Development of 16 microsatellite markers within the *Camassia* (Agavaceae) species complex and amplification in related taxa[1]

Theresa M. Culley[2,4], Ju-Fang Leng[3], Susan R. Kephart[3], Francis J. Cartieri[2], and Kathryn E. Theiss[3]

[2]Department of Biological Sciences, University of Cincinnati, 614 Rieveschl Hall, Cincinnati, Ohio 45221-0006 USA; and
[3]Department of Biology, Willamette University, 900 State Street, Salem, Oregon 97301 USA

• *Premise of the study:* The North American genus *Camassia* is an ecologically important group whose variability and evolution are little understood, being influenced by hybridization and geographic isolation. We developed microsatellite markers to investigate patterns of gene flow, population structure, and taxonomic relationships within this group.

• *Methods and Results:* Using a traditional approach with biotin-labeled probes, we developed 16 microsatellite primers in three species of *Camassia*: *C. howellii*, *C. leichtlinii*, and *C. quamash*. The number of alleles per locus averaged 3.94 per species, and levels of heterozygosity ranged from 0.000 to 1.00 and 0.033 to 0.917 for observed and expected heterozygosities, respectively. All primers amplified to varying extents in additional species (*C. angusta*, *C. cusickii*, *C. scilloides*) and in putative species in a related genus (*Hastingsia alba*, *H. atropurpurea*, *H. bracteosa*, *H. serpentinicola*).

• *Conclusions:* These microsatellite markers exhibit variation and are useful for ongoing studies of integrative taxonomy and population differentiation within this species complex.

Key words: Agavaceae; camas; *Camassia*; *Hastingsia*; integrative taxonomy; microsatellite.

Morphologically variable complexes of species represent formidable challenges for species delimitation and for understanding the mechanisms involved in their evolution. In these cases, genetic markers such as microsatellites can augment traditional systematic approaches by revealing patterns of gene flow, reproductive isolation, and potential cryptic variation between either sequence-invariant or morphologically indistinguishable groups (e.g., *Frullania* Raddi; Ramaiya et al., 2010). One taxonomically difficult group is *Camassia* Lindl., a monophyletic clade within the Agavaceae (Smith et al., 2008). This genus consists of six North American species of bulbous, spring perennials that are insect pollinated, ecologically important in diverse habitats, and culturally valuable as food plants of native peoples (Tomimatsu et al., 2009). Despite taxonomic distinctions supported by allozyme, morphological, and phenological criteria (Ranker and Schnabel, 1986; Uyeda and Kephart, 2006), species delimitation has been difficult, given high morphological variability and poor cpDNA resolution among both sympatric and geographically isolated taxa. Where the ranges of the most common western species (*C. quamash* (Pursh) Greene and *C. leichtlinii* (Baker) S. Watson) overlap, a total of 10 subspecies exist; potential hybridization among these taxa contributes additional complexity to species definition (Fishbein et al., 2010). To untangle the intricate taxonomic relationships within this difficult species complex, we developed 16 microsatellite markers for all *Camassia* species using a traditional approach. We aim to use these markers to examine the evolutionary processes and patterns within this genus, clarify taxonomic distinctions among morphologically similar taxa, and inform management decisions where *Camassia* is used in restoration and reintroduction efforts.

METHODS AND RESULTS

Microsatellite markers were developed in *Camassia* using the nonradioactive method of Glenn and Schable (2005) with individual samples from three different species. DNA was isolated with either cetyltrimethylammonium bromide (CTAB; following Doyle and Doyle, 1987) or a QIAGEN DNeasy Plant Mini Kit (QIAGEN, Culver City, California, USA) from three individuals of the following species: *C. howellii* S. Watson (sample HRA 5) from Hugo Rest Area in Grants Pass, Oregon (latitude 42.518433°, longitude −123.363021°), *C. leichtlinii* subsp. *suksdorfii* (Greenm.) Gould (sample BFV-12) from Butterfly Valley in Quincy, California (40.012283°, −120.9941190°), and *C. quamash* subsp. *maxima* Gould (sample BPP 99.85) from Bush Pasture Park in Salem, Oregon (44.929083°, −123.035666°). All DNA was treated with RNase, digested with the *Rsa*I restriction enzyme (New England Biolabs, Ipswich, Massachusetts, USA), ligated to the SuperSNX adapter pair (see Glenn and Schable, 2005), and bound to biotinylated oligos CA_{15}, GA_{15}, CAC_{10}, CTC_{10}, and CAG_{10} via PCR, before fragments were recovered with Dynabeads (Invitrogen, Carlsbad, California,

TABLE 1. Microsatellite loci information based on 16 primer pairs developed in *Camassia*.[a]

Locus	Primer sequences (5′–3′)	Repeat motif	Allele size (bp)	A	H_o	H_e	Mix	Color	GenBank accession no.	Species[b]
Camas01	F: TCCAAAAGCCATCCTTATGC R: TGGAAACTCAACGAGCAATG	$(GGA)_6$	166	2.00	0.197	0.257	A	NED	KC427110	*C. quamash*
Camas13	F: CTATTGGATTATAATTGATTGAGC R: ACTGCTGTTAATGTAGGAGGTGGTGATGA	$(GA)_{14}$	369	2.17	0.071	0.195	A	6-FAM	KC427111	*C. quamash*
Camas20	F: CGTCGTGATGACGAAAGTAGA R: CCCATTATTGCCATCACTGT	$(GTG)_5$	151	5.00	0.537	0.643	B	6-FAM	KC427112	*C. quamash*
Camas21	F: CATATAGAGTAACGGAACAACTGC R: TGTCAAGGTAAACATAGCTTTGC	$(AG)_{21}$	247	2.83	0.193	0.331	B	6-FAM	KC427113	*C. quamash*
Camas22	F: TCCACCAAGACATCCCTCTC R: CAGCAAGTGTGCATTCAAGG	$(CTC)_4$	153	1.17	0.167	0.097	B	VIC	KC427114	*C. howellii*
Camas29	F: CGGAGATGCATAGCAACAGA R: ATCCGTTTCTCCACCATAGC	$(GGAAGA)_{11}$	244	10.50	0.434	0.820	A	6-FAM	KC427115	*C. leichtlinii*
Camas33	F: CTAAACTATGTATCTTAACGAGATGC R: GACCTCGAATCCTTCCGGAATGC	$(CA)_6(GA)_{13}$	254	9.00	0.300	0.599	A	6-FAM	KC427123	*C. leichtlinii*
Camas34	F: CTCCCCCAGCTGAATCTACA R: GTCTGTAACTCGCCCAGCA	$(AAAATA)_3$	154	0.50	0.028	0.083	A	6-FAM	KC427124	*C. leichtlinii*
Camas45	F: CCTTCGACCAGGAGTTGAGA R: CATGTGCATGCAACTGTGTC	$(TCC)_5$	156	5.00	0.407	0.529	A	PET	KC427116	*C. leichtlinii*
Camas49	F: TGGTCATGGCTTCAGAGTTG R: TAATGCCCCTGCAGCTTAAC	$(TGG)_5$	238	4.83	0.262	0.370	B	PET	KC427117	*C. leichtlinii*
Camas51	F: TGCAGAAATATGAAAGCCACA R: TGATGATGGCCTCTCAACAG	$(GGA)_4$	237	2.17	0.299	0.323	C	NED	KC427118	*C. leichtlinii*
Camas56	F: TGGAGTAAGCACGAGAGCTG R: ACCCAGTCCAATCCTAGCAA	$(CCT)_8$	206	6.17	0.449	0.546	A	6-FAM	KC427119	*C. leichtlinii*
Camas62	F: AGCGAAAGGATCAAACCTCA R: CAAATAGAAGCGCGGCTAAC	$(GAA)_3$	209	2.00	0.090	0.303	A	PET	KC427120	*C. leichtlinii*
Camas79	F: AACTGCTGGGCGAGTTACAG R: TCGAGAGGCTTCTCTCCATC	$(GCT)_4$	224	3.00	0.143	0.170	A	VIC	KC427124	*C. leichtlinii*
Camas83	F: AGAGGGAAGAGGAGGTGGAG R: CCCTCAATGGATGCAGACTT	$(GCG)_6$	156	2.83	0.380	0.377	A	VIC	KC427121	*C. howellii*
Camas90	F: TGGTGATTATGATGACGACGA R: CATCAAAAGGGGGATCTTGAG	$(CA)_9$	187	3.83	0.137	0.360	B	PET	KC427122	*C. leichtlinii*

Note: A = number alleles per locus; H_e = expected heterozygosity; H_o = observed heterozygosity.

[a] Primer sequences, repeat motifs, fragment sizes, number alleles per locus, and observed and expected heterozygosities are all averaged across the six *Camassia* species studied.

[b] Species in which the primer was initially developed.

USA), and then ligated and transformed into *E. coli* TOP10 competent cells using the TOPO TA Cloning Kit (Invitrogen). Of the several hundred colonies obtained, 236 tested positive for containing inserts; these were amplified via PCR, using plastic toothpicks to transfer each colony directly to the PCR reaction wells. The PCR products were screened for size on 4% agarose gels; 192 products between 450 and 1200 bp were sequenced at the High-Throughput Genomics Center (Seattle, Washington, USA). Of these, 27 (14%) sequences contained a central repeat region and were deemed suitable for primer development. Primer pairs were then designed for each of these sequences using the Primer3 software program (Rozen and Skaletsky, 2000) with the following

TABLE 2. Descriptive statistics for individual microsatellite loci within each of the six sampled *Camassia* species.[a]

Locus	*C. angusta* ($n = 24$)			*C. cusickii* ($n = 10$)			*C. howellii* ($n = 19$)			*C. leichtlinii* ($n = 24$)			*C. quamash* ($n = 23$)			*C. scilloides* ($n = 35$)		
	A	H_o	H_e	A	H_o	H_e	A	H_o	H_e	A	H_o	H_e	A	H_o	H_e	A	H_o	H_e
Camas01	3	0.385	0.654	3	0.400	0.335	1	0.000	0.000	1	0.000	0.000	1	0.000	0.000	3	0.400	0.551
Camas13	1	0.000	0.000	—	—	—	1	0.000	0.000	3	0.000	0.461	6	0.391	0.679	2	0.033	0.033
Camas20	4	0.478	0.484	3	0.400	0.460	4	0.632	0.694	7	0.522	0.810	8	0.591	0.840	4	0.600	0.568
Camas21	3	0.174	0.294	—	—	—	1	0.000	0.000	6	0.333	0.592	4	0.308	0.660	3	0.343	0.439
Camas22	1	0.000	0.000	1	0.000	0.000	2	1.000	0.500	1	0.000	0.000	2	0.000	0.083	—	—	—
Camas29	10	0.273	0.816	6	0.900	0.710	13	0.526	0.856	10	0.278	0.809	9	0.154	0.811	15	0.474	0.917
Camas33	9	0.458	0.769	3	0.111	0.290	—	—	—	18	0.609	0.820	12	0.174	0.890	12	0.448	0.822
Camas34	2	0.167	0.498	—	—	—	1	0.000	0.000	—	—	—	—	—	—	—	—	—
Camas45	5	0.583	0.568	4	0.400	0.595	7	0.474	0.690	2	0.208	0.353	5	0.391	0.486	7	0.382	0.482
Camas49	7	0.100	0.673	1	0.000	0.000	11	0.842	0.813	3	0.200	0.261	1	0.000	0.000	6	0.429	0.471
Camas51	—	—	—	—	—	—	2	0.000	0.500	6	0.792	0.750	5	1.000	0.689	—	—	—
Camas56	7	0.583	0.819	2	0.300	0.255	4	0.158	0.240	3	0.118	0.381	10	0.870	0.849	11	0.667	0.734
Camas62	2	0.208	0.492	—	—	—	3	0.000	0.594	1	0.000	0.000	2	0.000	0.375	4	0.333	0.359
Camas79	3	0.133	0.127	—	—	—	1	0.000	0.000	3	0.063	0.119	5	0.348	0.307	6	0.313	0.465
Camas83	2	0.292	0.457	2	0.200	0.180	3	0.316	0.273	2	0.583	0.413	4	0.435	0.481	4	0.457	0.455
Camas90	4	0.087	0.542	3	0.100	0.515	3	0.158	0.148	6	0.261	0.750	2	0.043	0.043	5	0.171	0.162

Note: — = failed to amplify; A = number alleles per locus; H_e = expected heterozygosity; H_o = observed heterozygosity; n = sample size.

[a] Sample sizes are based on all populations surveyed for each *Camassia* species.

TABLE 3. Amplification success of all microsatellite primers across the following taxa of *Camassia* and *Hastingsia*: *C. angusta* (CAAN), *C. cusickii* (CACU), *C. howellii* (CAHO), *C. leichtlinii* (CALE), *C. quamash* (CAQU), *C. scilloides* (CASC), *H. alba* (HAAL), *H. atropurpurea* (HAAT), *H. bracteosa* (HABR), and *H. serpentinicola* (HASE).[a]

Locus	CAAN (*n* = 24)	CACU (*n* = 10)	CAHO (*n* = 19)	CALE (*n* = 24)	CAQU (*n* = 23)	CASC (*n* = 35)	HAAL (*n* = 2)	HAAT (*n* = 5)	HABR (*n* = 10)	HASE (*n* = 13)
Camas01	+	**	**	**	**	+	+	+	+	+
Camas13	**	—	**	**	**	**	**	**	**	**
Camas20	**	**	**	**	**	**	—	—	—	—
Camas21	**	—	**	**	+	**	—	**	**	**
Camas22	///	**	+[b]	**	**	///	+	**	+	+
Camas29	**	**	**	+	+	+	**	**	**	**
Camas33	**	**	—	**	**	**	**	—	—	—
Camas34	**	+	+	—	—	—	—	+	—	—
Camas45	**	**	**	**	**	**	**	**	**	**
Camas49	**	**	**	+	+	+	—	**	**	+
Camas51	—	—	+	**	**	—	—	—	—	—
Camas56	**	**	**	**	**[b]	**	**	**	**	**
Camas62	**	—	+	+	+	**	—	—	—	—
Camas79	+	—	+	+	**	+	**	**	**	**
Camas83	**	**	**	**	**	**	**	**	**	**
Camas90	**	**	**	**	**	**	—	**	**	**

Note: — = failed to amplify; /// = multiple peaks (not scored); + = inconsistent amplification; ** = consistent and strong amplification; *n* = sample size.

[a] The species in which each set of primers were initially developed are shaded in gray.

[b] Strong amplification of multiple peaks in some samples.

criteria: primer size 16–22 bp, 30–60% GC content, melting temperature 50–60°C, and fragment size 90–400 bp.

The 27 primer pairs were tested separately using three individuals representing different *Camassia* species in 10-µL PCR reactions composed of the following: 5 µL QIAGEN Master Mix (QIAGEN), 1 µL Primer mix (composed of 2 µM of each unlabeled forward and reverse primers), 3.8 µL of dH$_2$O, and 0.2 µL DNA (~10 ng). The PCR conditions, as suggested for the QIAGEN kit, were as follows: 95°C for 15 min, followed by 35 cycles each of 94°C for 30 s, 57°C for 45 s, and 72°C for 60 s, and then a final extension of 72°C for 30 min. Samples were analyzed at Cornell University's Life Sciences Core Laboratory Center on a 3730xl sequencer (Applied Biosystems, Carlsbad, California, USA) using the 500 LIZ internal size standard. Fragment analysis was then conducted with GeneMarker version 1.85 software (SoftGenetics, State College, Pennsylvania, USA). Nine of the 27 primers failed to amplify across all individuals or exhibited multiple banding patterns and were excluded from further analysis.

The resulting 18 primer pairs representing di-, tri-, and hexamer nucleotide repeats (Table 1) were then tested for each of the following six species of *Camassia* (Appendix 1), using samples from at least two populations whenever possible: *C. angusta* (Engelm. & A. Gray) Blank. (*n* = 24), *C. cusickii* S. Watson (10), *C. howellii* (19), *C. leichtlinii* (24), *C. quamash* (23), and *C. scilloides* (Raf.) Cory (35). Primers were also tested in related *Hastingsia* S. Watson species (Appendix 1), consisting of *H. alba* (Durand) S. Watson (2), *H. atropurpurea* Becking (5), *H. bracteosa* S. Watson (10), and *H. serpentinicola* Becking (13). For this analysis, primers were individually labeled with 6-FAM, PET, NED, or VIC, and placed in one of three primer mix solutions (see Table 1) for multiplex PCR using the QIAGEN Multiplexing Kit with the same PCR reaction amounts and thermocycler program as outlined above for testing individual primers. PCR products were sent to Cornell University with the same conditions as described above. At this point, another primer pair was discarded due to inconsistent amplification success. A second primer pair was also removed because it was based on the same library sequence as an existing primer (Camas33), exhibiting identical banding patterns but of different sizes (ca. 400 bp for Camas35 compared to ca. 250 bp for Camas33). Similarly, primers Camas34 and Camas79 were found to be based on the same library sequence but, as they amplified different portions and were not in linkage disequilibrium (see below), they were both included. Analysis of genetic variation of the 16 primer pairs was conducted in Genetic Data Analysis (Lewis and Zaykin, 2001) to quantify the number of alleles (*A*) per locus, as well as observed (*H*$_o$) and expected (*H*$_e$) values of heterozygosity, and tests of Hardy–Weinberg equilibrium and linkage disequilibrium.

All 16 primer pairs amplified in at least one *Camassia* species (Tables 2, 3), and three primer pairs amplified strongly across both genera (Camas45, Camas56, and Camas83), with an additional two pairs performing consistently across all *Camassia* taxa (Camas20 and Camas90). Each primer amplified to some degree within the species from which it was originally developed (Table 3), except for Camas34 which exhibited only very small peaks in *C. leichtlinii* but performed well within other taxa (Table 2). Within *Camassia* (Table 2), the number of alleles per locus ranged from one to 18, averaging 3.94 per species. Levels of heterozygosity for polymorphic loci in *Camassia* ranged from 0.000 to 1.00 and from 0.033 to 0.917 for *H*$_o$ and *H*$_e$, respectively. Up to eight loci were out of Hardy–Weinberg equilibrium in at least one population, reflecting in part the limited sampling, but a given locus did not deviate across all populations. The only primers in linkage disequilibrium were found in *C. angusta* (Camas29/56, Camas33/49, Camas29/49, and Camas62/90) and *C. leichtlinii* (Camas13/51, Camas20/51, Camas33/90, and Camas51/56). We are now using these primers to examine levels of genetic variation and population divergence within and among the entire *Camassia/Hastingsia* species complex.

CONCLUSIONS

Given the morphological and genetic variability as well as taxonomic challenges present in *Camassia* and *Hastingsia*, we anticipate that these microsatellite markers will be instrumental in helping to elucidate species delimitation, hybridization, and lineage differentiation within the complex. Consistent amplification of several markers across at least two genera suggest that they may be useful to further spur research in closely related taxa of the Chlorogaloideae and Agavaceae groups.

LITERATURE CITED

DOYLE, J. J., AND J. L. DOYLE. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* 19: 11–15.

FISHBEIN, M., S. R. KEPHART, M. WILDER, K. M. HALPIN, AND S. L. DATWYLER. 2010. Phylogeny of *Camassia* (Agavaceae) inferred from plastid *rpl16* intron and *trnD-trnY-trnE-trnT* intergenic spacer DNA sequences: Implications for species delimitation. *Systematic Botany* 35: 77–85.

GLENN, T. C., AND N. A. SCHABLE. 2005. Isolating microsatellite DNA loci. *Methods in Enzymology* 395: 202–222.

LEWIS, P. O., AND D. ZAYKIN. 2001. Genetic Data Analysis: Computer program for the analysis of allelic data. Version 1.0 (d16c). University of Connecticut, Storrs, Connecticut, USA. Website http://www.eeb.uconn.edu/people/plewis/software.php [accessed 14 June 2013].

RAMAIYA, M., M. G. JOHNSON, B. SHAW, J. HEINRICHS, J. HENTSCHEL, M. VON KONRAT, P. G. DAVISON, AND A. J. SHAW. 2010. Morphologically cryptic biological species within the liverwort *Frullania asagrayana*. *American Journal of Botany* 97: 1707–1718.

RANKER, T. A., AND A. F. SCHNABEL. 1986. Allozymic and morphological evidence for a progenitor-derivative species pair in *Camassia* (Liliaceae). *Systematic Botany* 11: 433–445.

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.

SMITH, C. I., O. PELLMYR, D. M. ALTHOFF, M. BALCÁZAR-LARA, J. LEEBENS-MACK, AND K. A. SEGRAVES. 2008. Pattern and timing of diversification in *Yucca* (Agavaceae): Specialized pollination does not escalate rates of diversification. *Proceedings of the Royal Society of London. Series B, Biological Sciences* 275: 249–258.

TOMIMATSU, H., S. R. KEPHART, AND M. VELLEND. 2009. Phylogeography of *Camassia quamash* in western North America: Postglacial colonization and transport by indigenous peoples. *Molecular Ecology* 18: 3918–3928.

UYEDA, J. C., AND S. R. KEPHART. 2006. Detecting species boundaries and hybridization in *Camassia quamash* and *C. leichtlinii* (Agavaceae) using allozymes. *Systematic Botany* 31: 642–655.

APPENDIX 1. Sampling information, locality details, and voucher information for analyzed populations of *Camassia* and *Hastingsia*.

Species (Species code)	Population (Population code)	*n*	Latitude	Longitude	Voucher collection no. (Herbarium)[a]
C. angusta (CAAN)	Anderson Prairie, KS (AND)	14	38.181420°	−95.264940°	Archibald 2011-5 (KANU)
	Otterbein, IN (OTT)	10	40.494493°	−87.124091°	Homoya & Dana 90-06-14-66 (INDNR)
C. cusickii (CACU)	Hells Canyon Overlook, OR (HC)	10	45.126024°	−116.836042°	Kephart 600 (WILLU)
C. howellii (CAHO)	Hugo Bend, OR (HUBS)	9	42.580549°	−123.374494°	Kephart 589 (WILLU)
	Hugo Rest Area (HRA)	10	42.518433°	−123.363021°	Dennis s.n. (WILLU)
C. leichtlinii subsp. *suksdorfii* (CALE)	Butterfly Valley, CA (BFV)	4	40.012283°	−120.994119°	Kephart & Theiss 604 (WILLU)
	Mt. Douglas Park, BC (DPK)	10	48.492650°	−123.350350°	Allen 1311 (WILLU)
	TNC Camassia Preserve, OR (TNC)	10	45.361261°	−122.619356°	Kephart 580 (WILLU)
C. quamash subsp. *maxima* (CALE)	Onion Peak, OR (ONP)	3	45.816047°	−123.884561°	Chambers 3175 (OSC)
	TNC Camassia Preserve, OR (TNC)	10	45.360717°	−122.618170°	Kephart 581 (WILLU)
	University of Victoria, BC (VIC)	10	48.460967°	−123.318900°	Allen 1310 (WILLU)
C. scilloides (CASC)	Biesecker Prairie, IN (BIE)	15	41.418433°	−87.468160°	Schnabel s.n. (WILLU)
	Gambill Goose Refuge, TX (GGR)	20	33.700120°	−95.651150°	Holmes, Singhurst & Mink 14513 (BAYLU)
H. alba (HAAL)	Butterfly Valley, CA (BFV)	2	40.01185°	−120.99238°	Hrusa and Dinsdale 7216 (CDA)
H. atropurpurea (HAAT)	Woodcock Fen, OR (WF)	5	42.12809 °	−123.69854°	Halpin 70 (HPSU)
H. bracteosa (HABR)	Howell's Fen, OR (HF)	10	42.23194°	−123.6583°	Lang 1802 (OSC)
H. serpentinicola (HASE)	Rough and Ready, Botanical Wayside, OR (RR)	13	42.0944°	−123.6835°	Lillico 409 (OSC)

Note: *n* = number of individuals, all sampled as silica-dried leaves collected in the field 2004–2012; where appropriate, prior vouchers are cited for rare or very small populations.

[a] Herbarium codes: BAYLU = Baylor University Herbarium; CDA = California Department of Food and Agriculture Herbarium; HPSU = Portland State University Herbarium; INDNR = Indiana Department of Natural Resources herbarium; KANU = University of Kansas Herbarium; OSC = Oregon State University Herbarium; WILLU = Willamette University Herbarium.