Novel Microsatellite Loci for Sebaea aurea (Gentianaceae) and Cross-Amplification in Related Species

Authors: Kissling, Jonathan, Bachmann, Olivier, Thali, Marco R., and Segarra-Moragues, José Gabriel

Source: Applications in Plant Sciences, 1(12)

Published By: Botanical Society of America

URL: https://doi.org/10.3732/apps.1300056

Applications in Plant Sciences 2013 1(12): 1300056

PRIMER NOTE

NOVEL MICROSATELLITE LOCI FOR *SEBAEA AUREA* (GENTIANACEAE) AND CROSS-AMPLIFICATION IN RELATED SPECIES[1]

JONATHAN KISSLING[2,3,6], OLIVIER BACHMANN[2], MARCO R. THALI[4], AND JOSÉ GABRIEL SEGARRA-MORAGUES[5]

[2]Institute of Biology, University of Neuchâtel, Rue Emile-Argand 11, 2000 Neuchâtel, Switzerland; [3]Department of Ecology and Evolution, University of Lausanne, 1015 Lausanne, Switzerland; [4]Ecogenics GmbH, Grabenstrasse 11a, 8952 Zürich-Schlieren, Switzerland; and [5]Centro de Investigaciones sobre Desertificación (CIDE-CSIC-UV-GV), Carretera de Moncada-Náquera, Km 4.5, E-46113 Moncada, Valencia, Spain

- *Premise of the study:* Microsatellite loci were developed in *Sebaea aurea* (Gentianaceae) to investigate the functional role of diplostigmaty (i.e., the presence of additional stigmas along the style).
- *Methods and Results:* One hundred seventy-four and 180 microsatellite loci were isolated through 454 shotgun sequencing of genomic and microsatellite-enriched DNA libraries, respectively. Sixteen polymorphic microsatellite loci were characterized, and 12 of them were selected to genotype individuals from two populations. Microsatellite amplification was conducted in two multiplex groups, each containing six microsatellite loci. Cross-species amplification was tested in seven other species of *Sebaea*. The 12 novel microsatellite loci amplified only in the two most closely related species to *S. aurea* (i.e., *S. ambigua* and *S. minutiflora*) and were also polymorphic in these two species.
- *Conclusions:* These results demonstrate the usefulness of this set of newly developed microsatellite loci to investigate the mating system and population genetic structure in *S. aurea* and related species.

Key words: diplostigmaty; Gentianaceae; microsatellites; multiplex; *Sebaea aurea.*

Sebaea aurea (L. f.) Roem. & Schult. is a diploid ($2n = 28$; Kissling et al., 2008) annual herb found in the Western Cape (South Africa). *Sebaea aurea* is a diplostigmatic species, with styles showing two stigmatic areas, one each at the apex and base of the styles (see Kissling et al., 2009a). This character is suspected to provide reproductive assurance and reduce seed discounting (i.e., the formation of self-fertilized seeds from ovules that, if they had not been self-fertilized, would have been cross-fertilized) and, therefore, might provide some evolutionary advantages (Kissling et al., 2009b).

To date, no polymorphic microsatellite markers have been developed in this species or in other species for which microsatellites could be potentially transferred to *S. aurea*. Therefore, the purpose of this study was to isolate and characterize microsatellite loci in *S. aurea* to conduct reproductive biology experiments and to estimate the selfing rate in a progeny array, which will contribute to a better understanding of the functional role of diplostigmaty. Furthermore, this set of microsatellite loci will contribute to the investigation of the distribution of genetic diversity across the species range.

METHODS AND RESULTS

Microsatellite isolation was carried out through two independent 454 pyrosequencing strategies. First, DNA was extracted from 20 individuals of *S. aurea* from the Helderberg Nature Reserve population (Appendix 1) using the DNeasy Plant Mini Kit (QIAGEN, Hombrechtikon, Switzerland) following the manufacturer instructions and then pooled and concentrated. Five hundred nanograms of total DNA was used to construct 454 genomic libraries by the sequencing service from the University of Valencia (Servicio Central de Soporte a la Investigación Experimental [SCSIE], Valencia, Spain) and shotgun sequenced on a GS Junior 454 sequencer (454 Life Sciences, a Roche Company, Barcelona, Spain). We obtained 120,157 reads with an average read length of 468.43 bp that were trimmed of adapter and low-quality regions and assembled into contigs using GS De Novo Assembler implemented in Newbler 2.5p1 (454 Life Sciences, a Roche Company, Madrid, Spain). Generated contigs and unique reads not assigned to contigs were subjected to BLAST analysis, and those matching organellar (chloroplast or mitochondria) sequences were discarded.

We screened all 64,422 unique reads and nonorganellar contigs with iQDD version 1.3.0.0 software (Meglécz et al., 2010). We set the script to identify all possible di-, tri-, tetra-, penta-, and hexanucleotide repeats with a minimum of five repeat units, as well as compound repeats, and to directly design primers using Primer3 (Rozen and Skaletsky, 2000). After discarding reads with too short flanking sequences, primers were successfully designed for 174 reads. Of these 73, 90, and 11 corresponded to di-, tri-, and tetranucleotide microsatellites, respectively.

Additionally, size-selected fragments from genomic DNA were enriched for microsatellites by Ecogenics GmbH (Zürich-Schlieren, Switzerland) using streptavidin-coated magnetic beads and biotin-labeled $(CT)_{13}$ and $(GT)_{13}$ repeat

[1]Manuscript received 22 June 2013; revision accepted 25 July 2013.

The authors thank Ophélie Gning and Anouk Sarr for providing technical assistance. We acknowledge grants to J.K. from the Swiss National Science Foundation (PA00P3_129140) and the Velux Stiftung (project no. 679) and to J.G.S.-M. from a postdoctoral research contract "Ramón y Cajal" from the Ministerio de Ciencia e Innovación (MICINN), Spain.

[6]Author for correspondence: jonathan.kissling@unine.ch

doi:10.3732/apps.1300056

probes. The simple sequence repeat (SSR)–enriched library was sequenced on a Roche 454 platform using the GS FLX Titanium reagents at Ecogenics GmbH. The 34,992 reads sequenced had an average length of 427 bp. Of these, 996 contained a microsatellite insert with a tetra- or a trinucleotide motif of at least six repeat units or a dinucleotide motif of at least 10 repeat units. Suitable primer design using Primer3 (Rozen and Skaletsky, 2000) was possible in 180 reads. Thirty-nine and 36 loci from the first and second sequencing strategies, respectively, were tested for functionality and polymorphism using at least seven individuals.

Amplifications were carried out in 10-μL reactions containing 1–3 ng of template DNA, 1× GoTaq Flexi Buffer (Promega Corporation, Neuchâtel, Switzerland), 1.5 mM $MgCl_2$, 0.2 mM dNTPs, 0.04 μM forward primer with a 5′ M13 tail, 0.16 μM 6-FAM-labeled M13 primer, 0.16 μM reverse primer, and 0.5 U HotStarTaq DNA Polymerase (QIAGEN). PCRs were performed in a Dyad Cycler (Bio-Rad GmbH, Cressier, Neuchâtel, Switzerland), and the PCR cycling conditions consisted of an initial activation step of 5 min at 95°C; followed by 30 cycles each of 45 s at 95°C, 60 s at 55°C, and 60 s at 72°C; eight cycles each of 45 s at 95°C, 60 s at 53°C, and 60 s at 72°C; and a final extension step of 30 min at 72°C. PCR products were run on an ABI3730xl Genetic Analyzer (Applied Biosystems, Rotkreuz, Switzerland) using GeneScan 500 LIZ as an internal size standard (Applied Biosystems). Amplified fragment lengths were assigned to allele sizes with GeneMapper (Applied Biosystems). This screening detected 16 polymorphic loci that consistently amplified (Table 1); however, only 12 with adequate sizes for multiplex amplification were selected to genotype the whole set of samples.

The 12 selected microsatellite loci were arranged into two multiplex reactions (I and II; Table 1). Amplifications were carried out in a 10-μL reaction using Type-it Microsatellite PCR kits (QIAGEN). The PCR cocktail included 5.0 μL 2× Type-it Master Mix, 1.0 μL primer mix (2 μM each forward fluorescent-labeled and unlabeled reverse primer; Table 1), 3.0 μL RNase-free water, and 1 μL template DNA (10–50 ng). The PCR program followed the manufacturer's conditions: an initial activation step at 95°C for 5 min; 28 cycles at 95°C for 30 s, 57°C for 90 s, and 72°C for 30 s; followed by a final extension step at 60°C for 30 min. PCR products were run and sized as in the aforementioned pilot study.

Genotypic data were obtained from three populations of *S. aurea* (Helderberg Nature Reserve, $N = 38$; Somerset West, $N = 25$; Paradyskloof, $N = 30$; Table 2, Appendix 1) for the 12 microsatellite loci (Tables 1 and 2). Number of alleles (A), observed heterozygosity (H_o), and unbiased expected heterozygosity (H_e) (Nei, 1978) were calculated with GENETIX version 4.05 (Belkhir et al., 2004). Linkage disequilibrium between pairs of microsatellite loci and between pairs of microsatellite loci and populations, using 10,000 permutations, were calculated with GENEPOP version 4.1.4 software (Rousset, 2008). This same software was used to estimate the frequency and 95% confidence intervals of estimated null allele frequencies of each locus in each population (Dempster et al., 1977). Of the 186 available pairwise comparisons between loci and populations, only seven showed significant linkage disequilibrium ($P < 0.05$), whereas of the 66 available pairwise comparisons between loci only two showed significant linkage disequilibrium ($P < 0.05$). Of these, only the pair Seba11-Seba13 was consistent between the Helderberg and Paradyskloof populations and the global estimate of linkage disequilibrium; however, none of the pairwise comparisons were significant after Bonferroni correction.

We detected a total of 164 different SSR alleles for the 12 polymorphic microsatellite loci in the three analyzed populations of *S. aurea*. The number of alleles ranged from a minimum of three alleles for locus Seba19849 to a maximum of 29 alleles for locus Seba13, and the mean number of alleles per locus was 13.67 ± 7.13. One locus (Seba19849) was monomorphic in the Somerset West population (Table 2). H_o ranged from 0.0 (locus Seba28206) to 0.868 (locus Seba13) in the Helderberg population, and unbiased H_e in the polymorphic loci ranged from 0.050 (locus Seba19849) to 0.945 (locus Seba28206) (Table 2). Almost half of the loci in each population showed significant deviation from Hardy–Weinberg equilibrium toward heterozygote deficiency, whereas the remaining loci in each population showed either nonsignificant heterozygote deficiency or excess. Global population inbreeding estimates gave significant heterozygote deficiency in all three populations. A likely explanation for this result may be the presence of null alleles. In fact, 17 out of the 36 loci per population comparison gave estimated frequencies of null alleles higher than 0.05. Nonetheless, local population substructure or high selfing rates within populations could also result in high inbreeding coefficient (F_{IS})

TABLE 1. Characteristics of 16 microsatellite loci developed in *Sebaea aurea*.

Locus	Primer sequences (5′–3′)[a]	Repeat motif	Allele size (bp)	T_a (°C)	Multiplex group	GenBank accession no.
Seba04	F: VIC-ATTCACATCGCTTACAGCCC	$(AG)_{10}$	130	56	I	KF218835
	R: AATTTAAGAACGTCGCCGC					
Seba10	F: PET-CCTTTATGTGCAACGGGAAG	$(CT)_8$	135	56	I	KF218836
	R: AATGATCCATGCATTCTGCC					
Seba11	F: 6FAM-TGCCTGTTCGACTTATGCAA	$(CT)_8$	136	56	I	KF218837
	R: GCCATGATTACTTCAGCCGT					
Seba13	F: VIC-CAAAGCACTAGCATTGCATGA	$(GA)_8$	317	56	I	KF218838
	R: GGAAAGGTGCGTTGTTGATT					
Seba17	F: 6FAM-AGTGGTACTCCGCCAACATC	$(AC)_7$	256	56	I	KF218839
	R: TCGTGAATTTGTAGCTTGCG					
Seba21	F: VIC-TTCTTGATGTGGCTCAGCAG	$(TG)_7$	259	56	I	KF218840
	R: GCAAACAGAGTTAGCATTGCAT					
Seba05119	F: NED-CGCCATTTCCAGACAACTCC	$(CGC)_7$	227	56	II	KF218845
	R: GAGGACGGAGTACGAGAACC					
Seba11349	F: VIC-ATGGGAGCAGGGGTTTACTG	$(TGC)_8$	231	56	II	KF218843
	R: TGAGCAGGTGGCCTATCATC					
Seba12491	F: ATCGCTTCAGAGCATTGTGG	$(TTG)_7$	248	56	NI	KF218847
	R: AGCCTTGAAAAGGAGATGCC					
Seba19849	F: 6FAM-TGTCGCAAAAATGATCTACGTTATG	$(ATA)_{12}$	174	56	II	KF218841
	R: CCGAAGAATGGATCGGGAAG					
Seba28414	F: NED-GGCTTTCGAACGGAATGGAG	$(TTG)_8$	136	56	II	KF218844
	R: ACTCCTCCTCAACCCAAACC					
Seba09440	F: CCTTACCCTTGTCCCTTCCC	$(GT)_{12}$	232	56	NI	KF218848
	R: AGAGACTAATATACTCAGCTTGCG					
Seba11773	F: GAAAGTAGCGGCGGTGTTG	$(CGG)_7$	247	56	NI	KF218849
	R: CTTCAGTCTCAGCTTTGCCC					
Seba28206	F: PET-ACAACATGCAACTTGACCATC	$(CT)_{12}$	176	56	II	KF218846
	R: CTCGTTGCTGGTGATTGAGG					
Seba31211	F: 6FAM-TCCACATCCCATTGGAGCAG	$(CA)_{15}$	250	56	II	KF218842
	R: CTGTGTTTTTAGCATGTGGGAG					
Seba31409	F: GCACAGTTGTAGGGCATTGAG	$(CA)_{15}$	241	56	NI	KF218850
	R: TCCGGTGAATCCTGCTAGAC					

Note: NI = not included in multiplex groups; T_a = annealing temperature.
[a]For each forward primer, the fluorescent label is indicated at the 5′ end.

TABLE 2. Results of initial primer screening of 12 polymorphic microsatellite loci in three populations of *Sebaea aurea*.[a,b]

	Helderberg Nature Reserve (N = 38)				Somerset West (N = 25)				Paradyskloof (N = 30)			
Locus	A	H_o	H_e	F_{IS}	A	H_o	H_e	F_{IS}	A	H_o	H_e	F_{IS}
Seba04	10	0.444	0.874	0.496***	7	0.200	0.858	0.776***	13	0.360	0.917	0.612***
Seba10	5	0.658	0.693	0.052^{ns}	5	0.720	0.764	0.059^{ns}	6	0.633	0.688	0.080^{ns}
Seba11	11	0.528	0.847	0.380***	11	0.571	0.893	0.366**	9	0.448	0.844	0.473***
Seba13	16	0.868	0.775	–0.122^{ns}	9	0.480	0.642	0.256**	18	0.833	0.829	–0.005^{ns}
Seba17	6	0.067	0.756	0.915***	7	0.182	0.806	0.778***	10	0.296	0.855	0.658***
Seba21	9	0.290	0.822	0.650***	8	0.160	0.777	0.797***	10	0.621	0.859	0.281^{ns}
Seba05119	7	0.278	0.824	0.669***	8	0.286	0.849	0.669***	8	0.280	0.813	0.660***
Seba11349	6	0.800	0.736	–0.090^{ns}	4	0.583	0.664	0.124^{ns}	5	0.633	0.691	0.085^{ns}
Seba19849	2	0.050	0.050	0.000	1	0.000	0.000	—	3	0.069	0.068	–0.009^{ns}
Seba28414	5	0.850	0.739	–0.156^{ns}	8	0.792	0.823	0.039^{ns}	6	0.500	0.774	0.358**
Seba28206	4	0.000	0.686	1.000***	6	0.227	0.610	0.633***	19	0.571	0.945	0.401***
Seba31211	5	0.750	0.745	–0.007^{ns}	5	0.522	0.693	0.251^{ns}	7	0.536	0.684	0.220*
Average	7.17	0.465	0.712	0.353***	6.58	0.394	0.698	0.443***	9.50	0.482	0.747	0.360***

Note: A = number of alleles per locus; F_{IS} = inbreeding coefficient; H_e = unbiased expected heterozygosity; H_o = observed heterozygosity; N = number of individuals.

[a] See Appendix 1 for population locality information.

[b] Deviations from HWE were statistically significant at *P < 0.05, **P < 0.01, ***P < 0.001; ns = not significant.

values. The analysis of progeny arrays would be required to definitely rule out or confirm the presence of null alleles.

The two multiplex reactions were also tested on three to six individuals of seven different species of *Sebaea* (Appendix 1) to assess for potential cross-amplification of SSR loci. Cross-amplification was only successful in *S. ambigua* Cham. and in *S. minutiflora* Schinz. All 12 microsatellite loci amplified in these two species and were polymorphic. Both of these species belong to the *S. aurea* clade, which exclusively contains the tetramerous *Sebaea* (Kissling et al., unpublished data). No successful amplification was obtained in any of the five remaining species for the 12 assayed loci.

CONCLUSIONS

Twelve microsatellite loci were characterized to investigate the reproductive biology and population structure in *S. aurea*. These markers will serve to estimate outcrossing rates in progeny arrays and therefore help to understand the reproductive function of secondary stigmas in *S. aurea*. Cross-species transferability experiments to *S. ambigua* and *S. minutiflora* expand the usefulness of this set of SSR loci to other species of *Sebaea*.

LITERATURE CITED

BELKHIR, K., P. BORSA, L. CHIKHI, N. RAUFASTE, AND F. BONHOMME. 2004. GENETIX 4.05, logiciel sous Windows™ pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5171, Université de Montpellier II, Montpellier, France.

DEMPSTER, A. P., N. M. LAIRD, AND D. B. RUBIN. 1977. Maximum likelihood from incomplete data via the *EM* algorithm (with discussion). *Journal of the Royal Statistical Society. Series B. Methodological* 39: 1–38.

KISSLING, J., Y.-M. YUAN, P. KÜPFER, AND G. MANSION. 2008. Cytogeography of Gentianaceae-Exaceae in Africa, with a special focus on *Sebaea*: The possible role of dysploidy and polyploidy in the evolution of the tribe. *Botanical Journal of the Linnean Society* 158: 556–566.

KISSLING, J., P. K. ENDRESS, AND G. BERNASCONI. 2009a. Ancestral and monophyletic presence of diplostigmaty in *Sebaea* (Gentianaceae) and its potential role as a morphological mixed mating strategy. *New Phytologist* 184: 303–310.

KISSLING, J., Y. YUAN, P. KÜPFER, AND G. MANSION. 2009b. The polyphyletic genus *Sebaea* (Gentianaceae): A step forward in understanding the morphological and karyological evolution of the Exaceae. *Molecular Phylogenetics and Evolution* 53: 734–748.

MEGLÉCZ, E., C. COSTEDOAT, V. DUBUT, A. GILLES, T. MALAUSA, N. PECH, AND J. F. MARTIN. 2010. QDD: A user-friendly program to select microsatellite markers and design primers from large sequencing projects. *Bioinformatics (Oxford, England)* 26: 403–404.

NEI, M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 89: 583–590.

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. 2000. 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.

Appendix 1. Species, collection locality, geographical coordinates, and herbarium voucher of species of *Sebaea* analyzed in this study. All specimens are deposited at the Université de Neuchâtel herbarium (NEU), Neuchâtel, Switzerland.

Species	Location	Geographical coordinates	Altitude (m)	Voucher no.	*N*
S. ambigua Cham.	South Africa: Western Cape, Pringle Bay	34°22.627′S, 18°49.814′E	40	Kissling and Zeltner 45/2005	4
S. aurea (L. f.) Roem. & Schult.	South Africa: Western Cape, Helderberg Nature Reserve	34°03′43″S, 18°52′24″E	139	Kissling 23/2010	38
	South Africa: Western Cape, Somerset West	34°02′54″S, 18°50′00″E	212	Kissling 33/2010	25
	South Africa: Western Cape, Paradyskloof	33°58′06″S, 018°52′43″E	233	Kissling 22/2010	30
S. exacoides (L.) Schinz	South Africa: Western Cape, Franschoek Pass along R45	33°55.041′S, 19°09.518′E	665	Kissling and Wuerfel 6/2008	4
S. marlothii Gilg	South Africa: Eastern Cape, Naude's Nek	30°43.125′S, 28°08.394′E	2513	Kissling and Zeltner 38/2005	4
S. micrantha Schinz	South Africa: Western Cape, Helderberg Nature Reserve	34°03.647′S, 18°52.258′E	156	Kissling 15/2010	4
S. minutiflora Schinz	South Africa: Western Cape, Gordon's Bay, unconstructed plot in front of BP station, along R44	34°08.513′S, 018°51.112′E	3	Kissling 14/2010	6
S. spathulata Steud.	South Africa: Western Cape, Pringle Bay	34°22.627′S, 018°49.814′E	4	Kissling and Zeltner 46/2005	4
S. thomasii Schinz	Lesotho: road to Sehlabathebe lodge	29°52.692′S, 029°05.342′E	2599	Kissling and Zeltner 26/2005	3
	Lesotho: ca. 1 km from Schelabathebe lodge	29°52.058′S, 029°06.966′E	2461	Kissling and Zeltner 29/2005	3

Note: *N* = number of samples analyzed.