Genetic variability and heterozygosity of cocoa accessions of Parinari (Pa) population, based on microsatellite markers

Milton Macoto Yamada*¹; Fábio G. Faleiro¹⁰; Uilson V. Lopes⁴; José L. Pires²; Alfredo Dantas Neto¹; Acassi Batista Flores¹ and Alessandra S.G. Faleiro¹
¹CEPLAC, Centro de Pesquisas do Cacau (CEPEC), Seção de Genética, Laboratório de Biotecnologia, Caixa Postal 7, CEP 45600-000, Itabuna, BA, Brazil; ² Seção de Genética, Centro de Pesquisas do Cacau (CEPEC, CEPLAC), CEP 45600-000, Itabuna, BA, Brazil; ³ Endereço atual: EMBRAPA Cerrados, C.P. 08223, CEP 73301-970, Planaltina, DF, Brazil. (* Corresponding Author. E-mail: macoto@cepec.gov.br)

ABSTRACT

The objective of this work was to study the genetic variability and the heterozygosity of accessions of the Parinari (Pa) series in the collection of the Active Germplasm Bank (BAG) of the Cocoa Research Center (CEPEC), CEPLAC, Bahia using microsatellite markers. In the present study, 34 accessions of Theobroma cacao L. were used, being 25 of the Parinari series. Twelve microsatellite loci were analyzed generating a total of 49 alleles. Results indicated genetic variability among the Parinari accessions with genetic distances varying from 0.00 to 0.64 and heterozygosity levels between 11 and 72%. About 40% of the accessions presented heterozygosity equal or greater than 50%.

KEY WORDS: Theobroma cacao L, germplasm, genetic diversity, molecular markers.

INTRODUCTION

Cocoa (Theobroma cacao L.) belongs to the Sterculiaceae family, which contains about 50 genus, the most important being Theobroma. Based on seed and fruit characteristics and the geographical distribution of cocoa, two racial groups are known (Criollo and Forastero). A third group (Trinitário), probably originated from the hybridization between Forastero and Criollo. Forasteros can be further subdivided into Upper and Lower Amazon, representing genetic materials responsible for 80% of the world cocoa production (Lerceteau et al., 1992; Marita, 1998).

The Upper Amazon accessions, particularly those of the Scavina (Sca) series, Iquitos Mixed Calabacillo (IMC), Nanay (Na), Pound (P) and Parinari (Pa) collected by Pound (1938), are important resistance sources to the main diseases of cocoa (Adomako and Adu-Ampomah, 2000; Luz et al., 1996; Umaharan et al., 2001 and Iwaro et al., 2001). Although reports exists of resistance to witches’ broom in some clones of this series (Bartley, 1983; Umaharan et al., 2001), such resistance has not been confirmed in the Active Germplasm Bank (BAG), of the Cocoa Research Center (CEPEC). However, resistance to Phytophthora sp., in fruits, has been observed in many of these clones (Iwaro et al., 2001; Luz et al., 1996; Pires et al., 1994). The accesses Pa 30 and Pa 150 were some of the few that presented resistance to the three main Phytophthora species (Luz et al., 1996). Also, under field conditions, Pa 300 presented low incidence to Phytophthora spp. (Pires et al., 1994).

Isoenzymatic and compatibility studies suggest that few trees contributed to the formation of the Parinari series, indicating little variability among accessions of this series (Yamada and Guries, 1994; Yamada et al., 1996). In these studies some alternative isoenzymatic alleles were not found and few alleles of incompatibility were observed within the population. It is suggested that seven trees contributed to the formation of the Parinari population (Lockwood and End, 1993).

Sounigo et al. (1995) evaluated the genetic diversity of 189 clones of 14 populations and 89 clones of 11 populations using isoenzymes and RAPD respectively, and showed a high degree of genetic diversity within the Scavina population and low
within P and IMC. The Parinari accessions presented low genetic diversity with isoenzymes and high diversity using RAPD (Christopher and Soungjo, 1995; Sankar and Soungjo, 1995; Soungjo et al., 1995). In the grouping analysis, P and Pa clustered together, being different from the other two groups, the Sca and the IMC accessions.

Bekele and Bekele (1996), using 28 quantitative and 26 qualitative morphologic descriptors and 100 accessions of 24 populations, showed that among the 11 groups formed, Pa accessions were represented in 3 different groups. Charters and Wilkinson (2000), using ISSR markers, clustered the 62 cocoa accessions studied in 14 groups, with the three accessions of Parinari series in the same group. In a more extensive study, Marita (1998) evaluating 270 accessions using RAPD demonstrated that the seven Parinari accessions clustered closely, except Pa 175.

The objective of this work was to study the genetic variability and heterozygosity of Parinari accessions from the germplasm collection at the CEPEC/CEPLAC based on microsatellite markers.

MATERIAL AND METHODS

Genetic material

In the present study 34 accessions of *Theobroma cacao* L were used, being 25 of the Parinari (Pa) series with two of them repeated in different rows of the collection, and with 7 suspects to be of the same geographical origin (Table 1).

DNA extraction

Leaves of each accession of *T. cacao* were collected for DNA extraction. The genomic DNA of each accession was extracted using the CTAB method (Doyle and Doyle, 1990), with some modifications (Araújo et al., 2000). After the extraction, the DNA concentration was estimated by spectrophotometer at 260 nm (Sambrook et al., 1989). Bands of total genomic DNA separated by electrophoresis in agarose gel of 0.8% were used as indicative of the integrity and purity of the extracted DNA. After quantification, the DNA samples of good quality were diluted to 10ng/mL.

Microsatellite markers

The amplification reactions for microsatellite were prepared in a total volume of 15 ml, including Tris-HCl 10 mM (pH 8.3), KCl 50 mM, MgCl2 2.4 mM, 150 mM for each one of the deoxinucleotides (dATP, dTTP, dGTP and dCTP), 3 pM for each one of the two primers (F and R), a unit of the enzyme Taq polymerase and approximately 30 ng of DNA. The amplifications were made in thermocycler, according to the following program: 4 minutes at 94 ºC + 10 cycles (30 seconds at 94 ºC + 60 seconds at 60 ºC -1 ºC to each cycle + 90 seconds at 72 ºC) + 30 cycles (30 seconds at 94 ºC + 60 seconds at 48 ºC + 90 seconds at 72 ºC) + 6 minutes for 72 ºC. After amplification, the temperature of the samples was reduced to 4 ºC. Twelve microsatellite pairs of primers were used (Table 2).

For each sample, 3 ml of a mixture of bromophenol blue (0.25%), glycerol (60%) and water (39.75%), were added. These samples were loaded in agarose gel 3% for separation of microsatellite fragments. The gel was submerged in TBE buffer (Tris-borate 90 mM, EDTA 1 mM) and electrophoretic separation was conducted for approximately four hours, at 100 volts. At the end of the run, the gels were stained with ethidium bromide and photographed under ultraviolet light.

Statistical analyses

The microsatellite markers generated were transformed to a coded numeric matrix, from which genetic distances were calculated between the genetic materials and cluster analyses done. The code of the molecular markers were 0 for absence of the allele, 1 for presence of one copy of the allele and 2 for presence of two copies of the allele.

The genetic distances were calculated according to the following formula:

$$DG_{ij} = 1 - \left( {\frac{\sum PLC_{ij}}{NTL_{ij}}} \right)$$

being:

$DG_{ij} =$ genetic distance between the accessions i and j;

$PLC_{ij} =$ proportions of coincidents alleles in each loci, being value 1 for 2 coincidents allele (in homozygous or heterozygous); 0.5 for one coincident allele and zero for nil coincident allele.

$NTL_{ij} =$ total number of loci without missing data to I or j.

Defined distance matrix, the multidimensional space
was reduced to bidimensional by the 'multidimensional scaling' method - MDS SAS and cluster analyses were done using the UPGMA method considering two groups (SAS, 1990).

The heterozygosity level was estimated as the ratio between the number of heterozygous locus and the total number of analyzed loci.

RESULTS AND DISCUSSION

The twelve microsatellite loci analyzed generated a total of 49 alleles, with an average of 4 alleles per locus (Table 2) and (Figure 1). Considering only Parinari accessions, this number drops to 44 alleles, with an average of 3.7 alleles per locus. Faleiro et al. (2001a) analyzing microsatellite loci in Upper Amazon accessions obtained an average of 7.5 alleles per locus. Risterucci et al. (2000), also analyzing microsatellites in cocoa, obtained an average of 5.6 alleles per locus. The low number of alleles per locus in our work is an indication of a certain genetic relationship among the accessions of the Parinari series.

The grouping analysis showed a certain relationship among the accessions of the Parinari series (Figure 2). The clones from 1 to 6, suspected of being Parinari, formed one group and the remaining known as being of the Parinari series formed another group.

The information about the origin of the Parinari accessions is scarce, but they seem to be originated from few trees, as indicated by Lockwood and End (1993). Most of the Parinari accessions should be genetically related as half-sibs or full-sibs, evidenced by the low number of incompatibility alleles (Yamada et al., 1996).

In isoenzyme studies, the alternative forms of MDH and IDH alleles (Yamada and Guries, 1994) were not observed, suggesting that the Parinari populations collected were not in Hardy-Weinberg equilibrium. In cocoa, pollination by insects is common, resulting in a predominance of full-sib families (Yamada and Guries, 1998). The absence of alternative forms of isoenzymes alleles and the presence of few compatibility alleles suggest that a small number of parents participated in the formation of that population. Also, the low number of microsatellite alleles per locus within this population is an indication of a low number of parents. Marita (1998) also found Pa clones very closely related in cluster analysis of 270 clones. Similar results were found in the case of Pa 56, Pa 88 and Pa 184, in studies with other cocoa accessions (Charters and Wilkinson, 2000).

The genetic materials 1 and 2 (accession 192) were collected in different rows of the germplasm collection, but they seem to be the same material (Figure 2 and genetic similarity 1.00, data not shown). CEPEC accessions (526, 527, 528 and 529) were clustered and are quite related among themselves. Those four CEPEC accessions were grouped and seem to be different from TSA 644 whose row in BAG/CEPEC was initially collected. These four accessions are different and do not seem to be related to Parinari accessions, as initially believed. The

Table 1. Average heterozygosity of cocoa accessions based on 12 microsatellites loci.

| Order | Accessions | Heter (%) | Line | Order | Accessions | Heter (%) | Line |
|-------|------------|-----------|------|-------|------------|-----------|------|
| 1     | 192?       | 42        | 120-P9 | 18    | Pa 195     | 50        | 123-P10 |
| 2     | 192?       | 42        | 442-P5 | 19    | Pa 285     | 25        | 445-P3  |
| 3     | CEPEC 526  | 58        | 774-P7 | 20    | Pa 294     | 36        | 777-P10 |
| 4     | CEPEC 527  | 50        | 745-P2 | 21    | Pa 30      | 45        | 441-P6  |
| 5     | CEPEC 528  | 50        | 746-P6 | 22    | Pa 300     | 42        | 736-P6  |
| 6     | CEPEC 529  | 58        | 747-P6 | 23    | Pa 303     | 45        | 737-P7  |
| 7     | TSA 644    | 58        | 230-P11| 24    | Pa 310     | 50        | 804-P3  |
| 8     | Pa 107     | 63        | M 256-B| 25    | Pa 44      | 58        | 770-P1  |
| 9     | Pa 120     | 58        | 775-P5 | 26    | Pa 46      | 42        | 748-P1  |
| 10    | Pa 13      | 33        | 116-P2 | 27    | Pa 4       | 20        | 769-P6  |
| 11    | Pa 137     | 72        | 776-P8 | 28    | Pa 51      | 11        | 771-P7  |
| 12    | Pa 148     | 50        | 443-P4 | 29    | Pa 7       | 45        | M249B   |
| 13    | Pa 15      | 50        | 444-P3 | 30    | Pa 70      | 36        | 772-P6  |
| 14    | Pa 150     | 67        | 121-P8 | 31    | Pa 81      | 36        | 440-P5  |
| 15    | Pa 16      | 42        | 118-P10| 32    | Pa 88      | 33        | M272A   |
| 16    | Pa 169     | 63        | 122-P8 | 33    | Pa 88      | 27        | 773-P5  |
| 17    | Pa 175     | 33        | M 250 A| 34    | Pa 107     | 63        | 774-P1  |
Table 2. Microsatellites used and respective number of alleles obtained.

| Name* | * Sequence of the primer (5′ 3′) | Structures repeated | No.of alleles |
|-------|---------------------------------|---------------------|--------------|
| MTcCIR3 | CATCCCCAGTATCTCATCCATTTTTCTATCA | (CT)20(TA)21 | 7 |
| MTcCIR4 | CGACTAAAAACCACACCATCA | (TC(TCT)2(TC)8 | 3 |
| MTcCIR6 | TTCCCTCTAAACTCCTAATTA | (TG)4(GA)13 | 6 |
| MTcCIR8 | CAGTCCCCATTTTACAT | (TC)5TT(TC)17 TTT(CT)4 | 4 |
| MTcCIR11 | TTTGGTGATTATTTGAG | (TC)13 | 6 |
| MTcCIR12 | TCTGACCCCAAACCTGTA | (CATA)2N18(TG)6 | 6 |
| MTcCIR13 | CAGTCTAAACACAGTGAG | (AG)13 | 4 |
| MTcCIR15 | CAGCCGCTCTTGTAG | (TC)19 | 3 |
| MTcCIR18 | GATAGCTAAGGGGATTTAGGA | (GA)12 | 2 |
| MTcCIR19 | CACAACCCGTGCTGATTAG | (CT)28 | 4 |
| MTcCIR24 | TTTGGGTTGATTCTTCTGAG | (AG)13 | 2 |
| MTcCIR26 | GCATTCTCATACATCATTC | (TC)4C(CT)4 TT(CT)11 | 2 |
| TOTAL | 12 loci | 49 alleles |

Identified and characterized by Lanaud et al. (1999).

Genetic materials 8 and 34 are classified as Pa 107, however, the results demonstrate that in fact they are different materials (genetic similarity 0.55, data not shown). Mistakes and mixtures in germplasm banks are not rare and deserve special attention by researchers (Figueira, 1998; Faleiro et al., 2001b; Sounigo et al., 2001; Saunders et al., 2001).

In spite of the fact that Parinari clones are related to each other, variability exists for morphology and size of the fruits, which can be explained by the divergence and heterozygosity of the parents that formed the Parinari population, making the progenies to segregate for those characteristics. In these cases, it is possible that the parents’ heterozygosity resulted in high variability of the progenies that is confirmed with heterozygosity data of the Parinari accessions (Table 1) that varied from 11% to 72%, being equal or more than 50% in 40% of the accessions. The average heterozygosity was 44.2%, which is higher than the 23.7% found by Crouzillat et al. (2000). Pa 51, Pa 4, and Pa 285 present lower levels of heterozygosity, while Pa 150, Pa 137, Pa 169 and Pa

Figure 1. This is the bands of primer 11, one of the most polymorphic. It was considered only the bottom part of the gel.
107 showed the highest levels.

The genetic variability of Parinari accessions is also confirmed by the genetic distances that varied from 0.00 to 0.64 (data not shown). Christhoper and Soungio (1995) found an average similarity index of 0.66 using RAPD. This data indicates a high genetic diversity within Parinari populations. Also, using morphological qualitative and quantitative descriptors, Parinari accessions were clustered into 3 different groups (Bekele and Bekele, 1996). Using isoenzymes, variability was not found, probably because the low number of enzyme systems used (Sankar and Soungio, 1995).

The accessions of Parinari series has a certain genetic relationship. However, data of genetics distance and heterozygosity obtained in this work demonstrate the existence of genetic variability.

ACKNOWLEDGEMENTS

This work is part of the BIOMOL project supported by the Common Fund for Commodities, supervised by the International Cocoa Organization and executed by the Government of Brazil (CEPLAC). IBECAU also contributed to the contracting Fábio G. Faleiro.

RESUMO

Variabilidade genética e heterozigose dos acessos de cacau da população parinari (Pa) baseado em marcadores microsatélites

O Objetivo deste trabalho foi para estudar a variabilidade genética e heterozigose dos acessos de cacau da série Parinari (PA) na coleção do Banco Ativo de Germoplasma (BAG) do Centro de Pesquisas do Cacau (CEPEC), CEPLAC, Bahia, usando marcadores microsatélites. No presente estudo, 34 acessos de Theobroma cacao L. foram usados, sendo 25 da série Parinari. Doze locos de microsatélites foram analisados gerando total de 49 alelos. Os resultados indicam variabilidade genética entre os acessos Parinari com distância genética variando de 0.0 a 0.64 e o nível de heterozigossidade entre 11 e 72%. Em torno de 40% dos acessos apresentaram heterozigose igual ou superior a 50%.

Figure 2. Graphic dispersion and cluster analysis of 34 cocoa accessions (Theobroma cacao L.) based on genetic distances calculated from microsatellite markers. The grouping analyses were made using the method UPGMA, of the SAS system, considering 2 groups.
REFERENCES

Adomako, B. and Adu-Ampomah, Y. 2000. Reflections on the yield of Upper Amazon cocoa hybrids in Ghana with reference to breeding for cocoa swollen shoot virus resistant. Cocoa Growers’ Bulletin. 52:33-45.

Araujo, I.S.; Bahia R.C.; Santos R.F.; Faleiro F.G. and Ahnert D. 2000. Otimização da extração e amplificação de DNA de Theobroma cacao L. visando a obtenção de marcadores moleculares RAPD. Genetics Molecular Biology. 2(suppl.):219-220. (resumo).

Bartley, B.G.D. 1983. Retrospect of research on resistance of cacao to Crinipellis perniciosa. p.32-35. In: Proceedings of the International Workshop on the witches’broom disease Crinipellis perniciosa (Stahel) Singer of cacao, 2nd, Amazônia, 1982.

Bekele, F. and Bekele I. 1996. A sampling of the phenetic diversity of cacao in the International Cocoa Gene Bank of Trinidad. Crop Sci ence. 36:57-64.

Charters, Y.M. and Wilkinson M.J. 2000. The use of self-pollinated progenies as ‘in-groups’ for the genetic characterization of cocoa germplasm. Theoretical and Applied Genetics. 100:160-166.

Christopher, Y and Sounigo, O. 1995. The use of RAPD for characterization and genetic assessment of cacao. p.38-51. In: Cocoa Research Unit. Report for 1995. The University of the West Indies, St. Augustine, Trinidad and Tobago.

Crouzillat, D.; Bellanger, L.; Rigoreau, M.; Bucheli P. and Pétiard,V. 2000. Genetic structure, characterization and selection of Nacional cocoa compared to other genetic groups. p.47-64. In: International Workshop on New Technologies and Cocoa Breeding, INGENIC 3. kota Kinabalu, Sabah, Malaysia.

Doyle, J.J. and Doyle, J.L. 1990. Isolation of plant DNA from fresh tissue. Focus. 12:13-15.

Faleiro, F.G.; Pires, J.L.; Lopes, U.V.; Yamada, M.M.; Moura, A.D.; Gardini,E.A.; Burbano, G.P.; Monteiro, W.R.; Marques,J.R.B.; Gramacho, K.; Faleiro A.S.G. and Santos, M.C.M. 2001a. Variabilidade genética de variedades de Theobroma cacao L. provenientes das amazônicas brasileira, equatoriana e peruana com base em marcadores microsatélites. In: Resumos do Congresso Nacional de Genética, 47ª, Aguas de Lindoa, 2001. Sociedade Brasileira de Genetica, Aguas de Lindoa. CD.

Figueira, A. 1998. Homonymous genotypes and misidentification in germplasm collection of Brazil and Malaysia. INGENIC Newsletter . 4:4-8.

Ivaro, A. D.; Singh, V.; Bharath, S.and Jugmohan, N. 2001. Germplasm evaluation at the International Cocoa Genebank, Trinidad for resistance to Phytophthora pod rot. p.30-33. In: Cocoa Research Unit. Report for 2000. The University of the West Indies, St. Augustine, Trinidad and Tobago.

Lanaud,C.; Risterucci, A.M.; Pieretti, I.; Falque M.; Bouet, A. and Lagoda,P.J.l. 1999. Molecular Ecology. 8:2141-2143.

Lerceateau, E.; Crouzillat D. and Petiard,V. 1992. Use of random amplified polymorphic DNA (RAPD) and restriction fragment length polymorphism (RFLP) to evaluate genetic variability within the Theobroma genus. p.332-344. In: Proceeding of the International workshop on conservation, characterization and utilisation of cocoa genetic resources in the 21st century. The Cocoa Research Unit, The University of West Indies, Port-of-Spain, Trinidad.

Lockwood, G. and End, M. 1993. History, technique and future needs for cacao collection. p.1-14. In: Proceeding of the Workshop on the Conservation, Characterization and Utilization of Cocoa Genetic Resources in the 21st Century. The Cocoa Research Unit, Port of Spain, Trinidad.

Luz, E.D.M.N.; Silva, S. D. V. M.; Yamada, M.M.; Lopes, U.V.; Braga M.C. T. and Brugnerotto, M.I.B. 1996. Selection of cacao genotypes resistant to Phytophthora capsici, P. palmivora and P. citrophthora in Bahia, Brasil. Fitopatologia Brasileira. 21(1):71-79.

Marita, J.M. 1998. Characterization of Theobroma cacao using RAPD-marker based estimates of genetic distance and recommendations for a core collection to maximize genetic diversity. M. S. Thesis.University of Wisconsin, USA.

Pires, J. L.; Luz, E.D.N.; Lopes U.V. and Bartley, B. G.D. 1994. Incidência natural de podridão-parda em
clones de cacauzeiro na Bahia, Brasil. p.871-877. In: Proceedings of the International Cocoa Research Conference, 11th, Yamousoukro, Ivory Coast, 1993.

Pound, J.F. 1938. Cacao and witches’broom disease (Marasmius perniciosus) of South America with notes on other species of Theobroma. Report on a visit to Ecuador, the Amazon Valley and Colombia,April 1937-April 1938. Trinidad, Port of Spain.

Risterucci, A.; Eskes, E.; Fargeas, D.; Motamayor, J. and Lanaud, C. 2000. Use of microsatellite markers for germplasm identity analysis in cocoa. p.25-33. In: International Workshop on New Technologies and Cocoa Breeding, INGENIC 3. Kota Kinabalu, Sabah, Malaysia.

Sambrook, J.; Fritsch E.F. and Maniatis, T. 1989. Molecular cloning: a laboratory manual. 2ed. Cold Spring Harbor, Cold Spring Harbor Laboratory, New York.

SAS INSTITUTE. 1990. SAS/STAT. User’s guide. Version 6. 4th ed. SAS, Cary.

Sankar, A. and Sounigo, O. 1995. The use of isozyme electrophoresis for characterization and genetic diversity assessment of cacao. p.52-63. In: Cocoa Research Unit. Report for 1995. The University of the West Indies, St. Augustine, Trinidad and Tobago.

Saunders, J.; Hemeida A.A. and Mischles, S. 2001. DNA fingerprinting programme for identification of Theobroma cacao accessions. p.108-114. In: International Workshop on New Technologies and Cocoa Breeding, INGENIC 3. Kota Kinabalu, Sabah, Malaysia.

Sounigo, O.; Christopher Y. and Sankar, A. 1995. Overview of the activities of the biochemistry laboratory. p.33-37. In :Cocoa Research Unit. Report for 1995. The University of the West Indies, St. Augustine, Trinidad and Tobago.

Sounigo, O.; Cristopher, Y.; Bekeke, F.; Mooledhar V. and Hosein, F. 2001. The detection of mislabelled trees in the International Cocoa Genebank, Trinidad (ICG, T) and options for a global strategy for identification of accessions. p.34-39. In: International Workshop on New Technologies and Cocoa Breeding, INGENIC 3. Kota Kinabalu, Sabah, Malaysia.

Umaharan, R.; Thevenin J.M. and Surujdeo-Maharaj, S. 2001. Evaluation of cocoa germplasm for resistance to witches’broom disease. p.30-33. In: Cocoa Research Unit. Report for 2000. The University of the West Indies, St. Augustine, Trinidad and Tobago.

Yamada, M. M.; Bartley, B.G.D.; Lopes, U. V. and Pinto, L.R.M. 1996. Herança do fator compatibilidade em Theobroma cacao L. II. Relações fenotípicas em genótipos adicionais do grupo Parinari (Pa). Agrotrópica. 8(2):51-52.

Yamada, M. M. and Guries, R.P. 1994. Variação genética de três sistemas isoenzimáticos em clones de cacau (Theobroma cacao L.) da serie Parinari. Agrotrópica. 6(1):27-29.

Yamada, M. M. and Guries, R.P. 1998. Mating system analysis in cacao (Theobroma cacao L.). Agrotrópica. 10(3):165-176.

Received: December 04, 2002;
Accepted: January 27, 2003.