Isolation and Characterization of Microsatellite Markers for an Endemic Tree in East Asia, Quercus variabilis (Fagaceae)

Authors: Wang, Xian, Li, Jing, and Li, Yong

Source: Applications in Plant Sciences, 3(6)

Published By: Botanical Society of America

URL: https://doi.org/10.3732/apps.1500032

Applications in Plant Sciences 2015 3(6): 1500032

PRIMER NOTE

ISOLATION AND CHARACTERIZATION OF MICROSATELLITE MARKERS FOR AN ENDEMIC TREE IN EAST ASIA, *QUERCUS VARIABILIS* (FAGACEAE)[1]

XIAN WANG[2], JING LI[2], AND YONG LI[2,3]

[2]College of Forestry, Henan Agricultural University, Zhengzhou 450002, People's Republic of China

- *Premise of the study:* Microsatellite markers of *Quercus variabilis* (Fagaceae) were isolated for population genetic and landscape genetic studies.
- *Methods and Results:* Roche 454 pyrosequencing combined with the magnetic bead enrichment protocol were used to isolate microsatellite markers for *Q. variabilis.* A total of 2121 microsatellites were identified from 63,851 individual sequence reads. One hundred microsatellite loci were selected to test primer amplification efficiency among 24 individuals from two wild populations. Among the 100 tested markers, 34 primer pairs were successfully amplified. Of these, 14 yielded polymorphic amplification products, whereas the remaining 20 loci were monomorphic. The number of alleles for polymorphic loci ranged from two to six, and the observed heterozygosity ranged from 0.042 to 0.750.
- *Conclusions:* These microsatellite loci will provide useful tools for further population genetic and landscape genetic studies on *Q. variabilis*.

Key words: 454 pyrosequencing; Fagaceae; magnetic bead enrichment; microsatellite markers; *Quercus variabilis*.

Quercus variabilis Blume (Fagaceae) is a tree native to East Asia found mainly in warm-temperate and subtropical regions (Chen et al., 2012). Previous studies on this species have focused on its physiology and ecology (Lei et al., 2013; Du et al., 2014). To date, only a few studies have reported on the genetic diversity and population structure of this species using isozyme, chloroplast DNA, and microsatellite markers (Zhou et al., 2003; Xu et al., 2004; Chen et al., 2012). However, the polymorphic isozyme loci are limited in number and the chloroplast DNA is maternally inherited, which limits the comprehensive understanding of its genetic diversity and population structure. Xu et al. (2004) developed 16 microsatellite loci for *Q. variabilis* using primer pairs from three other *Quercus* L. species; however, these loci do not meet the requirement of high-density markers in landscape genetics (Hall and Beissinger, 2014). Thus, developing a greater number of microsatellite loci for this species is necessary for population genetic and landscape genetic studies of *Q. variabilis* to progress. Microsatellite markers, or simple sequence repeats (SSRs), are widely used in population genetic and landscape genetic studies, providing useful information to plant natural resources management and promoting future long-term strategies (Ohtsuki et al., 2014; Lei et al., 2015). Next-generation sequencing allows for rapid development of a large number of SSR markers (Huang et al., 2014; Kumar et al., 2014). Here, we report the development of microsatellite markers for *Q. variabilis* using Roche 454 pyrosequencing combined with a magnetic bead enrichment protocol.

[1]Manuscript received 24 March 2015; revision accepted 21 April 2015.

Funding for this project was provided by the National Natural Science Foundation of China (31100272) and Henan Education Department Nature Science Research (2011B220004).

[3]Author for correspondence: liyongrui1@126.com

doi:10.3732/apps.1500032

METHODS AND RESULTS

Fresh leaves from 24 individuals of *Q. variabilis* were collected in two wild populations in Henan and Shandong provinces (Song Mountain: 34°28′08″N, 113°05′07″E; Meng Mountain: 35°29′34″N, 117°47′58″E) during May 2012. The samples were stored in silica gel until genomic DNA extraction and genotyping. Vouchers were deposited at the herbarium of Henan Agricultural University (Herbarium accession no. HHAU-2101-2112 [Song Mountain population] and HHAU-2113-2124 [Meng Mountain population]). Genomic DNA was extracted from approximately 30 mg of dried leaf tissue using a Tiangen Biotech (Beijing, China) DNA extraction kit (Plant no. DP305) following the manufacturer's protocol. A shotgun library was prepared by shearing 1 µg of genomic DNA using the DNA Library Preparation Kit (Roche Applied Science, Indianapolis, Indiana, USA) following the GS FLX+ library preparation protocol. The shotgun library was further enriched by eight 5′-biotinylated probes, namely, $(AG)_{10}$, $(AC)_{10}$, $(AAC)_8$, $(ACG)_8$, $(AAG)_8$, $(AGG)_8$, $(ACAT)_6$, and $(ATCT)_6$, and eventually sequenced on a Roche 454 GS FLX platform. A 1/16th run was performed using this pyrosequencing system. MISA software (Thiel et al., 2003) was used to search for microsatellites, while Primer3 software (Rozen and Skaletsky, 1999) was used to design primer pairs. Primer design parameters were set as follows: primer size = 18–27 nucleotides with 20 as optimum; primer melting temperature (T_m) = 55–65°C with 60°C as optimum; primer GC% = 40–60%; product size = 100–500 bp with 200 bp as optimum; and CG clamp = 3. PCRs were performed with a single microsatellite primer pair in a 30-µL reaction mixture containing 30 ng of genomic DNA, 0.2 mmol/L of each dNTP, 0.3 µmol/L of each primer, 3 µL 10× polymerization buffer, and 1 unit of *Taq* polymerase (Tiangen Biotech). PCR reactions were performed in a PTC-200 Thermal Cycler (MJ Research, Watertown, Massachusetts, USA) as follows: an initial denaturation at 95°C for 5 min; followed by 35 cycles at 94°C for 40 s, locus-specific annealing temperature (Table 1) for 45 s, and 72°C for 50 s; and a final extension at 72°C for 5 min. The PCR products were resolved on 8% native polyacrylamide gel and visualized by silver staining. The band size was determined in comparison with a 50-bp DNA ladder (TaKaRa Biotechnology Co., Dalian, Liaoning, China). Analyses of the number of alleles per locus (A), observed heterozygosity (H_o), expected heterozygosity (H_e), Hardy–Weinberg equilibrium (HWE), and linkage disequilibrium (LD) were performed using GENEPOP version 4.2 (Rousset, 2008). Analysis of the inbreeding coefficients (F_{IS}) and its significance testing were performed using FSTAT version 2.9.3.2 (Goudet, 1995).

TABLE 1. Primer sequences and characterization for 34 microsatellite loci isolated from *Quercus variabilis*.

Locus	Primer sequences (5′–3′)	Repeat motif	T_a (°C)	Allele size range (bp)	*A*	GenBank accession no.
Qv020	F: CATCGTCGAAGTCGAAGTCA	$(AAG)_6$	50	199–205	3	KP854060
	R: CTTCGATTTCACACGCTCAA					
Qv024	F: GGTTATCTCATTGCTCCTGAAAA	$(ATT)_5$	52	124–130	2	KP854064
	R: GTTGCCAACTCATCATCCCT					
Qv025	F: CCTGAGAGAGGGGTGTGAGA	$(AG)_6TGAT(GA)_8$	52	180	1	KP854065
	R: TCCTCCCCACAGACAAACTC					
Qv026	F: TCCATTGCCACACAACTCAT	$(GA)_{15}$	52	180–210	6	KP854066
	R: CCTCTAAGCCCCTCTCTCTCA					
Qv028	F: TAATTTGAGGCATGCACAGC	$(GA)_7$	48	245–255	3	KP854068
	R: GTGCCAAAATTGGGTTGAGT					
Qv030	F: CCCATTACCCGATCTTGCTA	$(CT)_6$	48	390	1	KP854070
	R: CTCACCGTCACAACAACCAC					
Qv033	F: CAAAACTTTTTGTTGCTGTGAA	$(AC)_7ATG(CA)_6$	48	229	1	KP854073
	R: GCCCAAACAAACATTCCATC					
Qv034	F: TAAGCCTGGATCCTCTGTGG	$(GAA)_6(GAT)_5$	48	249–267	2	KP854074
	R: CCTTTGGTGCTCTTGCTAGG					
Qv035	F: TCACAAACACAAAGAACACTCTCA	$(AC)_6$	54	116	1	KP854075
	R: TGGATCTGACGTGGCATTTA					
Qv036	F: GCCCTGGTAACGGGGTATTA	$(ATAA)_5$	50	395	1	KP854076
	R: CTTTTGAGGCCGAGAACAAG					
Qv040	F: TCCCCCTTACAAAAGGAATG	$(AC)_8$	50	400	1	KP854080
	R: TCACAAGTGTTTACATGCGCT					
Qv041	F: CGCTGGCCTTAACAGTGATT	$(AG)_6$	52	263–265	2	KP854081
	R: GGCCACTGACAATGAGGAAT					
Qv042	F: AGGCTCTGATACCATGTGGG	$(AG)_8$	50	314	1	KP854082
	R: GAAGCCATGTCCTTCAGCTC					
Qv048	F: AAAGGGGAAGAAAACCCTGA	$(AGA)_9$	48	189–195	2	KP854088
	R: CTGTTTTCCTCTTCACCCCA					
Qv049	F: AGAAACCCACAGCCACAAAC	$(CCT)_5$	48	255–264	2	KP854089
	R: ATTCCAGCATGTTCCAAAGC					
Qv056	F: CGGACCTTCCCATTACAGAA	$(AG)_7$	54	276	1	KP854096
	R: TGCGTGCTAGTAGGTTTGGA					
Qv060	F: GGCCTAGGAAGCTCTAGCTCA	$(GAA)_5$	52	354–363	2	KP854100
	R: GGCCTCTCTGGATGATGGTA					
Qv062	F: TCAACAGCCAACAGACAAGC	$(GA)_6$	48	308	1	KP854102
	R: GATGATTCTAATCTCACATCTCGG					
Qv063	F: TGGGAAAAATTAGTTTTGGGG	$(CAG)_5$	48	103	1	KP854103
	R: GATCCTGCTGGTGAGGATGT					
Qv070	F: TTATGAACTAAGGGCCCACG	$(AG)_8$	48	380	1	KP854110
	R: CTGAGGGTGCACTGTTCAGA					
Qv071	F: TAAAATGCCTGGAGTGGACC	$(ATG)_6$	54	368	1	KP854111
	R: AGCAGCCACCTCAGTAGGAA					
Qv075	F: TAGTCATCCTCGTTTTGGGC	$(TAG)_9$	50	325–348	3	KP854115
	R: CACCTTGAAGGAGACAAGGG					
Qv076	F: GAGGACCCAAATTTGAAGCA	$(GA)_6$	54	335	1	KP854116
	R: TGGATTCACCGATTCACTCA					
Qv078	F: AGGACCAAGGGTAGCTGGTT	$(GT)_7$	48	345	1	KP854118
	R: GGGTGACTGCACACGTTATG					
Qv081	F: ATTTGGTTTCAGCTCCCCTT	$(TTC)_{14}$	50	320	1	KP854121
	R: GCAATAGGCAAAGGATTGAGA					
Qv083	F: GGGATTGAACCTCAACCTCA	$(GAT)_6$	48	165–174	2	KP854123
	R: TTTGCATTTACTCCCTTGCC					
Qv084	F: CAATGTACTCGGGTTGGCTT	$(AT)_8$	48	173	1	KP854124
	R: CCTCCAGATGGTTGCAAGTT					
Qv089	F: GGGTGAGATTGAAAAGCCAA	$(AAG)_7(AGG)_5$	48	217–229	4	KP854129
	R: GCTTCCTCTTCCGCTTACCT					
Qv092	F: GAAATGTAAGGTGTATTGTAAAAATGG	$(TA)_9$	48	336–344	3	KP854132
	R: TCTTGACCCTATCGGACGAC					
Qv093	F: TCGTGCGAAATGAGACTTTG	$(TG)_7$	48	195	1	KP854133
	R: CGTGTTGCTAAAATATCGGTCA					
Qv094	F: AAGGCCGAAAGAGGAAGAAC	$(AG)_{14}…(GA)_{10}$	48	248–256	4	KP854134
	R: TCCTCATTTGTCACTCCCAA					
Qv095	F: GCATTTTGTCAAAGGAATAGCC	$(ACA)_{16}$	52	110	1	KP854135
	R: AGATTGCCAAAGGGGAGATT					
Qv096	F: AAGCACACGAATCCTCCAAA	$(AC)_8$	52	131	1	KP854136
	R: GATTTGGCCATCACAAAGGT					
Qv097	F: TGTTTTGCTTGGTTGGTTGA	$(AG)_{16}$	54	223	1	KP854137
	R: TCTTCTATGCTCACGGGCTT					

Note: *A* = number of alleles; T_a = annealing temperature.

TABLE 2. Genetic diversity parameters for two populations of *Quercus variabilis*.[a]

	Meng Mountain ($N = 12$)				Song Mountain ($N = 12$)			
Locus	A	H_e	H_o	F_{IS}	A	H_e	H_o	F_{IS}
Qv020	2	0.489	0.750	−0.571	3	0.540	0.750	−0.414
Qv024	2	0.042	0.042	0.000	2	0.159	0.167	−0.048
Qv026	6	0.667	0.167 (0.000)*	0.759	4	0.721	0.417 (0.044)*	0.433
Qv028	3	0.540	0.750	−0.414	3	0.453	0.583	−0.305
Qv034	2	0.489	0.750	−0.571	2	0.224	0.250	−0.125
Qv041	2	0.518	0.583	−0.132	2	0.480	0.583	−0.228
Qv048	2	0.228	0.250	−0.100	2	0.159	0.167	−0.048
Qv049	2	0.159	0.167	−0.048	2	0.344	0.417	−0.222
Qv060	2	0.290	0.333	−0.158	2	0.083	0.083	0.000
Qv075	3	0.413	0.500	−0.224	2	0.290	0.333	−0.158
Qv083	2	0.159	0.167	−0.048	2	0.344	0.417	−0.222
Qv089	4	0.656	0.583	0.115	3	0.301	0.333	−0.114
Qv092	3	0.552	0.417	0.254	2	0.392	0.500	−0.297
Qv094	4	0.601	0.500	0.175	4	0.645	0.500	0.233

Note: A = number of alleles; F_{IS} = analysis of the inbreeding coefficient; H_e = expected heterozygosity; H_o = observed heterozygosity.
[a]Geographic coordinates of the localities: Meng Mountain: 35°29′34″N, 117°47′58″E; Song Mountain: 34°28′08″N, 113°05′07″E.
*Significant deviation from Hardy–Weinberg equilibrium.

Our results suggest that next-generation sequencing technology combined with the magnetic bead enrichment protocol is a rapid and cost-effective means to identify large numbers of polymorphic microsatellite loci in *Q. variabilis.* By taking advantage of the Roche 454 sequencing platform, 63,851 individual sequence reads with an average length of 327 bp were obtained. A total of 2121 microsatellites were identified that contained enough flanking sequences for primer design. These sequences were deposited in GenBank (KP854041–KP856161). To test the amplification efficiency of these microsatellite loci, we randomly selected 100 loci to design primer pairs and perform PCRs. Sixty-six primer pairs were discarded for unsuccessful or nontarget amplification, whereas the remaining 34 primer pairs were tested for polymorphisms (Table 1). The 34 microsatellite sequences were compared with others in GenBank through BLAST searches. Qv040 and Qv089 matched significantly with the microsatellite CsCAT35 sequence (GenBank accession no. FJ868967) in *Castanea sativa* Mill. (Fagaceae) and REO_390 sequence (GenBank accession no. KJ207438) in *Q. rubra* L. (Fagaceae), respectively. Qv078 matched significantly with the mRNA sequence (GenBank accession no. XM_011077301) in *Sesamum indicum* L. (Pedaliaceae).

Fourteen of the 34 tested loci in 24 individuals from two *Q. variabilis* wild populations were polymorphic, with A ranging from two to six. H_e and H_o ranged from 0.042 to 0.721 and 0.042 to 0.750, respectively (Table 2). Locus Qv026 deviated significantly from HWE ($P < 0.05$) because of excessive homozygosity. F_{IS} varied from −0.571 to 0.759 (Table 2), and no significant negative F_{IS} values were observed. Three pairs of loci (Qv034 and Qv060, Qv034 and Qv041, Qv028 and Qv049) showed significant LD, indicating significant allelic associations between them. These new microsatellite loci will provide useful tools for further population genetic and landscape genetic studies on *Q. variabilis.*

CONCLUSIONS

In this study, we isolated 2121 microsatellite loci from *Q. variabilis.* One hundred microsatellite loci were randomly selected to design primer pairs and perform PCRs in two wild populations. Among these tested loci, 14 microsatellite markers were polymorphic. These new microsatellite loci increase the density of the microsatellite markers for this species, which will be helpful for further population genetic and landscape genetic studies on *Q. variabilis*.

LITERATURE CITED

CHEN, D., X. ZHANG, H. KANG, X. SUN, S. YIN, H. DU, N. YAMANAKA, ET AL. 2012. Phylogeography of *Quercus variabilis* based on chloroplast DNA sequence in East Asia: Multiple glacial refugia and mainland-migrated island populations. *PLoS ONE* 7: e47268.

DU, Q. R., J. P. LEI, J. F. LIU, P. C. WANG, W. F. XIAO, AND L. PAN. 2014. Eco-physiological response of *Quercus variabilis* seedlings to increased atmospheric CO_2 and N supply. *Chinese Journal of Applied Ecology* 25: 24–30.

GOUDET, J. 1995. FSTAT (version 1.2): A computer program to calculate F-statistics. *Journal of Heredity* 86: 485–486.

HALL, L. A., AND S. R. BEISSINGER. 2014. A practical toolbox for design and analysis of landscape genetics studies. *Landscape Ecology* 29: 1487–1504.

HUANG, D. N., Y. Q. ZHANG, M. D. JIN, H. K. LI, Z. P. SONG, Y. G. WANG, AND J. K. CHEN. 2014. Characterization and high cross-species transferability of microsatellite markers from the floral transcriptome of *Aspidistra saxicola* (Asparagaceae). *Molecular Ecology Resources* 14: 569–577.

KUMAR, S., N. SHAH, V. GARG, AND S. BHATIA. 2014. Large scale in-silico identification and characterization of simple sequence repeats (SSRs) from de novo assembled transcriptome of *Catharanthus roseus* (L.) G. Don. *Plant Cell Reports* 33: 905–918.

LEI, J. P., W. XIAO, J. F. LIU, D. XIONG, P. WANG, L. PAN, Y. JIANG, AND M. H. LI. 2013. Responses of nutrients and mobile carbohydrates in *Quercus variabilis* seedlings to environmental variations using in situ and ex situ experiments. *PLoS ONE* 8: e61192.

LEI, Y. K., W. WANG, Y. P. LIU, D. HE, AND Y. LI. 2015. Adaptive genetic variation in the smoke tree (*Cotinus coggygria* Scop.) is driven by precipitation. *Biochemical Systematics and Ecology* 59: 63–69.

OHTSUKI, T., T. SHODA, Y. KANEKO, AND H. SETOGUCHI. 2014. Development of microsatellite markers for *Vitex rotundifolia* (Verbenaceae), an endangered coastal plant in Lake Biwa, Japan. *Applications in Plant Sciences* 2: 1300100.

ROUSSET, F. 2008. GENEPOP'007: A complete reimplementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources* 8: 103–106.

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.

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

XU, X. L., L. A. XU, M. R. HUAN, AND Z. R. WANG. 2004. Genetic diversity of microsatellites (SSRs) of natural populations of *Quercus variabilis. Hereditas* 26: 683–688.

ZHOU, J. Y., J. Z. GUO, Z. S. YANG, AND W. H. ZHANG. 2003. Variation of peroxidase isozyme on natural populations of *Quercus variabilis. Journal of Northwest Forestry University* 18: 33–36.