GMR Genetics and Molecular Research
Online Journal
ISSN:1676-5680

Development and characterization of SSR markers from *Pinus massoniana* and their transferability to *P. elliottii*, *P. caribaea* and *P. yunnanensis*

Y.H. Feng[1,2], Z.Q. Yang[2], J. Wang[1], Q.F. Luo[1,2] and H.G. Li[1]

[1]Key Laboratory of Forest Genetics and Biotechnology,
Nanjing Forestry University, Nanjing, Jiangsu, China
[2]Timber Forest Research Division,
Guangxi Zhuang Autonomous Region Forestry Research Institute,
Nanning, Guangxi, China

Corresponding author: H.G. Li
E-mail: hgli@njfu.edu.cn

Genet. Mol. Res. 13 (1): 1508-1513 (2014)
Received December 6, 2012
Accepted June 26, 2013
Published March 12, 2014
DOI http://dx.doi.org/10.4238/2014.March.12.2

ABSTRACT. *Pinus massoniana* (Masson's pine) is a widespread tree species in central and southern China and northern Vietnam; it is valued for rosin and paper production. Despite the significant economic value of Masson's pine, little work has been done on its molecular genetics. We developed 318 SSR primers from genome sequences of *P. massoniana*, and we identified 10 polymorphic markers. The

number of alleles in the population of *P. massoniana* that we examined ranged from two to four, and the Shannon diversity index ranged from 0.150 to 1.133. Cross-species transferability of the 318 SSRs was also analyzed in the slash pine (*Pinus elliottii*), the Caribbean pine (*Pinus caribaea*) and the Yunnan pine (*Pinus yunnanensis*); 15, 10, and 10 primer pairs generated polymorphic amplification, respectively. These sets of polymorphic SSR markers will be useful for population genetics studies of *P. massoniana*, for genetic identification of interspecific hybridization, and for phylogeographic studies of *Pinus* spp.

Key words: *Pinus massoniana*; SSR; Cross-species transferability

INTRODUCTION

Masson's pine (*Pinus massoniana*) is native to a wide area of central and southern China and northern Vietnam, and it is a common tree species in plantation forestry for replacing or compensating for the loss of the natural forest in southern China. The plantations of Masson's pine are estimated to cover 6 million hectares of land, accounting for 15% of the total Chinese marketed timber (Zhang et al., 2012). With its high cellulose content and long fiber, Masson's pine wood has a multitude of uses, mostly for pulp and paper making, as well as for rosin production (Wilson, 1993; Chen et al., 1996; Cheng et al., 2004). Despite the substantial economic value of Masson's pine, progress in its molecular genetics is very limited. This is mostly attributed to the lack of ideal DNA markers in Masson's pine, even in the genus *Pinus*.

Undoubtedly, SSRs (single sequence repeats), co-dominant molecular markers, are currently ideal markers that are widely used in population genetics, construction of genetic maps and evolutionary studies of plants (Zietkiewicz et al., 1994; Cuadrado and Schwarzacher; 1998; Gonzalo et al., 2005). Therefore, the development of SSR markers for Masson's pine would provide useful tools to investigate population genetic structure, to construct genetic maps and to implement MAS (marker assistant selection) in this species. Furthermore, the set of SSR markers developed in Masson's pine would also have potential application value in the genetic identification of interspecific hybridization and the phylogeographic study of the genus *Pinus*. Here, we developed a set of SSRs derived from the genome of Masson's pine and evaluated their transferability in the genus *Pinus*.

MATERIAL AND METHODS

Plant material and DNA extraction

Altogether, 110 trees (genotypes) from four pine species were taken as plant material, of which 30 trees came from *Pinus massoniana* Lamb (Masson's pine), 30 trees from *P. elliottii* Engelm (slash pine), 30 trees from *P. yunnanensis* Franch (Yunnan pine), and 20 trees from *P. caribaea* Morelet (Caribbean pine).Young leaves were collected from each tree and stored in a refrigerator. All plant samples above were collected from

the pine gene pool in Nanning Forestry Research Institute (Nanning, Guangxi, China, 23°10'08" N, 107°59'40"E). Genomic DNA was extracted from young leaves using the CTAB method (Doyle and Doyle, 1987).

DNA sequencing and SSR identification

Genomic DNA sequences of *Pinus massoniana* Lamb were obtained through the Solexa sequencing technology (Cronn et al., 2008; Mardis, 2008). The SSRs were searched from the genome sequences using the SSRIT software (Temnykh et al., 2001). The screening criteria were set for the detection of di-, tri-, tetra-, penta-, and hexanucleotide motifs with a minimum of six, four, three, three, and three repeats, respectively. A total of 736 SSRs were found. SSRs that had low GC content or palindromic sequences in their flanking region were excluded from primer design. The major parameters for primer design were set as follows: primer length from 18 to 22 nucleotides with 20 as the optimum, PCR product size from 200 to 500 bp, optimal annealing temperature of 55°C, and GC content from 40 to 65% with 50% as the optimum. Finally, 311 locus-specific SSR primers were designed using the PRIMER 3 software (Steve and Helen, 2000).

PCR amplification, cross-species transferability and polymorphism analysis

To assess polymorphism, polymerase chain reaction (PCR) was performed in a 10 μL reaction solution containing 10-20 ng genomic DNA, 10 mM Tris-HCl, pH 8.0, 50 mM KCl, 2.5 mM $MgCl_2$, 0.2 mM dNTPs, 0.25 μM each primer, and 0.8 U Taq polymerase (Generay, Shanghai, China). The amplification protocol consisted of an initial denaturation at 94°C for 4 min, followed by 30 cycles of 15 s at 94°C, 15 s at the appropriate annealing temperature and 30 s at 72°C, ending with a final extension at 72°C for 15 min. Amplified products were separated on 8% denaturing polyacrylamide gels and visualized by silver staining. A 50-bp DNA ladder was used to identify alleles. The number of alleles (N_A) and Shannon's information index (I) were performed using Popgene1.32 (Yeh et al.,1999).

RESULTS AND DISCUSSION

Of the 318 SSR markers tested, 305 generated expected amplification products, and 10 of them were polymorphic. The genetic variability of the 10 polymorphic markers was estimated by genotyping thirty individuals of Masson's pine. Population genetic parameters were calculated through the PopGene32 software (Yeh et al., 1999). The number of alleles per locus ranged from two to four, with an average value of 2.4. The Shannon diversity index (I) (Lewontin, 1972) was from 0.150 to 1.133, with an average of 0.503 (Table 1).

To test cross-species transferablity, the same set of 318 SSR markers were amplified in thirty individuals of slash pine, twenty individuals of Caribbean pine, and thirty individuals of Yunnan pine. Thirty-four polymorphic SSR markers were characterized: 15 markers in slash pine, 10 markers in Caribbean pine, and 10 markers in Yunnan pine. The number of alleles per locus ranged from two to four, and the Shannon diversity index (I) per locus was from 0.087 to 1.144 (Table 1). The characteristics of 34 polymorphic SSR markers in *Pinus* are listed in Table 2.

This set of polymorphic SSR markers may serve as powerful tools not only for the identification of genetic structure and gene flow, the construction of genetic maps, and MAS in Masson's pine, but also for the identification of interspecific hybrids and phylogeographic studies in the genus *Pinus*.

Table 1. Results of cross-species amplification of thirty-four polymorphic SSR markers in *Pinus*.

Locus	Species	Sample size	N_A	I
PF322	Masson pine	30	2	0.678
PF402	Masson pine	30	3	0.66
PF408	Masson pine	30	3	0.581
PF460	Masson pine	30	2	0.572
PF463	Masson pine	30	4	1.133
PF464	Masson pine	30	2	0.15
PF492	Masson pine	30	2	0.604
PF569	Masson pine	30	2	0.251
PF576	Masson pine	30	2	0.204
PF615	Masson pine	30	2	0.199
PF310	Slash pine	30	2	0.2411
PF314	Slash pine	30	2	0.2868
PF377	Slash pine	30	3	0.7216
PF383	Slash pine	30	3	0.8647
PF402	Slash pine	30	3	0.3368
PF429	Slash pine	30	4	1.1439
PF489	Slash pine	30	4	1.0678
PF494	Slash pine	30	2	0.6474
PF505	Slash pine	30	4	0.7535
PF511	Slash pine	30	3	0.3887
PF549	Slash pine	30	2	0.5961
PF552	Slash pine	30	3	0.5524
PF585	Slash pine	30	3	0.5838
PF593	Slash pine	30	4	0.9941
PF615	Slash pine	30	2	0.5004
PF429	Caribbean pine	20	4	1.054
PF431	Caribbean pine	20	2	0.5
PF436	Caribbean pine	20	2	0.631
PF441	Caribbean pine	20	2	0.418
PF443	Caribbean pine	20	2	0.588
PF489	Caribbean pine	20	2	0.693
PF494	Caribbean pine	20	2	0.423
PF511	Caribbean pine	20	3	0.88
PF533	Caribbean pine	20	2	0.647
PF606	Caribbean pine	20	2	0.562
PF334	Yunnan pine	30	2	0.3251
PF403	Yunnan pine	30	2	0.5716
PF408	Yunnan pine	30	2	0.6109
PF463	Yunnan pine	30	2	0.0871
PF464	Yunnan pine	30	2	0.2573
PF492	Yunnan pine	30	2	0.5269
PF555	Yunnan pine	30	2	0.3326
PF557	Yunnan pine	30	2	0.2937
PF561	Yunnan pine	30	2	0.251
PF615	Yunnan pine	30	2	0.6906

N_A = number of alleles; I = shannon's information index.

Table 2. Characteristics of thirty-four microsatellite primers developed in *Pinus*.

Locus	Primer sequence (5'-3')	Repeat motif	Expected size (bp)	Ta (°C)
PF310	F: CGTCCCTCCCGTTTATTG R: GGTGACCTTGCTGCCTTG	$(TTTG)_4$	356	56
PF314	F: ATGCTTGCCTTATGACTTGACA R: CAGCACTACTATTGCAGGGAGA	$(CCAAA)_3$	336	57
PF322	F: CTGTGGCTATCTTTGACTCTGC R: GAAATTCTTGTGGTCGGATGTA	$(ACCT)_4$	411	57
PF334	F: GGGTGTATAGAGGGAAGGATTT R: ACAGAGGGGCTAGGTCAGG	$(TTGC)_4$	383	56
PF377	F: TGTTTCACCCACGCCAGTC R: GCCATTTCAAAAGAGGCAGAT	$(TAA)_5$	434	58
PF383	F: TGGGCGTAGGAGGGTTGT R: GGGCTTTCTTTGTGCTATTGG	$(AGTC)_5$	241	58
PF402	F: ATGCTCATAATGAAATGGGACT R: ATGCATTGCACTGCACGT	$(TGTCAG)_3$	203	55
PF403	F: ACGATTCTTGCCAACGCT R: GCTGGAACAATTCAAATTTTGT	$(AAATTG)_3$	314	56
PF408	F: TACAAAGGACTCCAGCAAAGTG R: GCGGATGTGCGAGGTTATG	$(TGAAT)_3$	302	57
PF429	F: GCTCCATGTTTGGAAGGG R: CCAGCCAGCGATCTAAGTAA	$(TTGC)_3$	401	55
PF431	F: TCCACATCTATGGGTGCTTG R: GCTTCATTGCTGAAAGGTCA	$(TCA)_4$	247	56
PF436	F: CAGGGAAGGAGACAAAACA R: GAGGGAAGAAGAAAGACATAAA	$(CCTC)_3$	207	52
PF441	F: CCACCAATGACATCAAGGAG R: TTACGAGTAAGCAAGTGACAGC	$(CAAT)_3$	324	55
PF443	F: TTTCCTTATCGCCCAAGT R: GGCCCAATGATTATCATACA	$(ATTT)_3$	295	52
PF460	F: AACCTCATCTGAAGAAGCCATA R: AGCAGCATTACCAGCAACATA	$(TGC)_4$	325	56
PF463	F: CTCTGGGTCCGTACTATCCG R: GAAGCAAAGGCGAGCAAA	$(ATCT)_3$	253	57
PF464	F: TTGCTCGCCTTTGCTTCT R: GCCTCCTTTACCACAGCCT	$(ATGC)_3$	233	56
PF489	F: GGGAAGCGAAAGTGATTATTG R: CCTGGGACTGAGACTGATTGA	$(AATA)_3$	446	57
PF492	F: TTATGTTGCGGATCAAGAATT R: ACATGGGCACAACTTGCTAT	$(GA)_7$	238	55
PF494	F: GCCCTTAATGGATTATTCTGC R: TTTCTACGCCTCCTCCTGTC	$(AGG)_4$	447	56
PF505	F: AGAGGAATAAGGTAAGGGATGG R: GAAAACCGCTTCAATGGC	$(TTAT)_3$	416	56
PF511	F: TCCTTCTGTTATTGTACCCTCC R: GATTGATTGTATTGCACCCAC	$(CAGG)_3$	372	55
PF533	F: TCCCGAAGAAAGGAACAC R: CATAGAACGCACGCAAAT	$(AAGC)_3$	224	52
PF549	F: GGGGTCAGGTTTGGCATC R: CGGGCTAAGCTAAGCAGGTA	$(CTGA)_3$	215	57
PF552	F: TGGTTTGGACATGGACTCAC R: AGATTTCCTTCAGAGGTTTCG	$(ATC)_4$	220	55
PF555	F: GGTGAGGGTAGTCGTCTGTCT R: CGATGAAGGGCAACTATGAT	$(CCTT)_3$	377	55
PF557	F: AGCACTCATCAACTTTCCAGC R: AAGATTAGATAAGTCCCGTCCC	$(AATG)_3$	440	56
PF561	F: TTATTGGGCAGGAGGACG R: TGGTTGCTTTAGATCGAACAG	$(GAAC)_3$	327	55
PF569	F: AACAGAAAGGAATCAAGTAGGC R: AATCACATGAAATGCTGGAAA	$(AGTA)_3$	308	55
PF576	F: CTTGGCGGCACTTATTGA R: ATGATATGATGGGGCTGGTA	$(CTTA)_3$	211	55
PF585	F: CTCCCGCTTTTCCTCCAC R: CCGTTTTCATTTCAGTCCTTG	$(CACT)_3$	338	57

Continued on next page

Table 2. Characteristics of thirty-four microsatellite primers developed in *Pinus*.

Locus	Primer sequence (5'-3')	Repeat motif	Expected size (bp)	Ta (°C)
PF593	F: AACTCCCCTTCCCCATACG R: ATTCCCGCCGACTCCTAA	$(CTTG)_3$	427	57
PF606	F: CAAGCGGAGTATGTCAGGTAG R: ATGTTTGTAGGTCGTTAGAGGG	$(AT)_6$	407	55
PF615	F: TAAATGATTGGCTATCGGAGAC R: CCCGCTCTGAAGATGTTGTC	$(ATTGAG)_4$ $(ACTA)_3(ACTC)_3$	399	56

Ta = annealing temperature.

ACKNOWLEDGMENTS

Research supported by the BaGui-Scholar Foundation, Guangxi Provincial Scientific and Technical Research Project (#GKG1123004-4A), Jiangsu Provincial Graduate Student Innovation Program (#CXZZ12_0541), the Program Development of Jiangsu Higher Education Institutions (PAPD), and State Financial Project for the Commercialization of Scientific & Technical Achievements (#2011TK090). We are grateful to Y.L. Huang for his assistance in plant sample collection. We also thank L.S. Li for genomic DNA sequencing and D.S. Wu for technical support.

REFERENCES

Chen TH, Wang ZR and Xu L (1996). Chem. Industry Forest Prod. Genetic variation of wood properties in *Pinus massoniana* Lamb and the utilization potential in papermaking industry. *Chem. Industry Forest Prod.* 16: 71-78.

Cheng XS, Chen WJ, Chen YP, Chen YX, et al. (2004). Preparation and Properties of HBS Lignin from Masson Pine. *Chem. Res. Chinese Univ.* 20: 225-228.

Cronn R, Liston A, Parks M, Gernandt DS, et al. (2008). Multiplex sequencing of plant chloroplast genomes using Solexa sequencing-by-synthesis technology. *Nucleic Acids Res.* 36: e122.

Cuadrado A and Schwarzacher T (1998). The chromosomal organization of simple sequence repeats in wheat and rye genomes. *Chromosoma* 107: 587-594.

Doyle JJ and Doyle JL (1987). Isolation of plant DNA from fresh tissue. *Focus* 12: 13-15.

Gonzalo MJ, Oliver M, Garcia-Mas J, Monfort A, et al. (2005). Simple-sequence repeat markers used in merging linkage maps of melon (*Cucumis melo* L.). *Theor. Appl. Genet.* 110: 802-811.

Lewontin RC (1972). The apportionment of human diversity. *Evolutionary Biol.* 6: 381-398.

Mardis ER (2008). The impact of next-generation sequencing technology on genetics. *Trends Genet.* 24: 133-141.

Steve R and Helen S (2000). Primer3 on the WWW for General Users and for Biologist Programmers. In: Bioinformatics Methods and Protocols: Methods in Molecular Biology. (Krawetz S and Misener S, eds.). Humana Press, Totowa, 365-386.

Temnykh S, DeClerck G, Lukashova A, Lipovich L, et al. (2001). Computational and experimental analysis of microsatellites in rice (*Oryza sativa* L.): frequency, length variation, transposon associations, and genetic marker potential. *Genome Res.* 11: 1441-1452.

Wilson LF (1993). China's masson pine forests: cure or curse. *J. Forestry* 91: 30-33.

Yeh FC, Yang RC and Boyle TBJ (1999). POPGENE version 1.32, Microsoft Window- Based Free Ware for Population Genetic Analysis. Computer Program and Documentation Distributed by University of Alberta and Centre for International Forestry Research, Alberta.

ZhangY, Yang Q, Zhou ZC and Jin GQ (2012). Divergence among masson pine parents revealed by geographical origins and SSR markers and their relationships with progeny performance. *New Forests* 44: 341-355.

Zietkiewicz E, Rafalski A and Labuda D (1994). Genome fingerprinting by simple sequence repeat (SSR)-anchored polymerase chain reaction amplification. *Genomics* 20: 176-183.