Short note: Development of Six EST-SSR Markers in Larch

By X. YANG, X. SUN and S. ZHANG*)

Research Institute of Forestry, Chinese Academy of Forestry, Key Laboratory of Tree Breeding and Cultivation of State Forestry Administration, Beijing 100091, People's Republic of China

(Received 21th January 2011)

Abstract

Six simple sequence repeats (SSR) markers were developed from expressed sequence tags (ESTs) in the genus *Larix*. Based on evaluation with 49 *L. kaempferi* genotypes, the number of alleles per locus ranged from two to four, and the expected (*He*) and observed (*Ho*) heterozygosity values were 0.225–0.694 and 0.201–0.656, respectively. The inbreeding coffcient (F_{IS}) for all loci were less than zero except that LAReSSR85 was 0.4383. All the six EST-SSR markers were transferable to *L. gmelini*, *L. olgensis* var *Koreana*, *L. principis-rupprechtii* and *L. olgensis*. BlastX analysis showed that five of the EST-SSRs were homologous to known genes. The six EST-SSR markers developed here can be valuable for biological applications in *Larix*.

Key words: Expressed sequence tag (EST), Simple sequence repeats (SSRs), *Larix kaempferi*, cross-species transferability.

The coniferous tree genus of larch (*Larix* Mill.) comprises 18 species and varieties, including ten species and one variety native to China (AGENDAE ACADEMIAE SINICAE, 1978). Most larch species are of considerable economic and ecological importance and have been largely planted in cold temperate and boreal regions in the northern hemisphere (KHASA et al., 2000). In China, for example, *L. gmelini* (Rupr.) Rupr., *L. olgensis* Henry and *L. principis-rupprechtii* Mayr. are commonly used in the north-eastern provinces in establishing fast-growing and high-yielding plantations, and the total of cultivation areas across the country have reached up to 1.5 (JI et al., 2004), 0.7 (LENG et al., 2004) and 0.15 (ZHANG et al., 2007) million hectares, respectively. Also, Japanese larch (*L. kaempferi* (Lamb.) Carr.), an exotic species introduced into China for more than 100 years, has been extended to the wide range around 25–40°N in the country, with a total area more than 0.33 million hectares (MA and SUN, 2008).

Simple-sequence repeats (SSRs), also known as microsatellites, are one of the most useful molecular marker types for genetic analysis. SSRs could be developed from two major sources, namely, enriched genomic libraries and expressed sequence tags (ESTs). In terms of cross-species transferability and utility, EST-SSRs are usually recognized to be advantageous over genomic SSRs (VARSHNEY et al., 2005). So far, though a number of genomic SSR markers have been reported in *Larix* (e.g. KHASA et al., 2000; ISODA and WATANABE, 2006; CHEN et al., 2009), few EST-SSR markers are available for the genus. Here we present the first report on the development and cross-species transferability investigation of larch EST-SSR markers.

SSRIT software (http://www.gramene.org/gramene/searches/ssrtool) was used to search for SSRs against a total of 1620 unique larch ESTs, including 62 ESTs from GenBank (http://www.ncbi.nlm.nih.gov/dbEST/) and 1558 developed by our group [868 from a suppression subtractive hybridization (SSH) library of *L. kaempferi* × *L. principis-rupprechtii* somatic embryos, 411 from a cDNA library of *L. kaempferi* stem and 279 from a SSH library of *L. kaempferi* × *L. olgensis* cuttings]. The minimum requirement for detection of di-, tri-, tetra-, penta- and hexa-nucleotide motifs was four repeats and 12 bp length, and 67 SSR-containing ESTs were then identified. Finally 45 primer pairs were designed using Primer Premier 5 program (Whitehead Institute for Biomedical Research, Cambridge, Mass) based on the following criteria: primer length 18–22 bp with 20 bp as the optimum, melting temperature (Tm) 50–65°C, GC content 40–65% and polymerase chain reaction (PCR) product size 100–300 bp.

The primer pairs were initially tested for PCR against 12 *L. kaempferi* genotypes (trees) sampled from a field trial. Genomic DNA was extracted from young needles as described by DOYLE and DOYLE (1990). PCR reaction of 20 µL consisted of 10× buffer (100 mM Tris-HCl pH8.3, 15 mM $MgCl_2$ and 500 mM KCl), 0.25 mM each dNTP, 0.2 µM each of the forward and reverse primers, 1 U Taq polymerase (Takara, Dalian, China) and 40 ng genomic DNA. PCR program was as following: 3 min at 94°C; 40 cycles of 30 s at 94°C, 45 s at Tm and 1 min at 72°C; and 7 min at 72°C. PCR products were checked through 6% polyacrylamide gel electrophoresis and silver staining. Finally six out of the 45 primer pairs were PCR successful across all the 12 genotypes and showed allelic polymorphisms (*Table 1*). Except for one case in LAReSSR14 whose EST sequence was downloaded from GenBank, the rest five EST-SSR markers were developed from our group's database and the EST sequences were deposited in dbEST of GenBank (accession numbers shown in *Table 1*). Moreover, in terms of SSR position in exons, LAReSSR12, LAReSSR14, LAReSSR19 And LAReSSR27,were found in coding sequence (CDS), whereas LAReSSR72 and LAReSSR85 were found in 3'UTR. (*Table 2*).

To further characterize the polymorphisms of the six EST-SSRs, 49 additional genotypes of *L. kaempferi* were tested as mentioned above. POPGENE version 1.32

*) Corresponding author: SHOUGONG ZHANG. Tel. +86 10 62889008. E-Mail: shougong.zhang@caf.ac.cn.

DOI:10.1515/sg-2011-0022
edited by Thünen Institute of Forest Genetics

Table 1. – GenBank accession, repeat motif, primer sequence, melting temperature (Tm) and target EST size of the six EST-SSR markers developed in *Larix*. The plus sign "+" indicates two or more discontinuous repeat motifs.

EST-SSR marker	GenBank accession	Repeat motif	Forward and reverse primers (5′–3′)	Tm (°C)	Target size (bp)
LAReSSR12	JG745369	$(att)_4 + (tgt)_4 + (gtggca)_4$	F:ATTATTGCCCTCTGAGTTTG R:ATTCACCCCAATCCCATC	56	131
LAReSSR14	AB251473	$(tcaggc)_5$	F:ACATTGAGCAGATGACCCAC R:ATGCGGAGGTTGAGTTGG	56	146
LAReSSR19	JG745370	$(cat)_4$	F:CCGAAATGAAGTCCGTGAG R:GCAGCAGCAAGTCCTAAAT	55	140
LAReSSR27	JG745371	$(agtcc)_4 + (gtcca)_6$	F:GGCTGAGGTTGCGAAAGA R:CAATTACATAAGTGGGACGAGA	56	142
LAReSSR72	JG745368	$(at)_6$	F:ATGGCTGTGGAAGCGAATA R:AAGGGATCACGAACTGAACTGG	60	168
LAReSSR85	JG771979	$(tac)_4$	F:TTTCGTATGGTCAAGTTCTG R:TGTCTATCCCAAGTCAGTCAT	52	172

Table 2. – Polymorphisms and function annotation of the six EST-SSR markers developed in *Larix*. NA, number of alleles per locus. *He*, the expected heterozygosity. *Ho*, the observed heterozygosity. Asterisks in F_{IS} indicate deviation from Hardy-Weinberg equilibrium (HWE) at 0.01 significance level, whereas NS denotes non-significance at 0.05 level.

EST-SSR marker	NA	*He*	*Ho*	F_{IS}	SSR position	Function annotation ($E \leq 10^{-5}$)	BlastX E-value
LAReSSR12	4	0.383	0.367	0.0308^{NS}	CDS	Chloroplast allene oxide cyclase precursor	3.0E-13
LAReSSR14	3	0.201	0.225	-0.1264^{NS}	CDS	Dehydrin	4.0E-13
LAReSSR19	2	0.328	0.367	-0.1308^{NS}	CDS	Legumin-like storage protein	6.0E-63
LAReSSR27	4	0.656	0.694	-0.0690^{NS}	CDS	Expansin 2	8.0E-56
LAReSSR72	4	0.504	0.551	-0.1039^{NS}	3'-UTR	Unknown	
LAReSSR85	3	0.551	0.306	0.4383^{**}	3'-UTR	ACR3	3.0E-65

(YEH et al., 1999) was employed to calculate the expected (*He*) and observed (*Ho*) heterozygosity as well as the inbreeding coefficient (F_{IS}) for linkage disequilibrium (LD) test. Two to four alleles were observed per EST-SSR marker, *He* and *Ho* values were 0.225–0.694 and 0.201–0.656, respectively (*Table 2*). F_{IS} values for six loci were from –0.1308 to 0.4383 (*Table 2*). Significant deviation from Hardy-Weinberg equilibrium was detected only in marker LAReSSR85 ($F_{IS} = 0.4383$), suggesting inbreeding or presence of null allele(s) at this locus. In addition, BlastX with NCBI database of non-redundant protein sequences showed that five of the EST-SSRs were homologous to known genes (*Table 2*).

Moreover, the six markers were investigated for cross-species transferability using 12 genotypes of each of the four species/varieties occurring naturally in China, *L. gmelini*, *L. olgensis* var *Koreana*, *L. principis-rupprechtii* and *L. olgensis*. PCR was performed as above and PCR products were electrophoresized through 6% polyacrylamide gel. All of the markers could be amplified successfully over the four taxa, except that LAReSSR14 produced slightly ambiguous band within 10 of the *L. gmelinii* genotypes. Thus, the six EST-SSR markers developed here may be highly transferable cross larch species and can be valuable for biological applications in *Larix*.

Acknowledgements

This work was supported by a grant from National "11.5" Key Project of China (2006BAD01A1401). The authors thank two anonymous reviewers for critical reading and valuable comments on an earlier version of the manuscript.

References

AGENDAE ACADEMIAE SINICAE (1978): Flora Reipublicae Popularis Sinicae: Tomus 7. Science Press, Beijing, China.

CHEN, C., C. LIEWLAKSANEEYANAWIN, T. FUNDA, A. KENAWY, C. H. NEWTON and Y. A. EL-KASSABY (2009): Development and characterization of microsatellite loci in western larch (*Larix occidentalis* Nutt.). Mol. Ecol. Resour. **9**: 843–845.

DOYLE, J. J. and J. L. DOYLE (1990): Isolation of plant DNA from fresh tissue. Focus **12**: 13–15.

DOI:10.1515/sg-2011-0022
edited by Thünen Institute of Forest Genetics

ISODA, K. and A. WATANABE (2006): Isolation and characterization of microsatellite loci from *Larix kaempferi*. Mol. Ecol. Notes **6**: 664–666.

JI, Y.-Z., L.-X. CHEN, B.-M. XUE and W.-Z. FENG (2004): Influence of fertilization on plant nutrition and physiological characteristics in larch plantations. Ecol. Environ. **13**: 217–219.

KHASA, P. D., C. H. NEWTON, M. H. RAHMAN, B. JAQUISH and B. P. DANCIK (2000): Isolation, characterization, and inheritance of microsatellite loci in alpine larch and western larch. Genome **43**: 439–448.

LENG, W.-F., H.-S. HE, R.-C. BU and Y.-M. HU (2007): Sensitivity analysis of the impacts of climate change on potential distribution of three larch (*Larix*) species in northeastern China. J. Plant Ecol. **31**: 825–833.

MA, C. G. and X. M. SUN (2008): Larch genetic improvement and its future development in China. World For. Res. **21(3)**: 58–63.

VARSHNEY, R. K., A. GRANER and M. E. SORRELLS (2005): Genic microsatellite markers in plants: features and applications. Trends Biotechnol. **23**: 48–55.

YEH, F. C., R. YANG and T. BOYLE (1999): POPGENE version 1.32. Microsoft window-based freeware for population genetic analysis. University of Alberta, Edmonton, Canada.

ZHANG, S.-G., W.-H. YANG, M.-X. LI, S.-Y. HAN, J.-H. WANG, X.-M. SUN, X.-S. WANG and L.-W. QI (2007): Studies on the pollen mother cell meiosis and pollen development of *Larix principis-rupprechtii*. Acta Phytotaxonomica Sinica **45**: 505–512.

DOI:10.1515/sg-2011-0022
edited by Thünen Institute of Forest Genetics