Chloroplast and Mitochondrial Microsatellites for Millettia pinnata (Fabaceae) and Cross-Amplification in Related Species

Authors: Wang, Yanling, Xie, Hongxian, Yang, Yi, Huang, Yelin, Wang, Jianwu, et al.

Source: Applications in Plant Sciences, 5(5)

Published By: Botanical Society of America

URL: https://doi.org/10.3732/apps.1700034

Applications in Plant Sciences 2017 5(5): 1700034

PRIMER NOTE

CHLOROPLAST AND MITOCHONDRIAL MICROSATELLITES FOR *MILLETTIA PINNATA* (FABACEAE) AND CROSS-AMPLIFICATION IN RELATED SPECIES[1]

YANLING WANG[2,3], HONGXIAN XIE[4], YI YANG[2,3], YELIN HUANG[4], JIANWU WANG[2,3,5], AND FENGXIAO TAN[2,3,5]

[2]Key Laboratory of Tropical Agro-Environment, Ministry of Agriculture, South China Agricultural University, Guangzhou 510642, Guangdong, People's Republic of China; [3]College of Natural Resources and Environment, South China Agricultural University, Guangzhou 510642, Guangdong, People's Republic of China; and [4]Guangdong Key Laboratory of Plant Resources and State Key Laboratory of Biocontrol, Sun Yat-sen University, Guangzhou 510275, Guangdong, People's Republic of China

- *Premise of the study:* Chloroplast and mitochondrial microsatellites were identified to study the population genetics of *Millettia pinnata* (Fabaceae).
- *Methods and Results:* Based on publicly available plastid genome sequence data of *M. pinnata*, 42 primer pairs were developed, of which 17 displayed polymorphisms across 89 individuals from four populations. For chloroplast loci, two to six alleles were recovered and the unbiased haploid diversity per locus ranged from 0.391 to 0.857. For mitochondrial loci, two to four alleles were recovered and the unbiased haploid diversity ranged from 0.264 to 0.740. Sixteen of the 17 screened markers could be successfully amplified in the related species *M. pulchra*.
- *Conclusions:* The 17 microsatellite markers developed here exhibited variation in *M. pinnata* and 16 presented transferability in the related species *M. pulchra*, suggesting that these markers will be valuable for genetic studies across *M. pinnata* and its related species.

Key words: chloroplast microsatellite; cross-amplification; Fabaceae; *Millettia pinnata*; *Millettia pulchra*; mitochondrial microsatellite.

Millettia pinnata (L.) Panigrahi (syn. *Pongamia pinnata* (L.) Pierre; Fabaceae) is an arboreal legume of the subfamily Papilionoideae and, more specifically, of the tribe Millettieae. According to Scott et al. (2008), this species is widely distributed in the Indian subcontinent and Southeast Asia, extending to Polynesia and northern Australia. Historically, this plant has been used as a source of traditional medicine, animal fodder, green manure, timber, fish poison, and fuel in India and neighboring regions (Satyavati et al., 1987). Seeds of *M. pinnata* contain oils that are inedible but useful for biodiesel, and thus it has received increasing attention as a sustainable biofuel crop in the past decade.

[1]Manuscript received 10 April 2017; revision accepted 24 April 2017.

The authors thank S. He and Y. Liu for their assistance in collecting plant materials. This study was supported by grants from the National Natural Science Foundation of China (31200466 and 41276107), the Natural Science Foundation of Guangdong Province (2015A030313136), the Guangdong Provincial Outstanding Young Teacher Training Foundation (YQ2014030), and the Jin Sui Plan Training Program of the College of Agriculture, South China Agricultural University (20160209 and 20160213).

[5]Authors for correspondence: wangjw@scau.edu.cn (J.W.), fxtan@scau.edu.cn (F.T.)

doi:10.3732/apps.1700034

Population genetic studies on *M. pinnata* have used molecular biological methods including amplified fragment length polymorphisms (AFLPs), three endonuclease (TE)-AFLP (Sharma et al., 2011), inter-simple sequence repeats (ISSRs) (Sahoo et al., 2010; Sharma et al., 2014), and the chloroplast *trnK/matK* and the internal transcribed spacer (ITS) region of nuclear ribosomal DNA (Hu et al., 2000, 2002; Arpiwi et al., 2013). Huang et al. (2016) developed nuclear simple sequence repeat (SSR) markers for *M. pinnata*; however, there is still a lack of plastid (chloroplast and mitochondrial) SSR markers capable of detecting high levels of polymorphism in this species. The uniparentally inherited characteristics of chloroplast and mitochondria can supply information on phylogenetic relationships between individuals because their lineages are not disturbed by recombination (Soranzo et al., 1999). In general, for the maternally inherited chloroplast, the higher sequence variations of SSR loci are distributed throughout the noncoding regions and the flanking regions are conserved (Powell et al., 1995), which makes it possible to monitor the population structure affected by pollen flow and seed-mediated gene flow (Provan et al., 2001). The search for mitochondrial SSR (mtSSR) loci might also be informative, although preliminary studies on plant mitochondrial microsatellites have shown little intraspecific variability (Soranzo et al., 1999). Therefore, plastid SSR markers can be effective for analyzing genetic diversity, population structure, paternity inheritance, and germplasm resource identification (Provan et al., 2001).

In this study, we developed a set of novel SSRs based on publicly available chloroplast and mitochondrial genome sequence data of *M. pinnata* to assess the genetic variation and population genetic structure of this species. Furthermore, we tested the transferability of these markers in the related species *M. pulchra* (Benth.) Kurz.

METHODS AND RESULTS

In this study, the complete chloroplast and mitochondrial genome sequence data of *M. pinnata* were downloaded from the National Center for Biotechnology Information (NCBI) GenBank database (GenBank accession no. JN673818.2 and JN872550.1, respectively). The SSR loci distributed throughout the *M. pinnata* chloroplast and mitochondrial genomes were screened using MISA software (Thiel et al., 2003). The SSR motifs contained one to five nucleotides with the minimum number of repeats as follows: 10 for mononucleotides, five for dinucleotides, four for trinucleotides, and three for tetranucleotides and pentanucleotides. A total of 97 repeat motifs were identified in the chloroplast and mitochondrial genomes, among which the most frequent types were mononucleotides (69 [71.1%]) and dinucleotides (17 [17.5%]), while tri- (5 [5.1%]) and tetranucleotide (6 [6.2%]) motifs were rare. Forty-two loci were selected at random to design primers using Primer3 (Rozen and Skaletsky, 1999), with the optimum conditions set at length of 22 bp (20–26 bp), temperature of 55–60°C, and product size range of 100–500 bp.

Eighty-nine individuals of *M. pinnata* from four natural populations (Appendix 1) were used to evaluate polymorphism of the target microsatellite loci. Genomic DNA from silica-dried leaves was isolated using the advanced cetyltrimethylammonium bromide (CTAB) method (Doyle, 1991). PCR amplifications were performed in a final volume of 15 µL, containing 30 ng of genomic DNA, 1× PCR buffer (10 mM Tris-HCl [pH 8.4] and 1.5 mM $MgCl_2$; TransGen Biotech Co., Beijing, China), 0.2 mM dNTPs (Bocai Biotech Co., Shanghai, China), 0.5 µM of each primer (BGI Sequencing Co., Beijing, China), and 0.5 units *EasyTaq* DNA polymerase (TransGen Biotech Co.). PCR reactions were conducted in a Bio-Rad PTC-200 thermal cycler (Bio-Rad Laboratories, Hercules, California, USA) under the following conditions: initial denaturation at 94°C for 4 min; followed by 35 cycles of 94°C for 1 min, 45 s at the specific annealing temperature for each primer pair (Table 1), and 72°C for 60 s; and a final extension of 10 min at 72°C. PCR products were detected using 1.0% agarose gel electrophoresis to test the utility of the primers. Finally, among the 42 selected primer pairs, 40 were successfully amplified but products from only 17 primer pairs exhibited clear SSR polymorphisms. Six individuals from the related species *M. pulchra* were used to evaluate the transferability of these polymorphic markers applying the 17 screened primers. With the Quant-iT PicoGreen dsDNA Reagent and Kit (including the 35–400-bp Range DNA Ladder; Invitrogen, Carlsbad, California, USA), the Fragment Analyzer Automated CE System (Advanced Analytical Technologies [AATI], Ames, Iowa, USA) was applied to perform SSR genotyping. Raw data were exported, and the number of alleles and allele sizes per locus were called using PROSize software (version 2.0, AATI). All sequences of cpSSR and mtSSR loci were deposited in GenBank, and their accession numbers are presented in Table 1. Because it would be difficult to score the mononucleotide microsatellites consistently, we

Table 1. Characteristics of 17 novel microsatellite markers developed in *Millettia pinnata*.

Locus	Primer sequences (5′–3′)	Repeat motif	T_a (°C)	ES (bp)	Position[a]	GenBank accession no.
POSSRB38[b]	F: TTAAGGAGGCCCCTAATGAAAT R: TTTTAGATACGGGCAGTAGGGA	$(T)_{10}…(A)_{10}$	55	258	*trnG-psaI* IGS	KY189098
POSSRB548[b]	F: ATTAATCGGGGATACACGACAG R: ATTCCGACAACTTCAGGAGAAA	$(T)_{10}…(A)_{13}$	55	191	*ycf3* CDS	KY189099
POSSR22[b]	F: ATTGCAGGTTAACCCCCTTC R: AAAATCGGCGGAGAAGTTTT	$(AT)_7$	55	190	*rpl20* intron	KY189089
POSSR30[b]	F: TCGTCGGTAAATTCAACGGT R: AGGGGATCCCTCTTGTTTTT	$(AT)_7$	55	360	*trnD* intron	KY189092
POSSR44[b]	F: GAATTTGTTTCTTCGTCTTTACAAAA R: GTGGATCAAGGCAGTGGATT	$(AT)_7$	55	173	*trnH* intron	KY189093
POSSR81[b]	F: GCTCCGTTCCATGTCTCATT R: TGAAAAATCATCGCAAACCTC	$(TA)_8$	55	275	*ycf1* intron	KY189096
POSSRB253[b]	F: AATTAGGCTCGATCAACTGGAA R: CTGTTCCTATCTAACGGAACGC	$(T)_{12}$	56	242	*ycf2-ycf15* IGS	KY411953
POSSRB426[b]	F: ATTCAGTATCCTGCCACGAAAT R: CGACCTTGCATTTTTCTACCTC	$(A)_{10}$	56	175	*ycf2-ycf15* IGS	KY411954
POSSRS6[b]	F: CTGTTCCTATCTAACGGAACGC R: AATTAGGCTCGATCAACTGGAA	$(A)_{12}$	56	242	*ycf2-ycf15* IGS	KY411955
POSSRS89[b]	F: TTCGTGATTCCTTGGTAAATCC R: AGGCCATTTTATCGACATGAGT	$(A)_{10}$	56	203	*atpF* CDS	KY411956
POSSRS217[b]	F: GATTACGATAAATTGGATCGGC R: ATCTTTCGAGATCCACCCTACA	$(T)_{11}$	56	168	*rps7* intron	KY411957
POSSRS225[b]	F: TTTCTTCCAACATAACAACCCC R: AAAGCGAGTCGACCACTAGAAC	$(T)_{15}$	56	152	*rps18* intron	KY411958
POSSRB218[c]	F: CGTTCCTTTTCTCTTGGACATC R: TGCATGGGTCTTCCTTTCTTAT	$(TATTA)_5$	55	168	*atp9* intron	KY189090
POSSRB310[c]	F: GAGTTTGAATTGACCCGGTTAG R: CTAGGCAGCTATCGGAAAGAAA	$(CT)_6$	55	242	*trnD-trnI* IGS	KY189091
POSSR53[c]	F: AGGCTGCTAGGAAGGAGGAC R: GGCACCATTTCAGAGAGGAG	$(TA)_6$	55	175	*ccmC-rps3* IGS	KY189094
POSSR61[c]	F: CGGTGAAGTACCCCCTTACA R: GCTTCGGTGTAAGCACCAGT	$(GA)_6$	55	242	*nad1-atp1* IGS	KY189095
POSSR95[c]	F: GTATGTACCCCGATTCGCAC R: GTCTAGGCTGATTTGGCAGG	$(CAAA)_5$	55	203	*Rrn26-mttB* IGS	KY189097

Note: *A* = number of alleles; CDS = coding sequence; ES = expected band size; IGS = intergenic spacer; T_a = annealing temperature.

[a]Position of each SSR is according to the NCBI sequence information of *M. pinnata* (GenBank accession no. JN673818.2 for the complete chloroplast genome and JN872550.1 for the complete mitochondrial genome).

[b]Loci located in the chloroplast.

[c]Loci located in the mitochondria.

TABLE 2. Characterization of 17 novel microsatellite markers in populations of *Millettia pinnata* and *M. pulchra.*[a]

	Millettia pinnata																*Millettia pulchra*			
	POPH (N = 26)				POLS (N = 25)				POCN (N = 14)				PONS (N = 24)				MVSD (N = 6)			
Locus	A	A_e	I	h_{unb}	A	A_e	I	h_{unb}	A	A_e	I	h_{unb}	A	A_e	I	h_{unb}	A	A_e	I	h_{unb}
POSSRB38[b]	4	3.159	1.209	0.711	3	2.323	0.944	0.593	3	2.882	1.079	0.703	4	3.032	1.190	0.699	3	2.571	1.011	0.733
POSSRB548[b]	5	4.225	1.487	0.794	6	4.496	1.614	0.810	3	2.579	1.004	0.659	5	3.945	1.452	0.779	3	2.000	0.868	0.600
POSSR22[b]	4	2.840	1.185	0.674	3	1.947	0.779	0.507	2	1.690	0.598	0.440	2	1.600	0.562	0.391	2	1.800	0.637	0.533
POSSR30[b]	4	3.314	1.266	0.726	3	2.462	0.974	0.620	2	1.960	0.683	0.527	3	1.646	0.675	0.409	2	2.000	0.693	0.600
POSSR44[b]	3	2.828	1.068	0.673	2	1.771	0.627	0.453	2	1.690	0.598	0.440	3	2.743	1.051	0.663	3	2.000	0.868	0.600
POSSR81[b]	4	2.704	1.170	0.655	4	3.307	1.290	0.727	2	1.849	0.652	0.495	4	2.969	1.200	0.692	2	1.800	0.637	0.533
POSSRB253[b]	4	2.704	1.170	0.655	3	1.947	0.779	0.507	4	2.390	1.055	0.626	4	3.245	1.228	0.723	3	2.571	1.011	0.733
POSSRB426[b]	5	3.347	1.345	0.729	4	2.376	1.085	0.603	4	2.882	1.195	0.703	6	5.434	1.738	0.851	2	1.800	0.637	0.533
POSSRS6[b]	5	2.522	1.173	0.628	6	5.631	1.760	0.857	4	3.267	1.277	0.747	6	3.556	1.501	0.750	2	2.000	0.693	0.600
POSSRS89[b]	4	3.414	1.307	0.735	3	2.828	1.068	0.673	4	3.379	1.272	0.758	4	2.717	1.179	0.659	3	2.571	1.011	0.733
POSSRS217[b]	5	3.453	1.401	0.740	5	3.307	1.315	0.727	4	3.267	1.254	0.747	5	3.550	1.397	0.751	2	1.800	0.637	0.533
POSSRS225[b]	4	3.073	1.220	0.702	5	3.342	1.356	0.730	5	3.500	1.390	0.769	5	3.789	1.471	0.768	2	1.471	0.500	0.400
POSSRB218[c]	4	3.414	1.307	0.735	2	1.676	0.593	0.420	2	1.324	0.410	0.264	3	2.743	1.051	0.663	3	2.571	1.011	0.733
POSSRB310[c]	3	2.770	1.058	0.665	3	1.900	0.807	0.493	2	1.960	0.683	0.527	3	2.323	0.960	0.594	2	1.800	0.637	0.533
POSSR53[c]	4	2.965	1.174	0.689	4	2.706	1.111	0.657	3	2.579	1.004	0.659	3	1.805	0.739	0.466	2	1.800	0.637	0.533
POSSR61[c]	3	2.104	0.901	0.547	3	2.990	1.097	0.693	2	1.960	0.683	0.527	3	2.667	1.028	0.652	—	—	—	—
POSSR95[c]	4	3.453	1.309	0.740	2	1.923	0.673	0.500	3	2.085	0.892	0.560	4	3.097	1.238	0.707	3	2.000	0.868	0.600

Note: — = not available; A = number of alleles per population; A_e = number of effective alleles per population; h_{unb} = unbiased haploid diversity; I = Shannon's information index; N = number of individuals analyzed.

[a]Voucher and locality information are provided in Appendix 1.

[b]Loci located in the chloroplast.

[c]Loci located in the mitochondria.

conducted direct sequencing of the PCR products using the BigDye Terminator v3.1 Cycle Sequencing Kit and ABI PRISM 3730 sequencer (Applied Biosystems, Waltham, Massachusetts, USA) with both forward and reverse primers, to verify the allelic variants tested in this study. Sequences were visualized and analyzed with the DNASTAR software package (DNASTAR, Madison, Wisconsin, USA).

The 17 selected primers exhibited high polymorphisms across 89 individuals of four *M. pinnata* populations. For each of these loci, the number of alleles per population, the number of effective alleles per population, Shannon's information index, and the unbiased haploid diversity (h_{unb}) of each microsatellite locus were calculated using GenAlEx version 6.5 (Peakall and Smouse, 2012). Among chloroplast loci, the number of alleles per locus per population varied from two to six, while h_{unb} ranged from 0.391 to 0.857. For mitochondrial loci, two to four alleles per locus per population were detected and h_{unb} ranged from 0.264 to 0.740 (Table 2). Subsequently, 16 of the 17 developed markers were successfully amplified in the related species *M. pulchra*, demonstrating their transferability (Table 2).

CONCLUSIONS

The 17 polymorphic SSR markers developed here proved useful in the evaluation of the genetic diversity of *M. pinnata*, and 16 showed high transferability within the related species *M. pulchra*. This set of novel polymorphic SSR markers will serve as a very useful tool for the genetic diversity analysis, clonal identification, and germplasm conservation of *M. pinnata* and its related species.

LITERATURE CITED

Arpiwi, N. L., G. Yan, E. L. Barbour, and J. A. Plummer. 2013. Genetic diversity, seed traits and salinity tolerance of *Millettia pinnata* (L.) Panigrahi, a biodiesel tree. *Genetic Resources and Crop Evolution* 60: 677–692.

Doyle, J. J. 1991. DNA protocols for plants. *In* G. M. Hewitt, A. W. B. Johnston, and J. P. W. Young [eds.], Molecular techniques in taxonomy, 283–293. Springer-Verlag, Berlin, Germany.

Hu, J. M., M. Lavin, M. F. Wojciechowski, and M. J. Sanderson. 2000. Phylogenetic systematics of the tribe Millettieae (Leguminosae) based on chloroplast *trnK/matK* sequences and its implications for evolutionary patterns in Papilionoideae. *American Journal of Botany* 87: 418–430.

Hu, J. M., M. Lavin, M. F. Wojciechowski, and M. J. Sanderson. 2002. Phylogenetic analysis of nuclear ribosomal ITS/5.8S sequences in the tribe Millettieae (Fabaceae): *Poecilanthe-Cyclolobium*, the core Millettieae, and the *Callerya* group. *Systematic Botany* 27: 722–733.

Huang, J. Z., X. H. Guo, X. H. Hao, W. K. Zhang, S. Y. Chen, R. F. Huang, P. M. Gresshoff, and Y. Z. Zheng. 2016. De novo sequencing and characterization of seed transcriptome of the tree legume *Millettia pinnata* for gene discovery and SSR marker development. *Molecular Breeding* 36: 75.

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.

Powell, W., M. Morgante, C. Andre, J. W. McNicol, G. C. Machray, J. J. Doyle, S. V. Tingey, and J. A. Rafalski. 1995. Hypervariable microsatellites provide a general source of polymorphic DNA markers for the chloroplast genome. *Current Biology* 5: 1023–1029.

Provan, J., W. Powell, and P. M. Hollingsworth. 2001. Chloroplast microsatellites: New tools for studies in plant ecology and evolution. *Trends in Ecology & Evolution* 16: 142–147.

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.

Sahoo, D. P., S. Aparajita, and G. R. Rout. 2010. Inter and intra-population variability of *Pongamia pinnata*: A bioenergy legume tree. *Plant Systematics and Evolution* 285: 121–125.

Satyavati, G. V., A. K. Gupta, and N. Tandon. 1987. Medicinal plants of India, vol. 2, p. 490. India Council of Medicinal Research, New Delhi, India.

Scott, P. T., L. Pregelj, N. Chen, J. S. Hadler, M. A. Djordjevic, and P. M. Gresshoff. 2008. *Pongamia pinnata*: An untapped resource for the biofuels industry of the future. *BioEnergy Research* 1: 2–11.

Sharma, S. S., M. S. Negi, P. Sinha, K. Kumar, and S. B. Tripathi. 2011. Assessment of genetic diversity of biodiesel species *Pongamia pinnata* accessions using AFLP and Three Endonuclease-AFLP. *Plant Molecular Biology Reporter* 29: 12–18.

Sharma, S. S., M. A. Islam, M. S. Negi, and S. B. Tripathi. 2014. Isolation and characterization of a first set of nine polymorphic microsatellite loci in *Pongamia pinnata* (Fabaceae). *Journal of Genetics* 93: e70–e74.

Soranzo, N., J. Provan, and W. Powell. 1999. An example of microsatellite length variation in the mitochondrial genome of conifers. *Genome* 42: 158–161.

Thiel, T., W. Michalek, R. K. Varshney, and A. Graner. 2003. Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). *Theoretical and Applied Genetics* 106: 411–422.

Appendix 1. Sampling information for the populations of *Millettia pinnata* and *M. pulchra* used in this study.

Species	Population code	Collection locality	Geographic coordinates	*N*	Voucher no.[a]
Millettia pinnata (L.) Panigrahi	POPH	Perhentian Islands, Malaysia	5°54′34.12″N, 102°44′42.43″E	26	Liu20150926
	POLS	Laem Son, Thailand	6°56′21.10″N, 99°42′43.64″E	25	Huang20160701
	POCN	Chilaw, Sri Lanka	7°33′43.16″N, 79°48′5.96″E	14	Liu20161001
	PONS	Tainan, Taiwan, China	23°1′1.2″N, 120°7′1.20″E	24	He20121208
Millettia pulchra (Benth.) Kurz	MVSD	Heishiding, Guangdong, China	23°27′43.91″N, 111°54′32.37″E	6	Shi20161210

Note: *N* = number of individuals sampled.

[a]All voucher specimens are deposited at the herbarium of Sun Yat-sen University (SYS), Guangzhou, China.