Forensic Science International 329 (2021) 111053

Contents lists available at ScienceDirect

Forensic Science International

journal homepage: www.elsevier.com/locate/forsciint

A Ge.F.I. – ISFG European collaborative study on DNA identification of *Cannabis sativa* samples using a 13-locus multiplex STR method

M. Di Nunzio[a,c,*], V. Agostini[d], F. Alessandrini[e], C. Barrot-Feixat[a], A. Berti[f], C. Bini[g], M. Bottinelli[h], E. Carnevali[i], B. Corradini[j], M. Fabbri[k], P. Fattorini[l], P. Garofano[m], S. Gino[n], A. Mameli[o], A. Marino[p], C. Previderè[q], C. Robino[r], C. Romano[s], P. Tozzo[t], A. Verzeletti[u], L. Buscemi[e], D. Gangitano[a], C. Di Nunzio[b,c]

[a] *University of Barcelona, Spain*
[b] *University of Catanzaro, Italy*
[c] *Biogem Scarl, Ariano Irpino, Italy*
[d] *University of Piemonte Orientale, Alessandria, Italy*
[e] *University of Marche, Ancona, Italy*
[f] *Carabinieri Scientific Investigation Department of Roma, Italy*
[g] *University of Bologna, Italy*
[h] *Molecular Diagnostics Laboratory, Gentilino, Switzerland*
[i] *University of Perugia, Italy*
[j] *University of Modena and Reggio Emilia, Italy*
[k] *University of Ferrara, Italy*
[l] *University of Trieste, Italy*
[m] *CAD, Anti Doping Center, Sanatorio San Luigi, Torino, Italy*
[n] *University of Piemonte Orientale, Novara, Italy*
[o] *Carabinieri Scientific Investigation Department of Cagliari, Italy*
[p] *Carabinieri Scientific Investigation Department of Parma, Italy*
[q] *University of Pavia, Italy*
[r] *University of Torino, Italy*
[s] *Carabinieri Scientific Investigation Department of Messina, Italy*
[t] *University of Padova, Italy*
[u] *University of Brescia, Italy*

ARTICLE INFO

Article history:
Received 2 July 2021
Received in revised form 29 September 2021
Accepted 6 October 2021
Available online 14 October 2021

Keywords:
Forensic plant science
Cannabis sativa
Short tandem repeats
DNA typing

ABSTRACT

Cannabis sativa is the most used controlled substance in Europe. With the advent of new and less restrictive European laws on cannabis sale for recreational use (including in Italy), an increase in indoor cannabis crops were observed. This increase was possible due to the availability of cannabis seeds through the internet market. Genetic identification of cannabis can link seizures and if in possession then might aid in an investigation. A 13-locus multiplex STR method was previously developed and validated by Houston et al. A collaborative exercise was organized by the Italian Forensic Geneticists – International Society of Forensic Genetics (Ge.F.I. – ISFG) Working Group with the aim to test the reproducibility, reliability and robustness of this multiplex cannabis STR kit. Twenty-one laboratories from three European countries participated in the collaborative exercise and were asked to perform STR typing of two cannabis samples. Cannabis DNA samples and the multiplex STR kit were provided by the University of Barcelona and Sam Houston State University. Different platforms for PCR amplification, capillary electrophoresis (CE) and genotyping software were selected at the discretion of the participating laboratories. Although the participating laboratories used different PCR equipment, CE platforms and genotyping software, concordant results were obtained from the majority of the samples. The overall genotyping success ratio was 96%. Only minor artifacts were observed. The mean peak height ratio was estimated to be 76.3% and 78.1% for sample 1 and sample 2, respectively. The lowest amount of −1 / + 1 stutter percentage produced, when the height of the parent allele was higher than 8000 RFU, resulted to be less than 10% of the parent allele height. Few common issues

* Correspondence to: Forensic Genetics Laboratory - Legal Medicine Unit, Department of Medicine, University of Barcelona, Spain.
E-mail address: michele.dinunzio@ub.edu (M. Di Nunzio).

https://doi.org/10.1016/j.forsciint.2021.111053

were observed such as a minor peak imbalance in some heterozygous loci, some artifact peaks and few instances of allelic drop-out. The results of this collaborative exercise demonstrated the robustness and applicability of the 13-locus system for cannabis DNA profiling for forensic purposes.

1. Introduction

C. sativa is the subject of one of the most common drug law offenses in Europe. Since 2014, cannabis accounted for almost 60% of an overall estimate of 1.6 million offenses including possession and trafficking [1]. Cannabis can be classified into legal fiber type (hemp) and illegal drug type (marijuana). Marijuana differs from hemp based on the high level of Δ9-tetrahydrocannabinol (Δ9-THC). In Italy, the possession of limited amounts of marijuana, i.e., containing 500 mg of Δ9-THC [2], is considered only a civil offense. On the contrary, trafficking and selling cannabis is prohibited and punished by law. On the other hand, support and promotion of hemp crops is allowed for textiles, food, cosmetic production and bioengineering industry [3].

The Scientific Working Group for the Analysis of Seized Drugs (SWGDRUG) recommended a series of analyses to confirm the presence of cannabis [4–7]. Nevertheless, none of these analytical tests are able to individualize cannabis plants or crop type. During 1990′s, different *C. sativa* DNA methods were developed to individualize and to determine the origin of plants for forensic purposes. Thus, various molecular techniques have been applied including, random amplified polymorphic DNA (RAPD), amplified fragment length polymorphism (AFLP) and short tandem repeats (STRs) [8–18]. STRs are the gold standard for human identification (HID) and as such they can also be applied to *C. sativa* for forensic purposes. Moreover, studies that included STRs *C. sativa* genotyping have reached promising results for forensic purposes. In 2016, a 13-locus cannabis STR method [10] was developed and validated according to guidelines of the International Society of Forensic Genetics (ISFG) [19] and the Scientific Working Group on DNA Analysis (SWGDAM) [20] recommendations for the use of non-human DNA. Although this STR method cannot distinguish hemp from drug type, a previous study showed that groups of cannabis seizures could be associated using phylogenetic analysis demonstrating its application for intelligence purposes [10]. Generally, when a new method is developed, inter-laboratory tests were carried out to test its robustness, reliability and reproducibility [21,22]. The objective of this work was to test this newly developed cannabis STR multiplex kit through an inter-laboratory exercise with different laboratories in Europe. Up to date, this is the first collaborative exercise on STR identification of *C. sativa*.

2. Materials and methods

2.1. Collaborative exercise design

After approval by the Italian Forensic Geneticists (Ge.F.I.) society, a collaborative exercise to test a 13-locus cannabis STR was organized. Twenty-one European laboratories from different institutions, using different DNA analysis platforms, participated in the exercise (Table 1). Sealed styrofoam boxes were prepared and shipped by courier to each participant laboratory. Each box contained: two cooling tablets, a self-sealing pre-PCR bag and a self-sealing post-PCR bag. The pre-PCR bag contained a PCR master mix tube (46.25 µL), a primer mix tube (6.25 µL) and sample set tubes in three separate bags. The sample set contained three aliquots (5 µL each) of *C. sativa* DNA (positive control, sample 1 and sample 2). The PCR master mix and the primer mix were prepared according to Houston et al. [10]. Lastly, an aliquot of 5 µL of allelic ladder was contained in the post-PCR self-sealing bag. Protocols for PCR amplification and DNA genotyping were also provided. Participants used their standard DNA genotyping platforms, as well as the interpretation and reporting guidelines. Some reagents used in this exercise were kindly provided by Dr. Rachel Houston (Sam Houston State University, Huntsville, TX, USA).

2.2. DNA quantitation of cannabis samples

Cannabis plant samples were kindly provided by the Instituto Toxicológico de Barcelona, Spain. Samples were previously quantified by the organizing laboratory. For this purpose, a cannabis DNA standard for the calibration curve was prepared according to

Table 1
Laboratory instrumentation, polymer, array, and software used in the inter-laboratory study.

Laboratory code	Laboratory type	Thermocycler	DNA sequencer	Size standard	Array	Polymer	Genotype software
1	Private	Veriti AB®	ABI 3500/3500 XL	LIZ®600	36 cm	POP-4™	GeneMapper® ID-X v 1.4
2	University	Veriti AB®	ABI 310	LIZ®500	36 cm	POP-4™	GeneMapper® ID-X v 1.1
3	University	9700 Thermocycler AB®	ABI 3500/3500 XL	LIZ®500	50 cm	POP-7™	GeneMapper® ID-X v 3.2
4	University	Mastercyler EP Eppendorf	ABI 310	LIZ®500	47 cm	POP-4™	GeneMapper® v 3.2.1
5	University	T100 Biorad	ABI 310	LIZ®500	36 cm	POP-4™	GeneMapper® ID-X v 3.2
6	University	Veriti AB®	ABI SeqStudio	LIZ®600	28 cm	POP-1™	GeneMapper® ID-X v 1.6
7	University	Veriti AB®	ABI 3130	LIZ®500	36 cm	POP-4™	GeneMapper® ID-X v 1.5
8	Private	Biometra T300	ABI 3500/3500 XL	LIZ®600	36 cm	POP-4™	GeneMapper® ID-X v 1.5
9	University	9700 Thermocycler AB®	ABI 3500/3500 XL	LIZ®500	36 cm	POP-4™	GeneMapper® ID-X v 1.4
10	University	MiniAmp AB	ABI 3130	LIZ®500	36 cm	POP-4™	GeneMapper® ID v 3.2
11	University	9700 Thermocycler AB®	ABI SeqStudio	LIZ®500	28 cm	POP-1™	GeneMapper® v 5
12	University	Mastercyler EP Eppendorf	ABI 310	LIZ®500	47 cm	POP-4™	GeneMapper® ID v 3.2.1
13	University	SimpliAmp AB®	ABI 3130	LIZ®500	36 cm	POP-4™	GeneMapper® ID-X v 3.2
14	Law Enforcement	Veriti AB®	ABI 3500/3500 XL	LIZ®600	36 cm	POP-4™	GeneMapper® ID-X v 1.6
15	Law Enforcement	9700 Thermocycler AB®	ABI 3500/3500 XL	LIZ®600	36 cm	POP-4™	GeneMapper® ID-X v 1.6
16	Law Enforcement	Veriti AB®	ABI 3500/3500 XL	LIZ®500	36 cm	POP-4™	GeneMapper® ID-X v 3
17	Law Enforcement	ProFlex PCR System	ABI 3500/3500 XL	LIZ®600	36 cm	POP-4™	GeneMapper® ID-X v 1.6
18	University	9700 Thermocycler AB®	ABI 310	LIZ®500	36 cm	POP-4™	GeneMapper® ID-X v 3.2
19	University	9700 Thermocycler AB®	ABI 3730	LIZ®500	50 cm	POP-7™	GeneMapper® ID-X v 3.2
20	University	9700 Thermocycler AB®	ABI 3500/3500 XL	LIZ®600	36 cm	POP-4™	GeneMapper® ID-X v 1.5
21	Core Lab	Mastercyler EP Eppendorf	ABI 3500/3500 XL	LIZ®600	36 cm	POP-4™	GeneMapper® ID-X v 1.4

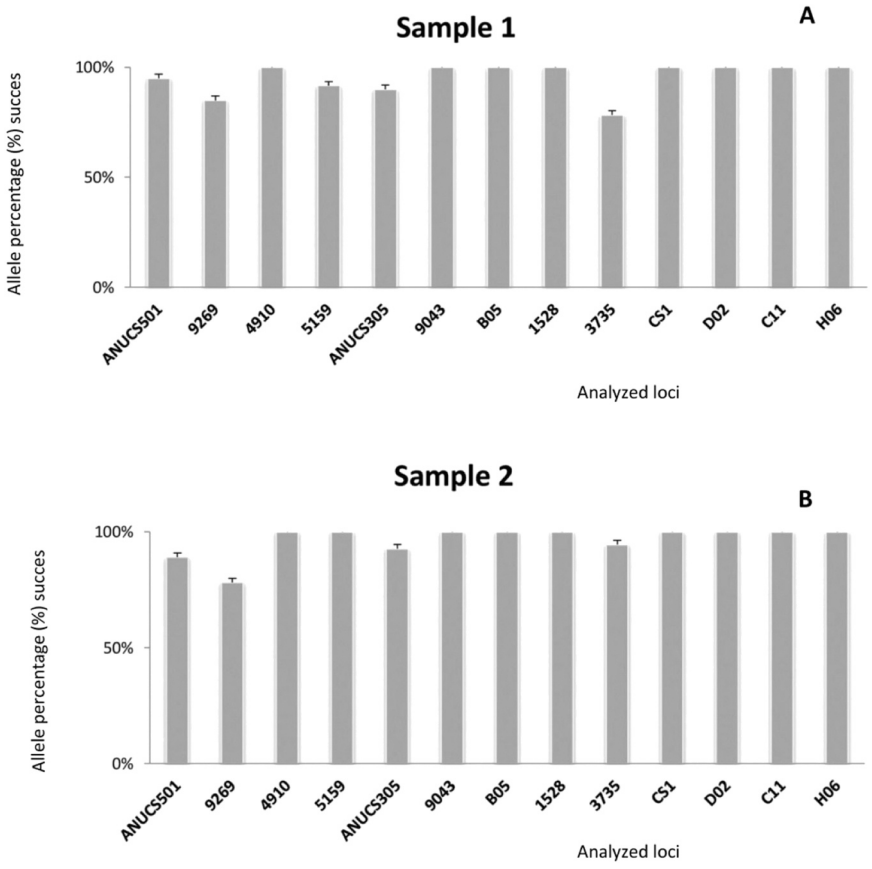

Fig. 1. Percentage of genotyping success for sample 1 (A) and for sample 2 (B), per locus between the participating laboratories, using the 13-locus multiplex STR C. sativa kit. Error bars represent standard error.

Table 2
Results of 13 cannabis STR typing of sample 1 of participating laboratories in the collaborative exercise.

Lab code	ANUCS 501	9269	4910	5159	ANUCS 305	9043	B05	1528	3735	CS1	D02	C11	H06
1	6,6	6,6	4,10	4,4.2	8,8	5,6	8,8	6,7	3,9	23,28	7,7	14,14	8,8
2	6,6	6,7[a]	4,10	4,4.2	8,8	5,6	8,8	6,7	3,OL[b]	23,28	7,7	14,14	8,8
3	6,6	6,6	4,10	4,4.2	8,8	5,6	8,8	6,7	3,9	23,28	7,7	14,14	8,8
4	6,6	6,6	4,10	4,4.2	8,8	5,6	8,8	6,7	3,9	23,28	7,7	14,14	8,8
5	6,6	6,6	4,10	4,4.2	8,8	5,6	8,8	6,7	3,9	23,28	7,7	14,14	8,8
6	6,6	NR	4,10	4,4.2	NR	5,6	8,8	6,7	NR	23,28	7,7	14,14	8,8
7	6,6	6,6	4,10	4,4.2	4,8[a]	5,6	8,8	6,7	3,OL[b]	23,28	7,7	14,14	8,8
8	6,6	6,6	4,10	4,4.2	8,8	5,6	8,8	6,7	3,9	23,28	7,7	14,14	8,8
9	6,6	6,6	4,10	4,4.2	8,8	5,6	8,8	6,7	3,3[c]	23,28	7,7	14,14	8,8
10	6,6	6,6	4,10	4,4.2	8,8	5,6	8,8	6,7	3,9	23,28	7,7	14,14	8,8
11	6,6	6,6	4,10	4.2,5[b]	8,8	5,6	8,8	6,7	3,9	23,28	7,7	14,14	8,8
12	6,6	6,6	4,10	4,4.2	8,8	5,6	8,8	6,7	3,9	23,28	7,7	14,14	8,8
13	NR	NR	4.10	4.2,5[b]	8,8	5,6	8,8	6,7	3,9	23,28	7,7	14,14	8,8
14	6,6	6,6	4,10	4,4.2	8,8	5,6	8,8	6,7	3,9	23,28	7,7	14,14	8,8
15	6,6	6,6	4,10	4,4.2	8,8	5,6	8,8	6,7	3,9	23,28	7,7	14,14	8,8
16	6,6	6,6	4,10	4,4.2	8,8	5,6	8,8	6,7	3,9	23,28	7,7	14,14	8,8
17	6,6	6,6	4,10	4,4.2	8,8	5,6	8,8	6,7	3,9	23,28	7,7	14,14	8,8
18	6,6	6,6	4,10	4,4.2	8,8	5,6	8,8	6,7	3,9	23,28	7,7	14,14	8,8
19	6,6	6,6	4,10	4,4.2	8,8	5,6	8,8	6,7	3,9	23,28	7,7	14,14	8,8
20	6,6	6,6	4,10	4,4.2	8,8	5,6	8,8	6,7	3,9	23,28	7,7	14,14	8,8

NR - No results obtained.
[a] False heterozygous locus, due to a wrong Bin Set.
[b] Wrong allele call, due to a wrong Bin Set.
[c] Missing one allele at heterozygous locus.

Houston et al. [10]. All samples used in the exercise were quantified by a real-time PCR equipment (7500 Real-Time PCR Systems, Applied Biosystems™, Carlsbad, CA, USA) using the SYBR™ Green PCR Mastermix (Applied Biosystems™, Carlsbard, CA, USA) and *C. sativa* specific primers, ANUCS304 (Thermo Fisher Scientific) [15]. For each calibration curve, the reference DNA sample was serially diluted in a

Table 3
Results of 13 cannabis STR typing of sample 2 of participating laboratories in the collaborative exercise.

Lab code	ANUCS 501	9269	4910	5159	ANUCS 305	9043	B05	1528	3735	CS1	D02	C11	H06
1	5,6	5.3,6	4,4	4,4.2	6,8	5,5	8,9	7,7	4,4	23,27	6,8	13,14	8,9
2	5,6	5.3,6	4,4	4,4.2	6,8	5,5	8,9	7,7	4,4	23,27	6,8	13,14	8,9
3	5,6	6,6[a]	4,4	4,4.2	6,8	5,5	8,9	7,7	4,4	23,27	6,8	13,14	8,9
4	5,6	5.3,6	4,4	4,4.2	6,8	5,5	8,9	7,7	4,4	23,27	6,8	13,14	8,9
5	NR	NR	4,4	4,4.2	NR	5,5	8,9	7,7	4,4	23,27	6,8	13,14	8,9
6	NR	NR	4,4	4,4.2	NR	5,5	8,9	7,7	NR	23,27	6,8	13,14	8,9
7	5,6	5.3,6	4,4	4,4.2	6,8	5,5	8,9	7,7	4,4	23,27	6,8	13,14	8,9
8	5,6	5.3,6	4,4	4,4.2	6,8	5,5	8,9	7,7	4,4	23,27	6,8	13,14	8,9
9	5,6	5.3,6	4,4	4,4.2	6,8	5,5	8,9	7,7	4,4	23,27	6,8	13,14	8,9
10	5,6	5.3,6	4,4	4,4.2	6,8	5,5	8,9	7,7	4,4	23,27	6,8	13,14	8,9
11	5,6	5.3,6	4,4	4,4.2	6,8	5,5	8,9	7,7	4,4	23,27	6,8	13,14	8,9
12	5,6	5.3,6	4,4	4,4.2	6,8	5,5	8,9	7,7	4,4	23,27	6,8	13,14	8,9
13	5,6	5.3,7[b]	4,4	4,4.2	6,8	5,5	8,9	7,7	4,4	23,27	6,8	13,14	8,9
14	5,6	5.3,6	4,4	4,4.2	6,8	5,5	8,9	7,7	4,4	23,27	6,8	13,14	8,9
15	5,6	5.3,6	4,4	4,4.2	6,8	5,5	8,9	7,7	4,4	23,27	6,8	13,14	8,9
16	5,6	5.3,6	4,4	4,4.2	6,8	5,5	8,9	7,7	4,4	23,27	6,8	13,14	8,9
17	5,6	5.3,6	4,4	4,4.2	6,8	5,5	8,9	7,7	4,4	23,27	6,8	13,14	8,9
18	5,6	5.3,6	4,4	4,4.2	6,8	5,5	8,9	7,7	4,4	23,27	6,8	13,14	8,9
19	5,6	5.3,6	4,4	4,4.2	6,8	5,5	8,9	7,7	4,4	23,27	6,8	13,14	8,9
20	5,6	5.3,6	4,4	4,4.2	6,8	5,5	8,9	7,7	4,4	23,27	6,8	13,14	8,9

NR - No results obtained.
[a] Missing one allele at heterozygous locus.
[b] Wrong allele call, due to a wrong Bin Set.

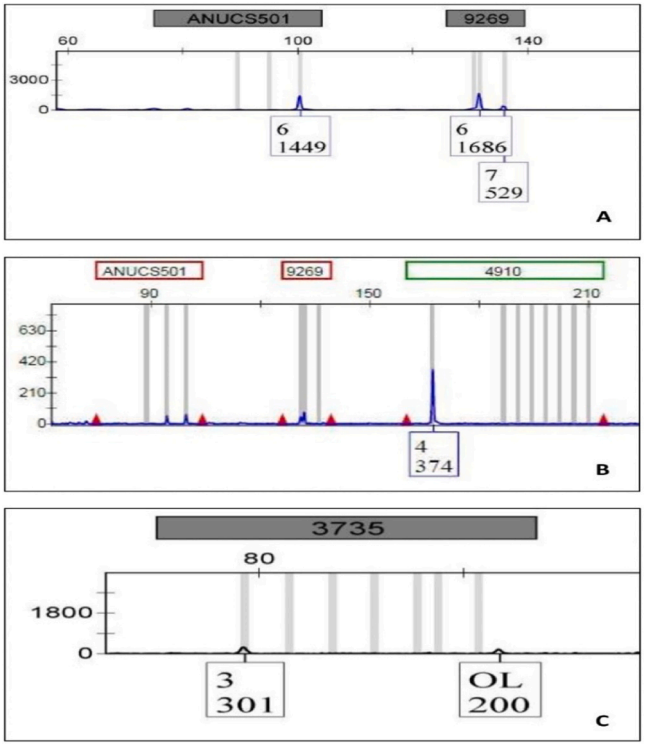

Fig. 2. Examples of typing errors made by participants laboratories: A) False heterozygous 9269 locus, due to an inappropriate allele call; B) No result for the ANUCS501 and 9269 loci, no allele call due to height peak imbalance between loci of the same channel and C) Wrong allele call in a 3735 locus, due to a wrong bin set. The real call was 3,9.

buffer of Tris-EDTA (TE), pH 8.0 to yield eight DNA samples with concentration ranging from 40 to 0.02 ng/μL. The thermal cycling parameters for every quantitation started with a 95 °C incubation for 10 min, and then followed by 40 cycles of 95 °C for 15 s, and 60 °C for 1 min

2.3. PCR multiplex amplification

The primer mix stock was, previously tested and then distributed by the core laboratory to all participants in aliquots of 6.5 μL. All multiplex PCR reactions were performed according to Houston et al. [23] using 0.4 ng of DNA template, the Type-IT® 2X PCR Master Mix (Qiagen) and a primer mix for 13 STRs using four fluorescent dyes (6-FAM®: ANUCS501, 9269, 4910, 5159; VIC®: ANUCS305, 9043; B05, 1528; NED®: 3735 and CS1; PET®: D02, C11 and H06) (Thermo Fisher Scientific). Different PCR amplification instruments were used by each participant laboratory (Table 1). PCR parameters were as follows: activation for 5 min at 95 °C, followed by 30 cycles of 30 s at 95 °C, 90 s at 60 °C, and 30 s at 72 °C, and a final extension of 30 min at 60 °C. All PCR reactions were carried out for one positive control, one negative control and two blind samples in triplicate.

2.4. Genotyping analysis

Capillary electrophoresis (CE) and amplified sample preparation were carried out according to Houston et al. [23]. Either LIZ®500 or LIZ®600 size standard with highly deionized (Hi-Di) formamide were used for CE runs. Running conditions were chosen based on standard HID STR genotyping protocols used by each laboratory. The different genetic analyzers used by each laboratory are displayed in Table 1. Different run parameters (i.e., oven temperature, pre-run condition, injection condition, run condition, capillary length, polymer type) were applied according to each laboratory protocol. For this reason, the majority of laboratories had to adjust their bin sets appropriately. Marker panels and allelic bins files for different versions of Genemapper® software (Thermo Fisher Scientific) were developed and provided by the core lab to each participant laboratory. Indeed, online technical support for the calibration of bins and markers according to different electrophoretic conditions was also provided. Analytical and stochastic thresholds were set according to each laboratory's protocols and interpretation guidelines. Participating laboratories provided a table with genotyping results and the raw data sample files (FSA files) with the printouts of the obtained electropherograms.

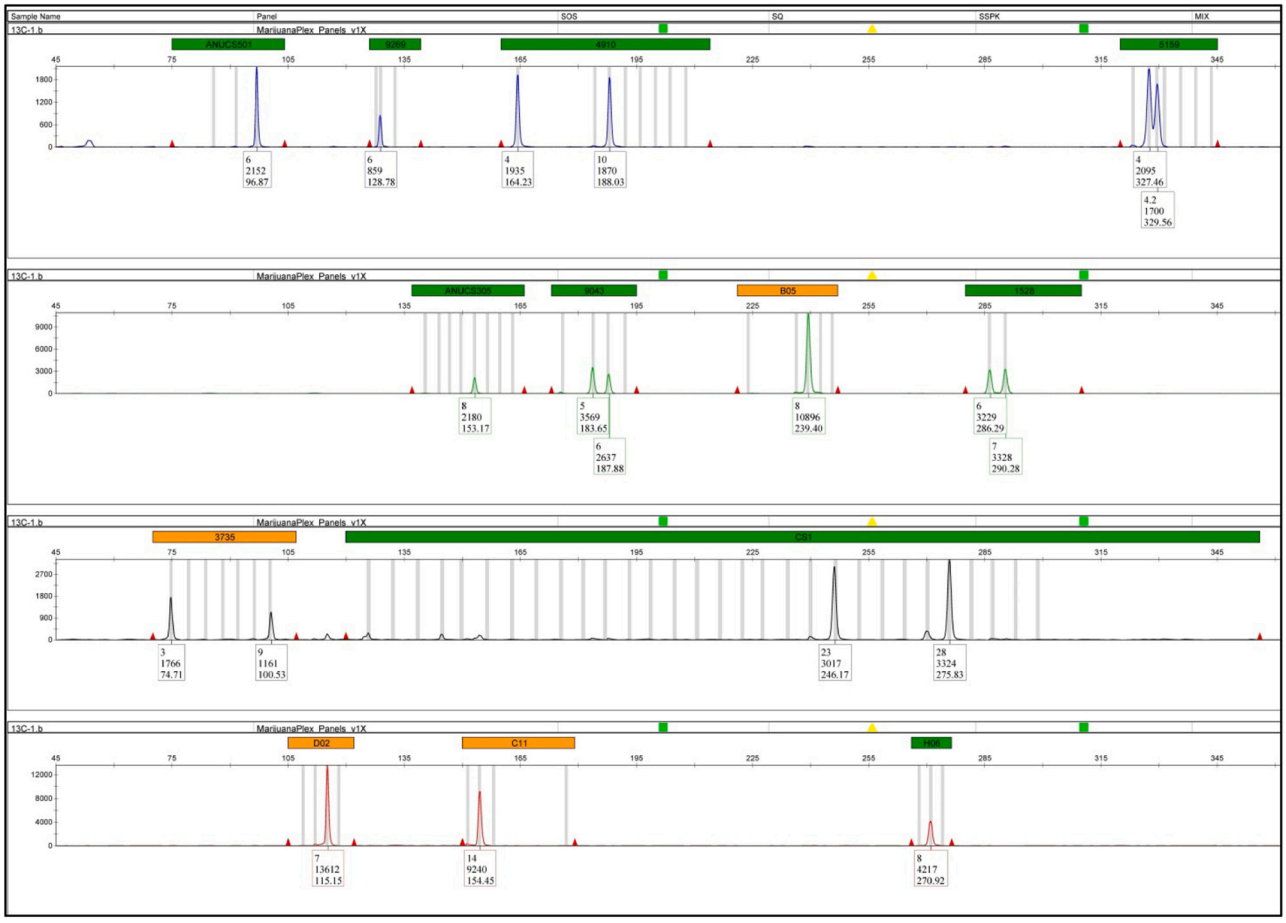

Fig. 3. Example of an electropherogram of analyzed loci by one of the twenty-one participant laboratories. Full concordance was observed with the core lab.

2.5. Analytical parameters determination: heterozygous peak height ratio, inter-loci balance, stutter ratio and analytical threshold

Heterozygous peak height ratio (PHR) was calculated for a given locus by dividing the peak height of an allele with a lower relative fluorescence units (RFU) value by the peak height of an allele with a higher RFU value in a heterozygous pair, and then multiplying this value by 100 to express the PHR as a percentage [24]. For the PHR, the mean, standard deviation (SD), median, minimum and maximum were calculated. The inter-loci balance was calculated as the ratio between the mean peak height for each locus and the mean peak height across all loci multiplied by 100.

Stutter percentage was calculated based on the peak height of the parent allele (stutter peak RFU/parent allele RFU x 100). Stutter peaks that were in-between two alleles on + 1/− 1 stutter position were regarded as −1 stutter of the longer allele. Stutter peak average was determined and the −1/+ 1 stutter mean obtained from each laboratory were estimated. As part of a new forensic kit in-house validation the assessment of −1 and + 1 stutter unit repeat filter application was required. To identify if a peak was a true allele or a stutter, we applied stutter ratio filters for different STR loci. Peaks below that filters were considered stutters [24]. Results from positive control, sample 1 and sample 2, each in triplicate, were utilized to determine the −1 and + 1 stutter ratio thresholds percentage. Each sample replicate was analyzed by the software GeneMapper ID-X (Thermo Fisher Scientific); no stutter filters were applied and a detection threshold was set to 30 RFU. RFUs from 3300 to 8000 data point were collected and analyzed. For the statistical analysis, GraphPad Prism software ver. 9.0.2 (San Diego, CA, USA) was used. Mean of Peak Height (PHLM) and Parent Allele Stutter (PHPASLM), −1 and + 1 Stutter Percentage Mean (MSLM and PSLM) were also calculated. PHLM, PHPASLM, MSLM and PSLM were used to calculate

Table 4
Observed peak height ratios (PHR) mean, median, minimum, and maximum at each heterozygous locus present in the cannabis STR system for sample 1 in triplicate obtained by 20 participating laboratories.

Marker	Observations	Mean PHR	SD	Median PHR	Minimum PHR	Maximum PHR	Inter-Loci balance
4910	20	78%	13%	79,48%	59%	99,97%	101,68%
5159	19	66%	19%	63,29%	27,92%	98,66%	85,93%
9043	20	81%	14%	69,94%	42%	97,83%	106,01%
1528	20	79%	17%	73,05%	47%	99,23%	104,10%
3735	16	77%	12%	79,47%	60%	98,60%	101,19%
CS1	20	77%	19%	70,00%	40,72%	99,29%	101,09%

SD: Standard Deviation; Minimum PHR: Minimum PHR observed value for that locus; Maximum PHR: Maximum PHR observed value for that locus; Median PHR: Min PHR/Max PHR; Mean Peak Height Ratio across all heterozygous loci: Medium value for Heterozygous loci was 76.27% for the sample 1, Inter Loci Balance: Mean PHR/MPH.

Table 5
Observed peak height ratios (PHR) mean, median, minimum, and maximum at each heterozygous locus present in the cannabis STR system for sample 2 in triplicate obtained by 20 participating laboratories.

Marker	Observations	Mean PHR	SD	Median PHR	Minimum PHR	Maximum PHR	Inter-Loci balance
ANUCS501	18	85%	8%	85,06%	71,28%	98,85%	106,47%
9269	17	88%	11%	81,85%	64,62%	99,09%	109,62%
5159	20	60%	17%	65,56%	33,58%	97,54%	74,70%
ANUCS305	19	83%	13%	82,43%	68,61%	96,26%	104,18%
B05	20	85%	11%	81,12%	62,77%	99,46%	106,88%
CS1	20	78%	13%	75,10%	50,28%	99,91%	98,03%
D02	20	84%	11%	75,15%	50,81%	99,49%	104,78%
C11	20	80%	13%	73,23%	47,03%	99,43%	100,26%
H06	20	76%	12%	72,54%	48,80%	96,29%	95,06%

SD: Standard Deviation; Minimum PHR: Minimum PHR observed value for that locus; Maximum PHR: Maximum PHR observed value for that locus; Median PHR: Min PHR/Max PHR; Mean Peak Height across all heterozygous loci: Medium value for Heterozygous loci was 79.87% for the sample 2, Inter Loci Balance: Mean PHR/MPH.

the Average of laboratories means and peak height (ALMPH) and the standard error of the mean (SEM) for all laboratories. The minimum peak height of parental allele with a stutter (MPHPAS) was also estimated.

To determine the limits of detection for a given analytical procedure, it is necessary to determine a Minimum Distinguishable Signal (MDS), the signal at which a peak can reliably be distinguished from noise. The MDS may be considered a relative fluorescence unit (RFU) or an analytical threshold (AT) for forensic purposes [25]. Furthermore, the detection limit at 99.7% confidence is based on the Gaussian distribution of noise peaks height and this does not consider a possible asymmetric distribution where, the correct 99.87% confidence should be applied. The background noise level detected of each instrument were compared, to evaluate their impact on the general background noise. Three negative samples (diH_2O) were amplified, run through capillary electrophoresis and analyzed by setting the GeneMapper™ software at 1 RFU as detection threshold. RFU data from 3300 to 8000 data point were collected. The GeneMapper data were exported to a txt file and then imported to an excel tool or GraphPad Prism software v. 9.0.2 (GraphPad software Inc., San Diego, CA) for the analysis. The background instrument peak heights (RFU) observed in each dye channel of each laboratory were compared. The following parameters were calculated for each laboratory: Maximum Peak Height (MPH), Average Peak Height (APH), Standard Deviation (SD), Limit of Detection (LOD), Limit of Quantitation (LOQ), Analytical Threshold (AT). APH and AT means were used for each dye channel to calculate the Standard Error of the Mean (SEM) for each laboratory.

3. Results and discussion

3.1. STR typing results

The percentage of genotyping success was estimated to be 96% (Suppl. Table 1 and Suppl. Table 2). Genotyping errors may be due to partial DNA degradation (during shipping and storage), amplification errors, stutter calculation errors, bin error and PCR artifacts such as allelic drop-out. All of these interpretation errors and artifacts can be resolved with an optimization of this methodology through an internal validation and subsequent interpretation guidelines. A locus was labeled discordant if one or more alleles were miscalled, for example, ANUCS305 locus resulted in about 10% of incorrect allele calls for both samples (Fig. 1). A locus was labeled concordant when all alleles were correctly called, for example 4910, 9043, B05, 1528, CS1, D02, C11 and H06 loci gave a 100% of consensus allele calls among the participant laboratories (Fig. 1). For sample 1, laboratories 2, 6, 7, 9, 11 and 13 reported incorrect allele call for a heterozygous STR loci: 9269 (laboratory 2), 5159 (laboratories 11 and 13), ANUCS305 (laboratory 7) and 3735 (laboratories 2, 7 and 9).

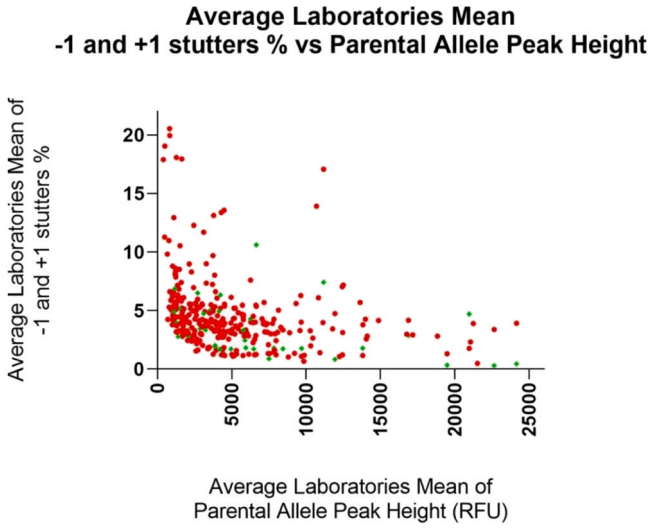

Fig. 4. Distribution of −1/+1 Stutters. Dots represent the average of each laboratory mean values of −1 / +1 stutter percentage of compared to the height of the relative allele (true peak). The% values of the stutters −1 are shown in red. The% values of the stutters +1 are shown in green. For interpretation of the references to colour in this figure, the reader is referred to the web version of this article

Table 6
Observed peak height ratios (PHR) mean, median, minimum, and maximum at each heterozygous locus present in the cannabis STR system for positive control in triplicate obtained by 20 participating laboratories.

Marker	Observations	Mean PHR	SD	Median PHR	Minimum PHR	Maximum PHR	Inter-Loci balance
ANUCS501	20	82%	12%	79,19%	61,00%	97,38%	110,77%
4910	20	79%	14%	77,38%	57%	98,14%	106,05%
ANUCS305	15	59%	23%	56,04%	17,97%	94,11%	79,63%
9043	20	82%	12%	78,69%	59%	98,70%	110,88%
B05	20	86%	11%	83,51%	67,18%	99,85%	115,71%
CS1	20	38%	8%	37,39%	19,67%	55,10%	50,69%
C11	20	85%	12%	80,50%	61,01%	99,99%	114,09%
H06	20	83%	9%	84,08%	69,18%	98,97%	112,18%

SD: Standard Deviation; Minimum PHR: Minimum PHR observed value for that locus; Maximum PHR: Maximum PHR observed value for that locus; Median PHR: Min PHR/Max PHR; Mean Peak Height across all heterozygous loci: Medium value for Heterozygous loci was 74.15% for the positive control, Inter Loci Balance: Mean PHR/MPH.

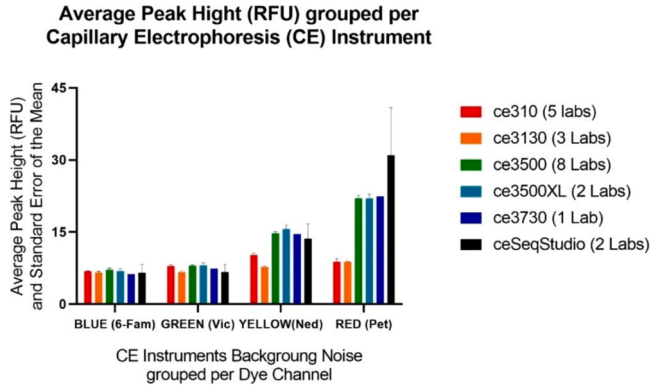

Fig. 5. Background-noise registered by different genetic analyzer instruments. Error bars represent standard error.

Moreover, laboratory 6 did not produced results for three loci: 9269, ANUCS305 and 3735. Laboratory 13 did not produce any result for loci ANUCS501 and 9269 (Table 2). For sample 2, laboratories 3 and 13 encountered a problem with incorrect allele call of the 9259 heterozygous STR locus. Moreover, laboratories 5 and 6 did not produced results on loci: ANUCS501, 9269 and ANUCS305. Indeed, laboratory 6 did not produce any result on locus 3735 (Table 3). The calculations of percentage of success were based on a total of 13 STR loci. Examples of typing errors are displayed in Fig. 2.

The negative controls did not show any evidence of external contamination. In the case when a laboratory failed to produce results for a specific STR locus or if only one allele at a heterozygous locus was obtained, it was considered an error and not a partial result for the purpose of the study. Results with concordant calls with those of the organizers were considered correct (Table 2 and Table 3). In summary, twenty laboratories submitted results for cannabis STRs. Eleven laboratories obtained full and concordant profiles for the two samples without errors, two laboratories obtained full and concordant profiles with one error, four laboratories completed the exercise with two errors, one laboratory committed three errors, while two laboratories completed the test with more than three errors. The success rates for cannabis STR typing ranged from 74.4% to 100% for sample 1 and from 69.2% to 100% for sample 2. Genotyping success calculations were based on the total number of loci tested by a laboratory in this study. An example of an electropherogram of this study is displayed in Fig. 3.

3.2. Peak height ratio

Peak heights for the heterozygous markers were averaged and the intra locus peak height ratio (PHR) was calculated. The final average PHR was calculate by the mean PHR for each heterozygous locus of the three samples, and resulted to be 78.2% (Table 4, Table 5 and Table 6). The mean inter-loci balance was estimated to be 100,00%.

3.3. Stutter ratio

Stutter statistics were calculated in triplicate for the positive control, sample 1 and sample 2, respectively (Suppl. Tables 3, 4 and 5). Stutters were observed when parent alleles were above 400 RFU heights. The core laboratory identified 116 "−1 stutters" and 23 "+ 1 stutters" from sample 1; for sample 2, the number of stutter were 100 and 26, respectively, and for the positive control 82 "−1 stutters" and 15 "+ 1 stutters" were observed. The stutter ratio was higher when the parent allele height increased. In this condition, the probability to observe the + 1 stutter allele was higher than to observe the −1 stutter allele. Smaller alleles lost their stutters, as in the case of alleles 4 and 6 in ANUCS501 locus for sample 1, sample 2 and positive control. Moreover, stutters seem to be specific for B05, CS1, D02 and C11 loci. Stutters were independent of the CE platform used by participating laboratories. In this study it was shown that the stutter percentage assumes a non-homogeneous distribution, due to the small amount of data analyzed. The −1/+ 1 stutter percentage raised when the parent peak height was lower than 800 RFU and it decreased when the parent peak height was higher than 1500 RFU. The lowest amount of − 1 / + 1 stutter percentage was produced when the height of the parent allele was higher than 8000 RFU with a percentage value less than 10 as shown in Fig. 4.

3.4. Analytical threshold study

The detection limit obtained by GraphPad Prism software v. 9.0.2 (GraphPad software Inc., San Diego, CA) at 99.7% confidence was based on the Gaussian distribution of noise peaks height and this did not consider a possible asymmetric distribution where, the correct 99.87% confidence should be applied. Average peak height and Analytical Threshold were smaller in the blue (6-FAM™) and in the green (VIC™) dye channel when compared to the yellow (NED™) and red (PET™) channels. Noise increased from blue dye channel to red dye channel. Background-noise data, listed by capillary electrophoresis platform, from each participant laboratory was reported in Suppl. Table 6 and in Fig. 5.

4. Conclusions

The present collaborative study of the Ge.F.I.-ISFG working group allowed to demonstrate the robustness and reproducibility of a 13-locus cannabis STR multiplex system for forensic DNA profiling. Overall, each participant, using various instrumentation, polymers, size standard and arrays, generated STR profiles for each of the 13 markers. Analysis of the electropherograms, by the core laboratory, showed a concordance of more than 96%. Data misinterpretation resulted in discordance in some markers; typing was more problematic with the analysis of the ANUCS501, 9269, ANUCS305 and 3735 loci. This problem will require further investigation and all of these discordant allele calls can be resolved with, an optimized bin set, more training and experience with this multiplex STR kit.

An important limitation of this study is that the analysis was conducted with cannabis DNA extracts instead of cannabis plant tissue. An important source of variation may occur during DNA extraction of different cannabis plant material. It is important to note that this STR kit cannot differentiate hemp from drug type; therefore, chloroplast DNA markers should be used for this purpose [26,27]. However, using STRs, cannabis seizures can be genetically associated through phylogenetic analysis of a previously established database. Lastly, a plant generated from clonal propagation can be genetically associated to its clones using this STR kit.

Future plans include: a) a collaboration between laboratories to test this kit with a new synthetic allelic ladder (in progress), to increase allele coverage and to aid in more accurate allele calls, b) studies of variation based on different extraction methods, types of tissue and various storage conditions and c) comparison of different databases. A written consensus standard for *C. sativa* authentication would be useful for the forensic community to establish rules and interpretation guidelines. Points to be considered for a future consensus standard include: allelic ladder, stutter filter recommendations, DNA quantitation methods, a comparative *C. sativa* STRs database, and troubleshooting. Finally, this inter-laboratory exercise can be considered a milestone in the identification of *C. sativa* samples.

CRediT authorship contribution statement

M. Di Nunzio: Methodology, Writing – original draft preparation, Writing – reviewing & editing, Investigation. **V. Agostini:** Investigation. **F. Alessandrini:** Investigation. **C. Barrot-Feixat:** Investigation. **A. Berti:** Investigation. **C. Bini:** Investigation. **M. Bottinelli:** Investigation. **E. Carnevali:** Investigation. **B. Corradini:** Investigation. **M. Fabbri:** Investigation. **P. Fattorini:** Investigation. **P. Garofano:** Investigation. **S. Gino:** Investigation. **A. Mameli:** Investigation. **A. Marino:** Investigation. **C. Previderè:** Investigation. **C. Robino:** Investigation. **C. Romano:** Investigation. **P. Tozzo:** Investigation. **A. Verzeletti:** Investigation. **L. Buscemi:** Investigation. **D. Gangitano:** Conceptualization, Writing – reviewing & editing. **C. Di Nunzio:** Methodology, Formal analysis, Investigation.

Conflict of Interest

None.

Acknowledgements

We would like to thank Doctor Albert Vingut López, director in charge at that time of the National Institute of Toxicology and Forensic Sciences (INTCF) Barcelonian Department (Spain), and Doctor Rachel Houston, from Sam Houston State University (Texas, USA), to help with this collaborative exercise providing reagents and cannabis DNA samples.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.forsciint.2021.111053.

References

[1] European Monitoring Center for Drugs and Drug Addiction (EMCDDA). European Drug Report., 2016. www.emcdda.europa.eu. (Accessed 23 September 2020).
[2] M. della S. e M della Giustizia, D.M. 11 aprile 2006, Italy, 2006 https://www.gazzettaufficiale.it/eli/id/2006/04/24/06A04031/sg (Accessed 23 September 2020).
[3] LEGGE 2 dicembre 2016 , n. 242, 2016. https://www.gazzettaufficiale.it/atto/stampa/serie_generale/originario (Accessed 30 April 2020).
[4] Scientific Working Group for the Analysis Of Seized Drugs (SWGDRUG) Recommendations recommendations include: code of professional practice education and training methods of analysis quality assurance united states department of justice drug enforcement administration executive office of the president office of national drug control policy counterdrug technology assessment center , 2019. www.swgdrug.org (Accessed 23 September 2021).
[5] A. Leghissa, Z.L. Hildenbrand, K.A. Schug, A review of methods for the chemical characterization of cannabis natural products, J. Sep. Sci. 41 (2018) 398–415, https://doi.org/10.1002/jssc.201701003
[6] G.T. Mitosinka, J.I. Thornton, T.L. Hayes, The examination of cystolithic hairs of Cannabis and other plants by means of the scanning electron microscope, J. Forensic Sci. Soc. 12 (1972) 521–529, https://doi.org/10.1016/S0015-7368(72)70717-3
[7] A.D. Jacobs, R.R. Steiner, Detection of the Duquenois-Levine chromophore in a marijuana sample, Forensic Sci. Int. 239 (2014) 1–5, https://doi.org/10.1016/j.forsciint.2014.02.031
[8] R. Gillan, M.D. Cole, A. Linacre, J.W. Thorpe, N.D. Watson, Comparison of Cannabis sativa by Random Amplification of Polymorphic DNA (RAPD) and HPLC of cannabinoids: a preliminary study, Sci. Justice J. Forensic Sci. Soc. 35 (1995) 169–177, https://doi.org/10.1016/S1355-0306(95)72658-2
[9] M. Kojoma, O. Iida, Y. Makino, S. Sekita, M. Satake, DNA fingerprinting of Cannabis sativa using Inter-Simple Sequence Repeat (ISSR) amplification, Planta Medica 68 (2002) 60–63, https://doi.org/10.1055/s-2002-19875
[10] R. Houston, M. Birck, S. Hughes-Stamm, D. Gangitano, Evaluation of a 13-loci STR multiplex system for Cannabis sativa genetic identification, Int. J. Leg. Med. 130 (2016) 635–647, https://doi.org/10.1007/s00414-015-1296-x
[11] H.M. Coyle, G. Shutler, S. Abrams, J. Hanniman, S. Neylon, C. Ladd, T. Palmbach, H.C. Lee, A simple DNA extraction method for marijuana samples used in Amplified Fragment Length Polymorphism (AFLP) analysis, J. Forensic Sci. 48 (2003) 2001207, https://doi.org/10.1520/jfs2001207
[12] H.J. Alghanim, J.R. Almirall, Development of microsatellite markers in Cannabis sativa for DNA typing and genetic relatedness analyses, Anal. Bioanal. Chem. 376 (2003) 1225–1233, https://doi.org/10.1007/s00216-003-1984-0
[13] S. Gilmore, R. Peakall, Isolation of microsatellite markers in Cannabis sativa L. (marijuana), Mol. Ecol. Notes 3 (2003) 105–107, https://doi.org/10.1046/j.1471-8286.2003.00367.x
[14] S. Gilmore, R. Peakall, J. Robertson, Organelle DNA haplotypes reflect crop-use characteristics and geographic origins of Cannabis sativa, Forensic Sci. Int. 172 (2007) 179–190, https://doi.org/10.1016/j.forsciint.2006.10.025
[15] C. Howard, S. Gilmore, J. Robertson, R. Peakall, Developmental validation of a Cannabis sativa STR multiplex system for forensic analysis, J. Forensic Sci. 53 (2008) 1061–1067, https://doi.org/10.1111/j.1556-4029.2008.00792.x
[16] C. Howard, S. Gilmore, J. Robertson, R. Peakall, A Cannabis sativa STR genotype database for Australian Seizures: forensic applications and limitations, J. Forensic Sci. 54 (2009) 556–563, https://doi.org/10.1111/j.1556-4029.2009.01014.x
[17] M.A. Mendoza, D.K. Mills, H. Lata, S. Chandra, M.A. Elsohly, J.R. Almirall, Genetic individualization of Cannabis sativa by a short tandem repeat multiplex system, Anal. Bioanal. Chem. 393 (2008), https://doi.org/10.1007/s00216-008-2500-3
[18] S. Köhnemann, J. Nedele, D. Schwotzer, J. Morzfeld, H. Pfeiffer, The validation of a 15 STR multiplex PCR for Cannabis species, Int. J. Leg. Med. 126 (2012) 601–606, https://doi.org/10.1007/s00414-012-0706-6
[19] A. Linacre , L. Gusmão , W. Hecht , A.P. Hellmann , W.R. Mayr , W. Parson , M. Prinz , P.M. Schneider , N. Morling , ISFG: Recommendations regarding the use of non-human (animal) DNA in forensic genetic investigations, 2011. https://doi.org/10.1016/j.fsigen.2010.10.017.
[20] Scientific Working Group on DNA Analysis Methods Validation Guidelines for DNA Analysis Methods SWGDAM Validation Guidelines for DNA Analysis Methods, (n.d.).
[21] J.L. Almeida, A. Dakic, K. Kindig, M. Kone, D.L.D. Letham, S. Langdon, R. Peat, J. Holding-Pillai, E.M. Hall, M. Ladd, M.D. Shaffer, H. Berg, J. Li, G. Wigger, S. Lund, C.R. Steffen, B.B. Fransway, B. Geraghty, M. Natoli, B. Bauer, S.M. Gollin, D.W. Lewis, Y. Reid, Interlaboratory study to validate a STR profiling method for intraspecies identification of mouse cell lines, PLoS One 14 (2019) 1–25, https://doi.org/10.1371/journal.pone.0218412
[22] D.Y. Wang, C.W. Chang, R.E. Lagacé, L.M. Calandro, L.K. Hennessy, Developmental validation of the AmpF℄STR ® identifiler ® plus PCR amplification kit: an established multiplex assay with improved performance, J. Forensic Sci. 57 (2012) 453–465, https://doi.org/10.1111/j.1556-4029.2011.01963.x
[23] R. Houston, M. Birck, S. Hughes-Stamm, D. Gangitano, Developmental and internal validation of a novel 13 loci STR multiplex method for Cannabis sativa DNA profiling, Leg. Med. 26 (2017) 33–40, https://doi.org/10.1016/j.legalmed.2017.03.001
[24] A.A. Westen, L.J.W. Grol, J. Harteveld, A.S. Matai, P. De Knijff, T. Sijen, Assessment of the stochastic threshold, back- and forward stutter filters and low template techniques for NGM, Forensic Sci. Int. Genet. 6 (2012) 708–715, https://doi.org/10.1016/j.fsigen.2012.05.001
[25] J. Bregu, D. Conklin, E. Coronado, M. Terrill, R.W. Cotton, C.M. Grgicak, Analytical thresholds and sensitivity: establishing RFU thresholds for forensic DNA analysis, J. Forensic Sci. 58 (2013) 120–129, https://doi.org/10.1111/1556-4029.12008
[26] R. Houston, M. Birck, B. LaRue, S. Hughes-Stamm, D. Gangitano, Nuclear, chloroplast, and mitochondrial data of a US cannabis DNA database, Int. J. Leg. Med. 132 (2018) 713–725, https://doi.org/10.1007/s00414-018-1798-4
[27] M.G. Roman, D. Gangitano, R. Houston, Characterization of new chloroplast markers to determine biogeographical origin and crop type of Cannabis sativa, Int. J. Leg. Med. 133 (2019) 1721–1732, https://doi.org/10.1007/s00414-019-02142-w