ONLINE RESOURCES

Identification and characterization of novel UniGene-derived microsatellite markers in *Podophyllum hexandrum* (Berberidaceae)

AKSHAY NAG[1], PANKAJ BHARDWAJ[1,2], PARAMVIR SINGH AHUJA[1] and RAM KUMAR SHARMA[1]*

[1]*Biotechnology Division, CSIR-Institute of Himalayan Bioresource Technology (IHBT; Council of Scientific and Industrial Research), Post Box 6, Palampur 176 061, India*
[2]*Present address: Department of Biotechnology, Central University of Punjab, Bathinda 151 001, India*

[Nag A., Bhardwaj P., Ahuja P. S. and Sharma R. K. 2013 Identification and characterization of novel UniGene-derived microsatellite markers in *Podophyllum hexandrum* (Berberidaceae). *J. Genet.* **92**, e4–e7. Online only: http://www.ias.ac.in/jgenet/OnlineResources/92/e4.pdf]

Introduction

Podophyllum hexandrum Royle (syn, *Sinopodophyllum hexandrum*; Berberidaceae), commonly known as Himalayan mayapple, is a rhizomatous species of great medicinal importance (Nag and Rajkumar 2011). Its distribution is confined to alpine regions of the Himalayas. In India *P. hexandrum* is found from Ladakh to Sikkim at altitudes of 3000–4200 m. It is known for its anticancer properties. The rhizomes and roots of *P. hexandrum* contain antitumour lignans such as podophyllotoxin, 4′-dimethyl podophyllotoxin and podophyllotoxin 4-O-glucoside (Tyler *et al.* 1988; Broomhead and Dewick 1990). Of these lignans, podophyllotoxin is the most important for its use in the semisynthesis of anticancer drugs, etoposide and teniposide (Issel *et al.* 1984). Podophyllotoxin acts as an inhibitor of microtubule assembly. These drugs are widely used in treatment of lung cancer, testicular cancer, neuroblastoma, hepatoma and other tumours. Podophyllotoxin also shows antiviral activity and it interferes with critical viral processes (Giri and Narasu 2000). Podophyllotoxin content of Himalayan mayapple is quite high (4.3%) compared to that of *P. peltatum* (0.25%), the most common species in the Americas (Jackson and Dewick 1984). While *P. hexandrum* has a wide region of distribution, within that region it appears mostly in valleys with secondary vegetation. In any population, the plant shows a kind of clumped distribution pattern. Earlier, *P. hexandrum* was used in folk medicine by local healers in small quantities, but commercialization of the plant for its medicinal attributes in recent years has increased demand and consequent exploitation. The size of the wild populations has been declining owing to overexploitation, habitat fragmentation, long dormancy, and low rate of natural regeneration. The population size of *P. hexandrum* in the Himalayas is very low (40–700 plants per location) and is declining rapidly each year. Some populations in certain pockets have virtually disappeared owing to anthropogenic activities and overexploitation (Bhadula *et al*. 1996). Therefore, *P. hexandrum* has been classified as an endangered species in India since 1987 (Nayar and Sastry 1987). Thus, there is a need to conserve genetic diversity of this prized medicinal plant, which may become extinct if reckless exploitation continues. Estimation of the level and distribution of genetic variation in endangered species is a primary objective in implementation of conservation programmes (Fritsch and Rieseberg 1996). Therefore it is necessary to evaluate the genetic variation from different regions for identification of elite germplasm with high genetic variability that can be used in conservation strategies.

Among the various molecular-marker technologies, microsatellites or simple sequence repeat markers (SSRs) are markers of choice because of multiple desirable characteristics. SSRs are arrays of short repetitive motifs (2–6 bp) that are distributed throughout the genome and have been utilized as a source of highly polymorphic and reproducible codominant markers. Generation of polymorphism at these sites is believed to be largely due to slippage of the template during replication, and this process results in an increase or decrease in the repeat number (Ellegren 2004). The high frequency at which mutations occur at these sites produces high degree of polymorphism, which is useful for population genetic analysis. Owing to these properties, microsatellite markers are widely used to make inferences about population structure and gene flow. They have

*For correspondence. E-mail: mrk_sharma@yahoo.com; ramsharma@ihbt.res.in.

Keywords. anticancer property; endangered plant; Himalayan plant; microsatellite markers; *Podophyllum hexandrum*.

also been used as disease markers, and in breeding programmes. Rapid increases in sequence information under various genome and EST projects have enriched the publicly available databases (http://ncbi.nlm.nih.gov). Non-redundant nucleotide sequences derived from these public databases have become a cost-effective source of microsatellite markers in several crop plants and rare species, in which marker identification was earlier based on tedious, labour-intensive nucleotide sequencing of clones from enriched genomic libraries.

Genetic diversity studies in *P. hexandrum* have largely been carried out using dominant molecular and phytochemical markers because microsatellite markers were not available (Nadeem *et al*. 2000; Sultan *et al*. 2008; Alam *et al*. 2008, 2009). Because of their limited resolution and dominant inheritance, use of dominant molecular and phytochemical markers can lead to an underestimation of recessive allele frequency in a population and hence a bias in estimates of genetic diversity and genetic differentiation (Nybom 2004). Therefore there is urgent need to identify highly polymorphic codominant microsatellite markers in *P. hexandrum*. The set of 20 novel *Podophyllum hexandrum* UniGene-derived microsatellite (PHUGMS) markers identified in the current study would enable future investigations on spatial genetic structure and population diversity in *P. hexandrum*.

Materials and methods

A total of 1084 FASTA formatted EST sequences in *P. hexandrum* were retrieved from the US National Center for Biotechnology Information (NCBI; http://www.ncbi.nlm.nih.gov/entrez) for subsequent data mining. A non-redundant (NR) expressed sequence data set of 26.94 kb was created by clustering random ESTs into 655 unigenes (195 contigs and 460 singletons) using SeqMan DNAStar Lasergene v7.1 (Dnastar, Madison, USA) using the search parameters reported earlier by Sharma *et al*. (2009). All the UniGenes were subsequently searched individually for presence of SSRs using repeatmasker (http://www.repeatmasker.org/cgi-bin/WEBRepeatMasker). Primers were designed using Primer3 software (http://www.genome.wi.mit.edu/genome_software/other/primer3.html). The major parameters for designing the primers were: primer length 18–27 (optimum 20 bp), PCR product size 125–300 bp, optimum annealing temperature 60°C, and GC content 40–80%. Amplification-based validation of the 47 primers was carried out in a test array of 15 accessions of *P. hexandrum*. The test array used in the present study comprised five individuals each of three different *P. hexandrum* populations namely Prashar (Mandi district), Khoksar (Lahul and Spiti district) and Bairagarh (Chamba district) of Himachal Pradesh, India, and are maintained at IHBT. A corresponding plant specimen of every individual has been preserved in herbarium according to the international code (PLP) at IHBT (table 1).

Table 1. Names of the locations of individuals used in the study and their herbarium voucher numbers. Five individuals from each location were preserved in the IHBT herbarium with the respective voucher number.

Location	Voucher no.
PRASHAR	PLP16512
KHOKSAR	PLP16514
BAIRAGARH	PLP16513

Extraction of total genomic DNA was done by CTAB method (Doyle and Doyle 1990), and polymerase chain reactions were performed as earlier reported by Sharma *et al*. (2009), with annealing temperature (T_a) for each PHUGMS primer as given in table 2. Amplification products were resolved on a 6% denaturing polyacrylamide gel (19:1 acraylamide:bisacrylamide) in 1× TBE buffer, visualized by silver staining (silver sequence staining reagents, Promega, Madison, USA), and sized using a 50-bp DNA ladder (MBI Fermentas, Vilnius, Lithuania). Analysis of molecular variance (AMOVA), was calculated using GenAlEx 6.4 (Peakall and Smouse 2006).

Results and discussion

The clustering of ESTs with DNAstar resulted in 41 unigenes, containing 47 SSRs. Nonredundant data represented 33 trirepeats, 9 tetrarepeats, 4 pentarepeats and 1 direpeat. Among the trirepeats, $(TTC)_n$, and $(GAT)_n$ were most abundant, followed by $(GGA)_n$, $(GAA)_n$ and $(CTT)_n$. Amplification-based validation of the 47 PHUGMS primer pairs amplified the expected amplicons in the target *P. hexandrum* DNA with 36 primer pairs. Of these, 20 PHUGMS markers were polymorphic among the tested populations.

PHUGMS markers identified in the current study were moderately to highly polymorphic and a total of 91 alleles (table 2). The number of alleles detected ranged from 2 to 9, with an average 4.55 alleles per locus. Expected (H_E) and observed heterozygosity (H_O) obtained using the Popgene software package (Yeh *et al*. 1997) ranged from 0.239 to 0.869 (av. 0.687) and 0.067 to 1.000 (av. 0.703), respectively. Analysis of molecular variance (AMOVA) using GenAlEx (Peakall and Smouse 2006) software revealed that slightly more variance was observed between populations (57%) than among populations (43%), which means that these markers are capable of distinguishing the variation between different populations at the molecular level and hence could be useful in resolving the genetic variation of Indian *Podophyllum* populations. Populations with high diversity that are better adapted to the environment need to be identified, and these may serve as germplasm of choice for conservation programmes.

In conclusion, the novel PHUGMS markers presented in this study show sufficient levels of polymorphism to be used

Table 2. Details of 20 microsatellite primer pairs identified in public EST database of *Podophyllum hexandrum*.

	Primer	Sequence	Repeat type	Size range (bp)	T_a (°C)	No. of alleles	Heterozygosity		Contributing EST's GenBank accession number*
							H_E	H_O	
1	PHUGMS03	F: 5′-AAACGATGTCTCTGCCGACT-3′ R: 5′CCAACAACAGACCGGATGAT-3′	$(GAAG)_3$	270–320	51	5	0.770	0.800	FK934424
2	PHUGMS06	F: 5′-CCTTGCTTGGCCATTAAAAA-3′ R: 5′-ATGAAAGGGGAGGGGTATTG-3′	$(AAG)_4$	260–290	50	6	0.818	0.733	GO254304
3	PHUGMS08	F: 5′-CCTCAGCACCAGACCTTTTC-3′ R: 5′-CCTCTCCTTCCAAGTCACCA-3′	$(ACAA)_3$	240–290	55	5	0.639	0.933	GR972406, R972402, GR972404, R972403, GR972401
4	PHUGMS10	F: 5′-AGGTGTCAAGCCAGAAGGAA-3′ R: 5′-ATATTCTACCCGGCCGTAGG-3′	$(TAT)_4$	230–270	53	2	0.239	0.267	FK934421
5	PHUGMS12	F: 5′-TCATCATCCTCCACTCTCCA-3′ R: 5′-TCGTCCATTCGATTGTGGTA-3′	$(TTC)_6$	150–180	50	2	0.460	0.067	GT152526
6	PHUGMS14	F: 5′-CAATGGCAGGCTATGGTTCT-3′ R: 5′-CCTATTCCGTCTCCTGGTCA-3′	$(TGC)_4$	210–240	53	3	0.618	0.733	FK934400, FF279338
7	PHUGMS21	F: 5′-GGAGGACGAAATCAACCAGA-3′ R: 5′-TTATCGGACCAACACAACCA-3′	$(TC)_{12}$	150–190	50	5	0.789	1.000	GW413540
8	PHUGMS23	F: 5′-CAACACACGACCACCATCTC-3′ R: 5′-GGAACCCTCAAAGGTGGACT-3′	$(TCT)_4$	210–250	55	4	0.697	0.333	GW413546
9	PHUGMS25	F: 5′-TCAATGTCCCACACAAGCAT-3′ R: 5′-TGACTGGTGACTGGTGAAGC-3′	$(GAA)_4$	224–260	51	3	0.687	0.933	FK934350, FF279329
10	PHUGMS31	F: 5′-AGGTACAGGGACACGCAAAG-3′ R: 5′-ATGAGCTCAAATGCCTCCTG-3′	$(CTT)_4$	220–270	53	5	0.614	0.133	GO254281, GO246447
11	PHUGMS34	F: 5′-CCGGGCAGGTAAATCAATAA-3′ R: 5′-AGCGAAAAGCTCTAGCCAAA-3′	$(GGA)_5$	210–265	50	4	0.644	0.607	FF279492, GT152380, GT152648, GW395890, GT152655
12	PHUGMS35	F: 5′-GAGGAGGAGGTAGAGGAAGTAGAA-3′ R: 5′-TTAGCTGCAGAGCGAAAAGC-3′	$(GAT)_4$	165–190	53	4	0.690	0.933	FF279492, GT152380, GT152648, GW395890, GT152655
13	PHUGMS36	F: 5′-TGATGAAGAGGAAGAAAAATGGA-3′ R: 5′-GGCCGAGGTACTTGTTCTCT-3′	$(GAT)_4$	320–390	53	4	0.708	0.867	FF279492, GT152380, GT152648, GW395890, GT152655
14	PHUGMS37	F: 5′-ATGGCAAATCTTGGGTTCAG -3′ R: 5′-GAGGTCCAGTGTTGCTGTCA-3′	$(GAA)_4$	230–260	53	3	0.690	0.667	GO248746
15	PHUGMS38	F: 5′-CCGGGCAGGTAAATCAATAA-3′ R: 5′-AGCGAAAAGCTCTAGCCAAA-3′	$(GGA)_7$	150–190	50	9	0.869	0.800	FF279492, GT152380, GT152648, GW395890, GT152655
16	PHUGMS40	F: 5′-CAGGTGCCATCCCAAGAC-3′ R: 5′-AGCCACCGCGCTACAATA-3′	$(GGA)_5$	190–275	50	6	0.816	0.667	FL640976, FL640953, FF279464
17	PHUGMS43	F: 5′-CCATTTAGGGTAAgGGGGTTT-3′ R: 5′-CCCTTGTCATGTGGAGGAGT-3′	$(TGG)_6$	160–180	53	4	0.710	0.867	GT152424, GT152602
18	PHUGMS44	F: 5′-AAGCTCCAGTTCAACCTCAA-3′ R: 5′-AACACCACCCAGTTCCTTGT-3′	$(CAGTT)_4$	237–270	50	6	0.772	1.000	GW420706
19	PHUGMS45	F: 5′-CACCCATGGATTCTTCTCC-3′ R: 5′-CATCAACGGCTGTGAAGAGA-3′	$(TTC)_6$	200–230	50	5	0.793	0.773	FF279417
20	PHUGMS47	F: 5′-TCCATCAGATCCAACTTCAA-3′ R: 5′-ACTCCTCTGTATAGCTTTGT-3′	$(CAA)_7$	200–195	50	6	0.708	0.933	GO248648

*GenBank accession numbers are of the original fragment.
PHUGMS, *Podophyllum hexandrum* UniGene microsatellite; T_a, annealing temperature, H_E, expected heterozygosity; H_O, observed heterozygosity.

for detailed population genetic studies and for evaluating genetic diversity, and therefore would be useful in the conservation and management of this medicinally important, severely endangered *P. hexandrum*.

Acknowledgement

Financial assistance received from Council of Scientific and Industrial Research, New Delhi, India, is gratefully acknowledged. This is IHBT Publication no. 3436.

References

Alam A., Gulati P., Gulati A. K., Mishra G. P. and Naik P. K. 2009 Assessment of genetic diversity among *Podophyllum hexandrum* genotypes of the North-western Himalayan region for Podophyllotoxin production. *Indian J. Bot.* **8**, 391–399.

Alam A., Naik P. K., Gulati P., Gulati A. K. and Mishra G. P. 2008 Characterization of genetic structure of *Podophyllum hexandrum* populations, an endangered medicinal herb of Northwestern Himalaya, using ISSR-PCR markers and its relatedness with Podophyllotoxin content. *Afr. J. Biotechnol.* **7**, 1028–1040.

Bhadula S. K., Singh A., Lata H., Kuniyal C. P. and Purohit A. N. 1996 Genetic resources of *Podophyllum hexandrum* Royle, an endangered medicinal species from Garhwal Himalaya, India. *Int. Plant Genet. Resour. Newsl.* **106**, 26–29.

Broomhead A. J. and Dewick P. M. 1990 Tumor inhibiytory aryitralin lignans in *Podophyllum versipelle*, *Diphylleia cymosa* and *Diphyllea grayi*. *Phytochemistry* **29**, 3831–3837.

Doyle J. J. and Doyle J. L. 1990 Isolation of plant DNA from fresh tissue. *Focus* **12**, 13–15.

Ellegren H. 2004 Microsatellites: simple sequences with complex evolution. *Nat. Rev. Genet.* **5**, 435–445.

Fritsch P. and Rieseberg L. H. 1996 The use of random amplified polymorphic DNA (RAPD) in conservation genetics. In *Molecular genetic approaches in conservation* (ed. T. B. Smith and R. K. Wayne), pp. 54–73. Oxford University Press, New York, USA.

Giri A. and Narasu M. L. 2000 Production of podophyllotoxin from *Podophyllum hexandrum*: a potential natural product for clinically useful anticancer drugs. *Cytotechnology* **34**, 17–26.

Issel B. F., Muggia F. M. and Carter S. K. 1984. *Etoposide (VP-16) - current status and new developments*. Academic Press, Orlando, USA.

Jackson D. E. and Dewick P. M. 1984 Biosynthesis of *Podophyllum* Lignanas – II. Interconversion of aryltralin lignan in *Podophyllum hexandrum*. *Phytochemistry* **23**, 1037–1042.

Nadeem M., Palini L. M. S., Purohit A. N., Pandey H. and Nandi S. K. 2000 Propagation and conservation of *Podophyllum hexandrum* Royle: an important medicinal herb. *Biol. Conserv.* **92**, 121–129.

Nag A. and Rajkumar S. 2011 Chromosome identification and karyotype analysis of *Podophyllum hexandrum* Roxb. ex Kunth using FISH. *Physiol. Mol. Biol. Plants* **17**, 313–316.

Nayar M. P. and Sastry A. R. K. 1987 *Red data book of Indian plants*, vol. 1. Botanical Survey of India, Calcutta, India.

Nybom H. 2004 Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. *Mol. Ecol.* **13**, 1143–1155.

Peakall R. and Smouse P. E. 2006 GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Mol. Ecol. Notes* **6**, 288–295.

Sharma R. K., Bhardwaj P., Negi R., Mohapatra T. and Ahuja P. S. 2009 Identification, characterization and utilization of unigene derived microsatellite markers in tea (Camellia sinensis L.). *BMC Plant Biol.* **9**, 53.

Sultan P., Shawl A. S., Ramteke P. W., Kour A. and Qazi P. H. 2008 Assessment of diversity in *Podophyllum hexandrum* by genetic and phytochemical markers. *Sci. Horti.* **115**, 398–408.

Tyler V. E., Brady L. R. and Roberts J. E. 1988 *Pharmacology*, 9th edition. Lea and Febiger, Philadelphia.

Yeh F. C., Yang R. C., Boyle T. B. J., Ye Z. H. and Mao J. X. 1997 POPGENE: the user-friendly shareware for population genetic analysis. Molecular Biology and Biotechnology Centre, University of Alberta, Canada. (http://www.ualberta.ca/~fyeh).

Received 12 September 2012; accepted 8 October 2012
Published on the Web: 14 February 2013