Conservation Genet Resour (2010) 2:195–199
DOI 10.1007/s12686-009-9133-5

TECHNICAL NOTE

Isolation and characterization of new microsatellite marker in *Taxus baccata* L.

Parvaneh Mahmoodi · S. Mojtaba Khayam Nekouei · Mohsen Mardi · S. Mostafa Pirseyedi · Mohammad Reza Ghaffari · Mahmood Ramroodi · Barat Ali Siahsar

Received: 10 October 2009 / Accepted: 20 October 2009 / Published online: 6 February 2010

Abstract *Taxus baccata* L. (English yew) is one of the most important medicinal tree species globally. It is well-known for its Taxol content. Here, we report the isolation and characterisation of 31 new polymorphic microsatellite loci from a repeat-enriched genomic library of *T. baccata* L. The genetic diversity of these loci was assessed in 48 individual samples of *T. baccata* L. All loci were variable: the number of polymorphic alleles per locus ranged from 2 to 9 (average 4.45). The observed and expected heterozygosities ranged from 0.15 to 1 and from 0.14 to 0.83, respectively. The loci were informative with polymorphic information content values that ranged from 0.21 to 0.82 (average 0.55). Nineteen of the 31 loci conformed to Hardy–Weinberg expectations. The loci identified in this study should provide useful tools to study the population structure and genetic diversity of *T. baccata* L. and promote its management and conservation.

Keywords SSR marker · Microsatellite · *Taxus baccata* L.

Taxus baccata L. (English yew) is one of the most important medicinal tree species globally (Brambilla et al. 2008). It is well known for its content of Taxol, which is an effective natural agent against ovarian and breast tumors as well as other types of cancer (Brambilla et al. 2008; Emami et al. 2007). *T. baccata* L. is a conifer native to Western, Central, and Southern Europe, Northwestern of Africa, Southwestern Asia, and Northern and Northwestern Iran (Lickl and Heinze 2001; Shanjani 2001). English yews have been found in the Hyracanian forest in Iran since the Tertiary period. This forest covers the northern expositions of the Elburs Mountains, where the unique ecotype has allowed the development of a rich diversity of tree species (Lessani 1999). Due to climate change and high levels of deforestation in Iran, this medicinal species, which is included in the "Red Books of threatened plants" is facing extinction and is only found in protected areas (Makhdoum 2008). The conservation of this endangered species depends on a detailed investigation of its genetic structure and diversity.

Molecular methods are being used increasingly to study genetic structure and diversity in plants. A few studies have been reported for yew that have used molecular markers such as internal transcribed spacers (ITSs; Spanu et al. 2006), random amplified polymorphic DNA (RAPD; Collins et al. 2003; Hilfiker et al. 2004) and allozymes (Lewandowski et al. 1995). However, the endophytic fungi that are found within *T. baccata* L. (Caruso et al. 2000) limit the study of genetic diversity using nonspecific markers. Simple sequence repeats (SSRs) are the most reliable, reproducible, and specific molecular markers for studies of genetic diversity. To date, only a limited number of SSRs (7 markers) are available in *T. baccata* L.; these have been reported by Dubreuil et al. (2008). The results of the present study should provide relevant new polymorphic SSR markers for the genetic conservation and management of *T. baccata* L.

Genomic DNA was extracted from a callus sample collected from the Pone Aram forest in Northern Iran using a

P. Mahmoodi · S. M. Khayam Nekouei · M. Mardi (✉) · S. M. Pirseyedi · M. R. Ghaffari
Department of Genomics, Agricultural Biotechnology Research Institute of Iran (ABRII), Mahdasht Road, Karaj, Iran
e-mail: mardi@abrii.ac.ir

M. Ramroodi · B. A. Siahsar
Department of Agronomy and Plant Breeding, University of Zabol, Zabol, Iran

Table 1 Characterization of 31 polymorphic microsatellite loci in English yew (*Taxus baccata* L.)

Locus name	Primer sequences (5′→3′)	Repeat motif	Allele no.	Allele size (bp)	H_o	H_e	HWE	PIC	GenBank Accession No.
ABRII-TB1	ACATTGGAACAAAGGAGTTG AGTTTG TGTCAGGTTCCCTA	(CT)17	3	390–410	0.6111	0.4937	0.633950	0.5003	GQ996553
ABRII-TB2	TCTCTCTCTCACTCACAA TAGTTTGTGTCAGGTTCCCT	(AC)3-(AC)6	4	382–396	0.4737	0.4438	0.945880	0.4600	GQ996553
ABRII-TB10	CCACCCAGATAAGACAGT AGTGAGTTCTTTGTTTGCAC	(CCAAAG)2	3	158–210	0.3684	0.3172	0.846378	0.3428	GQ996554
ABRII-TB13	TACTACATAGGAGGGGAG CACATTGACCTCGAAAGTC	(GT)6	3	204–210	0.7895	0.6415	0.012510	0.5996	GQ996555
ABRII-TB14	GACTATCCTCAACTGTCAAA CTATCAATTGTGACACACAG	(GT)13	7	120–146	1.0000	0.7744	0.555732	0.7285	GQ996556
ABRII-TB15	CTTGATTCTCATCTCTTCTG GTCTCCTCTTACTCCAAACT	(TC)15-(TG)13-(TG)5	9	153–191	0.7222	0.8397	0.138254	0.8251	GU002549
ABRII-TB22	ATGACTCACTAGATGATTGG GTGGGTATTCAGAACTTGTA	(TTGG)3	5	188–214	0.6316	0.5733	0.000014	0.5509	GU002550
ABRII-TB26	TACATAGATACCCAAAGTGC ATCTACACATATGCTCTTGC	(TGA)4	4	250–284	0.5882	0.5080	0.166965	0.5708	GU045477
ABRII-TB27	CAAGATATTCCAGTGTAGGA TAGTGTAGGCTCTGTTTTGT	(CTAATC)2	3	204–250	0.7895	0.5761	0.025840	0.5245	GU002551
ABRII-TB28	AAAGAGAATCCCTCCTTAC GATGGTATTGTACAGGAATC	(ATTTTT)2-(GT)4-(A)13-(C)12	2	230–236	0.4118	0.3369	0.327877	0.4414	GU045478
ABRII-TB30	CTAATCAGTTTGAGAAGGTG ACAAAGAGTCACTAAGAACG	(ATTTTT)2	3	184–240	0.4444	0.3667	0.738829	0.4276	GU002552
ABRII-TB31	GAGATGGATCTTTAGTTTCC TAGAGATCACCACAAACTTC	(CTCCTT)2	6	293–359	0.8947	0.6600	0.704773	0.6461	GU002553
ABRII-TB32	GACTACTTGCACACTCAGAC TGGAGAACAAGTTCACTCT	(CTCTTC)2	5	180–218	0.6842	0.5292	0.911504	0.5206	GU002554
ABRII-TB33	CAGCTTACATAAGCAAGAAC ATATGTGCACTCAGAGAAAG	(TGTATC)2	2	82–92	0.8500	0.5013	0.001394	0.3693	GU002555
ABRII-TB34	GACTGGAAAAACCTTCAC TTATGGAGGTAGACACTCTG	(A)13	9	204–286	0.7500	0.6551	0.000007	0.6023	GU002556
ABRII-TB35	AAGATCTTGATAGATAGGGG TATAGACCTTGTGTTTGGAC	(GAACTC)2-(CAAGTC)2	3	216–238	0.9500	0.6000	0.000055	0.4948	GU002557
ABRII-TB36	GAATGAGAAACCAACACC GATAAAGCAACTGTCTCAAG	(G)15	2	190–220	0.3684	0.3087	0.367263	0.3263	GU002558

Table 1 continued

Locus name	Primer sequences (5′→3′)	Repeat motif	Allele no.	Allele size (bp)	H_o	H_e	HWE	PIC	GenBank Accession No.
ABRII-TB37	CTTGAGACAGTTGCTTTATC CCAGGTTTAGACCAGTAGTA	(A)15-(T)9-(A)9-(G)9-(A)13	3	236–250	0.7500	0.4987	0.083931	0.3899	GU045479
ABRII-TB39	GTTGTGTGTATCAATGTGTG GTCTCTAATGTGTCCCCA	(ATC)4	2	214–217	0.1500	0.1423	0.769866	0.2137	GU045480
ABRII-TB40	AGGGTCTCATATCATTGG CCCAACTACACAATCTAAAC	(ACAA)4-(AC)26	7	130–260	0.7368	0.8307	0.183578	0.8016	GU045481
ABRII-TB41	CTCAAGAGGGCTTCTAAA ATAACACAACCTAGAGAGGG	(G)16	2	120–200	0.5263	0.3983	0.14235	0.3809	GU045482
ABRII-TB42	AAACAAGGAAAGAGTCTCC GTGGTATCGGTCACTAATAA	(T)16	5	180–212	0.6316	0.6245	0.000905	0.6094	GU002559
ABRII-TB45	CTATTGATGAGTCCTGAGAA ATGATAGAGAAGGGAGAAAC	(ATCATA)2	3	110–116	0.7500	0.6808	0.000012	0.5897	GU045483
ABRII-TB46	GTTGGAAAAGCCATAGATAG ATCTACTTCCCTACCACACT	(GGAG)3	2	265–285	0.4737	0.3713	0.206463	0.2879	GU002565
ABRII-TB47	GTAATCTCCAAAAGGGTATC CAAAACGTGATTGGTTC	(AAAGCA)2	3	205–213	0.2143	0.5423	0.011296	0.6147	GU045485
ABRII-TB50	ACAAAGACTATGAGCTATGC GAAAAGAGAATGTTGGGAG	(TC)8-(CT)13	4	270–330	1.0000	0.8317	0.081178	0.7171	GU045484
ABRII-TB53	AGTAAAGTTGGAATAGGTCC TCTAGTCCCAGATACAAAGA	(ACT)5	7	252–310	0.8125	0.8145	0.000000	0.8016	GU002560
ABRII-TB55	AAACCTACAGACAGTACCCT CTCCAAAACAAGGTAGAAG	(GGAGAC)2	4	220–238	0.5789	0.5875	0.094327	0.5708	GU002561
ABRII-TB56	AATGTAGAATGCTGAGTCAC CTATTCCATCACACACACTTT	(TG)6-(CTAGAT)2	7	204–270	0.9500	0.8000	0.027331	0.7478	GU002562
ABRII-TB57	GGGGAAAGTATACAGAAAAG CATGAGAGTCCAACTACATT	(ATTAGT)2	5	204–258	0.8947	0.7511	0.000006	0.7189	GU002563
ABRII-TB58	GTACTATGTGAAATCTTCGG ATGGATAAGCTCACACTAAC	(CTATCC)2-(TCTGTA)2	8	206–230	1.0000	0.7692	0.573944	0.7166	GU002564

H_O observed heterozygosity, H_E expected heterozygosity, *HWE* Hardy–Weinberg equilibrium test ($P < 0.05$), *PIC* polymorphic information content

DNeasy Plant Mini kit (Qiagen, Germany). A genomic library enriched for di- and tri-nucleotides was constructed using the fast isolation by AFLP of sequences containing repeats (FIASCO) protocol (Zane et al. 2002; Xu et al. 2006) with some modifications. A total of 250 ng genomic DNA was digested with *Mse*I (Fermentas, Germany) to give DNA fragments between 200 and 1,000 bp in length. The fragments were ligated to adapters and then amplified in two stages by PCR using *Mse*I primers to give numerous copies of each fragment. The genomic DNA fragments that contained SSRs were captured by hybridisation to biotinylated probes that consisted of di- and tri-nucleotide repeats [$(GC)_{17}$, $(AC)_{17}$, $(CT)_{17}$, $(AT)_{17}$, $(GT)_{17}$, $(ATT)_{10}$, and $(CTT)_{10}$], followed by binding to streptavidin-conjugated magnetic beads (BioMag®; Qiagen, Germany). Three non-stringent and three stringent washes were carried out with separation in a magnetic field. The recovered DNA fragments were amplified for 30 cycles using the *Mse*I primers. The PCR products were cloned into pGEM-T Easy vector (Promega, Germany), and transformed into *Escherichia coli* DH5*a*. Recombinant clones were identified by blue/white screening and restriction analysis (Eco*RI*; Fermentas, Germany). One hundred and seventy-two clones with inserts were purified using a plasmid extraction kit (Core-Bio, Canada) and sequenced (Macrogen Sequencing Service, Korea). Ninety-two of the clones contained microsatellite repeats. It was possible to design primers for 58 of these sequences that were unique, whereas the remaining 34 sequences were too close to the cloning site or were repetitive.

PCR amplification was performed on an ABI thermal cycler in a total volume of 15 µl, which included 20 ng DNA, 1.5 µl 10X PCR buffer, 1.2 µl $MgCl_2$ 25 mM, 0.5 µl dNTP mix 10 mM, 1 µl each primer 10 pmol/µl and 0.5 U Taq DNA polymerase (Fermentas, Germany). The following reaction conditions were used: 5 min at 95°C, followed by ten touchdown cycles of 30 s at 95°C, 45 s at 60°C (1°C lower per cycle) and 40 s at 72°C, and 25 cycles of 30 s at 95°C, 30 s at 50°C and 40 s at 72°C, with a final extension step of 7 min at 72°C. Amplified products were separated on 6% denaturing polyacrylamide gels and visualised by silver staining. A 50-bp DNA ladder (Fermentas, Germany) was used to identify the alleles.

The variability of these markers was analysed in 48 *T. baccata* L. individuals that were sampled at random from Northern and Northwestern Iran. POPGENE 32 (Yeh and Boyle 1997) was used to calculate the observed and expected heterozygosities and to evaluate deviation from Hardy–Weinberg equilibrium (HWE) and linkage disequilibrium between pairs of loci. All results were adjusted for multiple simultaneous comparisons using a sequential Bonferroni correction (Rice 1989). Polymorphic information content (PIC) was estimated using CERVUS v.2.0 (Marshall et al. 1998; Slate et al. 2000).

Out of 58 microsatellite loci isolated in *T. baccata* L., 31 (53%) were polymorphic (Table 1), whereas the remaining 27 loci (47%) showed monomorphic bands (20), multi-banding patterns (5) or no amplification (2). The monomorphic SSR loci should be tested in a larger sample population. The number of polymorphic alleles per locus ranged from 2 to 9 (average 4.45). The observed and expected heterozygosities ranged from 0.15 to 1 and from 0.14 to 0.83, respectively (Table 1). Nineteen of the 31 loci conformed to HWE. The observed departures from HWE were likely to be due to the Wahlund effect, pooling of samples from distant regions, and limited sample size. Significant genotypic linkage disequilibrium was found between some loci ($P > 0.05$). There was no evidence for allelic dropout or null alleles, as shown using the software MICRO-CHEKER (Van Oosterhaut et al. 2004). The loci were informative with PIC values that ranged from 0.21 to 0.82 (average 0.55).

The new polymorphic microsatellite markers identified in this study should provide useful tools to study the population structure and genetic diversity of *T. baccata* L. and promote its management and conservation.

References

Brambilla L, Romanelli A, Bellinvia M (2008) Weekly paclitaxel for advanced aggressive classic Kaposi sarcoma: experience in 17 cases. Br J Dermatol 158:1339–1344. doi:10.1111/J.1365-2133.2008.08517.x

Caruso M, Colombo AL, Fedeli L, Pavesi A, Quaroni S, Saracchi M, Bentrella G (2000) Isolation of endophytic fungi and actinomycetes taxane producers. Ann Microbiol 50:3–13

Collins D, Mill RR, Moller M (2003) Species separation of *Taxus baccata*, *T.* canadensis and *T. cuspidata* (Taxaceae) and origins of their reputed hybrids inferred from RAPD and cpDNA data. Am J Bot 90:175–182. doi:10.3732/ajb.90.2.175

Dubreuil M, Sebastiani F, Mayol M, Gonza′lez-Martı′nez SC, Riba M (2008) Isolation and characterization of polymorphic nuclear microsatellite loci in *Taxus baccata* L. Conserv Genet 9:1665–1668. doi:10.1007/s10592-008-9515-3

Emami SA, Asili A, Mohagheghi Z, Hassanzadeh MK (2007) Antioxidant activity of leaves and fruits of Iranian conifers. Evid Based Complement Alternat Med 4:313–319. doi:10.1093/ecam/nem011

Hilfiker K, Holderegger R, Rotach P, Gugerli F (2004) Dynamics of genetic variation in *Taxus baccata*: local versus regional perspectives. Can J Bot 82:219–227. doi:10.1139/b03-136

Lessani MR (1999) Yew, 1st edn. Research institute of forests and rangelands of Iran, Tehran, pp 7–13

Lewandowski A, Burczyk J, Mejnartowicz L (1995) Genetic structure of English yew (*Taxus baccata* L.) in the Wierzchlas Reserve: implications for genetic conservation. For Ecol Manag 73:221–227

Lickl E, Heinze B (2001) Eiben im Elburs–Ein kleines Vorkommen im Wald von Kheyrudkenar'. Der Eibenfreund 8:90–91

Makhdoum MF (2008) Management of protected areas and conservation of biodiversity in Iran. Int J Environ Stud 65:583–585

Marshall TC, Slate J, Kruuk LEB, Pemberton JM (1998) Statistical confidence for likelihood-based paternity inference in natural populations. Mol Ecol 7:639–655. doi:10.1046/j.1365-294x.1998.00374.x

Rice WR (1989) Analyzing tables of statistical tests. Evol Int J Org Evol 43:223–225. doi:10.2307/2409177

Shanjani P (2001) Quantitative and qualitativt of yew peroxidases in Arasbaran and Gorgan forest. Iran j bio 10:31–39

Slate J, Marshall TC, Pemberton JM (2000) A retrospective assessment of the accuracy of the paternity inference program CERVUS. Mol Ecol 9:801–808. doi:10.1046/j.1365-294x.2000.00930.x

Spanu I, Emiliani G, Paffetti D, Giannini R (2006) Proceedings of the 50th Italian society of agricultural genetics annual congress Ischia, Italy

Van Oosterhaut C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Mol Ecol Notes 4:535–538. doi:10.1111/j.1471-8286.2004.00684.x

Xu YQ, Wang Y, Li ZZ, Bao ZX, Zhou JF, Huang HW (2006) Characterization of polymorphic microsatellite loci in a traditional Chinese medicinal plant, *Gastrodia elata*. Mol Ecol Notes 6:316–318

Yeh FC, Boyle TJB (1997) Population genetic analysis of co-dominant and dominant markers and quantitative traits. Belg J Bot 129:157

Zane L, Bargelloni L, Patarnello T (2002) Strategies for microsatellite isolation: a review. Mol Ecol 11:1–16. doi:10.1046/j.0962-1083.2001.01418.x