AJ ACADEMIC JOURNALS
expand your knowledge

African Journal of Biotechnology

Full Length Research Paper

Diversity and genetic structure in natural populations of *Croton linearifolius* (Euphorbiaceae) based on molecular markers

Thalana Souza Santos Silva[1], Janaína Silva Freitas[1,2], Elisa Susilene Lisboa dos Santos[1,2] and Carlos Bernard Moreno Cerqueira-Silva[1,2*]

[1]Postgraduate Program in Environmental Sciences, Universidade Estadual do Sudoeste da Bahia, Itapetinga, BA, 45700-000, Brazil.
[2]Department of Exact and Natural Sciences, Universidade Estadual do Sudoeste da Bahia, Itapetinga, BA, 45700-000, Brazil.

Croton linearifolius Mull. Arg, an endemic species of Brazil, has insecticidal activity proven. To the detriment of its importance as a natural resource, the studies of this species are incipient, as well as the strategies applied to its management and conservation. The diversity and genetic structure in 61 individuals of *C. linearifolius* collected in the National Forest Contendas do Sincorá (NFCS) were estimated. Estimates were based on analysis of the amplification profile of nine combinations of pairs of resistance gene analogs (RGA) primers and eight inter simple sequence repeat (ISSR) primers. A total of 134 markers (81.3% polymorphic) were generated. Bayesian analysis indicated as most likely a structuring into two groups. Based on the molecular analysis of variance (AMOVA) it was possible to verify that 64% (p rand <0.01) of the variation occurs within the regions, and a significant amount (36%) (p rand <0.01) was attributed to variations between regions, indicating genetic structure between them. The AMOVA results were corroborated with the Principal Coordinate Analysis (PCoA), indicating an association between the distribution of variability and the geographical distribution of the collection regions. The probable mechanisms pollination and dispersion would justify, at least in part, the genetic structuring observed for *C. linearifolius*.

Key words: Conservation, genetic variability, genetic polymorphism, resistance gene analogs (RGA) marker, inter simple sequence repeat (ISSR) marker.

INTRODUCTION

The genus *Croton* L. (Euprobiaceae), with about 1,200 species, stands out the second largest in the Euphorbiaceae family, with most of its species dispersed in the tropical regions of the world (Govaerts et al., 2000). Brazil is considered one of the main centers of diversity of the genus *Croton*, with at least 350 species distributed

*Corresponding author. E-mail: csilva@uesb.edu.br.

Figure 1. General appearance of the species *Croton linearifolius* (A) male flower, fruit (B) and seed (C).
Source: Personal collection.

in different regions of the country (Berry et al., 2005). The highest concentration of species of this genus occurs in the Southeast and Northeast regions, with approximately 160 and 120 species, respectively (Cordeiro et al., 2015).

Species of the genus *Croton* are an important natural resource in the Northeast region of Brazil, being used for several purposes. Among the species listed popularly by having biological activity, some have tests that prove its medicinal potential, such as *Croton regelianus* (Antitumoral activity) (Bezerra et al., 2009), *Croton zehntneri* (Antimalarial activity) (Mota et al., 2012), *Croton campestris* (Gastric antiulcerogenic activity) (Júnior et al., 2014), *Croton cordiifolius* (Antinociceptive activity) (Nogueira et al., 2015), and/or insecticide, such as *Croton argyrophylloides, Croton nepetaefolius, Croton sonderianus* and *C. zehntneri* (Lima et al., 2006).

This study highlights the species *Croton linearifolius*, popularly known as "velame pimenta", with proven insecticidal potential against *Aedes aegypti* (Silva et al., 2014) and *Cochliomyia macellaria* (Silva et al., 2010),

which corroborates the traditional use indicated in the semi-arid region of Brazil (Silva et al., 2010, 2014). This species also presents diverse chemical composition, being rich in alkaloids, steroids and flavonoids, among other potentially promising compounds (Silva et al., 2010).

C. linearifolius is endemic in Brazil, with occurrence records in the states of Minas Gerais, Piauí, Tocantins and greater representation in Bahia (Cordeiro et al., 2015) (Figure 1), it has shrubby habit and can measure up to 2 m high. The plants of this species are monoecious, with lanceolate leaves and discolores, bisexual inflorescences, the fruits are dehiscent, the seeds measure about 4.0 mm long and 2.5 mm wide, with brown coloration (Lima and Pirani, 2008).

The exploitation of native species by the population is generally based on exploratory extraction, which can lead to a reduction of biodiversity (Mondini et al., 2009). Among the different components of biodiversity, genetic diversity is the basis for the evolutionary potential of a
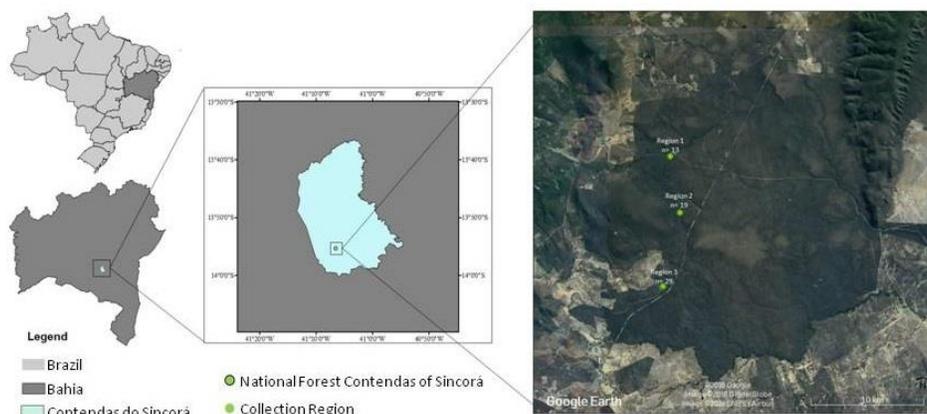
Figure 2. Geographical areas map (Brazil, Bahia State, Contendas do Sincorá city) were from collected samples of *Croton linearifolius.*

species and determines its chances of survival, reproduction and adaptation to possible environmental changes (Fleishman et al., 2001). Therefore, knowledge of the diversity and genetic structure of species that present potential as a natural resource is a prerogative for the elaboration of strategies for sustainable management and conservation of genetic resources. In this context, genetic-molecular markers are an important tool for population studies and have been used in the study of a wide range of organisms (Ibrahim et al., 2010).

Resistance gene analogs (RGA) and inter simple sequence repeat (ISSR) markers are examples of markers that do not require prior information on the genome of the species to be evaluated and enable a rapid and low cost characterization, especially useful for species not yet studied and with commercial interest still little explored.

In detriment of the ecological importance and economic potential of *Croton* species, population genetic studies are limited to a few species (approximately 1%), which hampers the development of management and conservation strategies that may help in the conservation and sustainable use of this biodiversity. Specifically for *C. linearifolius* the genetic-molecular studies are limited to the works of Scaldaferri et al. (2013, 2014) and Silva et al. (2018).

In this study, the genetic diversity and structure of natural populations of *C. linearifolius* were characterized and discussed, and present in the National Forest Contendas do Sincorá (NFCS), based on the application of RGA and ISSR molecular markers.

MATERIALS AND METHODS

Characterization of the biological material

The study was conducted from DNA samples from 61 individuals of *C. linearifolius* collected in three regions understood here as populations, present in the National Forest Contendas do Sincorá (NFCS), in the municipality of Contendas do Sincorá, Bahia, Brazil (Figure 2 and Supplemental Material 1). The exsicatas were deposited in the herbarium of the Universidade Estadual de Feira de Santana, under registration HUEFS 146620. The DNA samples were previously extracted from fresh leaves using the cetyltrimethylammonium bromide (CTAB) method (Doyle and Doyle, 1990) with modifications previously tested for *Croton* by Scaldaferri et al. (2013). The samples were deposited in the genomic DNA bank of the Laboratory of Applied Molecular Genetics (LAMG), at the Universidade Estadual do Sudoeste da Bahia (UESB), Itapetinga, Bahia, Brazil.

The integrity of the DNA samples was evaluated in 1% agarose gel (m/v) by electrophoresis (2 h in 90 V electric current) and visualized from loading dyes (bromophenol blue and xylene cyanol) and GelRed (Invitrogen Co., Carlsbad, CA, USA) according to manufacturer's specifications, in Kodak photodocumentation system, with incidence of ultraviolet (UV) illumination. To estimate the DNA concentration (ng/μL), the lambda molecular weight marker (Invitrogen Lambda DNA) was adopted as standard.

Molecular genotyping

Genotyping was performed based on 10 ISSR primers (Table 1) and 10 combinations of RGA primers pairs (Table 2) previously selected (Silva et al., 2018). PCR reactions were performed with a final volume of 16 μL containing 12 ng of DNA, 1.7 μL of 10X PCR buffer (LGC Biotechnology, São Paulo, Brazil), 1.0 μL of 50 mM MgCl$_2$ (LGC Biotechnology, São Paulo, Brazil), 1 μL of 2.5 mM dNTP mix (LGC biotechnology), 1 unit of Taq DNA polymerase (LGC Biotecnologia, São Paulo, Brazil), Milli-Q water and 1 μL primer.

In the amplification reactions, the following amplification programs were adopted: for primers RGAs; 5 min at 95°C; followed by 34 cycles (30 s at 95°C, 1 min at 37°C, 1 min and 20 s at 72°C); and 10 min at 72°C. For ISSR primers: 95°C for 5 min; followed by 34 cycles (94°C for 50 s, 48°C for 50 s, 72°C for 1 min); and 5 min at 72°C. All amplifications were performed in the LGMA of the UESB.

Aliquots (5 μL) of the amplification products were visualized from electrophoretic run, on 2% agarose gel (m/v) and 1x TBE solution, for approximately 2 h at 110 V. The visualization of the result of the electrophoretic run was performed with a Kodak photodocumentator, under the influence of ultraviolet light.

Table 1. Description of the 24 combinations of Resistance Gene Analogs (RGA) primer pairs (obtained from 23 primers) used in amplification tests in *Croton linearifolius.*

Combination	Code	Sequence	Reference
1	S1 AS1	GGT GGG GTT GGG AAG ACA ACG CAA CGC TAG TGG CAA TCC	
2	S1 AS2	GGT GGG GTT GGG AAG ACA ACG IAA IGC IAG IGG IAA ICC	
3	S1 AS3	GGT GGG GTT GGG AAG ACA ACG IAG IGC IAG IGG IAG ICC	Leister et al. (1996)
4	AS1 AS2	CAA CGC TAG TGG CAA TCC IAA IGC IAG IGG IAA ICC	
5	AS1 AS3	CAA CGC TAG TGG CAA TCC IAG IGC IAG IGG IAG ICC	
6	NBSF1 AS1	GGA ATG GGI GGI GIF GGI AAR AC CAA CGC TAG TGG CAA TCC	Yu et al. (1996) Leister et al. (1996)
7	NBSF1 AS2	GGA ATG GGI GGI GIF GGI AAR AC IAA IGC IAG IGG IAA ICC	Yu et al. (1996) Leister et al. (1996)
8	NBSF1 AS3	GGA ATG GGI GGI GIF GGI AAR AC IAG IGC IAG IGG IAG ICC	Yu et al. (1996) Leister et al. (1996)
9	S1 NBSR1	GGT GGG GTT GGG AAG ACA ACG YCT ACT TGT RAY DAT DAY YYT RC	Leister et al. (1996) Yu et al. (1996)
10	RGA1F RGA8R	AGT TTA TAA TTY EAT TGC T AGC CAC TTT TGA CAA CTG C	Kanazin et al. (1996) Kanazin et al. (1996)
11	NBSF1 NBSR1	GGA ATG GGI GGI GIF GGI AAR AC YCT ACT TGT RAY DAT DAY YYT RC	Yu et al. (1996) Yu et al. (1996)
12	RGA8F RGA8R	AGC GAC GAG AGT TGT ATT TAA G AGC CAC TTT TGA CAA CTG C	
13	RGA1R RGA1F	ACT ACG ATT CAA GAC GTC CT AGT TTA TAA TTY EAT TGC T	
14	RGA2R RGA1F	CAC ACG GTT TAA AAT TCT CA AGT TTA TAA TTY EAT TGC T	Kanazin et al. (1996)
15	RGA7R RGA1F	CCG AAG CAT AAG TTG GTG AGT TTA TAA TTY EAT TGC T	
16	RGA4R RGA4F	TAC ATC ATG TGT TAC CTC T TGT TAC TGC TTT GTT TGG TA	
17	RGA5R RGA5F	TCA ATC ATT TCT TTG CAC AA TGC TAG AAA AGT CTA TGA AG	

Table 1. Contd.

18	RGA6R	AAC TAC ATT TCT TGC AAG T
	RGA 6F	AGC CAA AGC CAT CTA CAG T
19	RGA1R	ACT ACG ATT CAA GAC GTC CT
	RGA 4F	TGT TAC TGC TTT GTT TGG TA
20	RGA4R	TAC ATC ATG TGT TAC CTC T
	RGA 1F	AGT TTA TAA TTY EAT TGC T
21	RGA2R	CAC ACG GTT TAA AAT TCT CA
	RGA 4F	TGT TAC TGC TTT GTT TGG TA
22	RGA2R	CAC ACG GTT TAA AAT TCT CA
	RGA 5F	TGC TAG AAA AGT CTA TGA AG
23	RGA2R	CAC ACG GTT TAA AAT TCT CA
	RGA 6F	AGC CAA AGC CAT CTA CAG T
24	RGA6R	AAC TAC ATT TCT TGC AAG T
	RGA 5F	TGC TAG AAA AGT CTA TGA AG

I= A/T/G/C; D= A/G/T; E=C/G; R= A/G; Y= C/T.

Table 2. Description of the 23 Inter Simple Sequence Repeat (ISSR) primers tested in *C. linearifolius.*

Primer Code	Sequence
DiCA3´G	5´- CAC ACA CAC ACA CAC AG- 3´
DiCA3`RG	5´- CAC ACA CAC ACA CAC ARG- 3´
DiCA3`YG	5´- CAC ACA CAC ACA CAC AYG-3`
DiGA3`C	5`- GAG AGA GAG AGA GAG AC- 3`
DiGA3` RC	5`- GAG AGA GAG AGA GAG ARC- 3`
DiGA3`T	5`- GAG AGA GAG AGA GAG AT- 3`
TriCAC3`RC	5`- CAC CAC CAC CAC CAC RC-3`
TriCAC3`YC	5`- CAC CAC CAC CAC CAC YC-3`
TriCAC5`CY	5`- CAC CAC CAC CAC CAC CY-3`
TriCAG3`RC	5`- CAC CAC CAC CAC CAC RC-3`
TriGTG3`YC	5`- GTG GTG GTG GTG GTG YC-3`
TriTGT3`YC	5`- TGT TGT TGT TGT TGT YC-3`
TriAAC3`RC	5`- AAC AAC AAC AAC AAC RC-3`
TriAAG3`RC	5`-AAG AAG AAG AAG AAG RC3`
TriACG3`RC	5`- ACG ACG ACG ACG ACG RC-3`
TriAGA3`RC	5`-AGA AGA AGA AGA AGA RC-3`
TriTGG3`RC	5`- TGG TGG TGG TGG TGG RC-3`
TriCGA3`RC	5`- CGA CGA CGA CGA CGA RC-3`
TriCGC3`RC	5`- CGC CGC CGC CGC CGC RC-3`
TriGAC3`RC	5`- GAC GAC GAC GAC GAC RC-3`
TriGCA3`RC	5`- GCA GCA GCA GCA GCA RC-3`
TriGCC3`RC	5`- GCC GCC GCC GCC GCC RC-3`
TriGGA3`RC	5`- GGA GGA GGA GGA GGA RC-3`

I= I=A/T/G/C; D= A/G/T; E=C/G; R= A/G; Y= C/T.

Table 3. Polymorphism obtained from nine combinations of Resistance Gene Analogs (RGA) primer pairs in the characterization of the 61 individuals of *Croton linearifolius.*

Combination	Code	Number of markers		Polymorphism (%)
		Total	Polymorphic	
2	S1+AS1	9	5	55.5
3	S1+AS3	5	2	40
6	NBSF1+AS1	6	1	16.6
14	RGA2R+ RGA1F	9	9	100
15	RGA7R+ RGA1F	4	4	100
17	RGA5R+ RGA5F	12	12	100
18	RGA6R+ RGA6F	3	3	100
19	RGA1R+ RGA4F	13	13	100
24	RGA6R+ RGA5F	8	8	100
Total	-	69	57	83%

Data analysis and genetic diversity estimation

The obtained band standards, with analysis of the gels by two evaluators, were considered for the construction of a binary data table, where zero (0) was assigned for the absence of bands and one (1) for presence. Based on the table, the percentage of polymorphism was calculated. In order to maintain only the markers that presented a percentage less than 20% of missing data (that is, data whose identification of presence or absence of bands cannot be determined with certainty).

Analyses of population structure were performed with the Bayesian method using the structure software, version 2.3.4 (Pritchard et al., 2000). Considering that the present study was conducted using natural populations, we used an admixture model with independent allele frequencies in each population. 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 20 simulations were performed to estimate each K. The ΔK ad hoc method described by Evanno et al. (2005) and implemented in the online tool Structure Harvester (Earl and vonHoldt, 2012) was used to estimate the most likely K in each set. After estimating the most likely K, the greedy algorithm implemented in CLUMPP v.1.1.1 (Jakobsson and Rosenberg, 2007) was used with a random input order and 1000 permutations to align the runs. Based on the posterior probability of membership (q) of a given accession belonging to a given group compared to the total number of groups (K), individuals were classified with q > 0.60 as a member of a given cluster, whereas for clusters with membership (q) values ≤ 0.60, the accession was classified as admixed.

Estimation of genetic similarity was performed based on the Jaccard coefficient, with the aid of the GENES program (Cruz, 2006). The genotype grouping was performed by the Neighbor Joining method, through the DARwin program (Perrier and Jacquemoud-Collet, 2009). Principal coordinate analysis (PCoA) and Molecular Variance (AMOVA), with 999 permutations were performed using the GenALEX v.6.5 program (Peakall and Smouse, 2012) in order to observe the genetic variation between and within the regions collection.

RESULTS AND DISCUSSION

A total of 134 markers were generated after the amplification reactions in 61 individuals of *C. linearifolius.*

After sorting the data, nine and eight RGA and ISSR primers were maintained in the analyses, respectively. The combinations of selected RGA primers produced a total of 69 markers (82.6% polymorphic) with an average number of 7.6 markers per primer pair. The combinations 18 (RGA6R + RGA6F) and 19 (RGA1R + RGA4F) generated the lowest (three) and the highest (13) number of brands, respectively (Table 3). The amplification reactions with the ISSR primers produced a total of 65 markers (80% polymorphic), with an average number of 8.1 marks per primer. The primers DiCA3`RG and TriAAC3`RC produced the lowest (6) and highest (10) number of marks, respectively (Table 4).

Although there are no studies with RGA markers for the genus *Croton*, results similar to the polymorphism observed with such markers for *C. linearifolius* were observed in different plant species, such as sugarcane (86.5%) (Jayashree et al., 2010) and rice (96%) (Ren et al., 2013). These results reinforce the potential of RGA markers to access genetic polymorphism. The polymorphism results obtained with the use of ISSR primers in *C. linearifolius* were similar to the results presented for other species of the genus *Croton*, such as *Croton tetradenius* (94.8%) (Almeida-Pereira et al., 2017) and *Croton heliotropiifolius* (94%) (Rocha et al., 2016).

Considering Bayesian analysis (based on Delta K values), the main structuring was observed in two gene pools (Delta K score = 1400) and possible sub-structuring of three or four gene pools (Delta K-400 and K ~200, respectively), as the most probable compositions for the 61 individuals of *C. linearifolius* (Figure 3). The histogram showing the level of structuring in two gene pools (Figure 4A) reveals a predominance of the gene pool represented by blue color in regions 1 (q = 0.78) and 2 (points 2, 3, 4 and 5) (Mean of q = 0.73), whereas in region 3 (points 6, 7, 8, 9 and 10) there is a predominance of the gene pool represented by green color (mean of q = 0.89). Considering a sub-structuring in three gene pools (Figure 4B), although there are no

Table 4. Polymorphism obtained from eight Inter Simple Sequence Repeat (ISSR) primer in the characterization of the 61 individuals of *Croton linearifolius.*

Primer code	Number of markers		Total polymorphism (%)
	Total	Polymorphic	
DiCA3`RG	6	4	66.6
DiGA3` RC	9	7	77.7
TriCAC3`RC	9	9	100
TriCAC3`YC	7	4	57.1
TriGTG3`YC	8	5	25
TriTGT3`YC	9	9	100
TriAAC3`RC	10	8	80
TriACG3`RC	7	6	85.7
Total	65	52	80
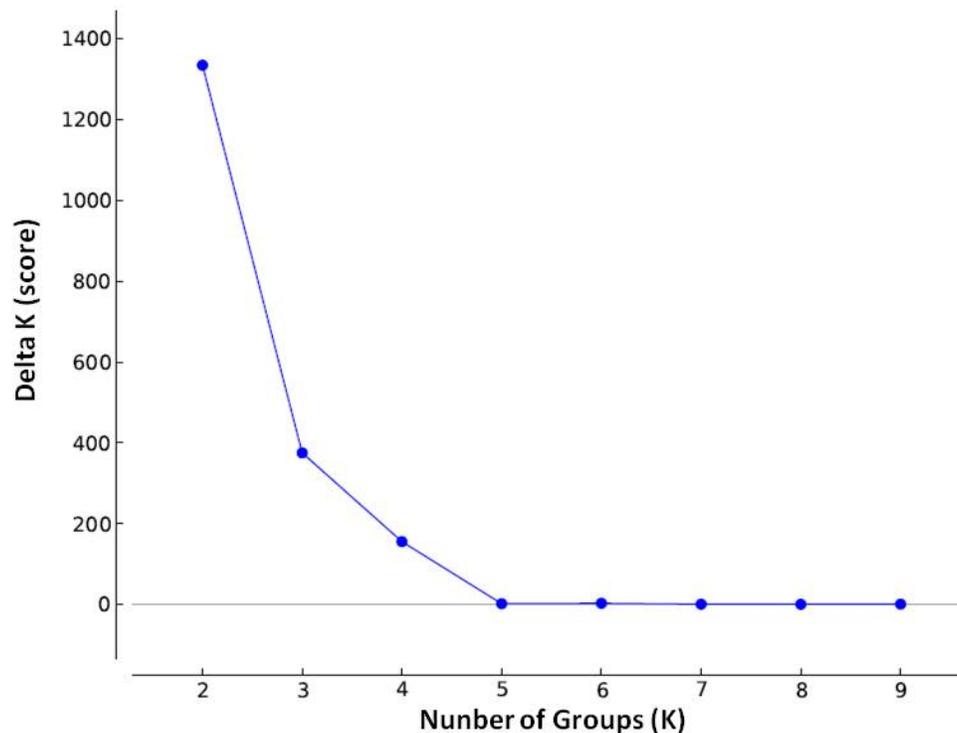
Figure 3. Numbers of genetic pool (clusters) inferred based on Bayesian analyses considering the most probable number of groups (K) estimated with the method described by Evanno et al. (2005).

significant changes in the composition of regions 1 (q = 0.82) and 3 (mean of q = 0.89), the change in composition of region 2 that presents the predominant presence of a distinct gene pool represented by red color (mean of q = 0.62).

The structure observed with Bayesian analysis was corroborated by the results obtained from the molecular analysis of variance (AMOVA), where it was possible to verify that 64% (p rand <0.01) of the variation occurs within the regions, and a significant amount (36%) (p rand <0.01) was attributed to differences between regions, which indicates genetic structuring between them. The AMOVA decomposition presented in the evaluation matrix pair to pair among the regions (Table 5) indicates a greater genetic distance between regions 1 and 3 (45.8%) and lower between regions 2 and 3 (29.7%). A similar result was observed for the species *Croton antisyphiliticus* which according to Oliveira et al.,
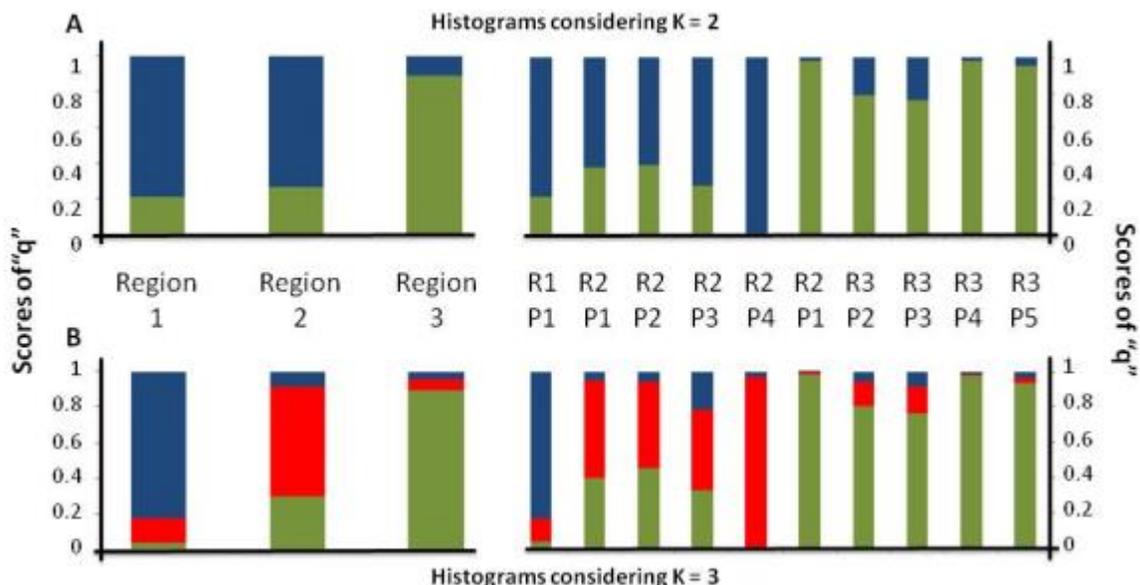
Figure 4. Histograms (based on Delta K values) representing the distribution of probable gene pools in the three regions (R) (upper and lower left) and collection points (P) (upper and lower right) of *Croton linearifolius*. Principal structure in two gene pools (Delta K score = 1400) (A) and sub-structuring in three gene pools (Delta K ~ 400 scores) (B). The colors used in the histograms represent the most probable ancestry of the group from which the individuals were derived.

Table 5. Matrix representative of the values of Φ (lower diagonal) and probability of occurrence (upper diagonal) estimated from the Molecular Analysis of Variance among the different collection regions of *Croton linearifolius*.

Correlation	Region 1	Region 2	Region 3
Region 1		0.001	0.001
Region 2	0.335		0.001
Region 3	0.458	0.297	

(2016) presented approximately 40% variation among the studied localities.

The genetic structure among plant populations is influenced, among other factors, by the effects caused by the dispersion of pollen and seeds, which are directly related to the connectivity among the populations, determining the gene flow rates (Zanella et al., 2012). Pollinators and seed dispersers that have effective dispersion over long distances decrease the likelihood of differentiation between populations, while dispersion at restricted distances has the opposite effect, promoting population genetic structuring (Loveless and Hamrick, 1984).

Although there are no records in the literature of studies on the reproductive biology of *C. linearifolius*, studies carried out with species of the same genus show that there is predominance of pollination by the wind and/or by different insects [*Croton suberosus* (Domínguez and Bullock, 1989), *Croton floribundus* and

Croton priscus (Passos, 1995), *Croton sarcopetalus* (Freitas et al., 2001), and *Croton urucurana* (Pires et al., 2004)]. Considering that insect-pollinated species tend to have lower gene flow rates and consequently greater genetic differentiation, since the distances traveled are relatively small (Loveless and Hamrick, 1984) it is expected that genus *Croton* will occur between populations.

The primary autochoric dispersion due to fruit dehiscence and secondary zoocoric dispersion due to associations with several ant species is a common strategy among species of the genus *Croton*. The presence of the elaiosomes is an attraction for the ants that act as secondary dispersers, thus reducing predation and competition between the seedlings (Webster, 1994; Passos and Ferreira, 1996; Leal et al., 2007; Lobo et al., 2011). The autochoric species tend to exhibit spatially limited seed dispersion, at maximum distances ranging from 3.4 (Passos and Ferreira, 1996) to 8.0 m (Narbona
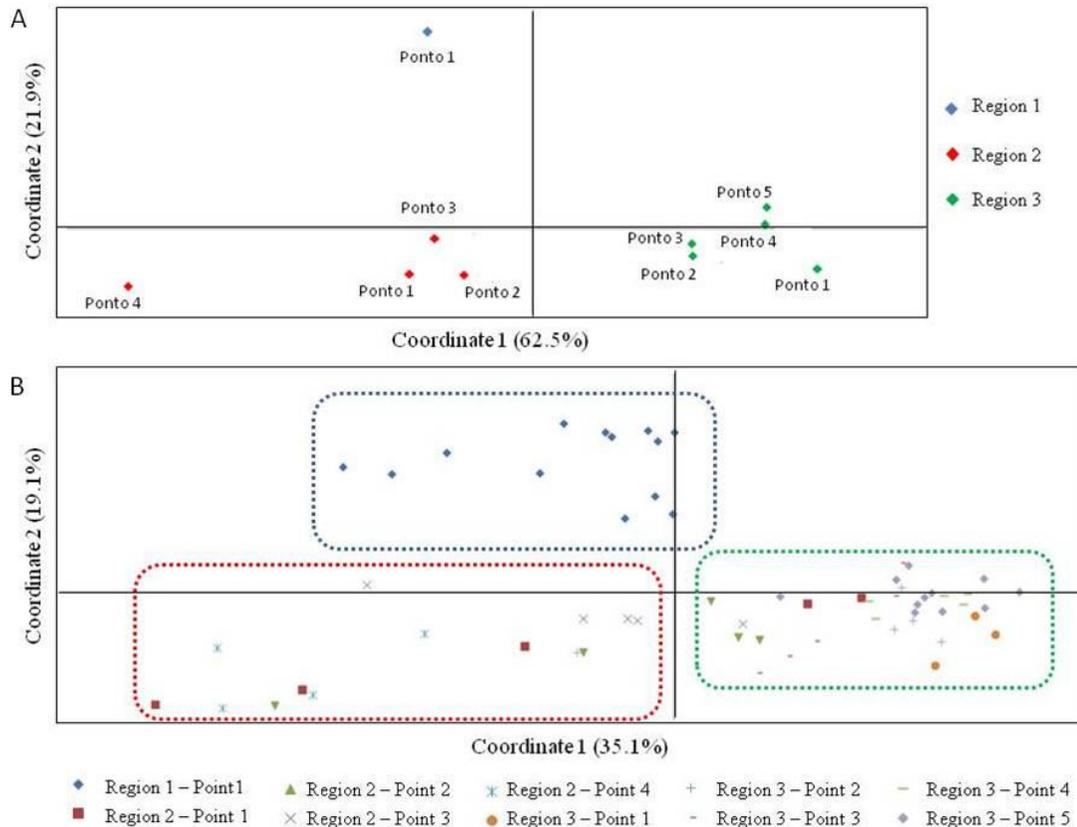
Figure 5. Dispersion graphic based on the main coordinate analysis (PCoA), represented by collection regions (A) and 61 individuals of *Croton linearifolius* (B). The colors of dashed lines correspond to those used to represent the predominant gene pools in the histograms obtained in structure (Figure 3).

et al., 2005). In addition, ants can travel up to 2.5 m away (Passos and Ferreira, 1996), which can also result in greater genetic differentiation between populations.

Therefore, the pollen and dispersion strategies observed for the genus and probable for *C. linearifolius* may be related to the low inferred gene flow between regions 1 and 3, based on STRUCTURE and AMOVA results, as well as the lower structure between these and region 2, which is located in the intermediate region. Similar structure, with existence of three gene pools and presence of greater differentiation between extreme collection points, was observed by Rocha et al. (2016) for *C. heliotropiifolius* from estimates made with ISSR and RAPD markers.

In addition to the mentioned factors, *C. linearifolius* is a monoic species, characterized by inflorescences with staminate (male) flowers in the terminal part and pistil (female) flowers in the lower part (Lima and Pirani, 2008). This arrangement of male and female flowers on an inflorescence may facilitate the occurrence of self-fertilization, a fact that has already been demonstrated in *C. floribundus* and *C. priscus* (Passos, 1995); *C. suberosus* (Domínguez and Bullock, 1989) and *C. sarcopetalus* (Freitas et al., 2001) that present flowers of similar morphology to *C. linearifolius*, which contributes to the increase of inbreeding.

The results obtained through the PCoA analysis (Figure 5) corroborate the hypothesis of sub-structuring in three gene pools obtained by Bayesian analysis (Figure 4B), corresponding to the geographical distribution of the three collection regions. Figure 5B shows the groupings of individuals delimited by rectangles of color corresponding to those used to represent the gene pools in the histogram. Several individuals collected in region 2 presented high q values for the characteristic gene pool of this region, represented by the red color in the histogram ($q \leq 0.60$) (Figure 4B and Supplementary Material 1). These results are in agreement with the collection regions and corroborate with the dispersion of individuals from that region on the PCoA plot (Figure 5B).

Based on the PCoA plot (Figure 5B) it is possible to observe the occurrence of individuals characteristic of region 2 for the gene pool characteristic of the individuals of region 3 (Supplemental Material 1), a fact that was corroborated by the molecular analysis of variance presented in the matrix of pair to pair evaluation among regions (Table 5), which indicated genetic structuring between individuals in regions 2 and 3 (approximately

30%). The observed result may be a reflection of the connectivity between the regions having the region 2 as intermediate, due to its geographical location between points 1 and 3.

In spite of the complementarity observed between the Bayesian analysis, PCoA Analysis and AMOVA results, there was no correspondence with the distance data projection (both by the Neighbor Joining grouping method and by the UPGMA) estimated by the Jaccard coefficient (data not presented). Considering that the choice of coefficients and clustering methods can directly influence the results of the projections (both in scatter plots and in dendrograms) (Scaldaferri et al., 2014; Cerqueira-Silva et al., 2009), it is probable that this absence of association is associated with the influence of the statistical method intrinsic to the procedures of distance estimation and/or clustering method.

Conclusion

The RGA and ISSR molecular markers were efficient in accessing polymorphic markers in *C. linearifolius*. The populations of *C. linearifolius* are structured in at least two gene pools and sub-structured in three and four gene pools, with a correspondence between the genetic structuring and the geographical distribution of the populations. The probable mechanisms of insect pollination and autochoric primary dispersion with secondary dispersion by ants would justify the genetic structuring observed for *C. linearifolius*, since these strategies would not be enough to promote the genetic flow among the collection regions. The results obtained contribute to the genetic-molecular knowledge of *C. linearifolius*, which may help in the strategies applied to the management and conservation of the species.

CONFLICT OF INTERESTS

The authors have not declared any conflict of interests.

ACKNOWLEDGEMENTS

REFERENCES

Almeida-Pereira CS, Silva AVC, Alves RP, Feitosa-Alcantara RB, Arrigoni-Blank MF, Alvares-Carvalho SV, Costa TS, White LAS, Pinto VS, Sampaio TS, Blank AF (2017). Genetic diversity of native populations of *Croton tetradenius* Baill. using ISSR markers. Genetic Molecular Research 16(2):2-12.

Berry PE, Hipp AL, Wurdack KJ, Ee BV, Riina R (2005). Molecular phylogenetics of the giant genus *Croton* and tribe Crotoneae (Euphorbiaceae sensu stricto) using ITS and trnL–trnF DNA sequence data. American Journal of Botany 92(9):1520-1534.

Bezerra DP, Marinho Filho JD, Alves AP, Pessoa C, De Moraes MO, Pessoa OD, Torres MC, Silveira ER, Viana FA, Costa-Lotufo LV (2009). Antitumor activity of the essential oil from the leaves of *Croton regelianus* and its component ascaridole. Chemistry and Biodiversity 6(8):1224-31.

Cerqueira-Silva CBM, Cardoso-Silva CB, Conceicao LD, Nonato JV, Oliveira AC, Corrêa RX (2009). Comparison of coefficients and distance measurements in passion fruit plants based on molecular markers and physicochemical descriptors. Genetics and Molecular Research 8(3):870-879. http://dx.doi.org/10.4238/vol8-3gmr614

Cordeiro I, Secco R, Carneiro-Torres DS, Lima LR DE, Caruzo MBR, Berry P, Riina R, Silva OLM, Silva MJDA, Sodré RC (2015). *Croton*. In: Lista de Espécies da Flora do Brasil. Jardim Botânico do Rio de Janeiro. Available in: [http://reflora.jbrj.gov.br/jabot/floradobrasil/FB17497]. Accessed 30 March 2015.

Cruz CD (2006). Programa Genes: Análise multivariada e simulação. Editora da Universidade Federal de Viçosa.

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

Domínguez CA, Bullock SH (1989). La reproducción de *Croton suberosus* (Euphorbiaceae) en luz y sombra. Revista de Biologia Tropical 37(1):1-10.

Earl DA, vonHoldt BM (2012). STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources 4(2):359-361.

Evanno G, Regnaut S, Goudet J (2005). Detecting the number of clusters of individuals using the software structure: A simulation study. Molecular Ecology 14(8):2611-2620.

Fleishman E, Launer AE, Switky KR (2001). Rules and exceptions in conservation genetics: Genetic assessment of the endangered plant *Cordylanthus palmatus* and its implications for management planning. Biological Conservation 98(1):45-53.

Freitas L, Bernardello G, Galetto L, Paoli AAS (2001). Nectaries and reproductive biology of *Croton sarcopetalus* (Euphorbiaceae). Botanical Journal of the Linnean Society 136(3):267-277.

Govaerts R, Frodin DG, Radcliffe-Smith A (2000). A world checklist and bibliography of Euphorbiaceae (and Pandaceae). Kew, Royal Botanic Gardens 2:417-536.

Ibrahim AA, Mohammad A, Bakir Haseeb AK, Farhan AH, Homaidan AA, Ali HB, Sadoon M, Mohammad S (2010). A Brief Review of Molecular Techniques to Assess Plant Diversity. International Journal of Molecular Sciences 11(5):2079-2096.

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(14):1801-1806.

Jayashree J, Selvi A, Nair NV (2010). Characterization of resistance gene analog polymorphism in sugarcane cultivars with varying levels of red rot resistance. Electronic Journal of Plant Breeding 1(4):1191-1199.

Júnior FE, de Oliveira DR, Boligon AA, Athayde ML, Kamdem JP, Macedo GE, da silva GF, De Menezes IR, Costa JG, Coutinho HD, Kerntopf MR, Posser T (2014). Protective effects of *Croton campestris* A. St-Hill in different ulcer models in rodents: evidence for the involvement of nitric oxide and prostaglandins. Journal of Ethnopharmacology 153(2):469-477.

Leal IR, Wirth R, Tabarelli M (2007). Seed dispersal by ants in the semi-arid Caatinga of northeast Brazil. Annals of Botany 99(5):885-894.

Lobo D, Tabarelli M, Leal IR (2011). Relocation of *Croton sonderianus* (Euphorbiaceae) seeds by *Pheidole fallax* Mayr (Formicidae): a case of post-dispersal seed protection by ants? Neotropical Entomology 40(4):440-444.

Loveless MD, Hamrick JL (1984). Ecological determinants of genetic

structure in plant populations. Annual Review of Ecology, Evolution, and Systematics 15:65-95.

Lima LR, Pirani JR (2008). Revisão taxonômica de *Croton* sect. Lamprocroton (Müll. Arg.) Pax (Euphorbiaceae s.s.). Biota Neotropica 82:177-231.

Lima MGAde, Maia ICC, Sousa BDde, Morais SMde, Freitas SM (2006). Effect of stalk and leaf extracts from Euphorbiaceae species on *Aedes aegypti* (Diptera, Culicidae) larvae. Revista do Instituto de Medicina Tropical de São Paulo 48(4):211-214.

Mota ML, Lobo LT, Costa JM, Costa LS, Rocha HA, Rocha E Silva LF, Pohlit AM, Neto VF (2012). In vitro and in vivo antimalarial activity of essential oils and chemical components from three medicinal plants found in northeastern Brazil. Planta Medica 78(7):658-64.

Mondini L, Noorani A, Pagnotta MA (2009). Assessing plant genetic diversity by molecular tools. Diversity 1(1):19-35.

Narbona E, Arista M, Ortiz PL (2005). Explosive seed dispersal in two perennial Mediterranean Euphorbia species (Euphorbiaceae). American Journal of Botany 92(3):510-516.

Nogueira Lde M, da Silva MR, Dos Santos SM, de Albuquerque JF, Ferraz IC, DE Albuquerque TT, Mota CR, Araújo RM, Viana GS, Martins RD, Havt A, Ximene RM (2015). Antinociceptive Effect of the Essential Oil Obtained from the Leaves of *Croton cordiifolius* Baill. (Euphorbiaceae) in Mice. Evidence-Based Complementary and Alternative Medicine 2015:1-7.

Oliveira TG, Pereira MAS, Coppede JS, França SdeC, Ming LC, Bertoni BW (2016). Genetic diversity analysis of *Croton antisyphiliticus* Mart. using AFLP molecular markers. Genetics and Molecular Research 15(1):1-8.

Passos LC (1995). Fenologia, polinização e reprodução de duas especies de *Croton* (Euphorbiaceae) em mata semidecidua. Dissertação (Mestrado). Instituto de Biologia, Universidade Estadual de Campinas, Brasil.

Passos L, Ferreira SO (1996). Ant dispersal of *Croton priscus* (Euphorbiaceae) seeds in a tropical semideciduous forest in Southeastern Brazil. Biotropica 28(4):697-700.

Peakall R, Smouse PE (2012). GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teach-ing and research-an update. Bioinformatics 28(19):2537-2539.

Perrier X, Jacquemoud-Collet JP (2006). DARwin Software. Available online: http://darwin.cirad.fr/.

Pires MMY, Souza LA, Terada Y (2004). Biologia floral de *Croton urucurana* Baill. (Euphorbiaceae) ocorrente em vegetação ripária da ilha Porto Rico, Porto Rico, Estado do Paraná, Brasil. Acta Scientiarum Biological Sciences 26(2):209-215.

Pritchard JK, Stephens M, Donnelly P (2000). Inference of population structure using multilocus genotype data. Genetics 155(2):945-959.

Ren J, Yu Y, Gao F, Zeng L, Xianjun L, Xianting W, Wengui Y, Guangjun R (2013). Application of resistance gene analog markers to analyses of genetic structure and diversity in rice. Genome 56(7):377-387.

Rocha TO, Freitas JS, Santos ESL, Scaldaferri MM, Gonçalves C, Cerqueira-Silva CBM (2016). Estimate of diversity and structure genetic in cassutinga (*Croton heliotropiifolius*) based on molecular markers. African Journal of Biotechnology 15(13):518-523.

Scaldaferri 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. African Journal of Biotechnology 12(30):4761-4766.

Scaldaferri 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. Genetics and Molecular Research 13(3):5566-5573.

Silva SLC, Carvalho MG, Gualberto AS, Carneiro-Torres DS, Vasconselos KCF, Oliveira NS (2010). Bioatividade do extrato etanóico do caule de *Croton linearifolius* Mull. Arg. (Euphorbiaceae) sobre *Cochliomyia macellaria* (Diptera: Calliphoridae). Acta Veterinaria Brasilica 4(4):252-258.

Silva SLC, Gualberto AS, Carvalho KS, Fries DD (2014). Avaliação da atividade larvicida de extratos obtidos do caule de *Croton linearifolius* Mull. Arg. (Euphorbiaceae) sobre larvas de *Aedes aegypti* (Linnaeus, 1762) (Diptera: Culicidae). Biotemas 27(2):79-85.

Silva TSS, Freitas JS, dosSantos ESL, Cardoso TdosS, Cerqueira-Silva CBM (2018). Caracterização e seleção de marcadores moleculares em *Croton linearifolius* Mull. Arg. como subsídio para estudos genéticos. Multi-Science Journal 1(10):4-8.

Zanella CM, Janke A, Palma-Silva C, Kaltchuk-Santos E, Pinheiro FG, Paggi GM, Soares LES, Bered F (2012). Genetics evolution and conservation of Bromeliaceae. Genetics and Molecular Biology 35(4):1020-1026.

Webster GL (1994). Synopsis of the genera and suprageneric taxa of Euphorbiaceae. Annals of the Missouri Botanical Garden 81:33-144.

Supplementary Material 1. Geographic coordinates and number of *Croton linearifolius* individuals sampled by regions and collection points in the Contendas do Sincorá National Forest, Bahia, Brazil.

Collection region	Collection points	Latitude	Longitude	Plant code	Individuals number
Region 1	1	-13.922	-41.117	UESB-Cl-1-13	13
Region 2	1	-13.943	-41.113	UESB-Cl-14-32	5
	2	-13.942	-41.112		5
	3	-13.942	-41.112		5
	4	-13.941	-41.112		4
Region 3	1	-13.97	-41.119	UESB-Cl-33-61	3
	2	-13.97	-41.119		5
	3	-13.97	-41.119		5
	4	-13.969	-41.119		5
	5	-13.97	-41.119		11