Sophora microphylla (Fabaceae) Microsatellite Markers and their Utility Across the Genus

Authors: Etten, Megan L. Van, Houliston, Gary J., Mitchell, Caroline M., Heenan, Peter B., Robertson, Alastair W., et al.

Source: Applications in Plant Sciences, 2(3)

Published By: Botanical Society of America

URL: https://doi.org/10.3732/apps.1300081

Applications in Plant Sciences 2014 2(3): 1300081

PRIMER NOTE

SOPHORA MICROPHYLLA (FABACEAE) MICROSATELLITE MARKERS AND THEIR UTILITY ACROSS THE GENUS[1]

MEGAN L. VAN ETTEN[2,5], GARY J. HOULISTON[3], CAROLINE M. MITCHELL[3], PETER B. HEENAN[3], ALASTAIR W. ROBERTSON[2], AND JENNIFER A. TATE[4]

[2]Institute of Agriculture and Environment, Massey University, Palmerston North 4442, New Zealand; [3]Landcare Research, Lincoln 7640, New Zealand; and [4]Institute of Fundamental Sciences, Massey University, Palmerston North 4442, New Zealand

- *Premise of the study:* Genus-specific microsatellite markers were developed for *Sophora* for population genetic and systematic studies of the group in New Zealand, and potentially elsewhere in the geographic range.
- *Methods and Results:* From sequencing a total genomic DNA library (using Roche 454), we identified and developed 29 polymorphic microsatellite markers for *S. microphylla* and *S. chathamica*. We tested 12 of these markers on 14 *S. chathamica* individuals and four *S. microphylla* populations. All loci amplified in both species and species-specific alleles occurred at seven loci. In *S. microphylla* populations, the observed and expected heterozygosities ranged from 0.000–0.960 and 0.000–0.908, respectively, with alleles per locus ranging from seven to 23.
- *Conclusions:* The developed markers will be valuable in studies of phylogenetics, population structure, mating system, and selection of provenances for restoration projects.

Key words: Fabaceae; genetic variation; simple sequence repeat markers; *Sophora microphylla*.

Sophora L. (Fabaceae) in New Zealand comprises eight closely related endemic species (Mitchell and Heenan, 2002) collectively known by the indigenous vernacular name kowhai. Kowhai nectar provides an important food source for New Zealand endemic passerine birds (Stewart and Craig, 1985; Spurr et al., 2011), and extracts of the leaves and bark are used by the indigenous Māori as remedies for various ailments. Although the eight species differ in morphological traits, habitat usage, and geographic distribution, both chloroplast and nuclear loci have shown little to no sequence variation among species, making it difficult to determine the relationship among them (Hurr et al., 1999; Heenan et al., 2001; Mitchell and Heenan, 2002). Microsatellites, due to their high variability, are useful markers for resolving phylogenetic relationships among closely related species as well as for population genetic analyses within species (Selkoe and Toonen, 2006). As no microsatellites have yet to be developed for any *Sophora* species within New Zealand, we used next-generation sequencing to develop and test polymorphic microsatellite markers for two species—the widespread *S. microphylla* Aiton and the range-restricted *S. chathamica* Cockayne—with the goal of developing 12 markers for use in phylogenetic and mating system analyses.

[1]Manuscript received 3 October 2013; revision accepted 8 November 2013.

Funding was provided for this work to Landcare Research via Ministry of Business, Innovation and Employment core funding and to A.W.R. and J.A.T. via the Massey University Research Fund. Duckchul Park carried out the library preparation and sequencing.

[5]Author for correspondence: m.vanetten@massey.ac.nz

doi:10.3732/apps.1300081

METHODS AND RESULTS

Total genomic DNA was extracted from fresh leaf samples of *S. chathamica* (voucher no. CHR 529909, deposited at Allan Herbarium, Christchurch, New Zealand) using the DNeasy Plant Mini Kit (QIAGEN, Hilden, Germany) according to the manufacturer's protocol. The DNA was used to create a shotgun multiplex identifier (MID) library and sequenced on a Roche 454 Junior Genome Sequencer (454 Life Sciences, a Roche Company, Branford, Connecticut, USA) using Roche Titanium chemistry following the method of Margulies et al. (2005) at Landcare Research (Auckland, New Zealand). The resulting library of 139,372 reads (average read length: 404 bp) and total number of 56.4 megabases was assembled into 18,811 contigs using Geneious 6.0 (Biomatters, Auckland, New Zealand). Putative chloroplast and mitochondrial sequences were identified and removed using a local BLAST search against complete sequences obtained from GenBank of *Arabidopsis thaliana* (L.) Heynh. (Brassicaceae), *Vigna radiata* (L.) R. Wilczek (Fabaceae), and *Carica papaya* L. (Caricaceae). The remaining sequences were searched for perfect di- to hexanucleotide repeat regions with at least five repeat units, using a tandem repeat search tool in Geneious (Phobos plugin; Mayer, 2010). We designed primers for a subset of repeats (those with no other repeats within 50 bp of the repeat region and few mononucleotide repeats) using Primer3 (Rozen and Skaletsky, 2000) as implemented in Geneious using the default settings except for: product size = 100–350 bp; primer size = 17 (minimum)–19 (optimal)–21 (maximum); melting temperature (T_m; °C) = 52–55–58; GC content (%) = 40–50–60; max T_m difference = 5°C; GC clamp = 1; max poly x = 4. Of the 780 primers designed, 48 pairs were chosen based on the number of repeats (6–14). For each of these primers, an M13 tag (CACGACGTTGTAAAACGAC) was added to the 5′ end of the forward primer and, for primers designed at Massey University, a PIG tail was added to the 5′ end of the reverse primer (GTTTCTT) to promote nontemplate (A) addition (Brownstein et al., 1996). The PIG tail was not added for those primers developed at Landcare Research (see Table 1).

These 48 primer pairs were tested on *S. microphylla* samples extracted using a modified cetyltrimethylammonium bromide (CTAB) protocol (Shepherd and McLay, 2011). PCR was performed in 10-μL reactions consisting of 1 μL 1:10 dilution DNA:H_2O (5–50 ng), 0.02 μM forward primer, 0.45 μM reverse primer, 0.45 μM M13 FAM-labeled primer, 1.5 mM $MgCl_2$, 1× buffer BD (Solis BioDyne, Tartu, Estonia), 250 μM of each dNTP, and 0.2–0.4 U FirePol *Taq* polymerase (Solis BioDyne). PCR conditions were: 95°C for 3 min; 35 cycles of 95°C for 30 s, 53°C for 40 s, and 72°C for 1 min; and extension at 72°C for 10 min. PCR products

TABLE 1. Characteristics of 29 polymorphic microsatellite loci isolated from *Sophora chathamica* sequences and screened in *S. microphylla*.

Locus	Primer sequences (5′–3′)[a]	Repeat motif	Allele size range (bp)	T_a (°C)	GenBank accession no.
Sop-42	F: CCATACCTGACACTTGCGG	$(AG)_9$	168–198	53	KF672193
	R: TTGAGTCCAACATGAATGGC				
Sop-248	F: TCCCGGAAATCTCATTCAAAGG	$(GTT)_{13}$	265–328	53	KF672195
	R: ACTCAAGGAGTTTAGGTAGCG				
Sop-308	F: TGAACCGCTATCTTTTCCC	$(AT)_9$	177–233	53	KF749284
	R: CCGCCACTGGTATTTTAAC				
Sop-336	F: CCCGCTCTATCACTTAAGCC	$(AT)_9$	312–316	63–53†	KF749294
	R: TTACGACTTTGTTAAGAGGCAC				
Sop-382	F: GCGTCATCAATGTATGATCG	$(AC)_9$	163–173	53	KF749285
	R: TCCTTTCTGTTCCGTAACG				
Sop-445	F: CCAAATGGAGGAAGAAGGGTATTC	$(AGG)_7$	182–197	53	KF672197
	R: AGCTTCAACGCCAAACATCC				
Sop-579	F: GACTAGACGTTTCACTATATGTCC	$(AT)_{10}$	276–286	53	KF749296
	R: CGTGGTGTCTCAGGTTACTTTC				
Sop-802	F: ACAAAGCCTCATACAGAGC	$(GTT)_{10}$	297–342	53	KF672187
	R: GAATGACCAAGGTATCGCC*				
Sop-803	F: TGGTTGAAACTTGGACTGG	$(TA)_8$	284–316	53	KF749289
	R: TTTTATAGTGGGCTTGGGC*				
Sop-805	F: CTCGTGGGATTGTCCTATG	$(GA)_8$	172–174	53	KF749290
	R: GTGAAAATGCTCATGAGGC*				
Sop-806	F: AGAGTACCCGTGGATATCC	$(AT)_9$	191–229	63–53†	KF672188
	R: GCGATCAAGACACACAAC*				
Sop-807	F: AGTGTACCTTGACGATTGTG	$(AT)_9$	317–355	53	KF672189
	R: TCAGTTGGTGAACTATCAAC*				
Sop-808	F: ATTCGCCTAGATTGGGATG	$(AG)_{11}$	281–357	63–53†	KF672190
	R: AATTGTCAGCACCGACTTC*				
Sop-813	F: TCACGACTCTACTGGACTG	$(AAT)_8$	161–182	53	KF749286
	R: TGAAACCCAACTTGTCACC*				
Sop-814	F: GACGACGTTTGGAAGTTTG	$(AC)_9$	167–215	53	KF672191
	R: GGATATGACTTTGTCCACTTC*				
Sop-816	F: ATGTCGGTGAATGGTGATC	$(TCTT)_8$	297–337	53	KF672192
	R: AGGAACACTTCACACATGG*				
Sop-817	F: GCCAAATAGAAATGCCTGG	$(ATT)_9$	131–152	53	KF749287
	R: TGTTTGGGTGTATCCATCC*				
Sop-818	F: ATGTCCTCCTCTACTCTGC	$(AT)_7$	176–180	53	KF749288
	R: ATGATCCAGCTCTCATGTG*				
Sop-820	F: ATAATTTCCGGCAAGGTGG	$(GT)_7$	225–229	53	KF749291
	R: CTAAATCAAAATGGGCCCG*				
Sop-822	F: AAAATGAAGACGGTGGGTG	$(TA)_6$	224–226	53	KF749292
	R: CCCTTCTGCAACTATCTGC*				
Sop-824	F: TATTTGGGATGGAGAACCC	$(AT)_8$	172–180	63–53†	KF749293
	R: GAGAATCGGGACTAGACAAC*				
Sop-825	F: ATCCTGCGAAATACGACAG	$(GA)_9$	186–204	53	KF672194
	R: GTCGGAGTCCATGATTCAG*				
Sop-828	F: AGTGGCTTGATCTTCAACC	$(TC)_9$	204–214	53	KF749295
	R: TACCCCGTGTTACGAAATC*				
Sop-831	F: CACTAGAGAATACGATTGCG	$(GA)_{12}$	185–249	53	KF672196
	R: GTATACGTTATACCCGCGC*				
Sop-834	F: TTGGGCCTACAATGTATGG	$(TCT)_9$	257–326	63–53†	KF672198
	R: CATGCTCATCTCCCAAGAG*				
Sop-835	F: GCTCTAACCCTTTCTCCAC	$(CT)_5$	233–257	53	KF749297
	R: TCAGAGAGAGGACAGTGTC*				
Sop-836	F: TCGGTAACCCTGAGTACTG	$(AT)_{10}$	253–339	53	KF749298
	R: ATCCCGGCCAAATAAAGAC*				
Sop-838	F: CTGCATCTGACTCGAATCC	$(TA)_7$	289–293	53	KF749299
	R: CAAACTGTGAGGAGACAGG*				
Sop-840	F: GAAAATGCAGGTTCCGTTC	$(GGC)_7$	210–216	53	KF749300
	R: CATGTCTATCACCGACTGG*				

Note: T_a = annealing temperature used in PCR.

[a] M13 tail (CACGACGTTGTAAAACGAC) added to the 5′ end of each forward primer.

* PIG tail (GTTTCTT) added to the 5′ end of each reverse primer.

† Touchdown PCR.

(0.20–1.00 µL) were added to 9 µL Hi-Di formamide (Applied Biosystems, Carlsbad, California, USA) and 1 µL CASS ladder (Symonds and Lloyd, 2004) for fragment sizing on an ABI3730 Genetic Analyzer (Applied Biosystems) by the Massey University Genome Service (Palmerston North, New Zealand). Alleles were visualized and scored using GeneMapper version 3.7 (Applied Biosystems).

Of the 48 loci tested, 29 loci were polymorphic, from which 12 were chosen (based on ease of scoring, good separation for coloading PCR products, and high number of alleles) to test on 88 individuals from four populations of *S. microphylla* and 14 individuals across the range of *S. chathamica* (Table 2, Appendix 1). PCR conditions were as described above except M13 primers were

TABLE 2. Results from screening 12 polymorphic markers in four populations of *Sophora microphylla* and 14 *S. chathamica* individuals.

	South Island									North Island				
	Ahuriri (n = 27)			Kowhai Point (n = 23)			Waimakariri (n = 22)			Vinegar Hill (n = 9)			*S. chathamica* (n = 14)	
Locus	A	H_o	H_e	A	H_o	H_e	A	H_o	H_e	A	H_o	H_e	A	P
Sop-42	5	0.593	0.521	8	0.478	0.525	5	0.409	0.619	4	0.556	0.593	2	1
Sop-248	11	0.792	0.863	11	0.810	0.872	12	0.818	0.817	6	0.667	0.753	7	0
Sop-445	4	0.538	0.655	5	0.364	0.549	5	0.571	0.703	5	0.778	0.623	2	0
Sop-802	10	0.960	0.842	11	0.571	0.836	13	0.682	0.822	4	0.778	0.722	7	1
Sop-806	6	0.619	0.659	16	0.429	0.908	10	0.800	0.791	4	0.889	0.685	6	0
Sop-807	5	0.385	0.436	10	0.471	0.806	10	0.476	0.816	5	0.222	0.759	8	0
Sop-808	12	0.667	0.889	6	0.591	0.716	9	0.737	0.816	7	0.556	0.821	4	2
Sop-814	5	0.440	0.782	5	0.526	0.702	5	0.700	0.714	3	0.444	0.537	4	1
Sop-816	5	0.852	0.755	5	0.773	0.756	5	0.682	0.640	4	0.778	0.623	3	0
Sop-825	5	0.750	0.747	5	0.682	0.718	8	0.714	0.816	7	0.778	0.821	5	1
Sop-831	12	0.240	0.844	13	0.714	0.893	12	0.467	0.873	6	0.571	0.684	9	4
Sop-834	7	0.292	0.620	5	0.538	0.743	11	0.273	0.884	1	0.00	0.00	4	1

Note: A = number of alleles; H_e = expected heterozygosity; H_o = observed heterozygosity; n = number of sampled individuals; P = number of alleles unique to *S. chathamica*.

labeled with FAM, NED, or VIC to allow coloading of PCR products when genotyping. Loci initially showing nonspecific amplification were reamplified using a touchdown PCR program (Table 1; initial denaturation 95°C for 3 min; 10 cycles of 95°C for 30 s, annealing temperature decreasing by 1°C each cycle starting at 63°C for 40 s, and 72°C for 1 min; 25 cycles of 95°C for 30 s, 53°C for 40 s, and 72°C for 1 min; final extension 72°C for 10 min). Additional information on the remaining 17 polymorphic loci can be obtained from the corresponding author.

For the *S. microphylla* populations, the numbers of alleles and observed and expected heterozygosities were determined using GenAlEx (Peakall and Smouse, 2006). We did not test for Hardy–Weinberg equilibrium because *S. microphylla* is a mixed-mating system species with potential for high selfing rates. Population measures were not estimated for *S. chathamica* because only one individual per population was available for genotyping. Instead, we calculated the number of alleles and the number of unique alleles found in *S. chathamica* but not in the *S. microphylla* populations sampled (Table 2). Voucher specimens were deposited in the Dame Ella Campbell herbarium (MPN) at Massey University and the Allan Herbarium (CHR) at Landcare Research (Appendix 1).

Of the 12 loci tested, all amplified in both *S. microphylla* and *S. chathamica*. The majority of alleles were shared between the two species, although at seven loci there were alleles specific to *S. chathamica*, four of which were outside the range found in *S. microphylla*. These differences suggest these loci could be phylogenetically informative, but greater sample sizes and broader range sampling are needed. All but one locus was polymorphic in all populations of *S. microphylla*, with the number of alleles ranging from one to 16. Observed heterozygosity ranged from 0.000–0.960 (average: 0.592), but was often lower than expected heterozygosity (range: 0.00–0.908; average: 0.720). This difference could be caused by a variety of processes including null alleles or violation of Hardy–Weinberg assumptions (e.g., high selfing rates). We checked for potential null alleles using MICRO-CHECKER version 2.2.3 (van Oosterhout et al., 2004) and found four loci with potential null alleles (Sop-808, 814, 831, and 834).

CONCLUSIONS

We designed and tested 48 primers for microsatellite loci derived from 454 sequencing, 29 of which were polymorphic within *S. microphylla*. We further tested 12 of the most polymorphic loci across the range of *S. chathamica* and for four populations of *S. microphylla*. The cross-compatibility between these two species suggests these markers could be successfully used in other closely related *Sophora* species, although the potential presence of null alleles should be explored further (e.g., genotyping parents and offspring; Dakin and Avise, 2004). The high polymorphism within populations and the species-specific alleles suggest the developed markers will be valuable in studies of population structure, dispersal, and species delineation, as well as for selection of populations for restoration projects.

LITERATURE CITED

BROWNSTEIN, M. J., J. D. CARPTEN, AND J. R. SMITH. 1996. Modulation of non-templated nucleotide addition by Taq DNA polymerase: Primer modifications that facilitate genotyping. *BioTechniques* 20: 1004–1010.

DAKIN, E. E., AND J. C. AVISE. 2004. Microsatellite null alleles in parentage analysis. *Heredity* 93: 504–509.

HEENAN, P. B., P. J. DE LANGE, AND A. D. WILTON. 2001. *Sophora* (Fabaceae) in New Zealand: Taxonomy, distribution and biogeography. *New Zealand Journal of Botany* 39: 17–53.

HURR, K. A., P. J. LOCKHART, P. B. HEENAN, AND D. PENNY. 1999. Evidence for the recent dispersal of *Sophora* (Leguminosae) around the Southern Oceans: Molecular data. *Journal of Biogeography* 26: 565–577.

MARGULIES, M., M. EGHOLM, W. ALTMAN, S. ATTIYA, J. BADER, L. BEMBEN, J. BERKA, ET AL. 2005. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437: 376–380.

MAYER, C. 2010. Phobos 3.3.11. Website http://www.rub.de/spezzoo/cm/cm_phobos.htm [accessed 4 January 2014].

MITCHELL, A. D., AND P. B. HEENAN. 2002. *Sophora* sect. *Edwardsia* (Fabaceae): Further evidence from nrDNA sequence data of a recent and rapid radiation around the Southern Oceans. *Botanical Journal of the Linnean Society* 140: 435–441.

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. Misener and S. A. Krawetz [ed.], Methods in molecular biology, vol. 132: Bioinformatics methods and protocols, 365–386. Humana Press, Totowa, New Jersey, USA.

SELKOE, K. A., AND R. J. TOONEN. 2006. Microsatellites for ecologists: A practical guide to using and evaluating microsatellite markers. *Ecology Letters* 9: 615–629.

SHEPHERD, L. D., AND T. G. B. MCLAY. 2011. Two micro-scale protocols for the isolation of DNA from polysaccharide-rich plant tissue. *Journal of Plant Research* 124: 311–314.

Spurr, E. B., S. Rod, and K. P. Tranter. 2011. Food preferences of the bellbird (*Anthornis melanura*) in native forest remnants on the Port Hills, Banks Peninsula, New Zealand. *Notornis* 58: 139–157.

Stewart, A. M., and J. L. Craig. 1985. Movements, status, access to nectar, and spatial organisation of the tui. *New Zealand Journal of Zoology* 12: 649–656.

Symonds, V. V., and A. M. Lloyd. 2004. A simple and inexpensive method for producing fluorescently labelled size standard. *Molecular Ecology Notes* 4: 768–771.

van Oosterhout, C., W. F. Hutchinson, D. P. M. Willis, and P. Shipley. 2004. MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* 4: 535–538.

Appendix 1. Locations and herbarium voucher information for *Sophora microphylla* and *S. chathamica* populations used in this study.

Population	Voucher specimen accession no.	Geographic coordinates
S. microphylla		
Ahuriri	MPN 47481	43°41′16.22″S, 172°35′16.36″E
Kowhai Point	MPN 47479	41°42′41.31″S, 173°6′48.62″E
Waimakariri	MPN 47480	43°27′36.01″S, 172°12′55.40″E
Vinegar Hill	MPN 47478	39°56′5.49″S, 175°38′34.18″E
S. chathamica		
Te Puna Point	CHR 493865	37°46′S, 174°56′E
Waiwera River	CHR 517103	36°32′24″S, 174°42′25″E
Puhoi River	CHR 517104A	36°30′52″S, 174°40′14″E
Puhoi River	CHR 517107A	36°30′52″S, 174°40′14″E
Wenderholm	CHR 517115	36°32′01″S, 174°42′45″E
Helensville	CHR 517116	36°39′43″S, 174°27′01″E
Te Mata Stream	CHR 517130	36°58′12″S, 175°29′54″E
Waitakaruru River	CHR 517131	37°17′09″S, 175°18′56″E
Waiwera River	CHR 517162	36°32′21″S, 174°42′33″E
Tawairoa Stream	CHR 527630	38°7′S, 174°54′E
Coromandel	CHR 527641	36°29′S, 175°23′E
Lake Whangape	CHR 546230	37°28′33″S, 175°3′40″E
Waitangi Stream	CHR 553890	36°33′25″S, 174°28′23″E
Kaukapakapa River	CHR 554078	36°38′1″S, 174°31′14″E

Note: CHR = Allan Herbarium at Landcare Research, Christchurch, New Zealand; MPN = Dame Ella Campbell Herbarium at Massey University, Palmerston North, New Zealand.