Isolation and Characterization of 22 EST-SSR Markers for the Genus Thujopsis (Cupressaceae)

Authors: Sato, Miyako, Hasegawa, Yoichi, Mishima, Kentaro, and Takata, Katsuhiko

Source: Applications in Plant Sciences, 3(2)

Published By: Botanical Society of America

URL: https://doi.org/10.3732/apps.1400101

Applications in Plant Sciences 2015 3(2): 1400101

PRIMER NOTE

ISOLATION AND CHARACTERIZATION OF 22 EST-SSR MARKERS FOR THE GENUS *THUJOPSIS* (CUPRESSACEAE)[1]

MIYAKO SATO[2], YOICHI HASEGAWA[2], KENTARO MISHIMA[3], AND KATSUHIKO TAKATA[2,4]

[2]Institute of Wood Technology, Akita Prefectural University, 11-1 Kaieisaka, Noshiro, Akita 016-0876, Japan; and [3]Forestry and Forest Products Research Institute, Forest Tree Breeding Center, 3809-1 Ishi, Juo, Hitachi, Ibraraki 319-1301, Japan

- *Premise of the study:* Expressed sequence tag–simple sequence repeat (EST-SSR) markers were developed from *Thujopsis dolabrata* var. *hondae* (Cupressaceae) using Illumina sequencing to investigate the genetic diversity and population structure of the genus *Thujopsis*.
- *Methods and Results:* Twenty-two primer pairs were developed from ESTs of *T. dolabrata* var. *hondae*. The primers amplified di- and trinucleotide repeat-containing sequences. Polymorphisms were assessed in 81 individuals from two *T. dolabrata* var. *hondae* populations and one *T. dolabrata* population. The number of alleles ranged from one to 17 per locus. The observed and expected heterozygosities ranged from 0.000 to 1.000 and from 0.000 to 0.926, respectively.
- *Conclusions:* These new EST-SSR markers will be useful in analyses of the genetic diversity and population structure of the genus *Thujopsis*.

Key words: Cupressaceae; expressed sequence tag; microsatellite; next-generation sequencing; *Thujopsis*.

The Cupressaceae clade has the broadest diversity of habitats and morphologies of any conifer family, and the genus *Thujopsis*, which belongs to the Cupressaceae, is considered to be one of the early diverging genera (Pittermann et al., 2012). *Thujopsis* is native to Japan and includes one species (*T. dolabrata* Siebold & Zucc.) and one northern variety (*T. dolabrata* var. *hondae* Makino) (Hayashi, 1960). The two varieties of *Thujopsis* are distinguished by variation in their cone morphology. The varieties also differ in their geographic ranges, although their distributions overlap in the central region of the Japanese archipelago (Hayashi, 1960). *Thujopsis dolabrata* has antimicrobial properties and demonstrates a strong antifeedant effect on termites (Inamori et al., 2006). Moreover, its essential oil is used as an antibacterial agent and as an aromatic substance. Thus, *T. dolabrata* is an important tree species in Japanese forestry and future forest tree breeding because of its superior wood properties. However, there have been few reports about genetic differences between the two varieties.

Recently, expressed sequence tag (EST)–based markers have been used increasingly in studies of genetic diversity and population structure in tree species (e.g., Fageria and Rajora, 2013). EST-based markers are less susceptible to null alleles than anonymous simple sequence repeats (SSRs). Moreover, because ESTs correspond to coding DNA, the flanking sequences of EST-SSRs are located in well-conserved regions across phylogenetically related species (Uchiyama et al., 2013). Nuclear SSR markers have been developed for *T. dolabrata* var. *hondae* (Mishima et al., 2012), and population genetics and phylogeographic analyses of this variety were performed using these markers (Higuchi et al., 2012). However, the nuclear SSR markers have not been tested on *T. dolabrata.* An analysis of the genetic diversity and population structure for *Thujopsis* plants in the Japanese archipelago is urgently required. Therefore, it is necessary to obtain molecular markers with high transferability within the genus *Thujopsis* that also exhibit a low frequency of null alleles. In this paper, we describe the development and characterization of 22 EST-SSR markers for the genus *Thujopsis* from expressed sequence data of *T. dolabrata* var. *hondae*.

METHODS AND RESULTS

One *T. dolabrata* var. *hondae* individual (voucher no. TF-K12-001) was used for the RNA sequencing experiment. Leaves and cambiums were sampled from a population in the Aomori Prefecture (Owani: 40°27′21″N, 140°34′08″E) and were immediately frozen in liquid nitrogen and stored at −30°C. The cetyltrimethylammonium bromide (CTAB) method was used to extract the total RNAs (Chang et al., 1993). A TruSeq RNA Sample Prep Kit (Illumina, San Diego, California, USA) was used to create the RNA sequencing library, according to the manufacturer's protocol. A HiSeq 1000 (Illumina) was used to sequence the library with 2 × 101-bp paired-end reads. More than 237.24 million paired-end raw reads were obtained. After removal of low-quality reads, 233.92 million clean reads remained (accession no.: DRA002435). Using the short reads assembly programs Velvet 1.2.08 (Zerbino et al., 2009) and Oases 0.2.08 (Schulz et al., 2012), the clean reads were assembled de novo into 76,377 contigs and 41,182 unigenes, from 100 to 14,834 bp, with a mean length of 1525 bp. MSATCOMMANDER 1.0.8 (Faircloth, 2008) was used to screen for microsatellite loci containing di- and trinucleotides. Primer3 (Rozen and Skaletsky, 1999) was then used to design PCR primers. The minimum number of repeats was set as nine and 11 for di- and trinucleotide repeats, respectively. All

[1]Manuscript received 20 October 2014; revision accepted 2 December 2014.

The authors thank S. Akiyama for assistance with the DNA extraction. This work was financially supported by a Grant-in-Aid for Scientific Research (B) (No. 25282069) of the Japan Society for Promotion of Science (JSPS).

[4]Author for correspondence: katsu@iwt.akita-pu.ac.jp

doi:10.3732/apps.1400101

TABLE 1. Characteristics of 22 EST-SSR loci for use in the genus *Thujopsis*.

Locus	Primer sequences (5′–3′)	Repeat motif	T_a (°C)	Allele size range (bp)	DDBJ accession no. 1[a]	DDBJ accession no. 2[b]
Tdest1	F: GCCTCCCTCGCGCCATCAGGATTTTCTGACAGGCTTTGTTCTC R: GTTTCTTAATTCCCAAGAGTGCTTATGAGTTC	$(CT)_{11}$	60	137–173	LC010288	LC010310
Tdest3	F: GCCTCCCTCGCGCCATCAGCGGCCCAGGTTTCTGTACTC R: GTTTCTTGCCCATTAAAGTCGGGTATTG	$(AT)_{11}$	60	169–186	LC010289	LC010311
Tdest11	F: GCCTCCCTCGCGCCATCAGTGGGATACATACTGCATTTGTTAGG R: GTTTCTTCTCCCCAAGCAAGTCACCAC	$(AT)_{12}$	60	138–161	LC010290	LC010312
Tdest14	F: GCCTCCCTCGCGCCATCAGCAGTAGACAATTTCTGCAAATCACC R: GTTTCTTTCCCTTTTGTTGGCATTATAGG	$(AG)_{12}$	60	157–185	LC010291	LC010313
Tdest17	F: GCCTCCCTCGCGCCATCAGGCTTTTGATGTCCGCTATATCCTC R: GTTTCTTGGAGATTCCAATGTTTGTCATGC	$(AG)_{12}$	60	159–168	LC010292	LC010314
Tdest21	F: GCCTCCCTCGCGCCATCAGGTCCATCCATTCTCACTCCAAAG R: GTTTCTTAGCAGACCCTATTTCACAGCATC	$(AG)_{13}$	60	231–274	LC010293	LC010315
Tdest24	F: GCCTCCCTCGCGCCATCAGATACCATACAGCTTTCAGCCAG R: GTTTCTTGCAGAACAAACGAATCAATGAGAG	$(AT)_{15}$	60	243–267	LC010294	LC010316
Tdest29	F: GCCTCCCTCGCGCCATCAGAAACGACTCTGCTGGATTTCAC R: GTTTCTTTTCCGCTCTTGATTTTCTCTCC	$(AC)_{16}$	60	219–246	LC010295	LC010317
Tdest35	F: GCCTCCCTCGCGCCATCAGAAGCTATTGACCCTTCTCAGGATAC R: GTTTCTTCCATGTTGAATTGTTCCCTTTC	$(CT)_{15}$	60	194–230	LC010296	LC010318
Tdest37	F: GCCTCCCTCGCGCCATCAGCCAAGCGACAGAAAACCATTC R: GTTTCTTTCAGTCTCTTCCTCCTCCTCCTC	$(ATC)_9$	60	164–176	LC010297	LC010319
Tdest38	F: GCCTCCCTCGCGCCATCAGTGACCATTCCTCCTCCTCCTC R: GTTTCTTCATGTTTGCAGTTGAGAGAAGACC	$(ACC)_9$	60	117–134	LC010298	LC010320
Tdest39	F: GCCTCCCTCGCGCCATCAGGCAGCACAGGAGAAGAAAGATG R: GTTTCTTACAACAGCCACAACGTGTCC	$(GCT)_9$	60	153–165	LC010299	LC010321
Tdest42	F: GCCTCCCTCGCGCCATCAGCTCCCTATCCCAACACCAACAC R: GTTTCTTTGCCTACCTATCCTTCTTCTTCTCC	$(ACC)_9$	60	226–255	LC010300	LC010322
Tdest43	F: GCCTCCCTCGCGCCATCAGGGTCCAATGCAGGTAATACAAGAAG R: GTTTCTTTCCCCGCCAAGATACTCAAC	$(CGG)_9$	60	137–153	LC010301	LC010323
Tdest44	F: GCCTCCCTCGCGCCATCAGTTTGGTGGTGGAGGTGGTG R: GTTTCTTCGCTTATGCCAAGCAGTCATC	$(GAT)_9$	60	134–137	LC010302	LC010324
Tdest45	F: GCCTCCCTCGCGCCATCAGTGAGGGTGGTGAGACAATTC R: GTTTCTTCAAGATTTGGAACTCCTGCAAC	$(GGT)_{12}$	60	211–236	LC010303	LC010325
Tdest49	F: GCCTCCCTCGCGCCATCAGGTGCCCTCAAAGTTACAGCAGTC R: GTTTCTTGCAATCACCTCATCCTCACTTC	$(GAT)_{10}$	60	233–248	LC010304	LC010326
Tdest52	F: GCCTCCCTCGCGCCATCAGTTCAGGAAGGCCAAGGAGAG R: GTTTCTTGATCCTCCTGCATCATTTTGTTC	$(GGT)_{11}$	60	239–251	LC010305	LC010327
Tdest53	F: GCCTCCCTCGCGCCATCAGCCAAAGCCCTTCCAGTAACATC R: GTTTCTTGATGGAATGAGTGAATCTCAGGAAC	$(CTT)_{13}$	60	244–284	LC010306	LC010328
Tdest54	F: GCCTCCCTCGCGCCATCAGCCCTGTATTATTCTCAACATCATCG R: GTTTCTTGGGATTCAGACAAGGGCAAG	$(CTT)_{11}$	60	182–203	LC010307	LC010329
Tdest56	F: GCCTCCCTCGCGCCATCAGCATTGCCCTTTGGAATATAGGATC R: GTTTCTTGTTGCCCATCTGCTCTTCTTC	$(AAG)_9$	60	153–165	LC010308	LC010330
Tdest58	F: GCCTCCCTCGCGCCATCAGCTGAACGGCGCCCTAATCTC R: GTTTCTTGCCCACTCCTCAAATCCAAC	$(AAG)_{13}$	60	151–180	LC010309	LC010331

Note: DDBJ = DNA Data Bank of Japan; T_a = annealing temperature.
[a] 1 = *T. dolabrata* var. *hondae*.
[b] 2 = *T. dolabrata.*

forward primers were fluorescently labeled using FAM (carboxyfluorescein) with the 454A adapter sequence (5′-GCCTCCCTCGCGCCATCAG-3′) at the 5′-end (Blacket et al., 2012). Additionally, all reverse primers were attached to a PIG-tail sequence (5′-GTTTCTT-3′) at the 5′-end of the sequence (Brownstein et al., 1996).

The CTAB method (Murray and Thompson, 1980) was used to extract genomic DNAs from needles sampled from 32 *T. dolabrata* plants in the Nagano Prefecture (Kiso: 35°43′38″N, 137°37′15″E) and 49 *T. dolabrata* var. *hondae* needles were sampled from two populations in the Aomori Prefecture (Owani: 40°27′21″N, 140°34′08″E; and Okoppe: 41°28′42″N, 140°57′10″E). These populations were located in the typical range of each variety. To confirm PCR amplification, we used eight DNA samples from four individuals per variety. PCR was performed in a final volume of 10 µL, containing 5 µL of 2× GoTaq Hot Start Colorless Master Mix (Promega Corporation, Madison, Wisconsin, USA), 1 µM of each primer, and 60 ng of template DNA. Reactions were performed with initial denaturation at 94°C for 2 min; followed by 40 cycles of 94°C for 30 s, 60°C for 30 s, and 72°C for 30 s; and then 72°C for 5 min using an ABI 9700 (Applied Biosystems, Foster City, California, USA). The PCR products were separated electrophoretically on 1% agarose gels and stained with GelRed Nucleic Acid (Nacalai Tesque, Kyoto, Japan). To confirm the presence of SSRs, we sequenced the PCR products from two DNA samples using an ABI PRISM 3100 Genetic Analyzer (Applied Biosystems). For fragment analysis, the PCR conditions were modified. The concentration of the primers were 0.15, 0.5, and 0.2 µM for the forward primer, reverse primer, and 454A-FAM primer, respectively, and the number of PCR cycles was 30. PCR products for 81 DNA samples from two varieties were electrophoresed on an ABI PRISM 3100 Genetic Analyzer (Applied Biosystems) and Geneious 7.0.4 software (Biomatters Ltd., Auckland, New Zealand; http://www.geneious.com/) was used to assess the fragment sizes. To characterize the microsatellite loci, CERVUS 3.0 software (Kalinowski et al., 2007) was used to calculate the number of alleles (A), the observed heterozygosity (H_o), the expected heterozygosity (H_e), and the frequency of null alleles (r). GenAlEx 6.501 (Peakall and Smouse, 2012) assessed the probability of identity (P_{ID}).

TABLE 2. The results of final primer screening on samples from three populations of the genus *Thujopsis.*

	Thujopsis dolabrata var. *hondae*										*Thujopsis dolabrata*					
	Okoppe (N = 27)					Owani (N = 22)					Kiso (N = 32)					
Locus	A	H_o	H_e	P_{ID}	r	A	H_o	H_e	P_{ID}	r	A	H_o	H_e	P_{ID}	r	Total A
Tdest1	12	1.000	0.897	0.026	–0.066	11	0.955	0.883	0.034	–0.052	12	0.938	0.898	0.024	–0.030	16
Tdest3	6	0.444	0.484	0.305	–0.002	2	0.409	0.333	0.508	–0.113	9	0.813	0.830	0.056	–0.001	10
Tdest11	12	0.963	0.860	0.040	–0.071	9	0.682	0.809	0.070	0.086	9	0.719	0.810	0.069	0.056	14
Tdest14	11	0.852	0.885	0.031	0.009	12	0.864	0.884	0.032	0.001	11	0.875	0.880	0.032	–0.005	14
Tdest17	8	0.778	0.747	0.102	–0.028	8	0.545	0.678	0.152	0.100	7	0.625	0.579	0.229	–0.073	8
Tdest21	17	0.852	0.903	0.023	0.023	11	0.818	0.770	0.080	–0.060	15	0.969	0.926	0.015	–0.031	21
Tdest24	8	0.852	0.832	0.058	–0.024	9	0.727	0.846	0.050	0.064	6	0.719	0.740	0.116	0.008	10
Tdest29	5	0.593	0.622	0.193	0.013	5	0.545	0.540	0.278	–0.008	1	0.000	0.000	1.000	ND	6
Tdest35	17	0.852	0.922	0.016	0.035	15	0.909	0.900	0.025	–0.024	13	0.813	0.850	0.044	0.010	23
Tdest37	4	0.333	0.372	0.438	0.062	4	0.409	0.411	0.398	–0.029	3	0.281	0.294	0.539	0.009	5
Tdest38	3	0.593	0.639	0.215	0.015	5	0.727	0.743	0.125	–0.001	4	0.625	0.615	0.234	–0.016	6
Tdest39	4	0.815	0.613	0.226	–0.159	3	0.682	0.654	0.206	–0.027	2	0.125	0.119	0.786	–0.023	4
Tdest42	6	0.519	0.558	0.277	0.011	5	0.636	0.630	0.210	–0.030	7	0.563	0.652	0.173	0.053	10
Tdest43	6	0.778	0.806	0.077	0.011	8	0.909	0.768	0.097	–0.100	4	0.406	0.455	0.342	0.080	8
Tdest44	1	0.000	0.000	1.000	ND	1	0.000	0.000	1.000	ND	2	0.031	0.031	0.940	–0.002	2
Tdest45	7	0.704	0.643	0.206	–0.058	4	0.545	0.579	0.249	0.024	2	0.094	0.091	0.833	–0.015	7
Tdest49	2	0.148	0.140	0.754	–0.030	3	0.091	0.090	0.834	–0.014	4	0.563	0.495	0.319	–0.082	5
Tdest52	2	0.926	0.507	0.376	–0.301	2	1.000	0.512	0.375	–0.333	2	0.281	0.246	0.604	–0.072	2
Tdest53	12	0.926	0.897	0.025	–0.025	10	0.818	0.885	0.033	0.030	9	0.719	0.783	0.080	0.039	14
Tdest54	3	0.370	0.545	0.325	0.169	3	0.455	0.563	0.274	0.078	4	0.094	0.632	0.212	0.743	5
Tdest56	3	0.778	0.645	0.212	–0.104	4	0.500	0.506	0.293	–0.014	4	0.531	0.563	0.257	0.018	4
Tdest58	5	0.519	0.524	0.320	–0.005	5	0.364	0.499	0.307	0.129	3	0.281	0.298	0.528	0.011	9

Note: A = number of alleles; H_e = expected heterozygosity; H_o = observed heterozygosity; N = sample size; ND = not determined; P_{ID} = probability of identity; r = null allele frequency.

Primer3 designed 58 primer pairs, of which 33 showed amplification for all eight samples in both varieties. For 29 of the 33 primer pairs, the SSR sequence in the PCR product was identified for both varieties. Finally, 22 of the 29 amplified primers showed clear fragment patterns, thus they were selected as the developed markers. These primer pairs show different fragment sizes and the same annealing temperature (Table 1). All 22 loci were polymorphic in both varieties. The observed number of alleles per population ranged from one to 17, H_o ranged from 0.000 to 1.000, and H_e ranged from 0.000 to 0.926. The P_{ID} ranged from 0.015 to 1.000 (Table 2). One of the loci (Tdest54) showed high r values relative to the other loci. These 22 EST-SSR markers had lower r and higher P_{ID} values compared with reported SSR markers developed from the genomic DNA of *T. dolabrata* var. *hondae* (Mishima et al., 2012).

CONCLUSIONS

In this study, we developed 22 EST-SSR markers for the two varieties of *Thujopsis*, using expressed sequence data of *T. dolabrata* var. *hondae*. These markers have two advantages: high ability to detect genetic polymorphisms in *Thujopsis* varieties, and a low null allele frequency. Accordingly, these EST-SSR markers will be valuable tools for investigating the genetic diversity and population structure of the genus *Thujopsis*. Moreover, these markers will help to advance breeding programs for species in the genus *Thujopsis.*

LITERATURE CITED

BLACKET, M. J., C. ROBIN, R. T. GOOD, S. F. LEE, AND A. D. MILLER. 2012. Universal primers for fluorescent labeling of PCR fragments—An efficient and cost-effective approach to genotyping by fluorescence. *Molecular Ecology Resources* 12: 456–463.

BROWNSTEIN, M. J., J. D. CARPTEN, AND J. R. SMITH. 1996. Modulation of non-templated nucleotide addition by Taq DNA polymerase: Primer modifications that facilitate genotyping. *BioTechniques* 20: 1004–1010.

CHANG, S., J. PURYEAR, AND J. CAIRNEY. 1993. A simple and efficient method for isolating RNA from pine trees. *Plant Molecular Biology Reporter* 11: 113–116.

FAGERIA, M. S., AND O. P. RAJORA. 2013. Effects of harvesting of increasing intensities on genetic diversity and population structure of white spruce. *Evolutionary Applications* 6: 778–794.

FAIRCLOTH, B. C. 2008. MSATCOMMANDER: Detection of microsatellite repeat arrays and automated, locus-specific primer design. *Molecular Ecology Resources* 8: 92–94.

HAYASHI, Y. 1960. Taxonomical and phytogeographical study of Japanese conifers. Norin-Shuppan, Tokyo, Japan.

HIGUCHI, Y., A. MATSUMOTO, Y. MORIGUCHI, K. MISHIMA, K. TANAKA, Y. YADA, K. TAKATA, ET AL. 2012. Genetic diversity and structure using microsatellite markers in natural and breeding populations of *Thujopsis dolabrata* var. *hondae. Journal of the Japanese Forest Society* 94: 247–251.

INAMORI, Y., Y. MORITA, Y. SAKAGAMI, T. OKABE, AND N. ISHIDA. 2006. The excellence of Aomori Hiba (Hinokiasunaro) in its use as building materials of Buddhist temples and Shinto shrines. *Biocontrol Science* 11: 49–54.

KALINOWSKI, S. T., M. L. TAPER, AND T. C. MARSHALL. 2007. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Molecular Ecology* 16: 1099–1106.

MISHIMA, K., T. HIRAO, A. WATANABE, AND K. TAKATA. 2012. Isolation and characterization of microsatellite markers for *Thujopsis dolabrata* var. *hondai* (Cupressaceae). *American Journal of Botany* 99: e317–e319.

MURRAY, M. G., AND W. F. THOMPSON. 1980. Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Research* 8: 4321–4325.

PEAKALL, R., AND P. E. SMOUSE. 2012. GenAlEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics (Oxford, England)* 28: 2537–2539.

PITTERMANN, J., S. A. STUART, T. E. DAWSON, AND A. MOREAU. 2012. Cenozoic climate change shaped the evolutionary ecophysiology of the Cupressaceae conifers. *Proceedings of the National Academy of Sciences, USA* 109: 9647–9652.

Rozen, S., and H. Skaletsky. 1999. Primer3 on the WWW for general users and for biologist programmers. *In* S. Misener and S. A. Krawetz [eds.], Methods in molecular biology, vol. 132: Bioinformatics methods and protocols, 365–386. Humana Press, Totowa, New Jersey, USA.

Schulz, M. H., D. R. Zerbino, M. Vingron, and E. Birney. 2012. Oases: Robust de novo RNA-seq assembly across the dynamic range of expression levels. *Bioinformatics (Oxford, England)* 28: 1086–1092.

Uchiyama, K., S. Fujii, W. Ishizuka, S. Goto, and Y. Tsumura. 2013. Development of 32 EST-SSR markers for *Abies firma* (Pinaceae) and their transferability to related species. *Applications in Plant Sciences* 1(2): 1200464.

Zerbino, D. R., G. K. McEwen, E. H. Margulies, and E. Birney. 2009. Pebble and rock band: Heuristic resolution of repeats and scaffolding in the Velvet short-read *de novo* assembler. *PLoS ONE* 4: e8407.