Isolation and Characterization of 28 Microsatellite Loci for a Korean Endemic, Lespedeza maritima (Fabaceae)

Authors: Jin, Dong-Pil, Cho, Won-Bum, Choi, In-Su, and Choi, Byoung-Hee

Source: Applications in Plant Sciences, 4(1)

Published By: Botanical Society of America

URL: https://doi.org/10.3732/apps.1500089

Applications in Plant Sciences 2016 4(1): 1500089

PRIMER NOTE

ISOLATION AND CHARACTERIZATION OF 28 MICROSATELLITE LOCI FOR A KOREAN ENDEMIC, *LESPEDEZA MARITIMA* (FABACEAE)[1]

DONG-PIL JIN[2], WON-BUM CHO[2], IN-SU CHOI[2], AND BYOUNG-HEE CHOI[2,3]

[2]Department of Biological Sciences, Inha University, Incheon 402-751, Republic of Korea

- *Premise of the study:* We developed microsatellite primers for *Lespedeza maritima* (Fabaceae), a Korean endemic shrub, and conducted cross-amplifications for closely related species.
- *Methods and Results:* We produced 28 polymorphic microsatellite markers through reference mapping of 300-bp paired-end reads obtained from Illumina MiSeq data. For 47 individual plants from two populations, the total alleles numbered two to 13, and observed and expected heterozygosities ranged from 0.067 to 0.867 and from 0.064 to 0.848, respectively. Most of these markers were well amplified in closely related species.
- *Conclusions:* In future research, the microsatellite markers described here will help reveal the taxonomic entity of this species.

Key words: endemic species; Fabaceae; hybridization; *Lespedeza maritima*; microsatellite; taxonomic entity.

Plants of species within *Lespedeza* Michx. sect. *Macrolespedeza* Maxim. (Fabaceae) are deciduous shrubs mainly distributed in the temperate forests of East Asian countries such as Korea, China, and Japan (Ohashi and Nemoto, 2014). Based on morphological and molecular studies (Lee, 1965; Akiyama, 1988; Han et al., 2010; Xu et al., 2012), hybridization is frequent among these species. Therefore, their taxonomic entities are often unclear.

One representative, *L. maritima* Nakai, is a Korean endemic that primarily occurs at low elevations on mountains along the southern coast of the Korean Peninsula. This species is distinguished from others within the genus by its unique morphological traits, e.g., leaflets with leathery surfaces and revolute margins, as well as abundant sprouting from the basal portion of the plant. Despite these distinctive characteristics, the taxonomic entity of *L. maritima* has been controversial. First described by Nakai (1923), this species has since then been regarded either as a hybrid between *L. cyrtobotrya* Miq. and *L. maximowiczii* C. K. Schneid. (Lee, 1965) or as a synonym of *L. thunbergii* (DC.) Nakai subsp. *thunbergii* (= *L. formosa* (Vogel) Koehne subsp. *velutina* (Nakai) S. Akiyama & H. Ohba) (Akiyama, 1988), thus making its boundary of taxonomic delimitation more ambiguous. Previous phylogenetic analysis (Xu et al., 2012) indicated low resolution for chloroplast DNA (cpDNA) and nuclear ribosomal internal transcribed spacer (nrITS) sequences among species. In plants such as those of *Quercus* L., codominant polymorphic microsatellite markers have been used to confirm that hybridization is frequent among those species (see, e.g., Lee et al., 2014). Advances in sequencing techniques, i.e., next-generation sequencing (NGS), have made it possible for researchers to obtain microsatellite libraries rapidly. Thus, microsatellite markers could be applied to resolve the above-mentioned debate about *L. maritima* and other *Lespedeza* species if they were available for that genus. Here, we used NGS to develop 28 microsatellite markers to clarify the taxonomic entity of *L. maritima*.

METHODS AND RESULTS

We sampled 47 individuals of *L. maritima* from Bogil Island (n = 32; 34°08′16″N, 126°31′49″E) and Namhae Island (n = 15; 34°44′33″N, 128°01′20″E), located in the South Sea of Korea. To avoid repeated collections from the same plant, we selected individuals that grew at least 5 m apart. The leaves were preserved with silica gel, and voucher specimens were deposited in the Herbarium of Inha University (IUI), Incheon, Korea (Appendix 1). Genomic DNA, used to acquire the microsatellite library, was extracted from the silica gel–dried leaves of one individual (voucher no. DP149122) with a DNeasy Plant Mini Kit (QIAGEN, Seoul, Korea). The extracted DNA (300 ng) was fragmented to 500 bp using a Covaris S220 (Covaris, Woburn, Massachusetts, USA). A TruSeq Nano DNA Library Preparation Kit (Illumina, San Diego, California, USA) was used for library preparation after sequencing on the Illumina MiSeq platform at Life Is Art of Science (LAS; Daejeon, Korea; http://lascience.co.kr/). In all, 5,379,227 pairs for 10,758,454 reads (300 × 300) were generated. Through SSR_pipeline version 0.951 (Miller et al., 2013), we selected the microsatellite library containing di- or trinucleotide repeats for which at least 10 or five, respectively, were obtained from total reads. Of the obtained microsatellite library, we initially discarded those that had flanking regions of less than 100 bp and more than 20 repeats. Then we attempted reference mapping of total paired reads to each remaining sequence using Geneious R 7.1.8 (Biomatters, Auckland, New Zealand). For this process, putative multicopy loci with exceptionally high coverage were removed, while certain other loci were selected based on the following criteria: (1) loci showing unique patterns that had two separate alleles and few variations at the site to which a primer was attached, and (2) no additional single nucleotide polymorphisms (SNPs) in the flanking region. From those, we designed 65 primer pairs with Primer3 (Rozen and Skaletsky, 1999) and added the M13(–21) sequence (5′-TGTAAAACGACGGCCAGT-3′) to the 5′ end of all forward primers intended for labeling

[1]Manuscript received 6 August 2015; revision accepted 15 September 2015.

This work was supported by funding for scientific research (grant KNA1-2-13, 14-2) from the Korea National Arboretum.

[3]Author for correspondence: bhchoi@inha.ac.kr

doi:10.3732/apps.1500089

TABLE 1. Characteristics of 28 microsatellite loci developed in *Lespedeza maritima*.

Locus	Primer sequences (5′–3′)	Repeat motif	Fluorescent dye[a]	Allele size range (bp)
LMS3	F: GAGGACAAATTAATCACAGC	$(AG)_{11}$	6-FAM	186–210
	R: GGAAAGAGTAGTCAAAGCA			
LMS5	F: TATAAGCCATTCAGAGGGA	$(AG)_{10}$	VIC	208–226
	R: GGCATAGCTGTAGTAAACA			
LMS6	F: CATAACAGTGTTGATAACCC	$(AG)_{11}$	NED	246–266
	R: TCAGTCGAGAAATATCTACG			
LMS7	F: CTAGCATTTCAATCAGTGG	$(AG)_{10}$	PET	231–243
	R: CTTGCTGCTATTGACTACA			
LMS8	F: CGAGTTCGAAACAGGATTA	$(AG)_{10}$	PET	258–270
	R: ATGTAGGAGGAGATGGTAC			
LMS11	F: ATAACCGTACAAAGGCAAG	$(AG)_{10}$	6-FAM	208–248
	R: CAGCCTGAATTCTCAAATC			
LMS17	F: TCATGAAAGTTAGTGTGGG	$(CT)_{10}$	6-FAM	305–327
	R: CATCATTGCAATCTGAGTG			
LMS18	F: GCTTATTGTTCTACCGAGT	$(CT)_{10}$	PET	210–226
	R: CATAGTTGTCCTCAAGAGAA			
LMS20	F: GTGATTGATATGGATGCCT	$(CT)_{11}$	VIC	197–217
	R: GAACAAGAAACAGAACACC			
LMS21	F: GGTACCCTTAACACGTAAT	$(CA)_{10}$	PET	291–299
	R: CACTTACTTACCCTAAGGC			
LMS22	F: ATTCCACTTTGGTCTGAAG	$(CT)_{11}$	6-FAM	241–251
	R: TACTCGGTAATAAAGCACG			
LMS24	F: CATGAATTGGGAGATGAGT	$(TC)_{12}$	NED	282–314
	R: CCTGGGTTATGCTAATTGA			
LMS25	F: TACTCTTCTTATTTGGCGG	$(TG)_{10}$	VIC	261–279
	R: CCTGAGAGGGAAAGAACTA			
LMS28	F: ACTAGCCACATACTCTCAA	$(GT)_{11}$	6-FAM	289–305
	R: GTAGGCATGCTAAGAATCT			
LMS33	F: CATAGCAATTGGACCTCAT	$(TC)_{11}$	VIC	283–303
	R: GTTGAGATCGGAAAGTGTA			
LMS34	F: GATCAAAGCAGTTTCAAGG	$(TC)_{10}$	NED	208–220
	R: GGAACAATACTCCCAAACA			
LMS39	F: CGACACACTTTCCAACTAT	$(TC)_{12}$	NED	309–327
	R: ACATAGAAGCTGTCATGAC			
LMS45	F: GAATTGACAAGTGGTGAAC	$(GCT)_6$	VIC	292–322
	R: TGGAGAAATTGTGAGTAGC			
LMS47	F: GCTCTTCACGAAATCAATC	$(CAT)_7$	PET	307–328
	R: CATTCGCATCTCAAAGAAG			
LMS48	F: GTGCATGCTAAGTTATCCA	$(ATC)_5$	VIC	169–176
	R: AGTCACACTCAGATGGGT			
LMS49	F: TATTCCATTCCTTGAACCC	$(CCA)_5$	6-FAM	275–281
	R: CAACATTGGCATATAGGGA			
LMS52	F: ATCTGAATATGTGCCTGAG	$(GCG)_7$	NED	231–252
	R: CTACCTTAACCCTCTTCCT			
LMS53	F: TGTTCCCTAAAGACCAAAG	$(ATC)_6$	PET	175–178
	R: ACACCACCTTACACATTAC			
LMS55	F: CTTAGAGGAGCCACATAAG	$(CAG)_6$	NED	171–186
	R: CATACCCACCAGCATTAC			
LMS57	F: AATAAAGCACCTCCATCTC	$(CAT)_6$	6-FAM	222–225
	R: CAATTTGGCCAACTATGAC			
LMS58	F: GTACATAATTTCGCACACC	$(ATC)_8$	VIC	248–260
	R: CTGGGTTCAAGTCTAAGTC			
LMS61	F: CTCTGATTCGTACACTGAA	$(ATG)_5$	VIC	227–233
	R: GAACGTCAATAGTCCTTCA			
LMS62	F: GAGCTTGGAATTGGTTACT	$(AGC)_5$	6-FAM	172–178
	R: GCATTTCCCATTAGAGTCT			

[a]Fluorescent dye used for fragment analysis.

with a fluorochrome. To test for amplification and polymorphism of each primer pair, we conducted PCR for 47 individuals of the two populations with a GeneAmp PCR System 2700 Thermal Cycler (Applied Biosystems, Foster City, California, USA). Each reaction mixture (10 µL total volume) contained 5 ng of DNA plus 5 µL of 2× Plus Mix (Dongsheng Biotech, Guangdong, China), which comprised 0.4 mM dNTPs, 2× PCR buffer with 4 mM $MgSO_4$, and 0.4 U·µL^{-1} of *Taq* DNA polymerase. The mixtures also contained the appropriate 0.08 µM forward M13(–21)-tagged primer, a 0.3 µM reverse primer, and a 0.3 µM M13(–21) labeled fluorescent marker (NED, PET, VIC, or 6-FAM). Conditions included initial denaturation at 94°C for 3 min; then 30 cycles at 94°C for 30 s, 53°C for 45 s, and 72°C for 45 s; with a final extension at 72°C for 7 min. Afterward, the PCR products were visualized on 2% agarose gels and resolved to genotype on an ABI 3730XL sequencer with GeneScan 500 LIZ Size Standard (Applied Biosystems). The sizes of the alleles were determined with GeneMapper 3.7 (Applied Biosystems). Using GenAlEx 6.5 (Peakall and Smouse, 2006), we estimated the total number of alleles (A_T) and number of alleles (A), plus values for observed heterozygosity (H_o), expected heterozygosity (H_e), and deviation from Hardy–Weinberg equilibrium (HWE)

TABLE 2. Genetic diversity values for 47 individuals of *Lespedeza maritima* across 28 microsatellite loci.

	Bogil Island (n = 32)			Namhae Island (n = 15)			Total (n = 47)
Locus	A	H_o	H_e[a]	A	H_o	H_e[a]	A_T
LMS3	9	0.625	0.785	4	0.533	0.744	10
LMS5	7	0.613	0.782**	6	0.667	0.762	8
LMS6	4	0.344	0.322	4	0.333	0.567	5
LMS7	4	0.344	0.297	5	0.467	0.438	5
LMS8	6	0.375	0.606*	2	0.467	0.464	6
LMS11	7	0.813	0.799*	6	0.200	0.756***	8
LMS17	6	0.406	0.515	3	0.357	0.472**	7
LMS18	5	0.344	0.723***	6	0.733	0.751	7
LMS20	6	0.688	0.715	8	0.800	0.811	9
LMS21	5	0.688	0.669	4	0.733	0.696	5
LMS22	5	0.500	0.500	4	0.400	0.651	5
LMS24	11	0.533	0.731***	9	0.400	0.822***	13
LMS25	8	0.533	0.747	8	0.714	0.758	9
LMS28	8	0.781	0.739	3	0.733	0.587	8
LMS33	9	0.710	0.848	8	0.533	0.836*	11
LMS34	6	0.844	0.764	4	0.400	0.560	7
LMS39	10	0.750	0.750	9	0.867	0.804	10
LMS45	8	0.844	0.763***	5	0.429	0.704*	8
LMS47	7	0.594	0.682	7	0.643	0.760	8
LMS48	3	0.375	0.310	3	0.533	0.464	3
LMS49	3	0.500	0.538	3	0.143	0.446**	4
LMS52	7	0.258	0.652***	4	0.308	0.648	8
LMS53	2	0.125	0.170	2	0.231	0.453	2
LMS55	5	0.531	0.612	3	0.600	0.540	6
LMS57	2	0.375	0.430	2	0.533	0.444	2
LMS58	4	0.375	0.359	3	0.533	0.527	4
LMS61	2	0.375	0.342	2	0.067	0.064	3
LMS62	3	0.355	0.599*	3	0.538	0.506	3

Note: A = number of alleles; A_T = total number of alleles; H_e = expected heterozygosity; H_o = observed heterozygosity; n = number of individuals sampled.
[a] Significant deviation from Hardy–Weinberg equilibrium: *P < 0.05, **P < 0.01, ***P < 0.001.

TABLE 3. Results of cross-amplification for the 28 microsatellite markers for *Lespedeza maritima* in four related species.

Locus	*L. bicolor*	*L. cyrtobotrya*	*L. maximowiczii*	*L. thunbergii* subsp. *thunbergii*
LMS3	+	+	+	+
LMS5	+	+	~	+
LMS6	+	+	+	+
LMS7	+	+	+	+
LMS8	~	+	–	~
LMS11	+	+	+	+
LMS17	+	+	+	+
LMS18	+	+	+	+
LMS20	+	+	+	+
LMS21	+	+	+	+
LMS22	+	+	+	+
LMS24	+	+	+	+
LMS25	+	+	+	+
LMS28	+	+	+	+
LMS33	+	+	~	+
LMS34	~	+	~	+
LMS39	–	–	–	+
LMS45	–	–	~	~
LMS47	+	+	+	+
LMS48	+	+	+	+
LMS49	+	+	+	+
LMS52	+	+	+	+
LMS53	+	+	+	+
LMS55	+	+	+	+
LMS57	+	+	+	+
LMS58	+	+	+	+
LMS61	+	+	+	+
LMS62	+	+	+	+

Note: + = successful amplification; – = failed amplification; ~ = successful amplification in only one individual sampled.

for each population. Null allele frequencies were calculated by MICRO-CHECKER version 2.2 (van Oosterhout et al., 2004). Additionally, we tested cross-amplification with four related species (*L. bicolor* Turcz., *L. cyrtobotrya*, *L. thunbergii* subsp. *thunbergii*, and *L. maximowiczii*) in successfully amplified primers, using two individuals per species. Among those species, the first three were sampled at Baegun Mountain (37°29′32″N, 126°30′58″E) while *L. maximowiczii* was collected at Biseul Mountain (35°42′16″N, 128°31′58″E). Sequences of the raw paired-end reads, successfully amplified and scored, are available in the SRA database of the National Center for Biotechnology Information (NCBI) (Bioproject no. PRJNA290063).

Of the 65 initially designed primer pairs, 45 were successfully amplified for 47 individuals from our two populations. Among those 45 primer pairs, 17 showed scoring errors. Finally, we selected 28 pairs, all of which were polymorphic (Table 1). The values for genetic diversity for the 47 individuals of *L. maritima* were as follows (Table 2): A_T varied from two to 13; H_o and H_e per locus ranged from 0.067 to 0.867 and from 0.064 to 0.848, respectively; eight loci (LMS5, LMS8, LMS11, LMS18, LMS24, LMS45, LMS52, and LMS62) showed significant deviation from HWE on Bogil Island, while six loci (LMS11, LMS17, LMS24, LMS33, LMS45, and LMS49) showed significant deviation from HWE on Namhae Island. Evidence of null alleles detected at most loci (except for LMS17) indicated significant deviations from HWE. Therefore, this lack of equilibrium could be explained by the presence of those null alleles. Our results from the cross-amplification showed that all loci, except LMS5, LMS8, LMS33, LMS34, LMS39, and LMS45, were well amplified for the four related species (Table 3).

CONCLUSIONS

We have developed a set of 28 polymorphic microsatellite markers for *L. maritima*. Most are well amplified in four related species within section *Macrolespedeza*: *L. bicolor*, *L. cyrtobotrya*, *L. thunbergii* subsp. *thunbergii*, and *L. maximowiczii*. Future studies will use these new markers as tools to reveal the taxonomic entity of *L. maritima*.

LITERATURE CITED

Akiyama, S. 1988. A revision of the genus *Lespedeza* section *Macrolespedeza* (Leguminosae). *University Museum, University of Tokyo, Bulletin* 33: 1–170.

Han, J. E., K. H. Chung, T. Nemoto, and B. H. Choi. 2010. Phylogenetic analysis of eastern Asian and eastern North American disjunct *Lespedeza* (Fabaceae) inferred from nuclear ribosomal ITS and plastid region sequences. *Botanical Journal of the Linnean Society* 164: 221–235.

Lee, J. H., D. P. Jin, and B. H. Choi. 2014. Genetic differentiation and introgression among Korean evergreen *Quercus* (Fagaceae) are revealed by microsatellite markers. *Annales Botanici Fennici* 51: 39–48.

Lee, T. B. 1965. The *Lespedeza* of Korea (1). *Bulletin of Seoul National University Forests* 2: 1–43.

Miller, M. P., B. J. Knaus, T. D. Mullins, and S. M. Haig. 2013. SSR_pipeline: A bioinformatic infrastructure for identifying microsatellites from paired-end Illumina high-throughput DNA sequencing data. *Journal of Heredity* 104: 881–885.

Nakai, T. 1923. Notulæ ad Plantas Japoniæ et Coreæ XXX. *Shokubutsugaku Zasshi* 37: 78.

Ohashi, H., and T. Nemoto. 2014. A new system of *Lespedeza* (Leguminosae tribe Desmodieae). *Journal of Japanese Botany* 89: 1–11.

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.

Rozen, S., and H. J. 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.

van Oosterhout, C., W. F. Hutchinson, D. P. Wills, and P. Shipley. 2004. MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* 4: 535–538.

Xu, B., X. F. Gao, and L. B. Zhang. 2012. Analysis of DNA sequences of six plastid and nuclear genes suggests incongruence, introgression, and incomplete lineage sorting in the evolution of *Lespedeza* (Fabaceae). *Molecular Phylogenetics and Evolution* 62: 346–358.

Appendix 1. Voucher and location information for *Lespedeza maritima* used in this study. Vouchers deposited at the herbarium of Inha University (IUI), Incheon, Korea.

Location	Geographic coordinates	Voucher no.	No. of individuals
Peak Gyeokja, Bogil-myeon, Wando-gun, Jeollanam-do, Korea	34°08′16″N, 126°31′49″E	DP149121–DP149152	32
Peak Gama, Mijo-myeon, Namhae-gun, Gyeongsangnam-do, Korea	34°44′33″N, 128°01′20″E	DP149017–DP149028, DP149030–DP149032	15