New polymorphic microsatellite markers for the endangered fern

*Ceratopteris thalictroides* (Parkeriaceae)

Won-Bum CHO, Eun-Kyeong HAN, Myounghai KWAK and Jung-Hyun LEE

Department of Biology Education, Chonnam National University, Gwangju 61186, Korea

1Plant Resources Division, National Institute of Biological Resources, Incheon 22689, Korea

(Received 31 May 2018; Revised 14 June 2018; Accepted 18 June 2018)

**ABSTRACT:** *Ceratopteris thalictroides* is a semi-aquatic fern with a circumtropical distribution. This species is designated internationally on the IUCN Red List as requiring at least some concern. Korean populations are of great concern for the species’ long-term survival, as they are at the northern limit of the species distribution. To establish an effective conservation strategy for those populations at the genetic level, we used the Mi-Seq platform to develop three sets of 25 polymorphic microsatellite markers for *C. thalictroides*, which is endangered in Korea. In populations sampled from Busan and Gochang, the number of alleles ranged from 2 to 13 (average of 5.64), and plants presented an expected heterozygosity of 0.000 to 0.860. These markers will be useful for evaluating the genetic status and conserving Korean populations of *C. thalictroides* more effectively.

**Keywords:** *Ceratopteris thalictroides*, conservation, microsatellite markers

*Ceratopteris thalictroides* (L.) Brongn. (Parkeriaceae) is a semi-aquatic fern with circum-tropical distribution, mainly growing in paddy fields, ponds, or marshes (Watano and Masuyama, 1994). This genus contains four species (Lloyd, 1974). Whereas three of those species are diploid plants, *C. thalictroides* is tetraploid (Hickok, 1977; Masuyama and Watano, 2005). Although *C. thalictroides* is widely distributed throughout tropical and subtropical regions, it is internationally considered to be of at least some concern, based on the IUCN Red List of Threatened Species.

In Korea, this species was first recorded at Suncheon and Gwangyang, in the southern regions of the Peninsula. Since then, more populations have been identified. This species is now designated as endangered in Korea and has been assigned legal protection for eight extant populations (Seocheon, Gunsan, Jeongeup, Buan, Iksan, Gochang, Gwangju, and Busan). These Korean populations are clearly critical because they represent the northern edge of the distribution range, possibly providing potential for further expansion of the species if those plants are able to adapt to selection pressures from such marginal environments (Kawecki, 2008). Therefore, efforts to conserve Korean populations are needed if we are to achieve long-term survival for *C. thalictroides*.

In China, researchers are also taking various steps to conserve the populations of *C. thalictroides*. There, this species has been designated as endangered in an effort to protect and manage its native habitats (Yu, 1999). In addition, its genetic variability has been analyzed by using various markers such as random amplified polymorphic DNA, inter-simple sequence repeat, and microsatellites (Dong et al., 2008; Yang et al., 2016). Although 30 microsatellite markers have been established for Chinese populations of *C. thalictroides* (Yang et al., 2016), they are not adequate for frequently examining genetic polymorphism. Furthermore, it is cost-effective for such markers to be organized as “sets” if researchers are to continue periodic genetic monitoring. Therefore, we have developed new polymorphic microsatellite sets for evaluating the genetic status of Korean populations of *C. thalictroides* and conserving those plants.

**Materials and Methods**

For these new microsatellite markers, total genomic DNA
was obtained from a fresh leaf collected from an individual plant of *C. thalictroides* growing at the Dongnim reservoir, Gochang, Korea. The extraction procedure involved an MG Plant Genomic DNA Extraction SV Miniprep Kit used according to the manufacturer’s protocol (Macrogen Inc., Seongnam, Korea). After the quality of the genomic DNA was checked by gel electrophoresis, we generated a shotgun library by using the Illumina Mi-Seq platform (LAS, Seoul, Korea). In all, 5,256,220 paired-ends with read lengths of approximately 300 bp were obtained. Based on screening with the SSR_Pipeline v.0951 infrastructure (Miller et al., 2013), we identified a 22,990 microsatellite motif with flanking regions larger than 100 bp that contained di-, tri-, or tetra-nucleotide repeats that were at least 12, 6, or 5 times, respectively, in size. Following the method of Cho et al. (2015), we assembled groupings of 50 to 100 reads that were reference guide-mapped on the whole raw reads, using the Geneious R10.1.3 software platform (Kearse et al., 2012). From this, we selected reads that had two separate alleles and no additional single nucleotide polymorphisms in the flanking region. After eliminating any duplication through *de novo* assembly, we used the final chosen reads to develop microsatellite markers. We then utilized the Primer 3 software in Geneious program R10.1.3 to design 81 primer pairs (18–22 bp long), using a melting temperature of 53–60 °C and GC contents of 35–65%.

The forward primers added three sets of M13 tag sequences (5′-CACGACGTTGTAAACGAC-3′, 5′-TGTGGAATTGTGAGCGG-3′, and 5′-CTATAGGGCACGCGTGGT-3′) with 6-FAM, VIC, and NED fluorescent dye, respectively. For multiplex polymerase chain reaction (PCR), each primer set comprised nine primers with lengths of 100–300 bp.

To test the effectiveness of these new microsatellite loci, we sampled 63 individuals of *C. thalictroides* from two populations (Dongnim reservoir, Gochang, *n* = 33; Is. Doonchido, Busan, *n* = 30). For each locus, PCR amplification was performed in a final volume of 5 µL that contained 15 to 20 ng of extracted DNA, 2.5 µL of 2× Multiplex PCR master mix (Qiagen, Valencia, CA, USA), 0.01 µM for the forward primer, 0.2 µM for the reverse primer, and 0.1 µM for the M13 primer. The PCR protocol included an initial denaturation for 15 min at 95°C; followed by 30 cycles, each consisting of denaturing for 30 s at 94°C, 1.5 min of annealing at 52°C, and extension for 1.5 min at 72°C; with a final extension for 10 min at 72°C. The PCR products were analyzed using an ABI 3730XL sequencer with GeneScan-500LIZ Size Standard (Applied Biosystems, Foster City, CA, USA). Allele sizes and the peaks of each sample were determined via Peak Scanner software 2 (Applied Biosystems). In analyzing the tetraploid microsatellite data, we were not able to make exact determinations of allele frequencies and patterns due to the uncertainty of multiple alleles (Dufresne et al., 2014). Therefore, we calculated the number of alleles (*N*<sub>a</sub>), the Shannon-Wiener Diversity Index (*H*′), and the expected heterozygosity (*H*<sub>E</sub>) using A4tetra 1.2 software (van Puyvelde et al., 2010), which was developed for analyzing codominant microsatellite data in tetraploid species.

### Results and Discussion

We have produced useful sets of microsatellite markers for analyzing genetic diversity and establishing conservation strategies for *Ceratopteris thalictroides*, an endangered plant. Of the 81 primer pairs studied here, 25 microsatellite markers were polymorphic, and 17 were found to be monomorphic in two of the examined populations (Table 1). All developed markers have been deposited in GenBank (Table 1). Among the 25 polymorphic microsatellite loci, the number of alleles per locus varied from 3 to 179 bp. The number of alleles per locus for each primer pair is shown in Table 1.

### Table 1. Characterization of 25 microsatellite loci for *Ceratopteris thalictroides*.

| Locus* | Primer sequence (5′–3’) | Repeat motif | *N*<sub>a</sub> | Allele size range (bp) | Fluorescent label | GenBank accession No. |
|--------|-------------------------|--------------|----------------|-----------------------|-----------------|----------------------|
| cet007 | F: AAAATCTGGCCACCGTTTG<BR>R: AAGATTGTAAGTGGCTG | (AG)<sub>7</sub> | 8 | 153–179 | NED | MH261329 |
| cet009 | F: TGAGTTGCTCATGATGTTGG<BR>R: AGACTCTTTTTTCTAGAGAC | (GAA)<sub>3</sub> | 3 | 232–241 | NED | MH261330 |
| cet017 | F: CAAGAGGAGGGCGAATAAC<BR>R: TCTCTCTCTCTGTTTCAGT | (AAG)<sub>6</sub> | 5 | 180–207 | VIC | MH261337 |
| cet023 | F: CTCAGACCTCTCTCTGGA<BR>R: CATTCGATACCAACAGCTA | (CTC)<sub>2</sub> | 5 | 179–189 | 6-FAM | MH261339 |
| cet028* | F: ACATGGCTTCTTATTCCTG<BR>R: TATGATCTTTGGTGCCAGCA | (AG)<sub>12</sub> | 4 | 153–179 | NED | MH261329 |

Korean Journal of Plant Taxonomy Vol. 48 No. 2 (2018)
Table 1. Continued.

| Locus<sup>a</sup> | Primer sequence (5′–3′) | Repeat motif | N<sub>x</sub> | Allele size range (bp) | Fluorescent label | GenBank accession No. |
|------------------|--------------------------|--------------|-------------|------------------------|------------------|----------------------|
| cet101           | F: ACTTTTGAAACTATTCCTGCTCG  R: TTGGTTGACATTGTGGTCTG      | (AC)<sub>3</sub>   | 2          | 98–102                | 6-FAM            | MH261347             |
| cet104           | F: TGCCATACAAAAGTCTAAGA  R: CCTACACTATGCACATGGA          | (TACA)<sub>6</sub> | 2          | 139–151               | VIC              | MH261348             |
| cet120           | F: GTCAGAATCCCGTCTGATTC  R: CAACAGAGGGACAGCTATT         | (GTAT)<sub>5</sub> | 5          | 222–238               | 6-FAM            | MH261352             |

Multiplex mix B

| cet010*          | F: TAAAAGCGATCCTACCTGAT  R: TAAAGCTTGGGAGAGGCTA          | (AG)<sub>6</sub>   | 9          | 143–165               | 6-FAM            | MH261331             |
| cet012*          | F: AGTCAAAAATCTGTCTAGG  R: TAGCGTGGCTATGTTCTTC           | (AG)<sub>5</sub>   | 5          | 227–241               | 6-FAM            | MH261332             |
| cet018*          | F: CAGGAGAGGAGAGATCGC  R: AGGAGAGACATCTTACTGCT         | (TC)<sub>2</sub>   | 7          | 242–254               | NED              | MH261334             |
| cet022           | F: TCATATCGCAATCAAAGAC  R: TGGAGAGAGAGGACATCTG          | (TC)<sub>7</sub>   | 7          | 177–189               | VIC              | MH261336             |
| cet025           | F: GGAGAGAAATGGTTGATGT  R: ATGACGAGATGAATGAGGTG         | (TG)<sub>7</sub>   | 8          | 175–189               | NED              | MH261338             |
| cet042*          | F: GAAAGACATACCTCTCTG  R: CAAGAAAGAGGAGGCT             | (GA)<sub>5</sub>   | 3          | 267–271               | VIC              | MH261342             |
| cet046           | F: GACACCATTCCATGCCAG  R: ATACATGCTGTTGCTGAGG           | (CT)<sub>7</sub>   | 3          | 185–189               | 6-FAM            | MH261343             |
| cet050           | F: CACATGAGCGAGTTGAGAG  R: TCATGGTGTGTTGGAGGAA       | (CT)<sub>5</sub>   | 3          | 209–233               | VIC              | MH261344             |
| cet108           | F: CCATCATTGAGCTCGAGAT  R: ACAGAGTGGACAAAGGTATT        | (TG)<sub>5</sub>   | 5          | 161–171               | NED              | MH261350             |

Multiplex mix C

| cet021*          | F: GCAACAACATAAGCGTCAAT  R: TGGTCACTGAGACATCAAAT       | (AC)<sub>6</sub>   | 6          | 240–254               | 6-FAM            | MH261335             |
| cet029*          | F: TACAGTGACAAATGCTCTCT  R: CTCATGAGATCTCTCCTTCT        | (GAA)<sub>4</sub>  | 4          | 213–222               | 6-FAM            | MH261340             |
| cet032*          | F: TTCTTCTCTGACCGATCTTT  R: AGTCTCTGAGGATCCTAATG      | (CAA)<sub>3</sub>  | 3          | 209–215               | VIC              | MH261341             |
| cet051           | F: GACGATGAGGACATATTGAT  R: GATAGGTCTGAGCGCATTTT      | (GA)<sub>2</sub>   | 14         | 291–315               | VIC              | MH261345             |
| cet053           | F: CTTGGGATGGAGAAAATGAT  R: CACTTCCTATGTCCTCCAAA       | (TC)<sub>3</sub>   | 7          | 226–242               | NED              | MH261346             |
| cet107           | F: GGCCTCTGTGGAATATGATT  R: TTCATGCTCTAAACTCAAACCG    | (TG)<sub>3</sub>   | 3          | 158–162               | NED              | MH261349             |
| cet113           | F: GTCGGACTTTTAATCCCAT  R: GGAAAAATTTGGGAGGAGC       | (CA)<sub>5</sub>   | 4          | 169–187               | VIC              | MH261351             |
| cet127           | F: TTGAACCTTGAGCATGAGAG  R: TGGCGAATATGACATACCTT      | (CT)<sub>9</sub>   | 9          | 257–275               | NED              | MH261353             |

N<sub>x</sub>, number of alleles.

<sup>a</sup>Reaction concentrations in PCR for primers were 0.01 for forward primer and 0.2 for reverse primer. Loci marked with asterisk (*) had reaction concentrations of 0.02 for forward primer and 0.4 for reverse primer.
ranged from 2 to 13, with an average of 5.64. Values for ($H_e$) were 0.000 to 0.860, while $H'$ ranged from 0.000 to 2.152 (Table 2). These results are similar to those reported from a previous study in China (Yang et al., 2016) that involved RAD tag sequencing ($N_A$, 4–10; $H_e$, 0.264–0.852). However, because the Korean Peninsula is a northern limit for the distribution of *C. thalictroides*, our findings suggest that the marker sets described here include considerable polymorphism.

### Acknowledgments

This work was supported by the National Institute of Biological Resources of Korea (Grant No. NIBR201703101).

### Conflict of Interest

Authors declare that there is no conflict of interest.

### Literature Cited

Cho, W.-B., I.-S. Choi and B.-H. Choi. 2015. Development of microsatellite markers for the endangered *Pedicularis ishidoyana* (Orobanchaceae) using next-generation sequencing. Applications in Plant Sciences 3: 1500083.

Dong, Y.-H., J.-M. Chen, G. W. Robert and Q.-F. Wang. 2008. Genetic variation in the endangered aquatic fern *Ceratopteris thalictroides* (Parkeriaceae) in China: implications from

---

**Table 2.** Summary of genetic parameters estimated from 25 microsatellite loci across 63 individuals sampled from the two Korean populations, Gochang (GC) and Busan (BS), Korea.

| Locus   | $N_a$ | $H_e$  | $H'$  | Sr       | $N_a$ | $H_e$  | $H'$  | Sr       |
|---------|-------|--------|-------|----------|-------|--------|-------|----------|
| cet007  | 2     | 0.114  | 0.229 | 153–155  | 8     | 0.826  | 1.856 | 153–179  |
| cet009  | 3     | 0.621  | 1.021 | 232–241  | 3     | 0.583  | 0.946 | 232–241  |
| cet010  | 7     | 0.718  | 1.425 | 143–163  | 8     | 0.810  | 1.825 | 145–165  |
| cet012  | 4     | 0.501  | 0.861 | 227–237  | 4     | 0.589  | 1.084 | 227–241  |
| cet017  | 5     | 0.685  | 1.307 | 210–222  | 2     | 0.320  | 0.500 | 213–222  |
| cet018  | 6     | 0.670  | 1.322 | 242–254  | 5     | 0.627  | 1.223 | 244–254  |
| cet021  | 5     | 0.622  | 1.190 | 240–254  | 4     | 0.643  | 1.130 | 240–252  |
| cet022  | 6     | 0.637  | 1.297 | 179–189  | 4     | 0.266  | 0.548 | 177–187  |
| cet023  | 4     | 0.691  | 1.244 | 180–207  | 4     | 0.692  | 1.256 | 180–207  |
| cet025  | 8     | 0.784  | 1.733 | 175–189  | 7     | 0.749  | 1.541 | 177–189  |
| cet028  | 3     | 0.601  | 0.994 | 179–189  | 3     | 0.604  | 1.010 | 179–189  |
| cet029  | 4     | 0.729  | 1.339 | 213–222  | 4     | 0.728  | 1.341 | 213–222  |
| cet032  | 2     | 0.497  | 0.691 | 209–215  | 3     | 0.564  | 0.906 | 209–215  |
| cet042  | 2     | 0.114  | 0.229 | 267–269  | 3     | 0.499  | 0.693 | 267–271  |
| cet046  | 2     | 0.165  | 0.305 | 187–189  | 3     | 0.615  | 1.025 | 185–189  |
| cet050  | 9     | 0.717  | 1.495 | 209–233  | 6     | 0.785  | 1.620 | 209–231  |
| cet051  | 11    | 0.859  | 2.124 | 295–315  | 13    | 0.860  | 2.152 | 291–313  |
| cet053  | 7     | 0.687  | 1.414 | 226–242  | 6     | 0.643  | 1.324 | 226–240  |
| cet101  | 2     | 0.030  | 0.077 | 98–102   | 1     | 0.000  | 0.000 | 98       |
| cet104  | 2     | 0.497  | 0.690 | 139–151  | 2     | 0.453  | 0.645 | 139–151  |
| cet107  | 2     | 0.030  | 0.077 | 160–162  | 3     | 0.183  | 0.368 | 158–162  |
| cet108  | 5     | 0.730  | 1.424 | 161–171  | 4     | 0.679  | 1.203 | 161–169  |
| cet113  | 2     | 0.497  | 0.691 | 169–185  | 4     | 0.637  | 1.163 | 169–187  |
| cet120  | 4     | 0.680  | 1.224 | 226–238  | 3     | 0.452  | 0.731 | 222–234  |
| cet127  | 5     | 0.640  | 1.208 | 265–273  | 7     | 0.803  | 1.771 | 257–275  |

$N_a$, number of alleles; $H_e$, expected heterozygosity; $H'$, Shannon-Wiener Diversity Index; Sr, size range (bp).
RAPD and ISSR data. Botanical Journal of the Linnean Society 157: 657–671.

Dufresne, F., M. Stift, R. Vergilino and B. K. Mable. 2014. Recent progress and challenges in population genetics of polyploid organisms: an overview of current state-of-the-art molecular and statistical tools. Molecular Ecology 23: 40–69.

Hickok, L. G. 1977. Cytological relationships between three diploid species of the fern genus *Ceratopteris*. Canadian Journal of Botany 55: 1660–1667.

Kawecki, T. J. 2008. Adaptation to marginal habitats. Annual Review of Ecology, Evolution, and Systematics 39: 321–342.

Kearse, M., R. Moir, A. Wilson, S. Stones-Havas, M. Cheung, S. Sturrock, S. Buxton, A. Cooper, S. Markowitz, C. Duran, T. Thierer, B. Ashton, P. Meintjes and A. Drummond. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28: 1647–1649.

Lloyd, R. M. 1974. Systematic of the genus *Ceratopteris* Brongn. (Parkeriaceae) II. Taxonomy. Brittonia 26: 139–160.

Masuyama, S. and Y. Watano. 2005. Cryptic species in the fern *Ceratopteris thalictroides* (L.) Brongn. (Parkeriaceae). II. Cytological characteristics of three cryptic species. Acta Phytophysica 56: 231–240.

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.

van Puyvelde, K., A. van Geert and L. Triest. 2010. ATETRA, a new software program to analyse tetraploid microsatellite data: comparison with TETRA and TETRASAT. Molecular Ecology Resources 10: 331–334.

Watano, Y. and S. Masuyama. 1994. Genetic differentiation in populations of the polymorphic fern *Ceratopteris thalictroides* in Japan. Journal of Plant Research 107: 139–146.

Yang, X. Y., Z. C. Long, A. W. Gichira, Y. H. Guo, Q. F. Wang and J. M. Chen. 2016. Development of microsatellite markers in the tetraploid fern *Ceratopteris thalictroides* (Parkeriaceae) using RAD tag sequencing. Genetics and Molecular Research 15: gmr7550.

Yu, Y. F. 1999. A milestone of wild plant conservation in China. Plants 5: 3–11.