ISSN 1022-7954, Russian Journal of Genetics, 2020, Vol. 56, No. 1, pp. 123–127. © Pleiades Publishing, Inc., 2020.

SHORT COMMUNICATIONS

Development of 14 Microsatellite Markers in the Endangered Relict Plant *Craigia yunnanensis* (Tiliaceae)

Y. L. Chen[a, b, c], J. Yang[a, b, *], and W. B. Sun[a, b, **]

[a] *Yunnan Key Laboratory for Integrative Conservation of Plant Species with Extremely Small Populations, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, 650201 China*

[b] *Key Laboratory for Plant Diversity and Biogeography of East Asia, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, 650201 China*

[c] *University of Chinese Academy of Sciences, Beijing, 100049 China*

**e-mail: yangjing@mail.kib.ac.cn*

***e-mail: wbsun@mail.kib.ac.cn*

Received January 17, 2019; revised February 12, 2019; accepted February 18, 2019

Abstract—*Craigia yunnanensis* (Tiliaceae) is an endangered relic tree in China. Its wild populations have been seriously damaged and fragmented due to severe human disturbance. In this study, we developed microsatellite markers in order to study the variation of the mating system of *C. yunnanensis*. We collected 19 067 214 unigenes sequences by shallow sequencing of the genome. A bioinformatics screening identified 15 235 unique and putative microsatellites, from which 590 novel microsatellite markers were developed. We designed 70 primer pairs and successfully amplified 14 of them in 30 individuals. The number of alleles per locus ranged from 1 to 12. The observed and expected heterozygosities ranged from 0.000 to 1.000 and 0.000 to 0.906, respectively. These microsatellite loci will enrich the genetic resources to develop functional studies and conservation strategies for this endangered relict species.

Keywords: endangered species, microsatellite, transcriptome sequencing, *Craigia yunnanensis*

DOI: 10.1134/S1022795420010032

Craigia yunnanensis W.W. Sm. and W.E. Evans (Tiliaceae) is a deciduous canopy tree distributed in southern China and northern Vietnam. Due to years of anthropogenic disturbances, the wild population of *C. yunnanensis* has reduced dramatically. The genus *Craigia* (Tiliaceae) was once widespread throughout the Northern Hemisphere, and abundant across Europe, North America and East Asia during the Tertiary period [1, 2]. Only *C. yunnanensis* survived in wild now and has been listed as a national key protected wild plants (Category II) in China (http://www.forestry.gov.cn/yemian/minglu1.htm) and evaluated as 'endangered' by Threatened Species List of China's Higher Plants [3]. In the IUCN (International Union for Conservation of Nature) Red List, *C. yunnanensis* is categorized as 'endangered' too (http://www.iucnredlist.org/details/32335/0). Extensive survey of natural habitats of *C. yunnanensis* located six small, remnant populations in Yunnan province [4]. We have studied the genetic diversity and population structure using the amplified fragment length polymorphism (AFLP) marker technique and found the genetic diversity of *C. yunnanensis* was moderate at the species level, but low at regional and population levels [5]. It was found that reproductive efficiency varied in populations which suffered different degrees of fragmentation and isolation, and that pollination is a possible weak link for *C. yunnanensis* under habitat degeneration. Genetic markers such as microsatellites have been successfully used in studying germplasm accessions, genetic linkage maps and genetic diversities throughout cultivars or wild plants, helping to develop conservation strategies for endangered species. It will be employed to study the variation of mating system in several populations of *C. yunnanensis*, which are under different levels of fragmentation, in order to analyze the impact of human disturbance on the process and the pattern of pollen dispersion. And this research could give scientific implications for in-situ conservation and renewal of *C. yunnanensis*.

Off the wide range of DNA markers in use, microsatellites or simple sequence repeat (SSR) markers are extensively employed in plant studies. SSR markers have the advantage of producing mostly codominant markers. By providing highly variable genetic markers at distinct loci, microsatellite analysis can be applied to evaluating pollen or seed dispersal distances and estimating effective population sizes [6]. Using SSR markers, Dawson et al. studied the natural population of an endangered species *Gliricidia sepium* and

assessed the pollen gene flow [7]; Yang et al. assessed parental relationships of the seedlings from a critical endangered maple, *A. yangbiense*, and proposed a conservation strategy for this species [8]. However, no microsatellite markers have been reported for *C. yunnanensis*. In this study, we isolated 14 novel microsatellite markers from the genome of *C. yunnanensis* using high-throughput sequencing techniques. The information in this study will be useful for genetic and evolutionary studies, providing important information for better management of *C. yunnanensis*.

In this study, the total genomic DNA from *C. yunnanensis* was used for genome sequencing. Genomic DNA was extracted from the leaves of 30 individuals from 5 populations using an improved CTAB (cetyltrimethylammonium bromide) method [9]. The leaf material was obtained include from Fadou population (23°22′ N, 104°46′ E; 1471 m), Xiajinchang population (23°10′ N, 104°49′ E; 1474 m), and other populations (Huguo, Jiangdong, Sudian) samples from the seedling of *Craigia yunnanensis* were planted in the greenhouse in Kunming Institute of Botany, Chinese Academy of Sciences. The concentration of each DNA sample was measured using a spectrophotometer at 260 and 280 nm.

Extracted DNA was used for a library preparation with a NEBNext Ultra II DNA Library Prep Kit for Illumina (NEB). Then, a normalized DNA library was constructed and subsequently sequencing was performed on the HiSeq X–Ten Benchtop Sequencer (Illumina) using the 2 × 150 bp read mode, produced 19067214 clean reads. CLCGenomics Workbench 7.5.1 (CLCBio) was used to run *de novo* assembly, resulting in 15235 contigs. SSRs were detected using QDD 2.1 Beta [10] with default settings. A total of 3203 SSRs were identified.

The microsatellite primers sets were designed and synthesized through Primer 5.0 software, which was used to check against potential primer dimers, hairpin structures and the occurrence of mismatches. Under following criteria: (1) primers' length ranged from 18 to 27 bases; (2) PCR product size ranged from 150 to 300 bp; (3) annealing temperature of 50 to 55°C; (4) a GC content of 40–60%.

The polymerase chain reaction (PCR) reactions were performed in 25 μL reaction, including 70–460 ng template DNA, 0.5 μL of each primer, 0.5 mM dNTP each, 2.5 μL 10× PCR Buffer [without Mg^{2+}, 100 mM Tris-HCl (pH 8.8), 500 mM KCl], 2 mM $MgCl_2$ and 0.2 U *Taq* polymerase/μL. PCR amplification was under the following conditions: 95°C for 3 min; 10 cycles of 95°C for 30 s, annealing at 60°C 30 s, and elongation at 72°C for 30 s; 20 cycles of 95°C for 30 s, annealing at 55°C 30 s, and extension at 72°C for 30 s, and a final extension step at 72°C for 6 min. Amplification products were detected by 1% agarose gel electrophoresis to assess if the expected size was produced for each primer. Some loci did not amplify in all samples although we adjusted the PCR conditions. These loci were excluded from further testing. The PCR products were separated and visualized using the QIAxcel capillary gel electrophoresis system (QLAGEN, Irvine, California, USA).

Forward primers were labeled with a fluorescent dye 5'6-FAM on the 5'-end. 3730XL (ABI, USA) was used to detect STR samples. The allele sizes and SSR data were assessed by Gene Mapper 4.0 (Applied Biosystems). The number of alleles per locus (*A*), the observed and expected heterozygosity (H_o and H_e), and the Hardy–Weinberg equilibrium were calculated in the five populations using the software GenAlEx6.5 [11]. The PIC was calculated for each locus using CERVUS version 3.0 [12]. Microsatellite DNA sequences underlying these analyses have been deposited to GenBank under the accessions MH712436–MH712449.

590 microsatellite loci were obtained in our research through sequencing, which were suitable for design of PCR primers using our study based on the *C. yunnanensis* genome. Seventy loci (37 out of two nucleotide repeat units, 23 out of three, 4 out of four, 3 out of five and 3 out of six) were selected as candidates, of which 14 microsatellites were characterized. Fluorescently labeled primers were further synthesized for these loci. The primers were tested for polymorphism on 30 individuals collected from five populations.

We used high-throughput Illumina HiSeq X-Ten Benchtop Sequencer to develop genomic SSR markers in *C. yunnanensis*. The number of assembly of reads of the long sequences was 15235 contigs. The bioinformatics pipeline identified 590 primer pairs representing putative microsatellite loci. Of the 70 primer pairs only 20 were successfully amplified. Fourteen microsatellite loci (Table 1) showed polymorphism, the amplified fragments were in the range of 143–347 bp. Fluorescently labeled primers were further synthesized for these loci. The results showed that the number of alleles per locus ranged from 1 to 12. The population level, H_o and H_e range from 0.000 to 1.000 and 0.000 to 0.906, respectively (Table 2). The PIC values were highest in population XJC (0.630–0.898) and lowest in population HG (0–0.610). For 5 of the 14 loci showed significant deviations from Hardy–Weinberg equilibrium ($P < 0.05$).

In conclusion, we identified 14 highly polymorphic microsatellite loci for *C. yannanensis*. These microsatellite markers may provide useful tools to assess genetic diversity, population structure, and gene flow for *C. yunnanensis*. Therefore, these primers are important for further molecular research on *C. yunnanensis*, and will help in building effective conservation strategies for this critically endangered species.

Table 1. Characteristics of 14 microsatellite loci developed for *Craigia yunnanensis* (Tiliaceae)

Locus	Primer sequences (5'–3')	Repeat motif	Size range, bp	T_a, °C	Gen Bank accession no.
cons_gr936	F: AAGCCACAAACATGGAGACC	$(GGGAA)_5$	135–155	55	MH712436
	R: GGTCAACCTCCACGAGAATG				
cons_gr1184	F: CAAATTCATGACTTCTTTGGG	$(TC)_{12}$	182–214	52	MH712437
	R: GTTTCCTCTTTCTGCATGGG				
cons_gr1484	F: CAGCCATGTGAGGAAATTGA	$(TC)_{12}$	145–183	55	MH712438
	R: GCACGAGTCGGTGTTAGTGA				
cons_gr2082	F: TTTGTGCAGCTCCAATAGGT	$(AG)_{11}$	186–202	53	MH712439
	R: TTTTCCGTGTTCAAGCAATC				
cons_gr2268	F: GTCCCCTAAAAGGAGGGAAA	$(AG)_{12}$	148–196	54	MH712440
	R: TCACTTTTCTCTTCTTGGCCT				
cons_gr4165	F: TTTCCACCCTCCCTTCTCTT	$(GAA)_{12}$	132–180	55	MH712441
	R: GTCCACGTCCAAGGCATAGT				
cons_gr240	F: CGTATTGGAAGTCCACACGA	$(CT)_{19}$	122–342	54	MH712442
	R: TCTCACCAGTCTGGCAACAA				
cons_gr720	F: CCTTAACAACTTGTCGTTTTGC	$(AT)_{15}$	311–343	53	MH712443
	R: CTGCTGTCACGGACACTGTT				
cons_gr1008	F: TCTTAGAGCCGCTTGGTTGT	$(GA)_{10}$	125–289	55	MH712444
	R: CTTATGAGCGCACCTCCTGT				
cons_gr4117	F: CCGCAGCTACCATGTCTGTA	$(CT)_{10}$	255–341	55	MH712445
	R: GACGTGGCAAGAATAAGGGA				
cons_gr4727	F: GCCGTCAAGAACACTTCTGTC	$(TA)_{10}$	161–557	55	MH712446
	R: ATGCCTTCGTTTGTATTCGG				
cons_gr4891	F: TTTCGCTTTCCTGCAAAACT	$(TC)_{13}$	198–242	55	MH712447
	R: TGCTCCAAATGAAACCAACA				
cons_gr4988	F: ATCATCAACCCTCGCAAAAC	$(AGGAAG)_5$	316–346	55	MH712448
	R: AGTGCAGGCACAGCAGTAAA				
cons_gr5464	F: TGGGAAATGATCAGGAAACTG	$(GA)_{13}$	123–221	55	MH712449
	R: GCATGTCATGAGAAACAAGAGTC				

T_a = annealing temperature.

FUNDING

The study was conducted at the Germplasm Bank of Wild Species in Southwest China, Kunming Institute of Botany, Chinese Academy of Sciences. This work was supported by the National Natural Science Foundation of China (Grants no. 31600439), National Key R&D Program of China (2017YFC0505200) and Key Laboratory Construction of Yunnan Science and Technology Talents and Platform Program (Grant no. 2018DG004).

Table 2. Polymorphism of the 14 microsatellite marker in five populations of *Craigia yunnanensis*

Locus	FD ($N = 5$)					HG ($N = 5$)					JD ($N = 5$)					SD ($N = 3$)					XJC ($N = 12$)				
	A	H_o	H_e	PIC	*P*	*A*	H_o	H_e	PIC	*P*	*A*	H_o	H_e	PIC	*P*	*A*	H_o	H_e	PIC	*P*	*A*	H_o	H_e	PIC	*P*
cons_gr936	4	0.600	0.660	0.610	0.340	1	0.000	0.000	0.000	—	1	0.000	0.000	0.000	—	1	0.000	0.000	0.000	—	5	0.750	0.688	0.630	0.715
cons_gr1184	5	1.000	0.740	0.701	0.556	2	0.600	0.500	0.375	0.655	4	0.800	0.660	0.610	0.446	4	1.000	0.722	0.671	0.609	8	0.833	0.858	0.841	0.617
cons_gr1484	4	1.000	0.656	0.605	0.677	1	0.000	0.000	0.000	—	3	0.800	0.640	0.563	0.019*	1	0.000	0.000	0.000	—	5	1.000	0.753	0.712	0.275
cons_gr2082	4	1.000	0.719	0.667	0.353	2	0.000	0.320	0.269	0.025*	1	0.000	0.000	0.000	—	1	0.000	0.000	0.000	—	6	0.917	0.757	0.721	0.439
cons_gr2268	7	1.000	0.844	0.825	0.521	2	0.800	0.480	0.365	0.136	3	0.600	0.460	0.410	0.821	2	0.500	0.375	0.305	0.637	10	1.000	0.872	0.858	0.375
cons_gr4165	6	1.000	0.760	0.730	0.451	3	0.200	0.620	0.548	0.107	4	0.600	0.700	0.645	0.339	3	1.000	0.611	0.535	0.392	6	1.000	0.748	0.709	0.493
cons_gr240	4	0.750	0.563	0.524	0.963	2	0.800	0.480	0.365	0.136	2	0.800	0.480	0.365	0.136	3	0.667	0.500	0.449	0.861	11	0.833	0.837	0.823	0.590
cons_gr720	6	1.000	0.813	0.786	0.679	4	1.000	0.660	0.610	0.544	2	0.200	0.500	0.375	0.180	2	0.667	0.444	0.346	0.386	12	1.000	0.906	0.898	0.706
cons_gr1008	4	1.000	0.750	0.703	0.677	1	0.000	0.000	0.000	—	3	0.600	0.460	0.410	0.821	5	1.000	0.778	0.744	0.532	7	1.000	0.778	0.713	0.038*
cons_gr4117	4	1.000	0.700	0.645	0.188	2	0.800	0.480	0.365	0.136	5	1.000	0.740	0.701	0.891	6	1.000	0.833	0.810	0.451	9	1.000	0.788	0.765	0.343
cons_gr4727	4	0.750	0.719	0.667	0.530	1	0.000	0.000	0.000	—	3	0.200	0.340	0.314	0.019*	2	0.333	0.278	0.239	0.729	12	0.833	0.823	0.806	0.166
cons_gr4891	6	0.800	0.820	0.794	0.451	2	1.000	0.500	0.375	0.025*	3	0.400	0.340	0.314	0.958	2	0.333	0.500	0.375	0.564	12	1.000	0.885	0.875	0.570
cons_gr4988	4	0.600	0.640	0.581	0.847	3	0.600	0.460	0.410	0.821	2	0.800	0.480	0.365	0.136	4	1.000	0.722	0.671	0.609	5	0.667	0.764	0.725	0.057
cons_gr5464	6	1.000	0.813	0.786	0.526	2	0.400	0.480	0.365	0.709	3	0.600	0.620	0.548	0.079	3	0.667	0.500	0.449	0.861	10	0.833	0.840	0.823	0.831

A—number of alleles; H_o—observed heterozygosity; H_e—expected heterozygosity; PIC—polymorphism information content; *N*—sample size. *P* value from exact tests for Hardy–Weinberg equilibrium. * Show significant deviation from Hardy–Weinberg equilibrium, * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

COMPLIANCE WITH ETHICAL STANDARDS

The authors declare that they have no conflict of interest. This article does not contain any studies involving animals or human participants performed by any of the authors.

REFERENCES

1. Kvaček, Z., Manchester, S.R., Zetter, R., and Pingen, M., Fruits and seeds of *Craigia bronnii* (Malvaceae—Tilioideae) and associated flower buds from the late Miocene Inden formation, lower Rhine basin, Germany, *Rev. Palaeobot. Palynol.*, 2002, vol. 119, nos. 3–4, pp. 311–324.
2. Jin, J.H., Kodrul, T.M., Liao, W.B., and Wang, X., A new species of *Craigia* from the Eocene Changchang formation of Hainan Island, China, *Rev. Palaeobot. Palynol.*, 2009, vol. 155, nos. 1–2, pp. 80–82.
3. Qin, H.N., Yang, Y., and Dong, S.Y., et al. Threatened species list of China's higher plants, *Biodiversity Sci.*, 2017, vol. 25, no. 7, pp. 696–744.
4. Gao, Z.R., Zhang, C.Q. and Milne, R.L., Size-class structure and variation in seed and seedling traits in relation to population size of an endangered species *Craigia yunnanensis* (Tiliaceae), *Aust. J. Bot.*, 2010, vol. 58, no. 3, pp. 214–223.
5. Yang, J., Gao, Z.R., Sun, W.B., and Zhang, C.Q., High regional genetic differentiation of an endangered relict plant *Craigia yunnanensis* and implications for its conservation, *Plant Diversity*, 2016, vol. 38, no. 5, pp. 221–226.
6. Ashley, M.V. and Dow, B.D., The use of microsatellite analysis in population biology: background methods and potential applications, *Exs*, 1994, vol. 69, no. 69, pp. 185–201.
7. Dawson, I.K., Waugh, R., Simons, A.J., and Powell, W., Simple sequence repeats provide a direct estimate of pollen-mediated gene dispersal in the tropical tree *Gliricidia sepium, Mol. Ecol.*, 2003, vol. 6, no. 2, pp. 179–183.
8. Yang, J., Zhao, L.L., Yang, J.B., and Sun, W.B., Genetic diversity and conservation evaluation of a critically endangered endemic maple, *Acer yangbiense*, analyzed using microsatellite markers, *Biochem. System. Ecol.*, 2015, vol. 60, pp. 193–198.
9. Liu, J. and Gao, L.M., Comparative analysis of three different methods of total DNA extraction used in *Taxus, Guihaia*, 2011, vol. 31, no. 2, pp. 244–238.
10. Meglécz, E., Costedoat, C., and Dubut, V., et al., QDD: a user-friendly program to select microsatellite markers and design primers from large sequencing projects, *Bioinformatics*, 2010, vol. 26, no. 3, pp. 403–404.
11. PeaKall, R. and Smouse, P.E., GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research: an update, *Bioinformatics*, 2012, vol. 28, pp. 2537–2539.
12. Kalinowski, S.T., Taper, M.L., and Marshall, T.C., Revising how the computer program Cervus, accommodates genotyping error increases success in paternity assignment, *Mol. Ecol.*, 2010, vol. 16, no. 5, pp. 1099–1106.