Development of SSR Markers for the Tropical Alpine Tree Species Leptospermum recurvum (Myrtaceae) on Mount Kinabalu in Borneo

Authors: Ando, Soichi, Kaneko, Shingo, Isagi, Yuji, Repin, Rimi, and Kitayama, Kanehiro

Development of SSR markers for the tropical alpine tree species *Leptospermum recurvum* (Myrtaceae) on Mount Kinabalu in Borneo[1]

Soichi Ando[2,5], Shingo Kaneko[3], Yuji Isagi[2], Rimi Repin[4], and Kanehiro Kitayama[2]

[2]Graduate School of Agriculture, Kyoto University, Kitashirakawa Oiwake-cho, Sakyo-ku, Kyoto 606-8502, Japan; [3]Faculty of Symbiotic Systems Science, Fukushima University, 1 Kanayagawa, Fukushima 960-1296, Japan; and [4]Sabah Parks, P.O. Box 10626, 88806 Kota Kinabalu, Malaysia

- *Premise of the study:* Nuclear microsatellite (simple sequence repeat [SSR]) markers were developed for the woody species *Leptospermum recurvum* found on Mount Kinabalu, Borneo, to facilitate investigation of the genetic structure and patterns of gene flow in relation to leaf phenotypic polymorphisms.
- *Methods and Results:* Eleven primer pairs were developed using the compound SSR marker technique. Ten of the 11 loci were polymorphic and their expected heterozygosity ranged from 0.04 to 0.83. Neither linkage disequilibrium nor departure from Hardy–Weinberg equilibrium were detected. All primer pairs also amplified the SSR loci of *L. polygalifolium*.
- *Conclusions:* These findings suggest the utility of these primers for investigating genetic structure and gene flow in *L. recurvum* and indicate their applicability to another species of *Leptospermum*.

Key words: *Leptospermum recurvum*; Mount Kinabalu; Myrtaceae; simple sequence repeat (SSR) markers.

Leptospermum recurvum Hook. f. (Myrtaceae) is a woody species distributed from 2700 to 4000 m on Mount Kinabalu in Borneo. This species has a wide habitat range from deep soils to denuded rocks, and varies in life form from a tree to a dwarf shrub to a cushion plant. All forms have continuous phenotypic polymorphisms with regard to leaf traits. These phenotypes often occur in the same place, but the composition of phenotypes varies among habitats. Elucidating the genetic structure and patterns of gene flow across phenotypes and habitats is the first step to understanding the evolutionary process behind this phenotypic differentiation. Lee and Lowry (1980) conducted an allozyme analysis, but they detected few polymorphisms among the five populations present in the distribution range despite large phenotypic variations. In such a situation, highly polymorphic simple sequence repeat (SSR) markers are required to determine genetic characteristics. Here, we present 10 polymorphic and one monomorphic SSR markers of *L. recurvum* and also describe successful amplification of the related species *L. polygalifolium* Salisb.

METHODS AND RESULTS

Prior to development of the markers, we collected seeds of *L. recurvum* from a population located at an elevation of 3300 m on Mount Kinabalu (6.05824°N, 116.56566°E) and cultivated the seedlings in our laboratory. Fresh leaves were

sampled from one of the seedlings for DNA extraction to construct an SSR library. The voucher specimen of this seedling was deposited in the Sabah Parks Herbarium at Kinabalu Park headquarters (KIN accession no. SNP28526). We developed SSR markers using an improved technique for compound SSR markers (Lian et al., 2006; Araki et al., 2011). An adapter-ligated DNA library was constructed as described previously (Lian et al., 2001). The fragments were amplified from an *Eco*RV (TaKaRa Biotechnology Co., Otsu, Japan) DNA library using an adapter primer, AP2 (5′-CTATAGGGCACGCGTGGT-3′), and a compound SSR primer, $(AC)_6(AG)_7$ or $(AC)_6(TC)_7$. The amplified fragments of 400–1000 bp were purified using a QIAquick Gel Extraction Kit (QIAGEN, Valencia, California, USA) and cloned using a QIAGEN PCR Cloning^{plus} Kit (QIAGEN) in accordance with the manufacturer's instructions. A total of 224 insert-positive clones were sequenced. We identified 33 distinct sequences containing an $(AC)_6(AG)_n$ or $(AC)_6(TC)_n$ compound SSR sequence at one end using ClustalX version 2.1 (Larkin et al., 2007). Based on these 33 sequences, specific primers were designed using Primer3 version 0.4.0 (Rozen and Skaletsky, 2000).

PCR amplifications were performed in a final volume of 5 μL using a QIAGEN Multiplex PCR Kit (QIAGEN) according to the standard protocol with a GeneAmp PCR System 2700 thermal cycler (Applied Biosystems, Carlsbad, California, USA). The combination of multiplex reactions is shown in Table 1. Compound SSR primers were labeled with fluorochromes, as shown in Table 1. PCR amplifications were performed under different cycling conditions based on the compound SSR primers: for the primers labeled with VIC or 6-FAM, initial denaturation was performed at 95°C for 15 min followed by 25 cycles of denaturation at 94°C for 30 s, annealing at 57°C for 1 min 30 s, extension at 72°C for 1 min, and a final extension at 60°C for 30 min; for the primers labeled with PET, the number of cycles was increased to 33, but the other conditions were unchanged. Size of the PCR products was determined using an ABI PRISM 3100 Genetic Analyzer (Applied Biosystems) and Genotyper software (Applied Biosystems) with GeneScan 500 LIZ (Applied Biosystems) as an internal size standard. Although the reliability and repeatability of compound SSR primers are controversial, we conducted repeated amplification of eight identical samples and obtained consistent results using the markers described here.

Eight individuals from the population located at 3300 m (GPS coordinates are shown in Table 2) were used in initial screening of the 33 primer pairs designed here. Of these 33 primer pairs, 11 exhibited clearly identifiable bands

TABLE 1. Characteristics of the 11 SSR loci isolated from *Leptospermum recurvum*.

Multiplex reactions*	Locus	Repeat motif	Primer sequences (5′–3′)	Fluorescent dye	T_a (°C)	Size range (bp)	GenBank accession no.
Multiplex genotyping 1							
a	Lrec005	$(AC)_6(AG)_9$	F: ACACACACACACAGAGAGAGAG R: TTTCAGGTGGAAATGATGAAT	6-FAM	57	171–191	AB761219
b	Lrec475	$(AC)_6(TC)_7$	F: ACACACACACACTCTCTCTCTC R: CCCAAGAGTGAGATTTGAAGAC	PET	57	98–104	AB761229
c	Lrec134	$(AC)_6(TC)_8$	F: ACACACACACACTCTCTCTCTC R: AGATCCCACTAATCAACCCTAA	VIC	57	176–182	AB761220
c	Lrec349	$(AC)_6(TC)_7$	F: ACACACACACACTCTCTCTCTC R: ATTTTCTCCAGCAGCTCTATGT	VIC	57	273–285	AB761226
c	Lrec368	$(AC)_6(TC)_{13}$	F: ACACACACACACTCTCTCTCTC R: AGCTCACCAACTTCCCTATTTA	VIC	57	87–111	AB761227
Multiplex genotyping 2							
d	Lrec264	$(AC)_6(AG)_{12}$	F: ACACACACACACAGAGAGAGAG R: ACGAAACCATAGAGAGATCGAG	6-FAM	57	168–204	AB761223
e	Lrec273	$(AC)_6(TC)_9$	F: ACACACACACACTCTCTCTCTC R: CTGCTATAAAAGGGGTCACTTC	PET	57	188–194	AB761224
e	Lrec401	$(AC)_6(TC)_7$	F: ACACACACACACTCTCTCTCTC R: ATATGGAGAGGAAAATGCAAAG	PET	57	90	AB761228
f	Lrec139	$(AC)_6(TC)_7$	F: ACACACACACACTCTCTCTCTC R: AAATTTCACGAACACTTCATTG	VIC	57	209–215	AB761221
f	Lrec206	$(AC)_6(TC)_7$	F: ACACACACACACTCTCTCTCTC R: GCTCACCTAAAAGAGATGAAGG	VIC	57	73–79	AB761222
f	Lrec296	$(AC)_6(TC)_7$	F: ACACACACACACTCTCTCTCTC R: CGAATTCTTAAGTGGCAAAAAT	VIC	57	100–108	AB761225

Note: T_a = annealing temperature.

* Letters represent sets of multiplex PCR performed to test the characteristics of the 11 loci (see Methods and Results).

(Table 1). We tested the characteristics of each locus with these 11 primer pairs using samples from four populations: three *L. recurvum* populations with various phenotypes from different habitats ranging from the lowest altitude (2700 m, forest), middle altitudes (3300 m, shrub), to the highest altitude (3900 m, alpine cushion community) on Mount Kinabalu, and one *L. polygalifolium* population found at 1900 m on Mount Kinabalu. Because the locus Lrec401 was monomorphic in all populations, this locus was excluded from the subsequent analyses. We found that all loci had more than one allele except for the locus Lrec139 in the *L. recurvum* population at the lowest altitude. Using GenAlEx version 6.3 (Peakall and Smouse, 2006), the observed heterozygosity, expected heterozygosity, and inbreeding coefficient values of *L. recurvum* were found to range from 0.04 to 0.94, from 0.04 to 0.83, and from −0.48 to 0.42 (except for the monomorphic locus Lrec139), respectively, and those for *L. polygalifolium* ranged from 0.07 to

0.67, from 0.29 to 0.84, and from −0.18 to 0.84, respectively. We also tested for departure from Hardy–Weinberg equilibrium (HWE) for all populations and loci as well as for linkage disequilibrium (LD) between all pairs of polymorphic loci using FSTAT version 2.3.9.2 (Goudet, 1995, 2001). We found that no population demonstrated significant departure from HWE, and that no pair of loci exhibited significant LD ($P < 0.05$ after Bonferroni correction).

CONCLUSIONS

The 10 primer pairs developed in this study successfully amplified the SSR loci of all three populations of *L. recurvum*, which include those from different habitats and of various phenotypes.

TABLE 2. Summary statistics of 10 SSR loci in three populations of *Leptospermum recurvum* and one population of *L. polygalifolium*.[a,b]

	L. recurvum												*L. polygalifolium* (n = 17)			
	Population A (n = 22)				Population B (n = 16)				Population C (n = 26)							
Locus	A	H_o	H_e	F_{IS}	A	H_o	H_e	F_{IS}	A	H_o	H_e	F_{IS}	A	H_o	H_e	F_{IS}
Lrec005	4	0.38	0.63	0.40	4	0.44	0.41	−0.07	3	0.42	0.34	−0.24	7	0.53	0.76	0.30
Lrec134	4	0.48	0.44	−0.08	4	0.25	0.37	0.33	2	0.04	0.04	−0.02	2	0.41	0.50	0.18
Lrec139	1	0.00	0.00	NA	2	0.06	0.06	−0.03	3	0.15	0.14	−0.07	3	0.20	0.29	0.31
Lrec206	2	0.25	0.40	0.37	2	0.50	0.47	−0.07	2	0.19	0.17	−0.11	2	0.07	0.44	0.84
Lrec264	10	0.75	0.83	0.10	9	0.94	0.80	−0.17	7	0.42	0.50	0.15	11	0.67	0.84	0.21
Lrec273	3	0.57	0.54	−0.05	3	0.80	0.54	−0.48	3	0.46	0.61	0.24	3	0.35	0.37	0.05
Lrec296	3	0.30	0.40	0.24	4	0.50	0.59	0.15	4	0.19	0.33	0.42	4	0.36	0.49	0.28
Lrec349	4	0.38	0.39	0.03	4	0.31	0.36	0.14	3	0.31	0.33	0.06	3	0.41	0.41	0.00
Lrec368	7	0.67	0.68	0.02	8	0.81	0.82	0.01	8	0.73	0.80	0.08	5	0.44	0.45	0.02
Lrec475	4	0.14	0.22	0.35	2	0.44	0.40	−0.08	3	0.54	0.40	−0.34	3	0.41	0.35	−0.18

Note: A = number of alleles per locus; F_{IS} = inbreeding coefficient; H_e = expected heterozygosity; H_o = observed heterozygosity; NA = not available because of a monomorphic locus.

[a] Population and voucher information: *Leptospermum recurvum*: Population A = 2700 m (6.04605°N, 116.55987°E), voucher no. SNP3208. Population B = 3300 m (6.05831°N, 116.56572°E), voucher no. SNP1597. Population C = 3900 m (6.07069°N, 116.56516°E), voucher no. 13896. *L. polygalifolium*: 1900 m (6.04186°N, 116.61804°E), voucher no. 233 (deposited as *L. flavescens*, a synonym of *L. polygalifolium*). All vouchers are deposited in the Sabah Parks Herbarium at Kinabalu Park headquarters. All specimens were collected by early workers in the study sites.

[b] No population or locus showed significant departure from Hardy–Weinberg equilibrium ($P < 0.05$ after Bonferroni correction).

These primer pairs will be useful for investigating patterns of gene flow and genetic structure of *L. recurvum* across habitats and phenotypes. In addition, our markers successfully amplified the loci of and detected polymorphisms in the related species *L. polygalifolium*; these markers are expected to have broader application.

LITERATURE CITED

ARAKI, K. S., S. KANEKO, Y. ISAGI, AND H. KUDOH. 2011. Isolation and characterization of microsatellite loci in a clonal herb, *Cardamine leucantha* (Brassicaceae). *American Journal of Botany* 98: e385–e387.

GOUDET, J. 1995. FSTAT version 1.2: A computer program to calculate F-statistics. *Journal of Heredity* 86: 485–486.

GOUDET, J. 2001. FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3). Available from http://www2.unil.ch/popgen/softwares/fstat.htm [accessed 24 October 2012].

LARKIN, M. A., G. BLACKSHIELDS, N. P. BROWN, R. CHENNA, P. A. MCGETTIGAN, H. MCWILLIAM, F. VALENTIN, ET AL. 2007. Clustal W and Clustal X version 2.0. *Bioinformatics (Oxford, England)* 23: 2947–2948.

LEE, D. W., AND J. B. LOWRY. 1980. Plant speciation on tropical mountains: *Leptospermum* (Myrtaceae) on Mount Kinabalu, Borneo. *Botanical Journal of the Linnean Society* 80: 223–242.

LIAN, C., Z. ZHOU, AND T. HOGETSU. 2001. A simple method for developing microsatellite markers using amplified fragments of inter-simple sequence repeat (ISSR). *Journal of Plant Research* 114: 381–385.

LIAN, C., M. A. WADUD, Q. GENG, K. SHIMATANI, AND T. HOGETSU. 2006. An improved technique for isolating codominant compound microsatellite markers. *Journal of Plant Research* 119: 415–417.

PEAKALL, R., AND P. E. SMOUSE. 2006. GenAlEx 6: Genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* 6: 288–295.

ROZEN, S., AND H. J. SKALETSKY. 2000. Primer3 on the WWW for general users and for biologist programmers. *In* S. Krawetz and S. Misener [eds.], Bioinformatics methods and protocols: Methods in molecular biology, 365–386. Humana Press, Totowa, New Jersey, USA.