

PRIMER NOTE

Isolation of microsatellite markers in *Cannabis sativa* L. (marijuana)

S. GILMORE*† and R. PEAKALL†

*Centre for Forensic Science, Canberra Institute of Technology, PO Box 826, Canberra ACT 2601, Australia, †School of Botany and Zoology, Australian National University, Canberra ACT 0200, Australia

Abstract

We have identified 15 variable microsatellite loci in *Cannabis sativa*. In 48 samples from five fibre crop seed accessions, we detected an average of 10 alleles per locus (range 2–28) with mean heterozygosity of 0.68 (range 0.28–0.94). Significant genetic differentiation was found between accessions ($F_{ST} = 0.12$, $P < 0.001$). These markers have utility for characterizing genetic diversity in cultivated and naturalized *Cannabis* populations.

Keywords: *Cannabis*, forensic, germplasm, hemp, microsatellites

Received 27 September 2002; revision received 14 November 2002

Cannabis sativa L. has been distributed worldwide as a source of fibre, feed, oils, medicines and intoxicants (Small & Cronquist 1976). Despite cultivation being restricted in many countries due to its psychoactive properties, there has been a resurgence of interest in *Cannabis* for its agronomic potential, particularly as a source of fibre. Illegal cultivation occurs for the drug marijuana and its derivatives, while naturalized or feral populations are widespread. Here, we report 15 microsatellite loci with primer sequences and evaluate their usefulness in population studies of *C. sativa*.

Seeds held by the Australian Federal Police were germinated and DNA was extracted by the cetyltrimethyl ammonium bromide-DNA extraction protocol (Doyle & Doyle 1990). An initial nonenriched microsatellite library was constructed, probed and positive colonies sequenced using the method of Scott *et al.* (2001). Sequencing products were run on the ABI PRISM 377 DNA sequencer and the DNA sequence edited in SEQUENCHER 3.0 (GeneCodes). Primers were designed using PRIMER 0.5 (Whitehead Institute for Biomedical Science). DNA fragments were labelled during polymerase chain reaction (PCR) with fluorescent dUTP nucleotides (Perkin-Elmer) before electrophoresis on an ABI PRISM 377 sequencer following the method of Peakall *et al.* (1998).

A second enriched library, constructed following the method of Fischer & Bachmann (1998), was obtained

commercially from the Centre for Identification and Diagnostics (Queensland University, Australia). One hundred recombinant clones were sequenced using the DYEnamic ET Terminator Cycle Sequencing kit (Amersham Pharmacia Biotech). Primer design was performed with PRIMER3 (Rozen & Skaletsky 1997) followed by the addition of the –21M13 (5'-TGTAACGACGGCCAGT) sequence at the 5'-tail for genotyping on an ABI PRISM Sequencer following the method of Schuelke (2000). This approach considerably reduces genotyping costs for small projects as a generic 'reporter' primer can be used for all loci, rather than a large number of primers requiring separate labelling.

Fifteen primer pairs were used to screen for polymorphism in *Cannabis*. When using fluorescent dUTPs, the 10 µL PCR contained 10 ng DNA, 10 mM Tris-HCl, pH 8.3, 50 mM KCl, 1 mM MgCl₂, 250 µM dNTPs, 0.2 µM each of forward and reverse primer, 0.04 µL fluorescently labelled dUTP (either 400 µM R110 or 100 µM TAMRA; ABI) and 0.75 U TaqGold DNA Polymerase (Perkin-Elmer). The reaction profile comprised a 10-min incubation at 90 °C to activate the TaqGold and then a cycle of 95 °C for 15 s, 52 °C for 15 s and 72 °C for 60 s, repeated 30 times. Following cycling, the reaction was held at 72 °C for 5 min, before a final 4 °C holding temperature. When fluorescently labelled –21M13 reporter primers were used in conjunction with –21M13 tailed amplification primers, the 10 µL reaction mix comprised 10 ng template DNA, 1× PCR buffer (QIAGEN; final 1.5 mM MgCl₂, but concentrations of Tris-HCl, KCl and (NH₄)₂SO₄ are proprietary), 0.2 µM dNTPs, 200 nM each of forward and labelled –21M13

Correspondence: S. Gilmore. Fax: +61-2-62074105; E-mail: simon.gilmore@cit.act.edu.au

Table 1 Primer sequences and characteristics for 15 *Cannabis sativa* microsatellite loci

Locus	Repeat motif	Primer sequence (5'–3')	T_a	Size of cloned allele (bp)	Allele size range (bp)	N_a	H_o	H_e	GenBank accession no.
ANUCS201	(GA) ₂₆	GGTTCAATGGAGATTCTCGT CCACTAAACCAAAAGTACTCTTC	52	187	161–223	18	0.51	0.86	AY167019
ANUCS202	(GA) ₂₀	AGGACCAATTTTGAATATGC AGAGAGGGAAGGGCTAACTA	52	161	147–185	14	0.63	0.86	AY167016
ANUCS203	(CT) ₅₀	GCTCTTCTTATTAATTCCTCCTT *GAATATGATAAGACACAATTCATT	54	195	169–267	28	0.67	0.94	AY167008
ANUCS204	(CT) ₂₆	TGGAAGATATGCAACTGGAG *AACGAAGATAAGCAGCAACA	54	150	128–184	14	0.14	0.87	AY167011
ANUCS205	(CT) ₂₁	TTGACTAACCGGCAAGATA *AAATTCAAACCGATTCTCAG	54	168	172–242	18	0.59	0.87	AY167010
ANUCS206	(AT) ₁₁	TCCAATAAGATTTTACAGTCG *TAACGGGTCTTTTGGGTATTT	54	167	159–167	4	0.38	0.58	AY167014
ANUCS301	(TTA) ₁₅	ATATGGTTGAAATCCATTGC TAACAAAGTTTCTGTGAGGGT	52	234	209–261	13	0.60	0.85	AY167020
ANUCS302	(CAA) ₇ –(CAA) ₄ ‡	AACATAAACCAACAACATGC ATGGTTGATGTTTTGATGGT	52	147	140–173	10	0.56	0.68	AY167021
ANUCS303	(GTG) ₇	TAATCAACAATGACAATGGC GATTAAGGTCCTCGACGATA	52	147	141–156	5	0.50	0.56	AY167022
ANUCS304	(TCT) ₈ TCA(TCT) ₇	TCTTCACTCACCTCCTCTCT *TCTTTAAGCGGGACTCGT	54	173	167–230	15	0.50	0.69	AY167018
ANUCS305	(TGG) ₁₀	AAAGTTGGTCTGAGAAGCAAT *CCTAGGAACCTTTCGACAACA	54	162	141–162	7	0.67	0.71	AY167009
ANUCS306	(GAT) ₃ –(GAT) ₆ §	ACTATTACTAAGCCTCCTCATCA *GTGGTAGTCTCATTGTTGGTG	54	95	92–95	2	0.67	0.49	AY167012
ANUCS307	(ACC) ₆	GGCTGAGTAGTCTAAGCTTCC *GGACTAGCCACCATCAGG	54	108	105–108	2	0.24	0.28	AY167015
ANUCS308	(TAA) ₃ –(AT) ₅ ¶	AGATGGTGTGGGTATCTTT *TGGTGCAGGTTTATACAATTT	54	186	177–203	8	0.18	0.39	AY167017
ANUCS501	(TTGTG) ₄	AGCAATAATGGAGTGAGTGAAC *AGAGATCAAGAAATTGAGATTCC	54	90	80–95	3	0.49	0.51	AY167013

*Primers constructed with a 5'–21M13 tail for genotyping.

‡§¶Six, 15 and 13 bp separate these microsatellite arrays, respectively.

T_a , Annealing temperature; N_a , number of alleles; H_o , observed heterozygosity; H_e , expected heterozygosity.

Number of alleles and heterozygosities were calculated for 48 individuals.

(either labelled with FAM, NED or VIC; Invitrogen), 50 nM reverse primer and 0.75 U Taq DNA polymerase (QIAGEN). The reaction profile comprised a predenaturation at 94 °C for 3 min then 25 cycles of 94 °C for 30 s, 54 °C for 30 s and 72 °C for 30 s for amplification with the locus-specific primers, followed by eight cycles of 94 °C for 30 s, 52 °C for 30 s and 72 °C for 30 s for amplification with the reporter primer. All amplifications were conducted in a Palm-Cycler CP002 (Corbett Research) and PCR products were run on an ABI PRISM 377 DNA sequencer. Fragment sizes were scored using GENESCAN™ and GENOTYPER™ software (Applied Biosystems). Population data were analysed using GENALEX Version 5 (Peakall & Smouse 2001).

The 15 primer pairs give reliable amplification across geographically separated populations of *C. sativa* (Table 1).

Analyses were performed on a set of 48 samples, representing five different fibre crop accessions. We detected an average of 10.7 ± 1.9 (mean \pm SE, range 2–28) alleles per locus with mean expected heterozygosity of 0.68 ± 0.05 (mean \pm SE, range 0.28–0.94) (Table 1). Significant genetic differentiation, determined by an analysis of molecular variance, was found between accessions ($F_{ST} = 0.12$, $P < 0.001$). Hardy–Weinberg tests were not conducted due to small within-population sample sizes.

The first 15 microsatellite markers developed for *C. sativa* are hypervariable and informative, allowing discrimination among individuals and accessions. Our results confirm the utility of these markers for characterizing genetic diversity in cultivated and naturalized *Cannabis* populations and will provide tools for germplasm management, plant breeding and forensic analyses.

Acknowledgements

We are grateful to Mojca Keglovic, Lauretta Grasso and Christine Hayes for laboratory assistance. The Australian Research Council, AFP, CIT and EcoFibre Industries provided funding through a Strategic Partnership with Industry Research and Training Scheme grant.

References

- Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. *Focus*, **12**, 13–15.
- Fischer D, Bachmann K (1998) Microsatellite enrichment in organisms with large genomes (*Allium cepa* L.). *Biotechniques*, **24**, 796–802.
- Peakall R, Gilmore S, Keys W, Morgante M, Rafalski A (1998) Cross-species amplification of Soybean (*Glycine max*) simple sequence repeats (SSRs) within the genus and other legume genera: implications for the transferability of SSRs in plants. *Molecular Biology and Evolution*, **15**, 1275–1287.
- Peakall R, Smouse P (2001) *GENALEX V5: Genetic Analysis in Excel*. Australian National University, Canberra, Australia. <http://www.anu.edu.au/BoZo/GenALEX/>.
- Rozen S, Skaletsky HJ (1997) PRIMER 3 code available at http://www-genome.wi.mit.edu/genome_software/other/primer3.html.
- Schuelke M (2000) An economic method for the fluorescent labeling of PCR fragments. *Nature Biotechnology*, **18**, 233–234.
- Scott IA, Hayes CM, Keogh JS, Morrison SF (2001) Isolation and characterisation of novel microsatellite markers from the Australian water skink *Eulamprus kosciuskoi* and cross-species amplification in other members of the species-group. *Molecular Ecology Notes*, **1**, 28–30.
- Small E, Cronquist A (1976) A practical and natural taxonomy for *Cannabis*. *Taxon*, **25**, 405–435.