Applications in Plant Sciences

ISOLATION AND CHARACTERIZATION OF 30 MICROSATELLITE LOCI FOR *CUNNINGHAMIA LANCEOLATA* (TAXODIACEAE)[1]

SI-SI WANG[2], YANG ZHANG[2], DE-CHEN LIU[2], XIAO-WEI SUN[2], RONG WANG[2], AND
YUAN-YUAN LI[2,3]

[2]School of Ecological and Environmental Sciences, Tiantong National Station of Forest Ecosystem, Shanghai Key Laboratory for Urban Ecology and Restoration, East China Normal University, Shanghai 200241, People's Republic of China

• *Premise of the study:* To quantify the population-level genetic characteristics of *Cunninghamia lanceolata* (Taxodiaceae), an important timber conifer, we developed 30 pairs of microsatellite primers based on the nuclear genome.
• *Methods and Results:* Using the streptavidin-biotin capture system, we developed 14 polymorphic and 16 monomorphic microsatellites. Polymorphisms were detected in 14 loci using 94 individual trees that were collected from three *C. lanceolata* populations in Hubei and Zhejiang provinces and in Chongqing Municipality, China. There were three to 30 alleles per locus, and the observed and expected heterozygosities ranged from 0.0313–0.8333 and from 0.0313–0.9246, respectively. Cross-species amplification showed that two to seven polymorphic loci were functional in three of the five related species that were collected.
• *Conclusions:* Our newly developed microsatellite primers provide neutral molecular markers that are beneficial to future studies of population genetics and germplasm conservation of *C. lanceolata*.

Key words: cross-amplification; *Cunninghamia lanceolata*; genetic diversity; microsatellite; Taxodiaceae.

Cunninghamia lanceolata (Lamb.) Hook. (Taxodiaceae), known as Chinese fir, is an evergreen, outcrossing, and long-lived conifer that is widely distributed in southern China and northern Vietnam. Because of its relatively low nutrient demands, fast rate of growth, and strong resistance to corrosion and insect attacks, *C. lanceolata* is an important timber source that has been cultivated for more than 2000 yr (Yeh et al., 1994). Its present plantations cover about 4 million hectares accounting for 20–25% of the total commercial production of timber in China (Bao and Jiang, 1998; Huang et al., 2005). Understanding its genetic background is therefore critical to selecting germplasm resources and managing forests.

Some codominant molecular markers have been reported for *C. lanceolata*, e.g., 10 pairs of polymorphic microsatellite primers based on the nuclear genome (Li et al., 2015), 28 polymorphic expressed sequence tag–simple sequence repeat (EST-SSR) markers (Wen et al., 2013), and 97 polymorphic SSR loci based on transcript data (Xu et al., 2016). Microsatellite loci located in noncoding regions are neutral and usually show higher mutation rates than those located in encoding regions of the genome (Charlesworth et al., 1994). Neutral markers can be used to study population genetic diversity that is not related to adaptive traits, and therefore can better reveal spatial genetic structure, gene flow, and historical events (e.g., bottlenecks and founder effects) that contribute to conservation of germplasms. However, detecting gene flow patterns, especially in fine-scale analyses such as parentage analyses, requires genetic resolution high enough to distinguish every individual and is thus dependent on a large number of neutral markers with high polymorphism. It is therefore necessary to develop more informative neutral molecular markers for *C. lanceolata*. Here, 14 polymorphic and 16 monomorphic microsatellite loci were isolated and characterized in the nuclear genome of *C. lanceolata* to facilitate future studies on population genetics and germplasm conservation.

METHODS AND RESULTS

Total genomic DNA was extracted from dried leaves of one *C. lanceolata* individual (located in Tiantong in Zhejiang Province, China [Appendix 1]) using the Plant Genomic DNA Extraction Kit (Tiangen, Beijing, China). Approximately 250 ng of DNA were digested with the restriction enzyme *Mse*I (New England Biolabs, Beverly, Massachusetts, USA), and fragments of 200–800 bp were fractionated. The fragments were linked with an *Mse*I-adapter pair (F: 5′-TACTCAGGACTCAT-3′; R: 5′-GACGATGAGTCCTGAG-3′). The diluted products were amplified by an *Mse*I-N primer (5′-GATGGTCCTGAGTAAN-3′) under the following conditions: an initial step at 95°C for 3 min, followed by 20 cycles of 30 s at 94°C, 1 min annealing at 53°C, and 1 min at 72°C. The products were hybridized with 5′-biotinylated probes (AG)$_{15}$ in a 250-μL reaction system at 48°C for 2 h. The hybridization products were adhered by streptavidin-coated magnetic beads (Promega Corporation, Madison, Wisconsin, USA). The washed and eluted DNA fragments were further amplified with the *Mse*I-N primer using the conditions given above for 30 cycles. The products were purified using a multifunctional DNA Extraction Kit (BioTeke, Beijing, China) and were ligated to a pMD19-T vector (TaKaRa Biotechnology Co., Dalian, China). Then, the products were transformed into *Escherichia coli*

TABLE 1. Characterization of 14 polymorphic and 16 monomorphic microsatellite loci developed in *Cunninghamia lanceolata*.[a]

Locus	Primer sequences (5′–3′)	Repeat motif	Allele size range (bp)	A	T_a (°C)	Fluorescent dye[b]	GenBank accession no.
CL22	F: TGGTAGTACTCGCAGGAAAT R: CAGAGAATGGACACAAACAG	$(GA)_5$	170	1	64	HEX	KY769227
CL34*	F: TGTCAAAGACTCCTGAGAAG R: GCAGCAATACACACACATAG	$(CT)_5$	176–190	5	61	6-FAM	KY769205
CL90*	F: GCAAAAGAGGTCCACAATAC R: ACTGCAAAGGAGATACGCTG	$(AG)_8$	214–222	5	65	TAMRA	KY769206
CL108*	F: GAGGTATAGAGTTTACGGTTTG R: TCCTATGTGTCTAGTCATGTCC	$(GA)_7$	108–124	6	64	HEX	KY769207
CL164	F: GCTGGATGAACCTGAGTTGATT R: GAAGTTTCTCTTTTGCTTTGGC	$(CT)_{12}$	124	1	63	6-FAM	KY769226
CL255	F: CAACAGAGATATTGCAGCA R: TTCACCTGTAAGGGTTTTC	$(AG)_{21}$	118	1	63	ROX	KY769229
CL256	F: GGTTGCTCAGGTTTGGA R: TGTCCCTCTACCTCTCA	$(AG)_{15}$	107	1	64	ROX	KY769230
CL278	F: GAGACTAAGCCTGGGAGAG R: GAGAGGGGATCGAAAGAGA	$(TC)_5$	114	1	65	ROX	KY769231
CL287	F: GCAGCAATACACATACACACAGAAT R: ACAAGTGTCAAAGACTCCTAAGAAG	$(AG)_9$	169	1	65	6-FAM	KY769220
CL288	F: GCAGCAATACACAACATAAG R: TACCGAATTGTCTCCAGATA	$(AG)_9$	176	1	62	6-FAM	KY769221
CL291	F: AATGGAGATATTGCAGC R: AATACATTGTGCGGGTT	$(AG)_{24}$	153	1	61	ROX	KY769232
CL295*	F: ATCATCAACACAATTAGCACCTGGAG R: GCAGATCATCGAACGTGAGTTAGCTT	$(CT)_8$	157–163	4	65	ROX	KY769208
CL343*	F: TGCAAGTGTGAATAGAACC R: GTAAAAGGGGAAAGGGAGT	$(GA)_{21}$	197–311	30	61	HEX	KY769209
CL346	F: ACACACAGAATGTAGGCAG R: ATTGTGCAGGTTTTGGAGT	$(AG)_{17}$	113	1	65	ROX	KY769233
CL389*	F: GGAGATTGTAAAATGGACTCTAGCC R: CTTTTTCTTGTTCTTTCGGAGGAGC	$(AG)_{15}$	147–171	10	63	6-FAM	KY769210
CL396	F: CCATGTGCCTCACTCTC R: GGTTAGGGGTTCAGGTT	$(CT)_5$	111	1	65	ROX	KY769234
CL540*	F: GGGTAGTGATCATGGAAGA R: AGAAAGCAGATATCGGTTG	$(TC)_{10}A(CA)_{10}$	118–160	19	63	ROX	KY769211
CL564*	F: TGAACTTGAACCTTGTGACTTAGC R: TACAAAACTGTGGGCTTGATGATG	$(CT)_{12}$	204–238	15	65	HEX	KY769212
CL586*	F: CAGCAAAGAAACGGTTATGGT R: GTTTTGTGGTATCCAACTAGG	$(TC)_9$	208–224	6	63	6-FAM	KY769213
CL631	F: TGGTGAGGAAGGATTCAGCCGACAG R: TCAGTTCCGGTTAGGCTCAGTACAC	$(TG)_5$	273	1	65	ROX	KY769222
CL653	F: AATGGAGGTATTGCAAC R: ACCTGTAAGGGTTTTCC	$(AG)_{20}$	114	1	59	ROX	KY769223
CL654*	F: TCTCTCCTCCCTTTGCTTACG R: CCATGCGTTGAAGAAGTATCG	$(TC)_8$	126–144	9	63	HEX	KY769214
CL723*	F: ATCTCTGTCTTTTGCACTCTC R: GGAATTATTGTTGGGGTTAGG	$(TC)_6$	211–227	5	63	TAMRA	KY769215
CL753	F: TAGAATCAACGCACAAGAAAGGC R: ACTCAAAAACATGACTCGGTAGC	$(GA)_9$	232	1	65	HEX	KY769236
CL761	F: CCTCTTATGACACATTGGT R: TTTTCAGATGACTCTCGGA	$(CT)_6$	138	1	59	HEX	KY769224
CL776*	F: ACTGCAAAGGAGATACGCTGAAGG R: GACGCAAAAGAGGTCCACAATACA	$(TC)_8$	221–225	3	65	TAMRA	KY769216
CL783	F: CTAGATACGAGTGTCGAAGA R: GCAATACACATACACACAGA	$(TC)_6$	183	1	65	HEX	KY769237
CL852*	F: CTAGTGTCCAAAAAAGAGCA R: GAGATATGAGTAGAATGAGG	$(CT)_9$	136–148	7	63	ROX	KY769217
CL871	F: TGGTCCGCGTTACAAGTATACATG R: ACTCTGCCCTTTTCACTATTCTGC	$(GT)_5$	228	1	65	HEX	KY769225
CL878*	F: CAGGGTAGCCTTTGAAACA R: GGCTCCATATAACAACATC	$(AG)_8$	147–163	7	64	ROX	KY769218

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

[a] All values are based on samples representing three populations located in Lichuan in Hubei Province, Lin'an in Zhejiang Province, and Wanxian in Chongqing Municipality, China.

[b] Fluorescent dyes (i.e., HEX, ROX, 6-FAM, and TAMRA) used for fragment analysis.

* Polymorphic microsatellite loci.

TABLE 2. Genetic properties of 14 newly developed polymorphic microsatellites of *Cunninghamia lanceolata*.[a]

Locus	Lichuan population				Lin'an population				Wanxian population			
	n	A	H_o	H_e	n	A	H_o	H_e	n	A	H_o	H_e
CL34	30	4	0.2000*	0.5847	25	3	0.2000*	0.5200	29	4	0.2069*	0.5850
CL90	32	5	0.6563	0.6880	32	4	0.3750	0.5709	30	4	0.5667	0.6356
CL108	32	5	0.1875	0.2336	32	4	0.3125	0.3021	30	3	0.2333	0.3203
CL295	32	4	0.5938	0.6047	32	4	0.6563	0.5843	30	4	0.5000	0.5701
CL343	32	16	0.1875*	0.8750	32	17	0.1563*	0.9246	29	14	0.2759	0.8209
CL389	30	7	0.2667	0.3819	32	5	0.1250	0.1220	30	5	0.2000	0.1904
CL540	29	15	0.3793*	0.9165	29	14	0.5172*	0.9201	28	12	0.5000	0.8935
CL564	32	12	0.8125	0.8705	32	11	0.6250	0.6999	30	10	0.8333	0.8621
CL586	31	3	0.1613	0.2089	32	6	0.2813	0.3105	30	4	0.2667	0.2706
CL654	32	7	0.5625	0.6002	32	6	0.6250	0.6969	30	6	0.6333	0.6418
CL723	32	5	0.0938*	0.4454	32	3	0.1563*	0.4043	30	4	0.1000*	0.5028
CL776	32	3	0.3438	0.3031	32	2	0.0313	0.0313	30	2	0.3667	0.3045
CL852	32	4	0.2500	0.2555	32	6	0.4688	0.3973	30	4	0.6000	0.4729
CL878	32	6	0.1875	0.2073	32	5	0.2813	0.4772	30	3	0.2000	0.1859

Note: A = number of alleles; H_e = expected heterozygosity; H_o = observed heterozygosity; n = number of individuals genotyped.

[a] Locality and voucher information for the populations are available in Appendix 1.

* Indicates significant deviation from Hardy–Weinberg equilibrium ($P < 0.05$).

strain JM109 (TaKaRa Biotechnology Co.) through transient thermal stimulation following the TaKaRa *E. coli* JM109 competent cell protocol.

A total of 1400 colonies were selected and tested by PCR with $(AG)_{10}$ and M13+/M13− as primers, producing 469 positive sequences. The positive PCR products were sequenced on an ABI 3730 DNA Sequence Analyzer (Applied Biosystems, Foster City, California, USA). We designed 263 primer pairs using the software Primer Premier 5.0 (PREMIER Biosoft International, Palo Alto, California, USA). The criteria for primer design were: (1) primer length between 18–28 bp and amplicon length between 100–300 bp; (2) melting temperature (T_m) in the range of 45–65°C; (3) GC content of the sequence between 40–60%, with no more than three Gs or Cs in the last five bases at the 3′ end of the primer without mismatch or secondary structures. Polymorphisms were detected using 24 individuals selected randomly from three populations from Lichuan (Hubei Province, China), Lin'an (Zhejiang Province, China), and Wanxian (Chongqing Municipality, China), according to the method proposed by Schuelke (2000). The nested PCR used three primers: a sequence-specific forward primer with an M13(−21) tail (5′-TGTAAAACGACGGCCAGT-3′) at its 5′-end, a universal M13(−21) primer labeled with a fluorescent dye, and a sequence-specific reverse primer. In the first 30 cycles, the forward primer was incorporated into the PCR products. Then these products were marked with the fluorescently labeled M13(−21) primer, which was incorporated during the following eight cycles at 53°C. Each reaction was run in a 20-µL system containing the following: 40 ng of DNA, 1× PCR buffer, 2.0 mM Mg²⁺, 0.2 mM of dNTPs, 0.1 µM M13(−21) primer labeled with HEX, ROX, or 6-FAM (Sangon Biotech, Shanghai, China), 0.025 µM forward primer with an M13(−21) tail, 0.1 µM reverse primer, and 2 units *Taq* DNA polymerase (Sangon Biotech). Thermocycling conditions were as follows: 5 min of denaturation at 94°C; 30 cycles of 30 s at 94°C, 45 s at 59–65°C (Table 1), and 45 s at 72°C; followed by eight cycles of 30 s at 94°C, 45 s at 53°C, and 45 s at 72°C; and a 10-min extension at 72°C. We genotyped the amplification products on an ABI 3730 automated sequencer using GeneScan 500 LIZ Size Standard (Applied Biosystems), and alleles were identified using GeneMapper 4.0 software (Applied Biosystems). Among the 263 primer pairs tested, 186 were not amplified in any samples, 32 did not produce clear and single-target bands, 12 could not be successfully amplified in the individuals used for testing, and three were duplicates of previous markers in Li et al. (2015). Finally, 14 polymorphic and 16 monomorphic microsatellite loci were obtained (Table 1).

All polymorphic loci were further characterized using 94 *C. lanceolata* individuals from the same three populations mentioned above. One of four fluorescent dyes (HEX, ROX, 6-FAM, TAMRA) labeled the forward primers. PCRs were set up in 10-µL reaction volumes containing 40 ng of template DNA, 1× PCR buffer, 2.5 mM Mg²⁺, 0.2 mM of each dNTP, 0.1 µM forward and reverse primer, and 1 unit of *Taq* DNA polymerase. We used the following conditions: 5 min of denaturation at 94°C; 30 cycles of 30 s at 94°C, 45 s at 61–65°C (Table 1), and 1 min at 72°C; and a final extension at 72°C for 10 min. The amplification products were scanned on an ABI 3730 sequencer using GeneScan 500 LIZ Size Standard (Applied Biosystems), and alleles were called and binned using GeneMapper 4.0 software (Applied Biosystems).

The number of alleles varied from three to 30 with an average of 9.4 using FSTAT 2.9.3 software (Goudet, 1995) (Table 1). The observed and expected heterozygosities ranged from 0.0313–0.8333 and 0.0313–0.9246, respectively, analyzed by the software TFPGA version 1.3 (Miller, 1997) (Table 2). After the sequential Bonferroni adjustment (Rice, 1989), only two loci (CL34 and CL723) displayed significant deviations from Hardy–Weinberg equilibrium ($P < 0.05$) in all three populations, and two loci (CL343 and CL540) deviated significantly from Hardy–Weinberg equilibrium in the Lichuan and Lin'an populations. No loci exhibited significant linkage disequilibrium after sequential Bonferroni adjustment by GENEPOP version 4.0 (Rousset, 2008). Null alleles were likely to be present in the loci CF34, CF343, CF389, CF540, and CF723 using MICRO-CHECKER version 2.2.3 (van Oosterhout et al., 2004).

We also tested the performance of 14 polymorphic primer pairs in five related species belonging to the same family (Taxodiaceae), using one to 10 individuals from each species (Table 3; Appendix 1). In *Metasequoia glyptostroboides* Hu & W. C. Cheng, successful amplifications occurred in seven loci (CL90, CL295, CL343, CL723, CL776, CL852, and CL878), of which three (CL90, CL723, and CL776) could also be amplified in *Cryptomeria fortunei* Hooibr. ex Otto & A. Dietr. and two (CL852 and CL878) in *Taxodium ascendens* Brongn. No polymorphic loci amplified successfully in *Cryptomeria japonica* (Thunb. ex L. f.) D. Don or in *Sequoia sempervirens* (D. Don) Endl. (Table 3). Unfortunately,

TABLE 3. Allele size ranges tested in five additional taxa for cross-amplification trials of SSR loci isolated from *Cunninghamia lanceolata*.

Locus	*Cryptomeria japonica* ($n = 1$)	*Cryptomeria fortunei* ($n = 5$)	*Metasequoia glyptostroboides* ($n = 10$)	*Sequoia sempervirens* ($n = 2$)	*Taxodium ascendens* ($n = 2$)
CL90	—	268	218–220	—	—
CL295	—	—	159–165	—	—
CL343	—	—	193	—	—
CL723	—	261	237–247	—	—
CL776	—	225	217–225	—	—
CL852	—	—	142	—	142
CL878	—	—	145–155	—	145

Note: — = primers could not be amplified.

we were unable to sample the only congeneric species of *C. lanceolata* (*C. konishii* Hayata), and thus the cross-amplification capability of these primers is likely underestimated.

CONCLUSIONS

Of the 30 markers reported here for *C. lanceolata*, 14 microsatellite loci showed a high level of polymorphism. These loci will be used to study population genetic diversity, gene flow, and mating systems. Combined with the previously isolated loci, these markers will facilitate the further investigation of parentage analyses and kinships between the planted and natural populations of *C. lanceolata*, all of which are relevant to germplasm conservation and forest management of this timber species.

LITERATURE CITED

Bao, F. C., and Z. H. Jiang. 1998. Wood properties of main tree species from plantation in China. China Forestry Publishing House, Beijing, China.

Charlesworth, B., P. Sniegowski, and W. Stephan. 1994. The evolutionary dynamics of repetitive DNA in eukaryotes. *Nature* 371: 215–220.

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

Huang, Z., Z. Xu, S. Boyd, and D. Williams. 2005. Chemical composition of decomposing stumps in successive rotation of Chinese fir (*Cunninghamia lanceolata* (Lamb.) Hook.) plantations. *Chinese Science Bulletin* 50: 2581–2586.

Li, Y. X., Z. S. Wang, J. K. Sui, Y. F. Zeng, A. G. Duan, and J. G. Zhang. 2015. Isolation and characterization of microsatellite loci for *Cunninghamia lanceolata* (Lamb.) Hook. *Genetics and Molecular Research* 14: 453–456.

Miller, M. P. 1997. Tools for population genetic analyses (TFPGA) 1.3: A Windows program for the analysis of allozyme and molecular population genetic data. Website http://www.marksgeneticsoftware.net/tfpga.htm [accessed 28 August 2017].

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. *Molecular Ecology Resources* 8: 103–106.

Schuelke, M. 2000. An economic method for the fluorescent labeling of PCR fragments. *Nature Biotechnology* 18: 233–234.

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.

Wen, Y., S. Ueno, W. Han, and Y. Tsumura. 2013. Development and characterization of 28 polymorphic EST-SSR markers for *Cunninghamia lanceolata* (Taxodiaceae) based on transcriptome sequences. *Silvae Genetica* 62: 137–141.

Xu, Y., R. Zheng, Z. Wang, Y. Wang, Z. Hong, L. Yang, Y. Lu, et al. 2016. Identification and characterization of genic microsatellites in *Cunninghamia lanceolata* (Lamb.) Hook. (Taxodiaceae). *Archives of Biological Sciences* 68: 417–425.

Yeh, F. C., J. Shi, R. Yang, J. Hong, and Z. Ye. 1994. Genetic diversity and multilocus associations in *Cunninghamia lanceolata* (Lamb.) Hook. from The People's Republic of China. *Theoretical and Applied Genetics* 88: 465–471.

Appendix 1. Locality information of *Cunninghamia lanceolata* and its related species used in this study. Voucher specimens were deposited in East China Normal University (HSNU), Shanghai, China.

Species	Population	Collection locality	Geographic coordinates	N
Cunninghamia lanceolata (Lamb.) Hook.	Tiantong	Tiantong, Zhejiang, China	29°48′19″N, 121°47′43″E	1
	Lichuan	Lichuan, Hubei, China	30°10′37″N, 108°37′03″E	32
	Lin'an	Lin'an, Zhejiang, China	30°19′14″N, 119°26′04″E	32
	Wanxian	Wanxian, Chongqing, China	30°39′43″N, 108°45′05″E	30
Cryptomeria japonica (Thunb. ex L. f.) D. Don	Shanghai	Shanghai Botanic Garden, Shanghai, China	31°08′48″N, 121°26′50″E	1
Cryptomeria fortunei Hooibr. ex Otto & A. Dietr.	Shanghai	Shanghai Botanic Garden, Shanghai, China	31°08′48″N, 121°26′50″E	5
Sequoia sempervirens (D. Don) Endl.	Nanjing	Nanjing Botanic Garden, Nanjing, China	32°04′15″N, 118°48′25″E	1
	Hangzhou	Hangzhou Botanic Garden, Hangzhou, China	30°15′19″N, 120°07′22″E	1
Metasequoia glyptostroboides Hu & W. C. Cheng	Lichuan	Lichuan, Hubei, China	30°10′22″N, 108°39′32″E	10
Taxodium ascendens Brongn.	Nanjing	Nanjing Botanic Garden, Nanjing, China	32°04′15″N, 118°48′25″E	2

Note: N = number of individuals.