Variability of *Myrciaria dubia* genotypes (Myrtaceae) in native populations of Roraima state

C.G. Souza[1], R.M.S. Paiva-Rodrigues[1], R.M. Bardales-Lozano[2], E.A. Chagas[3] and F. Granja[1]

[1] Laboratorio de Biologia Molecular, Universidade Federal de Roraima, Boa Vista, Roraima, Brasil
[2] Instituto de Investigaciones de la Amazonia Peruana (IIAP), Iquitos, Peru.
[3] Embrapa Roraima, Boa Vista, Roraima, Brasil

Corresponding author: C.G. Souza
E-mail: cledianegs@gmail.com

ABSTRACT. Camu – camu, *Myrciaria dubia* (Myrtaceae) is a native species of the Amazon Rainforest that has been attracting attention worldwide and arousing great interest in the food and pharmacological industries due to the high concentrations of ascorbic acid in its fruit, which is exported to several countries. Characterizing different materials of *M. dubia* by means of molecular markers allows integration of agronomic and molecular information to aid in the search for more promising varieties. We examined the genetic variability of 11 populations of this species distributed along the Branco River hydrographic basin in state of Roraima in northern Brazil. The populations were defined taking into account the origin of the subsample. The 55 sub-samples present in the Embrapa Roraima Germplasm Collection were evaluated using five *ISSR* initiators (UBC 811, UBC 812, UBC 817, UBC 868 and UBC 880). The five primers tested generated 64 fragments, with a 98% polymorphism rate. The greatest genetic variation was expressed within the populations (66.6%), while the lowest divergence was determined among the populations (33.4%) of the collection. There was a significant correlation between genetic and geographical distances

(Mantel test, r = 0.3%, P < 0.01). Analysis with the UPGMA method gave four subgroups showing that various individuals are genetically divergent and can be used in genetic breeding programs.

Key words: *ISSR* Markers; Camu-camu. Roraima; Genetic Variability

INTRODUCTION

Myrciaria dubia (Myrtaceae), known worldwide as camu-camu, is native to South America occurring in countries such as Peru, Guyana, Colombia, Venezuela and Brazil (Yuyama, 2011; Panduro et al., 2017). As it has the largest known fruit concentration of vitamin C (3 to 8 g in 100 g pulp) in the Amazon and contains compounds with antioxidant activity, the species is widely used in the northern region of the country and has high commercial potential (Suguino et al., 2001; Akter et al., 2011; Grigio et al., 2017). Its lyophilized pulp is mainly marketed to Europe, Japan and the United States, generating food, pharmacological and cosmetic products (Panduro et al., 2004, Akter et al., 2011).

Knowledge of genetic divergence is important step for efficient selection of species of commercial interest and in the process of domestication, such as *M. dubia*. Depending on the distribution of genetic variability, strategies can be defined for *in situ* conservation of populations with greater variation, selection of mother plants for implantation and supplementation of *ex situ* collections, as well as identification of the existence of gene flow (Martins, 1987; Negreiros et al., 2008; Nascimento et al., 2013).

In the northern Amazon, studies with morphological and agronomic markers have been performed with the aim of characterizing genetic variability to help improve the useful characteristics (Chagas et al., 2015; Lozano et al, 2016; Panduro et al., 2017). However, information generated through molecular markers is scarce for this species. Microsatellite markers such as *ISSR* (Inter Simple Sequence Repeats) can be used to determine the genetic distances allowing access to the genotype without influence of environmental conditions, as well as to verify the patterns of distribution of genetic variability within and among populations, and possible gene flow among them (Gonzales and Aguirre, 2007; Padmesh et al., 2012). To this end, we evaluated the genetic structure within and among 11 populations of *M. dubia* dispersed along the Branco river basin, with the use of *ISSR* markers.

MATERIAL AND METHODS

Study area

The hydrographic basin of the Branco river is located in the extreme north of Brazil, in the border areas with Venezuela and Guyana (Figure 1). The basin is the main drainage system of the state of Roraima, occupying about 191,000 km^2 of its territory (85%) being a sub-basin of the Negro river, which is a part of big Amazon System of rivers (Lemos et al., 2017). The Branco river basin is formed by the confluence of the rivers Tacuatu and Uraricoera, cutting almost the entire extension of Roraima in a

north-south direction. Its waters are classified as white because of the concentration of sediments in the period of heaviest rainfall incidence (IBGE, 1996).
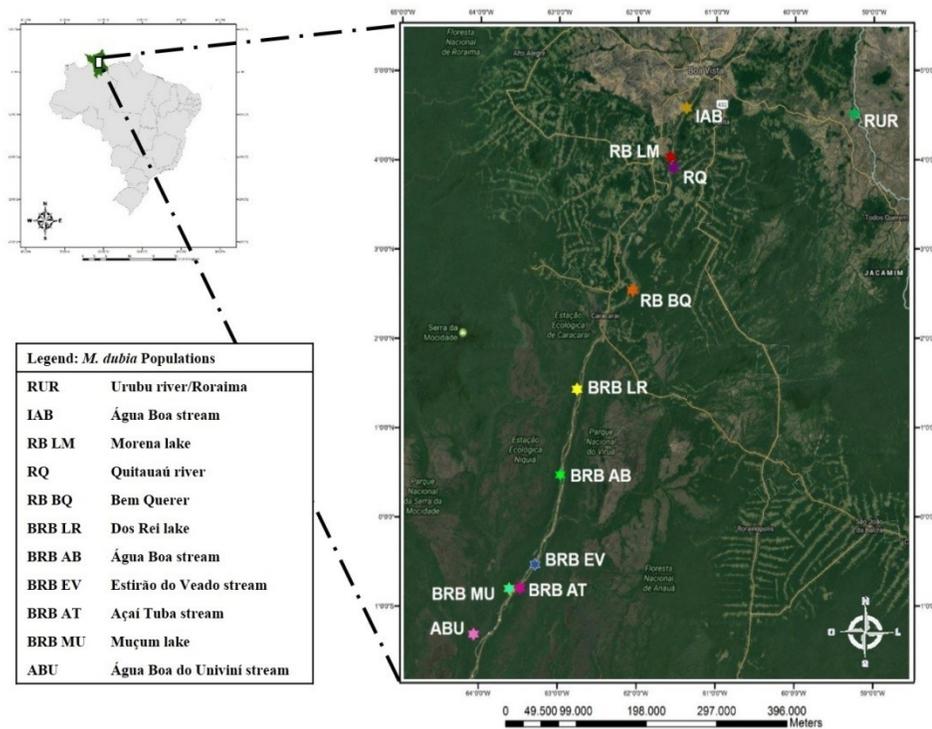

Figure 1. Geographic locations of the 11 different populations of *Myrciaria dubia* sampled from the Branco river basin, Roraima state.

Sampling

The genotypes of *M. dubia* used in this study are maintained at the Serra da Prata Experimental Field, Embrapa Roraima, located at Vicinal da Pratinha, km 07, Mucajaí municipality (02º22'36" N, 60º59'48.5" W). Plant cuttings were collected along the Branco River basin (Figure 1) and propagated by Chagas et al. (2015), establishing the Embrapa-RR Camu-Camu Germplasm Collection. These plants were transplanted in the form of cuttings directly from their natural habitat and because they are propagated vegetatively and not by seeds, they are identical to the mother plants from their original habitats. Each plant received a code with the abbreviation of the name of the river or stream of origin, identifying the populations under study (Table 1).

The maximum distance between the geographical coordinates occurs between the populations of the Urubu River (RUR) and Água Boa do Univiní (ABU), which is 309 km. From this material, genetic variability information was generated based on *ISSR* markers and correlated with the geographic distances of the populations.

Table 1. Populations of *Myrciaria dubia* collected in the state of Roraima, according to geographic coordinates.

Populations Identification	Water body of origin	Number of subsamples	Geographic coordinates	
BRB AB	Água Boa Stream / Branco River	5	1º9'52,56"N	61º20'21,90"W
ABU	Água Boa do Univini Stream	2	0º30'0"N	61º45'0"W
BRB AT	Açaí Tuba Stream	5	0º41'22,26"N	61º31'58,56"W
RB BQ	Bem Querer/ Branco River	9	1º55'16,86"N	61º0'26,10"W
BRB EV	Estirão do Veado Stream	4	0º47'10,86"N	61º27'11,46"W
IAB	Água Boa Stream / Mucajaí River	6	2º40'1,30"N	60º45'57,11"W
RB LM	Morena Lake	8	2º27'27,30"N	60º50'0,84"W
BRB LR	Dos Rei Lake	5	1º30'10,44"N	61º15'44,64"W
BRB MU	Muçum Lake	3	0º41'17,10"N	61º34'2,82"W
RQ	Quitauau River	3	2º25'44,22"N	60º45'57,11"W
RUR	Urubu River	5	2º12'28,73"N	60º2'28,62"W

Extraction of DNA from the *M. dubia*

For the extraction of the DNA, we collected young leaves from each individual. After collection, the leaves were stored in paper bags previously identified and transported in an insulated box with ice. In the laboratory, the material was stored in a freezer at -80ºC until extraction. DNA isolation was performed following the method of Doyle and Doyle (1987) modified. The quality and quantity of DNA present in the samples was analyzed by 0.8% agarose gel electrophoresis and by spectrophotometry in a CIRRUS 80 MB (FEMTO) apparatus.

Amplification of the extracted fragments

For the molecular variability analysis, we used the primers UBC 811, UBC 812, UBC 817, UBC 868 and UBC 880 for PCR. Amplification reactions were performed with a final volume of 25 μL, consisting of approximately 80 ng of genomic DNA, 2.0 μM of each initiator, 1X Reaction Buffer (pH 8.5), 1.0 mM $MgCl_2$, 1.0 mM dNTPs and 0.75 U of Taq Hot Start Polymerase (Promega). The annealing temperatures varied from 50 to 60°C.

The amplification reaction was performed in a GenCycler thermocycler (BYOSISTEMS) and its products were submitted to 1.5% agarose gel electrophoresis. Gels were visualized by excitation under ultraviolet light (302 nm) and documented by EasyDoc 200 digital photodocumentation System.

Data analysis

From the amplified bands, a binary matrix (absence = 0 / presence = 1) was constructed by observing the number and percentage of polymorphic bands per initiator, thus determining the amount of polymorphic locus. The matrix that was generated was used as the basis for analysis with the Computational Program in Genetics and Statistics GENES version 2017.3.25 (Cruz, 2006).

The minimum number of polymorphic bands required to obtain stable associations between the species was estimated through *bootstrap* analysis. The genetic similarity for each pair of individuals was carried out using resamples of different sizes with 10,000 permutations, performed with the aid of the software GENES (Cruz, 2006).

The distribution of genetic variability among and within the species populations was estimated through Molecular Variance Analysis (AMOVA) based on the method of Excoffier et al. (1992), carried out with GENES software (Cruz, 2006). The distance matrix was used for cluster analysis using a hierarchical method by the UPGMA technique. To verify the consistency of the clusters, 1,000 permutations were used. The coefficient of cophenetic correlation (r) was also calculated to verify the adequacy of the cluster pattern of the accessions. The consistency of the nodes and bifurcations was obtained using the bootstrap technique, with the aid of GENES (Cruz, 2006). To determine if there was any correlation between the geographic distances and the genetic distances estimated by the Nei and Li coefficient (1979), the Mantel test with 10,000 permutations was used in the statistical program R version 2017.3.4.0 with the vegan data package.

RESULTS

From the tags obtained on the gel a total of 64 fragments were generated using five *ISSR* primers (Table 2).

Table 2. Sequence of primers used for molecular analysis of 55 *Myrciaria dubia* subsamples. Annealing temperatures (AT), number of bands generated (NB), number of polymorphic bands (PB), percentage of polymorphic bands per initiator (% PB), Approximate amplitude (in base pairs) of the amplified fragments (Abp).

Initiator	Sequence (5' → 3')	AT (°C)	NB	PB	%PB	Abp
UBC 811	$(GA)_8C$	53.4	12	12	100	110 -900
UBC 814	$(CT)_8^a$	49.8	11	11	100	200 -900
UBC 817	$(CA)_8^a$	52.2	12	11	91.6	110 - 1600
UBC 868	$(GAA)_6$	49.8	16	16	100	250 -1250
UBC 880	$(GGAGA)_3$	49.2	13	13	100	150 - 1250
Total	-	-	64	63	98.32	250 -1300

The optimal number of markers occurred with a sample size of 51 fragments, reaching a level higher than 97% correlation. The percentage of polymorphism for all primers in the 55 subsamples of *M. dubia* analyzed was 98%. The AMOVA indicated that the greatest genetic variability is expressed among the individuals within the populations (66.6%), and a smaller divergence among the populations (33.4%) of *M. dubia* from the collection (P < 0.01).

The UPGMA cluster analysis allowed the separation of the 55 subsamples of *M. dubia* into two main groups we names G1 and G2, with a 86.4% consistency rate and four subgroups (I, II, III and IV) separated by cut point in the dendogram of 78.22% of observed average distance (Figure 2).
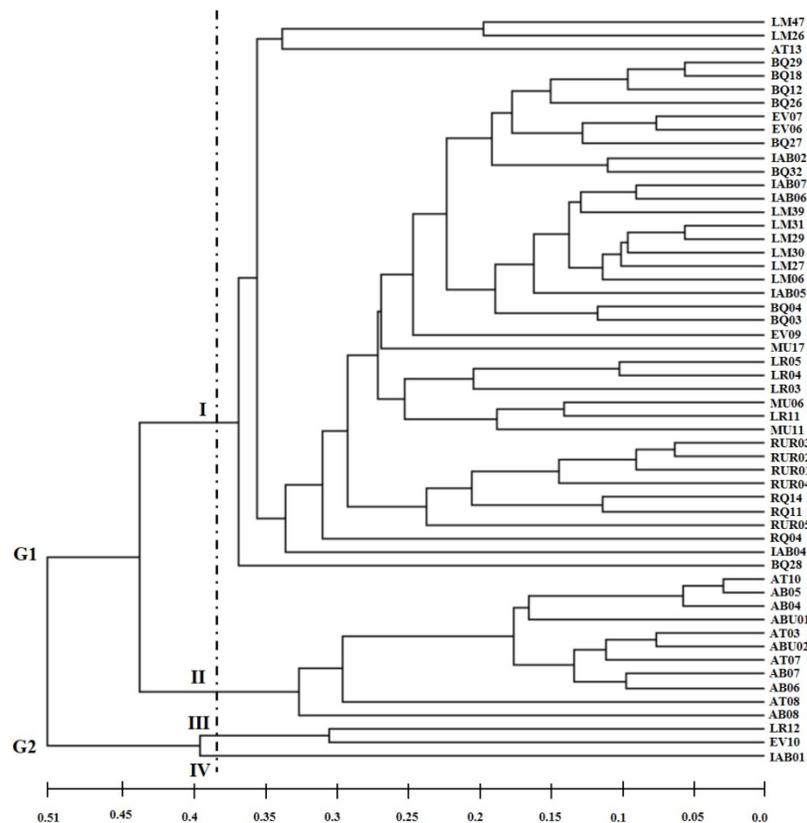
Figure 2. Grouping the 55 subsamples of *Myrciaria dubia* by the UPGMA method based on the distance matrix obtained by the complement of the similarity coefficient of Nei and Li (1979).

The subgroup I allocated 41 subsamples, with structuring of part of the accesses from the locations of Bem Querer, Morena lake, and Urubu river. The subgroup II presented all the subsamples from the locations of Água Boa stream and Água Boa do Univini river and four subsamples of the Acai Tuba stream. Subgroups III and IV were formed by only two and one subsample respectively, allocating the most divergent. Subsamples present in the same group are considered genetically closer. The coefficient of cophenetic correlation (CCC) was 86% correlation, 2.8 and 16.7 stress distortion. The correlation coefficient between geographical distances and genetic distances for the 55 subsamples of *M. dubia* was 0.3 with a significant result of $P < 0.01$.

DISCUSSION

Studies by Goulão and Oliveira (2001) describe that the use of *ISSR* markers are useful in the identification of cultivars of fruit species that usually present a long juvenile period due to high reproducibility, providing advantages over other methods. The *ISSR* markers were effective in determining the genetic variability of *M. dubia*, in the same way as in other studies applied to the family Myrtaceae, as described by

Oliveira et al. (2014) with *Psidium guajava* L. (guava), Alves et al. (2016) with *Myrcia ludiana* (Sw.) and Cruz et al. (2016) with *Plinia cauliflora* (Brazilian grape tree). This emphasizes that the use of molecular biology techniques in the evaluation of little studied species, such as *M. dubia*, allows the selection of genotypes with greater divergence for new studies of backcrossing of the species in a short period of time (Gois et al., 2014).

In a study using morphological characteristics and microsatellite markers performed in Peru (Smid et al., 2017), the authors identified a total of 91 alleles in 13 populations, concluding that the use of this tool, compared to the results of the morphological characteristics, showing a high level of genetic diversity among and within populations.

The inter- and intra-population genetic variability was similar to that found by Nunes et al. (2017) when analyzing 10 populations of *M. dubia* from Dos Reis lake, in Roraima, finding greater variability within (65%) than among populations (35%), with the use of 14 *ISSR* markers. The distribution patterns of genetic variability are highly correlated with the reproductive systems of plants (Martins, 1987). It is expected that cross-fertilized and long-lived species accumulate greater genetic variability within populations, with average values of 28% among populations when estimated by analysis of molecular variance based on dominant markers, as emphasized by Nybom and Bartish (2000). Considering that the values of variation among the populations of *M. dubia* were relatively high (33%), it is possible to infer that a partial form of self-fertilization occurs for this species, which has already been reported by Maués and Couturier (2002).

The genotypes AT08, AT13 (Açaí Tuba stream locality), AB08 (Água Boa stream locality), BQ28 (Bem Querer locality), EV10 (Estirão do Veado stream locality), LR12 (Dos Rei lake locality) and IAB01 (Água Boa stream locality) were grouped in an isolated manner, sometimes forming groups of a single individual. This characteristic is indicative of genetically divergent individuals in relation to the other genotypes of their group (Gois et al., 2014). Such genotypes, which presented greater genetic divergence in this study, are also part of the populations that have been highlighted in relation to the phenotypic characteristics analyzed by Chagas et al. (2015).

We concluded that the populations AB (Água Boa stream locality), AT (Açaí Tuba stream locality), EV (Estirão do Veado stream locality) and LR (Do Rei lake locality) contained the highest levels of ascorbic acid and that BQ populations (Bem Querer lacality) and IAB (Água Boa stream locality) were the most promising for average fruit weight. In this way, the crossing of this information indicates that the selection of combined subsamples can generate higher quality individuals, preserving or accentuating characteristics of agronomic interest, as pointed out by Rojas et al. (2011) and Nascimento et al. (2014).

The correlation of geographic and genetic distances suggests that there is a gene flow connecting the populations, which results in the genetic approximation of distinct populations and the high genetic variability within the populations (Slatkim, 1987). The grouping of subsamples from distinct regions, such as the Muçum lake (MU) that was allocated to the subgroup I, together with subsamples from alto Branco river, Quitauau

river (RQ) and Urubú river (RUR) confirm the existence of gene flow connecting the populations, through the dispersion of their seeds. As the dispersal system of the species occurs in a hydrochoric and zoocorical fashion, the populations are enriched with migratory flow of seeds carried by the ichthyofauna and the stream (Yuyama, 2011), in this way, the populations share alleles with each other maintaining genetic variability.

Gene flow is one of the evolutionary factors responsible for limiting the speciation process, restricting the effects of genetic drift and natural selection, and therefore divergence among populations through the incorporation of alleles throughout the occurrence of the species (Slatkim, 1987). The high genetic variability found in the Working Collection reflects the origin of the subsamples that were derived from natural populations under different ecological conditions. These conditions are a natural feature throughout the Branco river basin, where it presents areas of contrast represented by the contact between savannahs (lavrado - a type of savannah), forests and campinas / campinaranas (Carvalho, 2014), each population being adapted to specific environmental conditions of their place of origin.

CONCLUSIONS

The genetic variability among the populations of *M. dubia* that occur along the Branco river basin have a low correlation with the geographic distance, due to the mode of dispersion of the species to increase the gene flow among the populations. The analyses suggest that the individuals AT08, AB08, AT13, BQ28, EV10, LR12 and IAB01 are the most genetically divergent, making them promising for breeding programs because they accumulate high genetic variation.

ACKNOWLEDGMENTS

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

REFERENCES

Akter MS, Oh S, Eun JB and Ahmed M (2011). Nutritional compositions and health promoting phytochemicals of camu-camu (*Myrciaria dubia*) fruit: A review. *Food Res. Int.* 44: 1728-1732.

Alves MF, Nizio DAC, Brito FA, Sampaio TA, et al. (2016) Analysis of genetic diversity of a native population of *Myrcia lundiana* Kiaersk. Plants using *ISSR* markers. *Genet. Mol. Res.* 15: gmr15049198.

Carvalho TM (2014). Sistemas e Ambientes Denudacionais e Agradacionais, uma primeira aproximação para o Estado de Roraima, Norte da Amazônia. *Acta Geogr.* 8: 77- 98.

Chagas EA, Lozano RMB, Chagas, PC, Bacelar-Lima CG, et al. (2015). Intraspecific variability of camu-camu fruit in native populations of northern Amazonia. *Crop Breed. Appl. Biot.* 15: 265-271.

Cruz CD (2006). Programa GENES: Análise multivariada e simulação. Universidade Federal de Viçosa, Viçosa.

Cruz ES, Dantas ACVL, Carmo CD and Bastos LP (2016). Molecular characterization of jaboticaba tree genotypes located in the municipalities of Recôncavo of Bahia. *Rev. Bras. Frutic.* 38: 1-9.

Doyle JJ and Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem. Bull*. 19: 11-15.

Excoffier L, Smouse PE and Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*. 131: 479-491.

Gois IB, Ferreira RA, Silva-Mann R, Pantaleão SM, et al. (2014). Variabilidade genética em populações naturais de *Ziziphus joazeiro* Mart. por meio de marcadores moleculares RAPD. *Rev. Árvore*. 38: 621-630.

Gonzáles A and Aguirre X (2007). Inter Simple Sequence Repeats (*ISSR*s). In: Ecologia molecular (Eguiarte L, Souza V, Aguirre X, eds.). Instituto Nacional de Ecologia, México.

Goulão L and Oliveira CM (2001). Molecular characterisation of cultivars of apple (*Malus domestica* Borkh.) using microsatellite (SSR and *ISSR*) markers. *Euphytica*. 122: p.81-89.

Grigio ML, Chagas EA, Rathinasabapathi B, Chagas PC, et al. (2017). Qualitative evaluation and biocompounds present in different parts of camu-camu (*Myrciaria dubia*) fruit. *Afr. J. Food Sci.* 11: 124-129.

IBGE. Instituto Brasileiro de Geografia e Estatística. (1996). Anuário Estatístico do Brasil. Rio de Janeiro. 56: 1-11.

Lemos CEF, Vieira MIS and Ferraz RR (2017) Compartimentação hidrográfica da bacia de drenagem do Rio Branco, estado de Roraima, Brasil por meio da base hidrográfica ottocodificada. *Acta Geogr.* 11: 35-53

Lozano RMB, Chagas EA, Smirdele O, Rodriguez CA, et al. (2016) Genetic Divergence among Camu-Camu Plant Populations Based on the Initial Characteristics of the Plants. *J. Agr. Sci.* 8: 51-58.

Martins PS (1987). Estrutura populacional, fluxo gênico e conservação "in situ". *IPFE*. 35: 71-78.

Maués MME and Couturier G (2002). Biologia floral e reprodutiva do camu-camu (*Myrciaria dubia* (H.B.K.) McVaugh, Mytaceae) no Estado Pará, Brasil. *Braz. J. Bot*. 25: 441-448.

Nascimento WMO, Gurgel FL, Bhering LL, et al. (2013). Avaliações preliminares de parâmetros genéticos de acessos de *Myrciaria dubia* por marcadores fenotípicos. Embrapa Amazônia Oriental, Belém.

Nascimento WM, Gurgel FL, Bhering LL and Ribeiro OD (2014). Pré-melhoramento do camucamuzeiro: estudo de parâmetros genéticos e dissimilaridade. *Rev. Ceres*. 61: 538-543.

Negreiros JRS, Rodrigo SA, Álvares VS, Bruckner CH, et al. (2008). Divergência genética entre progênies de maracujazeiro-amarelo com base em características das plântulas. *Rev. Bras. Frutic*. 30: 197-201.

Nei M and Li WH (1979). Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc. Natl. Acad. Sci. USA*. 76: 5269-5273.

Nunes CF, Setotaw TA, Pasqual M, Chagas EA, et al. (2017). *Myrciaria dubia*, an Amazonian fruit: population structure and its implications for germplasm conservation and genetic improvement. *Genet. Mol. Res.* 16: gmr16019409.

Nybom H and Bartish I (2000). Effects of life history traits and sampling strategies on genetic diversity estimates obtained with RAPD markers in plant. *Perspect. Plant Ecol. Evol. Syst.* 3: 93-114.

Oliveira NNS, Viana AP, Quintal SSR, Paiva CL, et al. (2014). Análise de distância genética entre acessos do gênero *Psidium* via marcadores *ISSR*. *Rev. Bras. Frutic.* 36: 917-923.

Padmesh P, Mukunthakumar S, Vineesh PS, Skaria R, et al. (2012). Exploring wild genetic resources of *Musa acuminata* Colla distributed in the humid forests of southern Western Ghats of peninsular India using *ISSR* markers. *Plant Cell Rep.* 31: 1591-1601.

Panduro MP, Chagas EA, Davila EP, Rodriguez CA, et al. (2017). Selection of Superior Genotypes in 37 Clones of Camu-Camu by Repetitivity Analysis. *J. Agr. Sci.* 9: 175-187.

Panduro MP, Linares C, Mendoza H and Anguiz R (2004). Plan de mejoramiento genético de camu camu. Instituto de Investigaciones de la Amazonia Peruana, Iquitos.

Rojas S, Clement C, Yuyama K and Nagao EO (2011). Diversidade genética em acessos do banco de germoplasma de camu-camu (*Myrciaria dubia* [H.B.K.] McVaugh) do INPA usando marcadores microssatélites (EST-SSR). – *Cienc. Tecnol. Agropecuaria*. 12: 51-64.

Šmíd J, Kalousová M, Mandák B, Houška J, Chládová A, et al. (2017) Morphological and genetic diversity of camu-camu [*Myrciaria dubia* (Kunth) McVaugh] in the Peruvian Amazon. *PLOS One*. 12: e0179886.

Slatkin M (1987). Gene Flow and the Geographic Structure of Natural Populations. *Science*. 236: 787 - 792.

Suguino E, Araújo PSR and Simão S (2001). Cultivo do camu-camu (*Myrciaria dubia*). ESALQ, Piracicaba.

Yuyama K (2011). Melhoramento de camu-camu. In: Camu-camu (*Myrciaria dubia* (H.B.K.) McVaugh) (Yuyama K, Valente JP orgs.). CRV, Curitiba.

Zanatta CF and Mercadante AZ (2007). Carotenoid composition from the Brazilian tropical fruit camu–camu (*Myrciaria dubia*). *Food Chem*. 101: 1543 - 1549.