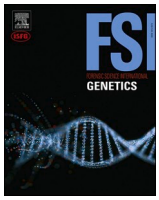
Forensic Science International: Genetics 68 (2024) 102971

Contents lists available at ScienceDirect

Forensic Science International: Genetics

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



Characterization and evaluation of nine *Cannabis sativa* chloroplast SNP markers for crop type determination and biogeographical origin on European samples

Michele Di Nunzio [a,*,1], Carme Barrot-Feixat [a], David Gangitano [b]

[a] Forensic Genetics Laboratory - Legal Medicine Unit, Department of Medicine, University of Barcelona, Spain
[b] Forensic & Legal Medicine Area, Department of Sociosanitary Sciences, School of Medicine, University of Murcia, Murcia, Spain

ARTICLE INFO

Keywords:
Forensic Plant Science - *Cannabis sativa* – Chloroplast DNA – Barcoding DNA – Single nucleotide polymorphisms

ABSTRACT

Cannabis sativa can be classified in two main types, according to psychotropic cannabinoid Δ9-tetrahydrocannabinol (Δ9-THC) content: the drug-type and the fiber-type. According to the European Monitoring Center for Drugs and Drug Addiction, most of the European Union countries consider the possession of cannabis, for personal use, a minor offense with possibility of incarceration. Despite of the model of legal supply (i.e., Spanish cannabis clubs, Netherlands coffee shops) or medical use (i.e., Italy), cannabis remains the most used and trafficked illicit plant in the European Union. Differentiating cannabis crops or tracing the biogeographical origin is crucial for law enforcement purposes. Chloroplast DNA (cpDNA) markers may assist to determine biogeographic origin and to differentiate hemp from marijuana. This research aims: to identify and to evaluate nine *C. sativa* cpDNA polymorphic SNP sites to differentiate crop type and to provide information about its biogeographical origin. Five SNaPshot™ assays for nine chloroplast markers were developed and conducted in marijuana samples seized in Chile, the USA-Mexico border and Spain, and hemp samples grown in Spain and in Italy. The SNapShot™ assays were tested on 122 cannabis samples, which included 16 blind samples, and were able to differentiate marijuana crop type from hemp crop type in all samples. Using phylogenetic analysis, genetic differences were observed between marijuana and hemp samples. Moreover, principal component analysis (PCA) supported the relationship among hemp samples, as well as for USA-Mexico border, Spanish, and Chilean marijuana samples. Genetic differences between groups based on the biogeographical origin and their crop type were observed. Increasing the number of genetic markers, including the most recently studied ones, and expanding the sample database will provide more accurate information about crop differentiation and biogeographical origin.

1. Introduction

Cannabis sativa is the most commonly consumed illicit drug in Europe, and also one of the numerous substances associated with drug-use offenses on the continent [1]. Despite the law offenses related to this drug rise, the rates of consumption, size of the illegal market, and policy ideas to cannabis vary noticeably across countries in Europe. Since the beginning of the century, the supply and use of cannabis for medical, industrial, or recreational purposes has been legalized in many non-European countries as in some states of the USA [2], Canada [3], New Zealand [4] and Uruguay [5]. In Europe, notable changes are also taking place in cannabis policy. The Nederlands with their "closed coffeeshop supply circuit" [1] were the forerunners, to countries' entities such as Spain's cannabis club [6], and Switzerland's pilot trials of legal cannabis sales in early 2023 [7].

Over the last ten years, the European Union has seen considerable cultivation of cannabis for industrial purposes. Indeed, it is legal to supply and cultivate authorized *C. sativa* plants for fiber and seeds if they have a low level of Δ9-tetrahydrocannabinol. These plant varieties, known as hemp, are laid down in the Council Directive 2002/57/EC [8]. Farmers can use only certified seeds of specified hemp varieties, with a THC content not exceeding 0.3 % [9]. Otherwise, plants with higher than 0.3 % levels of THC are illegal, and considered drugs. The cannabis drug type, also known as marijuana, is used for its psychotropic effect

* Corresponding author.
E-mail address: michele.dinunzio@ub.edu (M. Di Nunzio).
[1] ORCID: 0000–0001-6707–1960

https://doi.org/10.1016/j.fsigen.2023.102971
Received 1 August 2023; Received in revised form 10 October 2023; Accepted 20 October 2023
Available online 24 October 2023

[10–13] and for different medical purposes [14,15]. On the other hand, hemp, the non-drug strain, is used in fiber and food industries [16], as well as for oil production and skin cream treatments [17]. Therefore, it is imperative for the police to distinguish between illicit marijuana and legal hemp. Nowadays, police routine cannabis identification methods include cystolithic hair microscopic observation on plant leaves and buds [18], and the THC detection by most common cannabinoid quantification techniques such as gas chromatography/mass spectrometry (GC/MS) and high-performance liquid chromatography (HPLC) [19].

Besides chemical and morphological tests, molecular plant evidence analysis can be useful in criminal investigations. During the last 20 years, forensic science has been taking steps towards cannabis genetic identification [20–28]. Since cannabis acquired a fundamental role in forensic science, methodological discrimination between hemp and marijuana [29] was emphasized on the applications of many molecular biology techniques (i.e., whole plastome sequences, DNA chromatography, DNA markers, etc.) to identify and characterize cannabis samples.

The detection and the identification of cannabis plants, cannabis commercial seeds, and medical cannabis by using the high-resolution melting (HRM) analysis that was firstly suggested by Cowan and Elkins, and recently by Solano et al. and Anabalón et al. [30–32]. Alternatively, Yamamuro et al. established a simple and accurate cannabis DNA detection system using DNA chromatography [33]. The two-chromatography chip system was able to detect cannabis samples by PCR amplification of tetrahydrocannabinolic acid synthase gene sequence. Another DNA-based protocol for *C. sativa* identification was proposed by Kitamura et al. [34]. Cannabis samples, such as seed and resin, were successfully identified using the loop-mediated isothermal amplification (LAMP) assay [34]. Furthermore, as a lot of cannabis hybrids are trafficked, Matielo et al. proposed the analysis of whole cannabis plastome as identification tool of marijuana hybrids. However, due to the absence of studies and sequences deposited in the GenBank® database, this tool is still not able to determine the parental origin of some hybrids [35]. Another suitable method described of cannabis distinction was based on the chemotype [36]. Borroto et al. (2020) presented a genetic prediction system, based on the ratio of the THC and CBD, of 62 agricultural hemp cultivars [36].

The forensic botanical evidence is usually applied to link individuals or objects to a geographic area or to place a suspect in a crime scene. Modern plant molecular techniques can also determine where a plant specimen came from, or can identify a specific strain. Studies on *Cannabis sativa* are having an increased impact in forensic and molecular plant science. For example, groups of seized cannabis plants, genotyped with short tandem repeats (STRs), were associated using phylogenetic analysis [37–41].

Most markers for hemp vs marijuana differentiation were focused on the THC synthesis gene or genes that are involved in the THC/CBD chemotypes [42–47], though it has been demonstrated that the distinction between hemp and marijuana at a genome-wide level is not just limited to these genes [48]. To identify these differences, a DNA barcoding technique would be necessary. DNA barcoding is a molecular technique that uses specific regions of DNA to identify differences between species or populations. With this tool, inter- and intra-species genetic differences may be detected. For these reasons, interest has increased in forensic botany, in forensic genetics, and entomology for species identification [49–52]. Both organelle genomes, mitochondrial and chloroplast, are maternally inherited in *Cannabis sativa* [53]; but due to its gene content, conserved structure, and faster mutation rates [54], chloroplast DNA is the mostly appropriate for barcoding markers.

Moreover, as Gilmore *et al.* (2007) states [55], the chloroplast DNA haplotypes reflect crop characteristic and geographic origin of cannabis samples, definitively assisting in determination of biogeographical origin and crop type differentiation [56]. For this reason, plant barcoding research has been focused on the chloroplast genome (cpDNA) with the potential to identify *C. sativa* at the species level [27,57–61]. Several primer sets for plant DNA barcoding studies have been developed in the last thirty years [49,51,62–64] with the aim of characterizing plants species (analyzing inter-specific differences) or to distinguish between same species populations (analyzing intra-specific differences). Taberlet et al. [62], used six primers for the amplification of three chloroplast DNA markers (*trnT, trnL and trnF*) by PCR. The study was focused on plant population biology and species differences. Their three pairs of universal primers worked on most plant species including bryophytes, gymnosperms, and angiosperms demonstrating their usefulness for evolutionary studies at higher taxonomic levels. Demesure et al. [49] utilized eight cpDNA sequences of the highly conserved tRNA genes (*trnH, trnK, trnC, trnD, trnS, trnM, psaA and psbC*) to study oak trees belonging to the same species of the European *Quercus robur*. Santos and Pereira [51] evaluated the conservation degree of four chloroplast DNA regions (*atpF-atpH*, *psbA-trnH*, *trnL CD* and *trnL GH*) to detect differences in seven plant families (Asteraceae, Brassicaceae, Iridaceae, Orchidaceae, Poaceae, Rosaceae and Salicaceae).

C. sativa plastid genome barcoding regions have shown promising results in differentiating crop type and biogeographic origin [55,56,59–61]. Chloroplast and mitochondrial markers were investigated to identify and differentiate cannabis plants. For this purposes, a multi locus system of five "variable length homopolymer regions" (*rbcL-orf106, ccmp2, ccmp6* and *nad4*) was amplified by simple PCR assays [55]. The system was then modified and optimized to genotype eight markers located in seven cannabis organelles: five STR markers (*cscp001, cscp002, cscp003, cscp004* and *csmt001*) and three SNP markers (*cscp001*, *cscp005* and *csmt002),* both chloroplast and mitochondrial [56]. In recent years, attention has focused on chloroplast SNPs (single nucleotide polymorphisms) and the SNaPshot™ mini-sequencing technique. To date, ten SNPs located in four cpDNA hotspot regions (*rpl32-trnL, trnS-trnG*, *rps16* and *accD-psa*) have been evaluated. These SNPs were coupled with other types of markers (INDEL and hSTR) to improve crop discriminatory power and biogeographical origin information [59–61]. In this work, with the purpose of discriminating and providing origin information on cannabis samples, the SNaPshot™ technique was used to detect nine cpDNA SNPs: *trnH-psbA*, *psaB*, *petA-psbJ*, *psbE-petL*, *rps11*, *rpl36-rps8*, *ndhF*, *ndhD* and *rps15-ycf11*.

The NADH dehydrogenase subunit F gene (*ndhF*) was previously studied by Olmstead *et al.* and by Shannon *et al.* to demonstrate ordinal and familial relationships among plants [65,66]. Genes of the intergenic spacers *trnH-psbA* and *rpl36-rps8* were used by Wong *et al.* (2013) to differentiate medicinal Gentiana species and their adulterants [67]. The NADH dehydrogenase subunit D (*ndhD*), *petA*, *psbJ*, *rps15*, *ycf1*, *psbE* and *petL* genes were described by Shaw *et al.* (2014) to explain the utility of noncoding chloroplast DNA to differentiate angiosperms at the species level [68]. Sytsma et al. (2002), used the *ndhF* gene sequence to build a single most parsimonious tree based of 28 urticalean rosids taxas, including Cannabaceae [69]. The ribosomal protein 8 and 11 genes (*rps8* and *rps11*, respectively) were tested by Zhang et al. (2018) to determine pattern of day-length adaptation and latitude-distribution in *C. sativa* [70]. The abovementioned studies' results on plant genes and non-coding region were critical for our research aims because their capability to demonstrate ordinal and familiar relationship, differentiate plant used for medical use and for their geographical distribution. As Coyle et al. anticipated [71], the brand-new genetic techniques being developed for marijuana detection may allow for the identification of a geographic source to aid in the investigation of major marijuana growers and distributors. The objectives of this study were to examine the crop type and biogeographical discriminating power of the nine hotspot regions, which have nine SNP polymorphisms, and using them to explore genetic variability for the first time within species in *C. sativa* for forensic purposes.

This paper describes the development of five SNaPshot™ assays to genotype nine polymorphisms. Samples consisted of marijuana grown in Chile, Spain, and USA-Mexico border and hemp samples grown in Italy and Spain. Thus, by adding samples from Italy and Spain, this project

seeks to expand the geographic range of previous cannabis researches [55,56,59–61]. Furthermore, comprehensive haplotype analysis using genotype data from nine polymorphic markers was performed on a dataset of 106 samples representing our five different populations. Detecting country differences, or even better regional differences, in these hotspot DNA sequences could assist police investigation in determining whether samples represent legal or illegal cannabis, providing evidence to associate cannabis samples to each other, and aid in marijuana trafficking spotting into and within a country. Nowadays, the inability for law enforcement to unmistakably distinguish between hemp and marijuana is problematic in all marijuana prosecutions, from small amounts to trafficking substantial amounts of plant material. However, a tool that helps to differentiate cannabis legal crops from illegal crops would be critical to identifying safer cannabis products or to aiding police involved in a drug trafficking investigation to determine if the item being seized is really evidence of a crime.

2. Materials and methods

2.1. In silico analysis of C. sativa chloroplast published genomes

Five published chloroplast genome sequences of *C. sativa* cultivars were chosen for alignment, by using the NCBI GenBank® database: Dagestani (KR77995), Cheungsam (KR184827), Carmagnola (NC_026562), Yoruba Nigeria (NC_027223) and Yunma 7 (MW013540). Sequences were chosen to represent different cannabis varieties both for marijuana and hemp. FASTA file were imported in the Geneious Prime® software v.2022.0.1 (Biomatters, New Zealand) and aligned using the Mauve genome alignment tool to detect differences between sequences. Genic and inter-genic sequences with differences in bases were targeted as "hotspot" regions. Nine of these "hotspot regions", previously used to differentiate plants at the species level [65–69], containing a polymorphic SNP site were considered for the study: *trnH-psbA*, *psaB*, *petA-psbJ*, *psbE-petL*, *rps11*, *rpl36-rps8*, *ndhF*, *ndhD* and *rps15-ycf11* (Table 1).

2.2. Sample collection

Cannabis samples (N = 106) of marijuana and hemp from different sources were used for this research: Marijuana plant material grown in different cities from Spain (N = 50); Chilean Marijuana DNA extracts (N = 10), originated from several plants grown in various sites of the La Araucania region; Marijuana DNA extracts, coming from plants seized by U.S. Customs & Border Protection (U. S. Department of Homeland Security) at different sites of the USA-Mexican border (N = 6); Italian Hemp seeds (N = 20), coming from plants grown in the Le Marche region, were purchased online; and Spanish Hemp seeds (N = 20), coming from plants grown in the Xàtiva surrounding area (Valencia province). Spanish THC-positive plant material was obtained from the National Institute of Toxicology and Forensic Sciences (INTCF) - Barcelona Department (Spain) and Spanish marijuana seed shops. Hemp seeds were purchased from two companies: Alimento Completo Canarios (Xàtiva, VA, SP), and CBWeed (Forlí, FO, IT). A further 16 blind samples were analyzed. Marijuana seeds (N = 8) were purchased in two different regular seed shops in Barcelona. Spanish hemp seeds (N = 4), produced in the Valencia region, were taken from a bird-seed mix regularly sold in a Barcelona grocery store; while Italian hemp seeds (N = 4) were purchased from an online food market.

Table 1
Forward and Reverse primer list. Primers were used for the nine *C. sativa* chloroplast hot-spot regions amplification and sequencing. "Amplicon length bp" indicate the length in bp of the produced amplicon. "Assay" indicates in which of the five assays the cpDNA region was examined.

cpDNA Region		Primer name and sequences 5' – 3'	Amplicon length bp	Assay
ndhF	F	ACCACGATTATATGCCCAATC	132	1
	R	CAGTCAGTATAGCTTCGTTGGGA		
ndhD	F	ACCCCGTCAACAAAATGGGT	204	1
	R	TCCAATATTTGCGGGTTCCCT		
psaB	F	CTCCTTGGGAGGTGCCAAAT	162	2
	R	TTCGGAGTAAGCTCCTTGGC		
petA-psbJ	F	GTCCGATTTCAATGCCATGC	211	2
	R	TCGACACAAGAAAGGGTTGTG		
rpl36-rps8	F	ACGAACGGAGGCTCTGATTT	230	3
	R	CGACTAGAAGGAATCGGCGG		
psbE-petL	F	CGTTCTCCTGTGCTTCCAGA	163	3
	R	GGGAAATTTCGTTGCTGTGTCA		
trnH-psbA	F	CTTGATCCACTTGGCTACATCC	340	3
	R	CTTAGCTGCTCTAGAAGTTCC		
rps15-ycf11	F	AAGAAGGTTTCCGCACTCAT	121	4
	R	TCCATTCCATTCGGCTTGAT		
rps11	F	ACCAAGAAATTACCCGCCCC	240	5
	R	TCGAAAGGCCCTACTGTTCA		

2.3. Plant DNA extraction

Cannabis plant tissue fragments (leaf, stem, flower, or seed) were sampled. For each fragment, 10–20 mg was cut with a sterile blade, except for seeds that were processed entirely. Then, each fragment or seed was transferred, with a sterile tong, to a DWK Life Sciences Kontes™ Pellet Pestle™ (Fisher Scientific SL, Madrid, Spain) and finally was ground up by liquid nitrogen disruption. For each fragment, DNA extraction was performed using the DNeasy® Plant Mini Kit (Qiagen, Hilden, Germany) [72], according to the manufacturer's protocols.

2.4. Chloroplast DNA quantitation

The amount of cpDNA was determined, according to Houston et al. [56], by real-time PCR on a 7500 Real-Time PCR System (Applied Biosystem Hispania, Madrid, Spain) with SYBR™ Green PCR Master Mix (Thermo Fisher Scientific, Madrid, Spain) and chloroplast cannabis-specific primers *Cscp001* [55]. Previously, a *C. sativa* chloroplast DNA standard was prepared and serially diluted to generate calibration curves in triplicate; only results with $R^2 \geq 0.99$ were used. Each sample was quantified in triplicate.

2.5. SNaPshot™ assay development and genotyping

Primers from hotspot regions were designed using the Geneious Prime® software v.2022.0.1 (Biomatters, New Zealand) and Primer3 online software [73]. The Auto-dimer software [74] was used to detect any primer-primer interactions. Annealing temperatures were determined using empirical formulas [75] and a touchdown strategy, and finally confirmed by 1.5 % agarose gel electrophoresis with SYBR™ Safe DNA Gel Stain (Thermo Fisher Scientific). Amplification of each hotspot region was carried out using a GeneAmp® PCR System 9700 (Applied Biosystems). The PCR reaction mix (10 μL) consisted of: 6.25 μL of Type-it Microsatellite PCR Master Mix (Qiagen), 1.50 μL of 5 × Q solution (Qiagen), a 1 μL aliquot of sample DNA (20–40 pg) and 1.25 μL [2 μM] of primer mix (Thermo Fisher Scientific). Due to the TA rich nature of the cannabis chloroplast genome, primers design was limited and multiplex assays were generated when possible. PCR reactions were amplified in five assays as follows: 1 (Duplex: *ndhF, ndhD*); 2 (Duplex: *psaB, petA-psbJ*); 3 (Triplex: *rpl36-rps8, psbE-petL, trnH-psbA*); 4 (Singleplex: *rps15-ycf11*); and 5 (Singleplex: *rps11*). Forward and reverse primer sequences and assay formats are displayed in Table 1. Thermal cycling conditions consisted of an initial activation at 95 °C for 5 min, 30 s at 95 °C, 90 s at the optimal annealing temperature (60–63 °C), and 30 s at 72 °C repeated for 30 cycles and a final extension of 30 min at 60 °C. PCR products were then purified to remove primers and unincorporated deoxynucleotides by adding 2 μL of Exonuclease I (10 U/ μL, Thermo Fisher Scientific) and 5 μL of Calf Intestinal Alkaline

Phosphatase (CIAP) (1 U/μL, Promega Biotech, Spain) and incubated at 37 °C for 1.5 h and 75 °C for 30 min. Single base extension (SBE) was performed using the SNaPshot™ Multiplex Kit (Thermo Fisher Scientific) according to manufacturer's instructions [76]. SBE primers were designed using the Geneious Prime® software v.2022.0.1 (Biomatters, New Zealand), the Primer3 online software [73], and the Auto-dimer software [74]. Sequence and orientation of SBE primers is displayed in Table 2. Neutral DNA sequences of different lengths were added to SBE primers to allow spacing between SNP products [77]. SBE was conducted on the GeneAmp® PCR System 9700 (Applied Biosystems) thermocycler following manufacturer's instructions [76]. A second purification was conducted by adding 1 μL of CIAP followed by incubation at 37 °C for 1.5 h and 75 °C for 30 min. Capillary electrophoresis was performed either on a 3730 or 3730XL DNA analyzers (Applied Biosystems™) using the following run conditions: oven 66 °C; pre-run 15 kV, 180 s; injection 2 kV, 10 s; run 15 kV, 20 s; capillary length 50 cm; polymer POP- 7™; and dye set Any5Dye. A custom panel and bins set was developed for SNP analysis using GeneMapper ID v3.2 software (Applied Biosystems™).

2.6. Sanger sequencing

Sanger sequencing was performed to assess variability of each chosen polymorphic sites in the chloroplast genome: *ndhF, ndhD, psaB, petA-psbJ, rpl36-rps8, psbE-petL, trnH-psbA, rps15-ycf11* and *rps11*. Eighteen samples, from Italian hemp, Spanish hemp, Spanish marijuana, Chilean marijuana, and USA/Mexican marijuana were chosen from the hundred and six samples and sequenced using the BigDye™ Terminator v3.1 Cycle Sequencing Kit (Thermo Fisher Scientific) according to the manufacturer's instructions [78]. Primer design, primer-primer interactions, and annealing temperature were determined as described in the "SNaPshot™ assay development and genotyping" paragraph below. PCR amplification was performed on the GeneAmp® PCR System 9700 (Applied Biosystems) using the Type-it Microsatellite PCR Master Mix (Qiagen). Sequencing products were purified using paramagnetic bead protocol of the Mag-Bind® SeqDTR kit (Omega bio-tek, GE, USA), following manufacturers' protocol, and run on a 3730 DNA analyzer (Applied Biosystems™) using the following pre-set run conditions: oven 60 °C; pre-run 18 kV, 60 s; injection 1.6 kV, 8 s; run 19.5 kV, 1020 s; capillary length 50 cm; polymer POP-7™; and dye set Z. Sequences were analyzed and proof-read using Geneious Prime® software v.2022.0.1 (Biomatters, New Zealand) in 5'-> 3', as universal practice suggested by ISFG [79].

2.7. Statistical analysis

In order to expand the geographical area, additional samples from not analyzed geographical areas in previous similar studies (Italy and Spain) were genotyped in our research [55,56,59–61]. A total of ninety samples, from Italy and Spain, both marijuana and hemp were genotyped and combined with that samples witch origin was already tested, by using others markers, in others previous studies [56,59–61]. Based on the genotype SNaPshot™ analysis assays, samples that presented same haplotype were grouped together. The four-haplotype distribution through five population was displayed in Table 3, and haplotype data resulted by the combination of nine SNP markers. The geographical distribution and frequency of each of the four haplotypes through five populations, from different biogeographical areas, were displayed in a haplotype map. The haplotype data was constructed and analyzed by PopART software v. 1.7 [80].

For the nine cpDNA markers, the pairwise F_{ST} between each pair of the five populations were calculated using the exploratory population genetics software Arlequin v. 3.5.2.2 [81]. The F_{ST} pairwise genetic distance was used to identify and measure the four haplotype population differences based on genetic distance. Phylogenetic inferences were conducted using the Neighbor-joining (NJ) method in the Mega v. 11.0.11 software [82]. A distance matrix was generated using the NJ method with co-ancestry as genetic distance estimating the genetic distance between populations. In addition, and to visualize the relationship between populations, the haplotypes data were used to perform a Principal Component Analysis (PCA) as variables. The PCA was used to compare the analyzed cannabis groups by using the Past 4.03 software [83].

3. Results and discussion

3.1. Sanger sequencing

Eighteen of the one hundred and six samples were sequenced to confirm the previously predicted polymorphisms in the nine hotspot regions. The sequencing of the nine SNPs revealed that all loci were bi-allelic: *ndhF* (T/G), *ndhD* (T/A), *psaB* (T/G), *petA-psbJ* (C/A), *rpl36-rps8* (T/C), *psbE-petL* (T/G), *trnH-psbA* (G/T), *rps15-ycf11* (C/T) and *rps11* (T/G) (Table 3). As a result, SNaPshot™ genotyping assays were designed for all nine polymorphic loci. Lastly, nine novel sequences were submitted to GenBank® (accession numbers OP584332, OP584333, OP584334, OP584335, OP584336, OP584337, OP584338, OP584339, OP584340).

3.2. SNaPshot™ assay

Based on previously reported gene studies and their polymorphic content, nine hotspot regions in *C.sativa* were chosen to be analyzed for forensic purposes: *ndhF, ndhD, psaB, petA-psbJ, rpl36-rps8, psbE-petL, trnH-psbA, rps15-ycf11 and rps11*. Each of nine hotspot regions showed variability at the SNP site making them suitable for discriminating crop type and geographical origin. The polymorphisms are listed in Table 3. Five assays, three multi-plex and two single-plex, were developed to genotype five sample sets of cannabis samples coming from four different countries. All samples yielded full profiles for the five assays.

Table 2
SBE primer list. SBE primer sequences used for the amplification of the nine SNaPshot™ SBE markers, lower case letters indicate neutral sequences. In table appears each SNP position, primer orientation and length on consensus sequence. The nine sequences were submitted to GenbBank®.

Hot-Spot region	SBE Primer Sequence	SNP on consensus position bp	SBE Primer length bases	SBE Primer orientation	GenBank® Acc. Numb.
ndhF	cagtgacCGATTATATGCCCAATCATATATCACATTTA	110,357	38	F	OP584337
ndhD	cagtgacCTATTCCCAACCTCCAAGAAAAATT	116,350	32	F	OP584336
psaB	ctatTCCTGGATCCAGGGGGGAAT	39,630	24	R	OP584335
petA-psbJ	CTTCCTCTCATTCATTTTTTTGATACTTTG	63,645	30	R	OP584334
rpl36-rps8	cagtgacGGTTTAATTTCTGAACCACTTCCC	80,148	31	R	OP584333
psbE-petL	gcagtgacGGTACACTATTGACGATCTCACAAAGAT	64,999	36	R	OP584332
trnH-psbA	CTTTTATCTTGTCTAAAATTGAAAT	124	25	R	OP584338
rps15-ycf11	CCATTCGGCTTGATAATACTAATTTGAC	123,216	28	R	OP584339
rps11	CAGTGACATTACCCGCCCCCGTACATC	79,722	27	F	OP584340

Table 3
Haplotypes table. The four haplotype-distribution detected through five cannabis populations (N = 106). Rows hold the four haplotypes, and columns contains the nine hotspot regions evaluated in *C. sativa* chloroplast regions. Each of the nine locus resulted bi-allelic.

Haplotype	ndhD	ndhF	psaB	petA-psbJ	trnH-psbA	rpl36-rps8	psbE-petL	rps15-ycf11	rps11
1	A	G	C	G	A	A	C	A	G
2	T	T	A	G	A	A	C	G	T
3	T	T	A	T	C	G	A	G	T
4	T	T	A	G	C	A	A	G	T

An example of the resulting electropherogram is displayed in Fig. 1. Polymorphisms were observed in all nine plastid regions.

3.3. Haplotype analysis

Since Gilmore *et. al* advanced that their haplotype results were useful to provide important biogeographic information of *C. sativa* samples, the haplotype approach has been suggested as tool to predict features within plants [84–86]. Indeed, it has the advantage of combining the effects of multiple polymorphic sites to observe the variety, it may assist in individualization, and can also provide information about geographical distribution.

We included 106 cpDNA sequences of *C. sativa* in the analysis. A total of four different haplotypes (Table 3) were observed in this study, their geographic distribution between samples was constructed and a haplotype map was generated by PopART software (Fig. 2). Haplotype 1 was the most common, observed in 58.5% of all analyzed samples (N = 106) (Table 4). This haplotype was observed in all sample groups except Italian hemp and Spanish hemp. Marijuana samples from the same case number shared haplotypes without exception. Interestingly, Haplotype 2 (3.77% of samples) was shared by three different Chilean sample cases and one USA-Mexico border marijuana sample case. This haplotype

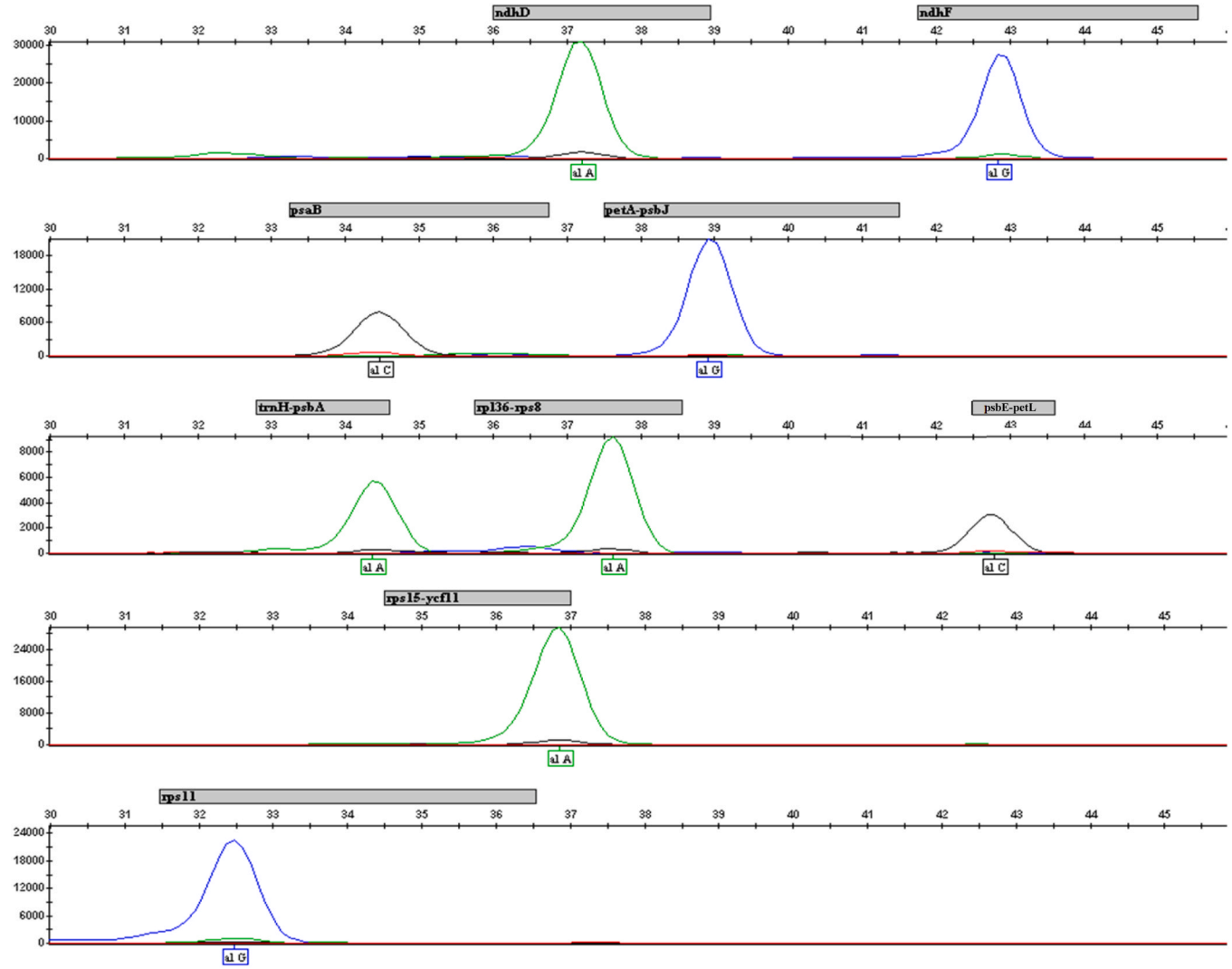


Fig. 1. Example of a full SNaPshot™ electropherogram of a cannabis sample (ID# 210321_10). The sample was run with an ABI 3730/3730-XL, using POP-7 on a 50-cm capillary length array. The electropherogram correspond to the nine SNPs genotype. The ndhD/ndhF duplex expressed an A and G respectively. While the second duplex, psaB/petA-psbJ, showed the presence of C and G respectively. The trnH-psbA/rpl36-rps8/psbE-petL triplex peaks were identified as A, A and C respectively. Lastly, the two single plex assays for rps15-ycf11 and rps11 showed the presence of an A and a G respectively. Alleles are named according to the base on the forward cpDNA strand.

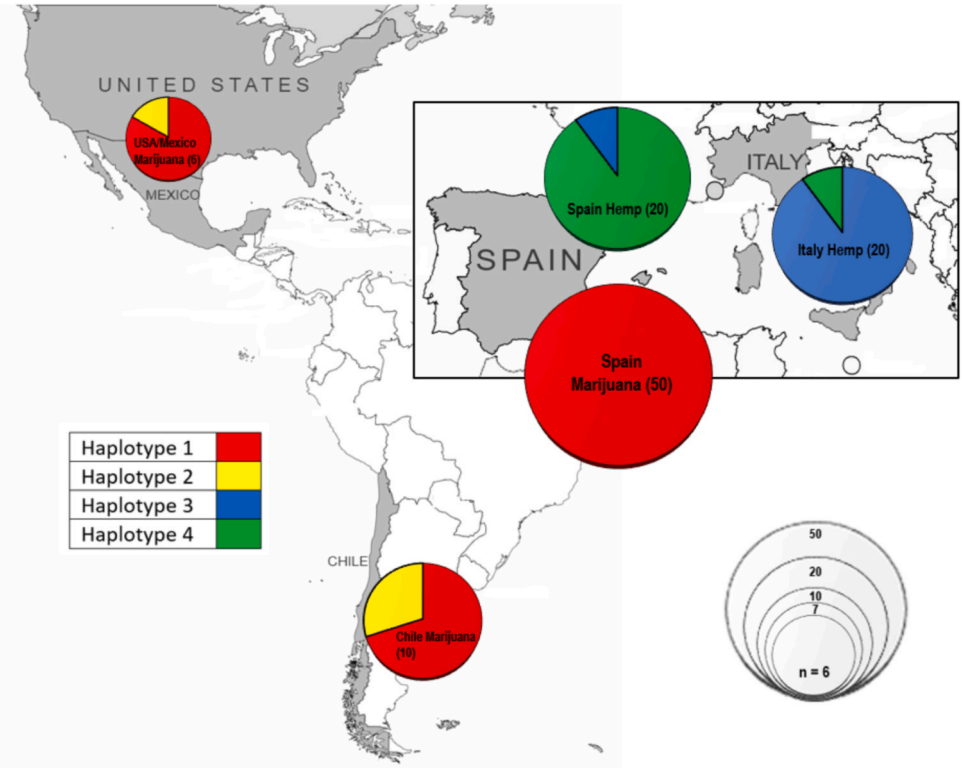


Fig. 2. Haplotype map. The map shows the haplotype proportion observed in the five sample groups (N = 106) with nine SNP markers in nine hotspot regions: trnH-psbA, psaB, petA-psbJ, psbE-petL, rps11, rpl36-rps8, ndhF, ndhD and rps15-ycf11. The legal crops are not sharing haplotypes with illegal crops. Showing genetic distinction, between hemp and marijuana samples, in these nine hotspot regions. Different colored portions in each pie chart represent haplotype frequencies. The size of each circle is proportional to the population size. The map has been generated with PopART software v.1.7.

Table 4
Crop type characteristics, geographical origin and haplotypes of samples (N = 106) used in SNapShot™ assay.

N	Crop Type	Country	Haplotype (N)*
20	Hemp	Italy	3 (18) 4 (2)
10	Marijuana	Chile	1 (7) 2 (3)
6	Marijuana	USA/Mexico border	1 (5) 2 (1)
50	Marijuana	Spain	1 (50)
20	Hemp	Spain	3 (2) 4 (18)

* Numbers in parentheses indicate number of cannabis samples "N" with that haplotype.

sharing was previously reported [56,59,60] with other cannabis chloroplast SNPs and could indicate that this USA-Mexico marijuana seizure was unrelated to any of the other cases, perhaps even originating in a different country. To note, all tested marijuana hybrid samples with a known concentration of THC > 0.3, or seeds sold as marijuana seeds (i.e., our THC range from 1.2 to > 20) showed one of these two haplotypes. Haplotypes 3 and 4 (at least 19% of samples each) were observed only in hemp samples grown in Italy and Spain. Haplotype differences were observed between hemp and marijuana samples. Among different crop types, Haplotypes 1 and 2 were only shared by marijuana samples. Two haplotypes (Haplotypes 3–4) were only observed in hemp. The haplotype proportions, mapped by geographic origin, are showed in Fig. 2.

The haplotype network (Fig. 3) shows a relationship of all haplotypes with few evolutionary steps between them. According to the haplotype network, based on the TCS algorithm, cannabis populations are divided in two main groups which shows some sign of segregation based on the geographical distribution of the populations. These two groups reflect a split between hemp and marijuana samples, as Italian hemp samples are grouping with sequences from Spanish hemp (Hap_3 and Hap_4). The marijuana samples from Chile and USA/Mexico are more closely related to those from illegal Spanish samples representing the second and first groups (Hap_2 and Hap_1). The TCS network results are displayed in Fig. 3. Results indicate that the nine polymorphic regions vary in the 106 analyzed samples by different crop type and geographical origin. According to authors, the massive Haplotype 1 sharing for the marijuana samples, may depend on the access and ease of buying seeds from all around the world. In this way, European cannabis genetics can easily travel to another continent and vice versa. At this time, this is the first study were SNPs SNaPshot™ assays were able to totally differentiate the hemp crop from the marijuana crop.

3.4. Statistical analysis

Genetic differentiation between any of two of the five populations was calculated by Arlequin v. 3.5.2.2 software, the F_{ST} value was used to pairwise the genetic distance analysis. Using the nine SNP markers in this investigation, Italian hemp showed F_{ST} value from 0.7689 to 0.9879 with all other populations with all p value < 0.05, and the Spanish hemp's F_{ST} value were from 0.7689 to 0.9851 with all other populations with all p value < 0.05. These results demonstrate that the genetic differences between each group to Italian and Spanish hemp was significant. Statistically significant differences ($p < 0.05$) happen between Italian hemp and Spanish hemp denoting that these markers can detect some differences between the same crop samples (hemp). Spanish marijuana had F_{ST} value smaller than others from 0.4465 to 0.5419 and statistical differences ($p < 0.05$) were detected between each group except for the USA-Mexican marijuana with a p value > 0.05 (0.1261), denoting that these markers could be useful to detect some differences in

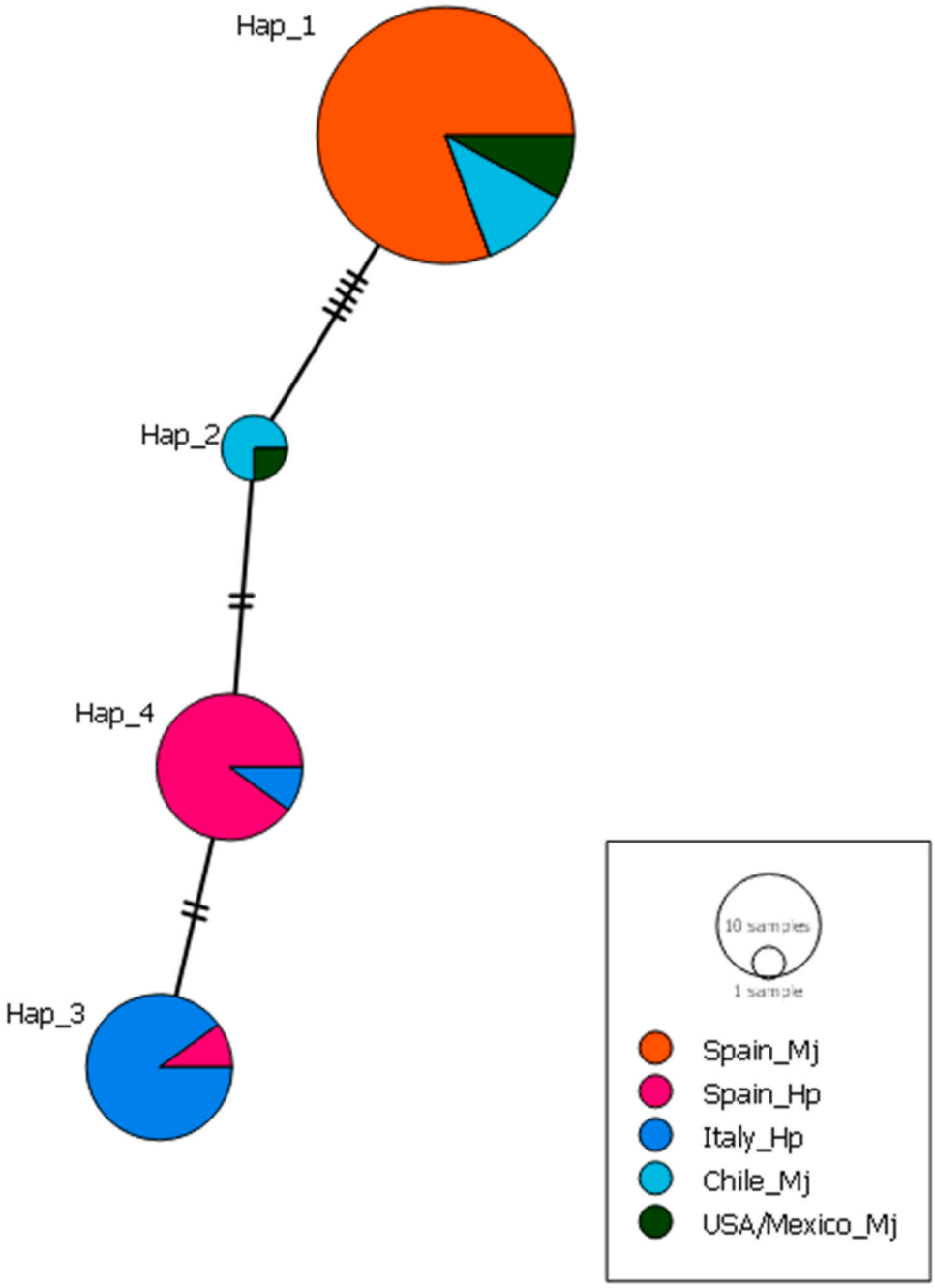


Fig. 3. TCS network constructed by PopART software. Haplotypes are represented by circles whose size are proportional to the number of individuals. Colors represents haplotype distribution in a geographic area. Haplotype 1: 50 samples from Spanish marijuana, 7 samples from Chilean marijuana and 5 from USA/Mexican marijuana. Haplotype 2: 3 samples from Chilean marijuana and 1 from USA/Mexican marijuana. Haplotype 3: 18 samples from Italian hemp and 2 from Spanish hemp. Haplotype 4: 18 samples from Spanish hemp and 2 from Italian hemp. Mutational steps between haplogroups are indicated by hatch marks.

same crop (marijuana) samples. Notable data came from the comparison of USA-Mexico Marijuana and Chilean Marijuana that showed a p value > 0.05 (0.9909), demonstrated that no statistically significant difference was detected between these two populations (Table 5).

Phylogenetic analysis was performed by the NJ tree method and pairwise comparison of the five populations (Fig. 4), using coancestry as genetic distance, which revealed the genetic association of two sets of populations (hemp and marijuana). The distance matrix comparing the five sample groups was showed in Table 6. The neighbor-joining method data yielded consistent results with the F_{ST} distance matrix, with genetic similarities existing between USA-Mexican marijuana, Chilean marijuana, and Spanish marijuana populations, and additional genetic similitudes for Italian hemp and Spanish hemp.

Since PCA is regarded as the best tool in the germplasm collection, because it helps to better understand the structure of the entire germplasm collection, it was performed to detect the most suitable variables among our five populations. Using the nine SNP markers, the PCA resulted just in one Principal Component (PC) with eigenvalues > 1 (1.62). This result only explains a small portion of variance between the used variables. The scatter plot was consistent with the F_{ST} distance matrix, revealing differences between hemp and marijuana groups. The hemp samples displayed genetic relatedness, sharing Haplotypes 3 and 4, demonstrating it to be the most distinct crop group. Instead the Chilean marijuana, the Spanish marijuana, and the USA/Mexican

Table 5
Comparison among five populations using the pairwise genetic-distance analysis based on Fixation Index (FST). Population-to-population comparison was calculated using Arlequin software v. 3.5.2.2.

Population	Italy Hemp	Spain Hemp	Spain Marijuana	USA-Mexico Marijuana
Spain Hemp	0,7689 (0,0000 ***)			
Spain Marijuana	0,9879 (0,0000 ***)	0,9851 (0,0000 ***)		
USA-Mexico Marijuana	0,9156 (0,0000 ***)	0,8934 (0,0000 ***)	0,4465 (0,1261)	
Chile Marijuana	0,8562 (0,0000 ***)	0,8136 (0,0000 ***)	0,5419 (0,0091 ***)	-0,1 (0,9909)

Probability values of FST are showed in parentheses
*** Statistically significant differences at 0,001 levels

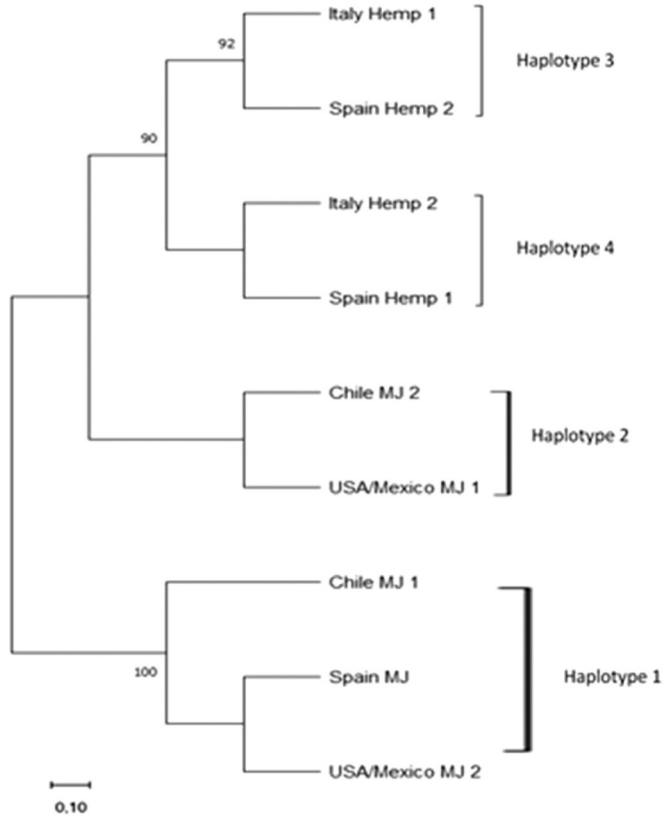


Fig. 4. Neighbor-joining tree depicting genetic distances among five *C. sativa* populations using cpDNA haplotypes; using co-ancestry as genetic distance. The NJ tree shows wide genetics differences between cannabis illegal crops (Haplotypes 1 and 2) and legal crops (Haplotypes 3 and 4). Bootstrap values greater than 89 (percentage occurrence in 1000 replicates) for internal nodes are given at each node (above or below the branches). Scale at the bottom represents 0.1 changes.

Table 6
Distance matrix from MEGA 11.1 software estimated using the Neighbor-Joining method with coancestry distance.

Population	USA/Mexico MJ	Spain MJ	Chile MJ	Italy Hemp
Spain MJ	0.483			
Chile MJ	0.483	0.483		
Italy Hemp	1.223	1.976	1.223	
Spain Hemp	1.223	1.976	1.223	0.131

marijuana clustered together in another big crop group, in which samples share Haplotypes 1 and 2. PCA showed a sharp distinction between Italian-Spanish hemp (Haplotype 3 and Haplotype 4) and the other sample groups (Fig. 5). Haplotype sharing was observed between European samples (i.e., Spanish marijuana) and some American samples (i.e., USA-Mexico marijuana and Chilean marijuana).

Plant species identification studies have used the characteristic of multi-locus barcoding to differentiate samples at the intra-species level [87,88], however, cannabis is becoming commonly accepted to be monotypic and consist only of a single species *C. sativa* [89,90]. Therefore, in forensic molecular plant studies, the capability of this barcoding has been studied to differentiate cannabis at an intra-species level [55,59–61]. Our study successfully developed three multi-locus assays (plus two single locus assays), increased the cpDNA markers database for Cannabis sativa intra-species differentiation, and demonstrated that the use of these markers is useful to differentiate samples by crop type.

3.5. Blind samples test

A blind test was performed by independent laboratory analysts. The sixteen blind samples provided full SNPs profiles and each profile matched with a previous detected haplotype. Eight samples were identified has marijuana and eight as hemp. The eight samples of hemp yielded a full profile, and matched with both Haplotype 3 and Haplotype 4. The four hemp seeds coming from the Spanish bird-seeds mix, resulted in Haplotype 4. Results suggested that they may have a Spanish or Italian origin, with a higher probability of coming from the Spanish area because haplotype 4 was found in 90 % of our hemp samples from the Valencian region and just in 10 % of our hemp samples from Italy (Table 3). The last four samples coming from the human food industry resulted in Haplotype 3. In this case, results may suggest that samples have a higher probability of coming from Italian hemp than Spanish hemp because the haplotype 3 was found in 90 % of our hemp samples from “Le Marche” region and just in 10 % (Table 3) of our Spanish hemp. Lastly, the other eight samples of marijuana also yielded a full profile that matched with the Haplotype 1, confirming just that these samples are part of the illegal crop type. In the future, additional markers may help to further identify marijuana seed origins. The blind sample test was showed in Table 7.

4. Conclusions

This project was focused on nine informative polymorphic SNPs with the purpose of determining crop type (drug versus hemp) and biogeographic origin. Five SNaPshot™ assays were developed to genotype our nine cannabis cpDNA polymorphic markers: *ndhF, ndhD, psaB, petA-psbJ, rpl36-rps8, psbE-petL, trnH-psbA, rps15-ycf11 and rps11*. Using SBE reactions, five groups from four different regions (Chile, USA-Mexico, Spain and Italy) were genotyped. All nine SNP markers were biallelic and variations were detected in all nine markers through the five populations. Combined haplotypes of each sample were used to determine the discriminating power of each markers. Haplotypes and statistical analyses, including phylogenetic analysis, pairwise comparisons and PCA, were performed to demonstrate differences and the relationship grade between populations. All statistical analyses showed a clear distinction between hemp and marijuana samples. Demonstrating that these markers are able to differentiate between these two crop types. Moreover, Chilean and USA/Mexican genetic similarities, were concordant with those of Houston et al. [56] and Roman et al. [59,60]. Based on recent studies [61] on the benefit of combining more loci for a better comprehensive analysis to distinguish cannabis crop type and biogeographical origin, it could be helpful to expand our cpDNA markers and sample database to get more exhaustive results on biogeographical origin (i.e., same country regions differences). Additionally, the development of a comprehensive molecular-analytical tool based on multiple cannabis cpDNA markers will help law enforcement

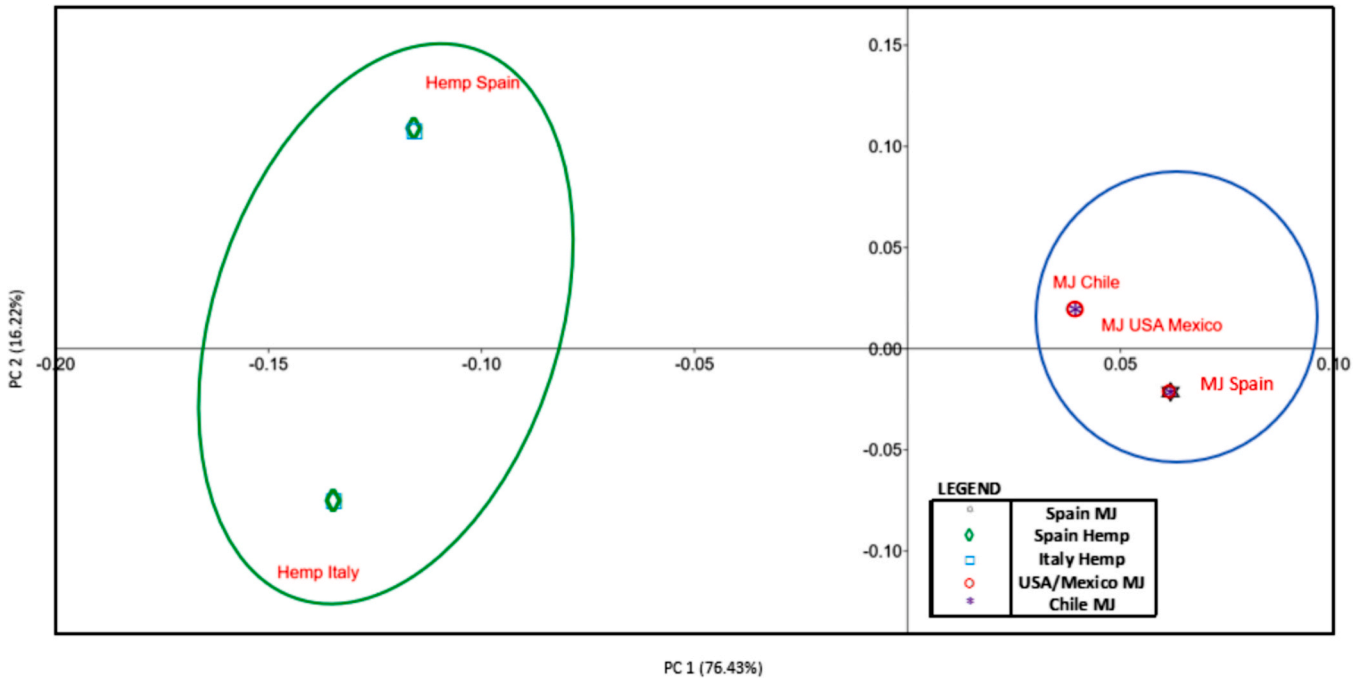


Fig. 5. Principal Component Analysis based on haplotype data (N = 106). PCA plot of the differentiation of cannabis samples analyzed by SNaPshot™ assays, using the nine SNPs found in both hemp and marijuana as discriminant tool. PCA plot was performed using Past3 software. Results showing a clear differentiation between crop type (Hemp vs Marijuana) with USA-Mexico, Chilean and Spanish marijuana being the most similar. Blue and green circles represent clusters of marijuana and hemp samples, respectively. The percentage of variance represented by each PC is shown between parentheses.

Table 7
Test results for blind samples. All marijuana samples from Spain shared haplotype 1. All hemp samples do not share haplotype with marijuana. Italian hemp seed showed haplotype 3. While, the haplotype 4 was expressed in hemp seeds from Spain.

N	Crop Type	Brand	Country of Origin	Haplotype
4	Marijuana	Sensi Seeds	Spain	1
4	Marijuana	Royal Queen Seeds	Spain	1
4	Hemp	Granzoo	Spain	4
4	Hemp	La Biologica	Italy	3

not only to provide leads and intelligence in cases related to drug trafficking, but also it can be applied to other criminal cases where cannabis samples are involved. In summary, this study demonstrated the applicability of genotyping nine cannabis cpDNA markers to differentiate crop type and adding, for the first time European samples to the forensic molecular plant study of *Cannabis sativa*. In a near future, with solid European collaboration between forensic genetics laboratories, it may also provide information regarding marijuana trafficking, and leads suggesting entry points into the European countries by sharing data for associations between cases.

Future studies

The results of the blind sample test demonstrated the applicability of our nine SNP markers. Samples yielded a full profile and were assigned to a proper haplotype. Finally, hemp samples had been assigned to different haplogroups that gave information about different geographical origin. To increase even more the differentiation system, also for the new low-THC level hybrids and mostly the biogeographical origin detection, a few improvements are needed: a, the expansion of our database including worldwide marijuana and hemp samples (from commercial or research seed banks); b, to develop a comprehensive genetic tool composed of multiple cpDNA SNP markers that will provide informative data related to biogeographic origin and crop differentiation [56,59,60] to be applied in forensic intelligence; and c, acquire samples from the > 700 hybrids currently on the market [91] to understand how the five assays cope with these new levels of complexity in samples.

Compliance with Ethical Standards

Not applicable.

Declaration of Competing Interest

The authors declare that they have no competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors would like to thank Dr. Rachel Houston and Mrs. Alejandra Figueroa for their kind contribution of marijuana DNA extracts. The authors would like to thank Dr. Albert Vingut López director in charge at that time of the National Institute of Toxicology and Forensic Sciences (INTCF) Barcelona Department (Spain) for kindly providing cannabis samples. The authors would also like to thank the Genomic Unit team (CCiTUB) of Parque Científico de Barcelona for their helpful technical support.

References

[1] E.M.C. for D. and D.A. (EMCDDA), Cannabis laws in Europe. Questions and answers for policymaking, 2023.
[2] J.J. Coleman, Did the 2018 Farm Bill ' s Hemp Provisions Decriminalize Marijuana ? 2 Passage of Agricultural Improvement Act (AIA) of 2018, 14, 2021 1–21.
[3] A. Potter, D. Weinstock, High Time: The Legalization and Regulation of Cannabis in Canada, Sask. Law Rev. 2020 (2020) 92–94.
[4] Minister of Justice, Cannabis legalisation and control bill - New Zealand, 2020. ⟨https://www.justice.govt.nz/assets/Documents/Publications/2020-Cannabis-Referendum-Public-Release-27-April.pdf⟩.

[5] P. de la República , Decreto N° 120/014, Eastern Asia University Academic Journal. 2014. ⟨https://www.impo.com.uy/diariooficial/2014/05/19/5⟩.
[6] BOCG, Proposición de Ley de regulación integral del cannabis, 2022.
[7] L. conseil federal Suisse, Ordonnance sur les assais pilotes au sens de la loi sur les stupefiants, 2021. 1–25. ⟨https://www.bag.admin.ch/dam/bag/en/dokumente/npp/cannabis/betm-pv-entwurf/entwurf_betm-pv_f.pdf⟩.
[8] European Commission, Council directive 2002/57/EC of 13 June 2002 on the marketing of seed of oil and fibre plants, Off. J. Eur. Union. (2002) 74–97. ⟨https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32002L0057⟩.
[9] European Parliament and the Council of the European Union, Regulation (EU) 2021/2115 OF the European parliament and of the council, Off. J. Eur. Union. 2021 (2021), 435/1-435/186, ⟨https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32021R2115⟩.
[10] D. Baker, G. Pryce, G. Giovannoni, A.J. Thompson, The therapeutic potential of cannabis, Lancet Neurol. 2 (2003) 291–298, https://doi.org/10.1016/S1474-4422(03)00381-8.
[11] M. Ben Amar, Cannabinoids in medicine: a review of their therapeutic potential, J. Ethnopharmacol. 105 (2006) 1–25, https://doi.org/10.1016/J.JEP.2006.02.001.
[12] L. Leung, Cannabis and its derivatives: review of medical use, J. Am. Board Fam. Med. 24 (2011) 452–462, https://doi.org/10.3122/JABFM.2011.04.100280.
[13] L.M. Borgelt, K.L. Franson, A.M. Nussbaum, G.S. Wang, The pharmacologic and clinical effects of medical cannabis, pharmacother, J. Hum. Pharmacol. Drug Ther. 33 (2013) 195–209, https://doi.org/10.1002/PHAR.1187.
[14] D.I. Abrams, C.A. Jay, S.B. Shade, H. Vizoso, H. Reda, S. Press, M.E. Kelly, M. C. Rowbotham, K.L. Petersen, Cannabis in painful HIV-associated sensory neuropathy: a randomized placebo-controlled trial, Neurology 68 (2007) 515–521, https://doi.org/10.1212/01.wnl.0000253187.66183.9c.
[15] J. Corey-Bloom, T. Wolfson, A. Gamst, S. Jin, T.D. Marcotte, H. Bentley, B. Gouaux, Smoked cannabis for spasticity in multiple sclerosis: a randomized, placebo-controlled trial, Cmaj 184 (2012) 1143–1150, https://doi.org/10.1503/cmaj.110837.
[16] A.G. Duque Schumacher, S. Pequito, J. Pazour, Industrial hemp fiber: a sustainable and economical alternative to cotton, J. Clean. Prod. 268 (2020), 122180, https://doi.org/10.1016/J.JCLEPRO.2020.122180.
[17] L. Nissen, A. Zatta, I. Stefanini, S. Grandi, B. Sgorbati, B. Biavati, A. Monti, Characterization and antimicrobial activity of essential oils of industrial hemp varieties (Cannabis sativa L.), Fitoterapia 81 (2010) 413–419, https://doi.org/10.1016/J.FITOTE.2009.11.010.
[18] 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. (1972), https://doi.org/10.1016/S0015-7368(72)70717-3.
[19] M. Pourseyed Lazarjani, S. Torres, T. Hooker, C. Fowlie, O. Young, A. Seyfoddin, Methods for quantification of cannabinoids: a narrative review, J. Cannabis Res. 2 (2020), https://doi.org/10.1186/s42238-020-00040-2.
[20] M. Kojoma, O. Iida, Y. Makino, S. Sekita, M. Satake, DNA fingerprinting of cannabis sativa using inter-simple sequence repeat (ISSR) amplification, Planta Med. 68 (2002) 60–63, https://doi.org/10.1055/s-2002-19875.
[21] 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.
[22] 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.
[23] R. Gilmore, S. Peakall, Isolation of microsatellite markers in Cannabis sativa L, Mol. Ecol. Notes 3 (2003) 105–107, https://doi.org/10.1046/j.1471-8286.2003.00367.x.
[24] 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.
[25] 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. (2009), https://doi.org/10.1007/s00216-008-2500-3.
[26] 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.
[27] V.H.G. Dias, A.S.D. Ribeiro, I.C.T. Mello, R. Silva, B.D. Sabino, R.G. Garrido, L. Seldin, R.S. Moura-Neto, Genetic identification of Cannabis sativa using chloroplast trnL-F gene, Forensic Sci. Int. Genet. 14 (2015) 201–202, https://doi.org/10.1016/j.fsigen.2014.10.003.
[28] L. Valverde, C. Lischka, S. Scheiper, J. Nedele, R. Challis, M.M. De Pancorbo, H. Pfeiffer, S. Köhnemann, Characterization of 15 STR cannabis loci: nomenclature proposal and SNPSTR haplotypes, Forensic Sci. Int. Genet. 9 (2014) 61–65, https://doi.org/10.1016/j.fsigen.2013.11.001.
[29] G.S. Gigliano, L. Cannabis sativa, Botanical problems and molecular approaches in forensic investigation, Forensic Sci. Rev. 13 (2001) 1–17.
[30] A.F. Cowan, K.M. Elkins, Detection and identification of Kratom (Mitragyna speciosa) and Marijuana (Cannabis sativa) by a real-time polymerase chain reaction high-resolution melt duplex assay, J. Forensic Sci. 65 (2020) 52–60, https://doi.org/10.1111/1556-4029.14167.
[31] J. Solano, L. Anabalón, A. Figueroa, D. Gangitano, ITS barcoding using high resolution melting analysis of Cannabis sativa drug seizures in Chile: a forensic application, Forensic Sci. Int. 316 (2020), https://doi.org/10.1016/j.forsciint.2020.110550.
[32] L. Anabalón, J. Solano, F. Encina-Montoya, M. Bustos, A. Figueroa, D. Gangitano, Cannabis seeds authentication by chloroplast and nuclear dna analysis coupled with high-resolution melting method for quality control purposes, Cannabis Cannab. Res 7 (2022) 548–556, https://doi.org/10.1089/can.2020.0168.
[33] T. Yamamuro, S. Miyamoto, M. Kitamura, T. Muro, Y.T. Iwata, H. Segawa, K. Kuwayama, K. Tsujikawa, T. Kanamori, H. Inoue, Development of simple and accurate detection systems for Cannabis sativa using DNA chromatography, Forensic Sci. Int. 291 (2018) 68–75, https://doi.org/10.1016/j.forsciint.2018.08.006.
[34] M. Kitamura, M. Aragane, K. Nakamura, T. Adachi, K. Watanabe, Y. Sasaki, Improved on-site protocol for the DNA-based species identification of Cannabis sativa by loop-mediated isothermal amplification, Biol. Pharm. Bull. 41 (2018) 1303–1306, https://doi.org/10.1248/bpb.b18-00272.
[35] C.B.D.O. Matielo, R.P.M. Lemos, D.S. Sarzi, L. de O. Machado, D.C. Beise, P.C. T. Dobbler, R.M. Castro, M.S. Fett, L.F.W. Roesch, F.A. de O. Camargo, V. M. Stefenon, Whole plastome sequences of two drug-type cannabis: insights into the use of plastid in forensic analyses, J. Forensic Sci. 65 (2020) 259–265, https://doi.org/10.1111/1556-4029.14155.
[36] E. Borroto Fernandez, V. Peterseil, G. Hackl, S. Menges, E. de Meijer, C. Staginnus, Distribution of chemical phenotypes (Chemotypes) in European agricultural hemp (Cannabis sativa L.) cultivars, J. Forensic Sci. 65 (2020) 715–721, https://doi.org/10.1111/1556-4029.14242.
[37] 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.
[38] M.S. Fett, R.F. Mariot, E. Avila, C.S. Alho, V.M. Stefenon, F.A. de Oliveira Camargo, 13-loci STR multiplex system for Brazilian seized samples of marijuana: individualization and origin differentiation, Int. J. Leg. Med. 133 (2019) 373–384, https://doi.org/10.1007/s00414-018-1940-3.
[39] L. de Oliveira Pereira Ribeiro, E. Avila, R.F. Mariot, M.S. Fett, F.A. de Oliveira Camargo, C.S. Alho, Evaluation of two 13-loci STR multiplex system regarding identification and origin discrimination of Brazilian Cannabis sativa samples, Int. J. Leg. Med. 134 (2020) 1603–1612, https://doi.org/10.1007/s00414-020-02338-5.
[40] M. Di Nunzio, V. Agostini, F. Alessandrini, C. Barrot-Feixat, A. Berti, C. Bini, M. Bottinelli, E. Carnevali, B. Corradini, M. Fabbri, P. Fattorini, P. Garofano, S. Gino, A. Mameli, A. Marino, C. Previderè, C. Robino, C. Romano, P. Tozzo, A. Verzeletti, L. Buscemi, D. Gangitano, C. Di Nunzio, A Ge.F.I. – ISFG European collaborative study on DNA identification of Cannabis sativa samples using a 13-locus multiplex STR method, Forensic Sci. Int. 329 (2021), 111053, https://doi.org/10.1016/j.forsciint.2021.111053.
[41] R. Xia, R. Tao, Y. Qu, X. Zhang, H. Yu, C. Yuan, S. Zhang, C. Li, Development and validation of a novel and fast detection method for cannabis sativa: a 19-plex short tandem repeat typing system, Front. Plant Sci. 13 (2022), https://doi.org/10.3389/fpls.2022.837945.
[42] E.P.M. de Meijer, M. Bagatta, A. Carboni, P. Crucitti, V.M.C. Moliterni, P. Ranalli, G. Mandolino, The inheritance of chemical phenotype in cannabis sativa L, J. Ind. Hemp. 8 (2003) 51–72, https://doi.org/10.1300/J237v08n02_04.
[43] M. Kojoma, H. Seki, S. Yoshida, T. Muranaka, DNA polymorphisms in the tetrahydrocannabinolic acid (THCA) synthase gene in “drug-type” and “fiber-type” Cannabis sativa L, Forensic Sci. Int. 159 (2006) 132–140, https://doi.org/10.1016/j.forsciint.2005.07.005.
[44] D. Rotherham, S.A. Harbison, Differentiation of drug and non-drug Cannabis using a single nucleotide polymorphism (SNP) assay, Forensic Sci. Int. 207 (2011) 193–197, https://doi.org/10.1016/j.forsciint.2010.10.006.
[45] K.U. Laverty, J.M. Stout, M.J. Sullivan, H. Shah, N. Gill, L. Holbrook, G. Deikus, R. Sebra, T.R. Hughes, J.E. Page, H. Van Bakel, A physical and genetic map of Cannabis sativa identifies extensive rearrangements at the THC/CBD acid synthase loci, Genome Res. 29 (2019) 146–156, https://doi.org/10.1101/gr.242594.118.
[46] T. Yamamuro, H. Segawa, K. Kuwayama, K. Tsujikawa, T. Kanamori, Y.T. Iwata, Rapid identification of drug-type and fiber-type cannabis by allele specific duplex PCR, Forensic Sci. Int. 318 (2021), https://doi.org/10.1016/j.forsciint.2020.110634.
[47] M.G. Roman, Y. Cheng, S. Kerrigan, R. Houston, Evaluation of tetrahydrocannabinolic acid (THCA) synthase polymorphisms for distinguishing between marijuana and hemp, J. Forensic Sci. (2022) 1–12, https://doi.org/10.1111/1556-4029.15045.
[48] J. Sawler, J.M. Stout, K.M. Gardner, D. Hudson, J. Vidmar, L. Butler, J.E. Page, S. Myles, The genetic structure of marijuana and hemp, PLOS One 10 (2015) 1–9, https://doi.org/10.1371/journal.pone.0133292.
[49] B. Demesure, N. Sodzi, R.J. Petit, A set of universal primers for amplification of polymorphic non-coding regions of mitochondrial and chloroplast DNA in plants, Mol. Ecol. 4 (1995) 129–134, https://doi.org/10.1111/J.1365-294X.1995.TB00201.X.
[50] G. Ferri, B. Corradini, F. Ferrari, A.L. Santunione, F. Palazzoli, M. Alu’, Forensic botany II, DNA barcode for land plants: which markers after the international agreement? Forensic Sci. Int. Genet. 15 (2015) 131–136, https://doi.org/10.1016/j.fsigen.2014.10.005.
[51] C. Santos, F. Pereira, Design and evaluation of PCR primers for amplification of four chloroplast DNA regions in plants, Conserv. Genet. Resour. 9 (2016) 9–12, https://doi.org/10.1007/S12686-016-0605-0.
[52] A.C. Raclariu, M. Heinrich, M.C. Ichim, H. de Boer, Benefits and limitations of DNA barcoding and metabarcoding in herbal product authentication, Phytochem. Anal. 29 (2018) 123–128, https://doi.org/10.1002/pca.2732.

[53] Q. Zhang, Y. Liu, Examination of the cytoplasmic DNA in male reproductive cells to determine the potential for cytoplasmic inheritance in 295 angiosperm species, Plant Cell Physiol. 44 (2003) 941–951, https://doi.org/10.1093/pcp/pcg121.
[54] N. Haider, Chloroplast-specific universal primers and their uses in plant studies, Biol. Plant. 55 (2011) 225–236, https://doi.org/10.1007/s10535-011-0033-7.
[55] 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.
[56] 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.
[57] A. Linacre , J. Thorpe , Detection and identification of cannabis by DNA, 91, 1998. pp. 71–76.
[58] M. Kohjyouma, I.J. Lee, O. Iida, K. Kurihara, K. Yamada, Y. Makino, S. Sekita, M. Satake, Intraspecific variation in Cannabis sativa L. based on intergenic spacer region of chloroplast DNA, Biol. Pharm. Bull. 23 (2000) 727–730, https://doi.org/10.1248/bpb.23.727.
[59] 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.
[60] M.G. Roman, R. Houston, Investigation of chloroplast regions rps16 and clpP for determination of Cannabis sativa crop type and biogeographical origin, Leg. Med. 47 (2020), 101759, https://doi.org/10.1016/j.legalmed.2020.101759.
[61] Y.C. Cheng, R. Houston, Evaluation of the trnK-matK-trnK, ycf3, and accD-psal chloroplast regions to differentiate crop type and biogeographical origin of Cannabis sativa, Int. J. Leg. Med. 135 (2021) 1235–1244, https://doi.org/10.1007/s00414-021-02518-x.
[62] P. Taberlet, L. Gielly, G. Pautou, J. Bouvet, Universal primers for amplification of three non-coding regions of chloroplast DNA, Plant Mol. Biol. 17 (1991) 1105–1109, https://doi.org/10.1007/BF00037152.
[63] S. Dumolin-Lapegue, M.H. Pemonge, R.J. Petit, An enlarged set of consensus primers for the study of organelle DNA in plants, Mol. Ecol. 6 (1997) 393–397, https://doi.org/10.1046/J.1365-294X.1997.00193.X.
[64] G. Forlani, D. Petrollino, A comprehensive molecular approach to the detection of drug-type versus fiber-type hemp varieties, Forensic Sci. Int. Genet. 52 (2021), 102464, https://doi.org/10.1016/j.fsigen.2021.102464.
[65] R.G. Olmstead, C.W. Depamphilis, A.D. Wolfe, N.D. Young, W.J. Elisons, P. A. Reeves, Disintegration of the scrophulariaceae, Am. J. Bot. 88 (2001) 348–361, https://doi.org/10.2307/2657024.
[66] S.L. Datwyler, G.D. Weiblen, On the origin of the fig: phylogenetic relationships of Moraceae from NDHF sequences, Am. J. Bot. 91 (2004) 767–777, https://doi.org/10.3732/ajb.91.5.767.
[67] K.L. Wong, P.P.H. But, P.C. Shaw, Evaluation of seven DNA barcodes for differentiating closely related medicinal Gentiana species and their adulterants, Chin. Med. 8 (2013) 1–12, https://doi.org/10.1186/1749-8546-8-16/FIGURES/7.
[68] J. Shaw, H.L. Shafer, O.Rayne Leonard, M.J. Kovach, M. Schorr, A.B. Morris, Chloroplast DNA sequence utility for the lowest phylogenetic and phylogeographic inferences in angiosperms: The tortoise and the hare IV, Am. J. Bot. 101 (2014) 1987–2004, https://doi.org/10.3732/ajb.1400398.
[69] K.J. Sytsma, J. Morawetz, J.Chris Pires, M. Nepokroeff, E. Conti, M. Zjhra, J. C. Hall, M.W. Chase, Urticalean rosids: circumscription, rosid ancestry, and phylogenetics based on rbcL, trnL-F, and ndhF sequences, Am. J. Bot. 89 (2002) 1531–1546, https://doi.org/10.3732/ajb.89.9.1531.
[70] Q. Zhang, X. Chen, H. Guo, L.M. Trindade, E.M.J. Salentijn, R. Guo, M. Guo, Y. Xu, M. Yang, Latitudinal adaptation and genetic insights into the origins of Cannabis sativa L, Front. Plant Sci. 871 (2018) 1–13, https://doi.org/10.3389/fpls.2018.01876.
[71] H.M. Coyle, T. Palmbach, N. Juliano, C. Ladd, H.C. Lee, An overview of DNA methods for the identification and individualization of marijuana, Croat. Med. J. (2003).
[72] Qiagen, DNeasy® Plant Handbook, 2020.
[73] A. Untergasser, I. Cutcutache, T. Koressaar, J. Ye, B.C. Faircloth, M. Remm, S. G. Rozen, Primer3 — new capabilities and interfaces, Nucleic Acids Res. 40 (2012), e115, https://doi.org/10.1093/NAR/GKS596.
[74] P.M. Vallone, J.M. Butler, AutoDimer: a screening tool for primer-dimer and hairpin structures, Biotechniques 37 (2004) 226–231, https://doi.org/10.2144/04372st03.
[75] N. Von Ahsen, C.T. Wittwer, E. Schü Tz, Oligonucleotide Melting Temperatures under PCR Conditions: Nearest-Neighbor Corrections for Mg 2, Deoxynucleotide Triphosphate, and Dimethyl Sulfoxide Concentrations with Comparison to Alternative Empirical Formulas, 2001. ⟨www.clinchem.org/content/vol47/issue11⟩ (accessed May 6, 2022).
[76] Applied Biosystems, ABI Prism® SNaPshot™ multiplex kit protocol, Appl. Biosyst. (2010) 1–42. ⟨https://tools.thermofisher.com/content/sfs/manuals/cms_041203.pdf⟩.
[77] J.J. Sanchez, C. Børsting, C. Hallenberg, A. Buchard, A. Hernandez, N. Morling, Multiplex PCR and minisequencing of SNPs—a model with 35 Y chromosome SNPs, Forensic Sci. Int. 137 (2003) 74–84, https://doi.org/10.1016/S0379-0738(03)00299-8.
[78] Thermo Fisher Scientific, BigDye™ Terminator v3.1 Cycle Sequencing Kit, Www. Thermofisher.Com. (2015) 9. ⟨www.impact-test.co.uk⟩.
[79] P. Gill, B. Brinkmann, E. D'Aloja, J. Andersen, W. Bar, A. Carracedo, B. Dupuy, B. Eriksen, M. Jangblad, V. Johnsson, A.D. Kloosterman, P. Lincoln, N. Morling, S. Rand, M. Sabatier, R. Scheithauer, P. Schneider, M.C. Vide, Considerations from the European DNA profiling group (EDNAP) concerning STR nomenclature, Forensic Sci. Int. 87 (1997) 185–192, https://doi.org/10.1016/S0379-0738(97)00111-4.
[80] J.W. Leigh, D. Bryant, POPART: full-feature software for haplotype network construction, Methods Ecol. Evol. 6 (2015) 1110–1116, https://doi.org/10.1111/2041-210X.12410.
[81] H.E.L. Excoffier, L, Lischer, Aorlequin v 3.5: an integrated software package for population genetics data analysis, Mol. Ecol. Resour. 10 (2010) 564–567.
[82] K. Tamura, G. Stecher, S. Kumar, MEGA11: molecular evolutionary genetics analysis version 11, Mol. Biol. Evol. 38 (2021) 3022–3027, https://doi.org/10.1093/molbev/msab120.
[83] P.D.R. Øyvind Hammer, David A.T. Harper, Past: paleontological statistics software package for education and data analysis, Palaeontol. Electron. 4 (2001) 9.
[84] M. Kanno, J. Yokoyama, Y. Suyama, M. Ohyama, T. Itoh, M. Suzuki, Geographical distribution of two haplotypes of chloroplast DNA in four oak species (Quercus) in Japan, J. Plant Res. 117 (2004) 311–317, https://doi.org/10.1007/s10265-004-0160-8.
[85] R. Kikuchi, P. Jae-Hong, H. Takahashi, M. Maki, Disjunct distribution of chloroplast DNA haplotypes in the understory perennial Veratrum album ssp. oxysepalum (Melanthiaceae) in Japan as a result of ancient introgression, N. Phytol. 188 (2010) 879–891, https://doi.org/10.1111/j.1469-8137.2010.03398.x.
[86] J. Lihová, H. Kudoh, K. Marhold, Genetic structure and phylogeography of a temperate-boreal herb, cardamine scutata (brassicaceae), in northeast Asia inferred from AFLPs and CPDNA haplotypes, Am. J. Bot. 97 (2010) 1058–1070, https://doi.org/10.3732/ajb.0900361.
[87] W.J. Kress, D.L. Erickson, A two-locus global DNA barcode for land plants: the coding rbcl gene complements the non-coding trnH-psbA spacer region, PLOS One 2 (2007), https://doi.org/10.1371/journal.pone.0000508.
[88] D.L. Erickson, J. Spouge, A. Resch, L.A. Weigt, W.J. Kress, DNA barcoding in land plants: developing standards to quantify and maximize success, Taxon 57 (2008) 1304–1316, https://doi.org/10.1002/tax.574020.
[89] A. Hazekamp, J.T. Fischedick, M.L. Díez, A. Lubbe, R.L. Ruhaak, in: H.-W. (Ben) Liu, L.B.T.-C.N.P.I.I. Mander (Eds.), 3.24 - Chemistry of Cannabis, Elsevier, Oxford, 2010, pp. 1033–1084, https://doi.org/10.1016/B978-008045382-8.00091-5.
[90] B.F. Thomas, M.A. ElSohly, in: B.F. Thomas, M.A.B.T.-T.A.C. of C. ElSohly (Eds.), Chapter 1 - The Botany of Cannabis sativa L., Elsevier, 2016, pp. 1–26, https://doi.org/10.1016/B978-0-12-804646-3.00001-1.
[91] D. Gloss, An overview of products and bias in research, Neurotherapeutics 12 (2015) 731–734, https://doi.org/10.1007/s13311-015-0370-x.