Microsatellite Primers in *Agave utahensis* (Asparagaceae), a Keystone Species in the Mojave Desert and Colorado Plateau

Charlee Byers², Peter J. Maughan², Jared Clouse², and J. Ryan Stewart²,³

²Department of Plant and Wildlife Sciences, Brigham Young University, 4105 Life Sciences Building, Provo, Utah 84602 USA

- **Premise of the study:** Utah agave (*Agave utahensis*) and its putative subspecies, *A. utahensis* subsp. *kaibabensis* and *A. utahensis* subsp. *utahensis*, are keystone species of the Mojave Desert and Colorado Plateau in the southwestern United States. Here we developed microsatellite markers to study population structure and genetic diversity of the two subspecies of *A. utahensis*.

- **Methods and Results:** We analyzed 22,386 454-pyrosequencing large contigs (>400 bp), derived from a genome reduction experiment consisting of *A. utahensis* accessions, for putative microsatellites. The use of unique multiplex barcodes for each of the *Agave* accessions allowed for the identification of putatively polymorphic microsatellites based solely on sequence alignment analysis. We report the characteristics of 11 polymorphic microsatellite loci based on a panel of 104 individuals from the two subspecies. The number of alleles per locus varied from three to eight, with an average of 5.5 alleles per locus. Observed and expected heterozygosity values ranged from 0.038 to 0.777 and 0.038 to 0.707, respectively.

- **Conclusions:** The microsatellites identified here will be invaluable for future studies of population structure, polyploidy, and genetic diversity across the species.

**Key words:** *Agave utahensis* subsp. *kaibabensis*; *Agave utahensis* subsp. *utahensis*; Asparagaceae; microsatellite; simple sequence repeat (SSR).

*Agave utahensis* Engelm. (Utah agave; Asparagaceae) and its putative subspecies, *A. utahensis* subsp. *kaibabensis* (McKelvey) Gentry and *A. utahensis* subsp. *utahensis*, play critical roles as keystone taxa in their native habitats. Utah agave contributes significantly to soil formation and the creation of oases of organic matter in the Mojave Desert and Colorado Plateau in the southwestern United States (Gentry, 1982). Their large inflorescences produce copious amounts of nectar for a wide array of pollinators, including bats, birds, bees, and hawkmoths (Slauson, 2001). The rise in frequent and intense fires across the southwestern United States (Gentry, 1982) could hasten its ecological decline if fire frequencies continue to increase.

Moreover, taxonomic questions exist concerning the subspecies ranking within this highly variable species complex. The subspecies ranking of the two Utah agave (subsp. *kaibabensis* and subsp. *utahensis*) is based solely on morphological traits, with subsp. *kaibabensis* having wider and greener leaves (3.0–5.5 cm) and larger rosettes (40–100 cm) than subsp. *utahensis* (1.0–3.0 cm, 15–40 cm, respectively) (Gentry, 1982; Reveal and Hodgson, 2003; Janeba, 2008). DNA marker information will help characterize the genetic diversity of the Utah agave species complex and confirm whether subspecies ranking is justified.

Despite its importance as a keystone species, no molecular tools have been developed to manage Utah agave or to investigate population structure, subspecies ranking, hybridization, or polyploidy within the species. We constructed primers to amplify microsatellite markers from both subspecies along a west-east transect across southwestern Utah and north-central Arizona.

**METHODS AND RESULTS**

The genomic reduction protocol used here is described in detail by Maughan et al. (2009) for single-nucleotide polymorphism discovery. Here we use the same methodology, but analyze the results for putatively polymorphic microsatellite sequences. The method is based on the conservation of restriction sites across individuals followed by genomic selection and en masse fragment sequencing via next-generation sequencing technology. Specific DNA barcodes are incorporated into the DNA fragment of each member of the screening panel, which are subsequently used to deconvolute the sequence read pool and to search for polymorphic microsatellite loci across the assembled DNA fragments. Leaf tissue was lyophilized and DNA extracted using a modified mini-salt protocol as reported by Todd and Vodkin (1996). In short, 450 ng of DNA from two *A. utahensis* subsp. *kaibabensis* and two *A. utahensis* subsp. *utahensis* agave accessions were separately double-digested for 1 h at 37°C with 3 units of the restriction enzymes *EcoRI* and *BfaI* (New England Biolabs, Beverly, Massachusetts, USA) in...
1× NEB4 restriction buffer. The resultant DNA fragments were immediately ligated with 1.5 μM 5′-TEG biotinylated/3′-phosphorylated EcoRI adapters and 15 μM 3′-phosphorylated BflI adapters (see Maughan et al., 2010 for adapter sequences) using 3 units of T4 ligase (New England Biolabs) at 16°C for 3 h. Non-biotin-labeled DNA fragments (BflI or BflI restriction fragments) were removed from the reaction using a biotin-streptavidin paramagnetic bead separation using M-280 streptavidin beads (Invitrogen, Carlsbad, California, USA) according to the manufacturer’s specifications. The remaining EcoRI-BflI and EcoRI-EcoRI fragments, with the biotin label still attached, were resuspended in 100 μL of 10:1 TE buffer (10 mM Tris, 1 mM EDTA, pH 8.0).

Four primer pairs, designed to be complementary to the EcoRI and BflI adapter sequences and to carry unique 5′-bar code sequences for each DNA sample. Amplification of each sample was performed in 50-μL PCR reactions using 1× Advantage HF 2 PCR Master Mix (Clontech, Mountain View, California, USA) and 0.2 μM of the MIDX-EcoRI and MIDX-BflI primer pairs, with the following thermocycling profile: 95°C for 1 min followed by 22 cycles of 95°C for 15 s, 65°C for 30 s, and 68°C for 2 min. Equimolar amounts of DNA for each of the PCR reactions were pooled and sequenced using standard protocols for 454-pyrosequencing as a service at the Brigham Young University DNA Sequencing Center (DNASC, Provo, Utah, USA) using a Roche-454 GS FLX instrument and Titanium reagents (454 Life sciences, Carlsbad, California, USA) as the size standard at the Brigham Young University DNAse. The data generated from the fragment analysis were scored using the microsatellite plugin in Geneious R6 (Drummond et al., 2011) and analyzed using default parameters of the computer program Arlequin version 3.5 (Excoffier and Lischer, 2010).

Four of the 15 preselected markers failed to amplify or produced poor, weak, or ambiguous amplification products. Eleven markers were polymorphic, producing a total of 61 alleles across both subspecies. The number of alleles per locus varied from a low of three alleles (BYUUS866) to a high of eight alleles (BYUUS164 and BYU8677), with an average of 5.5 alleles per marker (Table 2). Interestingly, two of the microsatellite markers (BYU4463 and BYU4988) exhibited individuals (~25%) that amplified between two and four alleles, presumably due to the amplification of two paralogous DNA sequences. The mean observed and expected heterozygosity values for A. utahensis subsp. utahensis and A. utahensis subsp. kaibabensis were 0.357 and 0.491 and 0.357 and 0.408, respectively (Table 2). Division of the diversity panel by putative subspecies ranking exhibited a between-subspecies genetic differentiation (FST) of 0.24. When all individuals were analyzed for Hardy–Weinberg equilibrium (HWE) on a subspecies basis, irrespective of subpopulation, several markers violated HWE (Table 2). This was likely due to population substructure (i.e., Wahlund effect). Indeed, when the data were analyzed for each of the eight subpopulations (Appendix 1), all microsatellites were in HWE for five of the populations, with only BYU4012 or BYU8677 violating HWE in the remaining three populations (note that the sample sizes for these populations are too small for definitive conclusions).

### Table 1. Characteristics of 11 polymorphic microsatellite primers developed for Agave utahensis.

| Locus  | Primer sequences (5′–3′) | Repeat motif | Allele size range (bp) | Tm (°C) | Multiplex set | BLASTX similarity (E-value)a | GenBank accession no. |
|--------|--------------------------|--------------|-----------------------|---------|--------------|----------------------------|----------------------|
| BYU3268 F: VIC-CCTCAATCAATAATTCTCA<br>R: TGAAGACCTCGATCATCC | (GCT)h | 138–147 | 57 | 2 | 1IAA-leucine resistant-like gene<br>XP_007044501.1 (2e-29) | KJ855288 |
| BYU3674 F: VIC-CATCAATATAAACATGAGAAGTTG<br>R: TTGCTACTTGAAGCTACCC | (ACA)h | 147–156 | 57 | 3 | Elongation factor 1-alpha<br>XP_003581610.1 (1e-24) | KJ855289 |
| BYU4012 F: VIC-CTTCTGAGGATAAAGTCTGGA<br>R: CCGCGGAGTAAAGGAGAG | (CAC)h | 131–140 | 57 | 1 | N/A | KJ855290 |
| BYU4463 F: VIC-CGTTGAATCCTGATGACC<br>R: AACCACGACACAGTGACGCC | (GAG)h | 178–199 | 57 | 4 | Allene oxide synthase1<br>NP_001105244.1 (1e-68) | KJ855291 |
| BYU4988 R: NED-GAGGCCTGGATATTTCTCCTG<br>R: TCTCCGATCTTACATCG | (TAT)h | 172–190 | 57 | 1 | N/A | KJ855292 |
| BYU5164 F: 6FAM-ATCGGTCCTCCACAGGGTTCA<br>R: TTGCTCTTCTCCAAGTG | (CCA)h | 155–176 | 57 | 1 | N/A | KJ855293 |
| BYU5866 F: 6FAM-ACAAGGCCCAACACAAATATAT<br>R: GGGCAATCCCTTCAATTTAACA | (AAT)h | 164–170 | 57 | 3 | N/A | KJ855294 |
| BYU7269 F: 6FAM-GATCCTCAATGCCTGACCTC<br>R: GAGCTGACCAAACTACC | (GTT)h | 155–167 | 57 | 2 | N/A | KJ855295 |
| BYU8490 F: 6FAM-CCAAATCTTGGACCAAATG<br>R: CAGAATAAGAGCCTCTTAAATTG | (GTT)h | 173–191 | 57 | 4 | N/A | KJ855296 |
| BYU8677 F: NED-AAAAAACAAAATGACCGCTAC<br>R: TGAAGATGCCGGATGATT<br>BYU9897 F: NED-AGTGCTTCTTTGGGCCCCCTT<br>R: CCGAAGTTGGGAGAGT | (AAT)h | 171–198 | 57 | 2 | N/A | KJ855297 |

Note: N/A = no significant associated hits to the database; Tm = annealing temperature.

*aBLASTX search against RefSeq database (E-value < 1e-15).
Table 2. Genetic properties, including number of alleles detected at each microsatellite marker and observed and expected heterozygosity values, for each subspecies.

| Locus | A. utahensis subs. | A. utahensis subs. | A. utahensis subs. |
|-------|---------------------|---------------------|---------------------|
|       | H_e | H_s | A     | H_e | H_s | A     |
| BYU3268 | 4  | 0.3333 | 0.3409 | 4  | 0.0384 | 0.0384 | 2  |
| BYU3674 | 4  | 0.3866 | 0.3863 | 4  | 0.3043 | 0.2695 | 3  |
| BYU4012 | 4  | 0.1621 | 0.3546 | 4  | 0.1600 | 0.5355 | 3  |
| BYU4463 | 7  | NT   | NT    | 7  | NT   | NT    | 4  |
| BYU4988 | 7  | NT   | NT    | 6  | NT   | NT    | 5  |
| BYU5164 | 8  | 0.5223 | 0.6615 | 8  | 0.7777 | 0.7031 | 4  |
| BYU5866 | 3  | 0.2343 | 0.4760 | 3  | 0.3913 | 0.4492 | 2  |
| BYU7269 | 5  | 0.5384 | 0.6832 | 5  | 0.4230 | 0.5316 | 3  |
| BYU8490 | 7  | 0.3000 | 0.4642 | 6  | 0.0416 | 0.0416 | 2  |
| BYU8677 | 8  | 0.4000 | 0.7067 | 7  | 0.3846 | 0.5399 | 4  |
| BYU9897 | 4  | 0.3333 | 0.3409 | 1  | 0.6956 | 0.5642 | 3  |
| Average | 5.5 | 0.3567 | 0.4905 | 5.0 | 0.3573 | 0.4081 | 3.2 |

Note: A = number of alleles; F_{ST} = differentiation between subspecies; H_e = expected heterozygosity; H_s = observed heterozygosity; n = number of individuals sampled; NT = not tested.

a Bold values indicate a significant deviation (P < 0.001) from Hardy–Weinberg equilibrium when analyzed by subspecies.

b Microsatellite markers BYU4463 and BYU4988 were not tested for H_s because more than two alleles were identified in several of the individuals analyzed.

c F_{ST} = 0.2407 (P < 0.0001).

CONCLUSIONS

We used a novel genomic reduction approach tied to 454-pyrosequencing and DNA sample barcoding to identify polymorphic marker loci in *A. utahensis*. Eleven microsatellite markers were validated on 104 accessions of *A. utahensis* subsp. *utahensis* and subsp. *kaibabensis*. These markers should prove invaluable for performing larger investigations of population structure, genetic diversity, and phylogenetic relationships in agaves.

APPENDIX 1

Geographic origin of the 104 *Agave utahensis* accessions studied. Plants sampled within a location were separated by a minimum of 10 m. Accessions were deposited in the Stanley L. Welsh Herbarium (BRY), Brigham Young University.

| Population location | Subspecies | N | Accession no. | Latitude  | Longitude |
|---------------------|------------|---|---------------|-----------|-----------|
| Silver Reef | utahensis | 14 | 624001 | 37°16′09.7″N | 113°22′41.3″W |
| Castle Cliff | utahensis | 15 | 624005 | 37°03′55.0″N | 113°53′16.0″W |
| Woodbury | utahensis | 17 | 624009 | 37°01′44.1″N | 113°52′11.7″W |
| Lime Kiln Canyon | utahensis | 15 | 624011 | 36°38′41.4″N | 114°01′05.5″W |
| Red Rock Spring | utahensis | 6 | 624016 | 36°33′06.6″N | 114°01′30.2″W |
| Ledges Parkway | utahensis | 11 | 624023 | 37°12′24.7″N | 113°35′59.8″W |
| Navajo Nation GC | kaibabensis | 16 | 624017 | 35°56′14.1″N | 111°41′28.4″W |
| Navajo Nation-Hoggan | kaibabensis | 10 | 624020 | 35°55′42.2″N | 111°34′44.7″W |

Note: N = number of individuals sampled at each location.

LITERATURE CITED

Arizaga, S., and E. Ezcurra. 2002. Propagation mechanisms in *Agave macroacantha* (Agavaceae), a tropical arid-land succulent rosette. *American Journal of Botany* 89: 632–641.

Drummond, A. J., B. Ashton, S. Buxton, M. Cheung, A. Cooper, C. Duran, M. Field, et al. 2011. Geneious v.5.4 created by Biomatters. Website http://www.geneious.com [accessed 20 August 2014].

Excoffier, L., and H. E. L. Lischer. 2010. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* 10: 564–567.

Gentry, H. S. 1982. *Agaves of continental North America.* University of Arizona Press, Tucson, Arizona, USA.

Jana, Z. 2008. *Agave utahensis* Cactus and Succulent Journal 80: 299–303.

Maughan, P. J., S. M. Yourstone, E. N. Jellen, and J. A. Udall. 2009. SNP discovery via genomic reduction, barcoding and 454-pyrosequencing in amaranth. *Plant Genome* 2: 260–270.

Maughan, P. J., S. M. Yourstone, R. L. Byers, S. Smith, and J. A. Udall. 2010. SNP genotyping in mapping populations via genomic reduction and next-generation sequencing: Proof-of-concept. *Plant Genome* 3: 166–178.

Reeves, J. L., and W. C. Hodgson. 2003. *Agave.* In Flora of North America Editorial Committee [eds.], *Flora of North America North of Mexico*, vol. 26. Oxford University Press, New York, New York, USA.

Rozen, S., and H. Skaltsky. 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.

Slauson, L. A. 2001. Insights on the pollination biology of *Agave* (Agavaceae). *Haseltonia* 8: 10–23.

Starch, J. E., D. Block, K. Boulez, S. E. Brenner, S. A. Cherwitz, C. Dagdigian, G. Fue llen, et al. 2002. The Bioperl toolkit: Perl modules for the life sciences. *Genome Research* 12: 1611–1618.

Thiel, T., W. Michalek, R. Varshney, and A. Graner. 2003. Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). *Theoretical and Applied Genetics* 106: 411–422.

Todd, J. J., and L. O. Vodkin. 1996. Duplications that suppress and deletions that restore expression from a chalcone synthase multigene family. *Plant Cell* 8: 687–699.

http://www.bioone.org/loi/apps