Molecular Biology Reports (2020) 47:2991–2996
https://doi.org/10.1007/s11033-020-05303-y

SHORT COMMUNICATION

Novel 28 microsatellite loci using high-throughput sequencing for an endangered species on *Metasequoia glyptostroboides* (Cupressaceae)

Jing-Wen Wang[1] · Tong-Lei Xu[1] · Ghenet Weldegebriel Sereke[1] · Rong Wang[1] · Yuan-Yuan Li[1,2]

Received: 16 November 2019 / Accepted: 31 January 2020 / Published online: 8 February 2020

Abstract
Metasequoia glyptostroboides is a living fossil and an endangered species listed in the International Union for Conservation of Nature (IUCN). Distinguishing the genotypes of all wild individuals of *M. glyptostroboides* is important to delimit management units and key germplasm resources. We characterized 28 novel polymorphic microsatellite loci using a streptavidin–biotin microsatellite-enriched library and Illumina high-throughput sequencing. Characteristics of each locus were tested using 140 individuals collected from five natural populations of *M. glyptostroboides*. The number of alleles per locus ranged from 3 to 20, with a mean number of about 8 alleles. The observed and expected heterozygosities in each population ranged from 0.0000 to 1.0000 and from 0.0000 to 0.8958, respectively. Four to nine loci were cross-amplified successfully in seven species of Cupressaceae. The novel SSR markers will provide a toolkit for DNA identification of all of the extant wild individuals guiding further conservation efforts of *M. glyptostroboides*.

Keywords Cupressaceae · Genetic structure · Germplasm resources · Illumina sequencing · *Metasequoia glyptostroboides*

Introduction

Metasequoia glyptostroboides Hu et Cheng (Cupressaceae) is the only existing species in this genus. *Metasequoia* was thought to originate in the Cretaceous. It was originally described from fossil material and was thought to have become extinct several million years ago [1]. Living individuals of *M. glyptostroboides* were discovered in central China in 1940s. *M. glyptostroboides* is considered to be a rare "living fossil" that requires conservation [2]. The International Union for Conservation of Nature (IUCN) has designated this flagship species [3, 4] as an endangered species. It has a limited distribution range (a very small region in central China) [5] and a small population size (ca. 5500 individuals) [6]. Given the critical role of genetic factors for long-term maintenance of endangered species [7], it is important to develop molecular markers to increase understanding of the genetic diversity and genetic structure of *M. glyptostroboides*.

Earlier studies on *M. glyptostroboides* identified 11 polymorphic microsatellite loci [8] and 23 EST-SSR markers [9]. These studies provided baseline genetic information. However, the genotypes described by these microsatellite markers are insufficient to distinguish all of the extant wild individuals of *M. glyptostroboides*, probably due to high level of inbreeding [2, 5]. A lack of sufficient genetic information on *M. glyptostroboides* has impeded the delimitation of management units and key germplasm resources. This is essential for formulating an effective conservation strategy. In this study, we characterized 28 novel polymorphic microsatellite markers by high-throughput sequencing. These polymorphic loci showed moderate to high levels of genetic variation and provide markers for future conservation genetic studies on *M. glyptostroboides*.

✉ Yuan-Yuan Li
yyli@des.ecnu.edu.cn

[1] Zhejiang Tiantong Forest Ecosystem National Observation and Research Station, Shanghai Key Laboratory for Urban Ecology and Restoration, School of Ecological and Environmental Sciences, East China Normal University, Shanghai 200241, China

[2] Institute of Eco-Chongming (IEC), Shanghai 200062, China

Materials and methods

Microsatellite library development and primer selection

Sequences of microsatellite loci were obtained by constructing a microsatellite-enriched genomic library following next-generation sequencing at Personalbio (Shanghai, China). First, the total genomic DNA was extracted from ~ 100 mg of fresh leaf tissues obtained from five wild individuals of *M. glyptostroboides* using a modified cetyltrimethylammonium bromide (CTAB) method [10]. Individuals were selected from three populations in the natural range (Guihua, Chatai, and Fanshen populations in Lichuan District, Hubei Province, China). The crude DNA extracts of the five individuals were purified using the GF-1 plant DNA extraction kit (Vivantis Technologies Sdn. Bhd., Subang Jaya, Malaysia) and then they were mixed equally. The mixed high-quality DNA was sheared to generate fragments of about 400 bp length to construct a genomic library following the Illumina preparation protocol. To enrich the DNA fragments containing microsatellites, SSR fragments in the genomic library were captured by selective hybridization with probes for ten dinucleotide repeat units ($(AG)_{10}$ and $(AC)_{10}$), eight trinucleotide repeats ($(AAC)_8$, $(ACG)_8$, $(AAG)_8$, and $(AGG)_8$), and six tetranucleotide repeats ($(ACAT)_6$ and $(ATCT)_6$), and were adhered by streptavidin-coated magnetic beads. The microsatellite-enriched library was sequenced on an Illumina MiSeq PE400 platform (Illumina, San Diego, CA, USA), which generated 250 bp paired-end reads. To obtain clean reads, the adapters and sequences < 50 bp in length and low-quality reads were removed.

The clean reads were then assembled to combined pairs using FLASH version 1.2.11 [11]. These pairs were used for detecting microsatellites by the MISA module [12]. After deleting the sequences with microsatellite flanking sequences (lengths < 20 bp), the remaining microsatellite sequences were clustered based on the criterion of sequence similarity > 98% using UCLUST version 1.2.22q [13]. Among them, we chose clusters with length polymorphism to develop microsatellite markers. Primers of 120 clusters (loci) were successfully designed using PRIMER3 version 2.3.6 [14].

PCR and fragment analysis

Availability of these selected microsatellite loci were tested using 28 randomly chosen individuals from two natural populations. After DNA extraction from silica gel-dried leaves using the plant genomic DNA kit (Tiangen, Beijing, China). The nest PCR was conducted according to the method proposed by Schuelke [15] in a 10-μL system composed of: 50 ng of template DNA, 1× PCR buffer, 2.0 mM Mg^{2+}, 0.2 mM dNTPs, 0.025 μM forward primer with an M13(-21) tail (5′-TGTAAAACGACGGCCAGT -3′) at its 5′-end, 0.1 μM reverse primer, 0.1 μM universal M13(-21) primer labeled with HEX, ROX, TAMRA or 6-FAM (Sangon Biotech, Shanghai, China), and one unit of *Taq* DNA polymerase (Sangon Biotech, Shanghai, China). Thermocycling conditions were performed as follows: an initial 5 min of denaturation at 94 °C, followed by 30 cycles of denaturation (94 °C for 30 s), annealing (54–65 °C for 45 s, depending on the specific locus, Table 1), and extension (72 °C for 45 s), followed by eight cycles of 30 s at 94 °C, 45 s at 53 °C and 45 s at 72 °C, and a final extension step at 72 °C for 10 min. PCR products that showed clear bands on 8% agarose gels were scanned on an ABI 3730 DNA sequence analyzer using GeneScan 500 LIZ size standard (Applied Biosystems, Foster City, CA, USA), and alleles were called and binned in GeneMapper 4.0 software (Applied Biosystems). Finally, PCR products in 28 of the 120 microsatellite loci were amplified successfully in all of the individuals, and none of them had previously been reported based on the results of BLAST searches in GenBank (Table 1).

The 28 polymorphic loci were further characterized using 140 individuals from five natural populations of Xiaohe, Guihua, Chatai, and Fanshen populations in Lichuan District, Hubei Province and the Shizhu population in Chongqing Municipality. We labeled the forward primers of these loci using fluorescent dyes (5′HEX, 5′ROX, 5′TAMRA and 5′6-FAM), and multiple PCRs were performed in a 10-μL reaction system containing 50 ng of template DNA, 1× PCR buffer, 2.5 mM Mg^{2+}, 0.2 mM of dNTP mixture, 0.15 μM forward and reverse primer, and one unit of *Taq* DNA polymerase, with the above-mentioned PCR protocol. Scanning of PCR products as well as calling and binning of alleles were implemented using the procedure describe above. We also tested cross-amplification in seven species of Taxodiaceae (now in the Cupressaceae) with one to four individuals, respectively, using the same extraction method and amplification conditions as described above.

Statistical analysis

Using TFPGA version 1.3 [16], the number of alleles per locus (*A*), observed heterozygosity (H_o) and expected heterozygosity (H_e) were determined. Polymorphic information content (*PIC*) by locus was estimated by Cervus version 3.0.7. Linkage disequilibrium was detected using FSTAT version 2.9.3 [17], and Hardy–Weinberg equilibrium was conducted with GENEPOP software [18]. We used MICRO-CHECKER version 2.2.3 [19] to estimate the null alleles.

Table 1 Characterization of 28 polymorphic microsatellite loci developed in *M. glyptostroboides*

Locus	Primer sequences (5–3′)	Repeat motif	T_a (°C)	*A*	Allele size range[a] (bp)	Fluorescent dye[b]	*PIC*	GenBank accession number
MG80	F: TCGAAATATACCTTGGGCGA R: CTTGCAGTGAAAGAAACACACA	$(AT)_9$	62	11	272–292	6-FAM	0.7400	MN599556
MG84	F: AATGCCCCTATGATTCTCCA R: CCATTCAAGCATCCATAGCC	$(AAG)_{13}$	59	11	255–309	ROX	0.8440	MN599557
MG90	F: TCTCCATCTTCCTTCTCGGA R: AGGCTTTGAGTGTTCGAGGA	$(CTT)_7$	59	3	229–238	ROX	0.2840	MN599558
MG94	F: AGACAAGAGGAGGAGGAGCC R: GCCACCCTTTGATCTGTCAT	$(GAT)_5$	65	5	225–267	6-FAM	0.0820	MN599559
MG95	F: ATGCATGTCCTTGAAAAGGC R: TGAGGATGGAAGGAGAGCAG	$(TTC)_{10}$	56	9	299–323	6-FAM	0.7460	MN599560
MG98	F: ATGCACATAAATGGTGGGGT R: TGTGAGTCTCAAATTCAAGGGA	$(TC)_{18}$	59	20	352–410	ROX	0.8370	MN599561
MG100	F: AGCCTAATGAAGGGCTTGAA R: ACCAAATCATCATCAAGGGC	$(AT)_9$	64	11	364–396	ROX	0.5620	MN599562
MG102	F: TCATAGACCCTACAAGATAGAAATCAA R: CATCTCCCACCCATGTTACC	$(AT)_7$	63	9	240–264	HEX	0.4920	MN599563
MG106	F: TGGGCGATAGCCTCATCTAT R: TTCCCACCTGGGTCACTAGA	$(AT)_8$	54	4	327–341	HEX	0.3890	MN599564
MG110	F: CCTCCGAAAAGAAAAAGAGGA R: AACTAGATGGGGTGGAAGCA	$(AAG)_9$	62	5	235–250	HEX	0.6900	MN599565
MG111	F: CCATAAACACAATCGGTACACA R: TGGTGAGGAAGTAGATGGGG	$(TA)_{10}$	63	20	257–301	6-FAM	0.8520	MN599566
MG112	F: TAGAGCATCATCAACGGCTG R: AAAGGCTCCAACAACTCTGG	$(AT)_7$	54	11	360–384	ROX	0.6400	MN599567
MG113	F: TTTCTCAAGTCCATGCTCCC R: CCCCTCCAAAAGTAAATTCTCC	$(TTC)_5$	63	3	448–460	6-FAM	0.0210	MN599568
MG116	F: CTTGTCCACCTCATCAGCCT R: TCATCAGGTGAACCAACCAA	$(TTC)_{21}$	59	13	281–344	ROX	0.5950	MN599569
MG117	F: TTGGGAAAGTTTGACACAAGG R: TTCTTCCATTGCCTTCATCC	$(AGA)_{10}$	61	7	431–461	HEX	0.6870	MN599570
MG121	F: AGAGGGAGAGGAGGTTGAGG R: AGGCTTGCCCTGATGATATG	$(GAG)_6$	62	8	181–220	ROX	0.6540	MN599571
MG124	F: AAACCAACACCCAAAATCCA R: AAGGCTTGTTGTTGGGAATG	$(ACA)_5$	63	4	293–329	6-FAM	0.3920	MN599572
MG130	F: TGTGGTCTTGACCTAAAATTTGATA R: TGGTAAATACAATCTACATGAAATGC	$(AT)_7$	59	3	212–226	HEX	0.3850	MN599573
MG136	F: GCATACATACAAGCATACATGCATAA R: TGTCCCACTTGGAATTGTTT	$(CATA)_6$	57	3	198–234	HEX	0.0140	MN599574
MG137	F: ATGCACACACACATGCACAC R: TGAGACACCAAGGTTGAAGG	$(CA)_6…(CA)_7$	63	6	277–307	HEX	0.2060	MN599575
MG140	F: ACAATGGGAAGGAGGAGGAT R: GTGCAGGGATGGTCTTGTCT	(GAG)7	58	3	240–261	ROX	0.1000	MN599576
MG141	F: GCATTCAAGTGCACTGCCTA R: TGGATGTGTGGAATGATGTG	$(TC)_{15}$	64	15	274–324	ROX	0.5590	MN599577
MG142	F: CTCATCCCAAGAAGACCGAG R: GGTTGTGGCTCCTGCATAGT	$(GGA)_6$	64	4	167–182	ROX	0.3960	MN599578
MG145	F: GCAAGCCCTGCTAAGAAATG R: TCATGGGAGTGTGTGTGCTT	$(GAA)_7$	59	5	360–384	TAMRA	0.3870	MN599579
MG146	F: TTAGGCAGTCTCGACTCGGT R: GGTACACCTGGAACTCCCCT	$(AT)_6$	65	15	357–391	HEX	0.8080	MN599580
MG150	F: TGAGGATGGAAGGAGAGCAG R: TTGTCCTTCAGCTGATACATCATT	$(AGA)_{12}$	63	7	133–157	ROX	0.7120	MN599581
MG156	F: AAGGAGGAAGAGAGGAGGAGAG R: CCCCAACATTTTGTCTCCAT	$(AGG)_5$	63	3	158–182	6-FAM	0.3540	MN599582
MG164	F: TAAGGAGGGGAAGGGGTAGA R: ATCCTCCTTCTCGACATTGC	$(AGG)_5$	59	3	129–135	HEX	0.1370	MN599583

A number of alleles, T_a annealing temperature, *PIC* polymorphic information content

[a]The values are based on samples collecting from five populations located in Xiaohe, Guihua, Chatai, and Fanshen populations in Lichuan District, Hubei Province and Shizhu population in Chongqing Municipality of China

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

Results and discussion

By Illumina sequencing of five individuals of *M. glyptostroboides* genomic DNA, 6,030,456 raw reads were obtained, with approximately 1510 million bases. The average percentage of GC was 34.88. After removal of adapters, reads < 50 bp in length and low-quality reads, 5,197,340 clean reads were obtained, which were then assembled into 1,651,083 combined pairs. A total of 410,388 microsatellites were identified. Among these, 255,417 (62.24%) microsatellites were dinucleotide repeat motifs, 40,245 (9.81%) were trinucleotide repeat motifs, 29,965 (7.30%) were tetranucleotide repeat motifs, 149 (0.04%) were pentanucleotide repeat motifs, and 10,540 (2.57%) were hexanucleotide repeat motifs. Among these microsatellite sequences, 173,281 clusters were obtained under the criterion of sequence similarity > 98%. However, only 280 of 173,281 clusters (0.16%) showed length polymorphism across the five individuals, indicating low variation in most of the SSR loci in *M. glyptostroboides*. Among the 280 clusters, 120 were designed successfully to be microsatellite primers. Given the conservatism of this species [8], this is a robust and functional high-throughput sequencing approach for microsatellite loci development.

In five populations, 28 polymorphism microsatellite loci were characterized, and they showed clean and reproducible amplification (Table 1). Signs of null alleles were present in ten loci (MG80, MG84, MG94, MG98, MG100, MG111, MG113, MG116, MG145, and MG146). We found no linkage disequilibrium between any two loci. However, significant deviation from Hardy–Weinberg equilibrium was prevalent in most loci (Table 2) using setting sequential Bonferroni adjustment, likely as the consequence of high levels of inbreeding [2, 5] or null alleles. The SSR

Table 2 Characterization of the 28 polymorphic microsatellite loci in five *M. glyptostroboides* populations

Locus	Shizhu ($N=32$)			Xiaohe ($N=33$)			Gguihua ($N=32$)			Chatai ($N=13$)			Fanshen ($N=30$)		
	A	H_o	H_e	A	H_o	H_e	A	H_o	H_e	A	H_o	H_e	A	H_o	H_e
MG80	9	0.5625*	0.7515	8	0.3030*	0.7152	7	0.3750*	0.7971	3	0.4615	0.5938	4	0.6333	0.6429
MG84	9	0.7188	0.8681	7	0.8485	0.8289	9	0.6563	0.7946	4	1.0000	0.7723	6	0.9667*	0.7938
MG90	2	0.6250	0.4365	2	0.0303	0.0303	3	0.4688	0.3924	2	0.4615	0.3692	2	0.4000	0.3254
MG94	3	0.0313	0.0620	2	0.0000	0.0597	3	0.0938	0.0918	2	0.0769	0.0769	2	0.1333	0.1266
MG95	8	0.7188	0.8244	6	0.5152	0.7478	7	0.7188	0.7500	5	0.8462	0.7015	4	0.8000	0.6808
MG98	9	0.3125*	0.7138	15	0.6364*	0.8946	15	0.6250*	0.8958	3	0.4615	0.6277	3	0.5333*	0.6689
MG100	8	0.3438*	0.5933	9	0.3333*	0.7030	6	0.3125*	0.7073	2	0.6154	0.4923	2	0.4667	0.3638
MG102	7	0.8750	0.7232	3	0.3636	0.3506	6	0.4688	0.4226	3	0.8462	0.5631	3	0.6667	0.5079
MG106	2	0.6563	0.4678	3	0.8788*	0.5366	2	0.6250	0.4762	2	0.8462	0.5077	3	0.8667*	0.5356
MG110	4	0.6563	0.7133	4	0.6364*	0.7315	5	0.5625	0.6900	3	1.0000	0.6492	3	0.9667*	0.6028
MG111	11	0.4063*	0.8433	14	0.8182	0.8909	12	0.7188	0.8512	5	0.3077	0.7108	6	0.5333	0.6153
MG112	8	0.7813	0.7495	8	0.8788	0.7077	7	0.5313	0.5898	3	0.9231	0.6246	4	0.9333*	0.6571
MG113	2	0.0000	0.0615	2	0.0303	0.0303	1	0.0000	0.0000	1	0.0000	0.0000	2	0.0000	0.0000
MG116	7	0.2813*	0.6959	6	0.6061*	0.6909	7	0.3438*	0.7059	4	0.6154	0.5323	3	0.2667	0.2390
MG117	6	0.6875	0.6786	7	0.5152	0.5995	4	0.5000	0.6751	5	0.6154	0.7754	4	0.7000	0.6650
MG121	4	0.5000	0.6186	3	0.6061	0.6252	4	0.5000	0.6186	4	0.3846	0.6615	6	0.4000*	0.6424
MG124	4	0.4063*	0.5670	2	0.4545	0.5002	2	0.6250	0.4583	2	0.7692	0.5169	2	0.6333	0.5079
MG130	2	1.0000*	0.5079	2	1.0000*	0.5077	2	0.9063*	0.5074	2	0.9231	0.5169	3	0.9000*	0.5384
MG136	1	0.0000	0.0000	2	0.0303	0.0303	2	0.0313	0.0313	1	0.0000	0.0000	1	0.0000	0.0000
MG137	3	0.2188	0.2039	4	0.3333	0.3641	5	0.2500	0.2326	2	0.0769	0.0769	2	0.0667	0.0655
MG140	1	0.0000	0.0000	1	0.0000	0.0000	2	0.0000*	0.1190	3	0.3846	0.4400	2	0.0000*	0.1266
MG141	9	0.4688	0.5942	7	0.6061*	0.6527	10	0.3750*	0.6558	5	0.4615	0.5138	5	0.2000*	0.4266
MG142	3	0.9688*	0.5233	2	0.9697*	0.5072	4	1.0000*	0.5531	2	1.0000*	0.5200	2	0.9667*	0.5079
MG145	3	0.2813	0.4385	4	0.2727*	0.5273	3	0.4063	0.3497	2	0.3846	0.5077	2	0.5000	0.4130
MG146	8	0.6250	0.8423	12	0.4242*	0.8480	12	0.5000*	0.8735	4	0.1538*	0.6185	4	0.1000*	0.3475
MG150	7	0.7813	0.7872	6	0.4848	0.7287	6	0.5938	0.7326	3	0.7692	0.6738	3	0.8000	0.6401
MG156	3	0.3438	0.4410	2	0.1818	0.3730	2	0.2813	0.3646	2	1.0000*	0.5200	2	1.0000*	0.5085
MG164	3	0.1563	0.2009	2	0.0606	0.1156	3	0.0625	0.1205	2	0.0000	0.2708	2	0.0000	0.0655

A number of alleles per locus, H_o observed heterozygosity, H_e expected heterozygosity, N number of individuals genotyped

*Deviation from Hardy–Weinberg equilibrium at $P < 0.05$ after Bonferroni correction

Table 3 Allele size ranges tested in seven related species for cross-amplification trials of 11 SSR loci isolated from *M. glyptostroboides*

Locus	*Sequoia sempervirens* (*N* = 4)	*Sequoiadendron giganteum* (*N* = 3)	*Taxodium ascendens* (*N* = 4)	*Taxodium distichum* (*N* = 2)	*Cryptomeria fortunei* (*N* = 3)	*Cryptomeria japonica* (*N* = 1)	*Cunninghamia lanceolata* (*N* = 2)
MG84	–	–	267–288	288	267–276	–	–
MG95	–	323	314–323	314–323	–	–	–
MG102	262	262	246–258	252–258	256	256	–
MG106	–	337	–	–	–	–	–
MG111	281	281	269	269	265	–	259–273
MG130	–	–	224	224	224	224	224
MG136	–	–	214	214	214	–	–
MG141	–	–	294	294	332–336	–	–
MG142	–	176	164–179	164–179	167–179	164	167–179
MG150	151	–	145–154	148–154	148–154	151	148–154
MG156	179	161–176	–	–	–	–	–

– No amplification with these primers

loci showed different genetic diversity patterns across the five natural populations of *M. glyptostroboides*. Overall, we observed 221 alleles across the 28 microsatellite loci. The number of alleles per locus ranged from 3 to 20, with an average of 7.9. *PIC* ranged from 0.014 to 0.852. H_o and H_e in each locus varied from 0.0000 to 1.0000 and from 0.0000 to 0.8958, respectively (Table 2). In addition, cross-species amplification showed that only four to nine loci were functional in seven species from other genera in Cupressaceae (Table 3). This is likely because of the great genetic divergence between *M. glyptostroboides* and these distantly related species [20].

We developed and characterized 28 novel and polymorphic microsatellite markers for *M. glyptostroboides*. These markers can be used to distinguish the microsatellite genotypes of all of the existing wild trees and give a unique DNA identification to each individual. These markers will also provide a genetic tool for describing the genetic diversity and genetic structure of *M. glyptostroboides* populations. Furthermore, this information is crucial for improving our understanding of gene flow, population dynamics, and germplasm resources of *M. glyptostroboides* and will aid in conservation and management efforts.

Acknowledgements The authors thank X.-W. Sun, X.-H. Yue, and S.-S. Wang for their supports for field surveys and laboratory experiments. We also thank Dr. S. R. Biswas and LetPub (www.letpub.com) for linguistic assistance during the preparation of this manuscript. This work was supported by the National Key Research and Development Program (2016YFC0503102) and the National Natural Science Foundation of China (31670540).

Author contributions J-WW wrote the manuscript, and collected and analyzed data. T-LX and GWS contributed to collection of data and revised the manuscript. RW helped improve the design of the experiment and critically revised the manuscript. Y-YL designed the experiment and critically revised the manuscript.

Data availability Raw sequencing data were deposited in the National Center for Biotechnology Information (NCBI) Sequence Read Archive (BioProject ID: PRJNA578834). Sequence information for the develop primers has been deposited to NCBI, and GenBank accession numbers are provided in Table 1.

Compliance with ethical standards

Conflict of interest The authors declare no conflict of interest.

Research involving human participants and/or plants This research involved endangered plants; studies were supported by the projects of Ministry of Science and Technology of the People's Republic of China and National Natural Science Foundation of China. Our sampling was permitted by Lichuan *Metasequoia* Management Office and Xingdoushan National Nature Reserve of China, and was in accordance with the ethical standards of the Regulations of the People's Republic of China on Wild Plants Protection (2017 Amendment PKULAW Version) (CLI.2.304108). During sampling, we only collected leaf tissues from each sampled individual and did not cause severe damage to wild populations of *Metasequoia glyptostroboides*.

References

1. LePage BA, Yang H, Matsumoto M (2005) The evolution and biogeographic history of *Metasequoia*. In: LePage BA, Williams CJ, Yang H (eds) The geobiology and ecology of Metasequoia. Springer, Heidelberg.
2. Li YY, Chen XY, Zhang X, Wu TY, Lu HP, Cai YW (2005) Genetic differences between wild and artificial populations of *Metasequoia glyptostroboides*: implications for species recovery. Conserv Biol 19:224–231. https://doi.org/10.1111/j.1523-1739.2005.00025.x

3. Verissimo D, MacMillan DC, Smith RJ (2011) Toward a systematic approach for identifying conservation flagships. Conserv Lett 4:1–8. https://doi.org/10.1111/j.1755-63X.2010.00151.x
4. Santarém F, Pereira P, Saarinen J, Brito JC (2019) New method to identify and map flagship fleets for promoting conservation and ecotourism. Biol Conserv 229:113–124. https://doi.org/10.1016/j.biocon.2018.10.017
5. Li YY, Ysang EPK, Cui MY, Chen XY (2012) Too early to call it success: an evaluation of the natural regeneration of the endangered *Metasequoia glyptostroboides*. Biol Conserv 150:1–4. https://doi.org/10.1016/j.biocon.2012.02.020
6. Farjon A (2010) A handbook of the world's conifers. Brill, Leiden
7. Spielman D, Brook BW, Frankham R (2004) Most species are not driven to extinction before genetic factors impact them. Proc Natl Acad Sci USA 101:15261–15264. https://doi.org/10.1073/pnas.0403809101
8. Cui MY, Yu S, Liu M, Li YY (2010) Isolation and characterization of polymorphic microsatellite markers in *Metasequoia glyptostroboides* (Taxodiaceae). Conserv Genet Resour 2:19–21. https://doi.org/10.1007/s12686-009-9132-6
9. Jin YQ, Bi QW, Guan WB, Mao JF (2015) Development of 23 novel polymorphic EST-SSR markers for the endangered relict conifer *Metasequoia glyptostroboides*. Appl Plant Sci 3:1500038. https://doi.org/10.3732/apps.1500038
10. Fan XX, Shen L, Zhang X, Chen XY, Fu CX (2004) Assessing genetic diversity of *Ginkgo biloba* L. (Ginkgoaceae) populations from China by RAPD markers. Biochem Genet 42:269–278. https://doi.org/10.1023/B:BIGI.0000034431.15308.57
11. Magoct T, Salzberg SL (2011) FLASH: fast length adjustment of short reads to improve genome assemblies. Bioinformatics 27:2957–2963. https://doi.org/10.1093/bioinformatics/btr507
12. Thiel T, Michalek W, Varshney R, Graner A (2003) Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). Theor Appl Genet 106:411–422. https://doi.org/10.1007/s00122-002-1031-0
13. UCLUST version 1.2.22q. https://www.drive5.com/uclust/downloads1_2_22q.html. Accessed 20 July 2017
14. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG (2012) Primer 3—new capabilities and interfaces. Nucleic Acids Res 40:e115. https://doi.org/10.1093/nar/gks596
15. Schuelke M (2000) An economic method for the fluorescent labeling of PCR fragments. Nat Biotechnol 18:233–234. https://doi.org/10.1038/72708
16. Miller MP (1997) Tools for population genetic analyses (TFPGA) 1.3: a Windows program for the analysis of allozyme and molecular population genetic data. https://www.marksgeneticsoftware.net/tfpga.htm. Accessed 20 Jan 2019
17. Goudet J (1995) FSTAT (version 1.2): a computer program to calculate F-statistics. J Hered 86:485–486
18. Rousset F (2008) GENEPOP'007: a complete re-implementation of the GENEPOP software for Windows and Linux. Mol Ecol Resour 8:103–106. https://doi.org/10.1111/j.1471-8286.2007.01931.x
19. van Oosterhout C, Hutchinson WF, Wills DP, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Mol Ecol Notes 4:535–538. https://doi.org/10.1111/j.1471-8286.2004.00684.x
20. Mao K, Milne RI, Zhang L, Peng Y, Liu J, Thomas P, Mill RR, Renner SS (2012) Distribution of living Cupressaceae reflects the breakup of Pangea. Proc Natl Acad Sci USA 109:7793–7798. https://doi.org/10.1073/pnas.1114319109