

## CHARACTERIZATION OF MICROSATELLITE MARKERS FOR PINEDROPS, *PTEROSPORA ANDROMEDEA* (ERICACEAE), FROM ILLUMINA MiSEQ SEQUENCING<sup>1</sup>

LISA C. GRUBISHA<sup>2,4,5</sup>, BAILEY A. NELSON<sup>2</sup>, NICHOLAS J. DOWIE<sup>3</sup>, STEVEN L. MILLER<sup>3</sup>,  
AND MATTHEW R. KLOOSTER<sup>2</sup>

<sup>2</sup>Centre College, 600 West Walnut Street, Danville, Kentucky 40422 USA; and <sup>3</sup>Department of Botany, University of Wyoming, 1000 E. University Ave., Laramie, Wyoming 82071 USA

- *Premise of the study:* *Pterospora andromedea* (Ericaceae) is a mycoheterotrophic plant endemic to North America with a disjunct distribution. Eastern populations are in decline compared to western populations. Microsatellite loci will allow comparison of genetic diversity in endangered to nonthreatened populations.
- *Methods and Results:* Illumina MiSeq sequencing resulted in development of 12 polymorphic microsatellite loci from 63 perfect microsatellite loci tested. One polymorphic locus was obtained from a traditional enrichment method. These 13 loci were screened across two western and two eastern populations. For western and eastern populations, respectively, number of alleles ranged from one to 10 and one to four, and observed heterozygosity ranged from 0.000 to 0.389 and 0.000 to 0.143.
- *Conclusions:* These are the first microsatellite loci developed for *Pterospora*. They will be useful in conservation efforts of the eastern populations and for examination of population genetic parameters at different geographic scales and comparison with mycorrhizal fungal hosts.

**Key words:** conservation genetics; endangered species; Illumina MiSeq; Monotropeoideae; mycoheterotroph; *Pterospora andromedea*.

Pinedrops, *Pterospora andromedea* Nutt., is a mycoheterotrophic plant and acquires carbon from a photosynthetic plant's mycorrhizal fungus (Leake, 1994). Four species in *Rhizopogon* Fr. subgenus *Amylopogon* (A. H. Sm.) Grubisha & Trappe form ectomycorrhizal symbioses with *Pinus* spp. and are fungal hosts to *P. andromedea* (Cullings et al., 1996; Bidartondo and Bruns, 2002; Dowie et al., 2011; Hazard et al., 2012; Grubisha et al., 2014b). *Pterospora andromedea* is a North American endemic in the subfamily Monotropeoideae (Ericaceae) and has a broad, disjunct distribution occurring in western and eastern regions (Bakshi, 1959). Eastern populations have always been rare compared to robust western populations; however, the eastern range has recently suffered population declines due to a variety of environmental and anthropogenic factors (Schori, 2002).

Population genetic studies of *P. andromedea* and the two primary *Rhizopogon* host species (*R. salebrosus* A. H. Sm. and *R. kretzeriae* Grubisha, Dowie & Mill.) that are currently being

conducted will provide information on evolution and maintenance of this symbiosis. Furthermore, population genetic analyses will be useful in assessing population viability that will aid in conservation efforts of both fungal host and plant. The microsatellite loci described here are the first developed for *P. andromedea*.

### METHODS AND RESULTS

Initial isolation of microsatellite loci followed the enrichment method of Glenn and Schable (2005) as described by Klooster et al. (2009). After cells were plated and incubated, hundreds of positive bacterial colonies with successful insertions were obtained. From these, 144 were individually selected and amplified using PCR with 99 (68%) of the inserts falling within the desired size limits of 500–1100 bp. These were then sequenced and screened for the presence of microsatellite regions. Of these sequenced products, 33 fragments possessed microsatellite loci consisting of di-, tri-, and tetranucleotide repeats ranging from five to 22 repeat units with suitable flanking sequences for primer design. Primers for these 33 loci were designed using Primer3 (Rozen and Skaletsky, 2000) with default parameters. These loci were screened for positive PCR amplification using agarose gel electrophoresis following Klooster et al. (2009). From the 33 loci tested, 19 were chosen for screening using fluorescently labeled (6FAM, VIC, PET, NED; Applied Biosystems, Foster City, California, USA) forward primers as described by Klooster et al. (2009). Fragment analysis was conducted using the GeneScan 500 LIZ Size Standard (Applied Biosystems) on an ABI 3730 DNA Analyzer (Applied Biosystems) by the Biotechnology Resource Center (BRC) at Cornell University. Allele sizes were called manually using the Microsatellite Plugin in Geneious version R6.1.5 (Drummond et al., 2011). Only one polymorphic locus was identified (Ptan64; Table 1).

Next-generation sequencing was used as an alternative method for acquiring a large quantity of genomic sequence data from which to identify microsatellite

<sup>1</sup>Manuscript received 4 August 2014; revision accepted 18 September 2014.

Funding for this research has been provided by the National Science Foundation (grant DEB-1050315 to M.R.K. and DEB-1050292 to S.L.M.). The authors thank Ken Cullings for assistance collecting plants in southwestern Wyoming.

<sup>4</sup>Author for correspondence: grubishl@uwgb.edu

<sup>5</sup>Current address: Department of Natural and Applied Sciences, University of Wisconsin–Green Bay, 2420 Nicolet Drive, Green Bay, Wisconsin 54311 USA.

doi:10.3732/apps.1400072

TABLE 1. Characteristics of 13 polymorphic microsatellite loci developed in *Pterospora andromedea*.

Locus	5' end-labeled dye	Primer sequences (5'–3')	Repeat motif	Allele size (bp)	GenBank accession no.
Ptan1	6FAM	F: CTCTCGGAACAGTCATTTCTCC R: CCGATTCAAGGTAGCACTGC	(ACAT) <sub>13</sub>	319	KJ617090
Ptan13	PET	F: ATTGGTTGTGGGACTGCATC R: CTGGGCATAGGGTGAAACTC	(CAGT) <sub>7</sub>	156	KJ617091
Ptan15	NED	F: TGGGAGTTGGAGCTAGGAAC R: AAAGCCGGTAACAGTTCCG	(ATTT) <sub>7</sub>	144	KJ617092
Ptan22	PET	F: TCCTCATAGCCTTCGTCTGTG R: ATTCTCTCTGCTCCCAACCG	(AG) <sub>12</sub>	100	KJ617093
Ptan23	VIC	F: AACCTGCGCATTTCTAAGCC R: GAGGGTGGTGAAAGGGTGAG	(CT) <sub>13</sub>	122	KJ617094
Ptan25	VIC	F: GTCATCGTCCCGGATTATGC R: CCTCCTCCACCAATGCTCTC	(ATGT) <sub>13</sub>	143	KJ617095
Ptan32	VIC	F: GTTAGGGCTCTCGGGTGATC R: GAGGGTAGAGATGGAGGCTATG	(CT) <sub>13</sub>	172	KJ617096
Ptan36	6FAM	F: CAAAGGAAGATGGCTCAAGTTC R: TTTCACGTCTGCGAGCAAC	(AT) <sub>12</sub>	167	KJ617097
Ptan50	NED	F: GCCGATCCCATTGACTGAG R: AACAAAGCTCCACAGGAATGC	(GGTT) <sub>8</sub>	366	KJ617098
Ptan52	VIC	F: GGGTGTGAGTCTTACGTG R: AGGCTACCATTATGCGAAACC	(GGGT) <sub>6</sub>	295	KJ617099
Ptan55	VIC	F: TAGCGCTTACTGGGTTGTG R: CAGGAGCTGCATCAACATGG	(AAACC) <sub>7</sub>	380	KJ617100
Ptan62	NED	F: GGCTGTTGATAGTGGCTGTC R: TTCTTAAAGCACGAATCTGCC	(GCGT) <sub>8</sub>	182	KJ617101
Ptan64	6FAM	F: TGAGGGTAGGATGGTCAATTA R: CGACGCGTACTTCAACCTTA	(AG) <sub>9</sub> GG(AG) <sub>3</sub> AA(AG) <sub>3</sub>	235	KJ617102

repeats. Silica-dried tissue (150 mg) from one plant collected outside of Laramie, Wyoming (41.25108°N, -105.41298°W) was ground in liquid nitrogen to a fine powder. Genomic DNA (gDNA) was isolated using the DNeasy Plant Mini Kit (QIAGEN, Valencia, California, USA). The Genomic DNA Clean & Concentrator Kit (Zymo Research Corporation, Irvine, California, USA) was used to concentrate approximately 600 µL of gDNA into a 30-µL volume. Library preparation and Illumina MiSeq sequencing (Illumina, San Diego, California, USA) were performed at the Advanced Genetic Technologies Centre (AGTC) at the University of Kentucky. Approximately 2 µg of gDNA was sheared using a Bioruptor NGS (UCD-600TS; Diagenode, Denville, New Jersey, USA) by sonicating at 4°C with six cycles of 5 s on and 90 s off. A TruSeq DNA Sample Preparation Kit version 2 (Illumina) was used to generate a paired-end library for Illumina MiSeq sequencing by L.C.G. at the AGTC. MiSeq sequencing resulted in 33,142,988 reads with an average length of 248 bases and 8,228,995,164 total bases. At the AGTC, raw sequence reads were filtered, reformatted, and trimmed using prinseq-lite.pl (Schmieder and Edwards, 2011) and entailed removing (1) duplicate reads, (2) reads with uncalled bases (N) >2% of read length, (3) reads with low-quality scores (Q < 17), (4) very short reads (average length - 2[standard deviation]), and (5) very long reads (average length + 2[standard deviation]). Contigs were assembled de novo using CLC Genomics Workbench version 5.1 (CLC bio, Aarhus, Denmark) by the AGTC producing 2,220,121 contigs, with an average size of 299 bp (N75 = 234, N50 = 330, N25 = 476) and a total of 663,389,430 bases. MSATCOMMANDER version 1.0.8 (Faircloth, 2008) identified the following number of perfect microsatellite repeats from assembled contigs: 12,151 dinucleotide with at least 12 repeat units, 456 trinucleotide with at least eight repeat units, 159 tetranucleotide with at least six repeat units, 65 pentanucleotide with at least six repeat units, and 56 hexanucleotide with at least six repeat units. Within MSATCOMMANDER, Primer3 (Rozen and Skaletsky, 2000) was used to generate primers for all loci using default parameters except that GC clamp = yes, maximum poly X = 3, and an optimum acceptable primer melting temperature of 60°C with a maximum difference of 2°C between forward and reverse primers. To avoid PCR product sizes at 250 bp that would be difficult to size correctly with the DNA standard (LIZ500), in this study PCR product size was set to 90–210 bp or 270–400 bp. The spreadsheet produced by MSATCOMMANDER was used to select a small subset of 63 loci with optimum primer conditions and loci that had no lowercase letters in the primer sequence that would indicate nucleotide mismatches in the assembled reads used to create the contigs.

Sixty-three loci were screened for positive PCR amplification using agarose gel electrophoresis against four *P. andromedea* isolates (two each from western and eastern regions; Table 2). PCR was conducted in a 10-µL reaction volume that included 0.1 µM of each forward and reverse primer, 1× CoralLoad PCR Buffer with 1.5 mM MgCl<sub>2</sub> (QIAGEN), 200 µM each dNTP, 2.5 units *Taq* DNA Polymerase (QIAGEN), and 1.0 µL of 1 : 10 diluted genomic DNA. Thermocycler parameters were 94°C for 3 min; 35 cycles of 94°C for 30 s, 55°C for 30 s, and 72°C for 30 s; followed by a final extension for 10 min at 72°C. PCR products were visualized in a 2% agarose gel stained with GelRed (Phenix Research Products, Candler, North Carolina, USA) on a UV transilluminator. Thirty-six loci that had one or two bright PCR products on the gels in the approximate expected size range (target size +40 bp, -20 bp) and no PCR bands outside of the expected size range were chosen for further screening.

Populations from Wyoming (*n* = 18), Washington (*n* = 17), Michigan (*n* = 7), and Quebec (*n* = 13) representing two western and two eastern populations, respectively, were chosen to screen the 36 loci (Table 2). Stem bract and root samples were immediately placed in Ziploc bags with 2–4-mm silica gel beads (Conservation Support Systems, Santa Barbara, California, USA) or preserved using the method of Dowie et al. (2011). Vouchers were deposited at the University of Michigan Herbarium (voucher no. MICH1485963, MICH1485964; Table 2). Plant DNA was isolated from stem bracts and/or roots. Stem bracts were homogenized using 2 × 0.5-mm ceramic beads in a FastPrep FP120 (Savant Bio101, Carlsbad, California, USA). DNA was isolated using the DNeasy Plant Mini Kit (QIAGEN). DNA from roots was isolated following Dowie et al. (2011) or Grubisha et al. (2014b). The forward primer for 36 loci was 5' end-labeled with one of four dyes: NED, VIC, 6FAM, or PET (Applied Biosystems; Table 1). PCR amplification was performed using the QIAGEN Multiplex PCR Kit in a 5-µL volume with 1× QIAGEN Multiplex PCR Master Mix, 50 nM each primer (exceptions noted below), and 0.75 µL of 1 : 10 diluted genomic DNA. Primers for Ptan1 were used at a concentration of 0.1 µM, 75 nM for Ptan13, Ptan22, and Ptan64, and 35 nM for Ptan25. Touchdown thermocycler conditions were: 95°C for 15 min; 10 cycles of 94°C for 30 s, 67°C for 90 s, decreasing 1°C each cycle, and 72°C for 30 s; 25 cycles of 94°C for 30 s, 57°C for 90 s, and 72°C for 30 s; with a final extension of 60 min at 60°C. Fragment analysis and genotyping were as described above.

Of the 36 loci tested from the Illumina data with fluorescently labeled primers, 12 (33%) were polymorphic (Table 1), nine (25%) were monomorphic (Table 3), 12 (33%) were not useable due to stutter or anomalous additional

TABLE 2. Genetic properties of 13 polymorphic microsatellite loci developed in *Pterospora andromedea*.<sup>a,b</sup>

Locus	Wyoming (n = 18)			Washington (n = 17)			Quebec (n = 13)			Michigan (n = 7)		
	A	H <sub>o</sub>	H <sub>e</sub>	A	H <sub>o</sub>	H <sub>e</sub>	A	H <sub>o</sub>	H <sub>e</sub>	A	H <sub>o</sub>	H <sub>e</sub>
Ptan1	8	0.389	0.848	8	0.000	0.834	4	0.000	0.689	3	0.000	0.484
Ptan13	2	0.111	0.457	2	0.000	0.499	1	0.000	0.000	1	0.000	0.000
Ptan15	1	0.000	0.000	1	0.000	0.000	1	0.000	0.000	2	0.143	0.143
Ptan22	6	0.294	0.768	6	0.176	0.800	2	0.000	0.517	1	0.000	0.000
Ptan23	2	0.278	0.246	5	0.059	0.725	2	0.000	0.517	2	0.000	0.440
Ptan25	5	0.389	0.598	6	0.118	0.788	1	0.000	0.000	1	0.000	0.000
Ptan32	3	0.167	0.367	5	0.118	0.578	3	0.000	0.591	1	0.000	0.000
Ptan36	3	0.167	0.522	4	0.118	0.594	1	0.000	0.000	1	0.000	0.000
Ptan50	2	0.278	0.500	4	0.118	0.405	1	0.000	0.000	1	0.000	0.000
Ptan52	2	0.167	0.157	3	0.059	0.415	1	0.000	0.000	1	0.000	0.000
Ptan55	1	0.000	0.000	1	0.000	0.000	1	0.000	0.000	2	0.000	0.440
Ptan62	1	0.000	0.000	4	0.000	0.656	2	0.000	0.517	2	0.000	0.440
Ptan64	3	0.167	0.427	10	0.176	0.847	3	0.000	0.665	3	0.143	0.670

Note: A = number of alleles; H<sub>e</sub> = expected heterozygosity; H<sub>o</sub> = observed heterozygosity; n = sample size.

<sup>a</sup>Vouchers for two populations were deposited in the University of Michigan herbarium: MICH1485963 (Washington), MICH1485964 (Quebec). Only a very small portion of plant roots were collected in Michigan and Wyoming and vouchers were not possible.

<sup>b</sup>Geographic coordinates for the populations: Wyoming = 44.55816°N, -110.39405°W and Washington = 48.54538°N, -119.07922°W. Due to the rare to endangered status of *Pterospora andromedea* in Michigan (46.6427°N, -89.1787°W) and Quebec (47.3166°N, -79.4533°W) locations are approximate.

peaks, and three (8%) had very weak to no amplification. For the 13 polymorphic loci (12 from Illumina data and one from the enrichment method), number of alleles, observed heterozygosity (H<sub>o</sub>), and expected heterozygosity (H<sub>e</sub>) were calculated in GenAlEx version 6.5 (Table 2; Peakall and Smouse, 2006, 2012). There were one to 10 alleles within western populations, and H<sub>e</sub> ranged from 0.000 to 0.848. Within the eastern populations there were one to four alleles, and H<sub>e</sub> ranged from 0.000 to 0.689. Genotype independence of loci across all pairs of loci within and among populations was tested using the Web-based version of GENEPOP 4.2 (Raymond and Rousset, 1995; Rousset, 2008). After Bonferroni correction (Rice, 1989), significant genotypic linkage disequilibrium was found in three populations: three pairs of loci in Quebec, two pairs of loci in Wyoming, and one pair of loci in Washington. The pairs of loci in linkage disequilibrium were not consistent across populations. When all populations were considered, seven pairwise comparisons had significant linkage disequilibrium and involved

two loci (Ptan22 and Ptan23) in two and four of the pairwise comparisons, respectively.

## CONCLUSIONS

The 13 polymorphic microsatellite loci developed here are the first for *P. andromedea*. Microsatellite loci for the two primary *Rhizopogon* hosts, *R. kretzerae* and *R. salebrosus*, were recently characterized (Grubisha et al., 2014a). These loci are currently being used in population genetic studies of *P. andromedea* and *Rhizopogon* mycobionts to examine genetic diversity and population genetic structure at different hierarchical levels. Furthermore, conservation genetic studies of the endangered eastern populations will provide baseline genetic data for management of populations.

TABLE 3. Characteristics of nine monomorphic microsatellite loci developed in *Pterospora andromedea*.

Locus	5' end-labeled dye	Primer sequences (5'–3')	Repeat motif	Allele size (bp)	T <sub>a</sub> (°C)	GenBank accession no.
Ptan6	PET	F: TCCAAATGACACCCAACATG R: GTTTGGATCTTTAGACCGAGG	(AT) <sub>12</sub>	136	57	KJ617103
Ptan11	PET	F: AACACAGGAGGACTAGAGG R: GGTTTACTGTTGGTAGGAGTTG	(AAC) <sub>12</sub>	181	57	KJ617104
Ptan20	NED	F: ACACGCTACGATAATTCCACG R: CAGGCTGCTTGGTGTAGAATG	(ATC) <sub>8</sub>	197	57	KJ617105
Ptan21	PET	F: TCACACAGCTACATTCCAG R: TTCCTCAGCAATGCCCAAAG	(AC) <sub>13</sub> (CA) <sub>4</sub>	112	57	KJ617106
Ptan24	6FAM	F: ACTAAGTCAGCCAGCATGTG R: GGATACGACAGCCTTTGGATC	(AT) <sub>15</sub>	127	57	KJ617107
Ptan29	PET	F: ACGTTGGTCATTCTGGTGGAG R: AGCCACACAGGACATCTAG	(AAT) <sub>8</sub>	285	57	KJ617108
Ptan31	NED	F: GTTGAGAGCTCCTTCATCACC R: GCCCTCGTCAAATCATCAAGG	(CTT) <sub>10</sub>	123	57	KJ617109
Ptan41	PET	F: AGGGTAGAAGGTTTGGAGGG R: AGGGCATAAGTTCTTGAAGCG	(AAAAT) <sub>8</sub>	339	57	KJ617110
Ptan57	NED	F: CCACGAAATGCAGTGTGAC R: AAGACGACAGCCAGACAATG	(ATCCCC) <sub>6</sub>	110	57	KJ617111

Note: T<sub>a</sub> = annealing temperature.

## LITERATURE CITED

- BAKSHI, T. S. 1959. Ecology and morphology of *Pterospora andromedea*. *Botanical Gazette (Chicago, Ill.)* 120: 203–217.
- BIDARTONDO, M. I., AND T. D. BRUNS. 2002. Fine-level mycorrhizal specificity in the Monotropoideae (Ericaceae): Specificity for fungal species groups. *Molecular Ecology* 11: 557–569.
- CULLINGS, K. W., T. M. SZARO, AND T. D. BRUNS. 1996. Evolution of extreme specialization within a lineage of ectomycorrhizal epiparasites. *Nature* 379: 63–66.
- DOWIE, N. J., J. J. HEMENWAY, S. M. TROWBRIDGE, AND S. L. MILLER. 2011. Mycobiont overlap between two mycoheterotrophic genera of Monotropoideae (*Pterospora andromedea* and *Sarcodes sanguinea*) found in the Greater Yellowstone Ecosystem. *Symbiosis* 54: 29–36.
- DRUMMOND, A. J., B. ASHTON, S. BUXTON, M. CHEUNG, A. COOPER, C. DURAN, M. FIELD, ET AL. 2011. Geneious v5.4 created by Biomatters. Website <http://www.geneious.com> [accessed 20 February 2014].
- FAIRCLOTH, B. C. 2008. MSATCOMMANDER: Detection of microsatellite repeat arrays and automated, locus-specific primer design. *Molecular Ecology Resources* 8: 92–94.
- GLENN, T. C., AND N. A. SCHABLE. 2005. Isolating microsatellite DNA loci. *Methods in Enzymology* 395: 202–222.
- GRUBISHA, L. C., J. D. BREWER, N. J. DOWIE, S. L. MILLER, S. M. TROWBRIDGE, AND M. R. KLOOSTER. 2014a. Microsatellite primers for fungi *Rhizopogon kretzerae* and *R. salebrosus* (Rhizopogonaceae) from 454 shotgun pyrosequencing. *Applications in Plant Sciences* 2(7): 1400029.
- GRUBISHA, L. C., N. J. DOWIE, S. L. MILLER, C. HAZARD, S. M. TROWBRIDGE, T. R. HORTON, AND M. R. KLOOSTER. 2014b. *Rhizopogon kretzerae* sp. nov.: The rare fungal symbiont in the tripartite system with *Pterospora andromedea* and *Pinus strobus*. *Botany* 92: 527–534.
- HAZARD, C., E. A. LILLESKOV, AND T. R. HORTON. 2012. Is rarity of pine-drops (*Pterospora andromedea*) in eastern North America linked to rarity of its unique fungal symbiont? *Mycorrhiza* 22: 393–402.
- KLOOSTER, M. R., A. W. HOENLE, AND T. M. CULLEY. 2009. Characterization of microsatellite loci in the myco-heterotrophic plant *Monotropa hypopitys* (Ericaceae) and amplification in related taxa. *Molecular Ecology Resources* 9: 219–221.
- LEAKE, J. R. 1994. Tansley Review, 69. The biology of mycoheterotrophic ('saprophytic') plants. *New Phytologist* 127: 171–216.
- PEAKALL, R. E., AND P. E. SMOUSE. 2006. GenAIEX 6: Genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* 6: 288–295.
- PEAKALL, R. E., AND P. E. SMOUSE. 2012. GenAIEX 6.5: Genetic analysis in Excel. Population genetic software for teaching and research: An update. *Bioinformatics (Oxford, England)* 28: 2537–2539.
- RAYMOND, M., AND F. ROUSSET. 1995. GENEPOP (version 1.2): Population genetics software for exact tests and ecumenicism. *Journal of Heredity* 86: 248–249.
- RICE, W. R. 1989. Analyzing tables of statistical tests. *Evolution* 43: 223–225.
- ROUSSET, F. 2008. GENEPOP'007: A complete reimplement of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources* 8: 103–106.
- ROZEN, S., AND H. J. 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.
- SCHMIEDER, R., AND R. EDWARDS. 2011. Quality control and preprocessing of metagenomic datasets. *Bioinformatics* 27: 863–864.
- SCHORI, A. 2002. *Pterospora andromedea* Nutt. (Pinedrops). New England Plant Conservation Program, Conservation and Research Plan for New England. New England Wild Flower Society, Framingham, Massachusetts, USA. Website <http://www.newfs.org/docs/pdf/Pterosporaandromedeaweb.pdf> [accessed 7 May 2014].