Torreya yunnanensis C. Y. Cheng & L. K. Fu (Taxaceae) is a Chinese endemic evergreen tree that is highly valued for its timber, as well as for its ornamental and ecological benefits. Wild resources of T. yunnanensis have declined sharply as a result of overexploitation and deforestation. It was listed as endangered in the China Species Red List (Wang and Xie, 2004) and is now rare in the field. Very little research has been done on this species, except for recent studies on chemical composition, cultivation techniques, and community structure (e.g., Hou et al., 2015; Fu et al., 2016; Li et al., 2016). Although conservation of this species is urgent, little is known about its genetic diversity, which is important for planning conservation strategy (Avise and Hamrick, 1996; Jiang et al., 1997).

Microsatellites, also called simple sequence repeat (SSR) markers, are a neutral molecular marker widely distributed in the nuclear genome of eukaryotes. Because of advantages such as high polymorphism, ability to facilitate genotyping, and low demand of DNA quality, microsatellites have been widely used as DNA markers in population genetic studies and parentage analysis. In this study, 16 polymorphic microsatellite markers were developed using Illumina 2 × 100-bp paired-end sequencing and bioinformatics screening. We believe these markers are useful for investigating genetic diversity and population structure for T. yunnanensis and other Torreya species.

**METHODS AND RESULTS**

A total of 64 T. yunnanensis individuals were sampled from five wild populations in this study (Appendix 1). The sample sizes were small for each population (ranged from 7–18) because T. yunnanensis is an endangered species and is rare in the field. Young leaves were collected and dry preserved with silica gel. Genomic DNA was extracted using the cetyltrimethylammonium bromide (CTAB) method (Doyle, 1987). A normalized DNA library was constructed using the TruSeq Stranded DNA Sample Preparation kit (Illumina Inc., San Diego, California, USA) for one sample from the Misha population. The normalized DNA library was sequenced using the Illumina HiSeq 2500 (Illumina Inc.). A total of 5,362,980,012 bp paired-end raw sequences were trimmed to remove adapter sequences and low-quality sequences using Trim Galore (http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/). De novo assembly based on trim reads was performed with Velvet 1.2.10 (Zerbino and Birney, 2008). A total of 50,381,156 trim reads with an average length of 98.7 bp (4,975,006,766 bp in total) were produced and de novo assembled into 283,057 contigs with an N50 length of 1823 bp. SSR detection was performed with MISA (Thiel et al., 2003) using the following criteria: ≥6 repeat units for dinucleotides, ≥5 for tri-, and ≥4 for tetra-, penta-, and hexanucleotides. A total of 4146 SSRs were identified. Of them, trinucleotide repeat motifs...
(50.19%) were the most common, followed by di- (44.19%), tetra- (3.14%), penta- (0.58%), and hexanucleotide (1.91%) repeat motifs. Primer pairs were designed using Primer3 (Koressaar and Remm, 2007; Untergasser et al., 2012) with the default conditions.

Fifty primer pairs amplifying SSRs containing dinucleotide or trinucleotide motifs were randomly selected and tested in four individuals. Sixteen primer sets displaying consistent amplification (Table 1) were selected for further polymorphism tests on 64 individuals. All PCRs were performed in 20-μL volume (50–100 ng of genomic DNA, 10 μL of 2× EasyTaq PCR SuperMix polymerase [TransGen Biotech Co., Beijing, China], 0.5 μM of each primer pair [the forward primer was fluorescently labeled with FAM, HEX, or TAMRA]) under the following conditions: 5 min denaturation at 95°C; 32 cycles of 30 s at 95°C, 1 min at specific annealing temperatures, and 1 min at 72°C (Table 1); and a final extension of 72°C for 10 min. PCR products were separated on an ABI PRISM 3730 Genetic Analyzer (Thermo Fisher Scientific, Waltham, Massachusetts, USA) with a GeneScan 500 LIZ Size Standard. Allele peaks were scored using GeneMarker (version 1.3; SoftGenetics, State College, Pennsylvania, USA) with a GeneScan 500 LIZ Standard (Applied Biosystems, Foster City, California, USA).

Hardy–Weinberg equilibrium was tested with GENEPOP version 3.4 (Rousset, 2008), and the P values were then tested using Bonferroni correction. The exact tests for genotypic linkage disequilibrium for each pair of loci in each population were performed with GENEPOP version 3.4. According to MICRO-CHECKER analysis, there is no evidence for scoring error due to stuttering, large allele dropout, or the existence of null alleles. Fourteen loci were polymorphic in five wild populations (Table 2) and two loci (Ty49984 and Ty61395) were monomorphic. Values for A, H, and F of 14 loci ranged from 2–12, 0.000–1.000, and 0.000–0.869 (Table 2), respectively, in five populations. The H and F values for all five populations, calculated from data from the 14 loci, ranged from 0.664–0.728 and 0.514–0.633 (data not shown); these values were higher than those found in T. jackii (H = 0.5012, F = 0.4830; Li, 2013), Taxus chinensis (Pilg.) Rehder (H = 0.107, F = 0.121; Vu et al., 2017), and Taxus cuspidata Siebold & Zucc. (H = 0.263, F = 0.028; Cheng et al., 2015). Significant deviation from Hardy–Weinberg equilibrium was found in loci Ty13986, Ty53899, Ty60665, and Ty72305 (P < 0.05) after Bonferroni correction. Most of the F values were negative, indicating low levels of inbreeding in these populations. There was no evidence of linkage disequilibrium among pairs of loci in the sample. The sequences of these microsatellite loci have been deposited in the GenBank database (Table 1), and the raw sequences from high-throughput sequencing have been deposited in the Sequence

### Table 1. Characterization of 16 *Torreya yunnanensis* microsatellite loci.

| Locus  | Primer sequences (5′–3′) | Repeat motif | Allele size range (bp) | A   | T<sub>a</sub> (°C) | GenBank accession no. |
|--------|--------------------------|--------------|-----------------------|-----|-------------------|----------------------|
| Ty12662| F: CAGGCCACATTAAATCCTGTA R: CTTCAACCCCGACTTGCTTA | (TA)<sub>a</sub> | 161–201               | 12  | 48                | KX359344             |
| Ty13451| F: GCTTGCCTGGAAACTTTACC R: GCATGCCAGGTCTCAGTTA | (GA)<sub>a</sub> | 203–221               | 6   | 56                | KX359348             |
| Ty13986| F: TGTCAGACCTAAACAGAGA R: TCATCTGCTTGTGCTCCTTT | (AG)<sub>a</sub> | 177–245               | 4   | 56                | KX359346             |
| Ty23232| F: CTTGACAGGACATTACAGG A: TGGGGGATCCTCGCTAT | (AT)<sub>a</sub> | 143–173               | 9   | 52                | KX359341             |
| Ty39589| F: GCATCGAGATGTTATCTTC | (TTC)<sub>a</sub> | 173–182               | 4   | 52                | KX359339             |
| Ty41047| F: TCCGTGAGAGCCTCTGCTG R: CAAAGGCTTCTGCAACAA | (AT)<sub>a</sub> | 203–251               | 4   | 51                | KX359349             |
| Ty49984| F: AGAGGGTCTTGAGGACT | (TA)<sub>a</sub> | 179                   | 1   | 55                | KX359342             |
| Ty53899| F: CATGCCCAGGCTCTCAATT R: CATGTTGAAGCACCCCTTGG | (GAA)<sub>a</sub> | 173–215               | 5   | 48                | KX359340             |
| Ty60665| F: TGGCTGCTCAATAGCAACA R: TGCAGGGCATATGCAAAA | (AT)<sub>a</sub> | 179–249               | 10  | 52                | KX359343             |
| Ty61395| F: TGTTGAAGGTTGGAACAA | (GA)<sub>a</sub> | 189                   | 1   | 53                | KX359345             |
| Ty72035| F: CAGATGCAAGAAGAAGACCA R: ATGGGCAAATATGGCCCTTTT | (AT)<sub>a</sub> | 125–173               | 5   | 50                | KX359336             |
| Ty74507| F: GCAGATTGAGAGCACTTTT R: AAGCGGGTTTTGCTCATTTC | (AT)<sub>a</sub> | 107–167               | 6   | 49                | KX359347             |
| Ty78313| F: AGCTTAGACCAATACACAGAA R: GGTTGCTTAAACCTTCAGGG | (ATA)<sub>a</sub> | 187–193               | 2   | 52                | KX359337             |
| Ty79607| F: TTTGAGATCACTGGGGAGAG R: GGGTTGTGACGCTTGGGAT | (GA)<sub>a</sub> | 215–307               | 18  | 52                | KX359351             |
| Ty80072| F: AAAGCTACAGCTGGTTGTCA R: TGCTAAAGCTCAACGCCATA | (AT)<sub>a</sub> | 145–197               | 10  | 51                | KX359338             |

Note: A = number of alleles; T<sub>a</sub> = annealing temperature.
Read Archive (SRA) of the National Center for Biotechnology Information (NCBI; accession no. SRR6428886).

Cross-amplification of these 16 primer sets was tested in five individuals in each of the following species: *T. fargesii* Franch., *T. grandis* Fortune ex Lindl., *T. jackii*, and *T. nucifera* (L.) Siebold & Zucc. Most primers also amplified in these species (Table 3).

### CONCLUSIONS

Genetic diversity in wild *T. yunnanensis* populations was shown to be high. The SSR markers that we developed were polymorphic and will be a useful tool to investigate the genetic diversity, population structure, and levels of gene flow, as well as to optimize breeding for *T. yunnanensis* and related species.

### ACKNOWLEDGMENTS

The authors thank Xin Hong (Anhui Normal University, Wuhu, Anhui Province, China) for sampling. The National Natural Science Foundation of China (no. 31400321) and the Natural Science Foundation of Anhui Province (11040606M77) provided financial support.

### LITERATURE CITED

Avise, J. C., and J. L. Hamrick. 1996. Conservation genetics: Case histories from nature. Chapman and Hall Press, New York, New York, USA.

Belkhir, K., P. Borsa, and L. Chikhi. 2001. GENETIX 4.02, logiciel sous Windows™ pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS-UPR, Université de Montpellier II, Montpellier, France.

Cheng, B. B., Y. Q. Zheng, and Q. W. Sun. 2015. Genetic diversity and population structure of *Taxus cuspidata* in the Changbai Mountains assessed by chloroplast DNA sequences and microsatellite markers. Biochemical Systematics and Ecology 63: 157–164.

Doyle, J. J. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* 19: 11–15.

Fu, R., H. Tao, and Y. Li. 2016. Analysis of growth behavior and reproductive technology of reya tree in *Torreya yunnanensis*. *Modern Horticulture (Chinese)* 10: 51.

Hou, Z., G. Wen, D. Zhou, M. Li, and F. Du. 2015. Floristic characteristics and flora of *Torreya yunnanensis* community. *Journal of West China Forestry Science (Chinese)* 44(3): 37–44.

Jiang, Z. G., K. P. Ma, and X. G. Han. 1997. Conservation biology. Science and Technology Press, Hangzhou, Zhejiang, China.

Koressaar, T., and M. Remm. 2007. Enhancements and modifications of primer design program Primer3. *Bioinformatics* 23: 1289–1291.

Li, H., F. Li, and S. Ma. 2016. Study on the molecular docking of PDE9A homologous 3JSW and Torreyunlignans derivatives. *Journal of Pharmaceutical Research (Chinese)* 35(9): 511–513.

Li, J. 2013. Multi-scale spatial genetic structure and population demographic history of *Torreya jackii* (Taxaceae), an endemic and endangered plant in China. PhD dissertation, East China Normal University, Shanghai, China.

Rousset, F. 2008. GENEPOP’007: A complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources* 8: 103–106.
Thiel, T., W. Michalek, R. K. Varshney, and A. Graner. 2003. Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (Hordeum vulgare L.). *Theoretical and Applied Genetics* 106: 411–422.

Untergasser, A., I. Cutcutache, T. Koressaar, J. Ye, B. C. Faircloth, M. Remm, and S. G. Rozen. 2012. Primer3: New capabilities and interfaces. *Nucleic Acids Research* 40: e115.

van Oosterhout, C., W. F. Hutchinson, D. P. M. Wills, and P. Shipley. 2004. MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* 4: 535–538.

Vu, D. D., T. T. X. Bui, M. T. Nguyen, D. G. Vu, M. D. Nguyen, X. Huang, and Y. Zhang. 2017. Genetic diversity in two threatened species in Vietnam: *Taxus chinensis* and *Taxus wallichiana*. *Journal of Forestry Research* 28(2): 265–272.

Wang, S., and Y. Xie. 2004. China Species Red List. Higher Education Press, Beijing, China.

Zerbino, D. R., and E. Birney. 2008. Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. *Genome Research* 18: 821–829.

### APPENDIX 1. Voucher information for *Torreya* species used in this study.

| Species                  | Voucher IDa | N   | Collection locality | Geographic coordinates |
|--------------------------|-------------|-----|---------------------|------------------------|
| *T. yunnanensis* C. Y. Cheng & L. K. Fu | Ty-001-XH   | 7   | Misha, Yunnan, China | 26°15′57″N, 99°38′08″E |
| *T. yunnanensis*         | Ty-002-XH   | 10  | Xiangtu, Yunnan, China | 26°15′22″N, 99°34′20″E |
| *T. yunnanensis*         | Ty-003-XH   | 12  | Baijixun, Yunnan, China | 27°23′50″N, 99°01′24″E |
| *T. yunnanensis*         | Ty-004-XH   | 17  | Baohe, Yunnan, China | 27°03′24″N, 99°20′37″E |
| *T. yunnanensis*         | Ty-005-XH   | 18  | Weideng, Yunnan, China | 27°07′50″N, 99°15′7″E |
| *T. fargesii* Franch.    | TF-001-XH   | 5   | Fangxian, Hubei, China | 31°49′57″N, 110°37′46″E |
| *T. grandis* Fortune ex Lindl. | Tg-001-XH   | 5   | Zhuji, Zhejiang, China | 29°42′28″N, 120°05′51″E |
| *T. jackii* Chun         | TJ-001-XH   | 5   | Xianju, Zhejiang, China | 28°46′32″N, 120°47′12″E |
| *T. nucifera* (L.) Siebold & Zucc. | TN-001-XH  | 5   | Hangzhou, Zhejiang, China | 30°06′56″N, 120°00′02″E |

Note: N = sample size.

*Vouchers deposited at Anhui Normal University. The voucher ID for each sample is sufficient to identify the exact voucher of these samples.

*XH in the voucher ID identifies the collector, Xin Hong.*