Characterization of Polymorphic Microsatellite Markers for Primula sikkimensis (Primulaceae) Using a 454 Sequencing Approach

Authors: Li, Chang-Han, Liu, Yun-Jiao, Zhang, Cai-Yun, Yan, Hai-Fei, Ge, Xue-Jun, et al.

Source: Applications in Plant Sciences, 4(7)

Published By: Botanical Society of America

URL: https://doi.org/10.3732/apps.1600015

Applications in Plant Sciences 2016 4(7): 1600015

PRIMER NOTE

CHARACTERIZATION OF POLYMORPHIC MICROSATELLITE MARKERS FOR *PRIMULA SIKKIMENSIS* (PRIMULACEAE) USING A 454 SEQUENCING APPROACH[1]

CHANG-HAN LI[2,4], YUN-JIAO LIU[2,4], CAI-YUN ZHANG[2], HAI-FEI YAN[3], XUE-JUN GE[3], AND GANG HAO[2,5]

[2]College of Life Sciences, South China Agricultural University, Guangzhou 510642, People's Republic of China; and [3]Key Laboratory of Plant Resources Conservation and Sustainable Utilization, South China Botanical Garden, Chinese Academy of Sciences, Guangzhou 510650, People's Republic of China

- *Premise of the study:* Microsatellite markers from *Primula sikkimensis* (Primulaceae) were developed for testing deep lineage divergence and speciation events.
- *Methods and Results:* A total of 3112 microsatellites were identified from 61,755 unique reads though 454 pyrosequencing technology. Twenty-nine microsatellite loci were selected for PCR amplification and polymorphic analyses. Among the 29 tested markers, 17 microsatellite loci were further used for genotyping in three wild *P. sikkimensis* populations. The number of alleles varied from one to eight, and the observed heterozygosity ranged from 0.111 to 1.000. Ten simple sequence repeat loci could be successfully cross-amplified in two *Primula* species. The transferability values were 76.5% in *P. florindae* and 58.8% in *P. alpicola*, respectively.
- *Conclusions:* These microsatellite markers will be valuable for testing the hypothesis of lineage divergence, genetic introgression, and cryptic speciation events between *P. sikkimensis* and its closely related taxa.

Key words: cross-amplification; deep lineage divergence; genetic introgression; microsatellites; *Primula sikkimensis*; Primulaceae.

The Himalayan region and the adjacent Hengduan Mountains of southwestern China, known as the Himalaya–Hengduan Mountains (HHM) region, have been designated as two of the world's 34 most important biodiversity hotspots (Myers et al., 2000). The HHM region is considered to be the cradle of many endemic plant groups (Li and Li, 1993) and the center for rapid radiation of several large alpine genera, such as *Primula* L., *Pedicularis* L., and *Rhododendron* L., as well as the center of the Sino-Himalayan floristic subkingdom (Wu and Wang, 1983). Its high species endemism is a likely product of high net diversification rates in the region, as seen in páramo hotspots evaluated by Madriñán et al. (2013). A number of studies have been devoted to the differences between the two parts of the HHM region (the Himalayas and the Hengduan Mountains), such as the direction of the mountain ranges, the time scale of the Qinghai–Tibet plateau (QTP) uplift process, and the effects of climate oscillations during the Quaternary (Favre et al., 2015). Correspondingly, the Sino-Himalayan floristic subkingdom in the HHM region has been recognized as including at least four subregions (Wu et al., 2011). However, it is not clear whether these differences between the Hengduan Mountains and the Himalayan regions have resulted in deep intraspecific lineage divergences and/or cryptic speciation in plant groups.

Primula sikkimensis Hook. (Primulaceae) is an endemic species in the HHM region (Hu and Kelso, 1996) and is the only species in *Primula* sect. *Sikkimensis* that is widely distributed in the region. It therefore provides a good example to examine the hypothesis of deep lineage divergence between the Himalaya and Hengduan mountains (Gao et al., 2007). Here, we developed a set of variable microsatellite markers using 454 pyrosequencing technology and further tested its cross-amplification in closely related taxa. These microsatellite markers will be important tools for surveying genetic divergence and cryptic speciation events in *P. sikkimensis* and its relatives.

METHODS AND RESULTS

Leaf samples of 62 individuals were collected in three populations from Chayu, Galongla, and Luding in China (Appendix 1). One individual of *P. sikkimensis* (sampled from Jiulong, China; Appendix 1) was used to isolate the microsatellite loci. Voucher specimens have been deposited at the herbarium of the South China Botanical Garden (IBSC), Guangzhou, Guangdong, China. Total DNA extraction of all samples was performed using a modified version of the cetyltrimethylammonium bromide (CTAB) protocol of Doyle and Doyle (1987). Microsatellite markers were isolated using a high-throughput genomic sequencing method as described by Wang et al. (2015). A shotgun library shearing 1 µg of genomic DNA was built using the DNA Library Preparation Kit (Roche Applied Science, Indianapolis, Indiana, USA) following the GS FLX+ library preparation protocol. The library was further enriched by hybridization with biotinylated

[1]Manuscript submitted 10 February 2016; revision accepted 16 March 2016.

The authors thank Xu Yuan and Wang Zheng-Feng for providing plant material and helping in data analyses. The project was supported by the National Natural Science Foundation of China (31500173) and the Guangdong Natural Science Foundation (2014A030310120).

[4]These authors contributed equally to this work.

[5]Author for correspondence: haogang@scau.edu.cn

doi:10.3732/apps.1600015

oligonucleotide probes $(AG)_{10}$, $(AC)_{10}$, $(AAC)_8$, $(ACG)_8$, $(AAG)_8$, $(AGG)_8$, $(ACAT)_6$, and $(ATCT)_6$ by Tóth et al. (2000) and Zane et al. (2002). The simple sequence repeat (SSR)–enriched libraries were then sequenced using a Roche 454 GS FLX DNA sequencing platform. In total, 61,755 unique reads were obtained with sizes ranging from 300 to 600 bp.

Microsatellite repeats in unique reads were identified by MISA software (Thiel et al., 2003). The SSR search was performed for di-, tri-, and tetranucleotides with a minimum of six, five, and five repeats, respectively, and a minimum product size of 100 bp. In total, 5377 unique reads with at least one microsatellite motif were obtained. Among these reads, 3112 unique reads, which had at least 50 bp in each flanking region for primer design, were chosen to filter the perfect SSR loci (sequences in these reads are available upon request). Then 29 loci were randomly selected to design primer pairs using Primer3 software (Rozen and Skaletsky, 1999). The minimum primer annealing temperature was set to 60°C, primer size was between 18–22 bp with an optimal size of 20 bp, and other settings were left at default values.

These primer pairs were initially tested for successful PCR amplification in three *P. sikkimensis* individuals from three separate populations. PCR reactions were performed on a PTC-200 Thermal Cycler (MJ Research, Watertown, Massachusetts, USA) with the following conditions: an initial denaturation at 94°C for 3 min; followed by 30 cycles at 94°C for 30 s, locus-specific annealing temperature (Table 1) for 45 s, and 72°C for 50 s; and a final extension at 72°C for 7 min. Amplicons were checked on 2% agarose gel stained with ethidium bromide.

In total, 20 primer pairs that generated specific amplification of corresponding PCR products were further resynthesized using fluorophore labeling (FAM or HEX) and used for amplification in the 62 individuals from the three populations. The same PCR conditions were used as described above. One microliter of the fluorescent PCR product was added into the mixture with 8.8 μL of formamide and 0.2 μL of GeneScan 500 LIZ Size Standard (Applied Biosystems, Life Technologies, Waltham, Massachusetts, USA). PCR products were subsequently run on an ABI PRISM 3130 Genetic Analyzer (Applied Biosystems). Genotypes were scaled by GelQuest software (version 3.2.1; SequentiX, Klein Raden, Germany). Seventeen of the 20 primers showed clear and robust genotype information. The microsatellite information and GenBank accession numbers are listed in Table 1.

Genetic diversity parameters, including allelic richness (*A*), observed and unbiased expected heterozygosity (H_o, H_e), and inbreeding coefficient (F_{IS}), were estimated by GenAlEx 6.5 (Peakall and Smouse, 2012). Deviations from Hardy–Weinberg equilibrium (HWE) at each locus were tested through GENEPOP 4.0.7 (Rousset, 2008) and are presented in Table 2. Numbers of alleles varied from one to eight; H_o and H_e ranged from 0.111 to 1.000 and 0.061 to 0.811, respectively; and F_{IS} ranged from −1.000 to 0.660. Twelve loci showed significant deviation from expectations under HWE (Table 2) because of an excess of homozygotes. Null alleles, inbreeding, Wahlund effect, and sampling effect (small population size) could all potentially cause deviations from HWE. Four loci with presence of null alleles were detected by MICRO-CHECKER (van Oosterhout et al., 2004).

In addition, we tested cross-amplification with two related species in sect. *Sikkimensis* (*P. alpicola* (W. W. Sm.) Stapf and *P. florindae* Kingdon-Ward). One individual of *P. alpicola* was sampled from Paizhen, Tibet (29°19′N, 95°19′E), and one individual of *P. florindae* was collected at Lulang, Tibet (29°42′N, 94°43′E) (Appendix 1). Primer transferability was considered successful when one clear distinct band in the expected size range was detected on 2% agarose. In total, 10 SSR loci could be successfully used in both *P. alpicola* and *P. florindae*, and only four loci could not be amplified in these two species. Specifically, the transferability values were 76.5% in *P. florindae* and 58.8% in *P. alpicola*, respectively.

CONCLUSIONS

In this study, 17 microsatellite markers were successfully developed for *P. sikkimensis*; these markers showed high polymorphism and could therefore be a powerful tool in population

Table 1. Characteristics of 17 microsatellite loci developed in *Primula sikkimensis*.

Locus	Primer sequences (5′–3′)	Repeat motif	Fluorescent dye	Allele size range (bp)	T_a (°C)	GenBank accession no.
S1	F: CCCTGTTCCAAGATTTGGTG R: AACTATCTGGCATGGATGGTC	$(AG)_{24}$	HEX	148–168	54	KU697616
S3	F: ATCTGTGTCCCAAACAACCC R: CCAAACAACCAAACAAGCCT	$(CTT)_{14}$	FAM	228–288	60	KU697617
S4	F: GCTCCTGATGGGTATTACGG R: CGCACACGGTTACTGTTTTG	$(AAC)_{13}$	HEX	148–191	60	KU697618
S5	F: GGGAGACCGATGGTTAAGGT R: CGTCGGTGTTGGTCCTCTAT	$(AG)_{18}$	HEX	106–124	59.8	KU697619
S9	F: CGGGTAGAGAGACAGCGTTC R: CCCTAGATCTCCAGCGAGTG	$(GA)_{17}$	HEX	124–131	60	KU697620
S12	F: ACTGCTCGATGATGGTTTCC R: ATGTTTCCGGACTGTTTCAA	$(GT)_{16}$	HEX	156–178	60	KU697621
S13	F: CAAAGACTCATTGACAACCGT R: GGCGGCTAATCTTGTGTAGG	$(AG)_{16}$	FAM	274–300	59.8	KU697622
S14	F: GACATGAAGAAACTGGAGACGA R: CGCTATGGCCGGTTATCTTA	$(AG)_{16}$	HEX	98–122	60	KU697623
S15	F: GATTGAGGAATGCGCAAAAT R: AAGCACTTGAGTTAAGCTAGCCA	$(CA)_{16}$	FAM	272–290	62.3	KU697624
S16	F: GCCAATACACACCTTCCACC R: TCTTAATGGGGGTTCTGGC	$(AG)_{16}$	FAM	260–282	60	KU697625
S17	F: AGGGGCATTTTGGTCATTTA R: GGGTAGCCGTCTCTCTCTCC	$(AC)_{15}$	HEX	118–130	64.9	KU697626
S18	F: CGTAAGGGTGCTTAAGCTGG R: GTCAAATGGCGTCGTATGTG	$(GT)_{15}$	HEX	160–182	60	KU697627
S21	F: GATTTGCAATAGCGAGAGCC R: GAGAGAGAGGCAGGCGAAC	$(CT)_{14}$	FAM	232–250	60	KU697628
S22	F: AAAGGGGAAGTCAGACGGTT R: CCGCCTTTTCTCCTCTCTCT	$(AG)_{12}$	HEX	138–158	60	KU697629
S23	F: TCATTTCGTTAAATTTTGTTTTCG R: CAAAATTGGGAGAGGCATGT	$(AC)_{16}$	FAM	198–218	59.8	KU697630
S24	F: AAGATCGACCCACGATCAAT R: TGTTTGATGTCGCGGTAACT	$(AG)_{15}$	FAM	252–268	60	KU697631
S29	F: GGGCATTTTGGTCATTTCAC R: GTGGTGGTGGCTGTTTCTCT	$(AC)_{11}$	FAM	204–260	60	KU697632

Note: T_a = annealing temperature.

TABLE 2. Estimated population genetic parameters (per nSSR locus) in three populations of *Primula sikkimensis* and cross-amplification in two related species.[a]

	Primula sikkimensis												Cross-amplification	
	XZCY (N = 18)			SCLD (N = 16)			XZGLL (N = 28)							
Locus	A	H_o	H_e	F_{IS}	A	H_o	H_e	F_{IS}	A	H_o	H_e	F_{IS}	Mean A	*Primula florindae*	*Primula alpicola*
S1	5	0.389	0.407	0.074	5	0.625	0.662	0.088	6	0.571	0.671*	0.166	8	+	+
S3	4	0.333	0.724*	0.559	8	0.438	0.711*	0.412	4	0.556	0.444	–0.234	10	+	+
S4	4	0.889	0.620*	–0.409	5	0.375	0.420*	0.139	4	0.607	0.470	–0.275	11	+	+
S5	2	0.278	0.239	–0.133	2	0.063	0.061	NA	2	0.222	0.198	–0.106	4	+	—
S9	3	0.056	0.156*	0.660	1	0.000	0.000	NA	5	0.286	0.413*	0.324	9	+	+
S12	5	0.500	0.452	–0.078	3	0.438	0.510	0.173	4	0.571	0.652*	0.141	7	+	+
S13	3	0.722	0.526	–0.348	2	0.188	0.170	–0.071	2	0.815	0.483*	–0.677	5	—	—
S14	3	0.500	0.508	0.044	6	0.750	0.715	–0.017	3	0.500	0.530	0.075	9	+	+
S15	2	0.611	0.486	–0.230	3	1.000	0.529*	–0.882	4	0.185	0.511*	0.649	8	+	+
S16	2	0.111	0.198	0.460	3	0.125	0.320*	0.630	5	0.643	0.735*	0.143	8	+	+
S17	2	0.563	0.482	–0.135	2	0.867	0.491*	–0.750	2	0.321	0.270	–0.174	3	—	—
S18	2	0.333	0.278	–0.172	5	1.000	0.648*	–0.519	2	0.286	0.245	–0.149	6	+	—
S21	1	NA	NA	NA	2	1.000	0.500*	–1.000	5	0.381	0.714*	0.486	6	—	—
S22	5	1.000	0.628*	–0.573	6	1.000	0.811	–0.203	4	0.714	0.504*	–0.401	9	+	+
S23	3	0.333	0.364	0.113	7	0.563	0.781*	0.310	2	0.357	0.459	0.239	11	+	—
S24	3	0.111	0.106	–0.015	2	0.625	0.469	–0.304	1	NA	NA	NA	4	+	+
S29	7	0.833	0.650	–0.256	2	0.750	0.469	–0.579	2	0.357	0.337	–0.043	10	—	—
Mean	3.294	0.445	0.401	–0.053	3.765	0.577	0.486	–0.186	3.353	0.434	0.449	–0.006	8.000		

Note: + = successful PCR amplification; — = unsuccessful PCR amplification; A = number of alleles per locus; F_{IS} = fixation index; H_e = expected heterozygosity; H_o = observed heterozygosity.

[a]Locality and voucher information are available in Appendix 1.

*Significant deviation from Hardy–Weinberg equilibrium ($P < 0.05$).

genetic studies. Cross-amplification of these microsatellite loci in two related *Primula* species (*P. alpicola* and *P. florindae*) was successful, which enables further studies to clarify underlying genetic introgression and cryptic speciation events between *P. sikkimensis* and its closely related taxa.

LITERATURE CITED

DOYLE, J. J., AND J. L. DOYLE. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* 19: 11–15.

FAVRE, A., M. PÄCKERT, S. U. PAULS, S. C. JÄHNIG, D. UHL, I. MICHALAK, AND A. N. MUELLNER-RIEHL. 2015. The role of the uplift of the Qinghai-Tibetan Plateau for the evolution of Tibetan biotas. *Biological Reviews* 90: 236–253.

GAO, L. M., M. MÖLLER, X. M. ZHANG, M. L. HOLLINGSWORTH, J. LIU, R. R. MILL, M. GIBBY, AND D. Z. LI. 2007. High variation and strong phylogeographic pattern among cpDNA haplotypes in *Taxus wallichiana* (Taxaceae) in China and North Vietnam. *Molecular Ecology* 16: 4684–4698.

HU, C. M., AND S. KELSO. 1996. Primulaceae. *In* Z. Y. Wu and P. H. Raven [eds.], Flora of China, vol. 15: Myrsinaceae through Loganiaceae. Science Press, Beijing, China, and Missouri Botanical Garden Press, St. Louis, Missouri, USA.

LI, X. W., AND J. LI. 1993. A preliminary floristic study on the seed plants from the region of Hengduan Mountain. *Acta Botanica Yunnanica* 15: 217–231.

MADRIÑÁN, S., A. J. CORTÉS, AND J. E. RICHARDSON. 2013. Páramo is the world's fastest evolving and coolest biodiversity hotspot. *Frontiers in Genetics* 4: 192.

MYERS, N., R. A. MITTERMEIER, C. G. MITTERMEIER, G. A. B. DA FONSECA, AND J. KENT. 2000. Biodiversity hotspots for conservation priorities. *Nature* 403: 853–858.

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.

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.

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.

TÓTH, G., Z. GÁSPÁRI, AND J. JURKA. 2000. Microsatellites in different eukaryotic genomes: Survey and analysis. *Genome Research* 10: 967–981.

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

WANG, X., J. LI, AND Y. LI. 2015. Isolation and characterization of microsatellite markers for an endemic tree in East Asia, *Quercus variabilis* (Fagaceae). *Applications in Plant Sciences* 3: 1500032.

WU, Z. Y., AND H. S. WANG. 1983. Plant geography of China. Science Press, Beijing, China.

WU, Z. Y., AND H. SU, Z. K. ZHOU, D. Z. LI, AND H. PENG. 2011. Floristics of seed plants from China. Science Press, Beijing, China.

ZANE, L., L. BARGELLONI, AND T. PATARNELLO. 2002. Strategies for microsatellite isolation: A review. *Molecular Ecology* 11: 1–16.

Appendix 1. Locality and voucher information for *Primula* individuals used in this study. Voucher specimens are deposited at the herbarium of the South China Botanical Garden (IBSC), Guangzhou, Guangdong, China.

Species	Population code	Collection locality	Geographic coordinates	Voucher no.
Primula sikkimensis	XZCY	Chayu, Tibet	27°00′N, 100°10′E	*Hao 934*
	SCLD	Luding, Sichuan	29°55′N, 102°3′E	*Hao 456*
	XZGLL	Galongla, Tibet	29°16′N, 95°05′E	*Wuxing s.n.*
	—	Jiulong, Sichuan	29°0′N, 101°30′E	Y2014163
Primula alpicola	—	Paizhen, Tibet	29°19′N, 95°19′E	*Hao & Xu 120195*
Primula florindae	—	Lulang, Tibet	29°42′N, 94°43′E	*Hao & Xu 120281*