GMR
The Original by FUNPEC-RP

Genetic diversity and population structure of cassava ethno-varieties grown in six municipalities in the state of Mato Grosso, Brazil

A.V. Tiago[1], E.S.S. Hoogerheide[2], E.C.M. Pedri[1], F.S. Rossi[3], E.S. Cardoso[1], J.M.A. Pinto[2], G.F. Pena[4] and A.A.B. Rossi[4]

[1] Programa de Pós-Graduação em Biodiversidade e Biotecnologia da Rede Bionorte, Universidade do Estado de Mato Grosso Carlos Alberto Reyes Maldonado, Alta Floresta, MT, Brasil
[2] Embrapa Agrossilvipastoril, Sinop, MT, Brasil
[3] Programa de Pós-Graduação em Biodiversidade e Agroecossistemas Amazônicos, Universidade do Estado de Mato Grosso Carlos Alberto Reyes Maldonado, Alta Floresta, MT, Brasil
[4] Universidade do Estado de Mato Grosso Carlos Alberto Reyes Maldonado. Faculdade de Ciências Biológicas e Agrárias, PPGBionorte, PPGBioAgro, PGMP, Laboratório de Genética Vegetal e Biologia Molecular, Centro de Pesquisa e Tecnologia da Amazônia Meridional, Alta Floresta, MT, Brasil

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

Genet. Mol. Res. 18 (4): gmr18357
Received May 17, 2019
Accepted August 14, 2019
Published October 30, 2019
DOI http://dx.doi.org/10.4238/gmr18357

ABSTRACT. Cassava is one of the main energy foods for millions of people, and has a great diversity of ethno-varieties that have specific characteristics often not found in commercial varieties. These constitute a gene pool and therefore a genetic resource that should be conserved and preserved. In this context, the objective of our study was to evaluate the genetic diversity and population structure of ethno-varieties of cassava grown in six municipalities of the state of Mato Grosso, with the aim of characterization and conservation of the varieties found in this area. The study was carried out with 157 samples of cassava. For the molecular analyses, 15 fluorochrome-labeled SSR loci were used. Microsatellite markers amplified a total of 158 alleles. The polymorphism information content for each locus varied from 0.132 (SSRY126) to 0.838

(SSRY47), with a mean of 0.680. The expected and observed heterozygosity showed an average of between 0.717 and 0.688, for SSRY126 and SSRY47, respectively. The heterozygosity values observed were higher than those expected in five of the six populations, generating negative values of the fixation index (-0.070). Among the six populations, Alta Floresta and Cuiabá had the highest percentage of polymorphic loci (100%). The groupings obtained by UPGMA, Structure and PCoA among the six populations were concordant in allocating the individuals into two genetic groups. We found considerable genetic diversity among the samples, evidenced by the high values in the diversity indices. These high values are possibly related to the management of the fields and the exchange of propagative material among the farmers. Therefore, it is proposed that both populations be conserved since they have potential that could be used for genetic improvement of this essential crop.

Key words: Genetic variability; *Manihot esculenta*; SSR markers

INTRODUCTION

South America is the main origin of cassava (*Manihot esculenta*), and Brazil, more precisely the southern Amazon region, is the focal point of that origin. It is one of the most important cultivated species in the world (Allem, 1994; Olsene Schaal, 2001; Pereira, 2015) and is one of the main sources of energy foods for millions of people, with considerable relevance, especially in the poorest countries (Fao, 2013; Ferreira, 2014).

It is used in human food, animal feed and processing by industry (starch, starch, flour). Sweet varieties are intended for human consumption, called *mandioca de mesa*, *macaxeira*, *aipim* or *cassava mansa.* Bitter varieties, called *cassava brava* due to the abundance of cyanide in their roots, are destined for industrialization (Ponte, 2008).

In Brazil, cassava is cultivated from the north to the south of the country due to its adaptability to climatic variations and can produce reasonable yields, even in areas with poor soils and unpredictable rainfall, as well as its resistance to pests and diseases (Cardoso and Souza, 2002; Oliveira, 2014).

Brazilian cassava production is mostly sustained by family-based agriculture and is part of the local economy, with a predominance of subsistence or regional marketing (Valle and Lorenzi, 2014). However, with the modernization of agriculture, the cultivation of commercial varieties has expanded, which has tended to replace the local varieties. The national participation of family farming in cassava production reached 87% (IBGE, 2006) and, in the state of Mato Grosso family farming accounts for more than 90% of cassava production, along with fruit and dairy farming (Embrapa, 2014).

Cassava presents a diversity of species and intraspecific diversity, due to the number of varieties within each of these species (Martins, 2005), and presents specific characteristics not found in improved materials (Cleveland et al., 1994). Therefore, the construction of a gene reservoir, which should be preserved, is fundamental as it may be used by breeders in breeding programs in the formation of new varieties or in the transmission of desirable characters to existing varieties (Valle, 1991; Faraldo et al., 2000).

Traditional farming has been reduced in recent years due to the migration from rural areas to urban areas, as well as the expansion of agricultural frontiers by large farmers who now dominate areas formerly occupied by smallholders, which has led to a drastic reduction of genetic diversity (Cleveland et al., 1994). This scenario is currently observed in the state of Mato Grosso; therefore knowledge about the genetic diversity among the species and within species of the populations grown on farms is fundamental for the conservation of the all cultivated species. If the greatest proportion of diversity resides among local populations, then they must be preserved, since they contain most of the genetic diversity (Hamrick and Godt, 1996).

Genetic variability of cassava is conserved in germplasm banks or in private study registries. Despite the recognized importance of the information in these data banks, cassava germplasm has still been little studied, and there is a shortage of information, such as that related to documentation, characterization and genetic diversity (Oliveira, 2010). In this context, DNA markers are important and efficient tools for determining genetic diversity in ethno-varieties of cassava, and several studies have reported the use of molecular markers in the study of genetic diversity among cassava varieties (Costa et al., 2013; Pereira, 2015; Ortiz et al., 2016; Tiago et al., 2016; Gonçalves et al., 2017; Pedri, et al., 2019). Among these molecular markers microsatellites, SSR (simple sequence repeats) are especially important, since they have many desirable attributes, such as multi-allelic nature, codominant inheritance, high reproducibility, relative abundance and random distribution in the genome, allowing genotyping of high yielding varieties (Varshney et al., 2005; Agarwal et al., 2008; Kalia et al., 2011). This tool becomes attractive and applicable to breeding programs (Turyagyenda et al., 2012), by assisting in the conservation of data, and the establishment of priority areas for conservation.

In view of the above, this study aimed to evaluate the genetic diversity and population structure of ethno-varieties of cassava grown on family farms in six municipalities in the state of Mato Grosso.

MATERIAL AND METHODS

Collection area

The collection was carried out on the farms of family farmers in six municipalities of the state of Mato Grosso, Brazil, and each municipality was considered a study population: Alta Floresta, Apiacás, Poconé, Cáceres, Cuiabá and Jangada (Figure 1).

Alta Floresta (9°57’00.8’’ S, 56°05’44.4’’ W) e Apiacás (9°33'52.1" S 57°23'38.9" W) are located in the Meso-region of Northern Mato Grosso and in the micro-region of Alta Floresta, MT. The system of the productive centers in the region is very similar, with dairy farming and milk being the main activity (Imea, 2017), together with family farming and the extraction of wood. Recently, several areas where livestock farming was being carried out are now replaced by agricultural areas, which now form the basis of the local economy (Bonini et al., 2013). The municipalities of Poconé, Jangada and Cuiabá are located in the territory called *Baixada Cuiabana*, in the Central-South Meso-region of Mato Grosso (Mda and Sdt, 2015). This region presents traditional characteristics in agriculture, as well as in cooking and vocabulary. Some communities have strong characteristics of peasantry, where agriculture is practiced in the traditional molds, mainly for subsistence, maintaining

significant agricultural diversity, with emphasis on the local varieties of cassava (Amorozo, 2010). The municipality of Cáceres is located in the Meso-region of Central South Mato Grosso (15º27' and 17º37' S and 57º00' and 58º48' W), and is part of the Upper Pantanal micro-region. The municipality's production system is based on agriculture, mining, fishing, plant extractivism, hunting and livestock farming (Pmsb, 2014).

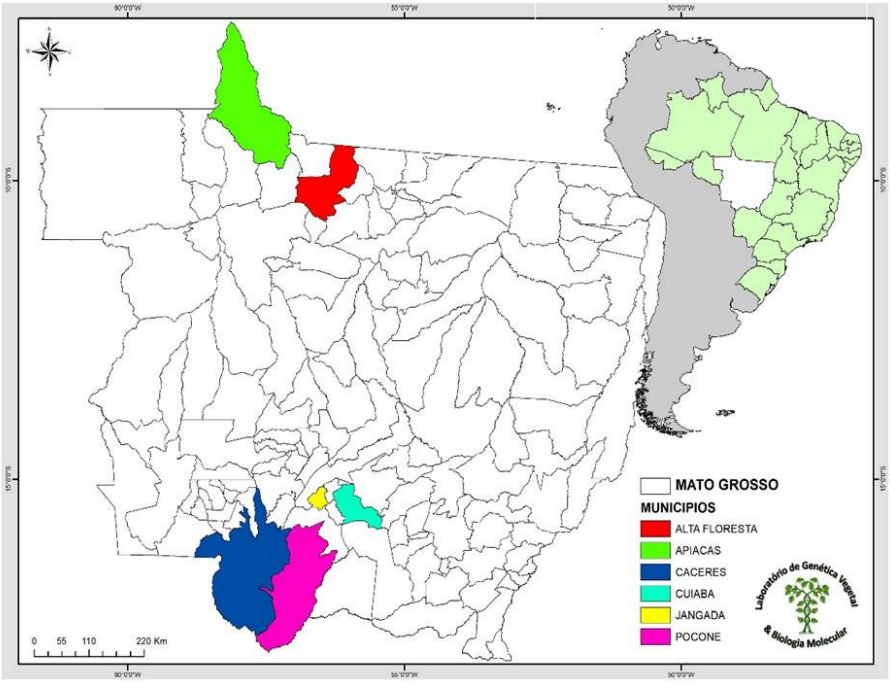

Figure 1. Location map of the municipalities of the state of Mato Grosso, Brazil, where the cassava collections were carried out.

During the expeditions, only cassava samples with different names were collected in each locality. The identification of ethno-variety was based on the knowledge of the farmers, that is, the name by which they knew the variety. In this context the term ethno-variety includes plants cultivated by farmers (local populations) in historically managed environments where biological diversity interacts with cultural diversity (Silva et al., 2001).

In order to provide accurate information on the collection points of each ethno-variety, we used GPS equipment (Global Positioning System), Garmin Etrex®.

Molecular characterization

Sampling and collection of plant material

Leaf tissue was collected from 157 cassava ethno-varieties, as follows: 29 samples from Alta Floresta; 17 from Apiacás; 26 from Cáceres; 45 from Jangada; 11 from Poconé and 29 from Cuiabá. The samples were inserted into 2.0 mL polypropylene tubes with loading buffer (containing 1 mL of saturated solution of NaCL-CTAB, 70g of NaCL, 3g of

CTAB dissolved in 200 mL of distilled water). The material was identified and taken to the Laboratory of Plant Genetics and Molecular Biology of the Southern Amazonian Technology Center (CEPTAM), Mato Grosso State University (UNEMAT), Alta Floresta campus, MT and to the Laboratory of Microbiology, Molecular Biology and Phytochemistry of Embrapa Agrossilvipastoril, Sinop-MT and stored in a freezer at -4ºC, until DNA extraction.

DNA extraction and SSR amplification

The DNA extraction of the samples was carried out at the Laboratory of Plant Genetics and Molecular Biology - UNEMAT, Alta Floresta, MT and at the Embrapa Agrossilvipastoril unit, Sinop, MT. The DNA was extracted from approximately 100 mg of leaf tissue based on the CTAB (CetylTrimethyl Ammonium Bromide) method described by Doyle and Doyle (1990), with the following modifications: STE buffer to macerate the leaves instead of liquid nitrogen, an increase in the concentration of polyvinylpyrrolidone (PVP) from 1% to 2% and from β-mercaptoethanol from 0.2% to 2% in the extraction buffer, in addition to a reduction in the incubation time at 65°C from 60 min to 30 min. The DNA concentration was estimated by spectrophotometry (NanoDrop 2000 – ThermoScientific) and its integrity verified in 1% agarose gel electrophoresis stained with *GelRed* (Biotium, Hayward, USA).

The amplification reactions were performed with microsatellite markers described by Chavarriaga-Aguirre et al. (1998) e Mba et al. (2001), the 15 (fifteen) primers used were labeled with fluorochrome 6-FAM (blue) and HEX (green) (Table 1). Polymerase chain reactions (PCR) were performed with a total reaction volume of 10µl, containing 1 µL of buffer [0.05% (w/v) bromphenol blue, 40% (w/v) sucrose, 0.1 M EDTA pH 8.0, 0.5% (w/v) sodium lauryl sulfate (SDS)]; 0.8 µL de dNTPs (2.5mM); 0.13 µL e 0.25 µL for each *primer [forward* and *reverse* (20 µM), respectively]; 0.2 µL of Taq DNA Polimerase (5 U); 0.25 µL of the HEX and FAM tagging (2 µM); 2 µl DNA, and ultra-pure Milli-Q® water to make up the total volume.

The amplifications were performed in thermocycler model T100 “Thermal Cycler” Bio-RAD, under the following conditions: denaturation at 94°C for 5 min; 30 cycles followed by denaturation at 94°C for 30 s; annealing temperature of 45°C for 45 s and 72°C for 45 s and eight cycles at 94°C for 30 s, 53°C for 45 s, 72°C for 45 s, and a final extension of 72°C for 10 min.

Amplification products were subjected to 1.5% agarose gel electrophoresis with 0.5X TAE buffer, constant voltage of 80 V for 40 min and visualized on an ultraviolet light transilluminator L-PIX Image (Loccus Biotecnologia). Then the samples, which presented bands on the agarose gel were sent to the Center for the Study of the Human Genome and Stem Cells, University of São Paulo (USP), for capillary electrophoresis genotyping in the Automatic analyzer DNA ABI 3130XL *Genetic Analyzer* (Applied Biosystems, Foster City, California, USA). The size of the amplified fragments was determined by comparison with a DNA of known size Rox 500 (APPLIED BIOSYSTEMS) by using the program Gene Marker® v. 2. 6. 3 (Softgenetics).

Table 1. Locus microsatellites used in the genotyping of 157 ethno-varieties of cassava cultivated in the state of Mato Grosso.

Locus	Fluorochrome	Motif	Classification	Amplification Range (pb)
SSRY-21**	FAM	$(GA)_{26}$	Simple perfect	172-212
SSRY-28**	HEX	$(CT)_{26}(AT)_3AC(AT)_2$	Imperfect composite	160-214
SSRY-27**	FAM	$(CA)_{14}$	Simple perfect	245-297
SSRY-35**	HEX	$(GT)_3GC(GT)_{11}(GA)_{19}$	Imperfect composite	174-310
SSRY-8**	FAM	$(CA)_{14}CT(CA)_2$	Simple imperfect	268-320
GAGG-5*	HEX	NP	---------	108-150
GA-12*	FAM	NP	----------	119-180
GA-21*	HEX	NP	----------	104-146
GA-131*	FAM	NP	----------	75-141
SSRY-43**	HEX	$(CT)_{25}$	Simple perfect	229-275
SSRY-47*	FAM	$(CA)_{17}$	Simple perfect	216-280
SSRY-126*	HEX	$(GT)_2T(GT)_5(GC)_4$	Imperfect composite	225-297
GA-136*	FAM	NP	----------	145-185
GA-140*	HEX	NP	----------	154-192
SSRY-40*	HEX	$(GA)_{16}$	Simple perfect	211-269

*Chavarriaga-Aguirre et al. (1998); **Mba et al. (2001); NP – Unpublished Motif; pb – pairs of bases.

Data analysis

Genetic diversity

The analysis of genetic diversity among ethno-variety was performed in two stages: considering the 157 individuals as belonging to a single population and later at the population level, considering the six regions sampled.

The estimation of genetic diversity, using the GDA program – Genetic Data Analysis (Lewis and Zaykin, 2001), was estimated by allele frequencies, number of alleles per locus (A), expected heterozygosity (H_e) and observed (H_o), on the Hardy-Weinberg equilibrium, assuming that the crosses occur randomly, therefore there is no endogamy, selection, migration, genetic drift or mutation (Kageyama et al. 2003), besides the index of fixation (f) and percentage of polymorphic loci (%P). The polymorphic information content (PIC) was determined with the aid of the program PowerMarker v.3.25 (Liu and Muse, 2005).

To determine the presence of rare alleles (RA) and exclusive alleles (EA), the criterion proposed by Cruz et al. (2011), where rare alleles were determined by expressing less than 0.05 in each sampled population and the number of exclusive alleles by counting the alleles present in only one of the populations. All these analyses were performed using the software GenAlEx 6.5 (Peakall and Smouse, 2012).

Population structure

The program GDA (Lewis and Zaykin, 2001) was used to estimate the genetic identity of Nei 1972, among the populations studied. With the PowerMarker program the genetic distance matrix of Nei (1972) was determined among the six populations sampled, and later imported into the program MEGA 6.5 (Kumar et al. 2004) for the construction of

the dendrogram via the UPGMA method (unweighted pair group method with arithmetic mean).

The program "Structure" 2.3.4 (Pritchard et al., 2000), based on Bayesian statistics, was used to infer the number of groups (K). The analyses were performed with 20 runs for each value of K, 200.000 "burn-ins" and 500.000 Markov chain Monte Carlo (MCMC) simulations. The K value ranged from 1 to 9. The variable K equals the number of genetically distinct populations. To define the most probable K, in relation to the proposed ones, the criteria described by Pritchard et al. (2010) and Evanno et al. (2005) were used, using the output files of Structure based on STRUCTURE HARVEST (Earl and Vonholdt, 2012) determined by the ΔK.

The organization of the genetic diversity was analyzed through the analysis of Principal Coordinates (PCoA) (Gower, 1966), which demonstrates the genetic distance between the individuals of a given population by the graphic representation and allows the identification of groups of individuals in two-dimensional or three-dimensional graphs, facilitating the visualization of the genetic structuring between the individuals and populations sampled. These results were obtained via the program GenAlEx 6.5 (Peakall and Smouse, 2012).

RESULTS AND DISCUSSION

Genetic diversity among cassava ethno-varieties

The 15 microsatellite markers amplified a total of 158 alleles in 157 cassava individuals, ranging from two (SSRY126) to 15 alleles (SSRY28 and SSRY47), with a mean of 10.5 alleles per locus (Table 2).

Table 2. Descriptive statistics of genetic diversity based on 15 microsatellite loci of 157 cassava ethno-varieties. (A = number of alleles per locus; H_e = expected heterozygosity; H_o = observed heterozygosity; f = allele fixation index; PIC= polymorphic information content).

Loci	A	H_e	H_o	f	PIC
SSRY21	12	0.820	0.845	-0.030	0.792
SSRY28	15	0.807	0.879	-0.090	0.787
SSRY27	12	0.800	0.664	0.171	0.768
SSRY35	12	0.662	0.547	0.173	0.620
SSRY8	10	0.710	0.826	-0.163	0.663
GAGG5	3	0.496	0.642	-0.295	0.400
GA12	11	0.812	0.822	-0.012	0.783
GA21	9	0.694	0.633	0.089	0.637
GA131	12	0.808	0.688	0.149	0.780
SSRY43	12	0.856	0.700	0.182	0.836
SSRY47	15	0.858	0.700	0.186	0.838
SSRY126	2	0.143	0.000	1.000	0.132
GA136	9	0.701	0.872	-0.244	0.642
GA140	12	0.757	0.729	-0.037	0.715
SSRY40	12	0.830	0.770	0.073	0.805
Mean	10.533	0.717	0.688	0.041	0.680

The expected heterozygosity (H_e) and observed (H_o) presented high values for most loci, with a mean between 0.717 and 0.688, respectively; the observed heterozygosity was higher than the expected heterozygosity in six loci of the 15 analyzed (Table 2). The obtained values indicate that the studied populations present genetic diversity, with the frequency of observed heterozygotes close to the expected heterozygotes. Carrasco (2012), working with 211 samples of cassava in three municipalities in the state of Mato Grosso, with 14 SSR loci, using the polyacrylamide technique, obtained a total of 49 alleles, with a mean of 3.79 alleles per loco, with observed and expected heterozygosity values of 0.60 and 0.59, respectively. Silva et al. (2016), analyzing 11 SSR loci, also based on the polyacrylamide technique, amplified 67 alleles in 22 cassava registers, with a mean of 6.09 alleles per loci, the expected and observed mean for heterozygosis being 0.65 and 0.61. Moura et al. (2013), working with 15 SSR loci and also using the polyacrylamide, technique found 75 alleles, with an average of five alleles per locus. The expected average heterozygosity was 0.66 and observed heterozygosity 0.61.

The results of the genetic diversity obtained in our study were higher than those found by Carrasco, Silva and Moura, but it can be observed that the values of heterozygosity were similar in the two studies, allowing for the detection of high levels of genetic diversity among the evaluated cassava ethno-varieties.

The estimate of the coefficient of fixation presented a mean value $f = 0.041$, indicating that the alleles of the study populations are not being fixed, either by the process of inbreeding or by any other event that distances the population from Hardy-Weinberg equilibrium.

The inbreeding of a given locus or population is null when the value of f is low or negative. When the value of f is very high, there is presence of inbreeding, and therefore the frequency of homozygotes is higher than expected (Templeton, 2006; Cruz et al., 2011; Maciel, 2014). Therefore, the value obtained for the heterozygosities, together with the value of f, with the exception of the locus SSRY126, indicate that among the cassava individuals analyzed, the process of inbreeding does not occur.

The polymorphic information content (PIC) for each loci presented a variation from 0.132 (SSRY126) to 0.838 (SSRY47), with a mean of 0.680 (Table 2). For Botstein et al. (1980), the molecular markers that present PIC values lower than 0.25 are considered to be less informative, those with values between 0.25 and 0.50 are classified as moderately informative and above 0.50 very informative.

In this study, 86.7% of the loci were obtained with values greater than 0.50, being therefore very informative, and indicated for a study of cassava genetic diversity.

The values for the allele frequencies are distributed in Figure 2. The frequency data ranged from 0.923 (SSRY126) to 0,003 (GA12; GA131; SSRY43; GA136; GA140). Among the 158 alleles found, 94 were considered rare, with frequency less than 5%, alleles with frequency greater than 5% were considered common, as recommended by Cruz et al. 2011. The frequencies of the rare alleles vary from 0.003 to 0.050, with a higher number of rare alleles for loci SSRY28 and SSRY47 (10) and lower for the locus GAGG5 (1). Locus SSRY126 did not present any rare alleles.

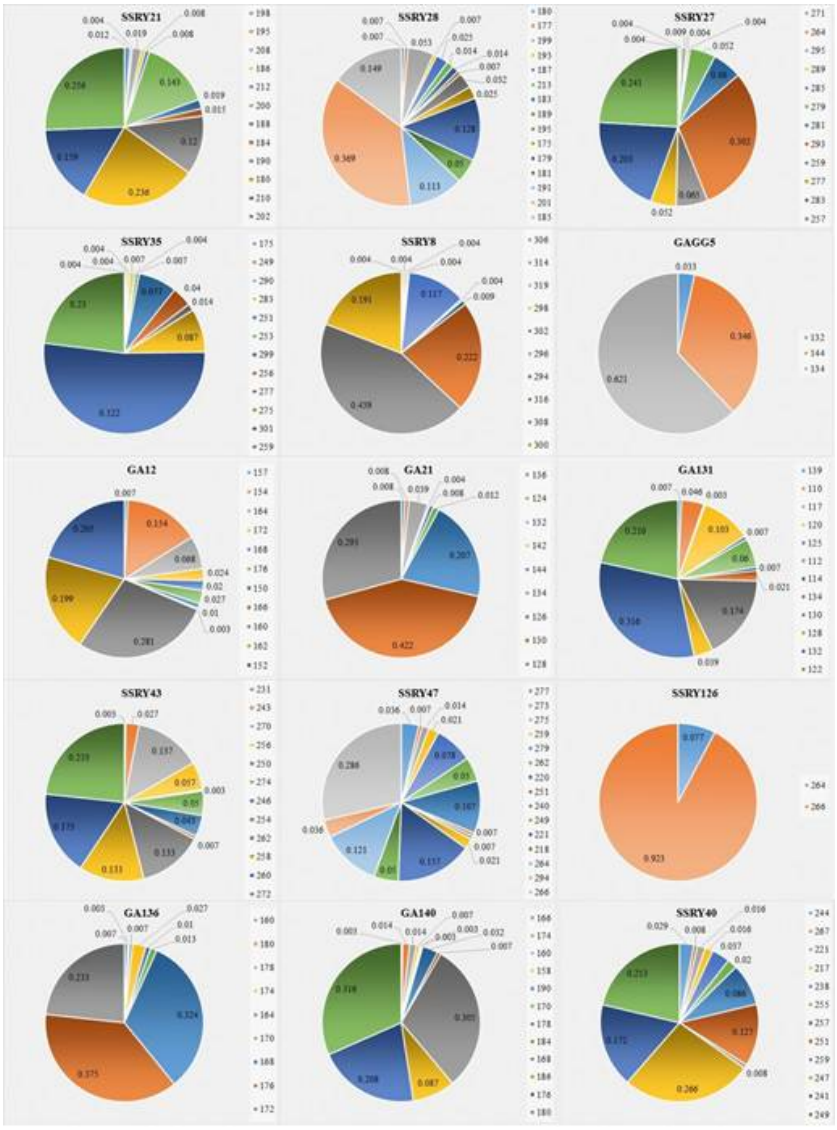

Figure 2. Distribution of allele frequencies of 15 microsatellite loci for 157 individuals of cassava.

Genetic diversity among populations of cassava ethno-varieties

Table 3 shows the number of rare and exclusive alleles distributed among the six populations. A total of 145 rare and exclusive alleles were found among these populations. The population of Jangada had the highest number of rare and exclusive alleles (37) and Poconé had the lowest number of alleles (8). In Jangada we observed 17 rare alleles, seven exclusive alleles and 13 rare-exclusive alleles.

The Jangada communities visited in this study depend basically on small-scale agriculture for subsistence, with the commercialization of only a small part of its

production. Cassava is the main item of cultivation of these communities, being mainly marketed in the form of flour (Oler, 2017). The history of the communities point to the emergence of *sesmaria* land, which is characterized by the production of food for family subsistence, not following the requirements of the market (Amaral, 2014). This may explain the high diversity of alleles found in the municipality of Jangada; that is, in these communities there is a greater number of varieties cultivated and conserved by farmers, making this population ideal as a priority area for the conservation of these ethno-varieties.

The population of Apiacás was the second population with the highest number of exclusive alleles, and rare-exclusive combined (13). This result is similar to what was found by Pedri et al. (2019), who evaluated the diversity of cassava ethno-varieties of three municipalities of Mato Grosso, including Apiacás. In their study, the authors observed that the population of Apiacás presented the highest number of rare-exclusive alleles in relation to the other municipalities, confirming, as in this study, that the Apiacás should be considered when dealing with conservation actions.

Table 3. Presence of rare and exclusive alleles in cassava based on 15 loci microsatellites in six municipalities in the state of Mato Grosso.

Population	No. of individuals	Total Alleles	RA	EA	RE	Total
Jangada	45	83	17	7	13	37
Alta Floresta	29	75	23	-	3	26
Cuiabá	29	84	16	4	3	23
Cáceres	26	77	14	1	4	19
Apiacás	17	83	19	10	3	32
Poconé	11	65	5	2	1	8
Total	157	467	94	24	27	145

Number of individuals in the population; Total alleles found in the population; RA: Rare allele; EA: exclusive allele; RE: Rare and exclusive allele; Total alleles: considering the rare, exclusive and rare and exclusive.

The number of individuals per population ranged from 45 (Jangada) to 11 (Poconé), with the number of alleles per locus between 5.6 (Cuiabá) and 4.6 (Poconé). The heterozygosity values observed were higher for five of the six populations compared to the expected ones, generating negative values for the fixation index (-0.070), however with a predominance of heterozygotes. The mean for expected and observed heterozygosity was 0.656 and 0.701, respectively (Table 4).

Table 4. Estimation of genetic diversity parameters by population. (N = number of individuals, A = average of alleles per locus, H_e = expected heterozygosity; H_o= observed heterozygosity; f = intrapopulation fixation index; %P = percentage of polymorphism).

Population	N	A	H_e	H_o	f	%P
Alta Floresta	29	5.000	0.610	0.654	-0.074	100
Apiacás	17	5.533	0.679	0.732	-0.081	93.3
Cáceres	26	5.200	0.653	0.720	-0.104	93.3
Jangada	45	5.533	0.647	0.629	-0.028	93.3
Poconé	11	4.649	0.665	0.752	-0.139	92.9
Cuiabá	29	5.600	0.680	0.716	-0.054	100
Mean	157	5.251	0.656	0.701	-0.070	95.5

Moura et al. (2016), who characterized rare cassava germplasm with 11 SSR loci, found similar results, with Ho values higher than He at all sites, and negative values for the fixation index.

The frequency of heterozygotes represents the existence of genetic variation, since individuals carry different alleles (Weir, 1996). The observed heterozygosity is an index of genetic diversity greatly influenced by the breeding system of the species. Thus, the high value of heterozygotes among ethno-varieties in all localities may be due to the process of sexual reproduction and subsequent incorporation of these materials in the plantations or by the drift effect, that is, the plantations would have been implanted from material with high heterozygosity and maintained via vegetative propagation over time (Faraldo et al., 2000).

Among the six populations analyzed, Alta Floresta and Cuiabá had the highest percentage of polymorphic loci (100%), and Poconé had a lower percentage of polymorphism (92.86%), but this was the population with the highest value for observed heterozygosity (H_o = 0.752). The average polymorphism among the populations was 95.48%, proving that the loci used had high informative power to detect genetic variability in cassava populations. Similarly, Carrasco (2012), obtained 95.00% polymorphism, and Mühlen et al. (2000) found 97.96% of polymorphic loci, both evaluating the genetic diversity of cassava registers based on microsatellite markers.

Population Structure

Table 5 presents an estimate of genetic identity Nei (1972) among pairs of populations. The most genetically identical populations are Poconé and Cáceres (0.865) and the less identical populations are Jangada and Apiacás (0.496), that is, these being the most dissimilar.

Table 5. Genetic identity of Nei (1972) based on 15 microsatellite loci among six populations of cassava: Alta Floresta (AF); Apiacás (AP); Cáceres (CA); Jangada (JA); Poconé (PC) and Cuiabá (CB).

Populations	AF	AP	CA	JA	PC	CB
AF	-	0.829	0.724	0.535	0.702	0.760
AP		-	0.827	0.496	0.780	0.750
CA			-	0.522	0.865	0.796
JA				-	0.530	0.583
PC					-	0.818
CB						-

The dendrogram generated on the basis of the UPGMA cluster among the six analyzed populations formed two groups (Figure 3), agreeing with the Bayesian grouping of the Structure (Figure 4). Group I (GI) of the dendogram consisted only of the population of Jangada, a municipality in the southern region, and was the most dissimilar group. Group II (GII) contained the other populations and revealed similarity between cassava ethno-varieties cultivated in the municipalities of the northern and southern regions of Mato Grosso state, which indicates that the exchange of genetic material made the populations more similar. In this grouping it is also possible to observe the formation of two subgroups by region. The first subgroup consists of populations from the northern region (Alta Floresta

and Apiácas) and the second subgroup of populations from the southern region (Cuiabá, Cáceres and Poconé).

The geographical position of the populations may be favoring the flow of exchange of manioc stalks, since a subgroup is found between the populations of Alta Floresta and Apiacás and another between Poconé, Cáceres and Cuiabá. The greater number of individuals sampled in the population of Jangada may have contributed to the distancing of this population in relation to the others. In addition, we can also cite the number of rare and exclusive alleles (37) (Table 3) found in the population, which indicates an isolation of the registers between the regions, that is, absence of gene flow between populations, which in the case of cassava, is summarized in the frequent exchange of stocks among farmers, whether from the same community, between municipalities or even between different states, in order to diversify the collection or to maintain it for future needs (Tiago, 2016; Oler, 2017).

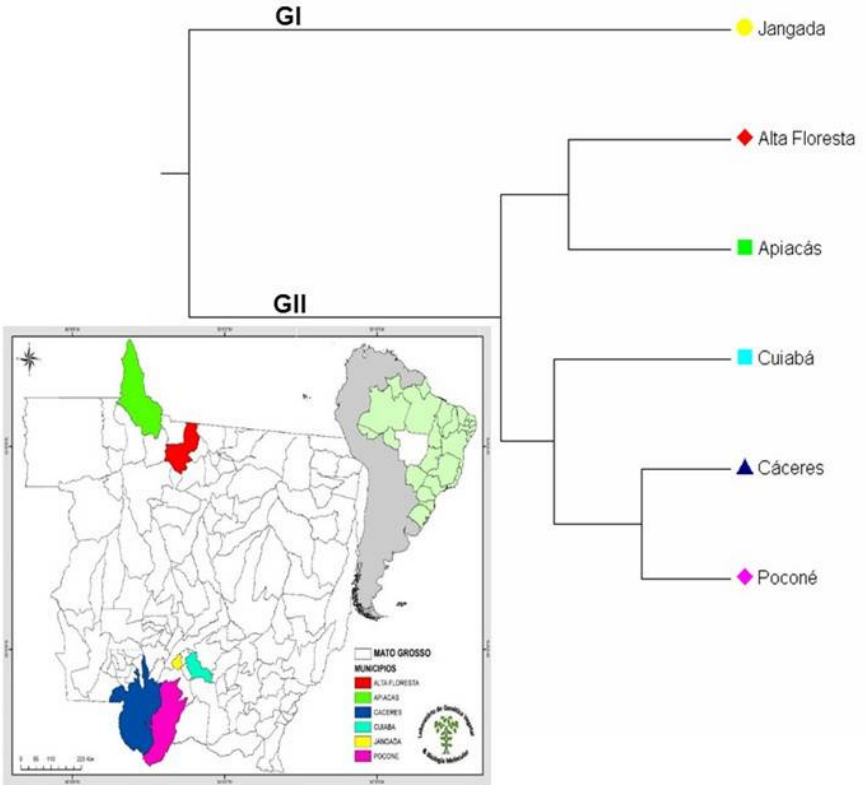

Figure 3. Dendrogram obtained by the UPGMA method from the genetic distance of Nei (1972), in six populations of cassava cultivated in the state of Mato Grosso.

The Bayesian grouping with the Structure Software did not separate the six analyzed localities, but it does show the formation of two genetic groups (K = 2) (Figure 4). It can be observed that even despite the formation of two groups, there is a mixture of alleles between the groups, demonstrated by the sharing of green colors between the red group and the red group among the green populations.

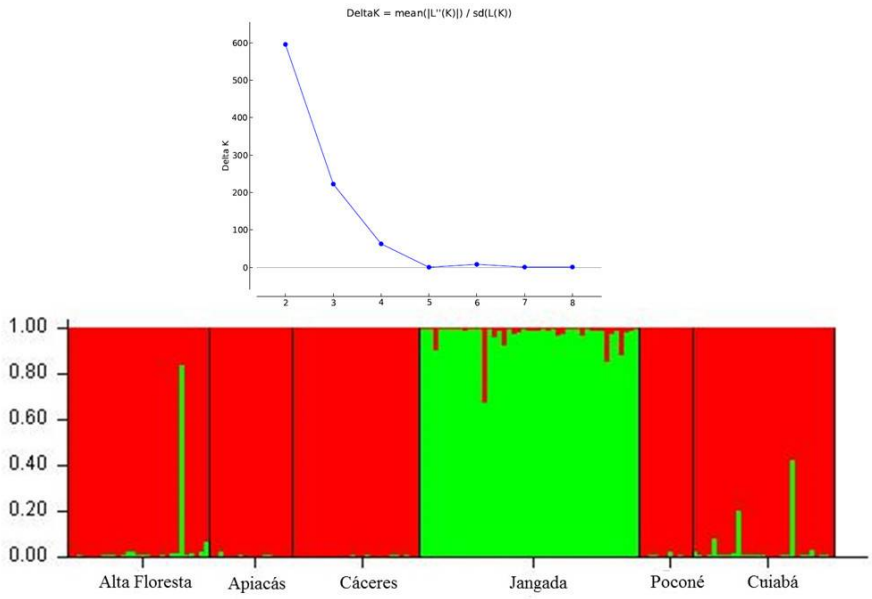

Figure 4. Analysis of the genetic structure of 157 individuals distributed in six cassava populations in different municipalities of the state of Mato Grosso from 15 microsatellite loci, assuming K = 2 (groups) according to the Structure Program. The same color for a different population indicates that they belong to the same group. Different colors in the same population indicate the percentage of alleles shared with each group.

Gonçalves et al. (2017) studying 51 cassava registers in the state of Minas Gerais, in four locations, obtained the formation of four groups according to ΔK, but also a mixture was observed among the four subpopulations, that is, registers from the same locality were allocated in different groups. Costa et al. (2013) working with 66 manioc accessions from the germplasm bank of Maringá, PR, observed the formation of two groups, which also had a mixture of genetic material between the groups formed. Ortiz et al. (2016), researching the genetic diversity and population structure of 121 registers distributed in three localities, found the formation of four groups (k = 4), that is, the traditional cultivars analyzed were divided into a larger number of groups than the number of groups collection sites. Therefore, the existence of a high number of alleles shared among samples from different localities, such as the one observed in this study, indicates a high degree of similarity between populations, allocating them in the same group, due to the selection and exchange of registers of cassava, which is a common practice among farmers (Emperaire and Peroni, 2007; Siqueira et al., 2009; Rimoldi et al., 2010).

The Principal Coordinates analysis (PCoA) contributed by giving the results found for analysis of genetic diversity and population structure. The first coordinate (PCoA1) explained 13.45% of the variation among individuals. The second coordinate (PCoA2) collaborated with 6.08% of the total variation. Together, they were able to explain 19.53% of the genetic variation (Figure 5). According to the analysis of UPGMA grouping and Structure program, PCoA also allocated the individuals into two groups, with the population of Jangada becoming more isolated.

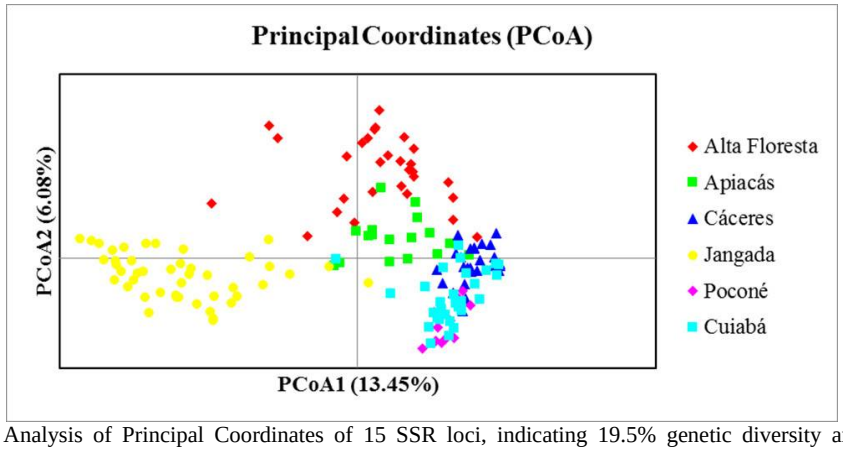

Figure 5. Analysis of Principal Coordinates of 15 SSR loci, indicating 19.5% genetic diversity among 157 cassava samples collected from six locations in the state of Mato Grosso.

The diversity of local cassava varieties, management and traditional knowledge make it possible to identify farmers as important maintainers of a significant part of the regional diversity of manioc (Marchetti, 2012). The materials studied here have the potential to be used in genetic improvement studies. In this case, the collection, preservation, maintenance and characterization of the ethno-varieties prevent genetic erosion of the crops, help to estimate the degree of kinship between accesses and ensure genetic variability for breeding programs (Cabral et al., 2001; Ribeiro et al., 2011). However, due to the recent socioeconomic transformations in the regions, the traditional agricultural activities have been negatively impacted, creating the need for studies and public policies aimed at the valorization of the management of the culture as important measures for the continuation of local agricultural practices and consequent maintenance of the local diversity of cultivated plants (Marchetti, 2012).

CONCLUSIONS

We found high genetic diversity among the cassava samples, evidenced by high values observed in the diversity indexes, with no inbreeding among the ethno-varieties. The populations are structured into two large genetic groups, caused by the isolation of the population of Jangada, due to the high values obtained in rare and exclusive alleles. The greatest diversity is distributed within each geographical region, proving the importance of the role of farmers in the flow of genetic material, since they promote both the introduction of cassava varieties and the diffusion of local varieties outside the communities. In order to preserve the high intra-population genetic diversity and the number of rare and exclusive alleles, it is proposed that all populations should be conserved, since they have the potential to be used for genetic improvement of this crop.

ACKNOWLEDGMENTS

This work was carried out with the support of the Coordination of Improvement of Higher Education Personnel - Brazil (CAPES) Financing Code 001 and the Amazon Fund/BNDES.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

REFERENCES

Agarwal M, Shrivastava N and Padh H (2008). Advances in molecular marker techniques and their applications in plant sciences. *Plant cell. rep*. 27: 617-631.

Allem AC (1994). The origin of *Manihot esculenta* Crantz (Euphorbiaceae). *Genet. resour. crop evol*. 41: 133-150.

Amaral CN (2014). Multifuncionalidades e etnoecologia dos quintais de agricultores tradicionais da baixada cuiabana: agrobiodiversidade e segurança alimentar. Doctoral thesis. Universidade Federal do Rio Grande do Sul, Porto Alegre.

Amorozo MCA and Viertler RB (2010). A abordagem qualitativa na coleta e análise de dados em etnobiologia e etnoecologia. In: Métodos e técnicas na pesquisa etnobiológica e etnoecológica (Albuquerque UP, Lucena RFP, Cunha LVFC, eds.). NUPPEA, Recife.

Bonini I, Pessoa MJG and Júnior SS (2013). Faces da produção agrícola na Amazônia mato-grossense: tipos de exploração, origem dos agricultores e impactos na conservação ambiental no município de Alta Floresta (MT). *Novos Cad. NAEA.* 16: 173-190.

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.

Cabral BLR, Souza JAD, Ando A, Veasey EA, et al. (2001). Isoenzymatic variability of cassava accessions from different regions in Brazil. *Sci. Agric.* 59: 521-527.

Cardoso CEL and Souza JS (2002). Importância, potencialidades e perspectivas do cultivo da mandioca na América Latina. In: Culturas de tuberosas amiláceas latino-americanas (Cereda MP, eds.). Fundação Cargill, Campinas.

Carrasco NF (2012). Diversidade genética de variedades tradicionais de mandioca (*Manihot esculenta* Crantz) cultivada em comunidades da Baixada Cuiabana em Mato Grosso por meio de microssatélites. Master's thesis. Escola Superior de Agricultura "Luiz de Queiroz", Piracicaba.

Chavarriaga-Aguirre PP, Maya MM, Bonierbale MW, Kresovich S, et al. (1998). Microsatellites in cassava (*Manihot esculenta* Crantz): discovery, inheritance and variability. *Theor. Appl. Genet.* 97: 493-501.

Cleveland DA, Soleri D and Smith SE (1994). Do folk crop varieties have a role in sustainable agriculture? *BioScience.* 44: 740-751.

Costa TR, Vidigal Filho PS, Gonçalves-Vidigal MC, Galván MZ, et al. (2013). Genetic diversity and population structure of sweet cassava using simple sequence repeat (SSR) molecular markers. *Afr. J. Biotechnol.* 12: 1040-1048.

Cruz CD, Ferreira MF and Pessoni LA (2011). Biometria aplicada ao estudo da diversidade genética. 1st edn. Suprema, Visconde do Rio Branco, Minas Gerais.

Doyle JJ and Doyle JL (1990). 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. *Conservation Genet. Resour*. 4: 359-361.

EMBRAPA (Empresa de Pesquisa Agropecuária Brasileira) (2014). 2ª Oficina de Concertação Estadual de Mato Grosso. Available at [https://www.embrapa.br/documents/1354377/2109296/Documento+base+CONTEXTUALIZA%C3%87%C3%83O.pdf/247bf759-27f9-4b4e-afad-1aa6cabd18d4?version=1.0]. Accessed August 05, 2019.

Emperaire L and Peroni N (2007). Traditional management of agrobiodiversity in Brazil: a case study of manioc. *Hum. ecol.* 35: 761-768.

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.

FAO (Organização das Nações Unidas para Agricultura e Alimentação) (2013). Save and Grow: Cassava. Available at [http://www.fao.org/3/a-i3278e.pdf]. Accessed July 26, 2019.

Faraldo MIF, Da Silva RM, Ando A and Martins PS (2000). Variabilidade genética de etnovariedades de mandioca em regiões geográficas do Brasil Genetic variability of landraces of cassava in geographical regions of Brazil. *Sci. Agric.* 57: 499-505.

Ferreira RCU (2014). Estrutura genética e populacional de acessos tradicionais de mandioca-de-mesa coletados em Campo Grande, Mato Grosso do Sul. Master's thesis. Universidade Estadual de Maringá, Paraná.

Genemarker SoftGenetics. Software PowerTools for Genetic Analysis. Available at [https://softgenetics.com/GeneMarker.php]. Accessed January 20, 2019.

Gonçalves TM, Vidigal Filho PS, Vidigal MCG, Ferreira RCU, et al. (2017). Genetic diversity and population structure of traditional sweet cassava accessions from Southern of Minas Gerais State, Brazil, using microsatellite markers. *Afr. J. Biotechnol.* 16: 346-358.

Gower JC (1966). Some distance properties or latent root and vector methods used in multivariate analysis. *Biometrika*. 53: 325-338.
Hamrick JL and Godt MJW (1996). Conservation genetics of endemic plant species. In: Conservation genetics, case histories from nature (Anvise JC and Hamrick JL, eds.). Chapman and Hall, New York.
IBGE (Instituto Brasileiro de Geografia e Estatística) (2006). Censo Agropecuário 2006. Available at [https://biblioteca.ibge.gov.br/visualizacao/periodicos/51/agro_2006.pdf]. Accessed August 5, 2019.
IMEA (Instituto Mato-Grossense de Economia Agropecuária) (2017). Mapa das macrorregiões do Imea. Available at [http://www.imea.com.br/upload/publicacoes/arquivos/justificativamapa.pdf]. Accessed May 27, 2019.
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: 93-107.
Kalia RK, Rai MK, Kalia S, Singh R, et al. (2011). Microsatellite markers: an overview of the recent progress in plants. *Euphytica*. 177: 309-334.
Kumar S, Tamura K and Nei M (2004). MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. *Brief. Bioinform*. 5: 150-163.
Lewis PO and Zaykin D (2001). Genetic Data Analysis: Computer program for the analysis of allelic data. Versão 1.0. Available at [http://lewis.eeb.uconn.edu/lewishome/software.html]. Accessed May 2, 2019.
Liu K and Muse SV (2005). PowerMarker: Integrated analysis environment for genetic marker data. *Bioinformatics*. 21: 21-29.
Maciel KDJS (2014). Análise da diversidade e divergência genética em clones de Eucalyptus spp. potencialmente importantes para Goiás. Master's thesis. Universidade Federal de Goiás, Goiás.
Marchetti FF (2012). Agricultura tradicional e a manutenção da agrobidiversidade em comunidades rurais do município de Santo Antonio de Leverger-MT. Master's thesis. Universidade Estadual Paulista, Rio Claro.
Martins OS (2005). Dinâmica evolutiva em roças de caboclos amazônicos. *Cad. estud. av*. 19: 209-220.
Mba REC, Stephenson P, Edwards K, Melzer S, et al. (2001). Simple sequence repeat (SSR) markers survey of the cassava (*Manihot esculenta* Crantz) genome: towards an SSR-based molecular genetic map of cassava. *Theor. Appl. Genet.* 102: 21-31.
MDA (Ministério do Desenvolvimento Agrário) and SDT (Secretaria de Desenvolvimento Territorial) (2015). Perfil territorial: Baixada Cuiabana – MT. Available at [http://sit.mda.gov.br/download/caderno/caderno_territorial_016_Baixada%20Cuiabana%20-%20MT.pdf]. Accessed June 04, 2019.
Moura EF, Farias Neto JT, Sampaio JE, Silva DT et al. (2013). Identification of duplicates of cassava accessions sampled on the North Region of Brazil using microsatellite markers. *Acta Amaz*. 43: 461-467.
Moura EF, Sousa NR, Moura MF, Dias MC, et al. 2016. Molecular characterization of accessions of a rare genetic resource: sugary cassava (*Manihot esculenta* Crantz) from Brazilian Amazon. *Genet. Resour. Crop Evol.* 63: 583-593.
Mühlen GS, Martins PS and Ando A (2000). Variabilidade genética de etnovariedades de mandioca, avaliada por marcadores de DNA Genetic diversity of cassava folk varieties assessed by DNA markers. *Sci. Agric.* 57: 319-328.
Nei M (1972). Genetic Distance between Populations. *Am. Nat*. 106: 283-292.
Oler JRL (2017). Etnobotânica e diversidade genética de mandioca (*Manihot esculenta* Crantz): a manutenção da agrobiodiversidade em comunidades tradicionais de Jangada, Mato Grosso, Brasil. Doctoral thesis. Universidade Estadual Paulista, Instituto de Biociências do Campus de Rio Claro, Rio Claro.
Oliveira AS (2014). Estudo da diversidade agrícola de raízes e tubérculos em assentamentos rurais no interior paulista. Master's thesis. Universidade Estadual Paulista, Rio Claro.
Oliveira MVC (2010). Caracterização molecular de clones de mandioca (Manihot esculenta Crantz) da Universidade Federal de Lavras. Doctoral thesis. Universidade Federal de Lavras, UFLA, Minas Gerais.
Olsen KM and Schaal BA (2001). Microsatellite variation in cassava (*Manihot esculenta*, euphorbiaceae) and its wild relatives: further evidence for a southern amazonian origin of domestication. *Am. J. Bot.* 88: 131-142.
Ortiz AHT, Rocha VPC, Moiana LD, Gonçalves-Vidigal MC, et al. (2016). Population structure and genetic diversity in sweet cassava cultivars from Paraná, Brazil. *Plant Mol. Biol. Rep*. 34: 1153-1166.
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.
Pedri ECM, Hoogerheide ESS, Tiago AV, Cardoso ES, et al. (2019). Genetic diversity of cassava landraces cultivated in northern Mato Grosso State, Brazil, using microsatellite markers. *Genet. Mol. Res.*18: gmr18315.
Pereira AA (2015). Diversidade genética, genômica e filogeografia de mandioca (*Manihot esculenta* Crantz): implicações para dispersão do cultivo ao longo dos principais eixos fluviais da bacia amazônica brasileira. Doctoral thesis. Escola Superior de Agricultura "Luiz de Queiroz", Piracicaba.
PMSB. Plano Municipal de Saneamento Básico. Prefeitura Municipal de Cáceres-MT: Diagnóstico/2014. Available at [http://www.caceres.mt.gov.br/downloads/produto2.pdf]. Accessed August 31, 2018.
Ponte CM (2008). A. Épocas de colheita de variedades de mandioca. Master's thesis. Universidade Estadual do Sudoeste da Bahia, Vitória da Conquista.

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

Pritchard JK, Wen W and Falush D (2010). Documentation for structure software: Version 2.3.4. Available at [http://burfordreiskind.com/wp-content/uploads/Structure_Manual_doc.pdf]. Accessed May 12, 2019.

Ribeiro MNO, Carvalho SP, Santos JB and Antonio RP (2011). Genetic variability among cassava accessions based on SSR markers. *Crop Breed. Appl. Biotechnol*. 11: 263-269.

Rimoldi F, Vidigal-Filho PS, Kvitschal MV, Gonçalves-Vidigal MC, et al. (2010). Genetic Divergence in Sweet Cassava Cultivars Using Morphological Agronomic Traits and RAPD Molecular Markers *Braz. Arch. Biol. Technol*. 53: 1477-1486.

Silva RM, Bandel G, Faraldo MIF and Martins OS (2001). Biologia reprodutiva de etnovariedades de mandioca. *Sci. Agric.* 58: 101-107.

Silva RS, Moura EF, Farias Neto JT, Sousa NR, et al. (2016). Genetic divergence among accessions of cassava (*Manihot esculenta* Crantz) sampled in the Tapajós region, State of Pará, using agronomic characters and microsatellite markers. *Semina: Ciênc. Agrár*. 37: 2989-3004.

Siqueira MV, Queiroz-Silva JR, Bressan EA, Borges A, et al (2009). Genetic characterization of cassava (*Manihot esculenta*) landraces in Brazil assessed with simple sequence repeats. *Genet. Mol. Biol.* 32: 104-110.

Templeton AR (2006). Population genetics and microevolutionary theory. A John Wiley & Sons Publication, USA.

Tiago AV (2016). Diversidade Genética e uso de Etnovariedades de Mandioca (*Manihot Esculenta* Crantz) Cultivadas em Propriedades Rurais no Município de Alta Floresta, Norte do Estado de Mato Grosso. Master's thesis, Universidade do Estado de Mato Grosso, UNEMAT, Alta Floresta.

Tiago AV, Rossi AAB, Tiago PV, Carpejani AA, et al. (2016). Genetic diversity in cassava landraces grown on farms in Alta Floresta-MT, Brazil. *Genet. Mol. Res.* 15: gmr15038615.

Turyagyenda LF, Kizito EB, Ferguson ME, Baguma Y, et al. (2012). Genetic diversity among farmer-preferred cassava landraces in uganda. *Afr. Crop Sci. J.* 20: 1530.

Valle TL and Geraldi IO (1991). Cruzamentos dialelicos em mandioca (*Manihot esculenta* Crantz.). Doctoral thesis. Universidade de São Paulo, Piracicaba.

Valle TL and Lorenzi JO (2014). Variedades Melhoradas de Mandioca como Instrumento de Inovação, Segurança Alimentar, Competitividade e Sustentabilidade: Contribuições do Instituto Agronômico de Campinas (IAC)[1]. *Cad. Ciên. Tecnol*. 31: 15-34.

Varshney KR, Graner A and Sorrells ME (2005). Genic microsatellite markers in plants: features and applications. *Trends biotechnol*. 23: 48-55.

Weir BS (1996). Genetic data analysis II. Sinauer Associates, USA.