Full Length Research Paper

Molecular genetic diversity in a core of cocoa (*Theobroma cacao* L.) clones with potential for selection of disease resistance, plant height and fruit production

Adeilson Batista Lins[1], Pabliane Ramos Lawinscky[1], Antônia Marlene Magalhães Barbosa[1], Fernanda Amato Gaiotto[2] and Ronan Xavier Corrêa[2]*

[1]Departamento de Ciências Agrárias e Ambientais, Universidade Estadual de Santa Cruz (UESC), Rodovia Ilhéus-Itabuna, km 16, CEP 45662-000, Ilhéus, Brazil.
[2]Departamento de Ciências Biológicas, Universidade Estadual de Santa Cruz, Rodovia Ilhéus-Itabuna, km 16, CEP 45662-000, Ilhéus, Brazil.

This study aimed to assess the genetic variability in groups of 11 clones of *Theobroma cacao* L., from different geographical regions, based on microsatellite markers, with the interest to characterize germplasm for breeding. The products of the amplification of these materials with 15 simple sequence repeat (SSR) markers were separated into ABI377 sequencer. The genotype encoded data were analyzed by means of Roger's genetic distances, which was employed in the main coordinate's analysis. The high heterozygosity observed in this group of clones (Ho = 0.7276) and genetic distances between pairs of clones (average 0.75) show that there is a high diversity among these clones. In cluster analysis, the cluster of Trinitario clones and hybrids was separated from the others according to their genealogy, while Forasteros were classified into different groups. The variability among these clones makes them important materials for parent selection in order to obtain hybrid progenies.

Key words: *Theobroma cacao,* genetic diversity, simple sequence repeat (SSR), germplasm, breeding.

INTRODUCTION

Theobroma cacao L. is relatively new in domestication (Dias, 2001) and its morphological features vary. The genus, *Theobroma* is typical of neotropical regions, with a natural distribution that covers the tropical rainforest in the western hemisphere, between latitudes 18° N and 15° S, spreading from southern Mexico to the Amazon Forest (Cuatrecasas, 1964). But this distribution may have been greatly influenced by pre-Columbian civilizations, which were responsible for the domestication of cacao and its distribution throughout Latin America over 2000 years of

cultivation. According to Venturieri (1993), there are 22 species of the genus, *Theobroma* which are restricted to tropical America. Nine of them are found in the Brazilian Amazon and are fit for cooking, and only five fit the preference for chocolate production. Among them, the species *T. cacao* and *Theobroma grandiflorum* S. stand out in the chocolate industry. The probable center of genetic diversity of cocoa, according to Cheesman (1944), covers the region of Napo, Putumayo and Caqueta, in the upper Amazon.

According to the characteristics of fruits and seeds and the geographic distribution, cocoa can be classified into three major racial groups: (i) the Forasteros, with flat, purple seeds, high hardiness and high yield potential (the Forasteros of the upper Amazon are vigorous, early maturing and resistant to certain pathogens; while the species found at the lower Amazon is known for the uniformity of its fruits, called Amelonados), (ii) the Criollo, which were originally cultivated by indigenous peoples of Mesoamerica (Mexico, Guatemala, Belize, Honduras and El Salvador) and have large seeds, rounded cotyledons and white or light violet when they are wet; (iii) and the Trinitario, natural hybrids between Forasteros and Criollo; although they are included in the group of Forasteros, they form a group that share intermediate characteristics (Dias, 2001; Bartley, 2005; Almeida and Valle, 2007).

Several studies on the genetic diversity of cacao sought to characterize and define genetic groups (Santos et al., 2005; Sounigo et al., 2005; Faleiro et al., 2004a, b; Dias, 2001; Risterucci et al., 2000; Lanaud et al., 1999; Lerceteau, 1997a, b; Russel et al., 1993). Some studies made the assumption that it is possible to differentiate the Criollo of Forasteros by isoenzymes (Ronning and Schnell, 1994) or by means of DNA markers (Laurent et al., 1993a, b). Attempts to categorize the genotypes of cocoa in the three racial groups were frustrated when using morphological descriptors (Engels, 1983) and isoenzymes (Lanaud et al., 1987) which prevented the separation into races of vegetables as initially proposed. The use of molecular markers could prove that the upper Amazon Forasteros, lower Amazon Forastero, Criollo and Trinitario have genetic differences. Studies of genetic diversity in *T. cacao* have been previously used mainly for demarcation of geographical regions (Figueira et al., 1994), discrimination of subgroups, determination of the ancestry of Criollo and modern Criollo, within the group Criollo (Motamayor et al., 2002) and delimitation of the breakdown of material Trinitario, Criollo and lower Amazon Forastero (Motamayor et al., 2003).

The international collections of germplasm are being characterized by their genetic diversity based on microsatellite markers. By means of the use of SSR markers to analyze hundreds of genotypes from commercial plantations and compare them with 246 genotypes from five collections of cocoa (germplasm banks and groups of parents of African cocoa clones), it was shown that there is great genetic differentiation among populations, except among the genebank of the West Africa and materials cultivated in countries of that region (Aikpokpodion et al., 2009). Among the 612 accessions from six groups of old cocoa genotypes, only 316 were kept for diversity analysis, since these markers allowed excluding the remaining accessions and identified as duplicates or misclassified. Such diversity analysis revealed strong structure of genetic diversity, consistent with the hydrographic division of the main rivers of the Peruvian Amazon (Zhang et al., 2009). Therefore, microsatellites have been instrumental in the management of germplasm banks and in understanding the structure of cocoa populations.

The planting of large commercial cocoa growing regions have been characterized by genetic diversity and possible involvement of different genetic groups in their composition. In Madagascar, SSR were used in diversity analysis of 27 cocoa clones and nine Trinitarians Criollo. This investigation revealed the distribution of Trinitario along a continuum between the Criollo and foreigners of the lower Amazon, except that the clones had high introgressions of alleles from Amazon. This shows that the Trinitario have as likely parent, the Amazonian individuals whose genetic diversity is low (Motamayor et al., 2002). Additionally, the hybrids with at least one parent of the high Amazon showed marked diversity. The low genetic diversity found in the family relationship between Trinitario and low Amelonados Amazon is the result of a few low Amazonian genotypes involved in the genealogy of Trinitario (Motamayor et al., 2003). Cocoa clones resistant to witches' broom grown in southern Bahia, Brazil, have high inter-genetic relationship due to the predominance of crossings of Scavina 6 with about a dozen other clones (Faleiro et al., 2001). In the analysis of 30 clones grown in this region, and four international clones of different geographic origins, it was observed that the genetic distances computed on the basis of microsatellite markers range from 0.13 to 0.71, revealing high genetic diversity among these materials (Faleiro et al., 2004a). In the analysis of 36 clones currently grown in southern Bahia, it was found that clones with potential tolerance to drought stress have genetic diversity comparable to international groups of clones such as TSH 1188 and CCN 51 (Bertolde et al., 2010). By examining 332 genotypes from farms in West Africa based on 12 SSR loci, the genetic variability and the presence of alleles associated with amelonado materials (potentially associated with high organoleptic qualities of cocoa) was shown. But this materials has been rarely used in these crop production in West Africa, highlighting the need to include local selections in germplasm banks, for conservation purposes (Aikpokpodion et al., 2009). The large-scale studies of genetic diversity of cacao has indicated more comprehensive ways to define the cacao genetic groups and eliminate access of misidentifications in the germplasm bank (Motamayor et al., 2008). In several other studies where cocoa is grown, also there

Table 1. Identification, geographical origin and type (racial groups), and height and disease resistance traits of 11 clones of *T. cacao*.

Clone	Name of clone	Origin	Historic	[a]Type	[b]Height	[c]Resist.	[d]Ref.
AMAZ 15	Amazon River	Peru-Ecuador	Wild	UA	M	VB	1
IMC 47	Iquitos Maraño (or Mixed) Calabacilo	Peru-Amazon	Wild	UA	M-H	Ce	2
SCA 24	Scavina	Peru	Wild	UA	L	VB	3
LCT EEN 37A	London Cocoa Trade Experimental Station of Napo	Ecuador	Wild	UA	M	VB	4
LCT EEN 162/1010	London Cocoa Trade Experimental Station of Napo	Ecuador	Wild	UA	M	VB	5
CCN 51	Collección Castro Naranjal	Ecuador	Cultivated	CH	M	VB	6
ICS 95	Imperial College Selection	Trinidad	Cultivated	T	L-M	Mo	7
TSH 1188	Trinidad Selection Hybrid	Trinidad	Cultivated	UH	H	VB, PP	8
COCA 3370/5	Coca River	Ecuador	Wild	UA	M-H	VB	9
GU 261	French Guyanese	French Guiana	Cultivated	LA	M-H	VB	10
H 56	Huallaga	Peru	Wild	UA	M-H	Mo	11

[a]Type: UA = Upper Amazon Forastero; CH = complex hybrid; T = Trinitario; UH = Upper Amazon Forastero Hybrids; LA = Lower Amazon Forastero. [b]Height: M = median; L = low; H = high. [c]Disease Resistance: VB = vassoura-de-bruxa; PP = podridão parda; mo monilíase; Ce = Ceratocystis. [d]References: 1 – Sounigo et al., 2005; Bekele and Bekele, 1994; ICGD, 2007; Dias, 2001; Pires, 2003. 2 – Dias, 2001; Crouzillat et al., 2000; ICGD, 2007; Pires, 2003. 3 – Bartley, 1994; Dias, 2001; Sounigo et al., 2005; ICGD, 2007; Pires, 2003. 4 – Sounigo et al., 2005; Risterucci et al., 2000; ICGD, 2007; Dias, 2001; Pires, 2003. 5 – Sounigo et al., 2005; Risterucci et al., 2000; ICGD, 2007; Dias, 2001; Pires, 2003. 6 – Neto et al., 2005; ICGD, 2007; Dias, 2001. 7 – Dias, 2001; Sounigo et al., 2005; Bartley, 1994; ICGD, 2007. 8 – Bartley, 1994; Dias, 2001; ICGD, 2007. 9 – Faleiro et al., 2004a; Bartley, 1994; ICGD, 2007; Dias, 2001. 10 – Lecerteau, 1997a; ICGD, 2007; Dias, 2001. 11 – ICGD, 2007; Dias, 2001; Pires, 2003.

are studies on the genetic diversity of local materials. These studies make it possible to better understand the genetic basis and potential use of these materials in improvement (Lima et al., 2013, Santos et al., 2015; Thondaiman et al., 2013). All these studies demonstrate that microsatellites are useful to characterize the genetic materials included in cocoa breeding programs around the word.

In the cocoa region of Bahia, Brazil, several local institutions develop research with cocoa tree, and groups of clones that best match the focus of improving their breeding programs. The germplasm characterization involves local and international clones for use in breeding programs. In this sense, this work was carried out to verify the genetic variability and clusters obtained with molecular data of 11 clones from different geographical regions, aimed to help the breeding programs of cocoa.

MATERIALS AND METHODS

The 11 clones of cocoa used in this study are from different geographical regions and are part of Brazilian breeding program (Table 1). These clones were *a priori* selected by the MCCS based on agronomic traits in order to hybridize and subsequently identify the progenies with combining ability for plant size, quality of products derived and to characterize routinely required as yield and disease resistance. Among clones considered promising, are some from the upper Amazon region that were used as parents of commercial hybrids in crosses with the lower Amazon Forastero (Amelonado), aiming at selection for combining ability for resistance to black pod (Tahi et al., 2006) and production components, such as weight per plant and seeds per fruit (Dias, 2001).

DNA samples were extracted from 300 mg of fresh leaves and healthy plants collected from each clone at intermediate age of maturity, using previously modified method (Doyle and Doyle, 1990; Bertolde et al., 2010). After extraction, all samples of DNA were purified with Prep-A-Kit Gene Mapris BIO-RAD according to the manufacturer's recommendations. Amplifications were

performed in a total volume of 13 µL of reaction solution with the concentrations previously described (Bertolde et al., 2010), in a thermal cycler programmed for 35 cycles according to the following program: 94°C for 2 min for DNA denaturation, 46°C or 51°C (temperature specific to primer) for 1 min for annealing of primer, 72°C for 1 min extension of the DNA molecule by Taq DNA polymerase. After these 35 cycles of amplification, a final step at 72°C for 7 min was performed.

The 15 microsatellite primers used for amplification are specific for *T. cacao* and were previously developed by Lanaud et al. (1999) and Risterucci et al. (2000). These primers are available at NCBI (2009) and are encoded as: AJ271827 (mTcCIR 35), AJ271942 (mTcCIR 37), AJ271945 (mTcCIR 43), AJ271946 (mTcCIR 44), AJ271953 (mTcCIR 54), AJ271956 (mTcCIR 57), Y16977 (mTcCIR 3), Y16978 (mTcCIR 2), Y16980 (mTcCIR 6), Y16982 (mTcCIR 8), Y16983 (mTcCIR 9), Y16984 (mTcCIR 10), Y16985 (mTcCIR 11), Y16986 (mTcCIR 12) and Y16987 (MTcCIR 13). Among these 15 primers, 10 were also used by other authors to characterize germplasm international collections of cocoa (Motilal et al., 2009).

Table 2. Description of microsatellite loci, observed heterozygosity (Ho) and length of the amplicons (pb) in 11 clones of *T. cacao*.

Loci	Patern*	A**	H_o	Length of amplicons (pb)
AJ271827	235	3	0.5413	222, 232, 238
AJ271942	210	5	0.7500	198, 203, 204, 205, 206
AJ271945	206	4	0.7107	248, 256, 258, 260
AJ271946	178	4	0.6016	149, 175, 183, 185
AJ271953	165	7	0.8395	229, 233, 235, 241, 243, 251, 255
AJ271956	253	7	0.8223	209, 211, 215, 217, 219, 227, 229
Y16977	249	9	0.8750	141, 147, 153, 155, 157, 159, 161, 163, 165
Y16978	254	4	0.6563	240, 250, 254, 256
Y16980	231	4	0.5826	248, 252, 254, 256
Y16982	301	5	0.6777	283, 291, 293, 295, 307
Y16983	143	7	0.8140	273, 275, 277,285, 287, 289, 295
Y16984	208	7	0.7603	198, 200, 202, 204, 206, 208, 210
Y16985	298	5	0.7521	292, 304, 306, 308, 318
Y16986	188	7	0.8182	188, 202, 204, 206, 212, 214, 210
Y16987	258	5	0.6983	142, 148, 156, 160, 164
Average	-	5.53	0.7267	-

*Number of base pairs (bp) in the fragment amplified with the primers AJ in clone Catongo (Lanaud et al., 1999) and with primers Y on IFC 5 clone (Risterucci et al., 2000). **A = number of alleles per primer generated by amplifying DNA from 11 different clones of *Theobroma cacao*.

The products of amplification were separated on ABI Prism 377 (Applied Biosystems) automated sequencer and genetically analyzed according to Bertolde et al. (2010). The genotype data generated by the programs Genescan and Genotype (Applied Biosystems) were coded into numerical matrices. The level of heterozygosity was calculated based on the ratio between the number of heterozygosis loci and the total number of loci analyzed. The matrix of modified Rogers' genetic distances (Goodman and Stuber, 1983) was used in the graphic dispersion analysis of data by principal coordinates (Gower, 1996), the program NTSYS-pcversion 2.0 (Rohlf, 1997). For each dendrogram, the value of cophenetic correlation was calculated between the genetic distance matrix and the matrix of cophenetic values.

RESULTS AND DISCUSSION

The 15 pairs of primers generated a total of 83 alleles and the group of analyzed clones showed an observed heterozygosity (Ho) of 0.7276 (Table 2). The number of alleles varied according to the primer: the most informative primer Y16977 (Ho = 0.87500) generated nine alleles and the less informative primer AJ271827 (Ho =0.5413) generated three alleles. The value of Ho in this group of clones was higher than that observed in studies with upper Amazon genotypes and clones grown in Africa, for which were reported Ho = 0.5790 (Aikpokpodion et al., 2009). The average number of alleles per primer (Na = 5.53) detected in these genotypes is lower than Na = 13.8 obtained in 35 commercial clones used in plantations in southern Bahia and resulting from crosses or from different places (Bertolde et al., 2010) and Na = 8.8 alleles detected in 60 clones sampled in different international collections of cocoa (Motilal et al., 2009). However, in the subgroup of 13 upper Amazon Forasteros, Motilal et al. (2009) found Na = 7.5, a value similar to that found in the present study. In the present study, eight out 11 clones analyzed was formed by upper Amazon clones. This lower number of alleles per locus is slightly higher to that found in other studies with natural populations (Na = 4.45) of hits from the Brazilian Amazon (Sereno et al., 2006).

Based on 15 microsatellite markers, high genetic variation among clones was detected (Table 3). The genetic distances ranged from 0.55900 (between genotypes COCA 3370 / 5 and AMAZ 15) to 0.89190 (between genotypes ICS 95 and SCA 24), averaging 0.75 for all combinations. The pairs of genotypes that had the greatest distances, around 0.8, were IMC 47 x LCTEEN 162/1010; SCA 24 x LCTEEN 162/1010; H 56 x LCTEEN 162/1010; IMC 47 x AMAZ 15; H 56 x AMAZ 15; GU 261 x AMAZ 15; SCA 24 x LCTEEN 37 A; GU 261 x LCTEEN 37 A; CCN 51 x SCA 24; H 56 x SCA 24; GU 261 x SCA 24; ICS 95 x SCA 24; H 56 x IMC 47; COCA 3370/5 x IMC 47; COCA 3370/5 X H 56; ICS 95 x H56. The other distances in the crosses had values ranging from 0.60 to 0.79. One of the smallest distances was observed between clone TSH 1188 and ICS 95 and average distances with IMC 67 24 and SCA 24. This result partially agrees with the genealogy of the clone, because TSH 1188 is the result of crosses between SCA 6 and ICS 1 IMC 67. A smaller distance between SCA 24 and IMC 47 shows tendency already expected for these two clones, since they were collected in the Amazon Forest of Peru (Pound, 1938) and are also classified as Upper Amazon or Wild.

Table 3. Genetic distances of 11 *Theobroma cacao* clones, estimated from 15 microsatellite alleles based on modified Rogers' coefficient.

Clone no.	Clone name	1	2	3	4	5	6	7	8	9	10	11
1	LCT 162/1010	0.0000										
2	AMAZ 15	0.6055	0.0000									
3	LCT 37 A	0.7853	0.6325	0.0000								
4	SCA 24	0.8466	0.7528	0.8062	0.0000							
5	IMC 47	0.8756	0.8266	0.7303	0.6583	0.0000						
6	H 56	0.8165	0.8165	0.7853	0.8466	0.8266	0.0000					
7	GU 261	0.8660	0.8292	0.8036	0.8036	0.7773	0.7906	0.0000				
8	COCA3370/5	0.6292	0.5590	0.6614	0.7500	0.8165	0.8416	0.7538	0.0000			
9	CCN 51	<u>0.7416</u>	<u>0.7416</u>	<u>0.7071</u>	<u>0.8756</u>	<u>0.7638</u>	<u>0.7638</u>	<u>0.7638</u>	<u>0.6922</u>	0.0000		
10	TSH 1188	<u>0.6831</u>	<u>0.5916</u>	<u>0.6325</u>	<u>0.7416</u>	<u>0.6831</u>	<u>0.7188</u>	<u>0.7638</u>	<u>0.6922</u>	0.6055	0.0000	
11	ICS 95	<u>0.7538</u>	<u>0.7230</u>	<u>0.7687</u>	<u>0.8919</u>	<u>0.7385</u>	<u>0.8118</u>	<u>0.7977</u>	<u>0.6892</u>	0.6571	0.6030	0.0000

Underlined numbers indicate the genetic distances between Amazon clones (1-8) and the hybrids (9-10) or Trinitario (11) clones.
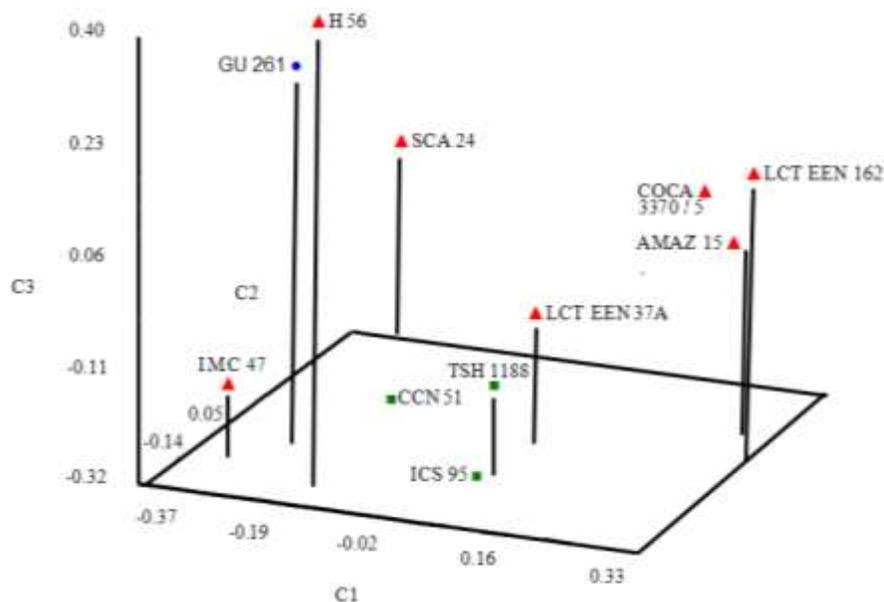
Figure 1. Principal coordinates' analysis of 11 clones of *Theobroma cacao* based on genetic distances from the data of 15 microsatellite primer pairs. (▲) Upper Amazon Forastero; (■) Hybrids and Trinitario; (●) Lower Amazon Forastero.

Distances between all Trinitario clones ranged from 0.60 to 0.66 and the distances between all Forasteros ranged from 0.56 to 0.88 (Table 3). Thus, the increased variability found between Forastero clones as compared to Trinitario is expected because the sample is formed by more Forastero than Trinitario. Additionally, the genealogy of Trinitario sampled here contributes to these differences. The following average genetic distance was observed among clones: 0.62 among clones Trinitario; 0.75 between Trinitario and Forastero; 0.74 among Forastero. Greater variability within the Forasteros as compared to Trinitarios was also observed in other genotypes based on molecular markers (Lecerteau et al.,

1997ab). It was found that the largest genetic distances are found both between clones belonging to the same racial group and between the two racial groups. Thus, the parent selection both between and within racial groups may be considered aiming for the selection of economically important traits present in this group of germplasm.

In cluster analysis, hybrids clones (TSH 1188 and CCN 51) and a Trinitario clone (ICS 95) stayed together, while the remaining clones were distributed into different plans (Figure 1). This pattern of group pattern shows that different materials do not form homogeneous groups according to the molecular diversity of racial groups,

since the apparent homogeneity of the three genotypes of Trinitario and hybrids of Trinitario can be explained by genealogy and not by racial group. The Forastero (LCTEEN 37 A) remained alone, but in an intermediate position between hybrids and a group formed by tree high Amazon clones (AMAZ 15, LCT EEN 162/1010 and COCA 3370/5). Indeed, the molecular markers alone are not a single security classification of racial groups as observed before by Lecerteau et al. (1997a) and Figueira et al. (1994). But, it is useful to show the genetic diversity of the germoplasm included in breeding.

In this study, when considering the classification of wild and domesticated clones, the formation of groups is also not clearly seen. Wild clones (AMAZ 15, LCT EEN 162/1010, LCT EEN37 A, SCA 24, IMC 47, COCA 3370/5, H 56) did not differ from the cultivated ones (TSH 1188, ICS 95, CCN 51 and GU 261). According to Figueira et al. (1994), the wild group is composed of upper Amazon Forastero, while all others are domesticated, since they suffer a lot of anthropogenic manipulation.

In this work, different clones from Ecuador were grouped with two clones from Trinidad (ICS 95 and TSH 1188) and one Peruvian clone (H 56). This result can be attributed to the structure of the populations grown in Ecuador which are offspring of parents known as "refractario", which in turn are derived from hybridization events between genotypes of three different possible regions and then selected for resistance to witches' broom (Bartley, 2001). Thus, the genetic structure of the Ecuadorian population can be attributed to (i) crosses between genotypes of Trinidad, Venezuela and Colombia, which were introduced in Ecuador in the late XIX century, (ii) National Cocoa and (iii) genotypes of Forastero from part of the Ecuadorian Amazon or Peru, at the middle Napo River (Bartley, 2001). Based on the grouping shown in this study, AMAZ 15 was allocated to the other Ecuadorians clones. Two clones from Peru (SCA 24 and IMC 47), lower Amazon clone of French Guiana (GU 261) are in the second group. According to Risterucci et al. (2000) and Sounigo et al. (2005), genotypes were classified by the acronym AMAZ which can also be classified as material of Peru or Ecuador. Sounigo et al. (2005) classified them as belonging to the Peruvian genotype AMAZ series, because the other genotypes were grouped in the same region of collection (SCA and IMC).

In conclusion, there is divergence between the materials that enable the selection of parents for hybridizations of crosses involving these genotypes. Also, there is no agreement among the clusters obtained at molecular level, geographic regions, level of domestication and racial groups for the germplasm in this work.

Conflicts of Interests

The authors have not declared any conflict of interests.

REFERENCES

Aikpokpodion PO, Motamayor JC, Adetimirin VO, Adu-Ampomah Y, Ingelbrecht I, Eskes AB, Schnell RJ, Kolesnikova-Allen M (2009). Genetic diversity assessment of sub-samples of cacao *Theobroma cacao* L collections in West Africa using simple sequence repeats marker. Tree Genet. Genomes. 5:699-711.

Almeida A-AF, Valle RR (2007). Ecophysiology of the cacao tree. Braz. J. Plant Physiol. 19:425-448.

Bartley BGD (1994) A Review of Cacao Improvement: Fundamentals, Methods and Results. Proceedings of the International workshop on Cocoa Breeding Strategies, Kuala Lumpur, Pp. 3-16.

Bartley BGD (2001) Refractario – An explanation of the meaning of the term and its relationship to the introductions from Ecuador in 1937. Ingenic Newsl. 6:10-15.

Bartley BGD (2005). The genetic diversity of cacao and its utilization. CABI, UK.

Bekele F, Bekele I (1994). Further studies on a concise list of morphological descriptors form cacao In: Proceedings of the International Workshop on New Technologies and Cocoa Breeding. Ingenic, Kuala Lumpur. Pp.155-165.

Bertolde FZ, Almeida A-AF, Corrêa RX, Loquercio LL, Gomes FP, Gaioto FA, Baligar VC (2010). Molecular physiological and morphological analysis of waterlogging tolerance in clonal genotypes of *Theobroma cacao* L. Tree Physiol. 30:56-67.

Cheesman EE (1944). Notes on the nomenclature classification and possible relationship of cocoa populations. Trop. Agric. 21:144-159.

Crouzillat D, Bellanger L, Rigoreau M, Bucheli P, Pétiard V (2000). Genetic Structure Characterization and selection of Nacional Cocoa compared to other genetic groups. In: Proceedings of the International Workshop on New Technologies and Cocoa Breeding Ingenic, Malaysia. pp. 48-66.

Cuatrecasas J (1964). Cacao and its allies, a taxonomic revision of the genus *Theobroma*. Contribuitions from the States National Herbarium. 35:379-614.

Dias LA (2001). Contribuições do melhoramento In. DIAS LA (Ed) Melhoramento genético do cacaueiro. Viçosa: Folha de Viçosa. Pp. 493-527.

Doyle JJ, Doyle JL (1990). Isolation of plant DNA from fresh tissue. Focus 12:13-15.

Engels JMM (1983). A systematic description of cacao clones III Relationships between characteristics and consequences for the cacao breeding. Euphytica 2:719-733.

Faleiro ASG, Faleiro FG, Lopes UV, Melo GRP, Monteiro WR, Yamada MM, Bahia RCS, Corrêa RX (2004a). Variability in cacao selected by producers for resistance to witches' broom based on microsatellite markers. Crop Breed. Appl. Biotechnol. 4:290-297.

Faleiro FG, Lopes UV, Yamada MM, Pires JL, Bahia RCS, Santos RC, Gomes LMC, Araújo IS, Faleiro ASG, Gramacho KP, Melo GRP, Monteiro WR, Valle RR (2001). Caracterização de variedade clonais de *Theobroma cacao* L com base em marcadores RAPD AFLP e Microssatélites. Agrostropica 13:79-86.

Faleiro GF, Pires JL, Monteiro WR, Lopes UV, Yamada MM, Piedra AG, Moura AD, Arévalo-Gardini E, Marques JRB, Gramacho KP, Faleiro ASG (2004b). Variability in cacao accessions from the Brazilian Ecuadorian and Peruvian Amazons based on molecular markers. Crop Breed. Appl. Biotechnol. 4:227-233.

Figueira A, Janick J, Levy M, Goldsbrough P (1994). Reexamining the classification of *Theobroma cacao* L using molecular markers. J. Am. Soc. Hortic. Sci. 119:1073-1082.

Goodman MM, Stuber CW (1983). Races of maize VI Isozyme variation among races of maize in Bolivia. Maydica 28:169-187.

Gower JC (1996). Some distant properties of latent root and vector methods used in multivariate analysis. Biometrika 53:325-338.

ICGD(2007). (Internation Cocoa Germoplasm Database): http://www.icgd.reading.ac.uk/.

Lanaud C, Risterucci AM, Pieretti I, Falque M, Bouet A, Lagoda PJL (1999). Isolation and characterization of microsatellites in *Theobroma cacao* L. Mol. Ecol. 8:2141-2152.

Lanaud C, Sounigo O, Amefia YK, Paulin D, Lachenaud P, Clement D (1987). New data on the mechanism of incompatibility in cocoa and its consequences on breeding. Café Cacao Thé 33:267-277.

Laurent V, Risterucci AM, Lanaud C (1993a). Varialbility for nuclear ribosomal genes within *Theobroma cacao*. Heredity 71:96-103.

Laurent V, Risterucci AM, Lanaud C (1993b). Chloroplast and mitochondrial DNA diversity in *Theobroma cacao*. Theor. Appl. Genet. 87:81-88.

Lerceteau E, Quiroz J, Soria J, Flipo S, Pétiard V, Crouzillat D (1997b). Genetic differentiation among Ecuadorian *Theobroma cacao* L accessions using DNA and morphological analyses. Euphytica 95:77-87.

Lerceteau E, Robert T, Pétiard V, Crouzillat D (1997a). Evaluation of the extent of genetic variability among *Theobroma cacao* accessions using RAPD and RFLP markers. Theor. Appl. Genet. 95:10-19.

Lima EM, Pereira NE, Pires JL, Barbosa AMM, Corrêa RX (2013). Genetic molecular diversity, production and resistance to witches' broom in cacao clones. Crop Breed. Appl. Biotechnol. 13:127-135.

Motamayor JC, Lachenaud P, e Mota JW, Loor R, Kuhn DN, Brown JS, Schnell RJ (2008). Geographic and Genetic Population Differentiation of the Amazonian Chocolate Tree (*Theobroma cacao* L). PLoS One. 3(10):e3311.

Motamayor JC, Risterucci AM, Lanaud C (2003). Cacao domestication II: progenitor germplasm of the Trinitario cacao cultivar. Heredity 91:322-330.

Motamayor JC, Risterucci AM, Lopez PA, Ortiz CF, Moreno A, Lanaud C (2002). Cacao domestication I: the origin of the cacao cultivated by the Mayas. Heredity 89:380-386.

Motilal LA, Dapeng Z, Umaharan P, Mischke S, Boccara M, Pinney S (2009). Increasing Accuracy and Throughput in Large-Scale Microsatellite Fingerprinting of Cacao Field Germplasm Collections. Trop. Plant Biol. 2:23-37.

NCBI - National Center for Biotechnology Information (2009). <https://www.ncbi.nlm.nih.gov/genbank/> Download in: 09 de November

Neto AD, Corrêa RX, Monteiro WR, Luz EDMN, Gramacho KP, Lopes U (2005). Caracterização de uma população de cacaueiro para mapeamento de genes de resistência à vassoura-de-bruxa e podridão-parda. Fitopatol. Bras. 30:380-386.

Pires JL (2003). Avaliação quantitativa e molecular de germoplasma para o melhoramento do cacaueiro com ênfase na produtividade qualidade de frutos e resistência a doenças 321p Universidade Federal de Viçosa 2003 Tesis Universidade Federal de Viçosa Viçosa.

Pound FJ (1938). Cacao and witches'broom disease (*Marasmius perniciosus*) of South America with notes on other species of *Theobroma* Yuille's Printery Port-of-Spain, 58 p.

Risterucci AM, Eskes B, Fargeas D, Motamayor JC, Lanaud C (2000). Use of microsatellite markers for germplasm identity analysis in cocoa In. Proceedings of the International Workshop on New Technologies and Cocoa Breeding. Ingenic 3 Malaysia 25-33.

Rohlf FJ (1997). NTSYS-pc numerical taxonomy and multivariate analysis system version 20. New York: Exerter Publications, P.103.

Ronning CM, Schnell RJ (1994). Allozyme diversity in a germplasm collection of *Theobroma cacao* L. J Heredity. 85:291-295.

Russel JR, Hosein F, Johson E, Waugh R, Powell W (1993). Genetic differentiation of cocoa (*Theobroma caca*o L) populations revealed by RAPD analysis. Mol. Ecol. 2:89-97.

Santos ES, Cerqueira-Silva CB, Mori GM, Ahnert D, Mello DL, Pires JL, Corrêa RX, de Souza AP (2015). Genetic Structure and Molecular Diversity of Cacao Plants Established as Local Varieties for More than Two Centuries: The Genetic History of Cacao Plantations in Bahia, Brazil. PLoS One 10(12):e0145276.

Santos RC, Pires JL, Lopes UV, Gramacho KPG, Flores AB, Bahia RC, Ramos HCC, Corrêa RX, Ahnert D (2005). Assessment of genetic diversity on a sample of cocoa accessions resistant to witches' broom disease base don RAPD and pedigree data. Bragantia 64:361-368.

Sereno ML, Albuquerque PSB, Vencovsky R, Figueira A (2006). Genetic diversity and natural population structure of cacao (*Theobroma cacao* L) from the Brazilian Amazon evaluated by microsatellite markers. Conserv. Genet. 7:13-24.

Sounigo O, Umaharan R, Christopher A, Sankar A, Ramdahin S (2005). Assessing the genetic diversity in the International Cocoa Genebank Trinidad (ICG T) using isozyme electrophoresis and RAPD. Genet. Resour. Crop Evol. 52:1111-1120.

Tahi GM, Kebe BI, N'Goran JAK, Sangare A, Mondeil F, Cilas C, Eskes AB (2006). Expected selection efficiency for resistance to cacao pod rot (*Phytophthora palmivora)* comparing leaf disc inoculations with field observations. Euphytica 149:35-44.

Thondaiman V, Rajamani K, Senthil N, Shoba N, Joel AJ (2013). Genetic diversity in cocoa (*Theobroma cacao* L.) plus trees in Tamil Nadu by simple sequence repeat (SSR) markers. Afr. J. Biotechnol. 12:4747-4753.

Venturieri GA (1993). Cupuaçu: a espécie sua cultura usos e processamento. Belém PA, 108 p.

Zhang D, Boccara M, Motilal L, Mischke S, Johnson ES, Butler DR, Bailey B, Meinhardt L (2009). Molecular characterization of an earliest cacao (*Theobroma cacao* L.) collection from Upper Amazon using microsatellite DNA markers. Tree Genet. Genomes 5:595-607.