Development and Characterization of 18 Polymorphic SSR Markers for Barthea barthei (Melastomataceae)

Authors: Huang, Guilian, Liu, Haijun, Sun, Hongbin, Liu, Ying, Zhou, Renchao, et al.

Source: Applications in Plant Sciences, 5(4)

Published By: Botanical Society of America

URL: https://doi.org/10.3732/apps.1600149

Applications in Plant Sciences 2017 5(4): 1600149

PRIMER NOTE

DEVELOPMENT AND CHARACTERIZATION OF 18 POLYMORPHIC SSR MARKERS FOR *BARTHEA BARTHEI* (MELASTOMATACEAE)[1]

GUILIAN HUANG[2], HAIJUN LIU[3], HONGBIN SUN[3], YING LIU[2], RENCHAO ZHOU[2], WENBO LIAO[2], AND QIANG FAN[2,4]

[2]State Key Laboratory of Biocontrol and Guangdong Key Laboratory of Plant Resources, Sun Yat-sen University, Guangzhou 510275, People's Republic of China; and [3]Rescue Center of Wildlife in Shenzhen, Shenzhen 518040, People's Republic of China

- *Premise of the study:* To examine population differentiation, simple sequence repeat (SSR) markers were developed in *Barthea barthei*, a shrub with a disjunct distribution in the southern mainland of China and Taiwan.
- *Methods and Results:* We used Illumina HiSeq technology to sequence a genomic library for SSR identification. Twenty-seven SSR loci were developed, of which 18 SSR loci were polymorphic in three populations composed of two varieties of *B. barthei*. At the population level, the number of alleles ranged from one to seven, and the observed and expected heterozygosity varied from 0 to 0.850 and from 0 to 0.809, respectively. Higher genetic differentiation between the two populations of *B. barthei* var. *barthei* (F_{ST} = 0.474) was observed relative to the two varieties (F_{ST} = 0.387 and 0.418, respectively).
- *Conclusions:* These polymorphic SSR markers may be useful for understanding phytogeographic history of *B. barthei*. Lower genetic differentiation between the two varieties than between the two populations of *B. barthei* var. *barthei* suggests that the taxonomic treatment may not hold.

Key words: *Barthea barthei*; disjunct distribution; genetic diversity; Melastomataceae; simple sequence repeat (SSR) marker.

Barthea Hook. f. (Melastomataceae) is a monotypic genus endemic to southern China. The only species, *B. barthei* (Hance ex Benth.) Krasser, is an evergreen shrub and has a disjunct distribution in southern mainland China (Guangdong Province, Guangxi Province, Fujian Province, Hunan Province, and Hong Kong) and Taiwan (Chen, 1984; Chen and Renner, 2007). There are two varieties for this species, *B. barthei* var. *barthei* and *B. barthei* var. *valdealata* C. Hansen. While *B. barthei* var. *barthei* is found throughout the range of the species, the variety *B. barthei* var. *valdealata* is confined to Shangsi County, Guangxi Province. The two varieties differ mainly in the width of the capsule wings (1 mm wide for *B. barthei* var. *barthei* vs. 2 mm for *B. barthei* var. *valdealata* [Chen, 1984; Chen and Renner, 2007]). Both varieties occur on forested mountain slopes at 500–1500 m elevation.

Patterns of disjunct distributions and their formative mechanisms have long been an important topic in the field of phytogeography. Disjunct distributions of plants can be used to reveal the relationships between floras of two or more regions. There are 35 genera of seed plants that have a southern mainland China–Taiwan disjunct distribution (Chen et al., 2012); however, when and how the disjunct distribution of these plants formed remains elusive (Chen et al., 2012; Ye et al., 2012). As a typical species with a disjunct distribution in southern mainland China and Taiwan, *B. barthei* can be used to address this phytogeographic question. Moreover, because the trait used to distinguish the two varieties of *B. barthei* shows substantial variation, the taxonomic treatment of them as two varieties is doubtful. Molecular data may help resolve these evolutionary or taxonomic questions. However, to our knowledge, there have been no molecular markers developed for *B. barthei* so far. In this study, we developed and characterized 27 nuclear simple sequence repeat (SSR) markers for *B. barthei* using paired-end reads (250 bp) generated using an Illumina HiSeq 2500 system (Illumina, San Diego, California, USA).

METHODS AND RESULTS

We sampled 20 individuals from each of three natural populations, namely, Nanling in Ruyuan, Guangdong Province (NA); Sanzhoutian in Shenzhen, Guangdong Province (SA); and Shiwandashan in Shangsi, Guangxi Province (SH) (Appendix 1). The SH population represents *B. barthei* var. *valdealata*, while the other two populations represent *B. barthei* var. *barthei*. Genomic DNA was isolated from silica-dried leaves using the cetyltrimethylammonium bromide (CTAB) method (Doyle and Doyle, 1987). We first constructed a genomic DNA library with 400-bp inserts from an individual of *B. barthei* collected in SA. The genomic DNA library was sequenced using an Illumina HiSeq 2500 system (approximately 1/10 Illumina lane) at Berry Genomics (Beijing, China). Initial quality filtering was performed with the sequencing company's in-house script *fastqc_adapter_pe*. Reads with (1) adapters, (2) >10% ambiguous base calls (Ns), or (3) >50% bases ≤5 in the Phred quality score were removed. A total

[1]Manuscript received 22 November 2016; revision accepted 14 February 2017.

This study was financially supported by the Guangdong Natural Science Foundation (2015A030302011), the Urban Management Bureau of Shenzhen Municipality (71020106 and 71020140), the Innovation of Science and Technology Commission of Shenzhen Municipality (JCYJ20150624165943509), and the Chang Hungta Science Foundation of Sun Yat-sen University.

[4]Author for correspondence: fanqiang@mail.sysu.edu.cn

doi:10.3732/apps.1600149

of 17.58 million 250-bp clean paired-end reads were obtained. The reads were then deposited in the National Center for Biotechnology Information (NCBI) Sequence Read Archive (accession no. SRR5130485). Each pair of reads was assembled into contigs using mothur version 1.37.6 (Schloss et al., 2009) with the parameters minlength = 350, maxlength = 500, minoverlap = 50, and mismatches = 2. We then clustered these contigs using CD-HIT version 4.6 (Li and Godzik, 2006) with the minimum identity of 98%. A total of 4,572,771 clusters were generated. When a cluster included more than three members, the longest one was used to seek SSR motifs. Using MISA software (Thiel et al., 2003), a total of 2535 SSR loci were detected from the 124,768 clusters with more than three members. Primer pairs for the 90 SSR loci with the longest dinucleotide repeats were designed with Primer3 (Rozen and Skaletsky, 1999).

To screen the SSR primers, we conducted PCR amplification using one individual each from the SA and SH populations in a final reaction volume of 20 μL containing ~20 ng DNA, 10 μL 10× PCR buffer with Mg^{2+}, 0.25 mM dNTPs, 1 μM each of forward and reverse primer, and 1 unit *EasyTaq* DNA polymerase (TransGen Biotech, Beijing, China). PCR was conducted for all primers using the following cycling program: 4 min of denaturation at 94°C; followed by 30 cycles of 40 s at 94°C, 40 s at the annealing temperature of each primer pair, and 60 s at 72°C; with a final extension of 8 min at 72°C. The PCR products were run in a 1.2% agarose gel to see if the expected size was obtained for each primer pair. Twenty-seven primer pairs designed for *B. barthei* produced single-band PCR products with the expected size (Table 1).

Table 1. Characteristics of 27 SSR markers developed for *Barthea barthei.*

Locus	Primer sequences (5′–3′)	Repeat motif	Allele size range (bp)	T_a (°C)	GenBank accession no.
BB01	F: CCTGGTTCAGTACAACCTGGG	$(TC)_{34}$	160–214	56	KY091625
	R: TGCCAGGGCCTACAATGAAG				
BB02	F: AACGTCTCCTCTCCCAGTCA	$(TC)_{28}$	174–196	58	KY091626
	R: CACCCTGACCTGACGAGAAC				
BB03	F: AGGTGGAGGATCAGGACATCA	$(TC)_{23}$	208–212	58	KY091627
	R: GTCTTGCTCCCCTCTAAAGCA				
BB04	F: GCTTGCCTTGGAGAGGTTCT	$(GA)_{22}$	200–216	58	KY091628
	R: GCAGGAACATCAGAGAAGAGCT				
BB05	F: ATGGCTCTTGATGCAGGCTT	$(AG)_{21}$	166–170	56	KY091629
	R: GAAGCGGTGTTGTGCATCTC				
BB06	F: GGAGCTCCGATAAACCAGCA	$(CT)_{20}$	200–204	58	KY091630
	R: ACAGAGCAACACCTCGCTTT				
BB07	F: TGTGACTCTTCGTTTCGTTCA	$(TC)_{11}$	204–214	58	KY091631
	R: GCTCGGATAGAGACTCGTGC				
BB08	F: GCAGCAGCTTCGGTAAGAAT	$(CT)_{11}$	238–260	58	KY091632
	R: GTATTGTCGCTTCACGCAAA				
BB09	F: ATCCCAAATGCTTCACAACC	$(AG)_{11}$	250–264	58	KY091633
	R: TCCCCTCATCATCTCTGCTT				
BB10	F: GGAGGTTTGTGACATTGCCT	$(AC)_{11}$	264–276	56	KY091634
	R: CAGAAGGGTCTTGCCAGTTC				
BB11	F: ACAAAACCCTTATCCCCACC	$(TG)_9$	194–202	56	KY091635
	R: TGTTATGACCCCATCCCCTA				
BB12	F: TGCTGCCCCTTTTCATATTT	$(AC)_9$	266–270	58	KY091636
	R: AGGTGCCCTGAGTAATGACC				
BB13	F: GGATGCTATGGAAAAAGCCA	$(AT)_8$	210–212	58	KY091637
	R: TTCAGGATCCTTTTTGGTGG				
BB14	F: AAAGAATTGGGGGTCCAAAC	$(AG)_8$	248–264	58	KY091638
	R: TCTTCACTTCACCAGCCCTT				
BB15	F: GAAACTAACTCCCCGCACAA	$(AG)_8$	204–210	58	KY091639
	R: ACGACGTTGGATGGTTTCTC				
BB16	F: CGGAAAAGTTTTGCACCAAT	$(AC)_8$	280–286	58	KY091640
	R: AGTAATTGCCGGAGGTGTTG				
BB17	F: GCTTCCTTGCTTTGATTTGG	$(AG)_7$	280–328	58	KY091641
	R: ATGGCTGAAATGCAGGAACT				
BB18	F: TCATGGGCTTCGCATATACA	$(AC)_7$	256–260	56	KY091642
	R: GAAATGCGTCGGTTTTGATT				
BB19	F: TCATGGGCTTCGCATATACA	$(TA)_7$	272	55	KY798206
	R: GAAATGCGTCGGTTTTGATT				
BB20	F: ACGTCTCGGAACGGAATATG	$(GA)_7$	254	55	KY798207
	R: GAGTAGCTGGATCAGAGCGG				
BB21	F: GAACCCATCCATTGGAAGAA	$(CT)_{20}$	204	56	KY798208
	R: TGCACATTGTTTTTCTGGTCTT				
BB22	F: GGAGCTCCGATAAACCAGCA	$(AT)_{10}$	254	56	KY798209
	R: ACAGAGCAACACCTCGCTTT				
BB23	F: GGTTTGATTTTCCATCTTTTTGA	$(AG)_{21}$	190	56	KY798210
	R: CTCGCGTGTCTACCATTTCA				
BB24	F: CGCCAGCGCCAATAGTAGTA	$(AT)_{10}$	200	56	KY798211
	R: CAAGGAGAGCCCCGAAACTT				
BB25	F: GTTCGCAGAAGCTGACACAA	$(AT)_{10}$	202	56	KY798212
	R: CCATAAAGGAGCCGAATGAA				
BB26	F: AATTATCCGGAGGATGCCTT	$(AT)_{10}$	200	56	KY798213
	R: CCCTACGTCGTTACATTCCG				
BB27	F: CCATAAAGGAGCCGAATGAA	$(AT)_{10}$	206	56	KY798214
	R: GTTCGCAGAAGCTGACACAA				

Note: T_a = annealing temperature.

TABLE 2. Genetic diversity at 18 SSR loci in three natural populations of *Barthea barthei*.[a]

	SH (N = 20)			NA (N = 20)			SA (N = 20)		
Locus	A	H_o	H_e[b]	A	H_o	H_e[b]	A	H_o	H_e[b]
BB01	4	0.400	0.676	7	0.500	0.809**	1	0.000	0.000
BB02	3	0.550	0.515	4	0.600	0.711	2	0.400	0.375
BB03	2	0.200	0.255	2	0.050	0.139**	2	0.300	0.255
BB04	2	0.200	0.375*	3	0.550	0.411	2	0.200	0.180
BB05	3	0.650	0.579	2	0.150	0.219	2	0.200	0.180
BB06	2	0.500	0.395	2	0.250	0.434	2	0.700	0.480*
BB07	2	0.700	0.455*	2	0.150	0.411*	2	0.300	0.329
BB08	1	0.650	0.439*	2	0.400	0.386	1	0.400	0.515**
BB09	1	0.250	0.219	2	0.250	0.501**	2	0.000	0.332**
BB10	5	0.600	0.765	2	0.050	0.049	3	0.000	0.335**
BB11	1	0.000	0.000	4	0.550	0.640**	2	0.850	0.489**
BB12	2	0.150	0.219	1	0.000	0.000	1	0.000	0.000
BB13	2	0.500	0.375	2	0.118	0.111	1	0.000	0.000
BB14	1	0.000	0.000	2	0.250	0.219	2	0.050	0.500**
BB15	1	0.000	0.000	1	0.000	0.000	2	0.150	0.139
BB16	1	0.000	0.000	1	0.000	0.000	1	0.100	0.095
BB17	1	0.000	0.000	1	0.350	0.289	1	0.100	0.096
BB18	1	0.000	0.000	2	0.100	0.095	1	0.053	0.051

Note: A = number of alleles; H_e = expected heterozygosity; H_o = observed heterozygosity; N = number of sampled individuals from each population.

[a] Population and locality information are provided in Appendix 1.

[b] Significant deviations from Hardy–Weinberg equilibrium after sequential Bonferroni corrections: * represents significance at the 5% nominal level; ** represents significance at the 1% nominal level.

To test the polymorphism level of the 27 primer pairs, PCR was conducted for all of the samples of *B. barthei* using the conditions mentioned above. We labeled the forward primers of the 27 primer pairs with the fluorescent dye FAM or HEX. Using GeneScan 500 ROX (Applied Biosystems, Waltham, Massachusetts, USA) as an internal size standard, we determined the fragment sizes of the PCR products on an ABI 3100 DNA Sequencer with Genotyper 4.0 (Applied Biosystems). Parameters of genetic diversity in each population and genetic differentiation among populations of *B. barthei* were calculated using GenAlEx version 6.5 (Peakall and Smouse, 2012).

Our results showed that 18 SSR markers were polymorphic in *B. barthei* (Table 2), and two to nine alleles were detected for these SSR loci at the species level. The observed heterozygosity of these polymorphic loci varied from 0 to 0.850, and the expected heterozygosity varied from 0 to 0.809. Of the 18 polymorphic SSR loci, three, five, and six in the SH, NA, and SA populations, respectively, exhibited significant deviations from Hardy–Weinberg equilibrium (Table 2). A measure of pairwise genetic differentiation between populations (F_{ST}) indicated that genetic differentiation between NA and SA populations was the highest (F_{ST} = 0.474), while lower genetic differentiation was observed between SH and NA populations (F_{ST} = 0.418) and between SH and SA populations (F_{ST} = 0.387). Therefore, our SSR data showed that genetic differentiation between the two varieties is lower than that between the two populations of *B. barthei* var. *barthei*.

CONCLUSIONS

This is the first set of molecular markers developed for *B. barthei*, an evergreen shrub with a disjunct distribution in southern mainland China and Taiwan. The 18 polymorphic markers may be useful for phytogeographic studies of *B. barthei* to reveal the formative mechanisms of the southern mainland China–Taiwan disjunct distribution. Lower differentiation between the two varieties than between allopatric populations of the variety *B. barthei* var. *barthei* suggests that the taxonomic division of *B. barthei* as two varieties may not hold.

LITERATURE CITED

CHEN, J. 1984. Melastomataceae. *In* C. Chen, H. Chang, R. Miau, and T. Hsu [eds.], Flora Reipublicae Popularis Sinicae, vol. 53(1), 152–162. Science Press, Beijing, China.

CHEN, J., AND S. S. RENNER. 2007. Melastomataceae. *In* Z. Y. Wu and P. H. Raven [eds.], Flora of China, vol. 13, 360–399. Science Press, Beijing, China, and Missouri Botanical Garden Press, St. Louis, Missouri, USA.

CHEN, Z. D., J. S. YING, AND A. M. LU. 2012. Disjunct distribution of seed plants between southwestern China and Taiwan island of China. *Chinese Bulletin of Botany* 47: 551–570.

DOYLE, J. J., AND J. L. DOYLE. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* 19: 11–15.

LI, W., AND A. GODZIK. 2006. Cd-hit: A fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics (Oxford, England)* 22: 1658–1659.

PEAKALL, R., AND P. E. SMOUSE. 2012. GenAlEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research–An update. *Bioinformatics (Oxford, England)* 28: 2537–2539.

ROZEN, S., AND H. SKALETSKY. 1999. Primer3 on the WWW for general users and for biologist programmers. *In* S. Misener and S. A. Krawetz [eds.], Methods in molecular biology, vol. 132: Bioinformatics: Methods and protocols, 365–386. Humana Press, Totowa, New Jersey, USA.

SCHLOSS, P. D., S. L. WESTCOTT, T. RYABIN, AND H. MARTIN. 2009. Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology* 75: 7537–7541.

THIEL, T., W. MICHALEK, R. K. VARSHNEY, AND A. GRANER. 2003. Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). *Theoretical and Applied Genetics* 106: 411–422.

YE, J., Z. CHEN, B. LIU, H. QIN, AND Y. YANG. 2012. Disjunct distribution of vascular plants between southwestern area and Taiwan area in China. *Biodiversity Science* 20: 482–494.

Appendix 1. Sampling information of *Barthea barthei* in this study. Vouchers were deposited at the Herbarium of Sun Yat-sen University, Guangzhou, Guangdong, China.

Taxon	Population code	Collection locality	Voucher information	Geographic coordinates	Altitude (m)	*N*
Barthea barthei (Hance ex Benth.) Krasser var. *barthei*	NA	Nanling, Ruyuan, Guangdong, China	*Q. Fan 14047*	24°52′33″N, 113°01′43″E	1128	20
Barthea barthei var. *barthei*	SA	Sanzhoutian, Shenzhen, Guangdong,China	*Q. Fan 14012*	22°37′31″N, 114°16′02″E	401	20
Barthea barthei var. *valdealata* C. Hansen	SH	Shiwandashan, Shangsi, Guangxi, China	*Q. Fan 14033*	21°52′59″N, 107°54′52″E	713	20

Note: *N* = number of individuals sampled.