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

Authors: Megan L. Van Etten, Gary J. Houliston, Caroline M. Mitchell, Peter B. Heenan, Alastair W. Robertson, et. al.
Source: Applications in Plant Sciences, 2(3)
Published By: Botanical Society of America
URL: https://doi.org/10.3732/apps.1300081
SOPHORA MICROPHYLLA (FABACEAE) MICROSATELLITE MARKERS AND THEIR UTILITY ACROSS THE GENUS

Megan L. Van Etten2,5, Gary J. Houliston3, Caroline M. Mitchell3, Peter B. Heenan3, Alastair W. Robertson2, and Jennifer A. Tate4

1 Institute of Agriculture and Environment, Massey University, Palmerston North 4442, New Zealand; 2 Landcare Research, Lincoln 7640, New Zealand; and 3 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.

METHODS AND RESULTS

Total genomic DNA was extracted from fresh leaf samples of S. chathamica (voucher no. CHR 529909, deposited at Allian 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. (Carcaceae). 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)–21 (optimal)–21 (maximum); melting temperature (Tm, °C) = 52–55–58; GC content (%) = 40–50–60; max Tm 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 (CAGCAGGTGGTAAACACGAC) 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 (GTGGTTTCTT) 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: H2O (5–50 ng), 0.02 μM forward primer, 0.45 μM reverse primer, 0.45 μM M13 FAM-labeled primer, 1.5 mM MgCl2, 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

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

Applications in Plant Sciences 2014 2(3): 1300081; http://www.bioone.org/doi/apps © 2014 Van Etten et al. Published by the Botanical Society of America.

This work is licensed under a Creative Commons Attribution License (CC-BY-NC-SA).
Applications in Plant Sciences 2014 2 (3): 1300081
Van Etten et al.—*Sophora microphylla* microsatellites

(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).

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

| Locus   | Primer sequences (5′–3′) | Repeat motif | Allele size range (bp) | \( T_a (°C) \) | GenBank accession no. |
|---------|--------------------------|--------------|------------------------|----------------|-----------------------|
| Sop-42  | F: CCATACCTGACACTTGGGG   | (AG)₉        | 169–198                | 53             | KF672193              |
|         | R: TTGACCACCATGATAGGCG   |              |                        |                |                       |
| Sop-248 | F: TCCCGGAATTCTTCAAAGAG | (GTT)₁₃      | 265–328                | 53             | KF672195              |
|         | R: ACTCACAGGTGATTACCGCC  |              |                        |                |                       |
| Sop-308 | F: TGACCGCTTAATTTCCCC   | (AT)₉        | 177–233                | 53             | KF749284              |
|         | R: CCGCCACCTGATTTTTAC   |              |                        |                |                       |
| Sop-336 | F: CCCGCTCTACATTAAGGCC  | (AT)₉        | 312–316                | 63–53¹        | KF749294              |
|         | R: TTACGGCTTTTAAAGGACAC |              |                        |                |                       |
| Sop-382 | F: GCGTCATCAATGTATGATCG | (AC)₉        | 163–173                | 53             | KF749285              |
|         | R: TCTCTTGCTTGGCTAGACC  |              |                        |                |                       |
| Sop-445 | F: CCAATGTCAGTCCCTTAC   | (AGG)₇       | 182–197                | 53             | KF672197              |
|         | R: ACTCAAGGAGTTAGATGG    |              |                        |                |                       |
| Sop-579 | F: GACTAGAGTGCTACTATG    | (AT)₁₀       | 276–286                | 53             | KF749296              |
|         | R: GGGTCTGACGTCTTTCTTC  |              |                        |                |                       |
| Sop-802 | F: ACAAGGCTCTTACACAGAG  | (GTT)₁₀      | 297–342                | 53             | KF672187              |
|         | R: GAAATGCAAGATATGACC    |              |                        |                |                       |
| Sop-803 | F: TGTTTGAAACTTGAGTGG    | (TA)₉        | 284–316                | 53             | KF749289              |
|         | R: TTTTATAGTGGGCTTGGGC   |              |                        |                |                       |
| Sop-805 | F: CTGTTGGATGTCTTGTAG    | (GA)₉        | 172–174                | 53             | KF749290              |
|         | R: GGAATGCTCAGTACAGACC   |              |                        |                |                       |
| Sop-806 | F: AGATGACCCTGCTATGACCC  | (AT)₁₀       | 191–229                | 63–53¹        | KF672188              |
|         | R: GCTGCGCTGCTACTTTC     |              |                        |                |                       |
| Sop-807 | F: AGGTAGCTCTCTACTGAG    | (AT)₉        | 317–355                | 53             | KF672189              |
|         | R: GTGATACTGCTTGGCTTGC   |              |                        |                |                       |
| Sop-808 | F: ATTCGGCTGTAAGTGATGC   | (AG)₁₁       | 281–357                | 63–53¹        | KF672190              |
|         | R: AATGTGCAGCAGCTTACC    |              |                        |                |                       |
| Sop-813 | F: TCAGAGCTCTCTACTGAG    | (AA)₉        | 161–182                | 53             | KF749286              |
|         | R: TGACTAACCTTCAGCTCC    |              |                        |                |                       |
| Sop-814 | F: GACAGCTTTTGAGAGTTTG   | (AC)₉        | 167–215                | 53             | KF672191              |
|         | R: GCAATAGACTTCTTGGAGAC  |              |                        |                |                       |
| Sop-816 | F: ATGTGCGTGTAAGTGATGC   | (TCTT)       | 297–337                | 53             | KF672192              |
|         | R: AGGAGAACTTCCACATGAG   |              |                        |                |                       |
| Sop-817 | F: GCCTAGTAAATAGCTCTTG   | (ATT)₉       | 131–152                | 53             | KF749287              |
|         | R: TGTGGGTGTATATCATCAG   |              |                        |                |                       |
| Sop-818 | F: ATGTGTCTCTCTTCTCTGC   | (AT)₁₁       | 176–180                | 53             | KF749288              |
|         | R: ATGATCGGCTTTCTCTCTG   |              |                        |                |                       |
| Sop-820 | F: ATATTTCCGTCGCAAAGTGG  | (GT)₉        | 225–229                | 53             | KF749291              |
|         | R: CTAATACAAATGGGGCGCG   |              |                        |                |                       |
| Sop-822 | F: AAAATGAGACACCGTTGGTG  | (TA)₉        | 224–226                | 53             | KF749292              |
|         | R: CCCCCTGCTGACTATCTGC   |              |                        |                |                       |
| Sop-824 | F: TATTTGGGATGAGAAACC    | (AT)₉        | 172–180                | 63–53¹        | KF749293              |
|         | R: GAAGATCGGCACTGACACC   |              |                        |                |                       |
| Sop-825 | F: ATCTTCGGGAAATAGACAG   | (GA)₉        | 186–204                | 53             | KF672194              |
|         | R: GTGCGTGGCATGTAGTACC   |              |                        |                |                       |
| Sop-828 | F: AGTGCTGGTATCTTCCACC  | (TC)₉        | 204–214                | 53             | KF749295              |
|         | R: TACCCTGGTATTCGCAACTC  |              |                        |                |                       |
| Sop-831 | F: CCTAGAAGATCATGGATGCG  | (GA)₁₂       | 185–249                | 53             | KF672196              |
|         | R: GTATCGGTATATCCGCGCG   |              |                        |                |                       |
| Sop-834 | F: TGTTTTGCTCAGACTGAGTC  | (TCT)₉       | 257–326                | 63–53¹        | KF749298              |
|         | R: AGCTGTGCTACTCTGCAAG   |              |                        |                |                       |
| Sop-835 | F: GCCGTAACCTTTTTCGCC    | (CT)₉        | 233–257                | 53             | KF749297              |
|         | R: TCAGAGGAGAAGGTGTCGC   |              |                        |                |                       |
| Sop-836 | F: TCCTGAACCTTGGATAGCTG  | (AT)₁₀       | 253–339                | 53             | KF749298              |
|         | R: ATCCCGGCAAATAAACAGAC  |              |                        |                |                       |
| Sop-838 | F: CTCGACTCAGTCGGATGAC   | (TA)₉        | 289–293                | 53             | KF749299              |
|         | R: CAAACTGTGAGGAGACAGGC  |              |                        |                |                       |
| Sop-840 | F: GAAATGACGCTTGCGTTC    | (GGC)₇       | 210–216                | 53             | KF749300              |
|         | R: CAGTGTCTATCACGGACTGG  |              |                        |                |                       |

**Note:** \( T_a \) = annealing temperature used in PCR.

¹M13 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.

Of the 48 loci tested, 29 loci were polymorphic, from which 12 were chosen (based on ease of scoring, good separation for co-loading 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 2 of 4

http://www.bioone.org/loi/apps
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. CARPENTER, 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 LAANGE, 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., S. EGHOLM, W. ALTMAN, S. ATTiya, J. BADER, L. BEMBIN, J. BERKA, ET AL. 2005. Genome sequencing in microfabricated high-density picolitre reactors. Nature 437: 376–380.

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

MITCHELL, A. D., and P. B. HEENAN. 2002. Sophora sect. Edwaradia (Fabaceae): Further evidence from nDNA 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, 365–368. Humana Press, Totowa, New Jersey, USA.

SELKOE, K. A., and R. J. TOONE. 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.
SPURRE, 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.

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

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

| Population       | Voucher specimen | 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               |
| Waipori River    | CHR 517103       | 36°32'24"S, 174°42'25"E        |
| Pahoi River      | CHR 517104A      | 36°30'52"S, 174°40'14"E        |
| Pahoi 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        |
| Waipori River    | CHR 517162       | 36°32'21"S, 174°42'33"E        |
| Taiaroa 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.

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.