Korean J. Pl. Taxon. 48(3): 195–200 (2018)
https://doi.org/10.11110/kjpt.2018.48.3.195
RESEARCH ARTICLE

pISSN 1225-8318
eISSN 2466-1546
Korean Journal of
Plant Taxonomy

Isolation and characterization of EST-SSR markers for *Astilboides tabularis* (Saxifragaceae), endangered species in Korea

Eui-Kwon JUNG, Dae-Hyun KANG[1], Ki-Oug YOO[2], Myounghai KWAK[3], Young-Dong KIM and Bo-Yun KIM[1]*

Department of Life Science, Hallym University, Chuncheon 24252, Korea
[1]Multidisciplinary Genome Institute, Hallym University, Chuncheon 24252, Korea
[2]Department of Biological Sciences, Kangwon National University, Chuncheon 24341, Korea
[3]Plant Resources Division, National Institute of Biological Resources, Incheon 22689, Korea
(Received 4 September 2017; Revised 17 September 2018; Accepted 28 September 2018)

ABSTRACT: Genetic assessments of rare and endangered species are among the first steps necessary to establish the proper management of natural populations. Transcriptome-derived single-sequence repeat markers were developed for the Korean endangered species *Astilboides tabularis* (Saxifragaceae) to assess its genetic diversity. A total of 96 candidate microsatellite loci were isolated based on transcriptome data using Illumina pair end sequencing. Of these, 26 were polymorphic, with one to five alleles per locus in 60 individuals from three populations of *A. tabularis*. The observed and expected heterozygosity per locus ranged from 0.000 to 0.950 and from 0.000 to 0.741, respectively. These polymorphic transcriptome-derived simple sequence repeat markers would be invaluable for future studies of population genetics and for ecological conservation of the endangered species *A. tabularis*.

Keywords: *Astilboides tabularis*, endangered species, EST-SSR markers, next-generation sequencing, conservation

Astilboides tabularis (Hemsl.) Engler is the only species in *Astilboides* (Saxifragaceae) known to be distributed in a cluster in the forests of river valleys of northeastern Korea and China (Jintang and Cullen, 2001; The Angiosperm Phylogeny Group et al., 2016). This species is a protected wild plant classified as endangered wildlife grade II by the Ministry of the Environment due to the possibility of the extinction of the population and/or reduction in the number of individuals by climate change (Ministry of the Environment of Korea, 2014). *A. tabularis* is a potential horticultural plant as an ornamental species given its large leaves (approximately 1 m in diameter) and beautiful panicles (Belyaeva and Butenkova, 2016; Choi et al., 2016). It also has a long history of usage as a medicinal plant for diabetes (Liu et al., 2016). Due to biological conservation efforts and given the ecological importance of *A. tabularis,* genetic diversity analysis studies have been conducted using AFLP and isozymes (Ku et al., 2006; Lee, 2008).

To the best of our knowledge, no microsatellite markers have been developed thus far for *A. tabularis* for population studies. Population genetics research provides insight into conservation and management plans for rare and threatened species (Ottewell et al., 2016). To assess the genetic diversity of *A. tabularis*, we developed expressed sequence tag–simple sequence repeat (EST-SSR) markers. These have been used as a powerful molecular tool for genetic diversity studies of many plant species (Yan et al., 2016; Wang et al., 2017).

Materials and Methods

For the construction of the RNA library, the total RNA was extracted from leaves of individuals representing *A. tabularis* from three populations (Voucher No. NIBRVP0000655607) (Table 1). RNA was extracted using RNeasy kits, version 2.2 (Illumina, San Diego, CA, USA) following the manufacturer's

*Author for correspondence: wken416@hallym.ac.kr

Table 1. Voucher information for the *Astilboides tabularis* populations sampled in this study. Vouchers were deposited in the Herbarium of the National Institute of Biological Resources (KB) and in the Herbarium of Hallym University (HHU), Republic of Korea. To prevent illegal collection, we did not disclose the exact locations of the sites.

Population	Locality	No.	Voucher No. (Herbarium)
Korea-1	Bonghwachi, Jeongseon	20	*NIBRVP0000655607* (KB)
Korea-2	Geomyongso, Taebaek	20	*KBY2017273* (HHU)
China	Fusong Xian, Jilin Sheng	20	*NIBRVP0000655609* (KB)

No., number of individuals sampled.

instructions, and was subsequently used for TruSeq cDNA library preparation and high-throughput Illumina HiSeq 100 bp paired-end de novo transcriptome sequencing. The analysis results reads were obtained and assembled into 102,884 unigenes with 7,476,378,742 read numbers. The de novo transcriptome assembly of these reads was performed using the short read assembling program Trinity r20140717 (Haas et al., 2013) with the default parameters. Microsatellites were detected using MIcroSAtellite (MISA) version 1.0.0 (Thiel et al., 2003) with thresholds of ten repeat units for dinucleotide and five repeat units for tri-, tetra-, penta-, and hexanucleotides. MISA identified 38,598 simple sequence repeats (SSRs), of which 96 loci were selected depending on (1) the number of SSR repeats, (2) a PCR product size of 150–500 bp, (3) an annealing temperature range of 55–60°C, and (4) a minimum GC content of 50% for further testing of *A. tabularis*. The primer sets were designed to flank the microsatellite-rich regions with a minimum of eight repeats using the Primer3 program (Rozen and Skaletsky, 1999).

We sampled 60 *A. tabularis* individuals from three wild populations (Table 1). All samples included in this study were collected in accordance with the requirements of permission and support from relevant authorities. To avoid collecting clones, we specified a distance of at least 2 m among individuals within each population. The voucher specimens were deposited in the National Institute of Biological Resources Herbarium (KB) and in the Herbarium of Hallym University (HHU) in Korea (Table 1). The locations of the sites have been withheld to prevent illegal collection. The utility of the 96 microsatellite markers was confirmed by PCR from each population of *A. tabularis* in a total volume 25 μL, containing 2.5 μL of 10× Ex Taq buffer (TaKaRa Bio Inc., Otsu, Japan), 2 μL of 2.5 mM dNTPs, 0.01 μM each of a forward and reverse primer, 0.1 μL of TaKaRa Ex Taq DNA polymerase (5 units/μL) (TaKaRa Bio Inc.), and 5–10 ng of template DNA. All PCRs were performed in a GeneAmp PCR System 9700 thermocycler (Applied Biosystems, Carlsbad, CA, USA) using the following program: initial denaturation at 98°C for 5 min followed by 30 cycles of denaturation at 95°C for 1 min, annealing at annealing at 59°C for 1 min with an extension at 72°C for 1.5 min, and a final extension step at 72°C for 10 min. Fluorescently labeled (HEX, FAM) PCR products were analyzed by an automated sequencer (ABI 3730XL) with the GeneScan 500 LIZ Size Standard (Applied Biosystems). Genotyping was performed using GeneMapper 3.7 (Applied Biosystems), and peaks were scored manually by visual inspections. Finally, we identified 26 polymorphic markers based on genotyping data, and functional annotations for these markers were performed on a subset of ESTs with BLASTX scores (E-value < 1×10^{-10}) using the gene ontology database (Table 2). The genetic parameters of polymorphic loci were assessed by calculating the number of alleles (A), the observed heterozygosity (H_o), and the expected heterozygosity (H_e) using GenAlEx 6.5 (Peakall and Smouse, 2012). Degrees of deviation from the Hardy-Weinberg equilibrium (HWE) were estimated with ARLEQUIN 3.5 (Excoffier and Lischer, 2010).

Results and Discussion

The results of the genetic diversity of 26 polymorphic markers are shown in Table 3. Overall, the 26 microsatellite loci were polymorphic, with the number of alleles per locus ranging from one to five (average 2.218). The H_o and H_e values ranged from 0.000 to 0.950 and from 0.000 to 0.741, respectively (Table 3). Some polymorphic loci significantly deviated from HWE, though this was not consistent across populations.

This study describes the first assembly and characterization of the leaf transcriptome of *A. tabularis* using the Illumina paired-end sequencing method. Twenty-six polymorphic markers were successfully amplified, revealing polymorphisms in *A. tabularis*. This work can serve as a basis for further studies on the genetic diversity and structure of *A. tabularis* and can provide essential information for devising effective conservation strategies.

Table 2. Characteristics of the 26 polymorphic microsatellite loci developed from *Astilboides tabularis.*

Locus	Primer sequence (5′-3′)	Repeat motif	Allele size range (bp)	GenBank accession No.	Putative function [Organism]	*E*-value
AT01	F: TCTGCCCTGATTGCACTTCA R: TCTCTCCTCTGCGTCATTGC	$(AG)_6$	220–224	MH476462	Not found	-
AT02	F: CAGTGAGAGACAGTGGCCTT R: ACGCCAAAACGATTGTGGTT	$(AG)_6$	223–227	MH476463	Unnamed protein product, partial [*Vitis vinifera*]	1.00E-99
AT04	F: CAAGCCTGCCTTCATCTTGC R: TGTTCCGAGGGATGTTGTGG	$(AG)_7$	220–222	MH476464	Not found	-
AT07	F: CAGAGGTGCCCACTTGGAAT R:GCTGGGATGAGGTTTCACCA	$(AG)_7$	227–229	MH476465	Not found	-
AT09	F: ACGTGCGTGTTACTTGAGTG R:GCAGAGCGAATTCGGAGAGA	$(TC)_7$	218–230	MH476466	Not found	-
AT12	F: GGAGGCTCTACTTCGTTGGG R:CTAGCTAGCACCCACAGGC	$(GA)_7$	231–237	MH476467	Conserved hypothetical protein [*Ricinus communis*]	9.00E-04
AT14	F: CAAGGACAATGGCACTTCCG R:TCCTCCCTCGTCATCCAGTT	$(AG)_7$	231–239	MH476468	Unnamed protein product, partial [*Vitis vinifera*]	2.00E-171
AT16	F: CATCTACCTCATCCCCACGC R:TGGTTTTGTGTTGGGCAACT	$(CT)_7$	228–238	MH476469	Cytochrome B5, n3,ATCB5-D,CB5-D [*Theobroma cacao*]	5.00E-56
AT24	F: TCGGCCTGTAGTGAGAGAGT R:ACCCGCCTAAAATCACCCAA	$(CT)_7$	246–252	MH939929	Not found	-
AT25	F: TGAAGTGCAGCAACAGAATTTGA R:AATGGGTCGGGTTTGGGAAA	$(TA)_6$	236–258	MH476470	Not found	-
AT29	F: AAGCCAACATTCTGCTTCGC R:CACCACTTCGATCCAACCCA	$(AG)_8$	249–255	MH939930	Not found	-
AT30	F: CATGCGCTTGTTCCGTACAG R:ACCCCGCTTTTTAGAGAGAGA	$(CT)_8$	252–256	MH476471	Transcription factor SPATULA isoform X2 [*Vitis vinifera*]	2.00E-06
AT32	F: GCAACGACGTCGATTTCCG R:TGCCCTAAAAATCACACTTCCG	$(CT)_7$	255–269	MH476472	Not found	-

Table 2. Continued.

Locus	Primer sequence (5′-3′)	Repeat motif	Allele size range (bp)	GenBank accession No.	Putative function [Organism]	*E*-value
AT33	F: GGCGTCTGGTCTTCGATCTT R:CGTCGTGTGTAAGCAAGCAG	$(TC)_7$	255–259	MH476473	Unnamed protein product, partial [*Vitis vinifera*]	0
AT43	F: AAACCATCTGAGCCCCTCAC R:GATTGTAACGCGCCGAAGAC	$(CT)_8$	262–276	MH939931	Uncharacterized protein TCM_015808 [*Theobroma cacao*]	1.00E-65
AT60	F: GAAGGTGTTGCTGATGAGCC R:ATTGCAACAACTGACACCGC	$(GCT)_6$	217–235	MH939932	Not found	-
AT63	F: CCATCTCACCATTCTCGCGA R:TCCATGGTTGCATTTGGGGA	$(CGC)_7$	235–244	MH476474	RING-H2 finger protein ATL3 [*Vitis vinifera*]	9.00E-80
AT65	F: GTGTTTCGGGTCGTGAGTCT R:GTTGAGGGACCTGACTGCAA	$(GAT)_5$	231–240	MH476475	Not found	-
AT66	F: TATCTCGCTAGGCCGGAGAT R:CGAGGAAGATTCGAGGCCAA	$(TGA)_6$	240–243	MH939933	Hypothetical protein PRUPE_2G009100 [*Prunus persica*]	2.00E-51
AT67	F: AATGAAGAGTCGTCGTCCCC R:TCCCATGCTGCCCATTACTT	$(CAC)_6$	239–251	MH476476	MLO-like protein 6 [*Vitis vinifera*]	5.00E-105
AT71	F: AGCCCTAACCGCCTTAATCG R:CGCCCCAGAGAAGATCGAAA	$(GCG)_7$	247–256	MH939934	Hypothetical protein B456_001G263100 [*Gossypium raimondii*]	5.00E-53
AT72	F: CGCGTCTCAAAATCGTCACC R:CGGAAGTTTTACGGCCACAG	$(GTT)_6$	242–254	MH476477	Uncharacterized protein LOC100257948 [*Vitis vinifera*]	8.00E-19
AT73	F: CTTCAGGGGAGTACGGAAGC R:CTGGCTCAAGCTTTCGGAAC	$(ACA)_6$	240–252	MH476478	Ethylene-responsive transcription factor 4-like [*Nelumbo nucifera*]	1.00E-28
AT74	F: GGCGTAGGGAGTACCACTTG R:CCCCCTCCTCTCCTCTTCAT	$(GAG)_5$	250–256	MH476479	Trihelix transcription factor ASIL2 [*Cucumis melo*]	2.00E-53
AT75	F: CGGCGAGGATTCAATGGAGA R:TCAAAAGCCGACGATCGTCT	$(TCC)_6$	251–254	MH939935	Hypothetical protein CICLE_v10032167mg [*Citrus clementina*]	4.00E-62
AT77	F: AGTGCACCTGGTGGAATGAG R:AGCGGAGGCATTGTTCTGAA	$(ACA)_6$	252–255	MH476480	Uncharacterized protein LOC100249661 [*Vitis vinifera*]	1.00E-144

Table 3. Genetic diversity in three *Astilboides tabularis* populations[a] based on the 26 newly developed polymorphic microsatellite markers (loci).

	Korea-1 (n = 20)			Korea-2 (n = 20)			China (n = 20)		
Locus	*A*	H_o	H_e[b]	*A*	H_o	H_e[b]	*A*	H_o	H_e[b]
AT01	1	0.000	0.000	1	0.000	0.000	3	0.100	0.265**
AT02	2	0.150	0.139	2	0.400	0.495	3	0.750	0.664
AT04	2	0.550	0.439	2	0.000	0.480**	2	0.167	0.498**
AT07	1	0.000	0.000	1	0.000	0.000	2	0.526	0.465
AT09	1	0.000	0.000	2	0.100	0.095	3	0.250	0.226
AT12	2	0.150	0.139	1	0.000	0.000	3	0.778	0.537
AT14	2	0.150	0.139	3	0.500	0.445**	2	0.053	0.229**
AT16	2	0.150	0.139	1	0.000	0.000	3	0.211	0.193
AT24	3	0.950	0.566**	3	0.750	0.636**	3	0.263	0.237
AT25	1	0.000	0.000	5	0.400	0.741**	2	0.059	0.057
AT29	2	1.000	0.500**	3	0.750	0.499*	3	0.684	0.608
AT30	1	0.000	0.000	2	0.000	0.495**	2	0.684	0.450*
AT32	1	0.000	0.000	3	0.850	0.611*	4	0.632	0.630
AT33	1	0.000	0.000	1	0.000	0.000	3	0.600	0.595
AT43	2	0.850	0.489**	1	0.000	0.000	3	0.111	0.285*
AT60	3	1.000	0.564**	2	0.550	0.499	4	0.737	0.560**
AT63	1	0.000	0.000	2	0.100	0.095	2	0.105	0.100
AT65	1	0.000	0.000	1	0.000	0.000	2	0.188	0.170
AT66	2	0.850	0.489**	1	0.000	0.000	1	0.000	0.000
AT67	2	0.053	0.051	1	0.000	0.000	4	0.222	0.398**
AT71	3	1.000	0.564**	3	0.400	0.586**	2	0.158	0.145
AT72	1	0.000	0.000	2	0.250	0.219	4	0.263	0.320
AT73	2	1.000	0.500**	3	0.944	0.628**	3	0.368	0.309
AT74	2	0.150	0.139	2	0.850	0.489**	3	0.556	0.475
AT75	2	0.850	0.489**	2	0.100	0.500**	2	0.526	0.432
AT77	2	0.100	0.095	2	0.600	0.420	1	0.000	0.000

n, number of individuals; *A*, number of alleles; H_o, observed heterozygosity; H_e, expected heterozygosity.
[a]Locality and voucher information are provided in Table 1. [b]Significant deviation from HWE after correction for multiple tests (*$p < 0.05$ and **$p < 0.01$).

Acknowledgments

This research was supported by Grant No. NIBR201703201 from the National Institute of Biological Resources under the Ministry of Environment, Republic of Korea.

Conflict of Interest

The authors declare that there are no conflicts of interest.

Literature Cited

Belyaeva, T. N. and A. N. Butenkova. 2016. Seed productivity and leaf anatomy of *Astilboides tabularis* (Hemsl.) Engl. International Journal of Pharma and Bio Sciences 7: 511–516.

Choi, J. S., J. H. Jeong and C. H. Lee. 2016. The effects of environmental conditions and chemical treatments on seed germination in *Astilboides tabularis* (Hemsl.) Engl. Korean Journal of Horticultural Science and Technology 34: 363–371.

Excoffier, L. and H. E. L. Lischer. 2010. ARLEQUIN suite ver. 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Molecular Ecology Resources 10: 564–567.

Haas, B. J., A. Papanicolaou, M. Yassour, M. Grabherr, P. D. Blood, J. Bowden, M. B. Couger, D. Eccles, B. Li, M. Lieber, M. D. MacManes, M. Ott, J. Orvis, N. Pochet, F. Strozzi, N. Weeks, R. Westerman, T. William, C. N. Dewey, R. Henschel, R. D. LeDuc, N. Friedman and A. Regev. 2013. *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nature Protocols 8: 1494–1512.

Jintang, P. and J. Cullen. 2001. *Astilboides*. *In* Flora of China. Vol. 8. Brassicaceae through Saxifragaceae. Wu, Z.-Y. and P. H. Raven (eds.), Science Press, Beijing and Missouri Botanical Garden Press, St. Louis, MO. Pp. 269–452.

Ku, Y., H. Oh, K. Shim, M. Kim and S. Lee. 2006. Growth characteristics, genetic diversity and conservation of endangered plants (I): the case of *Astilboides tabularis*, *Euchresta japonica*, *Echinosophora koreensis* and *Lilium cernuum*. National Institute of Environmental Research, Incheon, 60 pp. (in Korean)

Lee, W. K. 2008. Genetic variation on the endangered *Astilboides tabularis* and the rare *Rodgersia podophylla* of Saxifragaceae in Korea. MS Thesis, Sungkyunkwan University, Suwon, 28 pp. (in Korean)

Liu, Z., J. Zhai, N. Han and J. Yin. 2016. Assessment of anti-diabetic activity of the aqueous extract of leaves of *Astilboides tabularis*. Journal of Ethnopharmacology 194: 635–641.

Ministry of the Environment of Korea. 2014. Korean Red List of Threatened Species. National Institute of Biological Resources, Incheon. Pp. 223–234. (in Korean)

Ottewell, K. M., D. C. Bickerton, M. Byrne and A. J. Lowe. 2016. Bridging the gap: a genetic assessment framework for population-level threatened plant conservation prioritization and decision-making. Diversity and Distributions 22: 174–188.

Peakall, R. and P. E. Smouse. 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. Bioinformatics 28: 2537–2539.

Rozen, S. and H. Skaletsky. 1999. Primer 3 on the WWW for general users and for biologist programmers. Methods in Molecular Biology 132: 365–386.

The Angiosperm Phylogeny Group, M. W. Chase, M. J. M. Christenhusz, M. F. Fay, J. W. Byng, W. S. Judd, D. E. Soltis, D. J. Mabberley, A. N. Sennikov, P. S. Soltis and P. F. Stevens. 2016. An update of the Angiosperm Phylogeny Group classification for the orders and families flowering plants: APG IV. Botanical Journal of the Linnean Society 181: 1–20.

Thiel, T., W. Michalek, R. 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.

Wang, R. R.-C., S. R. Larson and K. B. Jensen. 2017. Differential transferability of EST-SSR primers developed from the diploid species *Pseudoroegneria spicata*, *Thinopyrum bessarabicum*, and *Thinopyrum elongatum*. Genome 60: 530–536.

Yan, H., Y. Zhang, B. Zeng, G. Yin, X. Zhang, Y. Ji, L. Huang, X. Jiang, X. Liu, Y. Peng, X. Ma and Y. Yan. 2016. Genetic diversity and association of EST-SSR and SCoT markers with rust traits in Orchardgrass (*Dactylis glomerata* L.). Molecules 21: 66.