Int J Legal Med (2012) 126:601–606
DOI 10.1007/s00414-012-0706-6

ORIGINAL ARTICLE

The validation of a 15 STR multiplex PCR for *Cannabis* species

Stephan Köhnemann · Johanna Nedele · Daniela Schwotzer · Julia Morzfeld · Heidi Pfeiffer

Received: 31 January 2012 / Accepted: 26 April 2012 / Published online: 10 May 2012

Abstract Trade and acquisition of *Cannabis* drugs are illegal in many countries worldwide; nevertheless, crimes related with these drugs are a major problem for the investigative authorities. With this manuscript, we want to introduce a 15 short tandem repeat (STR) *Cannabis* marker set that can be amplified in one PCR reaction. This multiplex PCR is specific to *Cannabis* species and combines highly informative STR markers. The 15 STR multiplex is easy to use and was validated according to common laboratory quality standards. Due to the fact that a lot of *Cannabis* plants are cultivated by clonal propagation and may show aneuploidy, polyploidy or multiple gene loci, it is not possible to apply biostatistics that follow the Hardy–Weinberg law. However, this multiplex will help the police to trace back trade routes of drug syndicates or dealers and it can help to link *Cannabis* plants to a crime scene.

Keywords *Cannabis* DNA short tandem repeat · STR

Introduction

Cannabis species are economic plants used for the production of food, fibres, oils and intoxicants [1]. Some of them belong to the most frequently used illicit drugs worldwide and are therefore prohibited by law in many countries [2]. The main psychoactive compound of drug *Cannabis* is Δ9-tetrahydrocannabinol (THC) which can be found in high concentrations in leaves and inflorescences.

Techniques to distinguish *Cannabis* on the DNA level were developed previously by different working groups [3–7]. Short tandem repeats (STRs) appear to have the best discrimination ability; nevertheless, none of these groups tried to combine more than six STRs in one multiplex PCR. STRs are repetitive sequences of up to six bases at a defined gene locus, found normally in the non-coding region of autosomal or gonosomal DNA. They should be flanked by a conserved DNA region so that it is possible to create homologues and species-specific primers for PCR amplification.

The idea of using DNA analysis on *Cannabis* was first triggered by the question of whether or not it would be possible to assign *Cannabis* samples to specific crime scenes/*Cannabis* plantations. It was assumed that this should be possible in cases in which *Cannabis* plants were clonally reproduced from so-called "mother plants". Due to the fact that a lot of *Cannabis* plants are bred by clonal propagation [5] and may show aneuploidy, polyploidy or multiple gene loci [8], it is not possible to apply Mendelian inheritance and Mendelian inheritance-based biostatistics like the Hardy–Weinberg law.

In recent years, forensic botany has become an interesting part of forensic molecular genetics [9, 10]. Relying on the topic of the introduction, this study has the aim to provide a user-friendly multiplex PCR that contains a high amount of *Cannabis* STR systems that have a good discrimination and multiplexing ability to help the police to trace back trade routes of drug syndicates or dealers and help them to link different *Cannabis* plants to a crime scene.

S. Köhnemann (✉) · J. Nedele · D. Schwotzer · H. Pfeiffer
Institute of Legal Medicine,
Röntgenstr. 23,
48149 Münster, Germany
e-mail: stephan.koehnemann@ukmuenster.de

J. Morzfeld
Federal Criminal Police Office,
Wiesbaden, Germany

Material and methods

Cannabis samples and DNA extraction

Cannabis samples (n=63) were taken from 41 police seizures and extracted by two different methods. The first extraction method was based upon the NucleoSpin Food kit (Macherey-Nagel, Düren, Germany), following the manufacturer's recommendations. However, the extraction was accomplished with 100 mg homogenised plant material instead of 200 mg plant material. The second extraction method was based upon the QuickExtract Plant DNA Extraction Solution (Epicentre Biotechnologies, Madison, WI, USA). Here, it was not necessary to homogenise the samples coming from leaves or inflorescence; 1 cm^2 of plant material was put directly into 100 µl extraction solution. The sample was then incubated at 65°C for 6 min, following 2 min at 98°C.

Some samples, depending on the quality of the first PCR results, were purified with the Qiagen PCR purification kit according to the manufacturer's recommendations. UV-Vis photospectrometry was accomplished to analyse the DNA concentration.

PCR primers and multiplex conditions

All samples were analysed at least twice by 15 different *Cannabis* multiplex STRs. The PCR reaction mix was run with 50 ng DNA, if not stated differently, in a total volume of 14.0 µl. The reaction mix contained 1.50 µl of $MgCl_2$ (50 mM), 0.2 µl of bovine serum albumin, 1.25 µl of AmpliTaq Gold buffer (10×), 0.15 µl AmpliTaq Gold, 1.25 µl of 10 mM dNTPs and each forward and reverse primer (changed according to [3, 5–7]) in a final concentration according to Table 1. The thermocycling started with 10 min at 95°C for the activation of the AmpliTaq Gold polymerase, followed by 25 cycles of (a) 95°C for 30 s, (b) 60°C for 30 s and (c) 72°C for 45 s. A final extension was held at 72°C for 30 min.

Validation studies

To test the sensitivity of the multiplex PCR, a short and a long STR systems, i.e. D02 and B01, were investigated at DNA amounts of 1,020, 510, 255, 128, 64, 32, 16, 8 and 4 ng per PCR reaction. The DNA was quantified via UV-Vis photospectrometry.

The drop-out rate was analysed for 1,020, 510, 255, 128, 64, 32, 16, 8 and 4 ng per PCR reaction in three separated runs. A drop-out was announced when an allele in the STR system was smaller than 50 relative fluorescence measured in units (RFU) or missing.

The primer specificity was tested on known drug and non-drug *Cannabis* samples, human samples, commonly traded tobacco, which is a possible contaminant in *Cannabis* samples, and *H. lupulus,* which belongs to the same plant family like *Cannabis*.

Table 1 Fifteen *Cannabis* STR multiplex primers

STR system	Reverse sequence	Forward sequence	Final primer amount per reaction [pm]	Size
E07 CANN1	GTG GTA GCC AGG TAT AGG TAG	HEX- CAA ATG CCA CAC CAC CTT C	0.500	105–111
ANUCS 302	ATG GTT GAT GTT TTG ATG GT	ROX- AAC ATA AAC ACC AAC AAC TGC	0.375	140–173
H09 CANN2	ACA CAT ACA GAG AGA GCC C -3′	FAM- CGT ACA GTG ATC GTA GTT GAG -3′	0.375	204–224
D02 CANN1	AGA AAT CCA AGG TCC TGA TGG	FAM- GGT TGG GAT GTT GTT GTT GTG	0.088	105–111
C11 CANN1	TGA ATT GGT TAC GAT GGC G	FAM- GTG GTG GTG ATG ATG ATA ATG G	0.125	150–175
B01 CANN1	CCA TAG CAT TAT CCC ACT CAA G -3′	FAM- TGG AGT CAA ATG AAA GGG AAC -3′	0.500	323–339
B05 CANN1	CCC CAA TCT CAA TCT CAA CCC -3′	HEX- TTG ATG GTG GTG AAA CGG C -3′	0.250	235–244
H06 CANN2	ACG TGA GTG ATG ACA CGA G -3′	HEX- TGG TTT CAG TGG TCC TCT C -3′	0.125	266–273
B02 CANN2	TGT TTT CTT CAC TGC ACC C -3′	ROX- CAA CCA AAT GAG AAT GCA ACC	0.750	163–172
H11 CANN1	CAG CGA ACA TTC ACT CTA GCT C -3′	FAM- GCA TGT GGT TGT TTC GTA CCC -3′	0.125	285–297
ANUCS305	AAA GTT GGT CTG AGA AGC AAT	HEX- AAA GTT GGT CTG AGA AGC AAT	0.300	141–162
ANUCS 308	TGG TGC AGG TTT ATA CAA TTT	HEX- AGA TGG TGT TGG GTA TCT TT	1.000	177–203
ANUCS 301	TAA CAA AGT TTC GTG AGG GT	ROX- ATA TGG TTG AAA TCC ATT GC -3′	1.500	209–261
CS1	TAA TGA TGA GAC GAG TGA GAA CG	ATTO501- AAG CAA CTC CAA TTC CAG CC	0.300	134–356
ANUCS 501	AGA GAT CAA GAA ATT GAG ATT CC	ATTO501- AGC AAT AAT GGA GTG AGT GAA C	0.175	80–95

All sequences are shown from 5′- to 3′-end. HEX, ROX, FAM and Atto501 are the dyes attached to the 5′-end. The final amount of PCR primers per PCR reaction is given in picomol

F forward, *R* reverse

The testing for reproducibility was accomplished with six drug *Cannabis* plants in triplicates. From each plant, DNA extracts from root, leaf and inflorescence were tested.

Results

DNA extraction

DNA extraction with the QuickExtract Plant DNA Extraction Solution is a very fast and simple method without removal of salts or other ingredients which inhibit PCR. It was not possible to get high quality typing results for the 15 STR multiplex system. Good analysis results and high levels of DNA of up to 100 ng per µl were achieved when the DNA was extracted with the NucleoSpin Food kit (for a representative electropherogram, see Fig. 1), especially when the plant material was grinded very well. In 100 mg leaf material, usually more than 20 ng DNA/µl was extracted in a total volume of 75 µl.

Sensitivity study

The RFU of both STR systems decreased in a logarithmic way, from the highest amount of DNA down to the smallest amount of DNA. The example for B01 of a heterozygous sample can be seen in Fig. 2. The R^2 for the regression models of allele 1 and allele 2 was 0.99.

It is recommended using at least 32 ng of DNA per PCR reaction to get high quality analysis results with an average RFU of 600 or higher, especially for long and heterozygous DNA systems.

Drop-out rate

There were no dropouts with a DNA amount of 32 ng DNA or more per PCR reaction. With 16 ng DNA per PCR reaction, we had two drop-outs in three runs over 15 STR-systems, i.e. 4.4 %. With 8 ng DNA or less, the drop-out rate was higher than 5 % and was falling below the detection limit of 95 % according to quality management standards. According to the sensitivity study,

Fig. 1 Fifteen *Cannabis* STR multiplex electropherograms. The peaks of the STR systems are labelled with the size and the height. Some dinucleotide repeat systems like H11 can produce broad shoulders. The sample was run with 50 ng DNA and analysed by GeneScan software

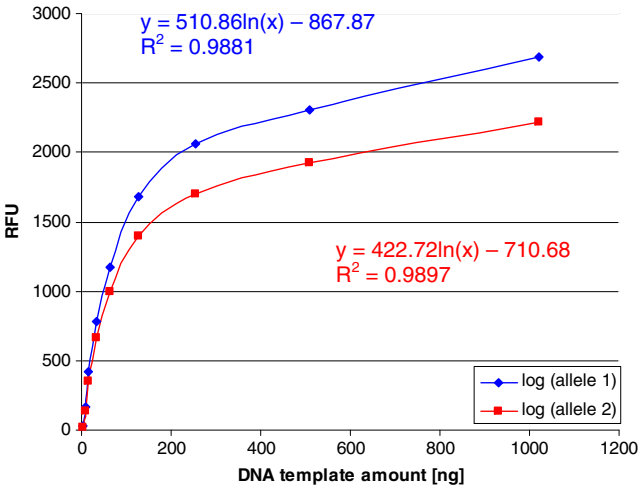

Fig. 2 Logarithmic regression model for STR system B01. It is obvious that there is a small heterozygote imbalance; however, both alleles have an obvious logarithmic model to calculate the RFU depending on the template amount of DNA where *x* is the amount of DNA and *y* is the RFU

it is recommended using at least 32 ng DNA per PCR reaction.

Species specificity

It was possible to get analysis results for drug and non-drug *Cannabis* (hemp) plants using the 15 STR *Cannabis* multiplex PCR. However, it was possible to distinguish between drug and non-drug *Cannabis* plants by sequencing the tetrahydrocannabinolic acid synthase gene (data not shown) [12, 13].

The human (*Homo sapiens*) DNA sample was tested with 50 ng DNA, and no amplification products were visible. Tobaccos from three different cigarette brands were tested, and in all cases the results were negative. The same was true for samples of *H. lupulus* which was verified by Sanger-sequencing species-specific intergenic parts of chloroplast DNA and BLAST analysis (data not shown). According to these results, it is unlikely that DNA from humans, cigarette tobacco or *H. lupulus* could contaminate the 15 *Cannabis* DNA analysis.

Reproducibility

From each plant, DNA extracts from root, leaf and inflorescence were tested and identical results were obtained.

Alleles and genotypes

It was possible to identify 78 alleles and 29 genotypes across 63 samples from 41 cases. The number of all different alleles found in the samples ranged from 11 at locus CS1 to three at loci D02, ANUCS308, B05 and H06 (see Table 2). In 22 samples, the common two allele heterozygosity was not found; instead, three alleles were detected in at least one STR locus. Similar results were already found by previous studies [8, 11, 14].

Table 2 Alleles found in 15 STR systems

STR system	Allele size [mer], rounded	STR system	Allele size [mer], rounded
D02-CANN1	104	B05-CANN1	239
	110		242
	113		245
C11-CANN1	151	H06-CANN2	266
	154		269
	158		272
	160	ANUCS501	97
	164		102
	167		107
	176		111
H09-CANN2	204	CS1	135
	207		178
	217		189
	221		201
H11-CANN1	289		218
	292		224
	294		253
	297		259
	302		276
	304		282
B01-CANN1	314		288
	321	ANUCS302	136
	325		142
	328		155
	338		158
	341		161
	345	B02-CANN2	168
	349		171
E07-CANN1	102		173
	104		177
	108		189
	111	ANUCS301	228
ANUCS305	141		231
	145		234
	148		237
	155		240
ANUCS308	185		246
	189		249
	201		254

The size of the alleles found was rounded

Discussion

With this manuscript, a 15 multiplex PCR is provided, giving the chance to analyse the common *Cannabis* STRs within one PCR reaction. This method is much faster and less labour intensive than the small multiplexes known from the literature. It is possible to use this assay for the analysis of regional *Cannabis* populations to help the police trace back trade routes of drug syndicates or dealers with maximised genetic information. This assay can also help to link *Cannabis* plants to a crime scene like an illegal *Cannabis* plantation. Due to the fact that *Cannabis* plants are mostly grown by clonal propagation, the limitation of the STR analysis is finding of DNA profiles that occur more often than others. Clonal propagation is the reason why it is not possible to apply Hardy–Weinberg biostatistics. However, it is possible to collect the different *Cannabis* DNA profiles found in a database. It is necessary to discuss the rules by which a *Cannabis* DNA profile is put into that database.

Due to the fact that the analysed samples of this study are real case samples, it was not known whether they are related or not before they were analysed. That is why the results are not representative for frequency estimations or the local *Cannabis* population. All cases are originating from a period of 3 months in which they were brought to the analysing laboratory for THC concentration measurements. According to the number of alleles, the overall distribution is an interesting aspect because another *Cannabis* STR study found the lowest number of alleles for *Cannabis* STR systems ANUCS501 and B02 [6]. Mendoza et al. [7] state that it was not possible to relate 98 *Cannabis* samples from 33 US states to each other. They used a six STR multiplex system and found 29 different alleles at locus ANUCS305, being their most diverse locus. The diversity of ANUCS305 in this study was relatively low, only four different alleles were found. These results suggest that there are possible geographic differences, although it may be difficult to predict the region from which a *Cannabis* plant was grown. Gilmore et al. [15] were analysing chloroplast and mitochondrial DNA loci. They found some *Cannabis* DNA haplotypes that may assist in the prediction of geographic origin, but they also recommended analysing more *Cannabis* STR loci to facilitate the prediction of geographic origin.

Some of the *Cannabis* samples had an identical profile, and they were coming from the same case, i.e. they were confiscated at the same crime scene. This suggests that these plants are bred by clonal propagation, which is a common technique to cultivate cuttings of mother plants that have a high potential for THC production. With this multiplex PCR, it is very well possible to exclude if two single plants are bred from the same mother plant or not. Nevertheless, one genotype was detected more often than other genotypes, even in different cases. More than one-third of all samples had this genotype (Fig. 1). This may be by chance, but more probably this is due to the fact that there is/was a very potential mother plant where a lot of daughter plants are derived from. These daughter plants now are themselves mother plants. The occurrence of one frequent genotype was also confirmed by the study of Howard et al. [6]. In 55 % of their *Cannabis*, samples they found an identical DNA profile.

Some *Cannabis* plants showed multiple alleles in at least one STR locus. Knight et al. [8] mentioned a reason for multiple alleles. They investigated the yield of *Cannabis* crops, and the genetic profiling was tested in three growing cycles of six different plants. Five *Cannabis* STR markers (ANUCS301, ANUCS302, ANUCS303, ANUCS304 and ANUCS306) were used to create genetic profiles. One plant, morphologically indistinguishable from the others, showed a tri-allelic pattern at two loci (ANUCS302, ANUCS303). Here, interbreeding could be excluded and the authors suggested the occurrence of polyploidy or locus duplication.

Adams et al. [16] wrote that polyploidy can be found in 57–70 % of flowering plants. However, polyploidy may also be induced artificially with colchicine treatment [17]. Colchicine is a toxic substance extracted from the roots of a certain *Colchicum* species. It inhibits cell mitosis, resulting in larger daughter cells with multiple chromosome sets. De Pasquale et al. [18] conducted studies where *Cannabis* plants were treated with colchicine which led to a THC increase of 166–250 %. These observations suggest that colchicine treatment of *Cannabis* is a common method to obtain polyploidy plants that produce high levels of THC. Another recent study [19] found that aneuploidy and the loss or gain of one chromosome can be induced by more than 400 substances. An actual study found that locus duplication may be a common part of plant evolution [20].

The detection of multiple alleles, clonal propagation and selective breeding makes an estimation of allele frequencies for *Cannabis* STRs extremely difficult (see also [19]). In 2005, an article by De Silva et al. [21] discussed the estimation of allele frequencies in polyploidy plants under certain patterns of inheritance. According to that article, horticulture plants often show polyploidy, but microsatellites cannot predict the number of alleles within these plants. The occurrence of polyploidy complicates the calculations of allele frequencies and the relative frequency of *Cannabis* DNA profiles. For future research, it is very important to know the reason for multiple alleles in *Cannabis* species.

In conclusion, this assay will be of help for the police to get genetic information of *Cannabis* plants. The police can use that information to link plants to a crime scene or to trace back trade routes of crimes related with *Cannabis*. For that reason, the *Cannabis* DNA profiles can be stored in a database to facilitate the search for identical plant profiles. However, finding identical *Cannabis* DNA profiles will also

limit the forensic application. The mode of breeding *Cannabis* plants by clonal propagation does not allow applying Hardy–Weinberg biostatistics and frequency estimations.

Acknowledgments The authors would like to thank Marianne Schürenkamp, Ulla Sibbing, Rachel Challis, Stefanie Scheiper and The Hemp Collaboration (THC) for the help and discussions.

Conflicts of interest The authors declare no conflict of interest.

References

1. Mercuri A, Accorsi C, Mazzanti M (2002) The long history of *Cannabis* and its cultivation by the Romans in central Italy, shown by pollen records from Lago Albano and Lago di Nemi. Veg Hist Archaeobot 11:263–276
2. Anderson P (2006) Global use of alcohol, drugs and tobacco. Drug Alcohol Rev 25:489–502
3. Linacre A, Thorpe J (1998) Detection and identification of *Cannabis* by DNA. Forensic Sci Int 91:71–76
4. Siniscalco Gigliano G, Caputo P, Cozzolino S (1997) Ribosomal DNA analysis as a tool for the identification of *Cannabis sativa* L. specimens of forensic interest. Sci Justice 37:171–174
5. Alghanim HJ, Almirall JR (2003) Development of microsatellite markers in *Cannabis sativa* for DNA typing and genetic relatedness analyses. Anal Bioanal Chem 376:1225–1233
6. Howard C, Gilmore S, Robertson J, Peakall R (2008) Developmental validation of a *Cannabis sativa* STR multiplex system for forensic analysis. J Forensic Sci 53:1061–1067
7. Mendoza MA, Mills DK, Lata H, Chandra S, ElSohly MA, Almirall JR (2009) Genetic individualization of *Cannabis sativa* by a short tandem repeat multiplex system. Anal Bioanal Chem 393:719–726
8. Knight G, Hansen S, Connor M, Poulsen H, McGovern C, Stacey J (2010) The results of an experimental indoor hydroponic *Cannabis* growing study, using the 'Screen of Green' (ScrOG) method—yield, tetrahydrocannabinol (THC) and DNA analysis. Forensic Sci Int 202:36–44
9. Bruni I, De Mattia F, Galimberti A, Galasso G, Banfi E, Casiraghi M, Labra M (2010) Identification of poisonous plants by DNA barcoding approach. Int J Legal Med 124:595–603
10. Ferri G, Alù M, Corradini B, Beduschi G (2009) Forensic botany: species identification of botanical trace evidence using a multigene barcoding approach. Int J Legal Med 123:395–401
11. Allgeier L, Hemenway J, Shirley N, Lanier T, Coyle HM (2011) Field testing of collection cards for *Cannabis sativa* samples with a single hexanucleotide DNA marker. J Forensic Sci 56:1245–1249
12. Taura F, Sirikantaramas S, Shoyama Y, Shoyama Y, Morimoto S (2007) Phytocannabinoids in *Cannabis sativa*: recent studies on biosynthetic enzymes. Chem Biodivers 4:1649–1663
13. de Meijer EP, Bagatta M, Carboni A, Crucitti P, Moliterni VM, Ranalli P, Mandolino G (2003) The inheritance of chemical phenotype in *Cannabis sativa* L. Genetics 163:335–346
14. Hsieh HM, Hou RJ, Tsai LC, Wei CS, Liu SW, Huang LH, Kuo YC, Linacre A, Lee JC (2003) A highly polymorphic STR locus in *Cannabis sativa*. Forensic Sci Int 131:53–58
15. Gilmore S, Peakall R, Robertson J (2007) Organelle DNA haplotypes reflect crop-use characteristics and geographic origins of *Cannabis sativa*. Forensic Sci Int 172:179–190
16. Adams KL, Wendel JF (2005) Polyploidy and genome evolution in plants. Curr Opin Plant Biol 8:135–141
17. Metz SG, Sharma HC, Armstrong TA, Mascia PN (1988) Chromosome doubling and aneuploidy in anther-derived plants from two winter wheat lines. Genome 30:177–181
18. De Pasquale A, Tumino G, De Pasquale RC (1974) Micromorphology of the epidermic surfaces of female plants of *Cannabis sativa* L. UN Bull Narc 26:27–40
19. Sharma CB (1990) Chemically induced aneuploidy in higher plants. Mutagenesis 5:105–125
20. Flagel LE, Wendel JF (2009) Gene duplication and evolutionary novelty in plants. New Phytol 183:557–564
21. De Silva HN, Hall AJ, Rikkerink E, McNeilage MA, Fraser LG (2005) Estimation of allele frequencies in polyploids under certain patterns of inheritance. Heredity 95:327–334