Isolation and Characterization of 13 Microsatellites for the Rare Endemic Shrub Tetratheca erubescens (Elaeocarpaceae)

Authors: Anthony, Janet M., Allcock, Richard J. N., and Krauss, Siegfried L.

Source: Applications in Plant Sciences, 4(2)

Published By: Botanical Society of America

URL: https://doi.org/10.3732/apps.1500102

Applications in Plant Sciences 2016 4(2): 1500102

PRIMER NOTE

ISOLATION AND CHARACTERIZATION OF 13 MICROSATELLITES FOR THE RARE ENDEMIC SHRUB *TETRATHECA ERUBESCENS* (ELAEOCARPACEAE)[1]

JANET M. ANTHONY[2,3,5], RICHARD J. N. ALLCOCK[4], AND SIEGFRIED L. KRAUSS[2,3]

[2]Kings Park and Botanic Garden, Botanic Garden and Parks Authority, Fraser Avenue, West Perth, 6005 Western Australia, Australia; [3]School of Plant Biology, The University of Western Australia, Nedlands, 6009 Western Australia, Australia; and [4]School of Pathology and Laboratory Medicine, Lotterywest State Biomedical Facility Genomics, University of Western Australia, Crawley, Western Australia 6009, Australia

- *Premise of the study:* Microsatellite markers were developed for the rare *Tetratheca erubescens* (Elaeocarpaceae) to assess genetic diversity and spatial structuring.
- *Methods and Results:* We generated ca. 2.7 million sequence reads using a Personal Genome Machine (PGM) semiconductor sequencer. Using the QDD pipeline, we designed primers for >12,000 sequences with PCR product lengths of 80–480 bp. From these, 30 primer pairs were selected and screened using PCR, from which 11 loci were found to be polymorphic and amplified reliably. For a sample of 95 plants from three populations, the number of alleles observed for these 11 loci ranged from two to seven and expected heterozygosity ranged from 0.06 to 0.72. No consistent evidence for null alleles or departure from Hardy–Weinberg equilibrium was found for any of the 11 loci.
- *Conclusions:* These markers will enable the quantification of genetic impact of proposed mining activities on the narrow endemic *T. erubescens*.

Key words: Elaeocarpaceae; microsatellite primers; Personal Genome Machine; shotgun sequencing; *Tetratheca erubescens.*

Tetratheca erubescens J. P. Bull (Elaeocarpaceae) is a recently described species (Bull, 2007) endemic to the Koolyanobbing Range (a banded ironstone formation) within the Coolgardie Biogeographic Region of the South West Australian Floristic Region (SWAFR). *Tetratheca erubescens* inhabits rock crevices containing red sandy loam soils among hill crests, steep slopes, cliffs, and associated rocky monoliths at altitudes of 445–450 m a.s.l., within parts of the range. Although the reproductive biology of the species is unknown, flower morphology and presentation are similar to the related banded ironstone endemic *T. paynterae* Alford, which displays buzz pollination by a suite of small native bee species, high outcrossing, and restricted pollen dispersal (Butcher et al., 2011). Seed dispersal is likely to be myrmecochorous due to the presence of a large elaiosome on the seed. Due to its isolation, low total number of individuals (ca. 6000 plants), a narrow distribution limited to less than 2 km, its ecological association with steep cliffs and associated rocky slopes that are rare habitat types in the region, and proximity to mining activity, it is listed as a Rare Flora under the Wildlife Conservation Act 1950 (WA) (Western Australian Minister for Environment, 2013). To assess the potential genetic impact of proposed mining activity on *T. erubescens*, microsatellite markers were developed to characterize genetic diversity and its spatial structure across the species' range.

[1]Manuscript received 3 September 2015; revision accepted 2 October 2015.
The authors thank Cliffs Natural Resources for funding this work.
[5]Author for correspondence: janet.anthony@bgpa.wa.gov.au

doi:10.3732/apps.1500102

METHODS AND RESULTS

Genomic DNA was extracted from fresh stem material of a single plant sampled at Koolyanobbing using a modified Carlson's method (Carlson et al., 1991) without the addition of β-mercaptoethanol to the lysis buffer and with the additional steps of potassium acetate following lysis incubation and a 5 M NaCl step followed by ethanol precipitation after the isopropanol precipitation step. Next-generation sequencing was performed on a Personal Genome Machine (PGM) semiconductor sequencer (Life Technologies, Carlsbad, California, USA) at the Lotterywest State Biomedical Facility Genomics Node in Perth, Western Australia. Briefly, 100 ng of DNA was sheared to approximately 300–400 bp using an S2 sonicator (Covaris, Woburn, Massachusetts, USA), and a single barcoded library was prepared using a NEBNext Ultra DNA Library Prep Kit (New England Biolabs, Ipswich, Massachusetts, USA). Size selection (insert sizes 330–360 bp) was performed by gel excision (E-Gel; Invitrogen/Thermo Fisher Scientific, Waltham, Massachusetts, USA), and the libraries were assessed and quantified using a Bioanalyzer 2100 (Agilent Technologies, Santa Clara, California, USA). The library was diluted to 9 pM for template preparation using a OneTouch 2 Template 400 kit (Life Technologies) and enriched. Sequencing was performed on a PGM using 850 flows (generating approximately 350–400-bp read lengths) on a 316 sequencing chip. After sequencing, signal processing, base-calling, and quality trimming were performed using the default settings of TorrentSuite 4.0, and library-specific FASTQ files were also generated. Sequencing resulted in >2.7 million reads, with a modal read length of 331 bp (total data output 647 Mb).

We used the QDD version 3.1 pipeline (Meglecz et al., 2014) to screen the raw sequences, remove redundant sequences, and design primers for >12,000 sequences with PCR product lengths of 80–480 bp. The default parameters of the program were used both for the screening steps and for primer design. The

resultant sequences were filtered to ensure that the primer was not overlapping the repeat sequence, there were no poly(A) or poly(T) runs for more than 7 bp within the sequence, and there was only one repeat motif between the primers. Because there were so many sequences containing microsatellites, we employed the suggestions of Meglecz as identified on the QDD website (http://net.imbe.fr/~emeglecz/qdd.html#choice) to choose primers from the primer table.

Thirty primer pairs were selected for initial screening using a CFX96 Touch Real-Time PCR Detection System (Bio-Rad Laboratories, Hercules, California, USA); in addition, 20 primer pairs previously developed for *T. paynterae* subsp. *paynterae* (Butcher and Krauss, 2009) were included in the screening. Amplification was performed using 5 µL SsoAdvanced SYBR Green Supermix (Bio-Rad Laboratories), 0.3 µM each forward and reverse primers, and 5–10 ng genomic DNA in a 10-µL reaction volume. Initial screening included reliable amplification of a single sample across a range of temperatures to determine the most appropriate annealing temperature, followed by evidence of polymorphism using the best temperature and eight samples from different populations across the range using precision melt cycles. The results were analyzed using Precision Melt Analysis software (Bio-Rad Laboratories). Next, fluorescently labeled primers were ordered to further test the optimal conditions for polymorphic loci on the ABI 3500 sequencer (Life Technologies). Subsequently, 13 primer pairs were selected to complete the study (Table 1). The remaining 17 loci did not amplify reliably.

For amplification of microsatellite loci, 10–20 ng of genomic DNA extracted from a minimum of 30 plants from each of three populations of *T. erubescens* was amplified by PCR with 2 µL 5× polymerase buffer containing dNTPs (Fisher Biotec, Wembley, Western Australia, Australia), 2 mM $MgCl_2$, 0.16 µM each reverse and fluorescently labeled forward primers (6-FAM, HEX, PET, and NED), and 0.5 units *Taq* polymerase (Fisher Biotec) in a 10-µL reaction volume. PCR was performed on a Veriti thermocycler (Life Sciences) using two amplification cycles: (1) an initial 2 min of denaturation at 94°C followed by 35 cycles of denaturation (94°C, 30 s), annealing (58°C, 30 s), and extension (72°C, 30 s), with a final extension of 30 min at 72°C (loci Te15, Te21, Te29, Te30); or (2) an initial 2 min of denaturation at 95°C followed by 35 cycles of denaturation (94°C, 30 s), annealing (58°C, 40 s), and extension (72°C, 30 s), with a final extension of 30 min at 72°C (loci Te10, Te23, Te25, TpB4, TpB11, TpC130, TpC131). PCR products were separated by capillary electrophoresis using an ABI 3500 Genetic Analyzer (Life Technologies), and allele sizes were determined using Geneious version 7.1 (Biomatters, http://www.geneious.com/). Multiple replicate runs were performed to ensure the accuracy of the final data set. Genetic diversity parameters were calculated for three populations using GenAlEx version 6.4 (Peakall and Smouse, 2006). Departure from Hardy–Weinberg equilibrium (HWE) was assessed for each locus and population by χ^2 tests, and the possibility of null alleles was checked using MICRO-CHECKER version 2.2.3 (van Oosterhout et al., 2004).

The number of alleles observed for the 11 polymorphic loci ranged from two to seven (Table 2), and the observed and expected heterozygosities ranged from 0.07 to 0.88 and 0.06 to 0.72, respectively (Table 2). Across three populations, there was no consistent evidence of null alleles, large allele dropout, scoring error due to stuttering, or departure from HWE for any of the 11 loci.

CONCLUSIONS

The 11 polymorphic microsatellites developed here will enable the characterization of the amount of genetic variation within *T. erubescens*, as well as the spatial structuring of that genetic variation. This information will enable a direct assessment of the potential genetic impact of removal of *T. erubescens* plants as a result of proposed mining; this is anticipated to inform the government regulatory process associated with the request to mine. These genetic data will also generate indirect estimates of gene flow within the species, from which inferences on seed and pollen dispersal can be made. These markers will also facilitate the direct estimation of mating systems and dispersal, through parentage assignment of seed. Ultimately, these markers will also facilitate the ongoing genetic monitoring of impacts on remaining plants from proposed mining activities. Stingemore et al. (2013) suggested that optimizing microsatellite markers following next-generation sequencing was time consuming; we found that with the assistance of real-time PCR and Precision Melt Analysis software the optimization of all loci was time and cost efficient.

TABLE 1. Characteristics of microsatellite primers developed for *Tetratheca erubescens.*

Locus	Primer sequences (5′–3′)	Fluorescent label[a]	Repeat motif	Allele size range (bp)	GenBank accession no.
Te10	F: CCAAGACCTTGCGGAACTTA R: GAGGGAATACGGACATTGGA	FAM	$(GAG)_{11}$	159–186	KM226784
Te15	F: ATAATTCAATTGGGCTGCCA R: TTGAATCACAGGCACTATGCA	VIC	$(TATT)_8$	119–143	KM226785
Te21	F: CGTGAAATTACAAGCATGGG R: TCTCAATTTCGGTCAGGAGG	PET	$(TATTT)_6$	124–139	KM226786
Te23	F: TGGTTTCAGGAAGTAATAGAAGCC R: TCAATTCGCCAAACAAATCT	VIC	$(CTT)_{10}$	104–122	KM226787
Te24	F: GCCATTAGACATCGTGCCAT R: TTAGACCGATTACCAAGGCTTC	NED	$(AGGAGC)_5$	200	KM226788
Te25	F: CTGTTTCAGGTGGATGCAAA R: TCCATCTCCATCACAATCGA	PET	$(ATA)_{10}$	151–178	KM226789
Te28	F: CCCTTAGTGATGCTGCAAGG R: CGTAAGTATGGTCCGGTCGT	FAM	$(ACA)_{10}$	137	KM226790
Te29	F: ATGCATCTTCATCCTTGGCT R: TTGGAATATGCTCTCGACGA	NED	$(TTA)_{10}$	119–135	KM226791
Te30	F: CCGGAATCACACCAACCTAC R: GGGTGGGATGTCAGAATCAC	FAM	$(GT)_{14}$	81–95	KM226792
TpB4+	F: AACACTAAACGAGGCAACTGTC R: AATCTACCCTCACCACCACA	FAM	$(GT)_{14}$	85–97	KM245050
TpB11*	F: ATGCACATTCAAGTGTAGACAC R: TTGACCTTGTTAAGAGCCATAT	NED	$(TC)_{13}$	164–172	EU350493
TpC130*	F: TAAACTTACCCAACGCACTCT R: GGATGAAACCACATTATGGAC	NED	$(ATC)_7$	150–179	EU350504
TpC131*	F: GGTGTTGACAAAGACCATGTT R: AGCAGGGAAATAGAGCCTG	HEX	$(ATG)_8$	138–150	EU350505

[a] Forward 5′ label.

+ Developed for *Tetratheca paynterae* subsp. *paynterae* and not previously published.

* Developed for *T. paynterae* subsp. *paynterae* (Butcher and Krauss, 2009); these were included to increase the number of polymorphic microsatellites.

TABLE 2. Results of primer screening in three populations (Te-B, Te-E, Te-J) of *Tetratheca erubescens*.[a]

	Te-B (n = 32)				Te-E (n = 33)				Te-J (n = 30)			
Locus	A	H_o	H_e	HWE[b]	A	H_o	H_e	HWE[b]	A	H_o	H_e	HWE[b]
Te10	2	0.50	0.50	n.s.	4	0.58	0.66	n.s.	4	0.67	0.72	n.s.
Te15	3	0.41	0.41	n.s.	5	0.76	0.61	n.s.	6	0.67	0.69	n.s.
Te21	2	0.16	0.19	n.s.	1	0	0	mono	1	0	0	mono
Te23	5	0.62	0.71	n.s.	7	0.70	0.58	n.s.	5	0.60	0.70	n.s.
Te24	1	0	0	mono	1	0	0	mono	1	0	0	mono
Te25	5	0.78	0.59	n.s.	4	0.88	0.69	***	5	0.43	0.55	***
Te28	1	0	0	mono	1	0	0	mono	1	0	0	mono
Te29	3	0.62	0.57	n.s.	3	0.39	0.45	n.s.	2	0.13	0.12	n.s.
Te30	2	0.53	0.43	n.s.	3	0.67	0.53	n.s.	2	0.07	0.06	n.s.
TpB4	3	0.31	0.37	n.s.	3	0.09	0.17	***	3	0.23	0.50	*
TpB11	3	0.44	0.55	n.s.	4	0.52	0.52	n.s.	4	0.67	0.63	n.s.
TpC130	2	0.25	0.22	n.s.	3	0.70	0.49	*	3	0.30	0.27	n.s.
TpC131	2	0.09	0.14	*	4	0.45	0.61	n.s.	2	0.37	0.49	n.s.

Note: A = number of alleles; H_e = expected heterozygosity; H_o = observed heterozygosity; HWE = Hardy–Weinberg equilibrium; n = number of individuals sampled.

[a] Geographic coordinates for the populations are: Te-B = 30.872833°S, 119.604250°E; Te-E = 30.873556°S, 119.609333°E; Te-J = 30.879944°S, 119.614944°E.

[b] Significant departures from HWE are: ***$P \leq 0.001$, **$P \leq 0.01$, *$P \leq 0.05$; n.s. = not significant; mono = monomorphic locus.

LITERATURE CITED

BULL, J. P. 2007. *Tetratheca erubescens* (Elaeocarpaceae), a new and geographically restricted species from the Coolgardie Biogeographic Region of south-western Australia. *Nuytsia* 17: 87–96.

BUTCHER, P. A., AND S. L. KRAUSS. 2009. Development of microsatellites from the rare ironstone endemic *Tetratheca paynterae* ssp. *paynterae* and cross-species amplification. *Molecular Ecology Resources* 9: 386–389.

BUTCHER, P. A., D. BRADBURY, AND S. L. KRAUSS. 2011. Limited pollen-mediated dispersal and partial self-incompatibility in the rare ironstone endemic *Tetratheca paynterae* subsp. *paynterae* increase the risks associated with habitat loss. *Conservation Genetics* 12: 1603–1618.

CARLSON, J. E., L. K. TULSIERAM, J. C. GLAUBITZ, V. W. K. LUK, C. KAUFFELDT, AND R. RUTLEDGE. 1991. Segregation of random amplified DNA markers in F_1 progeny of conifers. *Theoretical and Applied Genetics* 83: 194–200.

MEGLÉCZ, E., N. PECH, A. GILLES, V. DUBUT, P. HINGAMP, A. TRILLES, R. GRENIER, AND J. F. MARTIN. 2014. QDD version 3.1: A user-friendly computer program for microsatellite selection and primer design revisited: Experimental validation of variables determining genotyping success rate. *Molecular Ecology Resources* 14: 1302–1313.

PEAKALL, R., AND P. E. SMOUSE. 2006. GenAlEx 6: Genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* 6: 288–295.

STINGEMORE, J. A., P. G. NEVILL, M. G. GARDENER, AND S. L. KRAUSS. 2013. Development of microsatellite markers for two Australian *Persoonia* (Proteaceae) species using two different techniques. *Applications in Plant Sciences* 1: 1300023.

VAN OOSTERHOUT, C., W. F. HUTCHINSON, D. P. M. WILLS, AND P. SHIPLEY. 2004. MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* 4: 535–538.

WESTERN AUSTRALIAN MINISTER FOR ENVIRONMENT. 2013. Wildlife Conservation (Rare Flora) Notice 2013. *Western Australian Government Gazette* 170: September 2013.