups	2	20.194	0.056	2	<0.001
Within groups	393	1055.826	2.687	92	
Total	395	1076.020	2.743	100	

d.f. = degrees of freedom; SS = sum of squares; CV = coefficient of variation; TV = total variation; P = variation probabilities greater than the randomly observed values. Odds were calculated for 1000 random permutations.

The F_{ST} value, a differentiation parameter responsible for identifying the allelic frequencies and the proportion of total variation of allelic frequencies occurring between subpopulations, was 0.021. It is also used to measure the effects of population subdivisions and the subpopulation heterozygosity reduction caused by genetic drift. The F_{ST} values lower than 0.05 indicate little genetic differentiation and those higher than 0.5 confirm high genetic differentiation. The F_{ST} value of the herein studied population revealed little differentiation between groups, thus corroborating AMOVA results and suggesting efficient gene flow between individuals belonging to the studied *B. excelsa* population. Zucchi et al. (2005) assumed that population variations have direct implications for conservation, and it indicates that a larger number of variations need to be sampled when the F_{ST} value is high; on the other hand if it is low, a larger number of individuals per population should be sampled. Therefore, a large number of individuals had to be kept *in situ* or sampled *ex situ* for conservation in the current study.

The studied *B. excelsa* population presented genetic diversity and mean H_O higher than H_E. The population structure was divided into three groups, and the greatest diversity was observed within groups. Therefore, the population conservation demands that a large number of individuals be maintained *in situ* or sampled *ex situ* for conservation. The absence of population inbreeding suggests that the species is not reproductively isolated and presents effective gene flow.

Conflicts of interest

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

The authors wish to thank Conselho Nacional de Desenvolvimento Científico e

Tecnológico (CNPq Process #472988/2013-1). J.C. Cabral received a Master's scholarship from FAPEMAT/CAPES.

REFERENCES

Botstein D, White RL, Skolnick M and Davis RW (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum. Genet.* 32: 314-331.

Caixeta TE, Oliveira ACB, Brito GB and Sakiyama NS (2016). Marcadores moleculares. In: Tipos de marcadores moleculares. (Borém A and Caixeta E, eds.). Viçosa, UFV, 10-93.

Cavalcante MC, Oliveira FF, Maués MM and Freitas BM (2012). Pollination requirements and the foraging behavior of potential pollinators of cultivated Brazil nut (*Bertholletia excelsa* Bonpl.) trees in central Amazon rainforest. *Psyche (Camb., Mass.)* 2012: 1-9. https://doi.org/10.1155/2012/978019

Costa LS, Reiniger LRS, Heinzmann BM, Amaral LP, et al. (2015). Study of the genetic diversity and structure of a natural population of *Nectandra megapotamica* (Spreng.) Mez. using RAPD markers. *Genet. Mol. Res.* 14: 18407-18413. https://doi.org/10.4238/2015.December.23.28

Doyle JJ and Doyle JL (1987). Isolation of plant DNA from fresh tissue. *Focus* 12: 13-15.

Earl DA and VonHoldt BM (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Resour.* 4: 359-361. https://doi.org/10.1007/s12686-011-9548-7

Erickson DL, Hamrick JL and Kochert GD (2004). Ecological determinants of genetic diversity in an expanding population of the shrub *Myrica cerifera*. *Mol. Ecol.* 13: 1655-1664. https://doi.org/10.1111/j.1365-294X.2004.02139.x

Evanno G, Regnaut S and Goudet J (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* 14: 2611-2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Kageyama PY, Sebbenn AM, Ribas LA, Gandara FB, et al. (2003). Diversidade genética em espécies arbóreas tropicais de diferentes estágios sucessionais por marcadores genéticos. *Sci. For.* 64: 93-07.

Kalinowski ST, Taper ML and Marshall TC (2007). Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Mol. Ecol.* 16: 1099-1106. https://doi.org/10.1111/j.1365-294X.2007.03089.x

Machado FRB, Vinson CC, Silva VP and Ciampi AY (2002). Extração de DNA genômico de câmbios de espécies madeireiras tropicais. In: Anais 53° Congresso Nacional de Botânica, Recife, PE.

Mori SA and Prance GT (1990). Lecythidaceae-Part II: The zygomorphic-flowered New World genera (*Couroupita, Corythophora, Bertholletia, Couratari*, *Eschweilera,* & *Lecythis*), with a study of secondary xylem of Neotropical Lecythidaceae. New York Botanical Garden, New York, 366.

Müller CH, Figueirêdo FJC, Kato AK, Carvalho JEU, et al. (1995). A cultura da castanha-do-brasil. Embrapa/SPI. Embrapa - SPI. (Coleção Plantar, 23), Brasília, 65.

Peakall R and Smouse PE (2012). GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research - an update. *Bioinformatics* 28: 2537-2539. https://doi.org/10.1093/bioinformatics/bts460

Pritchard JK, Stephens M and Donnelly P (2000). Inference of population structure using multilocus genotype data. *Genetics* 155: 945-959.

Reis AM, Braga AC, Lemes MR, Gribel R, et al. (2009). Development and characterization of microsatellite markers for the Brazil nut tree *Bertholletia excelsa* Humb. & Bonpl. (Lecythidaceae). *Mol. Ecol. Resour.* 9: 920-923. https://doi.org/10.1111/j.1755-0998.2008.02481.x

Ribas LA and Kageyama PY (2004). Estrutura genética em uma população de *Trema micrantha* (L.) B. considerando diferentes estádios de vida. *Sci. For.* 65: 176-187.

Ribeiro MBN, Jerozolimski A, Robert P and Magnusson WE (2014). Brazil nut stock and harvesting at different spatial scales in southeastern Amazonia. *For. Ecol. Manage.* 319: 67-74. https://doi.org/10.1016/j.foreco.2014.02.005

Salomão RP (2009). Densidade, estrutura e distribuição espacial da castanheira-do- brasil (*Bertholletia excelsa* H. & B.) em uma floresta tropical de platô na Amazônia Setentrional. *Bol. Mus. Para. Emílio Goeldi Ciênc. Nat.* 4: 11-25.

Scoles R and Gribel R (2011). Population structure of Brazil nut (*Bertholletia excelsa*, Lecythidaceae) stands in two areas with different occupation histories in the Brazilian Amazon. *Hum. Ecol.* 39: 455-464. https://doi.org/10.1007/s10745-011-9412-0

Scoles R and Gribel R (2015). Human Influence on the Regeneration of the Brazil Nut Tree (*Bertholletia excelsa* Bonpl., Lecythidaceae) at Capanã Grande Lake, Manicoré, Amazonas, Brazil. *Hum. Ecol.* 43: 843-854. https://doi.org/10.1007/s10745-015-9795-4

Sebbenn AM, Kageyama PY, Siqueira ACMF and Zanatto ACE (2000). Taxa de cruzamento em populações de *Cariniana legalis* (Mart.) O. Ktze.: Implicações para a conservação e o melhoramento genético. *Sci. For.* 58: 25-40.

Soares FS (2016). Diversidade genética e propagação vegetativa de mangabeira (*Hancornia speciosa* Gomes). Tese de mestrado. Universidade do Estado de Mato Grosso. Available at [portal.unemat.br/media/files/Dissertacao_Fabiano%20Silva%20Soares%20-Final.pdf].

Souza GA, Carvalho MRO, Martins ER, Guedes RNC, et al. (2008). Diversidade genética estimada com marcadores ISSR em populações brasileiras de *Zabrotes subfasciatus*. *Pesq. Agr. Bras.* 43: 843-849. https://doi.org/10.1590/S0100-204X2008000700008

Sujii PS, Inglis PW, Ciampi AY, Solferini VN, et al. (2013). Isolation and characterization of microsatellite markers for *Bertholletia excelsa* (Lecythidaceae) population genetic analysis. *Genet. Mol. Res.* 12: 5278-5282. https://doi.org/10.4238/2013.November.7.2

Sujii PS, Martins K, Wadt LHO and Azevedo VCR (2015). Genetic structure of *Bertholletia excelsa* populations from the Amazon at different spatial scales. *Conserv. Genet.* 16: 955-964. https://doi.org/10.1007/s10592-015-0714-4

Wadt LHO, Kainer KA, Staudhammer C and Serrano R (2008). Sustainable forest use in Brazilian extractive reserves: Natural regeneration of Brazil nut in exploited populations. *Biol. Conserv.* 141: 332-334. https://doi.org/10.1016/j.biocon.2007.10.007

Yeh FC, Kang SS and Chung MG (1996). Evaluation of the natural monument populations of *Camellia japonica* (Thearaceae) in Korea based on allozyme studies. *Bot. Bull. Acad. Sin.* 37: 141-146

Zucchi MI, Pinheiro JB, Chaves LJ, Coelho ASG, et al. (2005). Genetic structure and gene flow of *Eugenia dysenterica* natural populations. *Pesq. Agr. Bras.* 40: 975-980. https://doi.org/10.1590/S0100-204X2005001000005

Zuidema PA and Boot René GA (2002). Demography of the Brazil nut tree (*Bertholletia excelsa*) in the Bolivian Amazon: impact of seed extraction on recruitment and population dynamics. *J. Trop. Ecol.* 18: 1-31. https://doi.org/10.1017/S0266467402002018