GMR

Brief Note

New set of microsatellites for Chinese tallow tree, *Triadica sebifera*

Y.F. Zhuang[1], Z.F. Wang[2] and L.F. Wu[1]

[1]La Jolla Country Day School, La Jolla, CA, USA
[2]Guangdong Provincial Key Laboratory of Applied Botany,
South China Botanical Garden, Chinese Academy of Sciences, Guangzhou, China

Corresponding author: Z.F. Wang
E-mail: wzf@scib.ac.cn

Genet. Mol. Res. 16 (2): gmr16029624
Received January 23, 2017
Accepted March 29, 2017
Published April 5, 2017
DOI http://dx.doi.org/10.4238/gmr16029624

ABSTRACT. Chinese tallow (*Triadica sebifera*) is an important crop and ornamental tree. After it was introduced into the USA, it gradually became a noxious invasive tree in south-eastern America since the middle of the 1900s. Because only six microsatellites were reported previously in *T. sebifera*, to better understand the genetic diversity and population dynamics of such species, we reported here 28 new microsatellite markers. For these 28 microsatellites, the number of alleles per locus ranged from 2-16. The expected heterozygosity and the expected heterozygosity corrected for sample size varied from 0.0796 to 0.9081 and from 0.0805 to 0.9176, respectively. These microsatellites will provide additional choice to investigate the genetic diversity and structure in *T. sebifera*.

Key words: Genetic diversity; Microsatellite marker development; Invasive species; Restriction site-associated DNA sequencing; *Triadica sebifera*

INTRODUCTION

Triadica sebifera belongs to the family of Euphorbiaceae. Because its taxonomy is in question, currently it has synonyms with the scientific name *Sapium sebiferum* (DeWalt et al., 2006). *T. sebifera* is a small- to medium-sized deciduous tree naturally distributed in China, Japan, and Vietnam. Being an important economic tree well known as Chinese tallow tree, its seeds can be used to produce vegetable oil (Yu, 1951). Therefore, it is also widely cultivated in Africa, America, Europe, and India.

T. sebifera is a fast-growing species. It grows on soils with humid conditions and tolerates flooding (Camarillo et al., 2015). However, because of its strong, deep taproot, it can withstand periods of drought (Scheld and Cowles, 1981). *T. sebifera* has a beautiful rounded or conical crown and colorful leaves in autumn. It is also remarkably free of or tolerant to insect pests and diseases (Wang et al., 2011). All these characteristics make *T. sebifera* a highly ornamental tree. The tree is monoecious with male and female flowers on the same plant, and high-quality honey can be made from the flowers. Besides its industrial applications, the seed oil can be used as purgative and emetic in China (Duke and Ayensu, 1985). The leave extract of *T. sebifera* has the effect of analgesia and anti-inflammation (Huang et al., 2004), while its roots are used for the treatment of snakebites and dyspepsia (Duke and Ayensu, 1985). The wood of *T. sebifera* is white and close-grained, suitable for carving and furniture (Xu, 2010).

Unlike the other cultivated plants, *T. sebifera* was introduced into America as an ornamental and crop tree (making soap from the seed oil) during the 1700s but gradually became an invasive species throughout south-eastern America since the middle of the 1900s. Toxins contained in its fallen leaves turn habitats unfavorable to the other native species (Cameron and Spencer, 1989; Saenz et al., 2013). These substances can change chemical properties of soil and litter negatively impact native vegetation. It then crowds out the native species once it establishes, forming monospecific forests and decreasing species diversity (Saenz et al., 2013). It can produce viable seeds as little as 3 years after germination. Seeds are eaten and spread by birds, which makes it uneasy to control (Aslan, 2011). Moving water also serves as an important mechanism for seed dispersal. Therefore, flooding is an important factor in accelerating its expansion (Henkel et al., 2016). Besides, it can also regrow from cut stumps and roots. So, once established, *T. sebifera* is virtually impossible to eliminate (Wheeler and Ding, 2014).

To better understand the genetic diversity and population dynamics of a species, microsatellite markers are used. Previously, however, only six microsatellites were reported in *T. sebifera* (DeWalt et al., 2006). Therefore, here we develop and characterize new microsatellite markers that will help manage this species within its native and invasive areas.

MATERIAL AND METHODS

We used the restriction site-associated DNA sequencing (RAD-seq) method to obtain whole genome DNA sequences. Based on these sequences, microsatellites were then identified and characterized.

We employed two individuals to construct RAD-seq libraries. Both individuals were from the Guangdong Province, China, one being from Lianzhou and the other from the South China Botanical Garden. Following the methods of Baird et al. (2008), we prepared RAD-seq libraries with the restriction enzyme *Eco*RI and 150-bp paired-end sequencing with

Illumina HiSeq 2500 genetic analyzer. For each individual, we obtained a total of 64,140,980 and 45,940,432 raw reads, respectively. By filtering PCR duplicates and low-quality reads, we finally got 45,793,922 and 29,735,782 clean reads, respectively. We assembled these sequences separately for each individual using the Rainbow 2.0.4 software (Chong et al., 2012). Then, we used Msatcommander 0.8.2 (Faircloth, 2008) to screen microsatellites from these assembled sequences. Particularly, we only chose sequences with their dinucleotide and trinucleotide motifs having at least nine and eight repeats. We pooled these screened sequences from each individual and further assembled them using CAP3 (Huang and Madan, 1999). We then randomly chose 57 microsatellite sequences that appeared in both individuals. We performed PCRs to test the availability of these microsatellites and PCR procedures were the same to Li et al. (2016). After checking PCR products on 2% agarose gel, we found that 51microsatellites could be successfully amplified with the right size. Because agarose gel could not separate polymorphic alleles in microsatellites, initially we used six individuals from Lianzhou to perform PCR and identify their polymorphism. After PCR amplification and running PCR products on ABI 3730 sequencer (Applied Biosystems, Carlsbad, CA, USA), we determined the microsatellite allele size with both ABI GeneMapper 4.1 and Gelquest 3.2.1.0 softwares (http://www.sequentix.de) to make sure that the results were determined consistently. Finally, we retained a total of 28 microsatellites that gave clear polymorphism. For these microsatellites, additional 18 individuals from Lianzhou were used to identify their full polymorphism. Overall, we used 24 samples to validate a total of 28 microsatellite loci.

Because *T. sebifera* is tetraploid, for each locus, we estimated the number of alleles manually and used ATETRA 1.1 (Van Puyvelde et al., 2010) to calculate the expected heterozygosity (H_E) and expected heterozygosity corrected for sample size (H_{EC}).

RESULTS AND DISCUSSION

For 28 microsatellite locus, we detected the highest number of alleles in locus WJ-55 with sixteen and the lowest in WJ-9 and WJ-45 with two each (Table 1). The H_E and H_{EC} varied from 0.0796 to 0.9081 and from 0.0805 to 0.9176, respectively.

Microsatellites are popular markers used at a wide range of genetic studies. Its polymorphism is originated from short repeats in the microsatellite sequence. These short repeats have in general 1-7 bp. For example, if one allele in a microsatellite locus contains ten 2-bp repeats, while the other allele contains eight 2-bp repeats, these two alleles will have a difference of 4 bp in length. We can then use gel electrophoresis methods to separate them and find different sizes (i.e., alleles). Although agarose gel is the easy way to separate different sizes of alleles in microsatellites, it has low resolutions. Therefore, using fluorescent primers with automatic capillary is one of the most popular ways for its high-throughput polymorphism assessment. Here we used ABI 3730 sequencer for the allele separation. Because different softwares use different algorithms to call allele sizes, which slightly resulted in different results (Covarrubias-Pazaran et al., 2016), we here used two softwares (GeneMapper and Gelquest) to compare the allele sizes, and we found that they produced identical results.

Nowadays, using RAD-seq, a next-generation sequencing (NGS) approach, to develop microsatellites has become a popular method (Zalapa et al., 2012). RAD-seq can identify a large number of microsatellite loci in both time and cost-efficient processes. As we have seen, only six microsatellite loci have been developed in *T. sebifera* using the traditional DNA-enrichment method (DeWalt et al., 2006). Due to thousands of researchers using RAD-seq to

Table 1. Characteristics of 28 microsatellite markers developed in *Triadica sebifera*.

Locus	Repeat motifs of individual applied to RAD sequencing	Primer sequences (5'-3')	Size range (bp)	N_A	H_E	H_{EC}	GenBank accession No.
WJ-6	$(GT)_9$	F: FAM-TGCCCTTCATGTGGAGCTG	244–254	3	0.4703	0.4753	KY428821
		R: TTCCACCTCAAAATCCCGC					
WJ-8	$(AG)_{11}$	F: FAM-GAGCTCTGTCCTGGTTGG	158-166	5	0.7796	0.7878	KY428822
		R:ACCCAAGAATTCACGAATAGC					
WJ-9	$(AG)_{10}$	F: FAM-AACCCGTAAAGGGCTTGC	190-192	2	0.0796	0.0805	KY428823
		R: CTGGTTCTCCTGGTTATCTATGC					
WJ-10	$(ATT)_{10}$	F: FAM-AAGGAATGGAGCGAAACGG	222-258	6	0.8796	0.8889	KY428824
		R:CCAATTGCGGCCATACTCG					
WJ-13	$(CTT)_{10}$	F: FAM-TCCGATCCAGTCCGTGTTG	176-197	5	0.7391	0.7469	KY428825
		R:GTGCGCGTGAGAGTGAATG					
WJ-15	$(CT)_9$	F: FAM-TGCTTGGGCTCTCTCTTACC	194-206	6	0.5924	0.5987	KY428826
		R: ACCAAAAGATGCATAGATAATGCC					
WJ-17	$(CTT)_9$	F: FAM-TCTCTCCTTCGCTCAACGG	147-159	5	0.6000	0.6063	KY428827
		R:TCCGGGATCGGTGGAATTG					
WJ-20	$(CT)_9$	F: FAM-GTTTGTGAAGAGGGGTGAGC	176-196	6	0.6797	0.6868	KY428828
		R:AGTTGCTGAAATCCATACCATACC					
WJ-24	$(ATT)_9$	F: FAM-TGAACCTCGAACAAAAGTCAG	119-173	11	0.6874	0.6946	KY428829
		R: GTCAT(C/A)ATAACTTCGCGGG					
WJ-25	$(AT)_9$	F: FAM-TAGCCATGTCCAAGCTCCG	124-136	6	0.7348	0.7425	KY428830
		R:CCCATAGTTAGTTGATTTTGTGGC					
WJ-26	$(TAA)_7$	F: FAM-GTCAGCAGGGGAGAGCAAC	158-167	5	0.5055	0.5109	KY428831
		R:AATGGACAAAATGGCGCAC					
WJ-27	$(CT)_{10}$	F: FAM-TGTATGGGAAGCGGCTCTG	260-268	5	0.6184	0.6249	KY428832
		R:ACTCCTGGGCTCCAACTTC					
WJ-30	$(CT)_{12}$	F: FAM-TAGCCAGGCCTATGCAATC	183-193	4	0.7244	0.7321	KY428833
		R:AAGTTGGTACCCTGGTCTG					
WJ-31	$(GA)_9$	F:FAM-TGCAAGATTGAAATCAGATAAACCC	323-327	3	0.6245	0.6311	KY428834
		R:AGAGCTTGGAAGGAGGCTG					
WJ-32	$(CT)_{10}$	F:FAM-AGATCTTGTTAATAGGCGACTGC	172-184	4	0.4333	0.4379	KY428835
		R:CTTTCCACTGCGCCTGAAG					
WJ-35	$(AG)_{16}$	F: FAM-AAGGAACCTGTTTGCTGGG	194-208	5	0.5043	0.5096	KY428836
		R:AAGTTCCGTTTCCACACGC					
WJ-37	$(CT)_9GCC(TG)_6$	F: FAM-GTCAGTCGTCACCATCATCAG	184-208	5	0.7010	0.7084	KY428837
		R:CTACGACGACGCAACCAAC					
WJ-39	$(AAT)_8$	F: FAM-ACTCTGCACTGGAAACAGAATG	235-241	3	0.6150	0.6215	KY428838
		R:AGTGGAAGCCTAAGGTGGG					
WJ-40	$(AG)_9$	F: FAM-GGAAACGGCGCCTATGATG	147-158	3	0.6649	0.6719	KY428839
		R:CCGGATGCACAAATGCTTTC					
WJ-41	$(AG)_{18}$	F:FAM-ACAAAGAAGTGAAAGAAAGAACCC	101-115	6	0.7412	0.7490	KY428840
		R:CAATGAATTTGGATTGGTGGC					
WJ-42	$(ATT)_{11}$	F: FAM-TCTTCGGGGAAACCGATCC	143-161	5	0.7849	0.7931	KY428841
		R:TGCTTTCAAAATGACACGGTTG					
WJ-45	$(AT)_9$	F: FAM-TTCATGTCTAAATTTCTCGGGC	188-190	2	0.1522	0.1538	KY428842
		R:TTGACCTCGCCATGACAAC					
WJ-46	$(GA)_{11}$	F: FAM-CAACAAAGGGACGAAGGGG	194-215	5	0.5572	0.5631	KY428843
		R:ACTGCCAGACTGCTCTCAC					
WJ-50	$(ATT)_8$	F: FAM-TGTGCTAGGAAGTAAAGGACAAG	167-191	9	0.8586	0.8676	KY428844

Annealing temperature was 55°C for all loci. N_A: number of alleles; H_E: expected heterozygosity; H_{EC}: expected heterozygosity corrected for sample size; F: fixation index.

develop microsatellites in both plants and animals recently, we here do not intend to summarize them all. Instead, compared to our previous study (Li et al., 2016) using the same procedures in *Bretschneidera sinensis*, a tertiary relict and endangered species in China, we identified 63 polymorphic loci from 170 candidate loci (37.1%) and here 28 usable microsatellite loci from 51 candidate loci (54.9%) in *T. sebifera* were obtained. It clearly demonstrated the high efficiency in microsatellite locus development in the present study.

Although the microsatellite loci developed by DeWalt et al. (2006) detected a higher number of alleles at each locus with 8-18 compared to ours with 2-16, they mixed samples from different resources (79 individuals collected in the USA, Australia, and China) while we only used 24 individuals from one population. Because sample sizes and origination contributed significantly to the allelic richness in the microsatellite loci (Greenbaum et al., 2015), relatively low number of alleles in the present study is reasonable. Actually, in DeWalt et al. (2011)'s later study, their six microsatellite loci could only detect 2.27, 2.21, and 2.39 a mean number of alleles per locus each in China, USA, and Australia, while our results detected 5.29, which is much higher than theirs. Higher alleles in microsatellite loci increase the resolutions of individual identification and population genetic analyses (Collevatti et al., 2001). However, because microsatellite markers are assumed to be neutral, they are very

suitable for much evolutionary (such as demographic histories) but not including adaptive inferences in species (Kirk and Freeland, 2011). SNP (single nucleotide polymorphisms) markers, identified in genes, could improve our understanding of such adaptive evolutionary potential (Kirk and Freeland, 2011). RAD-seq used here is also an efficient way to discover SNPs (Baird et al., 2008), but to validate SNPs is still costly.

Overall, these microsatellites will provide additional choice to investigate the genetic diversity and structure in *T. sebifera*, and especially to compare them in its native and introduced ranges. Also, natural hybridizations have been recently reported between *T. sebifera* and *T. cochinchinensis*, and between *T. sebifera* and *T. rotundifolia* (Wang et al., 2016). Microsatellites developed here will help to better observe such phenomenon for future plant breeding and management in this species.

Conflicts of interest

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

We thank G. Yun for assistance in sampling and laboratory.

REFERENCES

Aslan CE (2011). Implications of newly-formed seed-dispersal mutualisms between birds and introduced plants in northern California, USA. *Biol. Invasions* 13: 2829-2845. http://dx.doi.org/10.1007/s10530-011-9966-1

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

Camarillo SA, Stovall JP and Sunda CJ (2015). The impact of Chinese tallow (*Triadica sebifera*) on stand dynamics in bottomland hardwood forests. *For. Ecol. Manage.* 344: 10-19. http://dx.doi.org/10.1016/j.foreco.2015.02.013

Cameron GN and Spencer SR (1989). Rapid leaf decay and nutrient release in a Chinese tallow forest. *Oecologia* 80: 222-228. http://dx.doi.org/10.1007/BF00380155

Chong Z, Ruan J and Wu CI (2012). Rainbow: an integrated tool for efficient clustering and assembling RAD-seq reads. *Bioinformatics* 28: 2732-2737. http://dx.doi.org/10.1093/bioinformatics/bts482

Collevatti RG, Grattapaglia D and Hay JD (2001). High resolution microsatellite based analysis of the mating system allows the detection of significant biparental inbreeding in *Caryocar brasiliense*, an endangered tropical tree species. *Heredity (Edinb)* 86: 60-67. http://dx.doi.org/10.1046/j.1365-2540.2001.00801.x

Covarrubias-Pazaran G, Diaz-Garcia L, Schlautman B, Salazar W, et al. (2016). Fragman: an R package for fragment analysis. *BMC Genet.* 17: 62. http://dx.doi.org/10.1186/s12863-016-0365-6

DeWalt SJ, Siemann E and Rogers WE (2006). Microsatellite markers for an invasive tetraploid tree, Chinese tallow (*Triadica sebifera*). *Mol. Ecol. Notes* 6: 505-507. http://dx.doi.org/10.1111/j.1471-8286.2006.01296.x

DeWalt SJ, Siemann E and Rogers WE (2011). Geographic distribution of genetic variation among native and introduced populations of Chinese tallow tree, *Triadica sebifera* (Euphorbiaceae). *Am. J. Bot.* 98: 1128-1138. http://dx.doi.org/10.3732/ajb.1000297

Duke JA and Ayensu ES (1985). Medicinal Plants of China. Reference Publications, Inc. Michigan, USA.

Faircloth BC (2008). msatcommander: detection of microsatellite repeat arrays and automated, locus-specific primer design. *Mol. Ecol. Resour.* 8: 92-94. http://dx.doi.org/10.1111/j.1471-8286.2007.01884.x

Greenbaum G, Templeton AR, Zarmi Y and Bar-David S; PLOS ONE Staff (2015). Correction: Allelic richness following population founding events--a stochastic modeling framework incorporating gene flow and genetic drift. *PLoS One* 10: e0119663. http://dx.doi.org/10.1371/journal.pone.0119663

Henkel TK, Chambers JQ and Baker DA (2016). Delayed tree mortality and Chinese tallow (*Triadica sebifera*) population explosion in a Louisiana bottomland hardwood forest following Hurricane Katrina. *For. Ecol. Manage.* 378: 222-232. http://dx.doi.org/10.1016/j.foreco.2016.07.036

Huang BX, Huang ZQ, Xu XL, Wei JQ, et al. (2004). Analgesic and nti-inflammatory effect of extract of *Sapium sebiferum* leaves on animal model. *Zhongchengyao* 26: 476-479.

Huang X and Madan A (1999). CAP3: A DNA sequence assembly program. *Genome Res.* 9: 868-877. http://dx.doi.org/10.1101/gr.9.9.868

Kirk H and Freeland JR (2011). Applications and implications of neutral versus non-neutral markers in molecular ecology. *Int. J. Mol. Sci.* 12: 3966-3988. http://dx.doi.org/10.3390/ijms12063966

Li M, Chen HF, Wang ZF and Zhang S (2016). Isolation and characterization of polymorphic microsatellite markers in the endangered species *Bretschneidera sinensis* Hemsl. *Genet. Mol. Res*. 15: gmr.15038234.

Saenz D, Fucik EM and Kwiatkowski MA (2013). Synergistic effects of the invasive Chinese tallow (*Triadica sebifera*) and climate change on aquatic amphibian survival. *Ecol. Evol.* 3: 4828-4840. http://dx.doi.org/10.1002/ece3.857

Scheld HW and Cowles JR (1981). Woody biomass potential of the Chinese tallow tree. *Econ. Bot.* 35: 391-397. http://dx.doi.org/10.1007/BF02858589

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

Wang S, Chen Y, Yang Y, Wu W, et al. (2016). Phylogenetic relationships and natural hybridization in *Triadica* inferred from nuclear and chloroplast DNA analyses. *Biochem. Syst. Ecol.* 64: 142-148. http://dx.doi.org/10.1016/j.bse.2015.11.011

Wang Y, Huang W, Siemann E, Zou J, et al. (2011). Lower resistance and higher tolerance of invasive host plants: biocontrol agents reach high densities but exert weak control. *Ecol. Appl.* 21: 729-738. http://dx.doi.org/10.1890/09-2406.1

Wheeler GS and Ding J (2014). Is Chinese Tallowtree, *Triadica sebifera*, an appropriate target for biological control in the United States? *Invasive Plant Sci. Manag.* 7: 345-359. http://dx.doi.org/10.1614/IPSM-D-13-00061.1

Xu JS (2010). Cultivation and management of Chinese tallow tree. *Anhui Forest. Sci. Tech.* 1: 45.

Yu X (1951). Refining and utilization of tallow oil in *Sapium sebiferum. Chemical World* 3: 2-4.

Zalapa JE, Cuevas H, Zhu H, Steffan S, et al. (2012). Using next-generation sequencing approaches to isolate simple sequence repeat (SSR) loci in the plant sciences. *Am. J. Bot.* 99: 193-208. http://dx.doi.org/10.3732/ajb.1100394