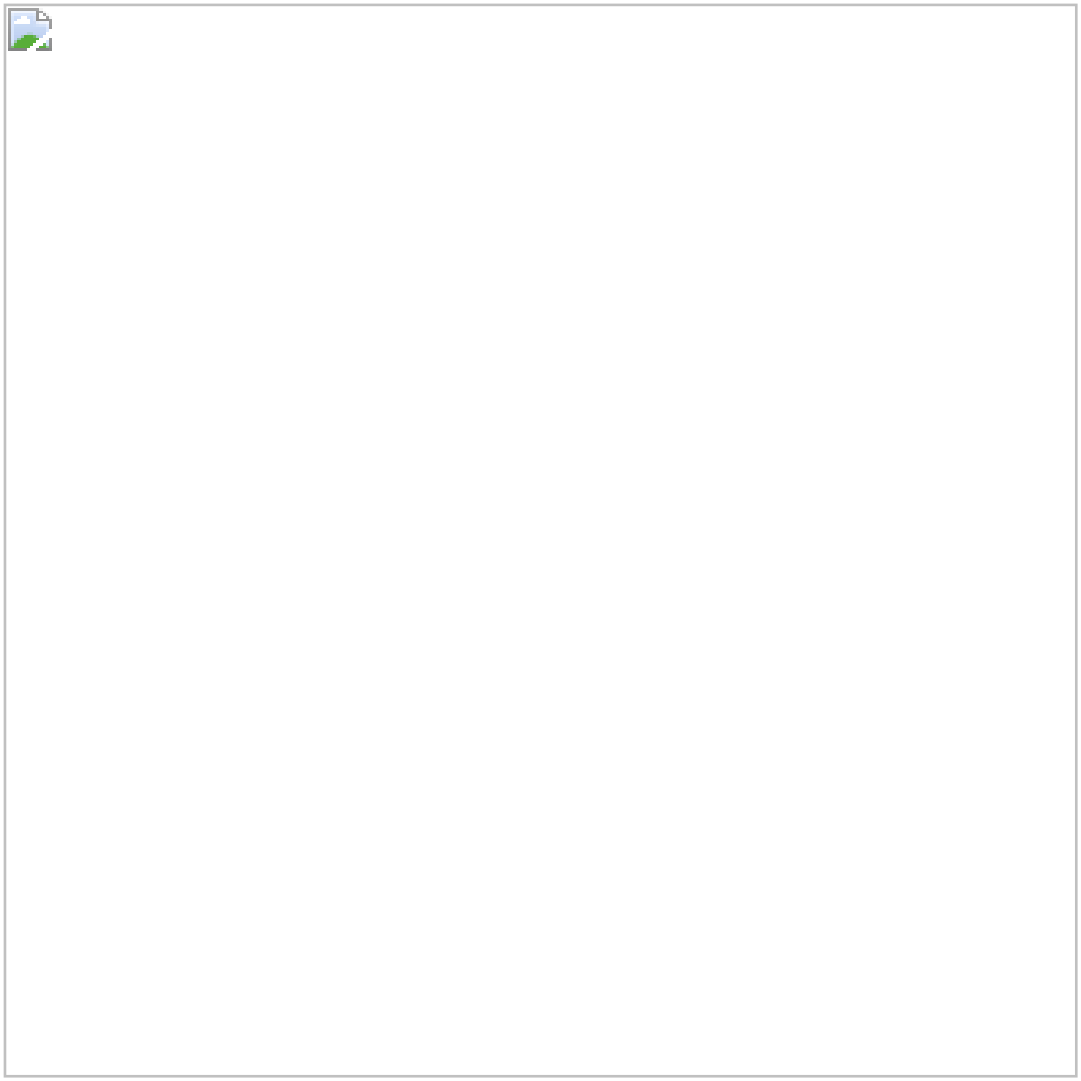

Development and Characterization of 10 Microsatellite Markers in Sagina nodosa (Caryophyllaceae)

Authors: Singh, Tarjinder, Edwards, Joan, and Maroja, Luana S.

Source: Applications in Plant Sciences, 2(1)

Published By: Botanical Society of America

URL: https://doi.org/10.3732/apps.1300064

Applications in Plant Sciences 2014 2(1): 1300064

PRIMER NOTE

DEVELOPMENT AND CHARACTERIZATION OF 10 MICROSATELLITE MARKERS IN *SAGINA NODOSA* (CARYOPHYLLACEAE)[1]

TARJINDER SINGH[2], JOAN EDWARDS[2], AND LUANA S. MAROJA[2,3]

[2]Department of Biology, Williams College, Williamstown, Massachusetts 01267 USA

- *Premise of the study:* We developed 10 novel microsatellite loci for *Sagina nodosa*, a diploid perennial arctic-alpine herb. To our knowledge, these are the first microsatellite loci for a *Sagina* species.
- *Methods and Results:* We performed a low-coverage 454 next-generation sequencing of enriched genomic fragments derived from one individual to generate a massive library of contigs containing potential polymorphic microsatellites. We present data for 10 novel polymorphic microsatellite loci containing di-, tri-, tetra-, and hexanucleotide repeats with two to nine alleles per locus assessed in 29 individuals.
- *Conclusions:* These polymorphic microsatellite loci in *S. nodosa* will provide insights on the population structure and life history of *S. nodosa* in Isle Royale and other North American populations.

Key words: 454 sequencing; arctic-alpine plant; knotted pearlwort; polymorphism; population genetics; *Sagina nodosa*.

Sagina nodosa (L.) Fenzl, commonly known as knotted pearlwort, is a small, diploid, perennial herb in the Caryophyllaceae family found predominantly in archipelagos and along exposed coastlines of the Northern hemisphere. The arctic-alpine plant is especially delicate, and on Isle Royale it usually inhabits cracks in rocks by the shore where it finds protection against cold waters, storms, and competition from other plants. Although stable populations exist in northern Canada and the arctic, the plant is classified as threatened in Michigan, USA (Slavick and Janke, 1993), and may be at risk or sensitive in the Canadian provinces of Alberta, New Brunswick, Saskatchewan, and Prince Edward Island (CFIA and NRCan/CFS, 2013).

Knotted pearlwort has two known modes of reproduction: it can cross- and self-pollinate to produce seeds for dispersal or it can propagate vegetatively (Crow, 1978; Magee and Ahles, 1999). However, to our knowledge, no genetic studies have been performed on the plants of this genus, and thus, our understanding of *Sagina* plants' life history is primarily observational. In addition, little is known about the population structure, genetic diversity, and the relative isolation of the North American subpopulations, particularly in the context of climate change and the potential disappearance of its habitats in the southern range its distribution.

We developed microsatellite markers to investigate the population structure and life history in isolated populations of *S. nodosa* at Isle Royale National Park in Michigan, USA. The Isle Royale population is particularly interesting as it provides a disjunct sample of individuals from the endangered southern population (Soper and Maycock, 1963). Furthermore, the numerous geographic barriers from the steep shoreline to the separation of islands provide natural ecological experiments that can provide insights on the life cycle and demographic history of *S. nodosa* (Given and Soper, 1981).

The identification of novel microsatellite loci traditionally requires the construction of a genomic library enriched for repeat motifs and the sequencing of individual clones to identify polymorphic loci (Zane et al., 2002). Recently new methods using next-generation sequencing increased the efficiency and speed of the process, while reducing costs (Santana et al., 2009). Here we performed a low-coverage 454 next-generation sequencing of enriched genomic fragments derived from one individual to generate a massive library of contigs containing potential polymorphic sequences (Abdelkrim et al., 2009). We present data for 10 novel microsatellite loci in *S. nodosa*, which will be used for landscape genetic analysis of the disjunct population of Isle Royale, Michigan, USA.

METHODS AND RESULTS

We collected fresh leaf material from *S. nodosa* individuals from eight islands located in the northeastern tip of Isle Royale National Park (Michigan, USA): the main island on Blake Point (3 individuals), Third Island (4 individuals), Long Island (5 individuals), North Government Island (6 individuals), Split Island (4 individuals), Edwards Island (3 individuals), South Government Island (3 individuals), and South Government Island Islet (1 individual). GPS coordinates for each individual can be found in Appendix 1. Small, nondestructive clippings of buds, stems, and leaves were placed into labeled air-tight plastic bags with Drierite, a strong desiccant. We then extracted genomic DNA with ArchivePure DNA Purification Kit (5 Prime, Gaithersburg, Maryland, USA) according to the manufacturer's protocol. A voucher specimen was deposited at the University of Michigan herbarium (MICH). The specimen is labeled "*Sagina*

[1]Manuscript received 31 July 2013; revision accepted 18 September 2013.

The authors thank the National Park Service for providing logistical support; E. Ury, E. Levy, R. Schoer, and R. A. Smith for providing assistance in the field; C. M. Tribble for help with genotyping; and S. Bogdanowicz for help with microsatellite 454 sequencing and analysis.

[3]Author for correspondence: lsm1@williams.edu

doi:10.3732/apps.1300064

nodosa, Michigan, Keweenaw Co., Isle Royale National Park, Edwards Island, Edwards, 31 Jul 2012 (MICH–1474911)."

One DNA sample from Isle Royale was then subjected to shotgun sequencing (1/16th run) using a Roche 454 Genome Sequencer FLX (454 Life Sciences, a Roche Company, Branford, Connecticut, USA) at the Evolutionary Genetics Core Facility (Ithaca, New York, USA). The library construction followed a modified protocol based on Hamilton et al. (1999). Following contig assembly with SeqMan Pro (Lasergene version 8.1.1; DNASTAR, Madison, Wisconsin, USA), we obtained 20,472 reads in FASTA format with an average read length of 341.3 bp and a total of 8,219,964 bp. The contigs were aligned under highly stringent conditions, requiring a minimum sequence length of 126 bp and with no more than 15 mismatched bases.

Using MSATCOMMANDER, we screened all assembled contigs for di-, tri-, tetra-, penta-, and hexanucleotide repeats, and ensured that each candidate contig had at least 40 bp in its flanking region and a minimum repeat length of 6 bp for dinucleotides and 5 bp for all other motifs (Faircloth, 2008). A total of 2287 sequences, or 11.17% of the total reads, had tandem repeats that satisfied our stringent conditions, with 765 (33.4%) dinucleotide repeats, 1059 (46.3%) trinucleotide repeats, 334 (14.6%) tetranucleotide repeats, 91 (4.0%) pentanucleotide repeats, and 38 (1.7%) hexanucleotide repeats.

We designed primers for 41 of the 2287 candidate contigs using BatchPrimer3 (You et al., 2008) with the following restrictions: amplification products must be within a size range of 100 to 500 bp, optimal melting temperature of 60°C (must be between 57°C and 62°C), optimal GC content of 50%, possession of at least one GC clamp, and reasonably low levels of self- and pair-complementarity. We then inspected each of the individual primers to ensure they did not amplify the same target microsatellite loci. Primers located near the ends of the contigs or in close proximity of microsatellite repeats were discarded.

For most primer pairs that amplified consistently (all except M33 and M35), we used a single-reaction nested PCR method (TP-PCR) to embed fluorescent dyes into PCR fragments for later genotyping (Schuelke, 2000). The tail 5′-CGAGTTTTCCCAGTCACGAC-3′ was appended to the 5′ end of our forward primers (Schuelke, 2000), and the tail 5′-GTTTCTT-3′ to the 5′ end of our reverse primers (Brownstein et al., 1996). A third fluorescently labeled primer, 6-FAM-5′-CGAGTTTTCCCAGTCACGAC-3′, was added to every reaction to embed the fluorescent dye 6-FAM label into the reaction products (Integrated DNA Technologies, Coralville, Iowa, USA). TP-PCR was performed with the following 25-μL PCR reaction mix: 1 μL of ~10–15 ng/μL genomic DNA, 11 μL of PCR-grade dH_2O, 10 μL of 2.5× HotMasterMix (5 Prime), 2.5 μL forward and reverse primer mix with a final concentration of 0.05 μM and 0.2 μL, respectively, and 0.5 μL of universal FAM-labeled primer with a final concentration of 0.2 μL. For more sensitive reactions, where yield was lower, the PCR Master Mix was replaced with the appropriate concentration of fresh *Taq* polymerase. Here, the PCR volume of 25 μL contained 1 μL of ~10–15 ng/μL genomic DNA, 15.85 μL of PCR-grade dH_2O, 0.5 μL of 0.2 mM dNTP, 0.15 μL of HotMaster Taq (5 Prime), 2.5 μL of 10× PCR buffer, 2.5 μL forward and reverse primer mix with a final concentration of 0.05 μM and 0.2 μM, respectively, and 0.5 μL of universal FAM-labeled primer with a final concentration of 0.2 μM. For loci M33 and M35, we used a 6-FAM fluorescently labeled forward primer (Integrated DNA Technologies, see Table 1 for primer sequence) and the Type-it Microsatellite PCR Kit (QIAGEN, Valencia, California, USA). For loci M33 and M35, each reaction (total volume of 12.5 μL) contained 2.5 μL of DNA, 6.25 μL of Type-it Multiplex PCR Master Mix (2×), 2.5 μL of DNase/RNase-free H_2O, and 1.25 μL primer mix (concentration of 0.2 μM).

All amplifications were carried out in the MultiGene Gradient PCR Thermal Cycler (Labnet International, Edison, New Jersey, USA). For the TP-PCR, the cycling started with initial denaturing at 94°C for 2 min; followed by 30 cycles of 94°C for 30 s, 56°C for 45 s, and 72°C for 1 min; followed by eight cycles of 94°C for 30 s, 53°C for 45 s, and 72°C for 1 min; and a final extension of 72°C for 10 min. For loci M33 and M35 (Type-it Microsatellite PCR Kit, QIAGEN), the cycling started with initial denaturing at 95°C for 5 min; followed by 28 cycles of 95°C for 30 s, 60°C for 90 s, and 72°C for 30 s; and a final extension of 60°C for 30 min. Prior to genotyping, all PCR products were visualized using gel electrophoresis.

PCR products were diluted in Hi-Di formamide (Applied Biosystems, Carlsbad, California, USA) and analyzed at the Cornell Life Sciences Core Laboratory Center on an ABI 4200 or 3730xl DNA Sequencer (Applied Biosystems) using the GeneScan 500 LIZ internal size standard (Applied Biosystems). For fragment analysis, we used GENEMAPPER version 3.7 (Applied Biosystems). All 29 individuals collected in Isle Royale were genotyped and analyzed using GENEPOP (Rousset, 2008) to quantify the allele frequencies and observed and expected heterozygosities at each locus.

Ten loci, all in linkage equilibrium, were polymorphic and included a balance of different microsatellite repeat motifs: four dinucleotide repeats, three trinucleotide repeats, two tetranucleotide repeats, and one hexanucleotide repeat (Table 1). In the 29 individuals genotyped, each locus had between two and nine total alleles and an observed heterozygosity ranging from 0.074 to 0.483 (Table 2). When individuals from all of the Isle Royale islands were analyzed as a single population, all loci significantly violated Hardy–Weinberg equilibrium (HWE), displaying a very low level of observed heterozygosity (Table 2). This is likely due to population substructure (i.e., Wahlund effect [Wahlund, 1928]) within the archipelago. Indeed, when we analyzed each island population separately, we found no deviation from HWE (although our sample sizes were too small for a definite conclusion). Alternatively, the low observed heterozygosity could be caused by significant clonal reproduction, but a larger analysis is needed to demonstrate the cause of HWE deviation.

CONCLUSIONS

We developed and characterized 10 novel microsatellite markers for 29 *S. nodosa* individuals from the disjunct Isle

TABLE 1. Characteristics of 10 microsatellite loci developed in *Sagina nodosa*.

Locus (Contig ID)	Primer sequences (5′–3′)	Repeat motif	Allele size range (bp)	T_a (°C)	GenBank accession no.
M2 (75)	F: AACGGCCAGATTCCAGTACA	$(TC)_9$	176–178	54	KF381173
	R: AGGATGTTGGAGCACGAGTT				
M8 (1355)	F: TACAATTTCTGAAACCACCGTTA	$(GGAT)_8$	168–176	56	KF381174
	R: CCACTTACAATCATGAACACACC				
M11 (2313)	F: GCAAGCATTTGGCCTACATT	$(TATC)_{11}$	237–261	56	KF381175
	R: TCCAAACATGCCATTTTCCT				
M14 (3637)	F: GAGGACAACATCGCTGAAGG	$(GA)_{10}$	233–239	56	KF381176
	R: CGCCATGGCTTTACTTTTGT				
M16 (3768)	F: GTCGATCTCTGTGGCATGAG	$(TTG)_6$	234–237	56	KF381177
	R: CCTTGTGTTGAGCTCCCTTC				
M25 (4289)	F: TTCCCTTACGATGCTTACGA	$(TTC)_{13}(TTA)_7$	239–251	56	KF381178
	R: TGTCGCGATTATCCAACAAA				
M28 (4395)	F: CGATAATCTCGTCGTCAAAGC	$(CTCTTT)_7$	167–179	56	KF381179
	R: CAAACCGACTCTGACCCAAT				
M33 (262)	F: ACCGTACGTGCACCTTCAA	$(ACA)_{13}$	240–252	56	KF381180
	R: ACAACCCACGCATTTCGTA				
M35 (3980)	F: TCTGTATGGTTCGTAACCTATGATG	$(AAG)_9$	240–256	56	KF381181
	R: CCAGAAACCCTTCTGAAACG				
M37 (4184)	F: ACCGTACGTGCACCTTCAA	$(AC)_{12}$	183–189	56	KF381182
	R: ACAACCCACGCATTTCGTA				

Note: T_a = annealing temperature.

TABLE 2. Genetic properties of the 10 microsatellite loci of *Sagina nodosa* among the 29 individuals sampled.

Locus	*A*	H_e	H_o	F_{IS}[a]
M2	3	0.492	0.074	0.8519
M8	3	0.563	0.172	0.6976
M11	9	0.836	0.483	0.4269
M14	4	0.618	0.138	0.7793
M16	2	0.494	0.138	0.7241
M25	6	0.770	0.241	0.6904
M28	3	0.498	0.207	0.5892
M33	4	0.623	0.178	0.7176
M35	5	0.634	0.250	0.6103
M37	4	0.526	0.241	0.5452

Note: *A* = number of alleles sampled; F_{IS} = inbreeding within populations; H_e = expected heterozygosity; H_o = observed heterozygosity.

[a] F_{IS} was calculated in GENEPOP (Rousset, 2008); all loci deviated significantly from HWE when populations were combined ($P < 0.0001$), but not when islands were analyzed individually.

Royale population using next-generation sequencing technology. The individuals are from eight different islands (subpopulations), and preliminary analyses suggest strong population substructure. To our knowledge, these are the first microsatellite loci for species in the *Sagina* genus. These novel markers will be used to genotype a larger population of individuals to provide insight on the population structure and life history of *S. nodosa* in Isle Royale and other North American populations.

LITERATURE CITED

ABDELKRIM, J., B. C. ROBERTSON, J. A. L. STANTON, AND N. J. GEMMELL. 2009. Fast, cost-effective development of species-specific microsatellite markers by genomic sequencing. *BioTechniques* 46: 185–191.

BROWNSTEIN, M. J., J. D. CARPTEN, AND J. R. SMITH. 1996. Modulation of non-templated nucleotide addition by tag DNA polymerase: Primer modifications that facilitate genotyping. *BioTechniques* 20: 1004–1008.

CFIA AND NRCAN/CFS. 2013. Plants of Canada database [online]. Website http://www.plantsofcanada.info.gc.ca/taxa.php?gid=1004196&lang=e [accessed 10 Sept 2013].

CROW, G. E. 1978. Taxonomic revision of *Sagina* (Caryophyllaceae) in North America. *Rhodora* 80: 1–91.

FAIRCLOTH, B. C. 2008. MSATCOMMANDER: Detection of microsatellite repeat arrays and automated, locus-specific primer design. *Molecular Ecology Resources* 8: 92–94.

GIVEN, D. R., AND J. H. SOPER. 1981. The Arctic-alpine element of the vascular flora at Lake Superior. National Museums of Canada, Ottawa, Canada.

HAMILTON, M. B., E. L. PINCUS, A. DI FIORE, AND R. C. FLEISCHER. 1999. Universal linker and ligation procedures for construction of genomic DNA libraries enriched for microsatellites. *BioTechniques* 27: 500–504.

MAGEE, D. W., AND H. E. AHLES. 1999. Flora of the Northeast: A manual of the vascular flora of New England and adjacent New York. University of Massachusetts Press, Amherst, Massachusetts, USA.

ROUSSET, F. 2008. GENEPOP'007: A complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources* 8: 103–106.

SANTANA, Q. C., M. P. A. COETZEE, E. T. STEENKAMP, O. X. MLONYENI, G. N. A. HAMMOND, M. J. WINGFIELD, AND B. D. WINGFIELD. 2009. Microsatellite discovery by deep sequencing of enriched genomic libraries. *BioTechniques* 46: 217–223.

SCHUELKE, M. 2000. An economic method for the fluorescent labeling of PCR fragments. *Nature Biotechnology* 18: 233–234.

SLAVICK, A. D., AND R. A. JANKE. 1993. The vascular flora of Isle Royale National Park, 3rd ed. Isle Royale Natural History Association, Houghton, Michigan, USA.

SOPER, J. H., AND P. F. MAYCOCK. 1963. A community of arctic–alpine plants on the east shore of Lake Superior. *Canadian Journal of Botany* 41: 183–198.

WAHLUND, S. 1928. Composition of populations and correlation appearances viewed in relation to the studies of inheritance. *Hereditas* 11: 65–106.

YOU, F. M., N. HUO, Y. Q. GU, M. C. LUO, Y. MA, D. HANE, G. R. LAZO, ET AL. 2008. BatchPrimer3: A high throughput web application for PCR and sequencing primer design. *BMC Bioinformatics* 9: 253.

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

Appendix 1. Locality information for *Sagina nodosa* individuals sampled in this study from Isle Royale National Park.

Individual ID	Latitude	Longitude	Elevation (m)	Island
224-2011	48.18811	−88.42351	597.29	BP
240-2011	48.18899	−88.42217	612.27	BP
245-2011	48.18728	−88.42492	598.077	BP
143-2011	48.17100	−88.43770	603.599	EI
203-2011	48.17165	−88.43434	607.539	EI
209-2011	48.17322	−88.43180	600.443	EI
176-2011	48.17592	−88.43745	575.213	LI
178-2011	48.17555	−88.43839	591.772	LI
181-2011	48.17695	−88.43556	602.021	LI
192-2011	48.17934	−88.43116	608.33	LI
195-2011	48.18047	−88.42902	612.27	LI
82-2011	48.17985	−88.41906	623.31	NG
86-2011	48.18026	−88.41858	609.117	NG
95-2011	48.17854	−88.42203	612.27	NG
96-2011	48.17901	−88.42100	613.058	NG
98-2011	48.17915	−88.42044	606.752	NG
99-2011	48.17947	−88.41978	610.696	NG
25-2011	48.17079	−88.41974	609.905	SG
34-2011	48.16748	−88.42570	615.423	SG
37-2011	48.16696	−88.42659	611.483	SG
51-2011	48.16883	−88.42611	598.868	SGI
109-2011	48.17619	−88.42807	599.656	SI
131-2011	48.17651	−88.42742	605.174	SI
133-2011	48.17766	−88.42520	604.386	SI
134-2011	48.17740	−88.42559	603.599	SI
152-2011	48.18301	−88.42416	617.001	TI
153-2011	48.18277	−88.42453	622.523	TI
154-2011	48.18259	−88.42494	617.792	TI
163-2011	48.18167	−88.42675	615.423	TI

Note: BP = Blake Point; EI = Edwards Island; LI = Long Island; NG = North Government Island; SG = South Government Island; SGI = South Government Island Islet; SI = Split Island; TI = Third Island.