Characterization of Polymorphic Microsatellite Markers in Pinus armandii (Pinaceae), an Endemic Conifer Species to China

Authors: Dong, Wan-Lin, Wang, Ruo-Nan, Yan, Xiao-Hao, Niu, Chuan, Gong, Lin-Lin, et al.

Source: Applications in Plant Sciences, 4(10)

Published By: Botanical Society of America

URL: https://doi.org/10.3732/apps.1600072

Applications in Plant Sciences 2016 4(10): 1600072

PRIMER NOTE

CHARACTERIZATION OF POLYMORPHIC MICROSATELLITE MARKERS IN *PINUS ARMANDII* (PINACEAE), AN ENDEMIC CONIFER SPECIES TO CHINA[1]

WAN-LIN DONG[2,3], RUO-NAN WANG[2,3], XIAO-HAO YAN[2], CHUAN NIU[2], LIN-LIN GONG[2], AND ZHONG-HU LI[2,4]

[2]Key Laboratory of Resource Biology and Biotechnology in Western China, Ministry of Education, College of Life Sciences, Northwest University, Xi'an 710069, People's Republic of China

- *Premise of the study: Pinus armandii* (Pinaceae) is an important conifer tree species in central and southwestern China, and it plays a key role in the local forest ecosystems. To investigate its population genetics and design effective conservation strategies, we characterized 18 polymorphic microsatellite markers for this species.
- *Methods and Results:* Eighteen novel polymorphic and 16 monomorphic microsatellite loci of *P. armandii* were isolated using Illumina MiSeq technology. The number of alleles per locus ranged from two to five. The expected heterozygosity ranged from 0.061 to 0.609 with an average of 0.384, and the observed heterozygosity ranged from 0.063 to 0.947 with an average of 0.436. Seventeen loci could be successfully transferred to five related *Pinus* species (*P. koraiensis*, *P. griffithii*, *P. sibirica*, *P. pumila*, and *P. bungeana*).
- *Conclusions:* These novel microsatellites could potentially be used to investigate the population genetics of *P. armandii* and related species.

Key words: cross-amplification; microsatellite markers; Pinaceae; *Pinus armandii*; polymorphism; population genetics.

Pinus armandii Franch. (Pinaceae) is an evergreen conifer tree species that is endemic to central and southwestern China (Fu et al., 1999). As a dominant species in warm- and cold-temperate forests, *P. armandii* plays a key role in the local ecosystems (Willyard et al., 2007; Liu et al., 2014). Previous studies of *P. armandii* focused mainly on its physiological ecology (Xiong et al., 2010; Yu et al., 2014), phylogenetic relationships, and phylogeographic structure (Liu et al., 2014; Li et al., 2015). In recent years, due to overcutting and destruction of natural habitats, the natural populations of *P. armandii* have been dramatically decreasing (Wang et al., 2014). It is important to gain knowledge of population genetic structure and genetic diversity of *P. armandii* to formulate effective conservation and management strategies. In addition, the closely related species *P. koraiensis* Siebold & Zucc., *P. griffithii* McClell., *P. sibirica* Du Tour, *P. pumila* (Pall.) Regel, and *P. bungeana* Zucc. ex Endl., which form a clade with *P. armandii* within subg. *Strobus* (D. Don) Lemmon (Liu et al., 2014; Li et al., 2015), are also important forest species in eastern Asia. In this study, we developed and characterized polymorphic microsatellite loci (simple sequence repeats [SSRs]) of *P. armandii* and its relatives to facilitate studies of their population genetics.

METHODS AND RESULTS

Genomic DNA was extracted from a fresh needle (specimen no.: WNU-NG-SX-2013-LZH-036) of *P. armandii* using the DNeasy Plant Mini Kit (QIAGEN, Hilden, Germany) and was sequenced using an Illumina MiSeq (Illumina, San Diego, California, USA) at Shanghai Genesky Biotechnologies (Shanghai, China) with 2 × 300-bp paired-end sequencing and MiSeq Reagent Kit version 3 (Illumina). A total of 6,783,777 clean reads were obtained after the adapter and low-quality sequences were removed. These clean reads were further assembled into 350,628 contigs using CLC Genomics Workbench version 7.5 (CLC bio, Aarhus, Denmark). The set of detailed parameters were: mismatch cost of 2, length fraction of 0.4, similarity fraction of 0.4, insertion cost of 2, deletion cost of 2, and a minimum contig length of 200 nucleotides. We extracted the contigs containing microsatellite markers with SciRoKo version 3.1 (Kofler et al., 2007), using default identification criteria used for mono-, di-, tri-, tetra-, penta-, and hexanucleotide repeats, with a minimum of 14, seven, five, four, four, and four repeats, respectively. In total, 887 microsatellite-containing contigs were obtained. Then, forward and reverse primers were designed with Primer Premier version 7.0 software (Clarke and Gorley, 2015). The criteria for primer design were as follows: (1) product size from 100 to 350 bp; (2) primer size from 18 to 25 bp with an optimum size of 20 bp; (3) primer melting temperate from 55°C to 63°C with an optimum temperature of 60°C; and (4) GC content of primers from 40% to 60%.

Fifty pairs of primers containing microsatellite repeats were randomly selected to test amplification efficiency and polymorphism in 52 individuals from three natural populations of *P. armandii* (Appendix 1). PCR amplification was performed in a 10-μL reaction volume containing 10 ng DNA template, 5 μL 2× polymerase mixture, 0.2 μM of each primer, and 3.6 μL ddH_2O. The PCR profiles were as follows: an initial denaturation of 5 min at 95°C; 35 cycles of denaturation of 30 s at 95°C, at the appropriate annealing temperature (Table 1) for 30 s, and an extension of 30 s at 72°C; followed by a final extension of 5 min at

[1]Manuscript received 19 June 2016; revision accepted 7 August 2016.

This work was supported by the National Natural Science Foundation of China (41101058, 31470400, J1210063), the Undergraduate Innovation and Entrepreneurship Training Programs of Northwest University (2016168), and the Program for Changjiang Scholars and Innovative Research Team in University (PCSIRT; no. IRT1174).

[3]These authors contributed equally to this work.

[4]Author for correspondence: lizhonghu@nwu.edu.cn

doi:10.3732/apps.1600072

TABLE 1. Characteristics of 34 microsatellite primers developed for *Pinus armandii*.

Locus[a]	Primer sequences (5′–3′)	Repeat motif	T_a (°C)	Allele size range (bp)	*A*	GenBank accession no.
Pa83	F: TAGTGTGGGAGTGGGAGGAA	$(AG)_{10}$	62	208–220	5	KU373058
	R: CCCACACCCTCTCCCTACTT					
Pa1539	F: AATTTTAGATGTAAAGCCTCATG	$(TA)_{12}$	53	204–210	2	KU373059
	R: TTGTGAACTAACTTTGGTGGG					
Pa2226	F: CATTGATCCTCAGCAGGTAG	$(TA)_{12}$	55	254–264	2	KU373060
	R: TATTGTTGTTTCATTCCCAC					
Pa2423	F: ATGACCAAATCACCCACAAA	$(ATTT)_4$	55	136–148	2	KU373061
	R: TTTGACTTGGGTCAAATCCC					
Pa26711	F: CAAGGTCAAGGTAAGGTTAAGGG	$(AACCTT)_5$	60	107–121	3	KU373062
	R: AAGGTTAAGGTTAAGGTTAGGTTAAGG					
Pa3553	F: AAGATTAAATCCCTAGCATCTACC	$(ATTT)_5$	59	341–361	5	KU373063
	R: TGTCCACGAGTTCTGCTCTGT					
Pa3701	F: TCATTACAGATGGCTGCGTC	$(AT)_{13}$	59	203–207	2	KU373064
	R: CCCAGTCGGAATCCTGTAAA					
Pa5960	F: TTACCCTAGCCACGACTATGC	$(GCCTA)_6$	55	204–209	3	KU373065
	R: GCTGCGTAAGGTTCGGTTAG					
Pa10136	F: CCATATGGTCACGCTACCTCT	$(TTA)_5$	59	289–298	4	KU373066
	R: TATGGAGTCAAGGTGGGAGC					
Pa11411	F: GAGAGGCCTGTCATGGAGTC	$(AGG)_6$	53	104–107	2	KU373067
	R: TAAAGGAGGCAGACCACGTC					
Pa12333	F: CCTTAACCTTAACCTTAACCTAACC	$(AACCTT)_7$	59	232–240	3	KU373068
	R: TTGACCTTGACGAAACCCTT					
Pa15326	F: CCCTTAACCTTAACCTTAACCTGAG	$(AAGGTT)_4$	59	138–144	4	KU373069
	R: CCCTAACCTTGACCAAACCC					
Pa118137	F: TACCAGTGCTCTTGGACTTGTGT	$(GAT)_8$	62	87–96	4	KU373070
	R: GAAAGTCACCATCCTCACCCTC					
Pa180916	F: CACATACACATCTATCTGCAAGC	$(AT)_{19}$	59	94–106	3	KU373071
	R: GTACCACCAGCTGATATTTGACA					
Pa5962	F: CCCTACCCATACACTACCCTAGC	$(CCTAA)_5$	65	238–244	3	KX254165
	R: AGGATGGTCTAGGATGGGCT					
Pa14201	F: TTCATAGGTTGTCAAGAAAGAGG	$(AT)_{12}$	57	232–238	2	KX254166
	R: AATAACAAGCCAAAGAAATCTCA					
Pa5418	F: AGGGCGTGACAGTTGGTATC	$(TTA)_8$	55	226–238	2	KX254167
	R: TGTCCTCCTCTTGCACAATG					
Pa8608	F: GGGTTTGGTCAAGGTTAGGG	$(TAAGGT)_4$	57	238–244	3	KX254168
	R: AAGGTTTGGTCAAGGTTAGGG					
Pa2181	F: GAGAGAGCGTGTATGTTTGGG	$(AG)_{20}$	62	216	1	KX254169
	R: TCATCTCTCTTTCCCTCCCTC					
Pa3455	F: ATGCTAGGCAAGGTAAGGCT	$(CTAGG)_4$	60	208	1	KX254170
	R: CCTATCCAATCGTAGCCCAA					
Pa5890	F: GGCTTGGGAGATTCTCGG	$(CTCTGC)_4$	56	150	1	KX254171
	R: GCAAAGAAGCAAATGAAGGC					
Pa6516	F: AAACATGGTGACCCAAGCAT	$(AAT)_9$	56	95	1	KX254172
	R: TTGAAGTCATCTTGTAATGTACTTGTC					
Pa9058	F: ACTTGGTAACCTTTCGCTTCT	$(TA)_{14}$	55	127	1	KX254173
	R: TGTGGATTTAAATGGAGATGAAA					
Pa9864	F: CCTTAACCTTAACCTAACCTTAACCT	$(GTTAAG)_6$	60	181	1	KX254174
	R: CCCTAACCTTGACCAAACCC					
Pa12494	F: AAGGACCTAGCCTTCTTGGG	$(TTGA)_6$	52	165	1	KX254175
	R: GCCCAATGGATTAATCTTCC					
Pa18101	F: TTGTTTGACACATCTAACAAGACC	$(TA)_{14}$	61	206	1	KX254176
	R: GATGGTTGAACTACATTTGGCA					
Pa19210	F: CACAATGTATCAATGGTCCG	$(AAT)_8$	60	330	1	KX254177
	R: ACAAGTGTTGAGTTAGGCGTAG					
Pa86828	F: GATTGGGGTTTATGAATGCTT	$(TG)_{12}$	59	172	1	KX254178
	R: AGAAAATAAACAATAGCGAGAGC					
Pa117430	F: AGAGATAGAAAGGGGGGGAG	$(AG)_{12}$	59	98	1	KX254179
	R: TTTGTCTCTTTATCTCACCCC					
Pa120817	F: CAACGATCCATGATGACCCTG	$(ACAT)_5$	56	204	1	KX254180
	R: TGCCTTGGCTATGTTGGGAA					
Pa832	F: CAATCTCTCCCCATTTCTATC	$(AATA)_6$	58	236	1	KX254181
	R: CCTCCCACTCCCACACTATC					
Pa101	F: GGAGACAGGGAGAGAGAGCA	$(GA)_{14}$	55	272	1	KX254182
	R: TAGGATAGGCTAGGCGAGGC					
Pa3849	F: GGGTGTTACACTAACCCAGCC	$(CCTAA)_4$	59	238	1	KX254183
	R: GCAACTCCTACTTCAGGTGTGT					

TABLE 1. Continued.

Locus[a]	Primer sequences (5′–3′)	Repeat motif	T_a (°C)	Allele size range (bp)	*A*	GenBank accession no.
Pa23367	F: GGGAGGGAAGAAGAAAGACA R: CCCTACCTCTCTCCACTCTCTCT	$(GA)_{14}$	60	240	1	KX254184

Note: *A* = number of alleles; T_a = annealing temperature.
[a]The first 18 primer pairs were determined to be polymorphic in *Pinus armandii.*

72°C. The PCR amplification products were separated in 10% nondenaturing polyacrylamide gels and were visualized by silver staining.

The allele sizes for each individual were automatically determined using Quantity One (Bio-Rad, Hercules, California, USA) with pBR322 DNA/MspI as DNA molecular-weight marker. The program GenAlEx version 6.501 (Peakall and Smouse, 2012) was used to evaluate various population genetic parameters of microsatellite loci, including the number of alleles per locus, expected and observed heterozygosity (H_e and H_o), and Hardy–Weinberg equilibrium (HWE). In addition, linkage disequilibrium (LD) among loci was detected using GENEPOP version 4.2.2 (Rousset, 2008). We also detected the null allele frequencies for each primer with MICRO-CHECKER version 2.2.3 (van Oosterhout et al., 2004).

In total, 34 primer pairs were successfully amplified with high-quality PCR products, with 18 of them exhibiting polymorphisms (Table 1). The number of alleles of these polymorphic primers ranged from two to five with an average of 2.4. H_e ranged from 0.061 to 0.609 with an average of 0.384, and H_o ranged from 0.063 to 0.947 with an average of 0.436. Two pairs of primers (Pa3553 and Pa118137) were found to deviate greatly from HWE, while we did not detect any LD between loci. This deviation might have been caused by insufficient sample size, nonrandom mating between individuals, migration, and/or natural selection of these two loci. In addition, no null alleles were detected for any locus in the current study. The detailed SSR characteristics are provided in Table 2.

To explore the broader utility of the SSR loci developed here, we amplified the primers in 20 individuals from five other species closely related to *P. armandii* (Appendix 1). Seventeen of the 18 primers produced robust, usually polymorphic DNA fragments across *P. koraiensis*, *P. griffithii*, *P. sibirica*, *P. pumila*, and *P. bungeana.* However, Pa3553 was not successfully amplified in *P. pumila* and *P. bungeana* (Table 3).

CONCLUSIONS

In the current study, we developed 18 polymorphic and 16 monomorphic loci for *P. armandii*, with allele numbers ranging from two to five for the polymorphic loci. These microsatellite markers will be useful for conservation genetic studies of *P. armandii*, such as those detecting genetic diversity and patterns of gene flow within and between populations. An assessment of their genetic information will also contribute to addressing how declining populations of *P. armandii* affect genetic diversity and gene flow, and will be useful more broadly in subg. *Strobus*.

LITERATURE CITED

CLARKE, K. R., AND R. N. GORLEY. 2015. PRIMER v7: User Manual/Tutorial, p. 296. PRIMER-E Ltd., Plymouth, United Kingdom.

FU, L. G., N. LI, AND R. R. MILL. 1999. Pinaceae. *In* C. Y. Wu and P. H. Raven [eds.], Flora of China, vol. 4, 11–52. Science Press, Beijing, China, and Missouri Botanical Garden Press, St. Louis, Missouri, USA.

KOFLER, R., C. SCHLOTTERER, AND T. LELLEY. 2007. SciRoKo: A new tool for whole genome microsatellite search and investigation. *Bioinformatics (Oxford, England)* 23: 1683–1685.

LI, Z. H., C. YANG, K. S. MAO, Y. Z. MA, J. LIU, Z. L. LIU, T. T. DENG, AND G. F. ZHAO. 2015. Molecular identification and allopatric divergence of the white pine species in China based on the cytoplasmic DNA variation. *Biochemical Systematics and Ecology* 61: 161–168.

TABLE 2. Locus-specific measures of genetic diversity across three populations of *Pinus armandii.*[a]

	YT population (*n* = 16)				NG population (*n* = 17)				XH population (*n* = 19)			
Locus	*A*	H_o	H_e	HWE[b]	*A*	H_o	H_e	HWE[b]	*A*	H_o	H_e	HWE[b]
Pa83	3	0.563	0.576	0.196	4	0.750	0.607	0.437	4	0.368	0.359	0.050
Pa1539	M	—	—	—	2	0.235	0.360	0.154	2	0.158	0.145	0.709
Pa2226	2	0.063	0.061	0.897	2	0.529	0.389	0.138	2	0.579	0.450	0.212
Pa2423	2	0.250	0.375	0.182	2	0.294	0.251	0.477	2	0.158	0.229	0.178
Pa26711	2	0.125	0.219	0.086	2	0.294	0.251	0.477	2	0.368	0.301	0.325
Pa3553	2	0.500	0.375	0.182	5	0.882	0.604	0.001*	2	0.526	0.388	0.120
Pa3701	M	—	—	—	2	0.176	0.327	0.058	2	0.474	0.450	0.820
Pa5960	2	0.250	0.375	0.182	2	0.235	0.208	0.582	3	0.474	0.450	0.648
Pa10136	M	—	—	—	4	0.412	0.389	0.354	3	0.158	0.148	0.987
Pa11411	2	0.500	0.375	0.182	2	0.588	0.415	0.086	2	0.316	0.388	0.418
Pa12333	2	0.375	0.469	0.424	2	0.500	0.375	0.212	3	0.421	0.481	0.844
Pa15326	3	0.250	0.225	0.955	2	0.294	0.251	0.477	4	0.316	0.393	0.141
Pa118137	4	0.875	0.609	0.000*	3	0.882	0.517	0.014	3	0.947	0.548	0.000*
Pa180916	M	—	—	—	M	—	—	—	3	0.105	0.101	0.996
Pa5962	2	0.688	0.498	0.128	2	0.706	0.498	0.086	3	0.579	0.536	0.628
Pa14201	2	0.313	0.404	0.364	2	0.471	0.484	0.906	2	0.368	0.494	0.267
Pa5418	2	0.625	0.430	0.069	2	0.412	0.389	0.812	2	0.684	0.478	0.060
Pa8608	2	0.188	0.170	0.679	2	0.647	0.500	0.225	2	0.474	0.494	0.855

Note: *A* = number of alleles; H_e = expected heterozygosity; H_o = observed heterozygosity; M = monomorphic fragment; *n* = number of individuals sampled.
[a]Locality and voucher information are provided in Appendix 1.
[b]*P* value of Hardy–Weinberg equilibrium test (*$P < 0.001$).

TABLE 3. Results of tests of cross-amplification of the 18 polymorphic microsatellite markers developed for *Pinus armandii* in each of five related *Pinus* taxa.[a]

Species name	*N*	Pa83	Pa1539	Pa2226	Pa2423	Pa26711	Pa3553	Pa3701	Pa5960	Pa10136	Pa11411	Pa12333	Pa15326	Pa118137	Pa180916	Pa5962	Pa14201	Pa5418	Pa8608
P. koraiensis	6	2	2	1	2	4	3	2	2	1	1	2	2	3	1	2	1	2	2
P. griffithii	5	3	3	2	2	3	1	1	2	1	3	1	1	2	1	2	2	1	2
P. sibirica	2	4	1	1	2	2	3	1	1	2	1	2	2	1	1	2	1	2	1
P. pumila	2	1	2	1	2	3	—	1	1	1	1	1	2	1	1	2	2	2	2
P. bungeana	5	1	1	1	2	3	—	2	1	1	2	1	2	1	2	2	1	2	1

Note: — = no amplification; *N* = number of individuals sampled.
[a]Numbers presented for each locus represent number of alleles observed.

LIU, L., Z. Z. HAO, Y. Y. LIU, X. X. WEI, Y. Z. CUN, AND X. Q. WANG. 2014. Phylogeography of *Pinus armandii* and its relatives: Heterogeneous contributions of geography and climate changes to the genetic differentiation and diversification of Chinese white pines. *PLoS ONE* 9: e85920.

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.

ROUSSET, F. 2008. GENEPOP'007: A complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources* 8: 103–106.

VAN OOSTERHOUT, C., W. F. HUTCHINSON, D. P. M. WILLS, AND P. SHIPLEY. 2004. MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* 4: 535–538.

WANG, T., S. Y. REN, Y. CHEN, Z. L. YUAN, L. X. LI, N. PAN, AND Y. Z. YE. 2014. Carbon storage dynamics of *Pinus armandii* forest at different diameter levels based on tree ring data in the Baotianman National Nature Reserve, central China. *Chinese Science Bulletin* 59: 3499–3507.

WILLYARD, A., J. SYRING, D. S. GERNANDT, A. LISTON, AND R. CRONN. 2007. Fossil calibration of molecular divergence infers a moderate mutation rate and recent radiations for *Pinus. Molecular Biology and Evolution* 24: 90–101.

XIONG, P., Z. F. XU, B. LIN, AND Q. LIU. 2010. Short-term response of winter soil respiration to simulated warming in a *Pinus armandii* plantation in the upper reaches of the Minjiang River, China. *Chinese Journal of Plant Ecology* 34: 1369–1376.

YU, F., D. X. WANG, X. F. YI, X. X. SHI, Y. K. HUANG, H. W. ZHANG, AND X. P. ZHANG. 2014. Does animal-mediated seed dispersal facilitate the formation of *Pinus armandii-Quercus aliena* var. *acuteserrata* forests? *PLoS ONE* 9: e89886.

Appendix 1. Voucher information for *Pinus* species used in this study. All vouchers were deposited at the Herbarium of the College of Life Sciences (WNU), Northwest University, Xi'an, China.

Species	Voucher specimen accession no.	Collection locality (Population code)	Geographic coordinates	*N*
P. armandii Franch.	WNU-YT-TB-2014-LZH-022	Yupu town, Tibet Province (YT)	29°37′15″N, 96°18′11″E	16
P. armandii	WNU-NG-SX-2013-LZH-036	Mt. Nangong, Shaanxi Province (NG)	32°13′48″N, 109°1′12″E	17
P. armandii	WNU-XH-SX-2013-LZH-087	Xunhua, Qinghai Province (XH)	35°48′56″N, 102°42′16″E	19
P. koraiensis Siebold & Zucc.	WNU-BS-PK-2013-LZH-049	Baishan, Jilin Province	41°56′24″N, 127°35′24″E	6
P. griffithii McClell.	WNU-JL-PG-2013-LZH-032	Jilong, Tibet Province	28°30′36″N, 85°13′12″E	5
P. sibirica Du Tour	WNU-BJ-PS-2013-LZH-008	Buerjing, Xinjiang Province	48°25′30″N, 86°6′4″E	2
P. pumila (Pall.) Regel	WNU-GH-PP-2013-LZH-003	Genghe, Neimenggu Province	52°21′55″N, 122°28′24″E	2
P. bungeana Zucc. ex Endl.	WNU-WZ-PB-2014-LZH-055	Wuzishan, Shaanxi Province	32°55′59″N, 107°49′59″E	5

Note: LZH = Zhonghu Li, collector; *N* = number of individuals sampled.