Validation of microsatellite loci for *Balfourodendron riedelianum* through analysis of Mendelian inheritance, genetic linkage, and genotypic linkage disequilibrium

B. I. Aguiar[1*], F. B. Gandara[1], M.L.M. Freitas[2], M.L.T. Moraes[3]; P.Y. Kageyama[1 †] and A.M. Sebbenn[2]

[1] Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Av. Pádua Dias, 11, Caixa Postal 9, Piracicaba, São Paulo State, 13418-900, Brazil.
[2] Instituto Florestal de São Paulo, CP 1322, São Paulo State, Brazil, 01059-970.
3 Universidade Estadual Paulista "Júlio de Mesquita Filho", Campus de Ilha Solteira, A. Brasil, 56, Ilha Solteira, São Paulo State, 15385-000, Brazil.
[†] in memorian

*Corresponding author: Bruna Ibanes Aguiar, E-mail: brunaibanes@yahoo.com.br

Abstract

Balfourodendron riedelianum is a tropical tree endemic to the Atlantic Forest biome that is currently endangered due to forest fragmentation and extensive exploitation. Three populations of the species are conserved *ex situ* in a provenance and progeny test at the Luiz Antônio Experimental Station, São Paulo State, Brazil. To verify if seven microsatellite loci developed for the species can be used as genetic markers in analyses focused on conservation and sustainable use strategies, leaf tissue samples were collected from the three provenances and from 17 seed trees in one provenance. We analyzed Mendelian inheritance and genetic linkage for the 17 seed trees and genotypic linkage disequilibrium for individuals from the three provenances. For six of the seven loci analyzed, all 17 seed trees showed heterozygosity. The inheritance and genetic linkage analyses were performed using respective locus, while the genotypic linkage disequilibrium analysis was performed for the seven loci. After Bonferroni correction, none of the 75 tests showed deviation from Mendelian segregation and genetic linkage, nor did we detect genotypic linkage disequilibrium. The results suggest that six of the seven loci can be used for population genetics studies on *B. riedelianum*.

Keywords: : *Atlantic Forest; Conservation genetics; Inheritance; Microsatellite markers; Tropical tree species*

Introduction

Balfourodendron riedelianum (Engl.) Engl. (Rutaceae) is endemic to the Atlantic Forest (Carvalho, 2004) and occurs naturally in Brazil, in the states of Espírito Santo, Mato Grosso do Sul, Minas Gerais, São Paulo, Paraná, Santa Catarina, and Rio Grande do Sul, as well as in Argentina and Paraguay. Currently, *B. riedelianum* is considered an endangered species (IUCN, 2015) as it occurs in the Brazilian biome most severely affected by forest fragmentation and anthropogenic activities (Ribeiro et al., 2009) and due to extensive exploitation of the species for its valuable wood.

To develop conservation and sustainable use strategies, as well as seed collection programs for environmental restoration, it is necessary to assess the genetic diversity, reproductive system, and gene flow of the species (Degen and Sebbenn, 2014). Microsatellite markers are an efficient tool in population genetics studies because they are robust, codominant, highly polymorphic, and relatively low cost (Nybom et al., 2014). However, in order to be used as genetic markers, it is necessary to determine the genetic linkage and equilibrium between loci and whether the inheritance pattern follows the laws of independent Mendelian segregation. Since genetic analyses assume regular Mendelian inheritance and an absence of genetic linkage and genotypic disequilibrium between pairs of loci, this information is essential to avoid bias in multilocus estimates (Manoel et al., 2015; Pupin et al., 2017).

DOI:10.2478/sg-2019-0015
edited by the Thünen Institute of Forest Genetics

The present study aimed to verify if the seven microsatellite markers developed for *B. riedelianum* follow the pattern of Mendelian segregation and show an absence of linkage and genotypic disequilibrium. This verification determines the suitability of the markers for use in population genetics and related research studies.

Materials and Methods

Study area and sampling

The study was carried out in a *B. riedelianum* provenance and progeny test established in 1986 at the Luiz Antônio Experimental Station, Luiz Antônio, São Paulo State, Brazil. The trial was established using open-pollinated seeds collected from 19 seed trees in three provenances: Alvorada do Sul (AS), Paraná State; and Gália (GA) and Bauru (BA), São Paulo State. For the Mendelian segregation and genetic linkage analysis, leaves and seeds of 17 seed trees from the AS provenance were collected, for a total of 235 samples. To analyze genotypic linkage disequilibrium, leaf tissue of 387, 359, and 354 individuals from provenances AS, GA, and BA, respectively, were sampled.

DNA extraction and microsatellite analysis

For nuclear DNA extraction, 30 mg of leaves dehydrated in silica gel were used for seed trees, and for seedlings we used 100 mg of fresh leaves (Inglis et al., 2016). DNA quantification was performed by spectrophotometry using NanoDrop® (Nanodrop/Thermo Fisher Inc.). The samples contained an absorbance ratio of 260/280 nm between 1.8 to 2.0 and 260/230 nm above 2.0. The genotyping of *B. riedelianum* microsatellite loci was performed by Heréditas/Genomax Tecnologia em DNA Ltda. Genotyping was conducted in a multiplex system on the ABI 3100 (Applied Biosystem), detection of alleles was performed by fluorescence, and allele size estimates were assessed using the Genescan software (Applied Biosystem). The values were analyzed with Genotyper (Applied Biosystems) to filter peaks and interpret the genotype of each individual.

Analysis of genotypic and linkage disequilibrium

The Gillet and Hattemer (1989) method was used to assess Mendelian inheritance of loci. This method compares the maternal heterozygous genotype with the segregation of its open-pollinated descendant progeny. It assumes that all loci have regular segregation, alleles follow a Mendelian segregation pattern, and the following expectations are met: i) regular meiotic segregation during ovule production; ii) random fertilization of ovules; and iii) no selection between the moment of fertilization and the genotyping of seedlings. The model further assumes a codominant relationship among all alleles and that all seedlings have at least one maternal allele. In the case of heterozygous mothers (for example: A_iA_j; $i \neq j$) the following expectations must be met: i) each individual within the progeny has at least one maternal allele (A_i or A_j); ii) the number of heterozygous individuals $A_iA_j(n_{ij})$ should be equal to the sum of the number of homozygous individuals $A_iA_i(n_{ii})$ and $A_jA_j(n_{jj}) : n_{ij} = n_{ii} + n_{jj}$ within the same progeny; and iii) the number of heterozygous individuals with allele i (A_iA_k; n_{ik}) must be equal to the number of individuals with allele j (A_jA_k; n_{jk}), thus, $k \neq i, j$.

The observed segregation for each heterozygous mother was compared to that expected for the 1:1 segregation hypothesis using the G-test (Sokal and Rohlf, 1981):

$$G1 = 2\left[n_{ij} \ln\left(\frac{n_{ij}}{E(n1)}\right) + (n_{ii} + n_{ij}) \ln\left(\frac{(n_{ii} + n_{jj})}{E(n1)}\right)\right],$$

where ln is the natural logarithm, $E(n1)$ is the expected number for the progeny genotype $A_jA_j(n_{jj})$ and $A_iA_i + A_jA_j(n_{ii} + n_{jj}) : E(n1) = 0.5(n_{ij} + n_{ii} + n_{jj})$; or $G2 = 2\left[n_{ik} \ln\left(\frac{n_{ik}}{E(n2)}\right) + n_{jk} \ln\left(\frac{n_{jk}}{E(n2)}\right)\right]$, where $E(n2)$ is the expected number of genotypes for alleles $A_iA_k(n_{ik})$ and $A_jA_k(n_{jk}) : E(n2) = 0.5(n_{ik} + n_{jk})$. To avoid false positives, the G-test was determined only when $n1$ and $n2 \geq 10$. Deviation of the G-test between the expected and observed segregation was determined statistically using the Bonferroni correction for multiple comparisons (95 %, α = 0.05).

To determine if the loci were genetically linked, a test was performed between pairs of loci using genetic information from parent trees that were heterozygous for two loci (A_iA_j, B_lB_m). Segregation was recorded in the progeny and the null hypothesis (H_0) was regular 1:1:1:1 Mendelian segregation. The hypothesis of regular segregation between pairs of loci was accepted or rejected based on the maximum likelihood of the G-test (Sokal and Rohlf, 1981), performed for each progeny:

$$G = 2\left[n_{il} \ln\left(\frac{n_{il}}{E(n)}\right) + n_{im} \ln\left(\frac{n_{im}}{E(n)}\right) + n_{jl} \ln\left(\frac{n_{jl}}{E(n)}\right) + n_{jm} \ln\left(\frac{n_{jm}}{E(n)}\right)\right], \text{ where}$$

n_{il}, n_{im}, n_{jl}, n_{jm} are the numbers of phenotypes observed for the phenotypes A_iB_l, A_iB_m, A_jB_l, A_jB_m; and $E(n)$ is the expected number of each genotype A_iB_l, A_iB_m, A_jB_l, A_jB_m calculated by: $E(n) = 0.25(n_{il}, n_{im}, n_{jl}, n_{jm})$. The Bonferroni correction was applied for multiple comparisons (95 %, α = 0.05).

Genotypic linkage disequilibrium was tested between pairs of loci only for adult individuals; progeny were not included in this analysis as all progeny have at least one maternal allele which creates bias in gene frequency estimates and may result in linkage disequilibrium. This analysis was performed using FSTAT (Goudet, 1995). The probabilities of significance of the test were obtained by permutation of the alleles between individuals with the Bonferroni correction for multiple comparisons (95 %, α = 0.05).

Results

Of the seven microsatellite loci, one locus did not show polymorphism. For six loci, no deviation was observed from the expected proportion of 1:1 Mendelian segregation after Bonferroni correction after 75 tests. However, this analysis was not performed for all seed trees due to excess of homozygosity in relation to expected ($0 : n$ for $n_{ij} : n_{ii} + n_{jj}$ or $n_{ik} : n_{jk}$). For Bri17 loci the analysis ($n_{ik} : n_{jk}$) was not performed for the families 1 (3 = 2:1), 10 (5 = 4:1) and 13 (4 = 1:3) due to small sample size (Table 1). Ten pairs of loci no deviation was detected in relation to that expected for independent segregation after 79 tests (Table 2). No linkage disequilibrium was detected for 21 pairs of loci after

Bonferroni correction (Table 3). This analysis strategy was used because the sample of the present study has a family structure, since all individuals must present at least one maternal allele resulting in bias in the analysis.

Table 1
Mendelian inheritance (1:1) for six microsatellite loci developed for *Balfourodendron riedelianum*.

Loci/seed tree	Maternal genotype	n	$n1$	$n_{ij} : n_{ii} + n_{jj}$	$G1$	$n2$	$n_{ik} : n_{jk}$	$G2$
Bri16								
1	275/285	16	11	1:10	8.55	5	0:5	NE
2	267/285	13	12	2:10	5.82	1	0:1	NE
5	275/285	16	13	3:10	3.98	3	1:2	0.33
6	275/285	16	13	7:6	0.08	3	1:2	0.33
9	267/285	13	12	2:10	5.82	1	0:1	NE
12	285/305	17	15	1:14	13.45	2	1:1	0
14	279/285	12	7	0:7	NE	5	3:2	0.20
15	275/285	23	17	8:9	0.06	6	4:2	0.67
17	275/285	8	8	5:3	0.51	0	0	-
Bri21								
1	197/201	16	10	2:8	3.85	6	4:2	0.67
2	191/197	13	10	2:8	3.85	3	0:3	NE
3	199/201	20	14	7:7	0	6	3:3	0
6	197/201	16	12	6:6	0	4	2:2	0
7	197/201	15	10	6:4	0.40	5	3:2	0.20
8	197/201	11	8	4:4	0	3	1:2	0.33
12	197/201	17	11	5:6	0.09	6	3:3	0
14	199/201	12	9	4:5	0.11	3	3:0	NE
15	191/201	23	20	7:13	1.83	3	0:3	NE
17	197/201	8	8	1:7	5.06	0	0	-
Bri4								
2	156/168	13	13	11:2	6.86	0	0	-
6	156/168	16	16	12:4	4.19	0	0	-
8	156/168	11	10	8:2	3.85	1	1:0	NE
9	156/168	13	13	11:2	6.86	0	0	-
11	156/168	8	7	5:2	1.33	1	1:0	NE
12	156/168	17	17	12:5	2.97	0	0	-
Bri10								
1	163/165	16	7	2:5	1.33	9	7:2	2.94
2	157/159	13	5	4:1	1.93	7	4:3	0.14
4	159/161	15	8	3:5	0.51	6	2:4	0.67
5	157/171	16	9	5:4	0.11	7	5:2	1.32
6	163/171	16	6	3:3	0	10	8:2	3.85
7	159/169	15	5	4:1	1.93	10	6:4	0.40
8	157/167	11	7	5:2	1.33	4	4:0	NE
9	159/167	13	9	3:6	1.02	4	2:2	0
11	157/159	8	7	2:5	1.33	1	0:1	NE
12	161/165	17	11	5:6	0.09	6	3:3	0
13	145/165	13	4	2:2	0	9	0:9	NE
14	157/165	12	5	2:3	0.20	7	2:5	1.32
15	157/159	23	10	5:5	0	13	7:6	0.07
Bri13								
6	116/118	16	16	6:10	1.01	0	0	-
7	116/118	15	15	4:11	3.40	0	0	-
8	116/120	11	10	4:6	0.40	1	1:0	NE
9	116/120	13	11	6:5	0.09	2	2:0	NE
14	116/120	12	12	6:6	0	0	0	-
Bri17								
1	253/255	16	13	4:9	1.97	3	2:1	NE
3	243/257	20	10	3:7	1.65	10	6:4	0.40
4	243/257	15	5	0:5	NE	9	5:4	0.11
5	239/999	16	5	0:5	NE	8	2:6	2.09
8	281/999	11	3	0:3	NE	8	4:4	0
10	255/277	14	9	2:7	2.94	5	4:1	NE
12	257/283	17	9	2:7	2.94	8	2:6	2.09
13	267/283	13	9	1:8	6.20	4	1:3	NE
14	255/267	12	3	0:3	NE	9	5:4	0.11
15	257/277	23	15	2:13	9.01	8	2:6	2.09

n : sample size; $G1$ and $G2$: G-test with one degree of freedom; NE: not estimated; to be considered statistically significant, $G1$ and $G2$ values after Bonferroni correction for the 75 tests performed should be greater than 17.36.

Table 3.
Probability of genotypic linkage disequilibrium between pairs of loci sampled for adult trees from each of the three provenances.

Loci pair	Alvorada do Sul	Gália	Bauru
Bri16xBri21	0.17202	0.83512	0.99643
Bri16xBri23	1.00000	0.24821	1.00000
Bri16xBri4	0.15417	0.28036	0.92083
Bri16xBri10	0.26012	0.90893	0.89345
Bri16xBri13	0.33333	0.01786	1.00000
Bri16xBri17	0.42143	0.53214	0.90833
Bri21xBri23	0.62560	0.41964	1.00000
Bri21xBri4	0.66845	0.54048	0.70655
Bri21xBri10	0.02321	0.47321	0.59107
Bri21xBri13	0.33155	0.34167	0.96429
Bri21xBri17	0.24524	0.74405	0.20595
Bri23xBri4	0.64286	1.00000	0.70119
Bri23xBri10	0.67619	0.56310	0.22560
Bri23xBri13	1.00000	0.20893	1.00000
Bri23xBri17	0.35536	0.49226	0.91607
Bri4xBri10	0.04524	0.56548	0.52440
Bri4xBri13	1.00000	1.00000	0.03869
Bri4xBri17	0.35238	1.00000	0.86071
Bri10xBri13	0.02262	1.00000	0.78452
Bri10xBri17	0.41964	1.00000	0.53452
Bri13xBri17	0.51131	0.10000	0.91607

The values represent the probability of genotypic imbalance after 1,680 permutations of alleles between individuals. Probability after Bonferroni correction $P = 0.000595$ (95 %, $\alpha = 0.05$).

Discussion

Six of the seven loci evaluated for heterozygous mothers presented Mendelian segregation and did not reject the null hypothesis of 1:1:1:1 segregation. No genotypic disequilibrium was detected for the seven loci. It is important to note that the results observed in the present study may be different from results based on natural *B. riedelianum* populations. Genetic linkage disequilibrium can be detected as a result of self-fertilization, correlated mating, mating between relatives, genetic drift, founder effect, and natural selection, and these effects tend to be more pronounced in populations experiencing increased fragmentation (Flint-Garcia et al., 2003; Gandara et al., 2014).

Table 2
G-test for the null hypothesis of independent segregation (1:1:1:1) between pairs of loci.

	Bri16xBri21	Bri16xBri17	Bri16xBri4	Bri16xBri10	Bri21xBri17	Bri21xBri4	Bri21xBri10	Bri4xBri10	Bri4xBri17	Bri10xBri17
1	0.33 (1)	1.47 (1)	9.55 (2)	7.22 (1)	2.23 (1)	11.93 (2)	10.03 (1)	0.31 (2)	8.87 (2)	11.22 (1)
2	9.50 (2)	11.05 (2)	2.28 (6)	3.97(2)	4.57 (3)	0.88 (6)	4.73 (2)	4.24 (6)	6.87 (6)	5.50 (2)
3	1.09 (6)	2.30 (5)	2.71 (8)	2.94(5)	4.05 (6)	0.86 (7)	3.59 (6)	3.29 (8)	0.56 (8)	3.23 (4)
4	2.30 (3)	5.64 (6)	6.67 (9)	10.68 (6)	2.09 (8)	0.73 (8)	0.67 (7)	1.86 (11)	2.39 (12)	0.50 (5)
5	1.44 (13)	3.20 (8)	1.40 (11)	6.59 (8)	3.62 (11)	1.07 (11)	0.89 (11)	2.66 (12)	-	7.86 (6)
6	2.40 (10)	6.59 (10)	5.68 (12)	4.25 (9)	3.13 (12)	1.79 (12)	5.69 (12)	-	-	3.33 (8)
7	1.41 (14)	4.58 (12)	-	8.54 (10)	7.35 (15)	-	2.04 (14)	-	-	6.12 (9)
8	7.60 (15)	4.029 (14)	-	3.31 (11)	-	-	1.72 (15)	-	-	5.83 (11)
9	2.30 (17)	10.19 (15)	-	2.23 (12)	-	-	2.69 (17)	-	-	1.61 (12)
10	-	-	-	2.65 (14)	-	-	-	-	-	14.98 (13)
11	-	-	-	2.28 (15)	-	-	-	-	-	0.50 (14)
12	-	-	-	1.58 (17)	-	-	-	-	-	4.43 (15)

Numbers in parentheses refer to the families analyzed; to be considered statistically significant, the values of the G-test after Bonferroni correction for the 79 tests performed should be greater than 17.75.

Conclusion

Six of the seven developed microsatellite loci for *B. riedelianum* showed no deviation from Mendelian segregation and did not present genetic linkage or genotypic linkage disequilibrium. Therefore, these loci can be used to quantify the genetic diversity, inbreeding, mating system, gene flow, and kinship of the species.

Acknowledgements

This study was supported financially by the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP, contract number: 2014/27143-9). We would like to thank Marcelo N. Ambrosano, Lya C. S. M. Pereira and Mayara O. Biliati for assistance in laboratory work and Dirceu de Souza for assistance in sample collection. Bruna Ibanes Aguiar was supported by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes: Finance Code 001). Alexandre M. Sebbenn and Miguel L.M. Freitas were supported by research scholarships from the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). Finally, we thank Dr. Evelyn Roberta Nimmo for editing the English of the manuscript.

References

Carvalho PER (2004) Pau-marfim Balfourodendron riedelianum. Circular técnica EMBRAPA 93:1–6.

Degen B, Sebbenn AM (2014) Genetics and tropical forest. In: Pancel L, Köhl M. Tropical Forest Handbook pp. 1–30. https://doi.org/10.1007/978-3-642-41554-8_75-1

Engel VL, Martins FR (2005) Reproductive phenology of Atlantic forest tree species in Brazil: an eleven year study. Tropical Ecology 46:1–16.

Flint-Garcia SA, Thronsberry JM, Bucler IV ES (2003) Structure of linkage disequilibrium in plants. Annual Review of Plant Biology 54:357–374. https://doi.org/10.1146/annurev.arplant.54.031902.134907

Gandara FB, Tambarussi EV, Sebbenn AM, Ferraz EM, Moreno MA, Ciampi AY, Vianello RP, Grattapaglia D, Kageyma PY (2014) Development and characterization of microsatellite loci for Cedrela fissilis Vell (Meliaceae), an endangerd tropical tree species. Silvae Genetica 63:240–243. https://doi.org/10.1515/sg-2014-0031

Gillet E, Hattemer HH (1989) Genetic analysis of isoenzyme phenotypes using single tree progenies. Heredity 63:135–141. https://doi.org/10.1038/hdy.1989.84

Goudet J (1995) Fstat Version 1.2. A computer program to calculate F-statistics. Heredity 86:485– 486. https://doi.org/10.1093/oxfordjournals.jhered.a111627

Inglis PW, Pappas M de CR, Grattapaglia D (2016) Protocolo de extração de DNA e RNA de alta qualidade para espécies ricas em compostos secundários. Comunicado Técnico 5p.

IUCN (2015) International Union for Conservation of Nature - Red List of Threatened Species. Version 2015-4. Available at: http://www.iucnredlist.org/search. Accessed in April 2016.

Manoel RO, Freitas MLM, Tambarussi EV, Cambuim J, Moraes MLT, Sebbenn AM (2015) Mendelian inheritance, genetic linkage, and genotypic disequilibrium at microsatellite loci in Genipa americana L. (Rubiaceae). Genetics and Molecular Research 14:8161–8169. https://doi.org/10.4238/2015.july.27.4

Nybom H, Weising K, Rotter B (2014) DNA fingerprint in botany: past, present and future. Investigative Genetics 5:1. https://doi.org/10.1186/2041-2223-5-1

Pupin S, Rosse LN, Souza ICG, Cambuim J, Marino CL, Moraes MLT, Sebbenn AM (2017) Analysis of Mendelian inheritance and genetic linkage in microsatellite loci of Eucalyptus urophylla S.T. Blake. Genetics and Molecular Research 16:1–9. https://doi.org/10.4238/gmr16039713

Ribeiro MC, Metzger JP, Martesen AC, Ponzoni FJ, Hirota MM (2009) The Brazilian Atlantic Forest: How much is left, and how is the remaining forest distributed? Implications for conservation. Biological Conservation 149:11431–1153. https://doi.org/10.1016/j.biocon.2009.02.021

Sokal RR, Rohlf FJ (1981) Biometry: The principles and practice of statistics in biological research. Copyright Ltd., New York.