ANALYSIS OF MICROSATELLITE MARKERS IN HEMP (*CANNABIS SATIVA* L.)

PRESINSZKA MARIA[1], STIASNA KLARA[1], VYHNANEK TOMAS[1], TROJAN VACLAV[1], MRKVICOVA EVA[2], HRIVNA LUDEK[3], HAVEL LADISLAV[1]

[1]Department of Plant Biology
[2]Department of Animal Nutrition and Forage Production
[3]Department of Food Technology
Mendel University in Brno
Zemedelska 1, 613 00 Brno
CZECH REPUBLIC

maria.presinszka@mendelu.cz

Abstract: Cannabis sativa L. (hemp) is one of the oldest cultivated plants used around the world for diverse applications. 22 genotypes of hemp were analyzed with 16 SSR markers (8 SSR markers "*ANUC*" and 8 SSR markers "*CAN*"). Used primers amplified 76 different polymorphic alleles with an average number of 4.75 alleles per locus. The number of alleles ranged from 1 (*CAN1690B*) to 7 (*ANUC204* and *CAN0110*). The diversity index (DI), the polymorphic information content (PIC) and the probability of identity (PI) were calculated. Values of diversity index ranged from 0 to 0.926 with an average 0.703, probability of identity from 0.004 to 1 with an average 0.141 and polymorphic information content from 0.926 to 0 with an average 0.688.Six SSR markers which reached values of DI and PIC higher than 0.8, can be used for studies of genetic variability. Dendrogram of similarity was constructed showing that genotype *Cannabis indica* 'Royal Caramel' is the most distant in our set of varieties. The industrial hemp varieties were separated from other genotypes. Results showed usefulness of microsatellite markers for detection of genetic diversity in *Cannabis.*

Key Words:Cannabis L., SSR markers, PCR, variability

INTRODUCTION

The genus *Cannabis* includes three different species: *Cannabis sativa*, *Cannabis indica* and *Cannabis ruderalis*, however botanists now generally agree that there is only a single highly variable species *C. sativa* (Hillig,Mahlbegr 2004). Hemp plants produce many different secondary metabolites such as cannabinoids, flavonoids, stilbenoids, alkaloids, lignanamides, and phenolic amides (Marks et al. 2009). The main psychoactive substance is Δ-9- tetrahydrocannnabinol (THC), and besides THC, another substance cannabidiol (CBD) is produced in high concentrations (Thichak et al. 2011). Two characteristic strains are distinguished: one is generally cultivated for fiber (hemp) and the other for drug use (marijuana) (Alghanim, Almirall 2003).

Genetic variability can be detected by different molecular markers such as restriction fragment length polymorphism (RFLP), random amplified polymorphic DNA (RAPD), amplified fragment length polymorphism (AFLP), or simple sequence repeat (SSR) (Varshney et al. 2005, Sarwat et al. 2012). SSR are the most useful markers, otherwise known as microsatellites, or short tandem repeats (STR) (Alghanim, Almirall 2003). Microsatellites are short repeats of DNA sequences with one to six nucleotides as the repeating unit. Sequence repeats show high levels of polymorphism between individuals (Zhang 2004). Because of these advantages, microsatellites have become well suited for a wide range of applications in genetic mapping, fingerprint and genotype identification, seed purity evaluation and germplasm conservation, genetic relatedness and paternity studies and markerassisted selection (Alghanim,Almirall 2003).

The aim was to detect genetic variability of *Cannabis* samples with a focus on the possibility to identify varieties of industrial hemp in the analyzed set for verification of the commodity during processing.

MATERIAL AND METHODS

Characterization of plant material

The used twenty-two genotypes of hemp (*Cannabis* L.) included six varieties of industrial hemp ("TE" - 'Finola', 'Tiborszálási', 'Tisza', 'Kompolti', 'Kompolti hybrid TC' and 'Carmagnola') from Czech Republic (Hempoint, Ltd. – Hana Gabrielová), four genotypes of *C. sativa* ("SA" – 'Arjan´s Haze', 'Amnesia Haze', 'Sour Diesel' and 'Presidential O.G.'), four genotypes of *C. indica* ("IN" – 'Great White Shark', 'Northern Light', 'Skunk' and 'Royal Caramel'), four hemp mixtures ("SM" – 'Special Queen', 'Opium', 'Mohan Ram' and 'Biddy Early'), and four genotypes of hemp hybrids ("HY" – 'Super Bud', 'Royal Medic', 'Big Bud XXL' and 'Jack Herrer Automatic') from Netherland (source of DNA – Dr. Arno Hazekamp) were used.

Experimental design for study of genetic variation

Genomic DNA of industrial hemp varieties was isolated from leaves using the isolation kit DNeasy Plant Mini Kit (Qiagen, GE). The DNA concentration was evaluated spectrophotometrically. 16 SSR markers: 8 SSR markers ("*ANUC*") (Gilmore, Peakall 2003) and 8 SSR markers ("*CAN*") (Gao et al. 2014) were used (Table 1). The reaction mixture for PCR of a total volume 25 µl contained 0.5 U *Taq* polymerase (Promega, USA), 1× aliquot buffer, 0.1mM of each dNTP (Promega, USA), 0.3 M of each primer and 30 ng of template DNA; the reaction conditions for PCR in T3 cycler (Biometra, Germany) according to Gilmore,Peakall (2003) for *ANUC* markers and Gao et al. (2014) for *CAN* markers were used. Useful step seems to be a control electrophoresis on 1.5% agarose gel (stained with ethidium bromide) with a fraction of the sample after amplification, which makes it possible to select usable samples for the separation on polyacrylamide gels. The amplification of SSR products was then visualized on 8% non-denaturating polyacrylamid (PAA) gels in TBE buffer (300 V) followed by staining with silver (0.2% $AgNO_3$). The resulting electrophoretograms were converted to binary matrices where presence (1) or absence (0) of the alleles were recorded. Alleles were evaluated by means of the statistical software FreeTree version 9.1 (Hampl et al. 2001) using the UPGMA (Unweight Pair Group Method with Arithemetic Mean) construction method and similarity coefficient according to Jaccard (1908). The software TreeView version 1.6 (Page, 1996) was used for the graphic visualization of the matrix. Three statistical parameters (diversity index - DI, probabilities of identity - PI and polymorphic information contents - PIC) were calculated according to Russellet al. (1997).

Table 1 Used SSR markers for Cannabis L.

Name	Repeat motif	Forward primer (5' - 3')	Reverse primer (5' - 3')
ANUC201	$(GA)_{26}$	GGTTCAATGGAGATTCTCGT	CCACTAAACCAAAAGTACTCTTC
ANUC202	$(GA)_{20}$	AGGACCAATTTTGAATATGC	AGAGAGGGAAGGGCTAACTA
ANUC203	$(CT)_{50}$	GCTCTTCTTATTAATTCCTCCTT	GAATATGATAAGACACAACTTCATT
ANUC204	$(CT)_{26}$	TGGAAGATATGCAACTGGAG	AACGAAGATAAGCACGAACA
ANUC205	$(CT)_{21}$	TTGACTAACCGGCAAAGATA	AAATTCAAAACCGATTCTCAG
ANUC301	$(TTA)_{15}$	ATATGGTTGAAATCCATTGC	TAACAAAGTTTCGTGAGGGT
ANUC302	$(CAA)_7$-$(CAA)_4$	AACATAAACACCAACAACTGC	ATGGTTGATGTTTTGATGGT
ANUC304	$(TCT)_8TCA(TCT)_7$	TCTTCACTCACCTCCTCTCT	TCTTTAAGCGGGACTCGT
CAN0039	$(CAT)_8$	GCAGCCATAGTCATGGTGTA	GTCATTGGAAAGACCAGCTT
CAN0093	$(GA)_{11}$	CAGTCTCTCAGATCAGACTACC	AGCGGCTAGCGTAACAGTAT
CAN0110	$(AT)_{10}$	GGGTAAAGCTTACGCAAAGT	AACAAACAGTTGGACACCCT
CAN0126	$(AATACC)_3(CAG)_6$*	GAGTAAGAGAAGGCGAACCA	CCTGTGTAACAGAAAACCCC
CAN0585	$(ACTTCTATT)_2T(CAAAAC)_3$	TCATCATCATCCCTCCCTAT	GGTCCATAGTTGGCTGATCT
CAN1347	$(CAA)_6(CATCATAAT)_2$	CAAACAGGGGAAAAGAGAGA	ATGAAGCGTTGGTACTAGGC
CAN1690B	$(AAC)_6(ATC)_7$	TGTTTCTAAGGCTCAGTCCC	GGCAAAGGTAAAGCAAGTGT
CAN2913	$(AAG)_7$	AGGAACACTTTGAAAGCGAG	CGGTCATCTACCTTGAGCTT

RESULTS AND DISCUSSION

Analysis of the microsatellite loci

A total of 16 SSR markers were used for analysis of a set of 22 *Cannabis* samples. In plant DNA one microsatellite locus is present on average every 33 kb (Alghanim,Almirall 2003). The differences in plants could be caused by variations in structure of the genome of various types (Cordeiro et al. 2000). In table 2 the size of alleles that ranged from 100–220 bp and number of alleles that fluctuated from 1 (*CAN1690B*) to 7 (*ANUC204* and *CAN110*) is presented. The size of alleles variedas reported also by Gilmore, Peakall (2003) and Gao et al. (2014). 76 different SSR marker alleles were found with an average of 4.75 alleles per locus. Analyzing another set of cannabis samples Alghanim, Almirall (2003) obtained similar results with an average of 4.7 alleles per locus.

Values of DI and PIC higher than 0.8 and value of PI lower than 0.06 was calculated for six SSR markers (Table 2). These markers are suitable for the identification of genotypes of industrial hemp. The best SSR marker with the highest DI and PIC (0.926) and lowest PI (0.004) was *CAN2913*. Conversely the SSR marker *CAN1690B* has a value of DI and PIC = 0 and PI = 1, which means it is not suitable for our study. Kayis et al. (2010) analyzed 22 markers of *Cannabis* from Turkey and determined in average values of PIC = 0.280 which is three times lower than our results. The lower value was caused by analyzing only samples of *Cannabis sativa*.

Table 2 Characteristics of analyzed microsatellite markers

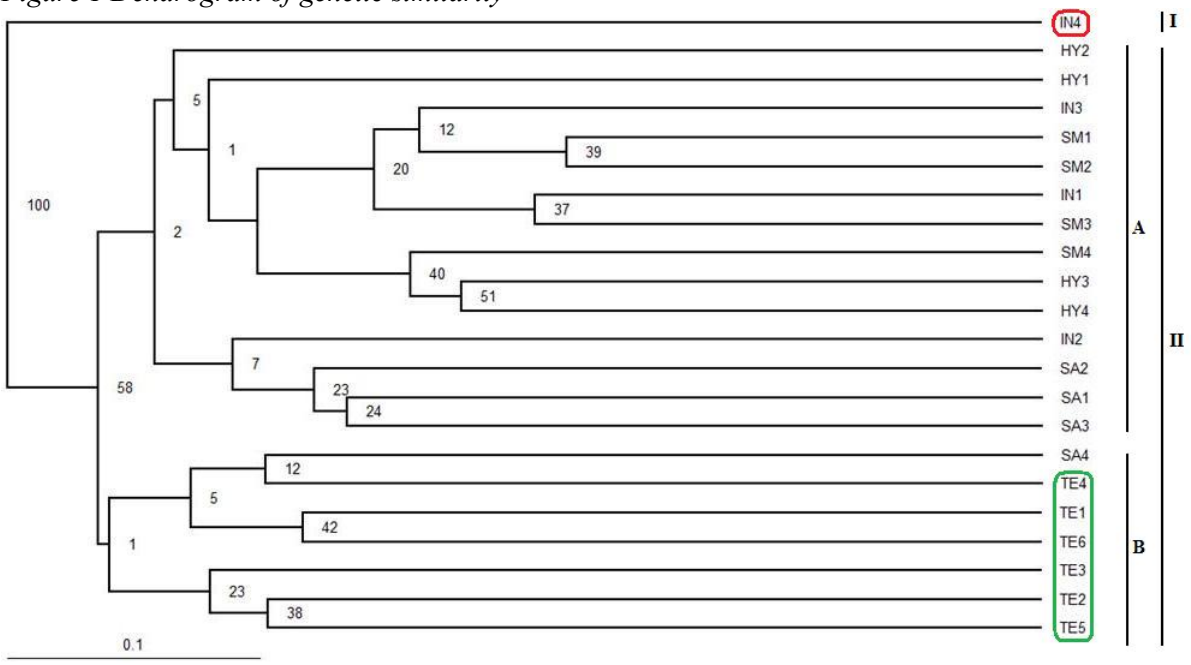
SSR marker	Size ofalleles (bp)	Number of alleles	DI	PI	PIC
ANUC201	140–200	6	0.690	0.090	0.680
ANUC202	140–200	5	0.730	0.090	0.720
ANUC203	120–200	5	0.801	0.052	0.784
ANUC204	100–180	7	0.780	0.040	0.770
ANUC205	120–180	5	0.839	0.027	0.830
ANUC301	200–260	4	0.576	0.200	0.539
ANUC302	140–160	6	0.770	0.040	0.760
ANUC304	140–200	4	0.740	0.070	0.720
CAN0039	200–240	4	0.856	0.020	0.852
CAN0093	200–230	5	0.874	0.012	0.874
CAN0110	100–130	7	0.907	0.007	0.906
CAN0126	160–180	4	0.662	0.147	0.620
CAN0585	200–220	4	0.611	0.216	0.542
CAN1347	200–220	4	0.491	0.248	0.490
CAN1690B	220	1	0.000	1.000	0.000
CAN2913	100–120	5	0.926	0.004	0.926
Average		4.750	0.703	0.141	0.688

Legend:bp – base pair, DI – diversity index, PI – probabilities of identity, PIC – polymorphic information contents

Genetic similarity of 22 genotypes of *Cannabis*

Similarity dendrograms and cluster analysis are suitable for description of genetic differences among genotypes (Saunders et al. 2001). For identification of differences between 22 genotypes of *Cannabis* a dendrogram (Figure 1) was therefore constructed. *Cannabis indica* ('Royal Caramel') (cluster I) is the most distant one from all other genotypes (cluster II). These genotypes were divided into two parts of subcluster A and B. The first subcluster IIA included: *C. indica*, *C. sativa* and the hybrids, and the second group (subcluster IIB): *C. sativa* and industrial hemp. It was possible in the dendrogram to differentiate industrial hemp (“TE”) from other types of *Cannabis*.

Figure 1 Dendrogram of genetic similarity



Legend: I and II – cluster, A and B – subcluster, TE1 – 'Finola', TE2 – 'Tiborszálási', TE3 – 'Tisza', TE4 – 'Kompolti',TE5 – 'Kompolti hybrid TC', TE6 – 'Carmagnola', SA1 - 'Arjan's Haze', SA2 – 'Amnesia Haze', SA3 – 'Sour Diesel',SA4 – 'Presidential O.G.', IN1 – 'Great White Shark', IN2 – 'Nothern Light', IN3 – 'Skunk', IN4 – 'Royal Caramel',SM1 – 'Special Queen', SM2 – 'Opium', SM3 – 'Mohan Ram', SM4 – 'Biddy Early', HY1 – 'Super Bud', HY2 – 'Royal Medic', HY3 – 'Big Bud XXL', HY4 – 'Jack Herrer Automatic'

CONCLUSION

In our study of genetic variability 16 SSR markers for *Cannabis* were evaluated. Number of alleles ranged from 1 to 7 alleles with an average 4.75 per locus. Six SSR markers suitable for identification of industrial hemp (*ANUC203*, *ANUC205*, *CAN0039*, *CAN0093*, *CAN0110* and *CAN2913*) and one uniform marker (*CAN1690B*) were detected in this study of genetic variability. The dendrogram displayed as the most distant sample *Cannabis indica* ('Royal Caramel'). From used genotypes the least variability was observed among industrial hemp varieties. The results seem to be useful for characterizing genetic diversity among *Cannabis* samples. In further genetic analyses of variability the set of used varieties and SSR markers should be expanded, which should improve the validity of the results.

ACKNOWLEDGEMENT

The authors thank to Hana Gabrielová and Arno Hazekamp for providing the experimental material and Barbora Balgová for help with experiments. This work was supported by IGA FA MENDELU No. TP 4/2015.

REFERENCES

Alghanim H. J., Almirall J. R. 2003. Development of Microsatellite Markers in *Cannabis sativa* for DNA Typing and Genetic Relatedness Analyses. *Analytical and Bioanalytical Chemistry*, 376: 1225–1233.

Cordeiro G. M., Taylor G. O., Henry R. J. 2000. Characterisation of Microsatellite Markers from Sugarcane (*Saccharum*sp.), a Highly Polyploid Species. *Plant Science*, 155: 161–168.

Gao C., Xin P., Cheng C., Tang Q., Chen P., Wang C., Zang G., Zhao L. 2014. Diversity Analysis in *Cannabis sativa* Based on Large-Scale Development of Expressed Sequence Tag-Derived Simple Sequence Repeat Markers. *PLoS ONE*, 9(10): e110638.

Gilmore S., Peakall R. 2003. Isolation of Microsatellite Markers in *Cannabis sativa* L. (marijuana). *Molecular Biology Notes*, 3: 105–107.

Hampl V., Pavlíček A., Flegr J. 2001. Construction and Bootstrap Analysis of DNA Fingerpriting-Based Phylogenetic Trees with a Freeware Program FreeTree: Application to Trichomonad Parasites. *International Journal of Systematic and Evolutionary Microbiology*, 51(4): 731–735.

Hilling K. W., Mahlberg P. G. 2004. A Chemotaxonomic Analysis of Cannabinoid Variation in *Cannabis* (Cannabaceae). *American Journal of Botany*, 91(6): 966–975.

Jaccard P. 1908. NouvellesRecherches sur la Distribution Florale. *Bulletin del la SociétéVaudoise des Sciences Naturelles,* 44: 223–270.

Kayis S. A., Hakki E. E., Pinarkara E. 2010. Comparison of Effectiveness of ISSR and RAPD Markers in Genetic Characterization of Seized Marijuana (*Cannabis sativa* L.) in Turkey. *African Journal of Agricultural Research*, 5(21): 2925–2933.

Marks M. D., Tian L., Wenger J. P., Omburo S. N., Soto-Fuentes W., He J., Gang D. R., Weiblen G. D., Dixon R. A. 2009. Identification of Candidate Genes Affecting Δ-9-Tetrahydrocannabinol Biosynthesis in *Cannabis sativa*. *Journal of Experimental Botany*, 60(13): 3715–3726.

Page R. D. M. 1996. Treeview: An Application to Display Phylogenetic Trees on Personal Computers. *Computer Applications in the Biosciences*, 12(3): 357–358.

Russell K., Fuller J., Young G., Thomas B., Taramino G., Maculay M., Waugh R., Powell W. 1997. Discriminating between Barley Genotypes Using Microsatellite Markers. *Genome,* 40(4): 442–450.

Sarwat M., Nabi G., Das S., Srivastava P. S. 2012. Molecular Markers in Medicinal Plant Biotechnology: Past and Present. *Critical Reviews in Biotechnology*, 32(1): 74–92.

Saunders J. A., Pedroni M. J., Penrose L. D. J., Fist A. J. 2001. AFLP Analysis of Opium Poppy. *Crop Science*, 41: 1596–1601.

Thichak S., Natakankitkul S., Chansakaow S., Chutipongvivate S. 2011. Identification of Drug-Type and Fiber-Type of Hemp (*Cannabis sativa* L.) by Multiplex PCR. *Chiang Mai Journal of Science*, 38(4): 608–618.

Varshney R. K., Graner A., Sorrells M. E. 2005. Genic Microsatellite Markers in Plants: Features and Applications. *Trends in Biotechnology*, 23(1): 48–55.

Zhang D-X. 2004. Lepidopteran Microsatellite DNA: Redundant but Promising. *Trends in Ecology and Evolution*, 19(10): 507–509.