

Full Length Research Paper

Estimate of genetic diversity in cassutinga (*Croton heliotropiifolius*) based on molecular markers

Tallita de Oliveira Rocha¹, Janaina Silva Freitas^{1,2}, Elisa Susilene Lisboa dos Santos², Murilo Marques Scaldaferri², Claudine Gonçalves de Oliveira² and Carlos Bernard Moreno Cerqueira-Silva^{1,2*}

¹Programa de Pós-graduação em Ciências Ambientais, Universidade Estadual do Sudoeste da Bahia, Itapetinga, BA, 45700-000, Brazil.

²Departamento de Ciências Exatas e Naturais, Universidade Estadual do Sudoeste da Bahia, Itapetinga, BA, 45700-000, Brazil.

Received 26 September, 2015; Accepted 25 January, 2016

The *Croton heliotropiifolius* (Euphorbiaceae family) is a shrubby plant that has attracted attention both by the need for conservation as for their pharmacological potential. The present study aimed to characterize the diversity and structure genetic of a population of *C. heliotropiifolius* present in a forest fragment in the city of Itapetinga, Bahia. Young leaves were collected from 41 individuals after DNA extraction were genotyped with 18 RAPD primers and 15 ISSR primers. Frequentist and Bayesian statistical methods were used to estimate the diversity and genetic structure, being observed a total of 164 polymorphic markers (mean 4 and 4.3 bands obtained with the use of RAPD and ISSR primers, respectively). Genetic diversity ranging between 0.12 and 0.48 and Bayesian method indicated by the existence of three probable gene pools ($K = 3$). Indications of association between spatial distribution of plants and the genetic structure were also observed, being likely that the dynamics of the seed carriage performed by the ants and the pollination by insects are related with the results observed. These results allow the beginning of a discussion about the diversity and genetics patterns of the species, since there are no studies of this nature for *C. heliotropiifolius*.

Key words: Caatinga, conservation genetics, genetic variability, inter simple sequence repeat (ISSR), random amplified polymorphic DNA (RAPD).

INTRODUCTION

The genus *Croton* consists of approximately 1300 species that are mostly shrubs and subshrubs (Zioldo, 2007). Several species of *Croton* spp. are known to have antioxidant and antimicrobial properties in potential, or for being used routinely as phytotherapy by local or regional

populations (in Caatinga biome) (Abreu et al., 2001). Among the species of croton that are used as medicinal plants, it is possible to cite the *Croton heliotropiifolius*, a species popularly known in northeastern Brazil as "cassutinga". This species is used to relieve

*Corresponding author. E-mail: csilva@uesb.edu.br Tel: +55 77 3261-8677. Fax: +55 77 3261-8609.

stomachaches in general, vomiting and diarrhea, apart from relieve fever (Randau, 2001), and also has important constituents for pharmacological and phytochemical studies, highlighting the presence of alkaloids and reducing sugars (Randau et al., 2004).

Despite its potential use as a medicinal plant and its wide dispersion in the Caatinga biome, genetic studies in *C. heliotropiifolius* are limited to testing methods for DNA extraction (Scaldfarri et al., 2013) and preliminary studies of genetic diversity (Scaldfarri et al., 2014). In contrast, there has been a growth in the number of associated studies with molecular techniques in plants, especially for genetic studies aimed at characterizing the diversity (Cerqueira-Silva et al., 2014). Genetic diversity consists in the existence of different phenotypic characteristics, as well as in molecular differences (enzymes, proteins and DNA sequences), present in individuals of a population (Frankham et al., 2008). In this context, Futuyma (2002) highlights that it is essential to carry out genetic studies to understand the different aspects of a natural population, such as diversity and genetic structure that contribute to an extensive discussion about the ecology of species.

The use of molecular markers in genetic characterization is routine in many laboratories and various molecular markers are available, however, despite the advancement and popularization of many molecular techniques in recent decades, Cerqueira-Silva et al. (2014) argue that the availability of resources and background information are still determining factors in the choice of markers used in genetic researches.

Considering the importance of genetic information, as well as the absence of molecular genetic diversity estimates for *C. heliotropiifolius*, the objective of this study was to characterize using RAPD (Random Amplified Polymorphic DNA) and ISSR (Inter Simple Sequence Repeat) markers, the diversity and structure genetic existent among 41 wild genotypes of this species that were collected in a fragment of native forest in Itapetinga city, Bahia, Brazil.

MATERIALS AND METHODS

Sample collection and genomic DNA extraction

The collection was carried out in a small hill in the city of Itapetinga, Bahia, Brazil, located in the southwestern region of the state. The vegetation observed in this area is typical of deciduous and semi-deciduous forest (Radam Brasil, 1981). In total were collected leaves of 41 wild genotypes of *C. heliotropiifolius*, representing three distinct regions along of the small hill (-15.256456 lat, -40.257637 long), as follows: 12 genotypes representative of the top, 16 genotypes representative of the middle region and 13 genotypes representative of the valley of the hill. Sample of this species was identified in the herbarium of the Universidade Estadual de Feira de Santana (UEFS, BA, Brazil) and its exciscatas have been properly deposited in this herbarium (under the codes A1-HUEFS189022, A2-HUEFS189023, A3-HUEFS189024, B1-HUEFS189025, B2-HUEFS189026, B3-HUEFS189027, C1-

HUEFS189028, C2-HUEFS189029, C3-HUEFS189030, D1-HUEFS189031, D2-HUEFS189032, D3-HUEFS189033).

The genomic DNA was isolated from fresh leaves using the cetyltrimethylammonium bromide (CTAB) method (Doyle and Doyle, 1990) with modifications previously tested for *Croton* by Scaldfarri et al. (2013). The quality of the DNA samples was assessed in agarose gel 1% (w/v) by electrophoresis (in a 90 V for 100 min) and visualized with a gel red staining buffer (Invitrogen), according to the manufacturer's specifications. In order to quantify the DNA concentration (ng/ μ L) we adopted an intact Lambda molecular weight marker as standard (undigested Lambda DNA).

Genotyping of samples

The DNA samples of the 41 genotypes were tested with 20 RAPD primers and 23 primers ISSR, adopting routines used by the research group in assays with the genus *Croton* in the Applied Molecular Genetics Laboratory of UESB (Scaldfarri et al., 2013, 2014). From these primers were selected those with better quality in the patterns of bands and better genetic repeatability.

The amplifications with both types of primers were conducted in a total volume of 15 μ L containing 15 ng of the DNA, PCR buffer 1X (20 mM Tris HCl [pH 8.4] and 50 mM KCl), 1.5 mM MgCl₂, 0.2 μ M of each dNTP, 1 μ M primer and 1 U of Taq DNA polymerase (Invitrogen, Carlsbad, California, USA). The amplification program adopted for PCR reactions were: 94°C for 5 min, followed by 34 cycles [94°C for 50 s, 48°C (ISSR reaction) or 34°C (RAPD reaction) for 50 s and 72°C for 1 min], with the final extension at 72°C for 5 min. The amplification products were separated by electrophoresis in 2% agarose gel (w/v) in TBE 1X running buffer at a constant voltage of 120 V for approximately 2 h.

Analysis of molecular data

Analysis was performed in duplicate and only patterns obtained clearly twice were scored. Presence or absence of fragments were recorded as 1 or 0 (respectively), and treated as binary characters. Resulting matrices of molecular data for all primers were subjected to multivariate statistical analyses, such as: estimate on the complement of genetic similarity ($dg_{ij} = 1 - sg_{ij}$; similarity = sg_{ij} and dissimilarity = dg_{ij}) based on the coefficient of Dice (1945), and clustering of genotypes by using the neighbor joining method. The statistical analyses were carried out with the assistance of the Genes software, Windows version (Cruz, 2006) and DarWin software (Perrier and Jacquemoud-Collet, 2006).

Analyses of population structure were performed with the Bayesian method using the STRUCTURE software, version 2.3.4 (Falush et al., 2003; Pritchard et al., 2000). We used an admixture model, and the burn-in period and replication numbers were set to 100,000 and 1,000,000, respectively, for each run. The number of groups (K) was systematically varied from 1 to 10, and 10 simulations were performed to estimate each K . We used the ΔK *ad hoc* method described by Evanno et al. (2005) and implemented in the online tool Structure Harvester (Earl and Vonholdt, 2012) to estimate the most likely K in each set.

After estimating the most likely K , we used the "greedy algorithm" implemented in CLUMPP software, version 1.1.1 (Jakobsson and Rosenberg, 2007) with a random input order and 1000 permutations to align the runs. The results were visualized using DISTRUCT software, version 1.1 (Rosenberg, 2004). Based on the posterior probability of membership (q) of a given individual belonging to a given group compared to the total number of groups (K), we classified individuals with $q > 0.60$ as a member of a given cluster, whereas for clusters with membership (q) values ≤ 0.60 , the individuals was classified as admixed.

Table 1. Primers used for obtaining RAPD markers with number total of bands (N° bands), number of polymorphic bands (N° PB) and percentage of polymorphism for 41 wild genotypes of *Croton heliotropiifolius*.

Primer*	Sequence (5'-3')	N° bands	N° PB	Polymorphism (%)
OPD-01	ACCGCGAAGG	4	4	100
OPD-02	GGAACCCAACC	4	4	100
OPD-03	GTCGCCGTCA	5	4	80
OPD-05	TGAGCGGACA	6	6	100
OPD-08	GTGTGCCCCA	4	4	100
OPD-10	GGTCTACACC	5	5	100
OPD-13	GGGGTGACGA	3	3	100
OPD-15	CATCCGTGCT	4	4	100
OPD-16	AGGGCGTAAG	3	3	100
OPD-18	GAGAGCCAAC	4	4	100
OPD-19	CTGGGGACTT	3	3	100
OPD-20	ACCCGGTCAC	4	4	100
OPE-01	CCCAAGGTCC	5	5	100
OPE-02	GGTGCGGGAA	5	5	100
OPE-03	CCAGATGAC	4	4	100
OPE-04	GTGACATGCC	3	3	100
OPE-05	TCAGGGAGGT	3	2	80
OPE-06	AAGACCCCTC	4	4	100

*Operon Technologies (<http://www.operon.com>).

RESULTS AND DISCUSSION

There were selected 18 RAPD and 15 ISSR primers that showed the better quality in the patterns of bands and genetic repeatability. Together, these 33 primers generated a total of 137 bands. The amplification reactions carried out with RAPD primers produced a total of 73 bands with an average number of four bands per primer. The extreme values oscillated from three bands (with use of the primers OPD-13, OPD-16, OPD-19, OPE-04 and OPE-05) to six bands (with use of the primer OPD-05) (Table 1). In turn, the amplification reactions carried out with ISSR primers produced a total of 64 bands with an average number of 4.3 bands per primer. The extreme values oscillated from two bands (with use of the primer TriCAC 5CY) to seven bands (with use of the primer DiCA 3G) (Table 2).

Although studies of this nature are still incipient to the genus *Croton*, and inexistent for *C. heliotropiifolius*, similar results were observed by Angelo et al. (2006) in *C. cajucara*. In this study, the authors used 10 RAPD primers and observed 71 bands, of which 66 were polymorphic. To species of others genus the results observed with RAPD and ISSR markers are very variable. As examples are the genetic diversity studies conducted by Ansari et al. (2012) using five ISSR primers to characterize 29 genotypes of *Tectona grandis*, being obtained 43 polymorphic bands and by Souza et al. (2008) from five ISSR primers to characterize 269 genotypes in 12

populations of *Zabrotes subfasciatus*, being obtained a total of 51 polymorphic bands.

The average distance revealed for the wild genotypes of *C. heliotropiifolius* on base in the molecular markers was $dg_{ij} = 0.26$, and maximum and minimum distances were $dg_{ij} = 0.48$ (among the genotypes located at the top and in the middle region of the hill) and $dg_{ij} = 0.12$ (among the genotypes located at the middle region and in the valley of the hill), respectively. This diversity can be considered relatively high, in view that for different accesses of *C. cajucara* the genetic diversity observed ranged from 0.07 to 0.25 (Angelo et al., 2006).

The dendrogram based on neighbor joining method showed the formation of different groups among the genotypes located at the top of the hill and the genotypes located in the valley of the hill (Figure 1). Genotypes collected in the middle region of the hill were grouped both among the genotypes of the top and among the genotypes of the valley of the hill, being possible that middle region of the hill being an convergence point for the genetic flow from the extreme portions of the hill. This hypothesis is supported by greater genetic diversity observed in the middle region (dg_{ij} in middle = 0.28) in relation to diversity found at the extremes of the hill (dg_{ij} in top = 0.24 and dg_{ij} in valley = 0.22).

In the Bayesian estimates (based on Delta K values) was observed the existence of three genetic pools as the most suitable structure for distribution and grouping of the 41 wild genotypes of *C. heliotropiifolius* evaluated (Figure

Table 2. Primers used for obtaining ISSR markers with number total of bands (N° bands), number of polymorphic bands (N° PB) and percentage of polymorphism for 41 wild genotypes of *Croton heliotropifolius*.

Primer*	Sequence (5'-3')	N° bands	N° PB	Polymorphism (%)
DiCA 3G	CACACACACACACAG	7	7	100
DiCA 3RG	CACACACACACACARG	5	4	80
DiGA 3C	GAGAGAGAGAGAGAG	3	3	100
DiGA 3RC	GAGAGAGAGAGAGARC	4	4	100
DiGA 3T	GAGAGAGAGAGAGAT	4	3	80
TriCAC 3RC	CACCACCACCACCACRC	3	3	100
TriCAC 5CY	CACCACCACCACCACCY	2	2	100
TriCAG 3RC	CAGCAGCAGCAGCAGRC	4	3	80
TriTGT 3YC	TGTTGTTGTTGTTGTYC	5	5	100
TriAAR 3RC	AACAACAACAACAACRC	4	4	100
TriAAG 3RC	AAGAAGAAGAAGAAGRC	5	5	100
TriACG 3RC	ACGACGACGACGACRC	4	3	80
TriAGA 3RC	AGAAGAAGAAGAAGARC	5	5	100
TriCGC 3RC	CGCCGCCGCCGCCGRC	5	5	100
TriGAC 3RC	GACGACGACGACGACRC	4	4	100

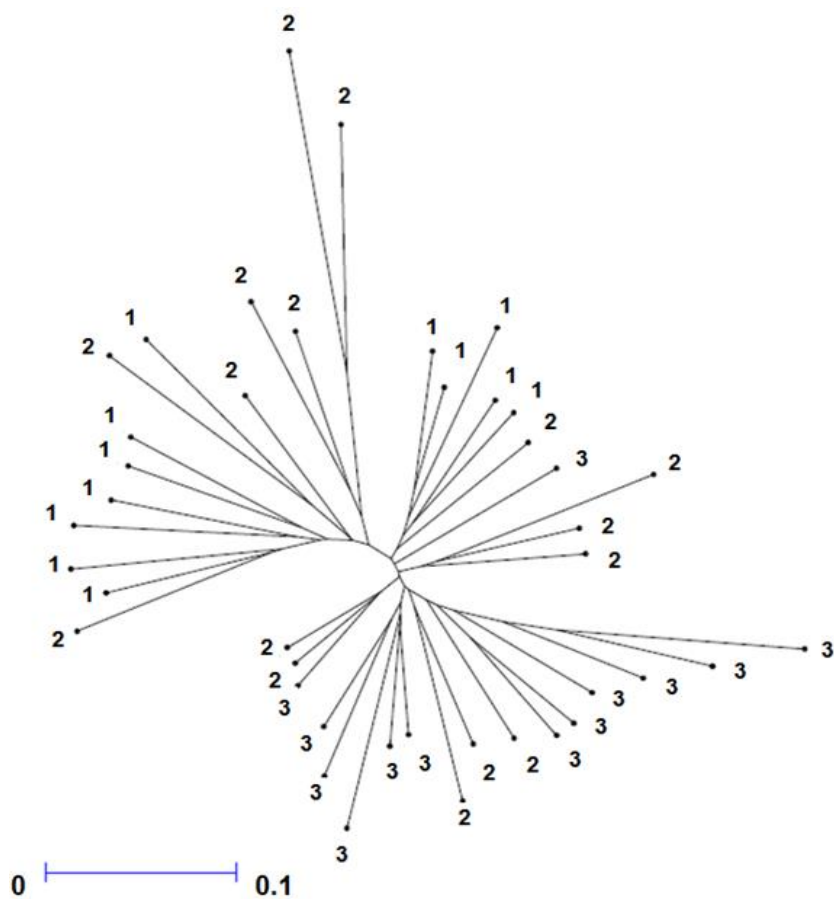


Figure 1. The neighbor-joining tree constructed using the coefficient of Dice (1945), from the genotyping data carried out with RAPD and ISSR markers in a population of 41 wild genotypes of *Croton heliotropifolius*. The numbers indicate the collection region of each genotype along hill (1 = top, 2 = middle, and 3 = valley of the hill).

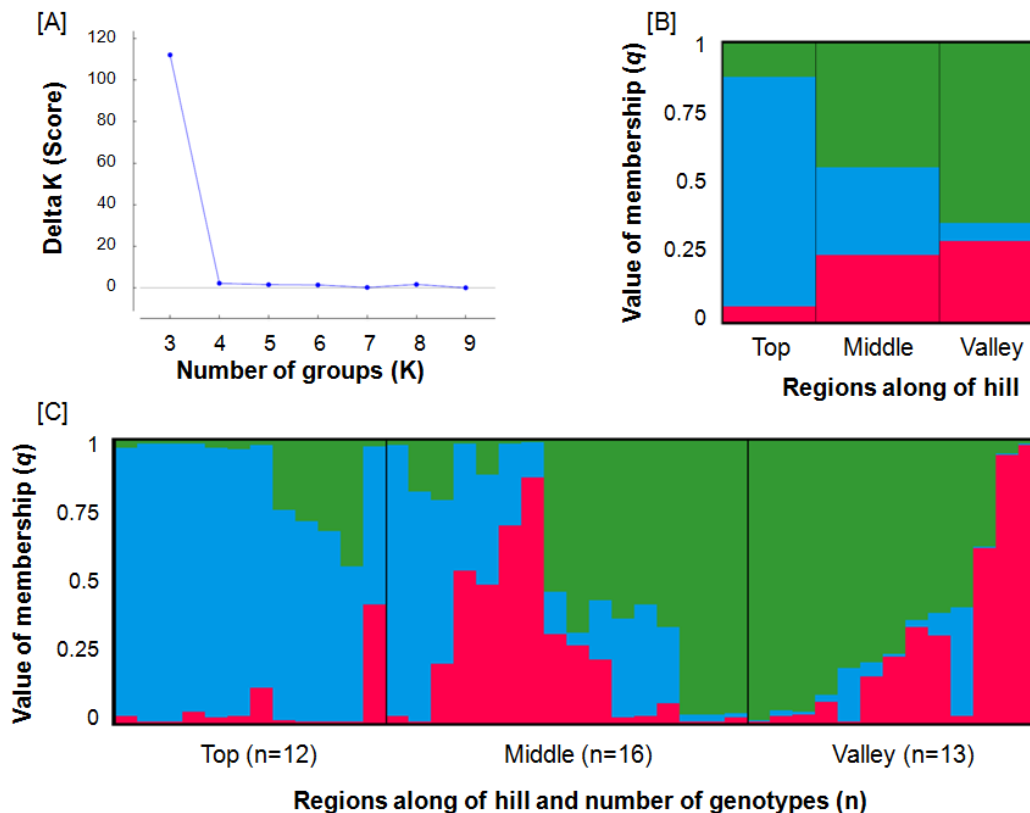


Figure 2. Numbers of genetic pool (clusters) inferred based on Bayesian analyses considering the most probable number of groups (K) estimated with the method described by Evanno et al. (2006) [A], well as histograms of the distribution of gene pools in the three regions of the hill [B] and in each of wild genotypes of *Croton heliotropiifolius* investigated [C]. Each column (histogram) represents the genotyping data (consensus) from each region of the hill [B] or from each wild genotype [C], and the colors used in the histograms represent the most likely ancestry of the cluster from which the accessions were derived.

2A). The histogram (Figure 2B and C) and dendrogram (Figure 1) indicate the existence of genetic structure among genotypes of *C. heliotropiifolius* along of the hill (*Serra da Torre*), corroborate with the hypothesis of occurrence of three genetic pools. At the top, and in the valley of the hill, the gene pools represented by the blue and green colors (respectively) are more predominant than the third gene pool represented by the red color (Figure 2B and C). In turn, in the middle region of the hill was observed a great mixture of the three gene pools, being evidenced in this region.

From the analysis, it can be inferred the occurrence of a barrier that prevent cross between the individuals of the top with those of the lowland region. However, a gene flow between individuals of the extremes with those who are in middle region seems to occur, which may be caused by pollinators and seed dispersors. Therefore, it is important to understand the pollination and seed dispersal mechanisms. In this context, in a study carried out by Leal et al. (2003) it was observed that for most species of the family Euphorbiaceae occurring in Caatinga the seed dispersal is carried out by ants. In this

study, it was found that the species *C. campestris* St-Hil, *C. argyophyllus* Kunth., and *C. blanchetianus* Baill., have diasporas with a structure rich in lipidic compounds that have an important role to attract ants. Also, Passos and Ferreira (1996) studied the seed dispersal in *C. priscus* Croizad and found about 11 different species of ants interacting with the seeds of the species.

Study with *C. sellowi*, another caatinga shrub, indicated a total of 19 species of floral visitors, represented mostly by insects, predominantly bees (Pimentel and Castro, 2009). In this context, *C. heliotropiifolius* is important in the diet of bees, as this species visited the flowers very often (Silva et al., 2014). Similar results were found in a study by Dominguez and Bullock (1989) for *C. suberosus*, where it was observed that the insects represent the main species of floral visitors.

Additionally, *C. heliotropiifolius* has a structure called elaiosome in its seeds that is rich in lipid and acts as an attractive for ants (Leal et al., 2003). Thus, once the plant seeds fall, ants feed on the elaiosome and the seed became exposed, sometimes close to the mother plant, which can contribute to a spatial structure (Bertagna,

2007). For *Mabea fistulifera* that also has elaiosome, ants of the genera *Atta* and *Acromyrmex*, carry the seeds for up to 4 m (Paternelli et al., 2004). This fact confirms the genetic structure observed in this study for the different regions of the hill (*Serra da Torre*).

Conclusion

The population of cassutinga (*C. heliotropiifolius*) considered in this study has its genetic diversity structure in three gene pools that are correlated with the spatial distribution of the species along of the sampled area (Itapetinga, Bahia, Brazil). Although further studies are needed to understand the factors responsible for this structure, it is likely that the dynamics of the seed carriage performed by the ants and the pollination by insects, especially the bees, are directly related.

Conflict of Interests

The authors have not declared any conflict of interests.

ACKNOWLEDGEMENTS

The authors are grateful to the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES (Brazilian Ministry of Education) for providing a fellowship to T.O.R. as part of her Masters in Environmental Sciences at Universidade Estadual do Sudoeste da Bahia (UESB).

REFERENCES

- Abreu AS, Barbosa PS, Muller AH, Guilhon GMSP (2001). Constituintes químicos do caule e das cascas do caule de *Croton pullei* var. *Glabor* (Euphorbiaceae). Revista Virtual de Iniciação Acadêmica da UFPA 1:2.
- Angelo PCS, Chaves FCM, Bizzo HR, Xavier JJB, Cruz JC, Lira MPS (2006). Genetic diversity in sacaca (*Croton cajucara* Benth.) accessed by RAPD markers. Rev. Bras. Pl. Med. 8:18-22.
- Ansari SA, Narayanan C, Wali SA, Kumar R, Shukla N, Rahangdale SK (2012). ISSR markers for analysis of molecular diversity and genetic structure of Indian teak (*Tectona grandis* L.f) populations. Ann. For. Res. 55:11-23.
- Bertagna M (2007). Variabilidade genética de *Croton floribundus* Spreng. (Euphorbiaceae) em seis fragmentos de Mata Atlântica do município de Campinas. Master's thesis, Universidade Estadual de Campinas.
- Brasil Ministério Das Minas E Energia/Secretária Geral (1981). Projeto RADAMBRASIL, Salvador: geologia, geomorfologia, pedologia, vegetação e uso potencial da terra. Rio de Janeiro: *Ministério de Minas e Energia/Secretária Geral* (Levantamento de Recursos Naturais).
- Cerqueira-Silva CBM, Nunes OJ, Santos ESL, Corrêa RX, Souza AP (2014). Genetic Breeding and Diversity of the Genus *Passiflora*: Progress and Perspectives in Molecular and Genetic Studies. Int. J. Mol. Sci. 15:14122-14152.
- Cruz CD (2006). Programa Genes: Análise multivariada e simulação. Editora da Universidade Federal de Viçosa.
- Dominguez CA, Bullock SH (1989). La reproduction de *Croton suberosus* (Euphorbiaceae) en luz y sombra. Rev. Biol. Trop. 37:1-10.
- Doyle JJ, Doyle JM (1990). Isolation of plant DNA from fresh tissue. Focus 12:11-15.
- Dice LR (1945). Measures of the amount of ecological association between species. Ecology 26:297-302.
- Earl DA, Vonholdt BM (2012). STRUCTURE HARVESTER: A website and program for visualizing structure output and implementing the Evanno method. Conserv. Genet. Resour. 4: 359–361.
- Evano G, Regnaut S, Goudet J (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol. Ecol. 14: 2611-2620.
- Falush D, Stephens M, Pritchard JK (2003). Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. Genetics 164:1567-1587.
- Frankham R, Ballou JD, Briscoe DA (2008). Fundamentos de genética da conservação. Sociedade Brasileira de Genética.
- Futuyma, DJ (2002). Evolução, ciência e sociedade. Sociedade Brasileira de Genética.
- Jakobsson M, Rosenberg NA (2007). CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics 23: 1801-1806.
- Leal RI, Tabarelli M, Silva JMC (2003). Ecologia e conservação da Caatinga. Ed. Universitária da UFPE 822.
- Passos L, Ferreira SO (1996). Ant dispersal of *Croton priscus* (Euphorbiaceae) seeds in a tropical semideciduous forest in southeastern Brazil. Biotropica 28:697-700.
- Perrier X, Jacquemoud-Collet JP, DARwin Software (2006). Available online: <http://darwin.cirad.fr/>.
- Paternelli EFO, Lucia TMCD, Martins SV (2004). Espécies de formigas que interagem com as sementes de *Mabea fistulifera* Mart.(Euphorbiaceae). Sociedade de Investigações Florestais. Rev. Árvore 28:733-738.
- Pimentel KGM, Castro CC (2009). Biologia reprodutiva de *Croton sellowii* Baill. (Euphorbiaceae) em uma restinga de Pernambuco. *Jepex*.
- Pritchard JK, Stephens M, Donnelly P (2000). Inference of population structure using multilocus genotype data. Genetics 155:945-959.
- Randau KP (2001). Estudo farmacognóstico (farmacobotânico e farmacológico) e atividade biológica de *Croton rhamnifolius* H.B.K e *Croton rhamnifolioides* Pax e Hoffm. (Euphorbiaceae). Master's thesis. Universidade Federal de Pernambuco. Recife-PE.
- Randau KP, Florêncio DC, Ferreira CP, Xavier HS (2004). Estudo farmacognóstico de *Croton rhamnifolius* H.B.K e *Croton rhamnifolioides* Pax e Hoffm. (Euphorbiaceae). Rev. Bras. Farmacogn. 14:89-96.
- Rosenberg NA (2004). Distruct: A program for the graphical display of population structure. Mol. Ecol. Notes 4:137-138.
- Scalaferrri MM, Freitas JS, Santos ESL, Vieira JGP, Gonçalves ZS, Cerqueira-Silva CBM (2013). Comparison of protocols for genomic DNA extraction from velame pimenta (*Croton linearifolius*), a native species to the Caatinga, Brazil. Afr. J. Biotechnol. 12:4761-4766.
- Scalaferrri MM, Freitas JS, Vieira JGP, Gonçalves ZS, Souza AM, Cerqueira-Silva CBM (2014). Comparison of methods for estimates of molecular genetic diversity in genus *Croton*: influence of coefficients, clustering strategies and data projection. Genet. Mol. Res. 13:5566-5573.
- Silva GAR, Bastos EM, Sobreira JAR (2014). Levantamento da flora apícola em duas áreas produtoras de mel no estado do Piauí. Enciclopédia Biosfera 10:3305-3316.
- Souza GA, Carvalho MRO, Martins ER, Guedes RNC, Oliveira LO (2008). Diversidade genética estimada com marcadores ISSR em populações brasileiras de *Zabrotes subfasciatus*. Pesq. Agropec. Bras. 43:843-849.
- Zioldo BD (2007). Estrutura genética em populações de *Croton floribundus* spreng. de diferentes fragmentos de mata. Master's thesis. Universidade Estadual de Londrina.