Genes Genet. Syst. (2016) **91**, p. 11–14

Microsatellite markers for the critically endangered elm species *Ulmus gaussenii* (Ulmaceae)

Qi-Fang Geng[1,2†], Jie Yang[1,2†], Jia He[1,2], Dan-Bi Wang[1,2], En Shi[1,2], Wei-Xiang Xu[1,2], Nasreen Jeelani[1,2], Zhong-Sheng Wang[1,2*] and Hong Liu[3,4]

[1]*School of Life Sciences, Nanjing University, 22 Hankou Road, Nanjing 210093, Jiangsu, China*
[2]*State Key Laboratory of Pharmaceutical Biotechnology, School of Life Sciences, Nanjing University, 22 Hankou Road, Nanjing 210093, Jiangsu, China*
[3]*International Center for Tropical Botany, Department of Earth and Environment, Florida International University, 11200 S. W. 8 Street, Miami, FL 33199, USA*
[4]*College of Forestry, Guangxi University, 100 Daxue East Road, Nanning 530004, Guangxi, China*

(Received 31 July 2015, accepted 17 November 2015; J-STAGE Advance published date: 1 April 2016)

The Anhui elm *Ulmus gaussenii* is listed as a critically endangered species by the International Union for Conservation of Nature and is endemic to China, where its only population is restricted to Langya Mountain in Chuzhou, Anhui Province. To better understand the population genetics of *U. gaussenii*, we developed 12 microsatellite markers using an improved technique. The 12 markers were polymorphic, with the number of alleles per locus ranging from two to nine. Observed and expected heterozygosities ranged from 0.021 to 0.750 and 0.225 to 0.744, respectively. The inbreeding coefficient ranged from –0.157 to 0.960. Significant linkage disequilibrium was detected for two pairs of loci, and significant deviations from Hardy-Weinberg equilibrium were found in nine loci. These microsatellite markers will contribute to the studies of population genetics in *U. gaussenii*, which in turn will contribute to species conservation and protection.

Key words: genetic analysis, microsatellite markers, Ulmaceae, *Ulmus gaussenii*

INTRODUCTION

Ulmus gaussenii W. C. Cheng (Ulmaceae) was named by Prof. Wanjun Zheng in 1955, and its natural range is restricted to the Inebriate Pavilion of the limestone Langya Mountain in Anhui Province, eastern China (Zheng et al., 1999). *U. gaussenii* (Anhui elm) is now one of the rarest and most endangered elm species, with approximately 26 mature trees known to survive in nearly 10 hectares of wild area (Sun, 2006). *U. gaussenii* is a medium deciduous tree and listed as a critically endangered species by the International Union for Conservation of Nature (Fu and Jin, 1992). To provide an effective conservation management plan for this critically endangered species, there is an urgent need to evaluate its genetic variation. Microsatellites, or simple sequence repeats (SSRs), are highly polymorphic, codominant markers that have been widely used in studies of population genetics (Zane et al., 2002; Hedrick, 2004; Abdul-Muneer, 2014).

Previous studies on *U. gaussenii* mainly focused on its habitat characteristics, physiology, cultivation and phylogenetics (Fu and Liu, 1999; Zheng et al., 1999; Fu and Huang, 2002; Sun, 2006, 2009; Han et al., 2011). However, nothing is known of its population genetics. To better understand genetic diversity, population genetic structure and mating systems in *U. gaussenii*, we have developed 12 polymorphic nuclear microsatellite markers using a suppression-PCR technique (Lian et al., 2001, 2003, 2006).

MATERIALS AND METHODS

Plant materials and DNA extraction We collected fresh leaves of 48 individuals of *U. gaussenii* from one population in Langya Mountain, Chuzhou, Anhui Province, China (32°17.1317′N, 118°17.0101′E). Samples were dried in silica gel and stored at room temperature until use. Total genomic DNA was extracted from the dried leaves using a modified cetyltrimethyl ammonium bromide method (Zhou et al., 1999).

Edited by Yoshihiko Tsumura
* Corresponding author. E-mail: wangzs@nju.edu.cn
† Qifang Geng and Jie Yang contributed equally to this work.
DOI: http://doi.org/10.1266/ggs.15-00053

Isolation of microsatellite markers SSR markers were isolated using an improved technique for isolating codominant compound SSR markers (Lian et al., 2001, 2003, 2006). Briefly, genomic DNA of *U. gaussenii* was digested with the blunt-end restriction enzymes AfaI, AluI, EcoRV and HaeIII. The restricted fragments were then ligated with a specific blunt adaptor (consisting of a 48-mer: 5′-GTAATACGACTCACTATAGGGCACGCGTG-GTCGACGGCCCGGGCTGGT-3′; and an 8-mer with the 3′-end capped with an amino residue: 5'-ACCAGCCC-NH_2-3′) using a DNA ligation kit (Takara Bio, Shiga, Japan). The ligated fragments were treated with ddGTP using AmpliTaq Gold (Applied Biosystems, Carlsbad, CA, USA) to block polymerase-catalyzed extension of the 8-mer adaptor strand.

Fragments flanked by an SSR at one end were amplified from the AfaI, AluI, EcoRV and HaeIII DNA libraries using one of the compound SSR primers $(AC)_6(TC)_5$, $(TC)_6(AC)_5$, $(AC)_6(AG)_5$ or $(TC)_6(TG)_5$ and an adaptor primer (5′-CTATAGGGCACGCGTGGT-3′). The amplified smeared fragments were cloned into the Trans *pEASY*-T1 Cloning vector system in accordance with the manufacturer's instructions (Beijing Transgen Biotech, Beijing, China). Recombinant clones were identified using blue/white screening on LB agar plates containing ampicillin, X-gal and IPTG. Insert-positive clones were amplified using the M13 forward and reverse primers. After electrophoresis on a 1.5% agarose gel, amplified fragments longer than 350 bp were sequenced using the M13 primer (Beijing Genomics Institute, Shenzhen, China). For each fragment containing the $(AC)_6(TC)_n$, $(TC)_6(AC)_n$, $(AC)_6(AG)_n$ or $(TC)_6(TG)_n$ compound SSR sequences at one end, a specific primer was designed using Primer3 software (Rozen and Skaletsky, 2000). The specific primer and corresponding compound SSR primer were used as a compound SSR marker.

To characterize the developed compound SSR markers, we performed a polymerase chain reaction (PCR) amplification in a 10-μl reaction mixture containing about 10 ng of template DNA, 1× AmpliTaq Gold 360 Master Mix (Applied Biosystems), 0.4 μl of 360 GC Enhancer (Applied Biosystems), and 0.5 μM of each specific primer and a fluorescent dye-labeled (6-FAM, VIC, NED or PET; Applied Biosystems) compound SSR primer.

PCR was performed with a Veriti Thermal Cycler (Applied Biosystems) under the following conditions: 10 min at 95 °C; 38 cycles of 30 s at 95 °C, 30 s at the annealing temperature of each designed specific primer, and 1 min at 72 °C; and a 7-min extension at 72 °C after the final cycle. The degree of polymorphism of SSR loci was surveyed by automated scanning detection using an ABI 3730 genetic analyzer (Applied Biosystems). Allele sizes were determined using GeneMapper analysis software version 4.0 (Applied Biosystems) with a LIZ-500 DNA size standard (Applied Biosystems).

Data analysis Null allele frequencies were determined with MICRO-CHECKER version 2.2.3 with the Oosterhout algorithm method (Van Oosterhout et al., 2004). The number of alleles per locus (N_A), observed heterozygosity (H_O) and expected heterozygosity (H_E) were calculated using GENALEX 6.5 (Peakall and Smouse, 2006). Deviation from Hardy-Weinberg equilibrium (HWE) and linkage disequilibrium (LD) were tested using Genepop version 4.2 (Raymond and Rousset, 1995; Rousset, 2008). The *P* values for HWE and LD were corrected for multiple comparisons by applying a sequential Bonferroni correction (Rice, 1989). Inbreeding coefficients (F_{IS}) were calculated using the FSTAT program version 2.9.3 (Goudet, 1995). The deviation from zero was tested in this population by 1000 per mutation tests with a sequential Bonferroni correction.

To evaluate whether the population had undergone recent bottlenecks, we adopted Wilcoxon's signed-rank test with 1,000 iterations using BOTTLENECK software version 1.2.02 (Piry et al., 1999). We ran BOTTLENECK using three possible mutation models: Infinite Allele Model (IAM), Stepwise Mutation Model (SMM) and Two Phase Model (TPM; 30% IAM and 70% SMM). These models are based on the principle that in a recently bottlenecked population, the observed heterozygosity is higher than the expected equilibrium heterozygosity estimated from the observed number of alleles (Cornuet and Luikart, 1996).

RESULTS AND DISCUSSION

A total of 172 clones with positive inserts were chosen and sequenced. One hundred and thirty-nine sequences were found to contain compound SSR motifs, of which 31 were discarded because the repeat was too close to one end of the insert to design a suitable primer or because the sequence was identical to that of another insert(s). Therefore, only 108 sequences [$(AC)_6(TC)_n$ (3), $(TC)_6(AC)_n$ (27), $(AC)_6(AG)_n$ (43) and $(TC)_6(TG)_n$ (35)] were suitable for designing specific primers. Using these primer pairs, 49 loci were successfully amplified and yielded a single band of the right size.

For further characterization, the 48 individuals from Langya Mountain were genotyped using the procedure described above. Twelve of the 49 loci were polymorphic. Allelic diversity per locus ranged from two to nine with an average of 4.83 (Table 1). The observed (H_O) and expected (H_E) heterozygosities ranged from 0.021 to 0.750 and 0.225 to 0.744, respectively. Inbreeding coefficient (F_{IS}) values ranged from –0.157 to 0.960 with an average of 0.340 (Table 1). Significant deviation from LD was detected in two pairs of loci (Ulmg30 and Ulmg42, Ulmg100 and Ulmg102) following Bonferroni correction ($P < 0.05$). Significant deviations from HWE were found at nine loci after the Bonferroni corrections ($P < 0.05$,

Table 1. Characteristics of 12 microsatellite loci isolated from *Ulmus gaussenii*

Locus	Primer sequence 5′-3′ (fluorescent dye)	Repeat motif	T_a (°C)	Size range (bp)	N_a	H_O	H_E	F_{IS}	r	GenBank accession no.
Ulmg27*	F: GAGTCTGCAAACGCACGATCACCT	$(TC)_6(AC)_8$	62	156–164	4	0.116	0.532	**0.786**	0.340	KT319055
	R: TCTCTCTCTCTCTCACACAC (FAM)								yes	
Ulmg30*	F: CGTGAGCAAAGTCGAGCTATTGCT	$(AC)_6(TC)_5$	62	169–177	5	0.396	0.576	**0.322**	0.132	KT319056
	R: ACACACACACACACTCTCTC (VIC)								yes	
Ulmg40*	F: TGTCAAAGGGTTGCTATTGTGAAT	$(AC)_6(TC)_5$	57	149–153	3	0.021	0.508	**0.960**	0.403	KT319057
	R: ACACACACACACACTCTCTC (NED)								yes	
ulmg42*	F: GTCTGTATGGGTCCTCATACAACT	$(TC)_6(AC)_6$	59	170–178	4	0.404	0.478	0.164	0.058	KT319058
	R: TCTCTCTCTCTCTCACACAC (FAM)								no	
ulmg44*	F: AGTAGTGTCCAAGAGATTATGGAG	$(TC)_6(AC)_5$	58	192–200	4	0.250	0.307	0.195	0.114	KT319059
	R: TCTCTCTCTCTCTCACACAC (VIC)								no	
Ulmg45*	F: GTGCTAAGATTAGCTTGGTGGTTA	$(TC)_6(AC)_5$	58	160–180	8	0.047	0.417	**0.891**	0.341	KT319060
	R: TCTCTCTCTCTCTCACACAC (VIC)								yes	
Ulmg61*	F: CCACAATCTCAAAGTCATACTCCT	$(AC)_6(AG)_5$	58	194–200	3	0.063	0.225	**0.729**	0.226	KT319061
	R: ACACACACACACACAGAGAG (VIC)								yes	
Ulmg65*	F: CCTGCATGAATGTCAATATA	$(TC)_5(TG)_8$	55	158–170	5	0.688	0.744	0.086	0.039	KT319062
	R: TCTCTCTCTCTCTCTGTGTG (PET)								no	
Ulmg100	F: CACTAATGTTTCAACTAGCT	$(TC)_6(AC)_{11}$	55	120–136	5	0.750	0.642	–0.157	–0.085	KT319063
	R: TCTCTCTCTCTCTCACACAC (PET)								no	
Ulmg102*	F: CTCAGATTGGAAGGCAAATA	$(TC)_6(AC)_6$	56	159–221	9	0.395	0.653	**0.404**	0.195	KT319064
	R: TCTCTCTCTCTCTCACACAC (VIC)								yes	
Ulmg107	F: GAAGGTGACCCGTTGTTAGG	$(TC)_6(AC)_{13}$	59	125–147	6	0.447	0.429	–0.030	–0.011	KT319065
	R: TCTCTCTCTCTCTCACACAC (PET)								no	
Ulmg108	F: TAAGAATACGGTGTTAGAACCTACA	$(TC)_6(AC)_6$	55	147–151	2	0.432	0.469	0.090	0.039	KT319066
	R: TCTCTCTCTCTCTCACACAC (PET)								no	
All					4.83	0.334	0.498	**0.340**		

T_a: annealing temperature of the primer pair; H_o: observed heterozygosity; H_E: expected heterozygosity; F_{IS}: inbreeding coefficient, with F_{IS} in bold indicating significant deviation from zero with a sequential Bonferroni correction; *: significant deviation from Hardy-Weinberg equilibrium following Bonferroni correction, with $P < 0.05$; r: null allele frequencies calculated with the Oosterhout algorithm (Van Oosterhout et al., 2004), with presence (yes) and absence (no) indicated below.

Table 1). The null allele frequencies as predicted from MICRO-CHECKER software ranged from –0.085 to 0.403 per locus (Table 1). The BOTTLENECK test indicated a recent population bottleneck in the *U. gaussenii* population under IAM, SMM and TPM (Wilcoxon test, $P < 0.05$), which may account for the significant LD and HWE deviations at these studied loci.

The significant deviation from LD may be due to demographic effects, such as a recent population bottleneck (as noted above) or genetic drift, rather than to a linkage relationship between these loci. The deviation from HWE may be due to several factors such as the presence of null alleles, a high level of inbreeding, a bottleneck effect, restricted population size, a founder effect, the Wahlund effect or genetic drift. In our case, for six loci (Ulmg27, Ulmg30, Ulmg40, Ulmg45, Ulmg61 and Ulmg102), the deviation from HWE as well as their significant F_{IS} might be attributed to null alleles, or to a high level of inbreeding. While the probable presence of null alleles at these six loci was inferred from MICRO-CHECKER, a high level of inbreeding in *U. gaussenii* could be the most important contributing factor for the deviation from HWE and high F_{IS} in these six loci (Islam et al., 2014). The restricted population size and paucity of reproductive individuals may result in a high level of inbreeding in the *U. gaussenii* population.

The remaining three loci (Ulmg42, Ulmg44 and Ulmg65) exhibiting deviation from HWE were not identified as possibly having null alleles and, in these cases, other factors should be considered to explain the deviation. Among the causes of HWE enumerated above, we consider inbreeding to be unlikely, as the F_{IS} values for these three loci were small. One of the remaining factors such as a bottleneck effect, restricted

population size, a founder effect, the Wahlund effect or genetic drift is more likely to cause the observed HWE deviation.

The 12 polymorphic microsatellite markers described here will be available for population genetic studies of *U. gaussenii*, which in turn will contribute to the conservation and protection of this endangered species.

This work was supported by the National Natural Science Foundation of China (Grant No. 31100270). We thank Prof. Javier Francisco Ortega of Florida International University and Associate Prof. Jingtao Sun of Nanjing Agricultural University for their kind comments and suggestions on this manuscript.

REFERENCES

Abdul-Muneer, P. M. (2014) Application of microsatellite markers in conservation genetics and fisheries management: recent advances in population structure analysis and conservation strategies. Genet. Res. Int. **2014**, 691759.

Cornuet, J. M., and Luikart, G. (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. Genetics **144**, 2001–2014.

Fu, L. G., and Jin, J. M. (1992) China Plant Red Data Book: Rare and Endangered Plants, Volume 1, pp. 692–693. Science Press, Beijing.

Fu, S. L., and Huang, B. L. (2002) Analysis on succession trend and physiological characters of secondary forest tree species in limestone mountain of eastern Anhui province. Scientia Silvae Sinicae **38**, 50–55.

Fu, S. L., and Liu, S. Q. (1999) Study on growth and physiological characters of rare tree *Ulmus chenmoui* during seeding. Journal of Fujian College of Forestry **19**, 253–255.

Goudet, J. (1995) FSTAT (Version 1.2): A computer program to calculate F-statistics. J. Hered. **86**, 485–486.

Han, X., Zhi, Y. B., Zhang, B. W., Zhou, Z. Z., He, J. Q., Wang, Z. L., Li, H. L., and Li, J. M. (2011) The phylogenetic relationships between *Ulmus gaussenii* Cheng and its close relatives in genus Ulmus based on the ITS and atpB-rbcL. Journal of Anhui University (Natural Science Edition) **35**, 98–107.

Hedrick, P. W. (2004) Recent developments in conservation genetics. For. Ecol. Manag. **197**, 3–19.

Islam, M. S., Lian, C. L., Kameyama, N., and Hogetsu, T. (2014) Low genetic diversity and limited gene flow in a dominant mangrove tree species (*Rhizophora stylosa*) at its northern biogeographical limit across the chain of three Sakishima islands of the Japanese archipelago as revealed by chloroplast and nuclear SSR analysis. Plant Syst. Evol. **300**, 1123–1136.

Lian, C. L., Zhou, Z. H., and Hogetsu, T. (2001) A simple method for developing microsatellite markers using amplified fragments of inter-simple sequence repeat (ISSR). J. Plant Res. **114**, 381–385.

Lian, C. L., Hogetsu, T., Matsushita, N., Guerin-Laguette, A., Suzuki, K., and Yamada, A. (2003) Development of microsatellite markers from an ectomycorrhizal fungus, *Tricholoma matsutake*, by an ISSR-suppression-PCR method. Mycorrhiza **13**, 27–31.

Lian, C. L., Wadud, M. A., Geng, Q. F., Shimatani, K., and Hogetsu, T. (2006) An improved technique for isolating codominant compound microsatellite markers. J. Plant Res. **119**, 415–417.

Peakall, R., and Smouse, P. E. (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol. Ecol. Notes **6**, 288–295.

Piry, S., Luikart, G., and Cornuet, J. M. (1999) BOTTLENECK: a computer program for detecting recent reductions in the effective population size using allele frequency data. J. Hered. **90**, 502–503.

Raymond, M., and Rousset, F. (1995) GENEPOP (Version 1.2): Population genetics software for exact tests and ecumenicism. J. Hered. **86**, 248–249.

Rice, W. R. (1989) Analyzing tables of statistical tests. Evolution **43**, 223–225.

Rousset, F. (2008) GENEPOP'007: a complete re-implementation of the GENEPOP software for Windows and Linux. Mol. Ecol. Resour. **8**, 103–106.

Rozen, S., and Skaletsky, H. (2000) Primer3 on the WWW for General Users and for Biologist Programmers. Humana Press, Totowa, NJ, USA.

Sun, Z. Y. (2006) Studies on the cultivating techniques of rare plant *Ulmus gaussenii*. Modern Agricultural Sciences and Technology, 15–16.

Sun, Z. Y. (2009) A study on cultivation experiments of rare trees, *Ulmus chenmoui* and *Ulmus gaussenii*. Anhui Forestry, 67–68.

Van Oosterhout, C., Hutchinson, W. F., Wills, D. P. M., and Shipley, P. (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Mol. Ecol. Notes **4**, 535–538.

Yoshida, Y., Honjo, M., Kitamoto, N., and Ohsawa, R. (2009) Reconsideration for conservation units of wild *Primula sieboldii* in Japan based on adaptive diversity and molecular genetic diversity. Genet. Res. **91**, 225–235.

Zane, L., Bargelloni, L., and Patarnello, T. (2002) Strategies for microsatellite isolation: a review. Mol. Ecol. **11**, 1–16.

Zheng, C. G., Wang, M. Y., and Sun, Q. X. (1999) The ecological characteristics of *Ulmus gaussenii* and its conservation measures. Resource Development & Market **15**, 204–205.

Zhou, Z. H., Miwa, M., and Hogetsu, T. (1999) Analysis of genetic structure of a *Suillus grevillei* population in a *Larix kaempferi* stand by polymorphism of inter-simple sequence repeat (ISSR). New Phytol. **144**, 55–63.