# GENETIC DIVERSITY AND SEX IDENTIFICATION IN *Genipa americana* L.

## [DIVERSIDAD GENÉTICA E IDENTIFICACIÓN SEXUAL EN *Genipa americana* L.]

**Ana Veruska Cruz da Silva[1]*, Evandro Neves Muniz[1], Camila Santos Almeida[2], Marina Ferreira da Vitória[2], Ana da Silva Ledo[1], Marília Freitas de Vasconcelos Melo[3], Allivia Rouse Carregosa Rabbani[4]**

*1. Researcher at Embrapa Coastal Tablelands, 49025-040, Aracaju-SE, Brazil. Email: ana.veruska@embrapa.br*
*2. Federal University of Sergipe, 49100-000, São Cristóvão-SE, Brazil*
*3. Faculty of Agricultural Sciences, State University of Sao Paulo, 18610-307, Botucatu-SP, Brazil*
*4. Federal Institute of Education, Science and Technology of Bahia, 45810000, Porto Seguro-BA, Brazil*
**Corresponding Author*

**SUMMARY**

Ten genotypes of *Genipa americana* L. located in Arauá (Sergipe, Brazil) were analyzed using RAPD markers. The genotypes were selected according to sex - five females and five males - in an attempt to find a molecular marker associated with the sex prediction. 74 DNA fragments were generated by 20 primers, 85% of which were polymorphic. Results show a quantitative genetic diversity value of 0.28 and a Shannon index of 0.42 in the population. The similarity among the specimens according to Jaccard's coefficient ranged from 0.22 to 0.75. The genetic diversity was relatively large among these individuals.

**Key words**: Tropical fruits; Genetic resources; Ecosystems.

**RESUMEN**

Diez genotipos de *Genipa americana* L. ubicados en Arauá (Sergipe, Brasil) fueron analizados usando marcadores RAPD. Los genotipos fueron seleccionados de acuerdo con el sexo en intento de encontrar un marcador sexual asociado a la predicción del sexo. 74 fragmentos de DNA fueron generados por 20 indicadores, el 85% de los cuales fueron polimórficos. Los resultados muestran un valor de diversidad genética cuantitativa de 0.28 y un índice de Shannon de 0.42. La similitud entre las especies de acuerdo con el coeficiente de Jaccard fue de 0.22 a 0.75. La diversidad genética entre esos individuos fue relativamente grande.

**Palabras clave:** Frutas tropicales; Recursos genéticos; Ecosistemas.

## INTRODUCTION

*Genipa americana* L. – Rubiaceae, commonly known as genipap, is a fruit species naturally occurring between Mexico and Brazil, with ecological, social and cultural importance. Its distribution is common in riparian areas, withstands long periods of immersion during the flood season the rivers. The species is highlighted by its multiplicity of uses. It is recommended for recovery of degraded areas, and its fruits are edible, with important quality attributes for foods industry, dye (Penalber *et al*., 1996) and traditional medicine to treat anemia, icterus, asthma, and liver and spleen problems (Conceição *et al*., 2011).

The species has high ecological plasticity occurring in several forest formations throughout tropical America, suggesting genetic diversity among individuals and populations, since differentiation between species is based on both genetic and environmental factors.

The sex determination in some fruits is still not well understood. The development of mechanisms of sex determination are extremely useful in plant breeding (Alstrom-Rapaport *et al*., 1998). In papaya, Souder *et al*. (1996) constructed a genetic linkage map using RAPD and defined two markers flanking for determining locus about 7 cM. Has a many farmers dissatisfied by the high rate of male plants in the areas of genipap, but research on the species is still limited.

Besides the analysis of genetic diversity, the techniques based on PCR have used to detect markers associated with sex in dioecious and bisexual plants. RAPD has been used in several studies, as in Guggal (*Commiphora wightii* R.) (Sanghamitra *et al*., 2010), *Hancornia speciosa* Gomes (Costa *et al*., 2011), *Myrciaria tenella* O. Berg (Pinheiro *et al*., 2011), *Cattleya labiata* (Pinheiro *et al*., 2012) among many other species.

The goal of this study was to evaluate the genetic diversity and try to identify the sex of brands associated genipap in a natural population located in Arauá, Sergipe, Brazil.

## MATERIAL AND METHODS

Previously characterized as to sex by local farmers, located in a natural population in the city - five male subjects (279, 281, 283, 285, 287) and five female individuals (280, 282, 284, 286, 288) were randomly selected of Arauá (latitude 11 º 17'23" and longitude 37 º 38'31"), Sergipe, Brazil. All subjects were georeferenced, and the healthy and ripe fruits were transported to the Seed Laboratory of Embrapa Coastal Tablelands (CPATC) in Aracaju-SE for the formation of seedlings in nursery.

DNA was isolated from 1g young leaves per plant as described by Doyle and Doyle (1990). For PCR-ISSR, 20 primers from Operon (Operon Technologies, EUA) and IDT (Integrated DNA Technologies, Germany) were used to screen for polymorphisms (Table 1). Polymerase chain reaction (PCR) amplifications were performed using a PTC-100 thermocycler (Programmable Thermal Controller, MJ Research Inc., Waltham, MA, United States), with reactions consisting of 5 min initial denaturation at 95ºC, followed by 45 cycles of denaturation at 94ºC for 1 min, primer annealing at 52ºC for 1 min and extension at 72ºC for 45 s, and one final extension step at 72ºC for 10 min.

Fragments were visualized on 2% agarose gel [1 x TBE – 89 mM Tris, 89 mM boric acid, 2.5 mM ethylenediamine tetraacetic acid (EDTA), pH 8.3] in a horizontal electrophoresis system (Sunrise, Gibco BRL) run at a constant voltage of 100 V for 60 min. The gel was stained with ethidium bromide solution (5 mg/mL) for 30 min. ISSR amplification products were visualized under ultraviolet light using a Gel Doc L-Pix imaging system (Loccus Biotecnologia, Brazil).

Only data from intensely stained, unambiguous and clear fragments were used for statistical analysis. ISSR markers were scored as present (1) or absent (0). The polymorphic information content (PIC) was calculated according to Ghislain (1999) and the marker index (MI) was determined as described by Zhao *et al*. (2007). A data matrix of ISSR scores was generated, and similarity coefficients were calculated using Jaccard's arithmetic complement index. A dendrogram was constructed using the Unweighted Pair Group with Arithmetic Mean (UPGMA) cluster algorithm. To determine dendrogram robustness, the data were bootstrapped with 10.000 replicates using FreeTree software (Pavlicek *et al*., 1999). Principal coordinates analysis method (PCoA) was performed using the software package XLSTAT (http://www.xlstat.com/) based on the calculated Jaccard's similarity coefficients. The Shannon diversity index (I) and genetic diversity (G) (Brown and Weir, 1983) were calculated using Genalex v.6.3 (Peakall and Smouse, 2005).

## RESULTS AND DISCUSSION

The primers used allowed the molecular characterization and study of the genetic diversity among individuals of genipap. However, no molecular markers associated with sex was observed. The 20 primers used in generating a total of 74 bands, of which 63 were polymorphic (85%) (Table 1). The primer showed the highest number of polymorphic bands was the X01 (6), while the A15 and W2 showed no polymorphic bands. The PIC value ranged from 0.00 to 0.48 and IM from 0.41 to 2.00. For the values of genetic structure, the Shannon index (I) averaged 0.42, and the genetic diversity index (G) presented a mean value of 0.28.

The percentage of polymorphism was found to consistently with Rabbani *et al*. (2012), using 12 primers in a natural population in the riparian area of the Lower São Francisco (Sergipe State) gave 74 polymorphic fragments, a total of 88 (84%).

**Table 1.** Number of polymorphic fragments (NPF), % polymorphism, polymorphic information content (PIC), marker index (MI) and genetic diversity (G) from 10 Genipa Americana L. individuals from Sergipe, Brazil.

| Primer | Sequence | NPF | %P | PIC | MI | I | G |
|---|---|---|---|---|---|---|---|
| A2 | TGC CGA GCT G | 5 | 60 | 0,27 | 0,54 | 0,39 | 0,27 |
| A3 | AGT CAG CCA C | 4 | 50 | 0,23 | 0,46 | 0,33 | 0,23 |
| A4 | AAT CGG GCT G | 5 | 100 | 0,35 | 1,76 | 0,52 | 0,35 |
| A8 | GTG ACG TAG G | 4 | 100 | 0,48 | 1,92 | 0,67 | 0,48 |
| A9 | GGG TAA CGC C | 5 | 100 | 0,30 | 1,50 | 0,47 | 0,30 |
| A10 | GTG ATC GCA G | 5 | 100 | 0,40 | 2,00 | 0,57 | 0,40 |
| A11 | CAA TCG CCG T | 5 | 60 | 0,19 | 0,57 | 0,30 | 0,19 |
| A13 | CAG CAC CCA C | 3 | 33 | 0,14 | 0,14 | 0,20 | 0,14 |
| A15 | TTC CGA ACC C | 0 | 100 | 0,00 | 0,00 | 0,00 | 0,00 |
| A16 | AGC CAG CGA A | 3 | 100 | 0,28 | 0,84 | 0,44 | 0,28 |
| A18 | AGG TGA CCG T | 3 | 100 | 0,29 | 0,87 | 0,45 | 0,29 |
| A20 | GTT GCG ATC C | 3 | 100 | 0,44 | 1,31 | 0,63 | 0,44 |
| IDT18 | GGA GGA GAG G | 3 | 33 | 0,12 | 0,12 | 0,22 | 0,12 |
| B18 | CCA CAG CAG T | 4 | 100 | 0,41 | 1,65 | 0,60 | 0,41 |
| X01 | CTG GGC ACG A | 6 | 100 | 0,27 | 1,60 | 0,43 | 0,27 |
| X03 | TGG CGC AGT G | 4 | 100 | 0,34 | 1,37 | 0,52 | 0,34 |
| W2 | ACC CCG CCA A | 0 | 100 | 0,00 | 0,00 | 0,00 | 0,00 |
| W4 | CAG AAG CGG A | 4 | 100 | 0,32 | 1,28 | 0,48 | 0,32 |
| W13 | CAC AGC GAC A | 4 | 100 | 0,45 | 1,79 | 0,64 | 0,45 |
| K20 | GTG TCG CGA G | 4 | 100 | 0,33 | 1,34 | 0,50 | 0,33 |
| Total | | 4 | 87 | 0,28 | 1,05 | 0,42 | 0,28 |

The Shannon index (I ≈ 0.42) can be considered low, as it was below 0.5. The 'I' varies from 0 to 1, and it is considered that the closer the value is to zero, the lower the diversity. The values found in this study are similar to those obtained by Allnut et al. (1999) for other tree species *Fitzroya cupressoides* with 0.42 to 0.56 and *Swietenia macrophylla*, with 0.27 to 0.41 (Gillies *et al*., 1999). The mean genetic diversity (G ≈ 0.28) obtained are in agreement with those found for other species. For *Trichilia pallida* Swartz ranged from 0.27 to 0.33 (Zimback *et al*., 2004) and *Aspidosperma polyneuron*, the average range is 0.28 (Torezan *et al*., 2005). The nonzero valuesfor L indicate that there is an excess of homozygotes or heterozygotes in the sample group (Barreira *et al*., 2012).

The high degree of polymorphism identified among the individuals studied is the heterozygosity for this marker, which can contribute to infer the genetic relationships (Table 2). According to Jaccard coefficient, the lowest genetic similarity was found between 285 and 279 individuals (0.22) and highest among 288 individuals and 287 (0.78).

There was the formation of three groups. The first consisted of four individual (280-283), the second with six (284-288) and third with only one individual (279), which was the most divergent within the population (Figure 1).

The similarity data are in accordance with the ACoP. In this study, the efficiency of RAPD markers was confirmed, because it was found that the percentage of variability accumulated in the first two axes was 52.26%, which confirms the genetic diversity among the individuals in this study. A bidimensional graphic was generated, and the parameters defined in the analysis reinforced the positioning of the different genotypes (279) (Figure 2).

It was possible to verify a natural clustering structure, depending on the genetic similarity between them evaluated. The use of more than one clustering method, due to differences in the ranking, optimization and ranking the groups, allows classification is complemented by the criteria that each technique uses, and prevents erroneous inferences are adopted in the allocation of materials, within a given genotype subgroup (Silva *et al*., 2012). To genipap, the most divergent individual was 279.

**Table 2.** Genetic similarity (%) (below diagonal) and estimated error (above diagonal) among 10 *Genipa americana* L. individuals.

| | 279 (M) | 280 (F) | 281 (M) | 282 (F) | 283 (M) | 284 (F) | 285 (M) | 286 (F) | 287 (M) | 288 (F) |
|---|---|---|---|---|---|---|---|---|---|---|
| 279 (M) | - | 0,33 | 0,25 | 0,27 | 0,32 | 0,25 | 0,22 | 0,28 | 0,25 | 0,23 |
| 280 (F) | 0,33 | - | 0,74 | 0,63 | 0,50 | 0,50 | 0,60 | 0,46 | 0,52 | 0,56 |
| 281 (M) | 0,25 | 0,74 | - | 0,64 | 0,51 | 0,51 | 0,58 | 0,47 | 0,53 | 0,60 |
| 282 (F) | 0,27 | 0,63 | 0,64 | - | 0,75 | 0,60 | 0,65 | 0,52 | 0,57 | 0,66 |
| 283 (M) | 0,32 | 0,50 | 0,51 | 0,75 | - | 0,50 | 0,54 | 0,46 | 0,57 | 0,58 |
| 284 (F) | 0,25 | 0,50 | 0,51 | 0,60 | 0,50 | - | 0,66 | 0,63 | 0,61 | 0,67 |
| 285 (M) | *0,22* | 0,60 | 0,58 | 0,65 | 0,54 | 0,66 | - | 0,53 | 0,58 | 0,71 |
| 286 (F) | 0,28 | 0,46 | 0,47 | 0,52 | 0,46 | 0,63 | 0,53 | - | 0,68 | 0,60 |
| 287 (M) | 0,25 | 0,52 | 0,53 | 0,57 | 0,57 | 0,61 | 0,58 | 0,68 | - | 0,78 |
| 288 (F) | 0,23 | 0,56 | 0,60 | 0,66 | 0,58 | 0,67 | 0,71 | 0,60 | *0,78* | - |

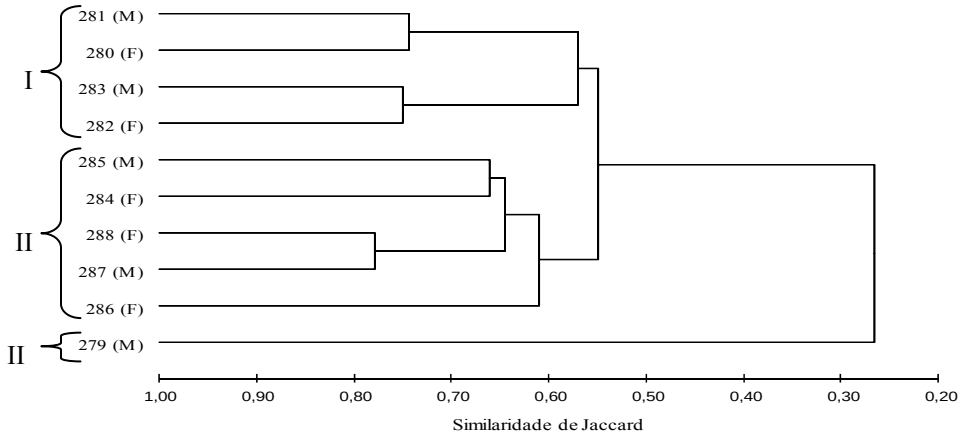


**Figure 1.** Genetic similarity dendrogram generated by Jaccard's similarity coeficiente and bootstrap analysis (10.000 x) and groups (I, II, III) for the 10 *Genipa americana* L. individuals using RAPD markers.

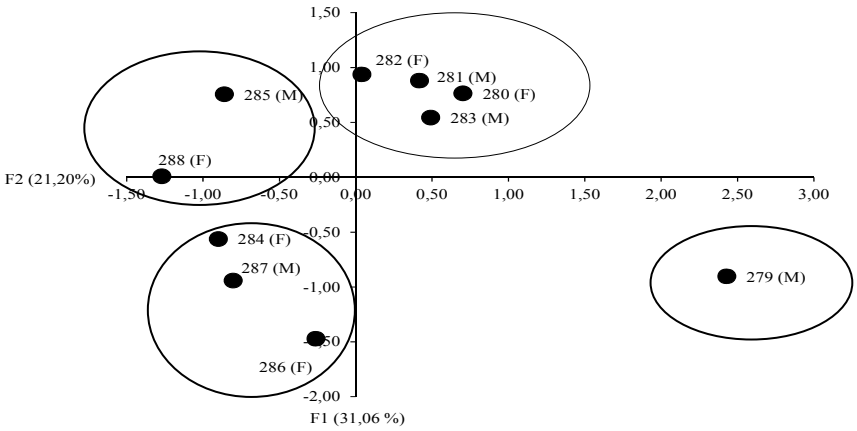


**Figure 2.** Principal Coordinates Analysis (PCoA) of 10 Genipa Americana L. individuals collected from natural population in Sergipe, Brazil based on RAPD markers data./HV

## CONCLUSION

This is the first report to predict sex genipap reported in northeastern Brazil. Unable to identify a specific molecular tag to sex using the primers tested, but our study is a contribution to the characterization of native genotypes. The genetic variation and the genetic relationships among wild genipap were efficiently determined using RAPD markers. The discrimination of genipap from Sergipe (Brazil) and identification of genotypes more genetically divergent may contribute to development of strategies for the implementation of conservation, breeding programs efforts and commercial exploitation.

## REFERENCES


Allnut, T.R.; Newton, A.C.; Lara, A.; Premoli, A.; Armesto, J.J.; Vergara, R. and Gardner, M. , 1999. Genetic variation in *Fitzroya cupressoides* (alerce), a threatened South American conifer. Molecular Ecology, Malden, 8 (6): 975-87.

Alstrom-Rapaport, C; Lascoux, A.; Wang, Y.Y.C.; Roberts, G. and Tuskan, G.A. 1998. Identification Of A Rapd Marker Linked To Sex Determination In The Basket Willow (*Salix Viminalis* L.). Journal Of Heredity, 89 (1): 44-49.

Barreira, S.; Sebbenn, A.M.; Scolforo, J.R.S. and Kageyama, P.Y. 2006. Diversidade genética e sistema de reprodução em população nativa de *Eremanthus erythropappus* (DC) MacLeish sob exploração. Scientia Forestalis, 71: 119-130.

Brown, A.H.D. and Weir, B.S. 1983. Measuring genetic variability in plant populations, in Isozymes in Plant Genetics and Breeding, Part A (Tanksley SD, Orton TJ, Editors). Elsevier Science Publishing, Amsterdam, p. 219-239.

Conceição, A.O.; Rossi, M.H.; Oliveira, F.F.; Takser, L. and Lafond, J. 2011. *Genipa americana* (Rubiaceae) fruit extract affects mitogen-activated protein kinase cell pathways in human trophoblast-derived BeWo cells: implications for placental development. Journal of Medicinal Food, 14: 483-494.

Costa, T. S.; Silva, A.V.C.; Ledo, A.S.; Santos, A.R.F. and Silva Júnior, J.F. 2011 Diversidade genética de acessos do banco de germoplasma de mangaba em Sergipe. Pesquisa Agropecuária Brasileira, 46 (5): 499-507.

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

Ghislain, M.; Zhang, D.; Fajardo, D.; Huamán, Z. and Hijmans, R.J. 1999. Marker assisted sampling of the cultivated Andean patato Solanumphureja collections using RAPD markers. Genetic Resources and Crop Evolution, 46: 547-555.

Gillies, A.C.M.; Navarro, C., Lowe, A.J., Newton, A.C.; Hernández, M.; Wilson, J. and Cornelius, J.P. 1999. Genetic diversity in Mesoamerican populations of mahogany (*Swietenia macrophylla*), assessed using RAPDs. Heredity, 83 (6): 722-32.

[Pavlicek, A.; Hrda, S. and Flegr, J. 1999. FreeTree – freeware program for construction of phylogenetic trees on the basis of distance data and bootstrap/jackknife analysis of the tree robustness. Application in the RAPD analysis of genus *Frenkelia*. Folia Biologica, 45: 97-99.

Peakall , R. and Smouse, P. E. 2005.GenAlEx 6: Genetic analysis in Excel., population genetic software for teaching and research. Australian National University, Canberra, Australia.

Penalber, T.J.A.; Sadala, M.A.C.; Castro, M.S. and Faria, L.J.G. 1996. Ensaios de extração e aplicação de corantes do fruto do jenipapeiro (*Genipa americana*). Revista Brasileira de Corantes Naturais, 2:129-135.

Pinheiro, L.R.; Almeida, C.S. and Silva, A.V.C. 2011. Diversidade genética de uma população natural de cambuizeiro e avaliação pós-colheita de seus frutos. Scientia Plena, 7 (6): 1-5.

Pinheiro, L.R.; Rabbani, A.R.C.; Silva, A.V.C.; Lédo, A.S.; Pereira, K.L.G.; Diniz, L.E.C. 2012. Genetic diversity and population structure in the Brazilian *Cattleya labiata* (Orchidaceae) using RAPD and ISSR markers. Plant Systematics and Evolution, 12.

Rabbani, A.R.C.; Silva-Mann, R. and Ferreira, R.A. 2012.Variabilidade genética de *Genipa americana* L. pertencente ao baixo curso do rio São Francisco. Revista Árvore, 36 (3): 401-409.

Sanghamitra S.; Geetha, K.A.; Hidayath, K.P. and Maiti, S. 2010. Identification of RAPD markers linked to sex determination in guggal [*Commiphora wightii* (Arnott.)] Bhandari. Plant Biotechnology Reports, 4 (1): 95-99.

Silva, A. V. C.; Santos, A.R.F.; Lédo, A.S.; Feitosa, R.B.; Almeida, C.S.; Silva, G.M. and Rangel, M.S.A. 2012. .Moringa genetic diversity from Germplasm Bank using RAPD markers. Tropical and Subtropical Agroecosystems, 15 (1): 31-39.

Souder, S.N.; Manshardt, R.M. and Stiles, J.I. 1996. A genetic linkage map of papaya based on randomly amplifies polymorphic DNA markers. Theoretical Applied Genetics, 93: 547-553.

Torezan, J.M.D; Souza, R.F.; Ruas, P.M.; Ruas, C.F.; Camargo, E.H. and Vanzela, A.L.L.

2005.Genetic variability of pre and post-fragmentation cohorts of *Aspidosperma polyneuron* Muell. Arg. (Apocynaceae). Brazilian Archives of Biology and Technology, 48 (2): 171-180.

Zhao, K.; Zhou, M.Q. and Chen, L.Q. 2007. Genetic diversity and discrimination of Chimonanthuspraecox (L.) link germplasm using ISSR and RAPD Markers. Hortscience, 42: 1144–1148.

Zimback, L.; Kageyama, P.Y.; Veiga, R.F.A. and Mello Júnior, J.R.S. 2004. Estrutura genética de populações de *Trichilia pallida* Swartz (Meliaceae) por marcadores RAPD. Scientia Forestalis, 65: 114-119.

*Submitted February 11, 2015 – Accepted March 27, 2015*