# Development of microsatellite markers in the tetraploid fern *Ceratopteris thalictroides* (Parkeriaceae) using RAD tag sequencing

**X.Y. Yang[1]*, Z.C. Long[2]*, A.W. Gichira[2], Y.H. Guo[1], Q.F. Wang[2] and J.M. Chen[2]**

[1]Laboratory of Plant Systematics and Evolutionary Biology, College of Life Sciences, Wuhan University, Wuhan, China
[2]Key Laboratory of Aquatic Botany and Watershed Ecology, Wuhan Botanical Garden, Chinese Academy of Sciences, Wuhan, China

*These authors contributed equally to this study.
Corresponding authors: Q.F. Wang / J.M. Chen
E-mail: qfwang@wbgcas.cn / jmchen@wbgcas.cn

Genet. Mol. Res. 15 (1): gmr.15017550
Received August 31, 2015
Accepted November 13, 2015
Published February 19, 2016
DOI http://dx.doi.org/10.4238/gmr.15017550

**ABSTRACT.** To understand the genetic variability of the tetraploid fern *Ceratopteris thalictroides* (Parkeriaceae), we described 30 polymorphic microsatellite markers obtained using the restriction site-associated DNA (RAD) tag sequencing technique. A total of 26 individuals were genotyped for each marker. The number of alleles per locus ranged from 4 to 10, and the expected heterozygosity and the Shannon-Wiener index ranged from 0.264 to 0.852 and 0.676 to 2.032, respectively. Because these 30 microsatellite markers exhibit high degrees of genetic variation, they will be useful tools for studying the adaptive genetic variation and sustainable conservation of *C. thalictroides*.

**Key words:** *Ceratopteris thalictroides*; Homosporous fern; Microsatellite marker; RAD tag sequencing; Tetraploid species

 

## INTRODUCTION

*Ceratopteris thalictroides* (L.) Brongn. (Parkeriaceae) is a semi-aquatic homosporous tetraploid fern. In China, the number of populations of *C. thalictroides* has declined rapidly due to the deterioration of primary habitats. As a consequence, the species is now considered to be endangered in China and is listed in the second category of key protected wild plants (Yu, 1999). In several other countries, including neighboring Vietnam and India, the species is also listed as endangered.

For conservation purposes, in recent years, genetic variation among Chinese *C. thalictroides* populations has been investigated using a variety of dominant genetic markers, e.g., random amplified polymorphic DNA, inter simple sequence repeats (Dong et al., 2008), and chloroplast DNA non-coding regions (Liao et al., 2011). However, these studies are still insufficient to define the adaptive population differentiation. Here, we report the development of polymorphic microsatellite markers from *C. thalictroides* using restriction site-associated DNA (RAD) tag sequencing, which will facilitate the ongoing studies of adaptive genetic variation and sustainable conservation for this endangered species.

## MATERIAL AND METHODS

A *C. thalictroides* individual from the Wuhan Botanical Garden was used as the source of DNA for this study. DNA was extracted from its fresh leaves using a Plant Genomic DNA Isolation kit (Tiangen, Beijing, China), following the manufacturer protocol. The RAD library was constructed according to the protocol described by Baird et al. (2008). This library was sequenced on an Illumina HiSeq 2000 Platform at Huazhong Agricultural University, generating 4.5 million DNA reads, with an average read length of 353 bp. A total of 650 microsatellite loci were identified from the resources using the MIcroSAtellite identification tool (Thiel et al., 2003) and 285 primers were successfully designed using Primer3 (http://biotools.umassmed.edu/bioapps/primer3_www.cgi).

A random selection of 115 of the designed primers were initially screened using total DNA isolated from the dried leaves of six *C. thalictroides* individuals. All forward primers were fluorescently labeled with FAM on the 5'-end. Polymerase chain reaction (PCR) amplifications were carried out in a volume of 20 µL containing 0.25 mM each dNTP, 2 µL 10X Taq buffer (10 mM Tris-HCl, pH 8.3, 1.5 mM $MgCl_2$, and 50 mM KCl), 1 mM each primer, 0.2 U Taq polymerase (TransGen Biotech Co., Beijing, China), and 25 ng DNA template. Amplification of genomic DNA was carried out using an ABI 2720 Thermal Cycler (Applied Biosystems, Foster City, CA, USA). The PCR profile was programmed with an initial denaturation of 5 min at 94°C, followed by 35 cycles of 30 s at 94°C, 30 s annealing at 53-55°C (depending on the type of primers), and 30 s extension at 72°C, with a final extension step of 10 min at 72°C. PCR products were separated using an ABI 3730 automated sequencer (Sangon Biotech, Shanghai, China) and visualized using the GeneScan system (Applied Biosystems).

Based on the initial screening results, 45 primer pairs were selected and used to genotype 26 individuals from two populations located in Baise, Guangxi Province (106°43'E, 23°33'N) and Yingde, Guangdong Province (113°37'E, 24°29'N). The number of alleles ($N_A$), expected ($H_E$) heterozygosity, and Shannon-Wiener index ($H'$) were calculated for each locus using ATETRA v.1.0 (Van Puyvelde et al., 2010).

## RESULTS AND DISCUSSION

We successfully amplified 30 polymorphic loci as shown in Table 1. $N_A$ ranged from 4 to

10, and $H_E$ and *H*' ranged from 0.264 to 0.852 and 0.676 to 2.032, respectively (Table 1). Because of the tetraploid nature of *C. thalictroides*, exact allele frequencies could not be determined. The test for linkage disequilibrium was not conducted in this study, due to the unknown allele dosage of partial heterozygotes. These 30 tested microsatellite markers exhibit high degrees of genetic variation, which will facilitate further investigation of adaptive genetic variation and conservation for *C. thalictroides* in China and adjacent countries.

**Table 1.** Primer sequences and characteristics of the 30 polymorphic microsatellite loci in 26 individuals of *Ceratopteris thalictroides* (the number in parentheses are standard deviations for $H_E$ and *H*').

| Locus | Primer sequence (5'-3') | Repeat motif | Allele size (bp) | Ta (°C) | $N_A$ | $H_E$ | *H*' | GenBank accession No. |
|---|---|---|---|---|---|---|---|---|
| Cer1 | F: ATGACAGTGCCGATGCCTAT<br>R: CTGGCCGGTTAAGTTGAGTC | $(GA)_{15}$ | 183-214 | 60 | 4 | 0.438 (0.031) | 0.805 (0.048) | KP858476 |
| Cer2 | F: ACAGGGCCAAAGCTAGTCAA<br>R: TACACACACACACGCACACA | $(GT)_6$ | 164-175 | 59 | 5 | 0.756 (0.005) | 1.457 (0.016) | KP858477 |
| Cer3 | F: GCCTATGGGTTTGGTGTGTC<br>R: CCGTGTGTGTGTGTCACAAG | $(GTTG)_6$ | 176-194 | 60 | 7 | 0.772 (0.007) | 1.698 (0.013) | KP858478 |
| Cer4 | F: AGGAAGGTGAGCAGTCTGAC<br>R: GGTGTTGTTGTTGTTGTTGTCC | $(ACA)_6$ | 94-99 | 59 | 5 | 0.444 (0.018) | 0.924 (0.031) | KP858479 |
| Cer5 | F: GGGGCAAGTCTCGTAAAAGA<br>R: CAAGCACTCTGTTCGCTCTG | $(TA)_6$ | 104-116 | 57 | 8 | 0.264 (0.000) | 0.676 (0.000) | KP858480 |
| Cer6 | F: TGGAAACCCACCGAACTTTG<br>R: GAACTCAGGAGACAGCAACG | $(GTT)_7$ | 118-124 | 58 | 8 | 0.814 (0.000) | 1.828 (0.000) | KP858481 |
| Cer7 | F: TGTGTGATGTGCTGTGTGTG<br>R: ACATGCAACCATCACAAGCC | $(TG)_6(GT)_8(TG)_7$ | 204-225 | 59 | 5 | 0.630 (0.026) | 1.242 (0.042) | KP858482 |
| Cer8 | F: TGTTACGGTGGTGTGGTGTA<br>R: ACACCACCACTGCCATTACT | $(GTC)_5$ | 270-294 | 59 | 4 | 0.705 (0.006) | 1.277 (0.018) | KP858483 |
| Cer9 | F: AGTCAAGAAGGCTACAGCGG<br>R: CGCTTAACCGTTTACCTATCG | $(GTCTT)_6$ | 160-174 | 60 | 4 | 0.467 (0.027) | 0.910 (0.044) | KT596676 |
| Cer10 | F: AAAGAAGATTTTTAGTTATCTGAATGC<br>R: CAAAAATAGCATAAGCTTCGGG | $(TA)_{10}$ | 390-438 | 57 | 6 | 0.761 (0.006) | 1.523 (0.011) | KT596677 |
| Cer11 | F: TCCCATCCAACCTAGTTTTCC<br>R: CTCGACGCAAGCTATTACCC | $(GA)_7$ | 280-298 | 60 | 5 | 0.714 (0.007) | 1.397 (0.013) | KT596678 |
| Cer12 | F: TGAGACTCCACGCTACATGC<br>R: ATCCTCCGTTGGTCTCCAGT | $(TG)_6$ | 280-320 | 60 | 4 | 0.501 (0.017) | 1.020 (0.026) | KT596679 |
| Cer13 | F: AACCCTGGAATCTTAGAGGAAAA<br>TGAGGCTTCCTTACCTTGCT | $(AG)_8$ | 180-198 | 59 | 6 | 0.454 (0.022) | 0.966 (0.032) | KT596680 |
| Cer14 | F: GCAGCCCACACTCCTACATC<br>R: AGAGAGAGGGAGTTGTGCCA | $(CT)_7$ | 220-260 | 60 | 10 | 0.837 (0.003) | 2.012 (0.006) | KT596681 |
| Cer15 | F: TCTTGGACATGGATATGGCA<br>R: TCACTTGTAGTGTCGTGTCGC | $(GA)_{13}$ | 200-235 | 60 | 9 | 0.740 (0.012) | 1.647 (0.025) | KT596682 |
| Cer16 | F: CGAGGCTTGGACTCTTCATC<br>R: CTACCAGAGGATTGGGAGCA | $(ATC)_5$ | 220-268 | 59 | 7 | 0.679 (0.017) | 1.372 (0.028) | KT596683 |
| Cer17 | F: TTGTCCACCATGTTCCTCCT<br>R: GTAGCCAGAATCATCGAGGC | $(TTC)_5$ | 110-128 | 60 | 7 | 0.462 (0.023) | 1.006 (0.033) | KT596684 |
| Cer18 | F: TGGGGTACGTGAGGTACGTT<br>R: GCCAAGTTTGGCCTCAAGTA | $(TG)_7(AG)_6(AGAA)_6$ | 130-148 | 60 | 6 | 0.803 (0.017) | 1.693 (0.009) | KT596685 |
| Cer19 | F: TGCTTCCATTGTGTTCCAAA<br>R: TCATGACTCCTTGAGCTCCC | $(CAT)_5$ | 150-165 | 60 | 10 | 0.823 (0.008) | 1.942 (0.020) | KT596686 |
| Cer20 | F: CGAAGCAAATGCATGACTGT<br>R: CCAGCAAATGAGGAAGTTAGTTT | $(CT)_8$ | 185-206 | 60 | 6 | 0.538 (0.021) | 1.139 (0.030) | KT596687 |
| Cer21 | F: TGCAGAGATAGCCACACCAC<br>R: TGAGTCAAAATTGCACCCAC | $(GAA)_5$ | 122-166 | 60 | 8 | 0.563 (0.013) | 1.304 (0.023) | KT596688 |
| Cer22 | F: TGCGTTATGCCTGCTCATAC<br>R: CTTCTTGTCCCATCATGCCT | $(TC)_6$ | 260-298 | 59 | 9 | 0.852 (0.002) | 2.032 (0.009) | KT596689 |
| Cer23 | F: TCCTTTCTCAATTCTCACTTTCG<br>R: GGAAATGCCCTTTTCTCCTC | $(TC)_6$ | 420-445 | 59 | 7 | 0.710 (0.014) | 1.487 (0.027) | KT596690 |
| Cer24 | F: CGCAGTTGACACACTCGTCT<br>R: CTGCAGGGATACGGAAACAT | $(TCT)_5$ | 300-320 | 60 | 7 | 0.774 (0.005) | 1.646 (0.012) | KT596691 |
| Cer25 | F: CATTTGGAAGGTGTTGCCTT<br>R: AGATTGTGCCCCATTGATGT | $(A)_{14}$ | 450-560 | 60 | 8 | 0.584 (0.016) | 1.222 (0.026) | KT596692 |
| Cer26 | F: AGCGGCGACCTACTCTGATA<br>R: CATTCATTATTGTGTGTTATGCTCC | $(GAA)_5$ | 160-178 | 60 | 6 | 0.790 (0.005) | 1.641(0.015) | KT596693 |
| Cer27 | F: CCCTGGCATCCTAAATTTCAA<br>R: TGCAGAGAACCAGTCATTCG | $(TA)_6$ | 280-320 | 61 | 10 | 0.788 (0.006) | 1.764 (0.019) | KT596694 |
| Cer28 | F: CTATGGCCAGGAAGAAGTCG<br>R: TCTCTCCCCATCCCCTATCT | $(AG)_6$ | 150-180 | 60 | 6 | 0.755 (0.011) | 1.597 (0.024) | KT596695 |
| Cer29 | F: ACAAACCAATAATTGCATTTTAGA<br>R: CATATGCAATGGGAAAATTCA | $(CT)_6$ | 240-268 | 57 | 10 | 0.673 (0.007) | 1.579 (0.014) | KT596696 |
| Cer30 | F: CAAAGAAAGAGAAAGGGTGCT<br>R: ACATTCTGCGCCTATTGTGT | $(AG)_6$ | 100-122 | 58 | 8 | 0.849 (0.002) | 1.958 (0.005) | KT596697 |

Ta = annealing temperature; $N_A$ = number of alleles observed; $H_E$ = expected heterozygosity; *H*' = Shannon-Wiener index.

 

## Conflicts of interest

The authors declare no conflict of interest.

## ACKNOWLEDGMENTS

We thank Yuan-Huo Dong for his kind assistance during fieldwork. Research supported by grants from the National Natural Science Foundation of China (#30800061 and #31270278).

## REFERENCES

Baird NA, Etter PD, Atwood TS, Currey MC, et al. (2008). Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS One* 3: e3376. http://dx.doi.org/10.1371/journal.pone.0003376

Dong YH, Chen JM, Robert GW and Wang QF (2008). Genetic variation in the endangered aquatic fern *Ceratopteris thalictroides* (Parkeriaceae) in China: implications from RAPD and ISSR data. *Bot. J. Linn. Soc.* 157: 657-671. http://dx.doi.org/10.1111/j.1095-8339.2008.00836.x

Liao YY, Yang XY, Motley TJ, Chen JM, et al. (2011). Phylogeographic analysis reveals two cryptic species of the endangered fern *Ceratopteris thalictroides* (L.) Brongn. (Parkeriaceae) in China. *Conserv. Genet.* 12: 1357-1365. http://dx.doi.org/10.1007/s10592-011-0236-7

Thiel T, Michalek W, Varshney RK and Graner A (2003). Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). *Theor. Appl. Genet.* 106: 411-422.

VAN Puyvelde K, VAN Geert A and Triest L (2010). atetra, a new software program to analyse tetraploid microsatellite data: comparison with tetra and tetrasat. *Mol. Ecol. Resour.* 10: 331-334. http://dx.doi.org/10.1111/j.1755-0998.2009.02748.x

Yu YF (1999). A milestone of wild plant conservation in China. *Plants* 5: 3-11.