Development and Characterization of 25 EST-SSR markers in Pinus sylvestris var. mongolica (Pinaceae)

Authors: Fang, Pan, Niu, Shihui, Yuan, Huwei, Li, Zhexin, Zhang, Yuncheng, et al.

Source: Applications in Plant Sciences, 2(1)

Published By: Botanical Society of America

URL: https://doi.org/10.3732/apps.1300057

Applications in Plant Sciences 2014 2(1): 1300057

PRIMER NOTE

DEVELOPMENT AND CHARACTERIZATION OF 25 EST-SSR MARKERS IN *PINUS SYLVESTRIS* VAR. *MONGOLICA* (PINACEAE)[1]

PAN FANG[2], SHIHUI NIU[2], HUWEI YUAN[2], ZHEXIN LI[2], YUNCHENG ZHANG[2], LU YUAN[2],
AND WEI LI[2,3]

[2]National Engineering Laboratory for Forest Tree Breeding, Key Laboratory for Genetics and Breeding of Forest Trees and Ornamental Plants of the Ministry of Education, College of Biological Science and Technology, Beijing Forestry University, Beijing 100083, People's Republic of China

- *Premise of the study:* A set of novel expressed sequence tag (EST) microsatellite markers was developed in *Pinus sylvestris* var. *mongolica* to promote further genetic studies in this species.
- *Methods and Results:* One hundred seventy-five EST–simple sequence repeat (SSR) primers were designed and synthesized for 31,653 isotigs based on *P. tabuliformis* EST sequences. The primer pairs were used to identify 25 polymorphic loci in 48 individuals. The number of alleles ranged from two to eight with observed and expected heterozygosity values of 0.0435 to 0.8125 and 0.0430 to 0.7820, respectively.
- *Conclusions:* These new polymorphic EST-SSR markers will be useful for assessing genetic diversity, molecular breeding and genetic improvement, and conservation of *P. sylvestris* var. *mongolica*.

Key words: expressed sequence tag; Pinaceae; *Pinus sylvestris* var. *mongolica*; polymorphism; primer pairs; transcriptome sequencing.

Pinus sylvestris L. var. *mongolica* Litv. (Pinaceae) is indigenous to the region north of the Greater Khingan Mountains and the Hulunbeier Grassland of the Inner Mongolia Autonomous Region (Zhu et al., 2006). This pine tree can live up to 150–200 yr and grow as tall as 15 m. It has strong cold resistance, enabling it to survive temperatures as low as –40°C, and is highly adaptable to various soil types with good growth on both barren and fertile land (Zhu et al., 2003). *Pinus sylvestris* var. *mongolica* is the best evergreen coniferous tree species for establishing windbreaks and providing sand fixation (Zhao et al., 2010). Due to these characteristics, it has economic and ecological benefits, and has been introduced and cultivated in many arid and semiarid regions of China, such as Chengde in Hebei Province (Zhao and Liu, 2007), Zhanggutai in Liaoning Province (Zeng et al., 2005), and Yulin in Shaanxi Province (Wang et al., 2009). This species originated from Honghuaerji in the Inner Mongolia Autonomous Region (Zhu et al., 2006). Molecular genetic studies have been few in number (Li et al., 2005), and no simple sequence repeats (SSRs) have been reported. To optimize the conservation and utilization of *P. sylvestris* var. *mongolica*, the development of expressed sequence tag (EST)–SSR markers is very useful for germplasm identification and research into the genetic diversity of this species.

Transcriptome sequencing is an efficient method for acquiring EST sequences. SSRs derived from EST sequences are more convenient and can be isolated with higher efficiency and at lower expense than genomic sequence SSRs (Wang et al., 2012). In a previous study, Niu et al. (2013) analyzed the evolution of genes in *Pinus* species and showed by clustering analysis that *P. sylvestris* var. *mongolica* is more closely related to *P. tabuliformis* Carrière than to three other *Pinus* species. Entries in PubMed (http://www.ncbi.nlm.nih.gov/pubmed) on Pinaceae SSRs account for only 0.42% of all entries related to SSRs to date (1 May 2013), and none of them focus on *P. sylvestris* var. *mongolica*. Because transcriptome sequence data are not available for *P. sylvestris* var. *mongolica*, we used the data available from *P. tabuliformis* to develop the markers described here. We developed and characterized 25 novel polymorphic EST-SSR markers for this species. These EST-SSR markers provide an important tool for the study of genetic diversity in *P. sylvestris* var. *mongolica*.

METHODS AND RESULTS

In total, 31,653 EST-SSR loci were identified in the transcriptome sequence data from the related species *P. tabuliformis* (SRA accession: SRA056887, http://www.ncbi.nlm.nih.gov/sra). The sequences were analyzed for potential SSRs using Simple Sequence Repeat Identification Tool (SSRIT) software (Temnykh et al., 2001; http://www.gramene.org/db/markers/ssrtool/). A set of 702 SSRs was identified that met a requirement for mono-, di-, tri-, tetra-, penta-, and hexanucleotide sequences with a minimum of 12, 6, 5, 5, 5, and 5 repeats, respectively. Among these, 175 SSRs were selected randomly for primer design, excluding the SSRs located at the loci termini. Primer pairs were designed using Primer Premier 5.0 software (PREMIER Biosoft International, Palo Alto, California, USA) (Wei et al., 2012) with the following criteria: primer lengths of 16–22 bp, GC content of 40–65%, annealing temperature (T_a) ranging from 40°C to 60°C, and a predicted PCR product size ranging from 100 to 500 bp.

Genomic DNA samples were isolated from the needles of 48 *P. sylvestris* var. *mongolica* plants using the advanced cetyltrimethylammonium bromide

---

[1]Manuscript received 25 July 2013; revision accepted 22 September 2013.

This work was supported by the National Natural Science Foundation of China (grant no. 31370657).

[3]Author for correspondence: bjfuliwei@bjfu.edu.cn

doi:10.3732/apps.1300057

(CTAB) method (Porebski et al., 1997). The samples were collected from a single seed orchard in Qigou, Hebei Province (41°0′13″N, 118°27′38″E) and deposited at the National Engineering Laboratory for Forest Tree Breeding, China (NELFTB). All trees in the seed orchard derived from Honghuaerji in the Inner Mongolia Autonomous Region. PCR amplifications were performed in 20-μL volumes that included 50–80 ng of genomic DNA, 5 μM concentrations of each primer, and 10 μL 2× PCRMaster Mix consisting of 0.1 unit/μL *Taq* DNA polymerase, 4 mM $MgCl_2$, and 0.4 mM dNTP (Aidlab Co. Ltd., Beijing, China). The PCR reactions were performed in a Veriti Dx 96-well Thermal Cycler (Applied Biosystems, Foster City, California, USA) under the following conditions: initial denaturation at 95°C for 5 min followed by 35 cycles at 95°C for 30 s, annealing for 45 s at the optimal temperature for each primer pair, and 72°C for 30 s, with a final extension of 10 min at 72°C (Table 1). PCR products were resolved on 6% polyacrylamide denaturing gels using an HT-CX01 gel sequencing cell (Hongtao Jiye Technology Development Co. Ltd., Beijing, China). SSR patterns were visualized by silver staining. The SSR fragment sizes were estimated by comparison with DNA marker I (Aidlab Co. Ltd.).

One hundred seventy-five EST-SSR primer pairs were synthesized (Shanghai Sangon Co. Ltd., Beijing, China). Fifty-six primer pairs were identified that yielded stable, clear, and repeatable amplicons in *P. sylvestris* var. *mongolica*. The other primer pairs were unstable or gave no product. The 56 primers corresponded to 31 loci that were monomorphic (data not shown) and 25 loci that were polymorphic (Table 1). The polymorphic SSR loci were analyzed with POPGENE version 1.32 software (Yeh et al., 1999) for the number of alleles per locus (A), observed heterozygosity (H_o), expected heterozygosity (H_e), and fixation index (F_{IS}). Detailed data are shown in Table 2. The A values ranged from two to eight with a mean of 3.12. The H_o and H_e values were 0.0435–0.8125 and 0.0430–0.7820 with averages of 0.3412 and 0.4027, respectively. The F_{IS} values ranged from –0.2877 to 0.6773 with an average of 0.1175. Hardy–Weinberg equilibrium (HWE) and linkage disequilibrium using Bonferroni correction were tested for every locus. The following loci deviated significantly ($P < 0.002$) from HWE:

TABLE 1. Characteristics of 25 EST-SSRs developed in *Pinus sylvestris* var. *mongolica*.

Locus	Primer sequences (5′–3′)	Repeat motif	Expected size (bp)	T_a (°C)	GenBank accession no.
lw_isotig00542	F: AACAGGAGCATATCAATCAA R: GTGGCATTCTACAAGCAATT	$(T)_{40}$	257	55	KF501186
lw_isotig04204	F: CTCCGTTTGGGTTGTGTTTG R: ATCCTTGCCGCCAGATTTGT	$(CGGCT)_5$	230	55	KF501187
lw_isotig04600	F: TCAGGGAAAATGTAGGAAAATG R: AATCTGTTGTTGTGGGACTTGA	$(CAG)_{10}$	305	55	KF501188
lw_isotig06440	F: GGGACAAGGGACATCG R: TGGAGACTTCGGGTGC	$(AGGTTG)_5(AGGCTG)_6$	298	55	KF501189
lw_isotig07383	F: CAAACAAAAAACAGTCTGCA R: ATCGTCATCATCATCGTCAC	$(GAT)_8$	191	55	KF501190
lw_isotig10603	F: CAAAATCGTCTACTTCTCCCCC R: CAAAGCAAAAGAACTCCAACGA	$(CAG)_7$	196	55	KF501191
lw_isotig17679	F: TTGTTTGCCACATTGTTGCC R: CAAACCACCGCTGCTTCTAA	$(TTAA)_5$	277	55	KF501192
lw_isotig21953	F: ATGGTGTGTTTGAAGCGGAA R: ATTGCAGCCACTGGTGTCTT	$(ATGGG)_7$	208	55	KF501193
lw_isotig26230	F: GGGCATTACATAAACACGGG R: TGCCCTTGAGCATTTGATTA	$(TA)_{10}$	260	55	KF501194
lw_isotig27940	F: GCAGGCAACAACAAAAGTGACA R: AGCAATCGAGTGGCAAATCTTC	$(TGGA)_5$	231	55	KF501195
lw_isotig00080	F: CGGGCAAAATGACCGAAG R: TGGAGGAGGTAGAGGGGG	$(CCG)_6$	177	55	KF501196
lw_isotig00081	F: TGCGGAAGGCGTGAGTAG R: TGGAGGAGGTAGAGGGGG	$(CCG)_6$	290	58	KF501197
lw_isotig01420	F: TCCGTGACCCTATTACGT R: CGATTAGTTGCTTGCCTT	$(CTG)_5$	174	50	KF501198
lw_isotig02138	F: ATGCATCTTGTCCTCTCT R: TTCCTGATTCACACTCCC	$(AG)_6$	124	42	KF501199
lw_isotig02347	F: CTCGTCCTTCTTGTCCGC R: GCTATTGCTCCACTTGCC	$(TG)_7$	198	50	KF501200
lw_isotig03088	F: CATTTGGTTGACTTTGTT R: TTGTAGTGAGATCTGTGC	$(GA)_6$	235	45	KF501201
lw_isotig04931	F: TAGACCTCATCACAAACT R: ACAAAAACGAATACAAAT	$(AC)_6$	132	40	KF501202
lw_isotig02842	F: GTGATGGTGTGGTGGCTGTA R: TCCTTTGTGGGAGATTGGTG	$(AGA)_5$	229	55	KF501203
lw_isotig04195	F: GAGATCACCGAAACAACAAAA R: TACAAGTCCCAGCAAAACAAT	$(GAG)_5$	189	55	KF501204
lw_isotig04306	F: GCCATTTTTTTCTTCTCTCCT R: GGTCGGTTTCTGAATTTCTAA	$(TCC)_7$	196	55	KF501205
lw_isotig05123	F: TGTGCGTATAGGAGGTGGAG R: ATGAAAGGTGACAAAGCGGT	$(GAG)_6$	166	55	KF501206
lw_isotig06215	F: TCAGGTGCTTACCCCTTTTC R: TGGCAGCTATTCCAGTCTTT	$(CAA)_5$	275	55	KF501207
lw_isotig11166	F: ACACACACTGAGCTCCAATTT R: AGTCCCACCTCTGCTGATACA	$(TA)_7$	137	55	KF501208
lw_isotig12667	F: CCAAGGTGAAAAGGAAATGA R: TTCTGACAGGGAGCGACTGA	$(CA)_6$	199	55	KF501209
lw_isotig20215	F: AGAGGTGATCGCAGTCAAAGA R: TTCAAAAAGACCAAACCGTAG	$(TA)_7$	186	55	KF501210

Note: T_a = annealing temperature.

TABLE 2. Allelic diversity of 25 polymorphic EST-SSR loci in *Pinus sylvestris* var. *mongolica*.

Locus	A	H_o	H_e	F_{IS}
lw_isotig00542	2	0.4348	0.4816	0.0873
lw_isotig04204	2	0.4375	0.3454	−0.2800
lw_isotig04600	3	0.1053	0.1021	−0.0447
lw_isotig06440	3	0.2083	0.1932	−0.0897
lw_isotig07383	3	0.3696	0.4728	0.2097
lw_isotig10603	2	0.4375	0.4086	−0.0821
lw_isotig17679	3	0.4375	0.4432	0.0025
lw_isotig21953	7	0.6250	0.7820	0.1924
lw_isotig26230	3	0.3958	0.4629	0.1360
lw_isotig27940	3	0.1778	0.5571	0.6773
lw_isotig00080	3	0.2979	0.2919	−0.0313
lw_isotig00081	3	0.2128	0.2276	0.0553
lw_isotig01420	3	0.7872	0.6179	−0.2877
lw_isotig02138	2	0.0571	0.1590	0.6354
lw_isotig02347	2	0.0435	0.0430	−0.0222
lw_isotig03088	2	0.4545	0.5057	0.0909
lw_isotig04931	4	0.4348	0.6350	0.3078
lw_isotig02842	2	0.1250	0.3789	0.6667
lw_isotig04195	4	0.4468	0.4221	−0.0699
lw_isotig04306	3	0.5000	0.4781	−0.0569
lw_isotig05123	2	0.1458	0.1366	−0.0787
lw_isotig06215	2	0.0625	0.0612	−0.0323
lw_isotig11166	5	0.2917	0.6252	0.5286
lw_isotig12667	2	0.2292	0.4998	0.5366
lw_isotig20215	8	0.8125	0.7371	−0.1140
Average	3.12	0.3412	0.4027	0.1175

Note: A = number of alleles; F_{IS} = fixation index; H_e = expected heterozygosity; H_o = observed heterozygosity.

lw_isotig27940, lw_isotig02138, lw_isotig04931, lw_isotig02842, lw_isotig11166, and lw_isotig12667. No linkage disequilibrium ($P < 0.002$) was detected among any loci.

To identify potential functions of the 25 SSR-associated unigenes, the sequences were aligned with the GenBank database using the BLASTX program (http://blast.ncbi.nlm.nih.gov/Blast.cgi) (Yang et al., 2012). The *E*-value was limited to 0–1.0E^{-5}. Gene Ontology (GO) was also used to predict functions of the unigenes (http://geneontology.org/). Eighteen sequences were found to have potential functions by BLASTX or GO analysis. These sequences showed significant homology to protein sequences from *Picea sitchensis* (Bong.) Carrière, *Picea glauca* Voss, *Selaginella moellendorffii* Hieron., *Vitis vinifera* L., *Cucumis sativus* L., and *Zea mays* L. The potential functions were mainly related to ionic bonding, oxidation–reduction processes, and feedback regulation (Table 3).

CONCLUSIONS

Very few SSR markers for *P. sylvestris* var. *mongolica* have previously been reported. Here we have developed 25 novel EST-SSR polymorphic markers for this species. The 25 markers provide an efficient tool for investigating population genetic diversity in different environments, as well as illuminating infraspecific phylogeography, mating systems, and gene flow in different populations. These new EST-SSRs will facilitate studies on molecular breeding, genetic improvement, and conservation of *P. sylvestris* var. *mongolica*.

LITERATURE CITED

LI, H. Y., J. JIANG, G. F. LIU, X. J. MA, J. X. DONG, AND S. J. LIN. 2005. Genetic variation and division of *Pinus sylvestris* provenances by ISSR markers. *Journal of Forest Research* 16: 216–218.

NIU, S. H., Z. X. LI, H. W. YUAN, X. Y. CHEN, Y. LI, AND W. LI. 2013. Transcriptome characterisation of *Pinus tabuliformis* and evolution of genes in the *Pinus* phylogeny. *BMC Genomics* 14: 263.

POREBSKI, S., L. BAILEY, AND B. BAUM. 1997. Modification of a CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components. *Plant Molecular Biology Reporter* 15: 8–15.

TEMNYKH, S., G. DECLERCK, A. LUKASHOVA, L. LIPOVICH, S. CARTINHOUR, AND S. MCCOUCH. 2001. Computational and experimental analysis of microsatellites in rice (*Oryza sativa* L.): Frequency, length variation, transposon associations, and genetic marker potential. *Genome Research* 11: 1441–1452.

WANG, H. B., P. PAN, AND B. S. GAO. 2009. Key technologies of drought-resistant afforestation with *Pinus sylvestris* var. *mongolica* in Maowusu Sand Land. *Journal of Northwest Forestry University* 24: 70–73.

WANG, L. H., Y. X. ZHANG, X. Q. QI, Y. GAO, AND X. R. ZHANG. 2012. Development and characterization of 59 polymorphic cDNA-SSR markers for the edible oil crop *Sesamum indicum* (Pedaliaceae). *American Journal of Botany* 99: e394–e398.

WEI, Z. Z., L. B. LUO, H. L. ZHANG, M. XIONG, X. WANG, AND D. ZHOU. 2012. Identification and characterization of 43 novel polymorphic EST-SSR markers for Arum lily, *Zantedeschia aethiopica* (Araceae). *American Journal of Botany* 99: e493–e497.

YANG, A. H., J. J. ZHANG, H. TIAN, AND X. H. YAO. 2012. Characterization of 39 novel EST-SSR markers for *Liriodendron tulipifera* and cross-species amplification in *L. chinense* (Magnoliaceae). *American Journal of Botany* 99: e460–e464.

YEH, F. C., R. C. YANG, AND T. BOYLE. 1999. POPGENE version 1.32: Microsoft Windows–based freeware for population genetic analysis, quick user guide. Center for International Forestry Research, University of Alberta, Edmonton, Alberta, Canada.

ZENG, D. H., G. S. CHEN, F. S. CHEN, Q. ZHAO, AND X. Y. JI. 2005. Foliar nutrients and their resorption efficiencies in four *Pinus sylvestris* var. *mongolica* plantations of different ages on sandy soil. *Scientia Silvae Sinicae* 41: 21–27.

ZHAO, X. B., AND G. Z. LIU. 2007. A review of studies of introduction cultivates and afforestation technology on *Pinus sylvestris* var. *mongolica* in sandy area. *Journal of Northwest Forestry University* 22: 86–89.

ZHAO, X. B., S. P. SU, AND Y. R. FU. 2010. Updating and transformation technology of inefficient wind-breaking and sand-fixing forests in Yulin Sandy Land. *Journal of Northwest Forestry University* 25: 104–106.

ZHU, J. J., Z. P. FAN, D. H. ZENG, F. Q. JIANG, AND T. MATSUZAKI. 2003. Comparison of stand structure and growth between artificial and natural forests of *Pinus sylvestiris* var. *mongolica* on sandy land. *Journal of Forest Research* 14: 103–111.

ZHU, J. J., H. TAN, H. Z. KANG, AND M. L. XU. 2006. Natural regeneration characteristics of *Pinus sylvestris* var. *mongolica* forests on sandy land in Honghuaerji, China. *Journal of Forest Research* 17: 177–184.

Table 3. Potential functions of the SSR-associated sequences in *Pinus sylvestris* var. *mongolica*.

Locus	BLAST top hit accession no.	BLAST top hit description [organism]	*E*-value	GO_ID	Putative gene function
lw_isotig00542	None	None	None	None	None
lw_isotig04204	ABK21059.1	Unknown [*Picea sitchensis*]	2E-64	None	None
lw_isotig04600	XP_002273895	PREDICTED: uncharacterized protein LOC100267221 [*Vitis vinifera*]	2E-47	None	None
lw_isotig06440	None	None	None	None	None
lw_isotig07383	XP_004154913	PREDICTED: protein RCC2-like [*Cucumis sativus*]	1E-25	None	None
lw_isotig10603	None	None	None	None	None
lw_isotig17679	None	None	None	None	None
lw_isotig21953	ADE76095.1	Unknown [*Picea sitchensis*]	7E-12	None	None
lw_isotig26230	None	None	None	None	None
lw_isotig27940	None	None	None	None	None
lw_isotig00080	ABA54143.1	Putative glycine-rich protein [*Picea glauca*]	1E-11	GO:0046872	Metal ion binding
				GO:0008270	Zinc ion binding
				GO:0006355	Regulation of transcription, DNA-dependent
				GO:0003676	Nucleic acid binding
lw_isotig00081	ABA54143.1	Putative glycine-rich protein [*Picea glauca*]	1E-11	None	None
lw_isotig01420	ACN39897.1	Unknown [*Picea sitchensis*]	2E-157	GO:0055114	Oxidation-reduction process
				GO:0046872	Metal ion binding
				GO:0020037	Heme binding
				GO:0016705	Oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen
				GO:0016491	Oxidoreductase activity
				GO:0009055	Electron carrier activity
				GO:0005506	Iron ion binding
lw_isotig02138	XP_002971210.1	Hypothetical protein SELMODRAFT_171829 [*Selaginella moellendorffii*]	3E-107	GO:0046872	Metal ion binding
lw_isotig02347	XP_002990606.1	Hypothetical protein SELMODRAFT_448108 [*Selaginella moellendorffii*]	2E-57	None	None
lw_isotig03088	XP_002266814.1	PREDICTED: CCA-adding enzyme [*Vitis vinifera*]	1E-118	GO:0016779	Nucleotidyltransferase activity
				GO:0006396	RNA processing
				GO:0003723	RNA binding
lw_isotig04931	ABR16534.1	Unknown [*Picea sitchensis*]	2E-143	None	None
lw_isotig02842	ADE76527.1	Unknown [*Picea sitchensis*]	0	None	None
lw_isotig04195	ABK21301.1	Unknown [*Picea sitchensis*]	7E-80	None	None
lw_isotig04306	ABR17562.1	Unknown [*Picea sitchensis*]	1E-79	None	None
lw_isotig05123	ABK22664.1	Unknown [*Picea sitchensis*]	6E-165	None	None
lw_isotig06215	ABQ51222.1	R2R3-MYB transcription factor MYB6 [*Picea glauca*]	1E-127	GO:0006355	Regulation of transcription, DNA-dependent
				GO:0005634	Nucleus
lw_isotig11166	ABK23767.1	Unknown [*Picea sitchensis*]	2E-23	GO:0009055	Electron carrier activity
lw_isotig12667	None	None	None	None	None
lw_isotig20215	DAA51826.1	TPA: hypothetical protein ZEAMMB73_014853 [*Zea mays*]	2E-14	GO:0055114	Oxidation-reduction process
				GO:0051536	Iron-sulfur cluster binding
				GO:0050660	Flavin adenine dinucleotide binding
				GO:0046872	Metal ion binding
				GO:0016614	Oxidoreductase activity, acting on CH-OH group of donors
				GO:0016491	Oxidoreductase activity
				GO:0009055	Electron carrier activity
				GO:0005506	Iron ion binding
				GO:0003824	Catalytic activity

Note: GO_ID = Gene Ontology ID.