Novel Microsatellite Development and Characterization for Phacelia formosula (Hydrophyllaceae)

Authors: Riser, James P., Schwabe, Anna L., and Neale, Jennifer Ramp

Source: Applications in Plant Sciences, 5(7)

Published By: Botanical Society of America

URL: https://doi.org/10.3732/apps.1700030

Applications in Plant Sciences 2017 5(7): 1700030

PRIMER NOTE

NOVEL MICROSATELLITE DEVELOPMENT AND CHARACTERIZATION FOR *PHACELIA FORMOSULA* (HYDROPHYLLACEAE)[1]

JAMES P. RISER II[2], ANNA L. SCHWABE[2,3], AND JENNIFER RAMP NEALE[2,4]

[2]Denver Botanic Gardens, 909 York Street, Denver, Colorado 80206 USA

- *Premise of the study:* Microsatellite primers were developed to characterize genetic diversity and structuring in the genus *Phacelia* (Hydrophyllaceae) and to further conservation efforts for *P. formosula*.
- *Methods and Results:* Fifteen novel microsatellite primers were developed for *P. formosula*. These were characterized for genetic variation in three separate *P. formosula* populations. Two to nine alleles were found per locus. Overall observed heterozygosity and expected heterozygosity ranged from 0.000 to 0.800 and 0.000 to 0.840, respectively. Additionally, these loci were successfully amplified and showed polymorphism in *P. gina-glenneae* and a potential new *Phacelia* species.
- *Conclusions:* These microsatellite markers will be useful in assessing genetic diversity, structuring, and gene flow within and among populations of the rare *P. formosula*, in addition to related *Phacelia* species. These markers will provide important genetic data needed for appropriate conservation and management of these rare plants.

Key words: Colorado; conservation genetics; Hydrophyllaceae; microsatellite; *Phacelia*; *Phacelia formosula*.

Phacelia Juss. (Hydrophyllaceae) is a speciose genus with approximately 167 (USDA NRCS, 2017) species in the United States, predominantly in western states. We follow the taxonomy of the Boraginales Working Group (Luebert et al., 2016) in conserving *Phacelia* within the family Hydrophyllaceae as opposed to in a subfamily of the Boraginaceae (e.g., APG IV, 2016). In addition to the ongoing debate regarding the status of the Hydrophyllaceae, this maintains agreement with the recent *Flora of Colorado* (Ackerfield, 2015). Given the number of species in both the genus and family, we expect that these markers will have broad applicability for conservation and population-level studies. Additionally, there are many rare and locally endemic species in *Phacelia* (34 species with a G1 or G2 rank; NatureServe, 2017). Whereas previous population genetic studies in *Phacelia* used cpDNA (Levy et al., 1996) or allozymes (Levy and Neal, 1999), we developed the first primers specifically for population-level assessments in the genus.

Phacelia formosula Osterh. (North Park phacelia) is a rare endemic found only in the North Park basin in Jackson County, Colorado, USA. Within this area, *P. formosula* is found in scattered small populations restricted to soils derived from the Coalmont Formation (U.S. Fish and Wildlife Service, 2011). An understanding of the genetic diversity and distribution of *P. formosula* would be extremely useful in guiding management and conservation actions. Currently, these data are lacking for *P. formosula*, as well as other *Phacelia* species.

Here we report the development and characterization of 15 novel microsatellite loci for *Phacelia*, all of which were tested for polymorphism in *P. formosula*. Additionally, we cross-amplified these loci in a presumably closely related species, *P. gina-glenneae* N. D. Atwood & S. L. Welsh, and in a recently discovered population of uncertain specific status (*Phacelia* sp. in Table 1).

METHODS AND RESULTS

Microsatellite development using DNA extracted from silica-dried *P. formosula* leaf tissue was conducted by Ecogenics GmbH (Balgach, St. Gallen, Switzerland). Microsatellite content of the genomic DNA fragments was enriched via biotin-labeled tetranucleotide (GTAT, GATA, AAAC, and AAAG; Roche 454 platform [Basel, Basel-Stadt, Switzerland] with GS FLX Titanium reagents) and dinucleotide (CT and GT; Illumina MiSeq platform [San Diego, California, USA] using the Nano 2 × 250 version 2 format) repeats using magnetic streptavidin beads. The enrichments were multiplexed with additional species and produced libraries with 4264 and 13,858 reads (respectively), which were assessed for microsatellites using Primer3 (Rozen and Skaletsky, 1999). The tetranucleotide reads averaged 415 bp in length with 151 reads containing a tetra- or trinucleotide microsatellite insert ≥6 repeat units. The dinucleotide reads averaged 402 bp in length, and 1502 reads contained a dinucleotide microsatellite insert ≥10 repeat units. Suitable primer design was possible in 83 of the tetranucleotide reads and 918 of the dinucleotide reads, of which 24 and 16 reads (respectively) were tested for functionality and polymorphism in seven samples using the methods of Schuelke (2000). The assessment resulted in 40 loci, which were then narrowed to 15 by the authors based on multiplex potential. These loci were multiplexed in two panels for data collection and analyses (Table 1).

For locus amplification within populations, total genomic DNA was extracted from silica-dried leaf tissue at Denver Botanic Gardens using the Omega E.Z.N.A. DNA Mini Kit (Omega Bio-tek, Norcross, Georgia, USA; short protocol with both elution steps). Amplification was carried out at the Nevada Genomics Center (Reno, Nevada, USA) using two PCR panels with different

---

[1] Manuscript received 30 March 2017; revision accepted 14 May 2017.

Funding to support this work was provided by the U.S. Fish and Wildlife Service Agreement F09AC00281 and the Gladys Cheesman Evans Endowment of the Denver Botanic Gardens.

[3] Current address: School of Biological Sciences, University of Northern Colorado, Greeley, Colorado 80639 USA.

[4] Author for correspondence: nealejr@botanicgardens.org

doi:10.3732/apps.1700030

Table 1. Characteristics of 15 microsatellite loci developed in *Phacelia formosula* from Jackson County, Colorado, USA, including annealing temperature for two separate PCR panels and individual primer concentrations.

Locus	Primer sequences (5′–3′)	Repeat motif	T_a (°C)	Primer concentration (μM)	Fluorescent label	Allele size range (bp)	GenBank accession no.
Phafor_00006	F: GAGTTCGCGGACAAAGATGC R: CTTGTGATCGCGTCCTAAGC	$(TATG)_8$	63.8	0.75	6-FAM	189–197	KP281305
Phafor_00246	F: CCGCCATCTCCTTCTTCCAC R: AGCAATTCGCATTAGGTCCG	$(TACA)_{10}$	63.8	2	NED	203–222	KP281306
Phafor_00567	F: AAAACGCACCGTCTCACTTG R: TTCCGCATAAAACGCCATCG	$(ATAC)_{11}$	62	2	VIC	222–260	KY442304
Phafor_00650	F: ATACTGAAACCGCCGCAATC R: TTACCGCCTTCAAACCCAAC	$(TGTA)_7$	63.8	2	PET	194–222	KP281307
Phafor_00745	F: AGGGTCGACACCATACTTCTC R: GCAGACCAGGCTGTTTAGTG	$(TTCT)_8$	63.8	2	NED	253–271	KP281308
Phafor_01477	F: GCAAGCAATGAACACACTCC R: GTCGTATCAAAAGGCATCATTGG	$(CAA)_7$	62	2	6-FAM	83–96	KP281309
Phafor_01499c	F: TGCAAAAGAGATCCATCCACG R: AGGCCTATTCTGTCCCACATC	$(GA)_{14}$	62	2	PET	228–240	KP281303
Phafor_01817c	F: ATCGGGTTACAGAGATGGGC R: GCCCGTTGGGAGAACAAAAG	$(AG)_{14}$	63.8	1.5	VIC	110–133	KY442307
Phafor_02245	F: TCCCATGTTAGCGAGTAGGC R: TGCGTAATCTTGACTTCTAACGAG	$(ATC)_8$	62	4	6-FAM	208–223	KP281310
Phafor_02638	F: GATCAGCGATCAGTGCGTG R: CAAATAGCACCCAACTCCACC	$(TATG)_{12}$	62	4	NED	220–249	KP281312
Phafor_02824	F: CCGTTGCTGATGTCCTTGTC R: CAACATGGTTGCATGCTTCAG	$(TATC)_{18}$	63.8	4	6-FAM	262–288	KY442308
Phafor_03037	F: AACACATCCATGTTGCACCG R: TGGGTCACAGTTAGGTGATG	$(TATC)_7$	63.8	1.5	VIC	231–239	KP281311
Phafor_03754	F: GGGTTAAGTACTGTCATCTTGTACG R: ATTGAACTTCCACTATCGACTTG	$(TATG)_9$	62	2	NED	165–189	KP281313
Phafor_05461c	F: TTTGTCCAGACGCAAACGAC R: AAGCAGTCGTGAGGAAAACG	$(GT)_{12}$	62	2	6-FAM	155–167	KP281302
Phafor_13597s	F: TGGTGCTAAAGAATATCTGCCG R: ATTGGAAGTGACTAGAACAAGC	$(CT)_{14}$	62	1	VIC	106–140	KP281304

Note: T_a = annealing temperature.

annealing temperatures to maximize multiplexing effectiveness. Final PCR volume was 11 μL: 1 μL aliquot of panel mix (containing: forward primers [labeled with a universal M13 tail: 5′-TGTAAAACGACGGCCAGT-3′], reverse primers [primer concentration varied by loci, see Table 1], a fluorescently labeled [6-FAM, NED, PET, or VIC] 5′ tag, and 20 ng of DNA template) and 10 μL of QIAGEN Multiplex PCR Mastermix (QIAGEN, Hilden, Germany). Reaction conditions were as follows: an initial 15-min 95°C denaturing step; followed by 40 amplification cycles of 95°C for 30 s, 62°C (panel 1) or 63.8°C (panel 2) for 45 s, and 72°C for 45 s; followed by a final elongation step at 72°C for 30 min using a GeneAmp 9700 thermocycler (Applied Biosystems, Carlsbad, California, USA). PCR products were diluted to an appropriate concentration determined by PicoGreen dilution tests. One microliter of diluted PCR product was added to 10 μL of HiDi Formamide with the size standard GeneScan 500 LIZ (Thermo Fisher Scientific, Waltham, Massachusetts, USA) and 7 μL of molecular-grade water followed by electrophoresis on an ABI Prism 3730 DNA Analyzer (Applied Biosystems) at the Nevada Genomics Center. Genotype data were visualized and fragment sizes scored at Denver Botanic Gardens using Geneious version 6.0.6 (Kearse et al., 2012).

GenAlEx version 6.3 (Peakall and Smouse, 2006) was used to calculate observed heterozygosity (H_o) and expected heterozygosity (H_e) and to test for deviation from Hardy–Weinberg equilibrium (HWE). GENEPOP (Raymond and Rousset, 1995; Rousset, 2008) was used to test for linkage disequilibrium for each pair of loci in each population. ML-NullFreq (Kalinowski and Taper, 2006; 10,000 replicates) was used to estimate the frequency of null alleles.

All 15 microsatellite loci were variable and polymorphic in all three of the *P. formosula* populations. The number of alleles per locus ranged from two to nine (JC1, $N = 30$), from two to seven (JC2, $N = 30$), and from two to eight (JC3, $N = 30$). The H_o and H_e at JC1 ranged from 0.133 to 0.800 and 0.124 to 0.840, respectively. At JC2, H_o and H_e ranged from 0.033 to 0.700 and 0.064 to 0.707, respectively. At JC3, the H_o ranged from 0.000 to 0.700 and H_e ranged from 0.067 to 0.742 (Table 2). Three loci (Phafor_00567, Phafor_02245, and Phafor_02824) at the JC1 population and four loci at both the JC2 population (Phafor_00567, Phafor_01817c, Phafor_02245, and Phafor_03754) and the JC3 population (Phafor_00246, Phafor_00567, Phafor_01817c, and Phafor_02245) showed significant deviation from HWE (Table 2). After Bonferroni correction, no evidence of significant linkage disequilibrium was detected. Heterozygote deficiencies, possibly indicating the presence of null alleles, were detected for five loci at the JC1 population (Phafor_00006, Phafor_00567, Phafor_01817c, Phafor_02245, and Phafor_02824), six loci at the JC2 population (Phafor_00567, Phafor_1477, Phafor_01817c, Phafor_02245, Phafor_3754, and Phafor_13597s), and seven loci at the JC3 population (Phafor_00246, Phafor_00567, Phafor_00745, Phafor_01817c, Phafor_02245, Phafor_02824, and Phafor_05461c). All 15 microsatellite loci were successfully cross-amplified in both *P. gina-glenneae* (GC, $N = 30$) and a newly discovered *Phacelia* population (LC, $N = 30$) of uncertain specific status. All of the amplified loci were polymorphic in the *P. gina-glenneae* population while two loci (Phafor_03037 and Phafor_01817c) were monomorphic in the *Phacelia* sp. population (LC) (Table 2).

CONCLUSIONS

The novel microsatellite markers described here are the first developed not only for *Phacelia*, but also for the Hydrophyllaceae. These markers will be valuable for investigating population genetic structure in *Phacelia* and potentially other genera within Hydrophyllaceae. Knowledge of genetic diversity present within and among the scattered populations of the rare *P. formosula* will be used to better manage the known populations to ensure their future persistence. Additionally, these markers will be useful for assessing genetic diversity in a newly discovered population of *Phacelia* that is morphologically similar to *P. formosula* but occurs in different habitats more than 40 km away. Investigating potential gene flow between this new population and existing *P. formosula* populations will be helpful in inferring its specific status. Cross-amplification in *P. gina-glenneae* demonstrates the utility of these markers in assessing genetic diversity in other species of *Phacelia*. Our results indicate the presence of potential null alleles. Several methods (Chapuis and Estoup, 2007) and programs can be used to detect and account for null alleles in population-level analyses (such as Kalinowski and

TABLE 2. Genetic characterization of 15 newly developed polymorphic microsatellite loci in *Phacelia formosula*, *P. gina-glenneae*, and a *Phacelia* population of uncertain specific status.[a]

	Phacelia formosula												*Phacelia* sp.				*Phacelia gina-glenneae*			
	JC1				JC2				JC3				LC				GC			
Locus	*A*	H_o	H_e	HWE[b]	*A*	H_o	H_e	HWE[b]	*A*	H_o	H_e	HWE[b]	*A*	H_o	H_e	HWE[b]	*A*	H_o	H_e	HWE[b]
Phafor_00006	2	0.367	0.486	0.178	3	0.133	0.126	0.985	2	0.467	0.498	0.732	3	0.267	0.500	0.002**	2	0.033	0.033	0.926
Phafor_00246	5	0.200	0.269	0.109	3	0.200	0.240	0.080	2	0.000	0.067	0.000***	2	0.233	0.339	0.087	2	0.267	0.391	0.081
Phafor_00567	4	0.267	0.366	0.001***	4	0.095	0.543	0.000***	4	0.069	0.250	0.000***	3	0.045	0.278	0.000***	3	0.000	0.653	0.003**[c]
Phafor_00650	4	0.633	0.678	0.968	3	0.100	0.096	0.994	3	0.433	0.429	0.752	2	0.167	0.206	0.295	4	0.400	0.408	0.106
Phafor_00745	2	0.133	0.124	0.696	2	0.379	0.414	0.649	2	0.069	0.128	0.013	3	0.200	0.283	0.297	11	0.667	0.622	0.885
Phafor_01477	4	0.733	0.674	0.678	2	0.067	0.064	0.850	3	0.633	0.499	0.391	2	0.033	0.095	0.000***	2	0.233	0.206	0.469
Phafor_01499c	6	0.733	0.745	0.986	5	0.600	0.523	0.976	4	0.567	0.599	0.531	6	0.833	0.789	0.890	6	0.767	0.776	0.296
Phafor_01817c	4	0.367	0.448	0.106	2	0.033	0.095	0.000***	4	0.400	0.433	0.029*	1	0.000	0.000	—[c]	5	0.500	0.536	0.875
Phafor_02245	2	0.133	0.231	0.020*	3	0.100	0.239	0.000***	3	0.286	0.506	0.006**	3	0.417	0.352	0.645	2	0.300	0.255	0.334
Phafor_02638	7	0.633	0.684	0.274	4	0.367	0.321	0.959	4	0.700	0.665	0.540	5	0.586	0.646	0.965	6	0.567	0.574	0.718
Phafor_02824	6	0.286	0.641	0.000***	6	0.333	0.347	1.000	7	0.667	0.742	0.368	3	0.000	0.667	0.112[c]	3	0.000	0.472	0.000***
Phafor_03037	3	0.367	0.445	0.671	3	0.433	0.515	0.615	3	0.367	0.403	0.255	1	0.000	0.000	—	4	0.600	0.446	0.480
Phafor_03754	5	0.633	0.628	0.955	3	0.500	0.496	0.000***	4	0.367	0.474	0.458	2	0.100	0.095	0.773	5	0.267	0.656	0.000***
Phafor_05461c	3	0.333	0.380	0.378	5	0.700	0.637	0.682	4	0.321	0.406	0.295	2	0.300	0.339	0.524	5	0.500	0.572	0.759
Phafor_13597s	9	0.800	0.840	0.725	7	0.567	0.707	0.437	8	0.667	0.681	0.831	8	0.767	0.780	0.944	6	0.700	0.773	0.885

Note: *A* = number of alleles per population; H_e = expected heterozygosity; H_o = observed heterozygosity; HWE = *P* values of χ^2 test of Hardy–Weinberg equilibrium.

[a] Voucher and locality information are provided in Appendix 1.

[b] Significant deviation from Hardy–Weinberg equilibrium expectations noted ($^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$).

[c] Results based on >50% missing data due to poor amplification.

Taper, 2006; van Oosterhout et al., 2006), and we encourage their use with these markers. These microsatellite markers constitute a valuable tool for fine-scale genetic investigations in the genus *Phacelia*, as well as conservation of rare *Phacelia* species.

LITERATURE CITED

Ackerfield, J. 2015. Flora of Colorado. Botanical Research Institute of Texas, Fort Worth, Texas, USA.

APG IV. 2016. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. *Botanical Journal of the Linnean Society* 181: 1–20.

Chapuis, M. P., and A. Estoup. 2007. Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution* 24: 621–631.

Kalinowski, S., and M. Taper. 2006. Maximum likelihood estimation of the frequency of null alleles at microsatellite loci. *Conservation Genetics* 7: 991–995.

Kearse, M., R. Moir, A. Wilson, S. Stones-Havas, M. Cheung, S. Sturrock, S. Buxton, et al. 2012. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics (Oxford, England)* 28: 1647–1649.

Levy, F., and C. L. Neal. 1999. Spatial and temporal genetic structure in chloroplast and allozyme markers in *Phacelia dubia* implicate genetic drift. *Heredity* 82: 422–431.

Levy, F., J. Antonovics, J. E. Boynton, and N. W. Gillham. 1996. A population genetic analysis of chloroplast DNA in *Phacelia. Heredity* 76: 143–155.

Luebert, F., L. Cecchi, M. W. Frohlich, M. Gottschling, C. M. Guilliams, K. E. Hasenstab-Lehman, H. H. Hilger, et al. 2016. Familial classification of the Boraginales. *Taxon* 65: 502–522.

NatureServe. 2017. NatureServe Explorer: An online encyclopedia of life [web application]. Version 7.0. NatureServe, Arlington, Virginia, USA. Website http://explorer.natureserve.org [accessed 15 February 2017].

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.

Raymond, M., and F. Rousset. 1995. GENEPOP (version 1.2): Population genetics software for exact tests and ecumenicism. *Journal of Heredity* 86: 248–249.

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

Rozen, S., and H. Skaletsky. 1999. 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.

Schuelke, M. 2000. An economic method for the fluorescent labeling of PCR fragments: A poor man's approach to genotyping for research and high-throughput diagnostics. *Nature Biotechnology* 18: 233–234.

USDA, NRCS. 2017. The PLANTS Database. National Plant Data Team, Greensboro, North Carolina, USA. Website http://plants.usda.gov [accessed 17 February 2017].

U.S. Fish and Wildlife Service. 2011. 5-Year review of *Phacelia formosula* (North Park phacelia), December 2011. U.S. Fish and Wildlife Service, Western Colorado Field Office, Grand Junction, Colorado, USA.

van Oosterhout, D., D. Weetman, and W. F. Hutchinson. 2006. Estimation and adjustment of microsatellite null alleles in nonequilibrium populations. *Molecular Ecology Notes* 6: 255–256.

Appendix 1. Voucher and general location information for five *Phacelia* populations used in this study.

Species	Population	Locality[a]	N	Voucher (Accession no.)[b]
Phacelia formosula Osterh.	JC1	Jackson County, Colorado, USA	30	*M. Islam 1487* (KHD00062092)
	JC2	Jackson County, Colorado, USA	30	*N. D. Atwood 33622* (BRY-V 0050698)
	JC3	Jackson County, Colorado, USA	30	*N. D. Atwood 33558* (BRY-V 0050700)
Phacelia sp.	LC	Larimer County, Colorado, USA	30	*M. Islam 1489* (KHD00062091)
Phacelia gina-glenneae N. D. Atwood & S. L. Welsh	GC	Grand County, Colorado, USA	30	*M. Islam 12-271* (KHD00051791)

N = number of individuals sampled.

[a] Detailed location information has been omitted due to the protected status of these species.

[b] One voucher was collected from each sampled population. Vouchers were deposited at the Kathryn Kalmbach Herbarium (KHD), Denver Botanic Gardens, Denver, Colorado, USA, or the S. L. Welsh Herbarium (BRY), Brigham Young University, Provo, Utah, USA.