

SHORT REPORT

Open Access



Development and characterization of nuclear microsatellite markers for *Eremanthus erythropappus* and their transferability across related species

Lucas Fernandes Rocha^{1*} , Natália Ribeiro Paula², Alison Gonçalves Nazareno³ and Dulcinéia de Carvalho²

Abstract

Background: We developed simple sequence repeats (SSR) for *Eremanthus erythropappus* (DC.) MacLeish, an endangered tree species endemic to the Brazilian Savanna and Atlantic Forest biomes, and tested their transferability to two closely related *Eremanthus* species.

Results: Using a genomic library enriched with tandem repeat motifs, we identified 16 primer pairs, and characterized them in two populations. Nine primers amplified the expected size fragments and seven SSRs were polymorphic, providing a total of 38 alleles and an average of 4.22 alleles per marker. The polymorphic information content (PIC) ranged from 0.44 to 0.94 with an average of 0.65. The average observed heterozygosity across all loci varied from 0.61 to 1.00. The observed (H_o) and expected (H_e) heterozygosity within the two populations varied from 0.65 to 1.00 and from 0.31 to 1.00, respectively.

Conclusions: These newly developed SSR markers are a powerful tool for population genetic analyses and may be useful in studies on species ecology, evolution, and taxonomy.

Keywords: *Eremanthus erythropappus*, Simple sequence repeats, Genetic diversity, *Eremanthus incanus*, *Eremanthus glomerulatus*

Background

Eremanthus erythropappus (DC.) MacLeish is an endemic tree species native to the high-elevation mountains of the Brazilian Savanna and Atlantic Forest biomes [12, 13]. Regenerants of the *Eremanthus* genus are light-demanding, as is common with pioneer species [3, 21], and *E. erythropappus* normally occurs in shallow, rocky soils with low fertility [17]. Historically in Brazil, the species has been over harvested for use as fence posts in rural areas due to the durability of the wood against

weathering [22, 23], and for the extraction of a sesquiterpene alcohol, α -bisabolol, which is used in many pharmaceutical products because of its anti-inflammatory, antibacterial, skin-smoothing, and wound healing properties [10, 24].

Microsatellite markers, also known as simple sequence repeats (SSR), are often used to study the genetic diversity of plants and animals. SSR sequences are distributed throughout the euchromatin and eukaryote genome [11, 19] and exhibit high polymorphism and repeatability across laboratories [28]. Microsatellite loci present the highest information content among all classes of molecular markers [27]. These markers can be analyzed using polymerase chain reaction (PCR), allowing for a highly informative evaluation of a large number of loci, as well

*Correspondence: lucasrochamg@gmail.com

¹ Department of Forest Science, São Paulo State University, UNESP, Botucatu, São Paulo state 18.610-034, Brazil
Full list of author information is available at the end of the article



© The Author(s) 2020. This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

as assessing effects related to species population genetics, breeding programs, and germplasm conservation [18]. The objective of this study was to develop a set of microsatellite markers to be used to evaluate the genetic diversity and structure of *E. erythropappus* and other species of the *Eremanthus* genus.

Results

From the initial 16 primer pairs, nine amplified the expected size fragment while the remaining seven either did not amplify or generated multiple bands indicating non-specific amplification. The optimum annealing temperature was 53 °C for all nine primers. Of the nine amplified primers, seven were polymorphic (Ere02, Ere03, Ere07, Ere08, Ere10, Ere13, and Ere14) and two were monomorphic (Ere04 and Ere09). The observed and expected heterozygosities ranged from 0.31 to 1.00 and from 0.38 to 0.91, respectively. The total number of alleles was 38 and the number of alleles per polymorphic marker ranged from 2 to 18 (Table 1). The mean within-population inbreeding coefficient (F_{IS}) was -0.23. The polymorphic information content (PIC) ranged from 0.44 to 0.94, with an average of 0.65. The combined exclusion probability is sufficient to perform a paternity/maternity exclusion analysis among breeding populations using the seven polymorphic loci. Thus, for the first parent, the exclusion probability reached 0.98 and for the second parent it reached 0.99. None of the tested loci showed significant deviation from Hardy–Weinberg equilibrium, nor did they exhibit a high frequency of null alleles ($\bar{r}_d < 0.5$; $P > 0.05$). Six SSRs (Ere03, Ere07, Ere08, Ere09, Ere13, and Ere14) cross amplified in *E. incanus* and *E. glomerulatus*.

Discussion

The development of SSR markers for *E. erythropappus* enables the application of new genetic research into this endemic and overexploited tree species. Whereas *E. erythropappus* presents relevant economic interest and remarkable ecological importance, new studies will enable us to analyze the genetic diversity, gene flow and also possible processes of inbreeding and clonality. Overall, the genetic diversity estimates found herein were considerably higher than those found in previous studies on the species using inter-simple sequence repeat (ISSR) markers [6, 7, 15].

Although SSR markers are sometimes identified as monomorphic through agarose electrophoresis, they could be polymorphic considering the resolution of capillary electrophoresis (CE). While MetaPhor™ Agarose (4 bp) offers a high resolution, CE presents a better resolution and greater separation efficiency (2 bp) [25]. Thus, the two monomorphic primers developed herein may also be valuable for studying genetic parameters [14]. The evidence of previous researches suggests that the high information content found in SSR markers may enable their use as primers.

Conclusions

We developed nine microsatellite markers, of which seven are polymorphic and two are monomorphic. These markers can inform new research on population genetics, genetic diversity, spatial genetic distribution, as well as the sustainability of forest management practices employed for this species. Additionally, these primer pairs may be an important tool to assist breeding programs of this species. The analysis of cross-amplification

Table 1 Genetic characterization of the nine developed microsatellites for *Eremanthus erythropappus* and its transferability to *Eremanthus incanus* and *Eremanthus glomerulatus*

Locus	PIC	Itamonte (n = 21)				Lavras (n = 21)				Transferability	
		Na	Ho	He	FIS	Na	Ho	He	FIS	<i>E. incanus</i>	<i>E. glomerulatus</i>
Ere02	0.61	3	1.00	0.56	-0.78	3	1.00	0.61	-0.63	-	-
Ere03	0.68	5	1.00	0.60	-0.67	9	0.65	0.56	-0.16	+	+
Ere04	0.44	1	0.00	0.00	0.00	1	0.00	0.00	0.00	-	-
Ere07	0.53	2	0.85	0.49	-0.74	2	0.83	0.49	-0.71	+	+
Ere08	0.67	2	0.95	0.49	-0.92	3	0.31	0.38	0.19	+	+
Ere09	0.51	1	0.00	0.00	0.00	1	0.00	0.00	0.00	+	+
Ere10	0.94	11	0.84	0.84	0.27	18	0.81	0.91	0.11	-	-
Ere13	0.79	3	0.48	0.48	1.00	5	0.00	0.64	1.00	+	+
Ere14	0.67	3	0.58	0.58	-0.54	3	0.62	0.43	-0.45	+	+

Geographic coordinates for the populations are Itamonte = 22°16'45.00"S 44°46'25.60"W and Lavras = 21°19'51.85"S 44°57'54.76"W; All two populations are located in the Minas Gerais state, Brazil

PIC Polymorphic content information, Na Number of alleles, Ho observed heterozygosity; He expected heterozygosity, FIS Inbreeding coefficient; +, successful PCR amplification; -, unsuccessful PCR amplification

confirmed the possibility of applying these markers in genetic studies of *E. incanus* and *E. glomerulatus*.

Methods

We used a (GA)_n and (CA)_n microsatellite-enriched library based on Billotte et al. [1]. As such, leaf tissue samples from *E. erythropappus* adult trees were collected and preserved in silica gel. Total genomic DNA was extracted using the CTAB method according to Doyle and Doyle [4]. Thirty ng of genomic DNA was digested using the *RsaI* restriction enzyme (Promega, Madison, Wisconsin, USA), and fragments were ligated to adapters (*Rsa21* 5'-CTCTTGCTTACGCGTGGACTA-3' and *Rsa25* 5'-TAGTCCACGCGTAAGCAAGAGCACACA-3'). For the enrichment of SSRs sequences, we used (CT)₈, (GT)₈, and (TTC)₈ repeats using biotinylated microsatellite probes, and the target fragments were captured by the use of streptavidin-coated magnetic beads (Promega Corporation, Madison, Wisconsin, USA). The *Rsa21* and *Rsa25* adapter sequences were used as primer templates for the amplification of fragments. The microsatellite fragments were ligated to a pGEM-T Easy Vector System (Promega Corporation, Madison, Wisconsin, USA). The plasmids were introduced into *Escherichia coli* XL1-Blue strains, and transformed cells were plated on Petri dishes with Luria-Bertani (LB) agar medium containing ampicillin (100 µg ml⁻¹) and X-galactosidase (5-bromo-4-chloro-indolyl-β-D-galactoside) (50 µg ml⁻¹). The

recombinant colonies were sequenced using an ABI 377 automated sequencer and the Big Dye Terminator Kit (Applied Biosystems, Vienna, Austria). We found 16 positive clones that contained microsatellite sequences with at least five tandem repeats. Primer pairs were designed using the software Primer 3 [20] with a product size ranging from 100 to 300 base pairs (bp), primer size from 18 to 22 bp, GC % from 40 to 60, and annealing temperature from 57 to 60 °C.

Polymerase chain reaction (PCR) was performed by screening each primer pair through 10 annealing temperatures (between 46 and 55 °C) for 21 individuals from two different *E. erythropappus* populations. The final volume of each reaction was 15 µl using 30 ng of template DNA added to 12 µl reaction mixture containing 3.33 mM IB Phosneutria buffer (consisting of 100 mM Tris-HCl pH 8.4; 500 mM KCl; 1% Triton X-100; 15 mM MgCl₂), 1.5 mM MgCl₂, 0.28 mM of each dNTP, 1 U Taq polymerase, and 0.22 mM of each primer (forward and reverse). The temperature regime was assessed separately for each primer pair; as such, we tested a total of 17 temperatures (from 46 to 62 °C) for six individuals from two populations using MJ Mini™ Thermal Cycler (Bio-Rad, Singapore). The optimal PCR profile used for the amplification of each microsatellite consisted of an initial 3 min at 94 °C, followed by 30 cycles of denaturation at 94 °C for 30 s, annealing temperature (Table 2) for 30 s, extension at

Table 2 Characteristics of nine microsatellite loci developed for *Eremanthus erythropappus* from Minas Gerais State, Brazil

Locus	Primer sequences (5'-3')	Repeat motif	Size	Tm (°C)	GenBank
ERE02	F: TCTTGCTTACGCGTGTGACT R: TGCATCCACTCCAATCACTT	(GA)21	119	53	MK075833
ERE03	F: GAAGGGAGACATCGGAAGAA R: ACGGAACGGAGAAGAAGAAA	(CTT)5; (CTT)9; (CTT)10; (CTT)8	232	53	MK075834
ERE04	F: CAGTGAGGGGAAGGAGAAT R: CCTCCACTATAGGGCGGAAT	(CTT)37	398	53	MK075835
ERE07	F: GCGTGGGACTAACCCATTC R: ACCTGTTGGTAAAAGGATGC	(CTT)9	120	53	MK075836
ERE08	F: GAGCCTTCCATGGGAGTAGG R: TGGGAGGGAGAAAATTGAACA	(AGC)5	238	53	MK075837
ERE09	F: GCTTACGCGTGGGACTAACT R: GCGTGGACTAGGAAAACGAA	(CA)3; (GA)8; (GTA)6	269	53	MK075838
ERE10	F: GATCATCGCCATGAAGGAAT R: CAGTGAGGGGAAGGAGAAT	(GAA)3; ATA; (GAA)27	244	53	MK075839
ERE13	F: GAGACCCTGGCTGTCTTCAT R: GCGTTGAGTTTCGGAGAAGT	(CT)6; (CA)6	378	53	MK075840
ERE14	F: CATCGATTTGGAGGCTTCAT R: TGCTTACGTGTGCTCTTGCT	(CT)11; (AT)8; (GT)18	207	53	MK075841

Locus name, primer sequence (F: forward, R: reverse), repeat motif, fragment size (base pair), Temperature of melting (Tm), and GenBank accession numbers are shown

72 °C for 1 min, and a final extension at 72 °C for 7 min. Amplifications were performed using a MJ Mini™ Thermal Cycler.

Additionally, we sampled 42 individuals from population 1 (Itamonte: 22°16'45.00"S 44°46'25.60"W) and population 2 (Lavras: 21°19'51.85"S 44°57'54.76"W) to evaluate the SSR primer pairs. Voucher specimens were deposited in the ESAL herbarium of the Federal University of Lavras (UFLA), Brazil. Amplification was conducted using the thermocycler GeneAmp PCR System 9700. We applied the same reaction components and PCR thermal cycle used in the validation process. We separated the PCR products using a 3% high-resolution MetaPhor™ agarose (Lonza, Rockland, Maine, USA) stained with GelRed™. Allele sizes were estimated by comparison to a 10-bp DNA Ladder standard (Invitrogen, Carlsbad, California, USA). Individuals that failed to amplify at a minimum of three primer pairs were excluded.

To analyze genotyping errors due to the presence of null alleles, stuttering, and allele dropout, we applied the Brookfield 1 method [2] using 1000 permutations and the Micro-Checker 2.2.3 software [26]. We estimated allele richness using the MSA software [5]. The number of alleles per locus (N_A), observed heterozygosity (H_O), and expected heterozygosity (H_E) for each population and locus according to the Hardy–Weinberg equilibrium were calculated using GenAEx 6.4 [16]. The within-population inbreeding coefficients (F_{IS}) were determined using FSTAT 2.9.3.2 [8], applying a Bonferroni correction for multiple comparisons. The probability of non-exclusion for each locus, the combined probability of paternity exclusion, and the PIC were calculated using CERVUS 3.0 [9].

We also tested for cross-amplification into two other species, *E. incanus* and *E. glomerulatus*. We amplified the DNA of 10 individuals for each species from one population located in the Beautiful River Falls Ecological Park, Lavras, Minas Gerais State, Brazil (21°19'44.98"S; 44°58'24.58"W). PCR reactions and electrophoresis were performed following the same protocol described above. Subsequently, we compared amplification bands to identify the primers that showed amplification patterns for the two species.

Abbreviations

PCR: Polymerase chain reaction; CTAB: Cetyltrimethylammonium bromide; SSR: Simple sequences repeats; PIC: Polymorphic information content; DNA: deoxyribonucleic acid; H_O : Observed heterozygosity; H_E : Expected heterozygosity; F_{IS} : Inbreeding coefficient.

Acknowledgements

We are thankful to Eve R Nimmo for English review.

Authors' contributions

LFR, DC and NRB performed the laboratory work; DC, LFR and AGN analyzed the raw data and contributed with the statistical and genetical analysis. DC also administered the funds of this research. All authors read and approved the final manuscript.

Funding

We are grateful to FAPEMIG (Fundação de Amparo à Pesquisa do Estado de Minas Gerais) for financial support (CRA APQ-02641-14). This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹ Department of Forest Science, São Paulo State University, UNESP, Botucatu, São Paulo state 18.610-034, Brazil. ² Department of Forest Sciences, Laboratory of Tree Genetic Conservation, Federal University of Lavras, UFLA, Lavras, Minas Gerais state 37.200-000, Brazil. ³ Department of Biology, Federal University of Minas Gerais, UFMG, Belo Horizonte, Minas Gerais state 31.270-901, Brazil.

Received: 6 August 2019 Accepted: 29 June 2020

Published online: 07 July 2020

References

- Billotte N, Lagoda P, Risterucci AM, Baurens FC. Microsatellite-enriched libraries: applied methodology for the development of SSR markers in tropical crops. *Fruits*. 1999;54:277–88.
- Brookfield J. A simple new method for estimating null allele frequency from heterozygote deficiency. *Mol Ecol*. 1996;5:453–5.
- Carvalho PER. Espécies florestais brasileiras: recomendações silviculturais, potencialidade e uso da madeira. Brasília: EMBRAPA; 1994.
- Doyle JJ, Doyle JL. Isolation of plant DNA from fresh tissue. *Geog Genet*. 1987;12:13–5.
- Dieringer D, Schlotterer C. Microsatellite analyser (MSA): a platform independent analysis tool for large microsatellite data sets. *Mol Ecol Notes*. 2003;19:167–9.
- Estopa RA, Souza AD, Moura MCO, Botrel MCG, Mendonça EG, Carvalho D. Diversidade genética em populações naturais de candeia (*Eremanthus erythropappus* (DC.) MacLeish). *Sci Forestalis*. 2006;70:97–106.
- Freitas VLO, Lemos-Filho JP, Lovato MB. Contrasting genetic diversity and differentiation of populations of two successional stages in a neotropical pioneer tree (*Eremanthus erythropappus*, Asteraceae). *Genet Mol Res*. 2008;7:388–98.
- Goudet J. FSTAT (version 1.2): a computer program to calculate F-statistics. *J Hered*. 1995;86:485–6.
- Kalinowski ST, Taper ML, Marshall TC. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Mol Ecol*. 2007;16:1099–106.
- Kamatou GP, Viljoen AM. A review of the application and pharmacological properties of α -Bisabolol and α -Bisabolol-rich oils. *J Am Oil Chem Soc*. 2010;87:1–7.
- Katti MV, Ranjekar PK, Gupta VS. Differential distribution of simple sequence repeats in eukaryotic genome sequences. *Mol Biol Evol*. 2001;18:1161–7.
- Loeuille B, Lopes JC, Pirani JR. Taxonomic novelties in *Eremanthus* (Compositae: Vernonieae) from Brazil. *Kew Bull*. 2012;67:1–9.

13. MacLeish NF. Revision of *Eremanthus* (Compositae: Vernonieae). *Ann Missouri Bot Gard.* 1987;74:265–90.
14. Nazareno AG, Reis MS. The same but different: monomorphic microsatellite markers as a new tool for genetic analysis. *Am J Bot.* 2011;98(10):265–7.
15. Pádua JAR, Brandão MM, Carvalho D. Spatial genetic structure in natural populations of the overexploited tree *Eremanthus erythropappus* (DC.) MacLeish (Asteraceae). *Biochem Syst Ecol.* 2016;66:307–11.
16. Peakall RO, Smouse PE. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Mol Ecol Notes.* 2006;6:288–95.
17. Pérez FMJ, Scolforo SJR, Mello JM, Borges RLF, Camolesi JF. Sistema de manejo para a candeia-*Eremanthus erythropappus* (DC.) MacLeish a opção do sistema de corte seletivo. *Cerne.* 2004;10:257–73.
18. Powell W, Morgante M, Andre C, Hanafey M, Vogel J, Tingey S, Rafalski A. The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. *Mol Breed.* 1996;2:225–38.
19. Phumichai C, Phumichai T, Wongkaew A. Novel chloroplast microsatellite (cpSSR) markers for genetic diversity assessment of cultivated and wild *Hevea rubber*. *Plant Mol Biol.* 2015;33:1486–98.
20. Rozen S, Skaletsky H. Primer3 on the WWW for general users and for biologist programmers. *Bio Methods Prot.* 2000;132:365–86.
21. Scolforo JRS, Loeuille BFP, Altoé TF. Caracterização da candeia. In: Scolforo JRS, Oliveira AD, Davide AC, editors. *O manejo sustentável da candeia: o caminhar de uma nova experiência florestal em Minas Gerais.* 1st ed. Lavras: Editora UFLA; 2012.
22. Scolforo JRS, Oliveira AD, Acerbi Junior FW. Inventário Florestal de Minas Gerais: equações de volume, peso de matéria seca e carbono para diferentes Fitofisionomias da Flora Nativa. 1st ed. Lavras: Editora UFLA; 2008.
23. Silva MA, Melo JM, Scolforo JR, Czanck Júnior L, Andrade IS, Oliveira AD. Análise da distribuição espacial da candeia (*Eremanthus erythropappus* (DC.) MacLeish) sujeita ao sistema de manejo porta-sementes. *Cerne.* 2008;14:311–6.
24. Sousa OV, et al. Antinociceptive and anti-inflammatory effects of the essential oil from *Eremanthus erythropappus* leaves. *J Pharm Pharmacol.* 2008;60:771–7.
25. Swerdlow H, Gesteland R. Capillary gel electrophoresis for rapid, high resolution DNA sequencing. *Nucleic Acids Res.* 1989;18(6):1415–9.
26. Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P. MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Mol Ecol Notes.* 2004;4:535–8.
27. Varshney RK, Graner A, Sorrells ME. Genic microsatellite markers in plants: features and applications. *Trends Biotechnol.* 2005;23:48–55.
28. Vieira ML, Santini L, Diniz AL, Munhoz CD. Microsatellite markers: what they mean and why they are so useful. *Genet Mol Biol.* 2016;39:312–28.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

