International Journal of Legal Medicine (2019) 133:373–384
https://doi.org/10.1007/s00414-018-1940-3

ORIGINAL ARTICLE

13-loci STR multiplex system for Brazilian seized samples of marijuana: individualization and origin differentiation

Mauro Sander Fett[1,2] · Roberta Fogliatto Mariot[1] · Eduardo Avila[2,3] · Clarice Sampaio Alho[3] · Valdir Marcos Stefenon[4] · Flávio Anastácio de Oliveira Camargo[1]

Received: 30 August 2018 / Accepted: 9 October 2018 / Published online: 15 October 2018

Abstract

It is known that *Cannabis* in Brazil could either originate from Paraguay or be cultivated in Brazil. While consumer markets in the North and Northeast regions are maintained by national production, the rest of the country is supplied with *Cannabis* from Paraguay. However, the Brazilian Federal Police (BFP) has exponentially increased the seizure number of *Cannabis* seeds sent by mail. For this reason, the aim of the study was to assess the 13-loci short tandem repeat (STR) multiplex system proposed by Houston et al. (2015) to evaluate the power of such markers in individualization and origin differentiation of *Cannabis sativa* samples seized in Brazil by the BFP. To do so, 72 *Cannabis* samples seized in Brazil by BFP were analyzed. The principal coordinate analysis (PCoA) and probability identity (PI) analysis were computed. Additionally, the *Cannabis* samples' genotypes were subjected to comparison by Kruskal-Wallis *H*, followed by a multiple discriminant analysis (MDA). All samples analyzed revealed a distinct genetic profile. PCoA clearly discriminated the seizure sets based on their geographic origin. A combination of seven loci was enough to differentiate samples' genotypes, and the PI for a random sample is approximately one in 50 billion. The *Cannabis* samples were 100% correct as classified by Kruskal-Wallis *H*, followed by an MDA. The results of this study demonstrate that the 13-loci STR multiplex system successfully achieved the aim of sample individualization and origin differentiation and suggest that it could be a useful tool to help BFP intelligence in tracing back-trade routes.

Keywords *Cannabis sativa* L. · Genotype · Short tandem repeat · Marker

Mauro Sander Fett and Roberta Fogliatto Mariot contributed equally to this work.

✉ Roberta Fogliatto Mariot
rfmariot@gmail.com

[1] Soil Science Department, Agronomy Faculty, Federal University of Rio Grande do Sul (UFRGS), Avenida Bento Gonçalves, 7712, Porto Alegre, Rio Grande do Sul CEP 91540-000, Brazil

[2] Setor Técnico-Científico, Superintendência Regional do Departamento de Polícia Federal no Rio Grande do Sul, Avenida Ipiranga, 1365, Bairro Azenha, Porto Alegre, Rio Grande do Sul CEP 90160-093, Brazil

[3] Bioscience Faculty, Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS), Avenida Ipiranga, 6681, Bairro Partenon, Porto Alegre, Rio Grande do Sul CEP 90619-900, Brazil

[4] Interdisciplinary Center of Biotechnology Research, Federal University of the Pampa (UNIPAMPA), Rua Aluízio Barros Macedo, Br 290, km 423, Bairro Piraí, São Gabriel, Rio Grande do Sul CEP 97300-000, Brazil

Introduction

Cannabis continues to be the most widely cultivated, produced, trafficked, and consumed illicit drug worldwide [1]. Reports by Member States to the United Nations Office on Drugs and Crime over the period 2009–2014 indicate that the South American countries of Colombia and Paraguay are important sources of the Cannabis herb sold in international markets [2].

The Brazilian Federal Police (BFP) is the main Brazilian law enforcement agency acting nationally to eradicate illicit drugs. However, no detailed studies about *Cannabis* production, distribution, and consumption in Brazil are available. It is known that *Cannabis* in Brazil could either originate from Paraguay or be cultivated in Brazil. The South, Southeast, and West Central Brazilian geopolitical regions are supplied by *Cannabis* cultivated in Paraguay and trafficked from that country. According to the National Anti-Drug Secretariat (*Secretaria Nacional Antidrogas*—SENAD), Brazil receives 85% of its whole *Cannabis* production from Paraguay [3].

While the South, Southeast, and West Central regions of Brazil are supplied with *Cannabis* from Paraguay, consumer markets in the North and Northeast regions are maintained by a national marijuana production. A specific region in Northeast Brazil is an important *Cannabis* producer, known as the Marihuana Polygon (*Polígono da Maconha*) because of its numerous *Cannabis* plantations. Marihuana Polygon is located between Bahia and Pernambuco states, comprising the cities of Petrolina, Juazeiro, Cabrobró, Salgueiro, Floresta (among others), the São Francisco river basin, and the north of Bahia State (Moreira 2005). Approximately 40,000 workers are estimated to be directly and indirectly involved in *Cannabis* plantation, while the marijuana economy reportedly yields around US$52 thousand annually in this region [4].

Although in recent years *Cannabis* drug trafficking forms have been changing, BFP has exponentially increased its seizure number of *Cannabis* seeds sent by mail. This change in the *Cannabis* trafficking scenario may be occurring due to BFP efforts to eradicate large-scale plant cultivation on farms and a shift from the outdoor setting to indoor cultivation, a tendency also observed in Europe [5, 6].

Additionally, licitly and illicitly produced *Cannabis* in jurisdictions that have legalized recreational *Cannabis* use can be used to supply illicit *Cannabis* markets of neighboring jurisdictions. However, the extent to which smuggling has increased as a result of *Cannabis* legalization is difficult to evaluate. Likewise, officials in Argentina and Brazil voiced concern following the legalization of *Cannabis* use in Uruguay [1].

Despite its important cultivation as a source of food, fiber, and medicine, and its global status as the most used illicit drug, the *Cannabis* genus has an inconclusive taxonomic organization and evolutionary history [7]. As academic and commercial interest grows, and as local decriminalization in specific areas or countries occurs, governments, industry, and the research community will encounter a rising demand for taxonomic information to help resolve disputes, establish registered cultivars, perform quality control analysis, and create reliable centralized databases of *Cannabis* information [8].

The usual ways of cultivating *Cannabis* include both clonal and sexual propagation. Clonal propagation results in genetically identical plants [9], which facilitates DNA analysis to link seized drugs to the region where the plant was cultivated. Unfortunately, clonal propagation is not the method adopted for illegal cultivation in South America. The BFP operations have verified the presence of seeds probably used to generate new plantations in the same region or other regions. However, those plantations are located in remote and very small places, making its tracking by police difficult. But plants originating from seed cultivation could still be tracked back by DNA analysis.

The sequencing analysis of the ribulose-1,5-bisphosphate carboxylase/oxygenase (*rbc*L) gene did not differentiate twelve *Cannabis* seizure samples from Southeastern Brazil [10]. On the other hand, many studies have been using short tandem repeats (STR) to genotype *Cannabis* for forensic purposes and have yielded promising results [11–14]. Houston et al. (2015), however, following the International Society of Forensic Genetics (ISFG) and Scientific Working Group for the Analysis of Seized Drugs (SWGDAM) guidelines, demonstrated the applicability of a 13-loci STR system for *Cannabis sativa* [15].

The present study aimed to use the 13-loci STR multiplex system proposed by Houston et al. (2015) to evaluate the power of such markers in individualization and origin differentiation of *Cannabis sativa* L. samples seized in Brazil by the BFP.

Material and methods

Cannabis samples

From Northeastern Brazil, 36 *Cannabis* samples were obtained. Twenty-four plants from eight different locations, at least two plants per location, were seized in 2015 (N15), and 12 *Cannabis* plants from four different locations, three plants per location, were seized in 2017 (N17).

Additionally, 31 *Cannabis* samples cultivated by BFP from foreign seeds seized by postal authorities (PF) were also collected. Individual *Cannabis* stem and/or flower fragments were cut for those 31 samples. The commercial seed brands were identified by their labels: Northern Light, Special Kush, Special Queen #1, Royal Bluematic, Sour Diesel, Blue Mystic, Royal Caramel, Cream Caramel Auto, Souvenir, and Dutch Passion. Some seeds were not labeled. According to seizing documentation, the seeds' origins were mostly European, sold online by Dutch companies.

Finally, five pressed *Cannabis* samples seized in Southern Brazil by regular drug enforcement authorities (APR) were obtained. In this particular case, individual plants could not be collected. Likely, these samples originated from Paraguay, according to police investigation.

All 72 samples analyzed in the present study were seized and/or cultivated by the BFP. All samples were naturally dried at room temperature, without any preparation.

DNA isolation and quantification

DNA isolation was performed using DNeasy Plant Mini Kit (Qiagen, Valencia, CA) and Dneasy mericon Food Kit (Qiagen, Valencia, CA) following the manufacturer's protocol [16, 17], except for the initial amount of sample on Dneasy mericon Food Kit, in which ≤ 20 mg was used, instead of ≤ 200 mg. In order to verify which kit worked better, DNA from 15 *Cannabis* samples were isolated using both kits.

DNA purity and concentration were assessed by absorbance measurements using a NanoDrop Lite Spectrophotometer (Thermo Fisher Scientific, Waltham, MA). For integrity evaluation, 1 μL of DNA was migrated by electrophoresis (20 min at 80 V) in agarose gel (0.8% agarose, 1× TBE) stained with GelRed (Biotium, Hayward, CA).

Cannabis DNA sample concentrations were quantified by Qubit dsDNA HS Assay Kit (Invitrogen Carlsbad, California, USA).

PCR multiplex reaction

The *Cannabis* 13-loci STR multiplex conducted was based on a previous study [15]. Thirteen previously published *Cannabis* microsatellites (E07 CANN1, ANUCS 302, H09 CANN2, D02 CANN1, C11 CANN1, B01 CANN1, B05 CANN1, H06 CANN2, ANUCS 305, ANUCS 308, ANUCS 301, CS1, and ANUCS 501) were used in this study (Table 1). Amplification of these markers was performed via polymerase chain reaction (PCR) using the Multiplex PCR Kit (Qiagen, Valencia, CA) on a VERITI 96-well Thermo Cycler (Thermo Fisher Scientific, Waltham, MA, USA). PCR reactions were prepared at a 12.5-μL volume using 2 ng of template DNA. The PCR master mix consisted of 6.25 μL of 2× QIAGEN Multiplex PCR Master Mix (Qiagen, Valencia, CA), 2.60 μL of 10× Primer mix, 1.25 μL of 5× Q-(Qiagen, Valencia, CA), and 0.4 μL of 8 mg/mL bovine serum albumin (BSA) (Sigma-Aldrich St. Louis, MO).

Forward primers were labeled with four different fluorescent dyes (FAM, PET, NED, and VIC, Life Technologies); final optimal concentrations of forward and reverse primers are shown in Table 1. PCR cycling conditions were as follows: activation for 5 min at 95 °C, followed by 7 cycles of 30 s at 95 °C, 90 s at 60 °C, and 30 s at 72 °C; 5 cycles of 30 s at 95 °C, 90 s at 59 °C, and 30 s at 72 °C; 5 cycles of 30 s at 95 °C, 90 s at 58 °C, and 30 s at 72 °C; 5 cycles of 30 s at 95 °C, 90 s at 57 °C, and 30 s at 72 °C; 5 cycles of 30 s at 95 °C, 90 s at 56 °C, and 30 s at 72 °C; 5 cycles of 30 s at 95 °C, 90 s at 55 °C, and 30 s at 72 °C; and a final extension of 30 min at 60 °C. Every set of PCR reaction included one negative and at least one positive control. The positive control DNA samples 1-D1 and 4-A2 were kindly provided by Dr. David Gangitano (Department of Forensic Science, College of Criminal Justice, Sam Houston State University, Huntsville, TX).

Single PCR experiments followed by 2 μL of PCR product migrated by electrophoresis (40 min at 80 V) in denaturing agarose gel (1% agarose, 1× TBE) stained with GelRed (Biotium, Hayward, CA) were previously performed for at least ten DNA samples and for the 13 primer pairs to confirm the expected amplicon length for each primer locus.

Fragment analysis

Fragment separation and detection of PCR Multiplex products were carried out in a 3130 Genetic Analyzer (Applied Biosystems).

An aliquot of 1 μL of PCR product was added to 9.5 μL Hi-Di Formamide® and 0.5 μL LIZ® 500 Size Standard (Applied Biosystems). Samples were then denatured for 5 min at 95 °C, loaded on the 3130 Genetic Analyzer (Applied Biosystems), and run using the following conditions: oven 60 °C; prerun 15 kV, 180 s; injection 1.2 kV, 23 s; run 15 kV, 1200 s; capillary length 36 cm; polymer POP-4™; and dye set G5.

A bin set and an allelic ladder were also kindly provided by Dr. David Gangitano. The allelic ladder was included on each injection to ensure accurate genotyping.

Table 1 Characteristics of 13 *Cannabis* STR markers used in this study based on Houston et al., 2015

Marker	Dye	STR motif	Type of repeat	Observed alleles	Primer c oncentration (μM)
D02	FAM	(GTT)	Simple	6, 7, 8	0.08
C11	FAM	(TGA)x(TGG)y	Compound/Indel	13, 14, 15, 21	0.10
H09	FAM	(GA)	Simple	11, 12, 16, 18, 19, 20, 21, 22, 23, 24	0.17
B01	FAM	(GAA)x(A)(GAA)y	Complex	10, 13, 14, 15	0.19
E07	VIC	(ACT)	Simple	7, 8, 9	0.62
305	VIC	(TGG)	Simple	4, 8, 9	0.17
308	VIC	(TA)	Simple	3, 5, 6, 7, 8, 9, 10, 11, 12, 13	0.27
B05	VIC	(TTG)	Simple	7, 8, 9, 10	0.06
H06	VIC	(ACG)	Simple	7, 8, 9	0.15
501	NED	(TTGTG)	Simple	4, 5, 6, 7	0.21
CS1	NED	(CACCAT)	Simple	10, 11, 12, 14, 16, 17, 21, 23, 24, 25, 26, 27, 28, 29, 30	0.29
302	PET	(ACA)x(ACA)y(ACA)z	Compound	22, 29, 31, 33, 35, 36, 37	0.17
301	PET	(TTA)	Simple	6, 15, 16, 17, 19, 20, 21, 22, 23, 24, 25	0.62

Genotyping was performed using GeneMapper v. 4.0 software (Applied Biosystems). The analytical threshold was set at 150 relative fluorescence units (RFUs) as recommended by Houston et al. (2015) [15].

The allele's nomenclature used in the present work was developed by Houston et al. (2015) [15] following Valverde et al. (2014) [18] and ISFG recommendations from human-specific STR loci [19, 20].

Single PCR reactions

With the view to having a full DNA profile and checking whether homozygote *Cannabis* samples previous genotyped using the 13-loci STR multiplex system could be heterozygotes, single PCR reactions for B01 CANN1, ANUCS308, and ANUCS301 locus for all 72 *Cannabis* samples were done.

The single PCR reactions were also prepared at a 12.5-μL volume, but using 1 ng of template DNA. An aliquot of DNA (1 μL) from each sample was added to 11.5 μL of PCR master mix. The PCR master mix consisted of 6.25 μL of Taq PCR Master Mix (Qiagen, Valencia, CA), 0.625 μL of primer forward and reverse, and 4 μL of distilled water provided with Taq PCR Master Mix. Single PCR cycling was as follows: activation for 5 min at 95 °C, followed by 30 cycles of 30 s at 95 °C, 90 s at 55 °C, and 30 s at 72 °C, and a final extension for 30 min at 60 °C.

A volume of 5 μL of single PCR product from B01 CANN1, ANUCS308, and ANUCS301 locus was added to a new tube and properly homogenized just before the fragment analysis. The fragment analysis of the PCR product pool was carried out as previously described.

Sensitivity

To establish the sensitivity of the 13-loci STR multiplex system, DNA from two different *Cannabis* individuals were amplified by the following amounts: 20 ng, 10 ng, 2 ng, 1.5 ng, 1 ng, and 0.5 ng per 12.5 μL of PCR reaction volume. The two *Cannabis* samples were selected based on their high heterozygosity. Each series was amplified in triplicate and electrophoretically resolved using the same parameters described above [21].

Mixture study

Pressed *Cannabis* samples seized in the South of Brazil displayed a profile with more than one individual (see "Results and discussion"). In order to interpret those data and evaluate contamination, a mixture study was needed.

Two different mixture sets using DNA from four different *Cannabis* samples were normalized for 2 ng using miliQ water and tested using the following ratios: 19:1, 9:1, 3:1, 1:1, 1:3, 1:9, and 1:19 [21].

The mixture series were amplified in triplicate and electrophoretically resolved using the same parameters described above [21].

Allele 10 of the locus B01 CANN1 sequence

An off-ladder and off-bin peak was observed in 32 of the 36 *Cannabis* samples from Northeastern Brazil, very close to locus B01 CANN1. For these reasons, single PCR product from four homozygous samples that exhibit this peak close to locus B01 CANN1 was sequenced to check whether it could be considered a new, undescribed allele. A specific primer pair was designed using Primer 3 software from NCBI (https://www.ncbi.nlm.nih.gov/tools/primer-blast/). PCR amplification and cycling sequencing were carried out on a VERITI 96-well Thermo Cycler (Thermo Fisher Scientific, Waltham, MA). BigDye Direct Cycle Sequencing Kit (Applied Biosystems) was used and followed the manufacturer's protocol. Samples were run on a 3130 Genetic Analyzer (Applied Biosystems) or a 3500 Genetic Analyzer (Applied Biosystems) under the following conditions: oven 60 °C; prerun 15 kV, 60 s; injection 1.2 kV, 23 s; run 15 kV, 1200 s; capillary length 36 cm; polymer POP-4TM; and dye set Z. Data analysis was performed on the SeqScape v 3.0 (Applied Biosystems).

Allele 10 of the locus B01 CANN1 sequences data has been submitted to the GeneBank databases under accession numbers MH520119, MH520120, and MH520121.

Genetic parameters

The program GenAlEx 6.5 [22, 23] was employed to compute several genetic parameters, in order to determine the usefulness of the STR markers for forensic analyses.

Firstly, the list of private alleles, the total number of alleles (A), the effective number of alleles ($A_E = {}^{1}/_{\sum p_i}$), observed heterozygosity (H_O), expected heterozygosity ($H_e = 1-\sum p_i^2$), Shannon index of diversity ($I = -\sum p_i log_2 p_i$), and polymorphic information content ($PIC = 1-\sum_{i=1}^{n} p_i^2$) were estimated for each locus, considering all samples ($n = 72$) as a single population.

A principal coordinate analysis (PCoA) was computed based on individual genetic pairwise distance ($D = (S_1 - S_2)^2$, where S_1 is the size of allele 1 and S_2 the size of allele 2) for discriminate individuals based on their genotypic information, while Nei's standard genetic distance [24] and F_{ST} were used for a pairwise comparison between seizures. An analysis of molecular variance (AMOVA) was performed to determine the amount of differentiation observed among and within

groups, using 10,000 permutations of microsatellite genotypes to test significance, while 1000 permutations were employed for determining the significance of the among-groups difference estimated.

A probability identity (PI) analysis was used to determine the minimal number of loci needed to determine the genotypic identity of two individuals randomly sampled within each set, based on the combination of STR loci (locus 1; locus 1+ locus 2; locus 1+ locus 2+l locus 3, and so on). This analysis estimates the average probability that two independent samples will have the same identical genotype.

The power of exclusion was estimated as $PE_i = H^2[1-(1-H)H^2]$ for each STR locus and as $PE_{comb} = 1 - \prod(1 - PE_i)$ for the combined set of markers [25]. As complement of the PI, the power of discrimination of the combined set of markers was estimated as $PD_{comb} = 1 - PI$ [26].

Multiple discriminant analysis

Cannabis sample genotypes were subjected to comparison by Kruskal-Wallis *H*, followed by a multiple discriminant analysis (MDA) in the backward mode using the origin of *Cannabis* sample (groups previously defined) as dependent variable and nonmetric, and each sample genotype as independent variable.

The multivariate discriminant function is based on Wilks' lambda (Λ^*) value from the analysis of variance, where the criterion used by the statistical model is the minimization of Λ^* (Eq. 1). Λ^* reaches 1 if all the group means are the same while a low Λ^* value means that the variability within the groups is small compared to the total variability.

$$\Lambda^* = \frac{|W|}{|B+W|} \quad (1)$$

where $|W|$ is the determinant of the matrix of sums of squares due to the error, while $|B+W|$ represents the determinant of the matrix of the total sum of squares. At each step, the property that minimized the overall Wilks' Lambda was entered. The maximum significance of *F* to enter a property was 0.01. The minimum significance of *F* to remove a property was 0.01. The data were carried out on Statistica, version 12, Statsoft Inc. Software.

Different genotypes, obtained by 13 STR loci, were analyzed jointly in order to select the minimum number of markers able to discriminate the *Cannabis* sample groups. A MDA for the alleles of each genotype/sample was applied, considering their geographic origin (seized in Northeastern Brazil in 2015 and 2017, seized in the South of Brazil, and plants cultivated by BFP using seized mail seeds).

In order to process the data, each marker was decompound according to the presence of their alleles. So, if a sample displayed allele 11 on locus H09, the number "11" was added to the data matrix; on the other hand, if the sample does not display allele 11 on locus H09, the number "0" was added to the data matrix instead.

Results and discussion

Quality and amount of isolated DNA

DNA was extracted from all *Cannabis sativa* samples, including pressed samples seized from Southern Brazil. However, the DNA isolated by Dneasy mericon Food Kit demonstrated a higher quality and concentration when compared with DNA isolated by DNeasy Plant Mini Kit. The average amount plus the standard deviation of 15 DNA samples was 14.5 ± 9.03 ng/μL in a total volume of 100 μL using Dneasy mericon Food Kit and 2.68 ± 1.37 ng/μL in a total volume of 100 μL using DNeasy Plant Mini Kit. Also, the purity of the 15 isolated DNA samples assessed by 280/260 ratio was 1.454 ± 0.175 using DNeasy Plant Mini Kit and 1.795 ± 0.068 using Dneasy mericon Food Kit.

It should be noted that an adequate amount and quality of DNA was isolated with a low amount of plant material and no sample preparation was necessary. Those issues will be very important for future lab workflow design, since most *Cannabis* seizures from Southern Brazil are pressed, very dry, and seem to have very-low-quality samples. Despite that, employed extraction methods provided an adequate amount and quality of DNA obtained.

STR multiplex

The *Cannabis* STR multiplex system previously reported by Köhnemann et al. (2012) [14] and optimized by Houston et al. (2015) [15] was used as reference for the present work with some modifications such as a PCR reaction using Multiplex PCR Master Mix (Qiagen, Valencia, CA) and different PCR cycling conditions, primer, and DNA concentration.

A total of 72 samples of *C. sativa* were amplified using the 13-loci STR multiplex system tested, although only 33 samples (40.3%) showed a full DNA profile. The loci affected by locus dropout were B01 CANN1 (37.5%), ANUCS308 (27.8%), and ANUCS301 (54.2%). Houston et al. (2015) also observed similar loci dropout [15].

Houston et al. (2015) experimentally determined the annealing temperatures of these three problematic markers to check whether weak primer binding and eventually primer-primer interaction were the cause of allele dropout. The annealing temperatures of markers ANUCS301, ANUCS308, and B01 CANN1 were 53, 55, and 55 °C, respectively [15].

For this reason, instead of using a fixed annealing temperature of 60 °C on the multiplex PCR cycle, a touchdown of the annealing temperature from 60 to 55 °C was proposed, as

described in "Material and methods." With touchdown cycling, unspecific amplification was no longer observed; however, lower temperatures necessary to reach the theoretical annealing of ANUCS301 (annealing temperatures for touchdown cycling ranging from 60 to 53 °C) resulted in many unspecific peaks on obtained electropherograms.

In order to attain full DNA profiling and check whether homozygote *Cannabis* samples previously genotyped using the 13-loci STR multiplex system could be heterozygotes, single PCR reactions for B01 CANN1, ANUCS308, and ANUCS301 were performed for all 72 samples.

When the pools composed of single PCR reactions were analyzed, an off-ladder and off-bin peak was observed in almost all *Cannabis* samples from Northeastern Brazil, very close to locus B01 CANN1. The peak size, in base pairs, was compatible with an allele comprising ten B01 CANN1 motif tandem repetitions. Since no such allele was previously described, single PCR product from homozygous samples that exhibit this peak close to locus B01 CANN1 were sequenced to check whether it could be considered a new, undescribed allele.

Sequencing analysis of three different Northeastern Brazil samples showing homozygous pattern for such marker confirmed the occurrence of ten tandem repeats of the trinucleotide motif in these samples. Therefore, the hypothesis of a new, not previously described allele, was corroborated. Due to very high frequencies in Northeastern Brazil samples, the addition of the allele 10 bin to B01 CANN1 marker bin set was implemented.

As reported by Houston et al. (2015), only the H09 STR marker showed some difficulties for automatic allele calling due to high stutter peaks [15]. Such a feature is probably due to the dinucleotide STR motif presented by this marker.

Besides the above-described technical issues, effective individualization of *Cannabis* drug samples was achieved using only the data provided by the 13-loci STR multiplex system, since all 72 samples analyzed revealed a distinct genetic profile, including those plants collected in the same location. These distinct genotypes were confirmed when data from single PCR pools with ANUCS301, ANUCS308, and B01 CANN1 markers were added to the samples profiles.

Sensitivity

For sensitivity testing, 20 ng, 10 ng, 2 ng, 1.5 ng, 1 ng, and 0.5 ng of DNA from two different *Cannabis* individuals were amplified. The sensitivity and optimal input of DNA determined for the 13-loci STR multiplex were 1.0 ng and 2.0 ng, respectively. Allele dropout and peak imbalance were displayed when template DNA was at 0.5 ng. Split peaks were observed when template DNA was at or above 10 ng.

Mixture study

A DNA profile is generally considered to be comprised of more than one individual if three or more alleles are present at one or more loci and/or the peak height ratios between a single pair of allelic peaks for one or more loci are below the empirically determined appropriate threshold for heterozygous peak height ratio(s) [27].

As previously mentioned, *Cannabis* seizures from Brazil are usually pressed. Also, the five samples seized from Southern Brazil displayed electropherograms with more than two peaks in some loci, suggesting the presence of DNA from more than one plant. Budowle et al. (2009) recommended that a laboratory must define within its standard operating protocol the specific elements necessary to make reliable allelic and nonallelic peak assignments [27]. Then a mixture study was done with the intent to interpret those data.

The ratio mixtures 19:1, 9:1, 1:9, and 1:19 showed severe allele dropouts for the minor contributor. Instead, the ratio mixtures 3:1 and 1:3 displayed peak imbalances. Mixture profiling patterns studied are extremely important when Brazilian seized samples are considered, since most apprehended drugs in nationwide drug enforcement operations consist of pressed, mixed samples specially prepared to facilitate smuggling and trafficking. Such samples are usually composed of two or more different *Cannabis* specimens, presenting genetic profiles eventually including more than two alleles for some markers. Although mixture analysis can present significant challenges in this situation, especially due to unknown contributor ratios, our analysis indicates that, for almost all of our samples, major contributor profiles can be easily distinguished from minor donors. Efforts to improve profile quality, mainly regarding achievement of proper peak balance among different markers or distinct alleles within a single marker presenting heterozygous patterns, are necessary for proper identification of mixture occurrence and eventual single contributor profile discrimination.

Critical analysis of peak balance is crucial when more than two alleles are observed for any marker in a single profile. Phenomena leading to this kind of behavior can be related not only to mixed vegetable samples, but to variation in a particular plant chromosomic ploidy instead, since tetraploid *Cannabis* samples have been described in scientific literature, with higher occurrence found among inbred varieties and cultivars designed for high cannabinoid content or improved hallucinogenic effect [28–32]. When such a feature is present, evaluation regarding potential ploidy variability or sample mixture occurrence is essential. Unbalanced peak heights inside any particular marker can be an indication of mixed sample occurrence, in contrast to balanced profiles expected to originate from tetraploid plants. Once again, a correct

Table 2 Genetic parameters estimated for all samples ($n = 72$) based on 13 analyzed *Cannabis* STR loci

	D02	C11	H09	B01	E07	305	308	B05	H06	501	CS1	302	301
A	3	3	10	4	3	4	10	4	3	4	13	6	11
Ae	2.281	1.946	4.233	2.692	1.813	1.625	5.592	2.004	1.988	1.390	9.274	2.848	3.469
I	0.925	0.836	1.706	1.095	0.763	0.647	1.881	0.943	0.825	0.563	2.355	1.207	1.514
Ho	0.431	0.292	0.507	0.306	0.333	0.333	0.347	0.556	0.333	0.208	0.708	0.565	0.514
He	0.566	0.489	0.769	0.633	0.452	0.387	0.827	0.505	0.500	0.283	0.898	0.654	0.717
PIC	0.562	0.486	0.764	0.629	0.448	0.385	0.821	0.501	0.497	0.281	0.892	0.649	0.712
PE	0.272	0.207	0.503	0.337	0.178	0.134	0.593	0.219	0.216	0.074	0.727	0.359	0.432
PD	0.728	0.680	0.913	0.787	0.636	0.562	0.944	0.709	0.674	0.464	0.978	0.818	0.869

A = mean number of alleles; *Ae* = mean effective number of alleles; *I* = mean Shannon index of diversity; *Ho* = mean observed heterozygosity; *He* = mean expected heterozygosity; *PIC* = polymorphic content index; *PE* = power of exclusion; *PD* = power of discrimination

adjustment of multiplex PCR reaction chemistry and conditions is necessary, in order to minimize errors in peak balance evaluation and to reliably determine causes for multiallelic profile observations.

Genetic parameters

The main genetic parameters of forensic interest computed for each locus were calculated based on allele frequency and are

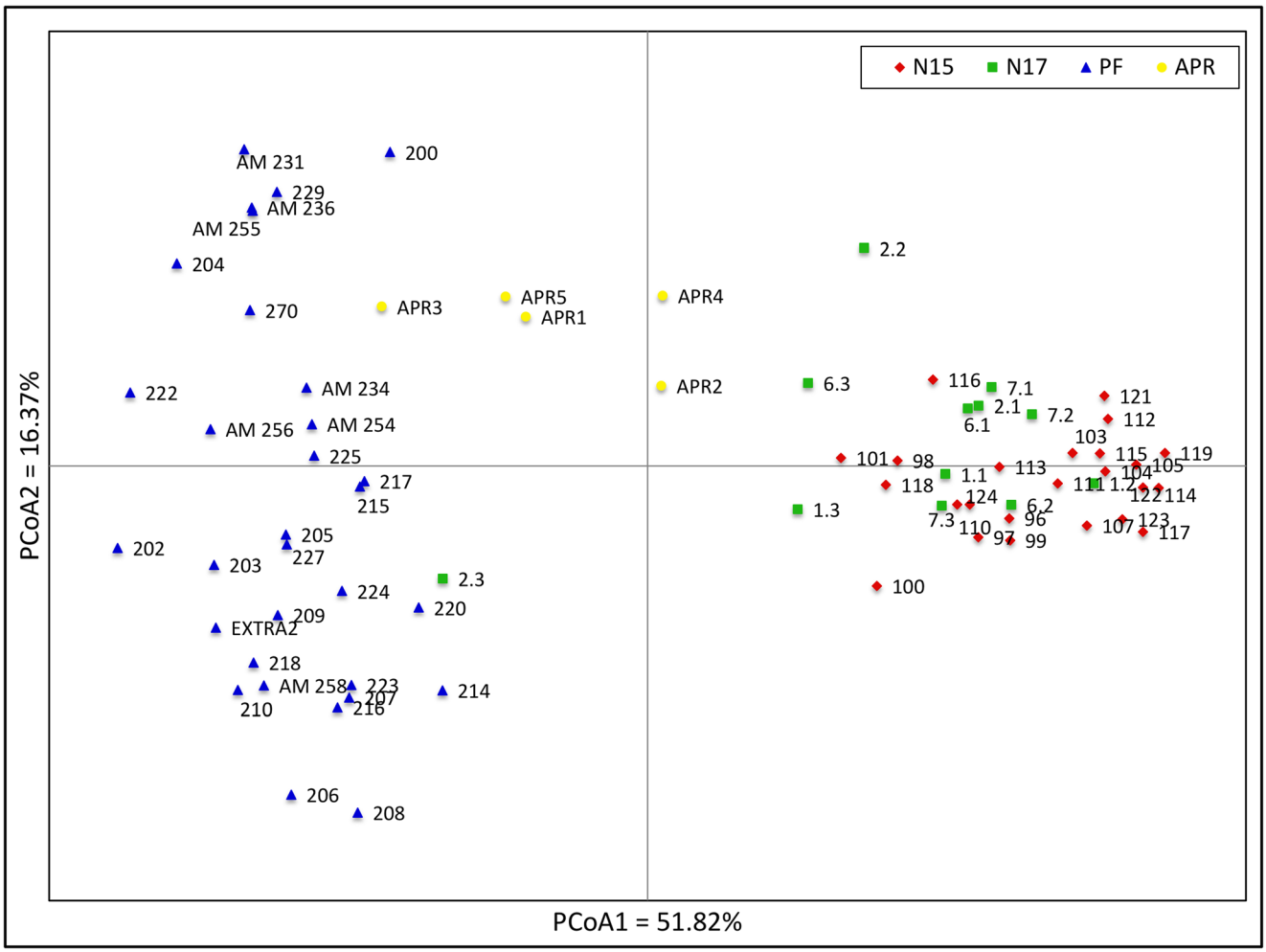

Fig. 1 Principal coordinate analysis (PcoA) of the 72 samples based on genotypes of 13 analyzed *Cannabis* STR loci. N15 = *Cannabis* samples obtained in crops from Northeastern Brazil in 2015; N17 = *Cannabis* samples obtained in crops from Northeastern Brazil in 2017; PF = *Cannabis* samples cultivated from foreign seeds seized by Brazilian post services; APR = regular drug enforcement seizures from Rio Grande do Sul State, Brazil

Table 3 Pairwise Nei's genetic distance (below diagonal) and F_{ST} (above diagonal) for the seizure sets based on genotypes of 13 analyzed *Cannabis* STR loci

Group	N15	N17	PF	APR
N15		0.031	0.224	0.167
N17	0.012		0.155	0.111
PF	0.616	0.446		0.100
APR	0.343	0.228	0.224	

N15 = *Cannabis* samples obtained in crops from Northeastern Brazil in 2015; N17 = *Cannabis* samples obtained in crops from Northeastern Brazil in 2017; PF = *Cannabis* samples cultivated from foreign seeds seized by Brazilian postal services; APR = regular drug enforcement seizures from Rio Grande do Sul State, Brazil

summarized in Table 2. The highest number of alleles, level of heterozygosity, Shannon index of diversity, polymorphic content index, power of exclusion, and power of discrimination were observed for locus CS1, as reported by Houston et al. (2015) [15], followed by loci 308 and H09.

Allelic patterns observed in the studied samples matched the profiles published by Houston et al. (2015) concerning the size of the alleles [15], except by allele 10 from locus B01, which was observed in the Brazilian samples belonging to N15 and N17 groups.

All five APR samples from pressed marijuana seized in Southern Brazil presented more than two alleles for some loci, since these samples are composed of a mixture of unknown plants. In order to enable analysis of these samples together with the remaining dataset, the two alleles with higher RFU values in the total sampling were chosen. Similarly, five samples from leaf or seed seizures presented multiple alleles in at least one STR locus: one sample from apprehended PF seed displayed three alleles in two loci (302 and 501) and four alleles in CS1 locus. For these samples, the two alleles with higher RFU values were also used as the true genotype. A preliminary analysis using different combinations of alleles revealed that the results are the same for all cases, independent of the arrangement of alleles employed (data not shown).

Köhnemann et al. (2012) [14], Knight et al. (2010) [33], and Hsieh et al. (2003) [34] found similar results for locus 302, and Shirley et al. (2013) [35] for locus CS1 (called by NMI01). The multiallelic pattern is suggestive of polyploidy, where multiple copies of the genome are present, a situation common in plants with no adverse effects. However, other explanations of multiallelic patterns include locus duplication and aneuploidy.

It is important to mention that those five sample genotypes were obtained from plant leaf material, with those multiallelic loci exhibiting peaks of approximately equal height. These observations along with the absence of additional peaks at the remaining loci support the conclusion that this profile is not due to cross-contamination or mixed samples, such as pressed samples seized in Southern Brazil.

Overall, the PCoA based on the 13 STR markers (Fig. 1) clearly discriminated the seizure sets based on their geographic origin: the first principal component (PCoA1) expressed with 51.82% of the explained variance, and the second principal component (PCoA2) expressed with 16.37% of the explained variance.

Samples from Northeastern Brazil (N15 and N17), although apprehended in different years, were grouped together, with the exception of one sample (sample 2.3 from the 2017 seizure).

The samples from PF seeds apprehended by the BFP formed a distinct group. Such seeds are usually purchased

Table 4 Analysis of molecular variance (AMOVA) of 13 analyzed *Cannabis* STR loci

Source	Df	SS	MS	Est. Var.	%
Among groups	3	231.163	77.054	4.343	38***
Within groups	68	484.170	7.120	7.120	62
Total	71	715.333		11.463	100

***P (rand ≥ data) = 0.001

Table 5 List of private alleles for each seizure set of 13 analyzed *Cannabis* STR loci

Group	Locus	Allele	Freq
N15	H09	22	0.042
N15	H09	23	0.021
N15	501	7	0.021
N17	302	35	0.083
N17	301	23	0.042
N15 + N17	B01	10	0.819
PF	H09	24	0.033
PF	305	9	0.032
PF	308	3	0.048
PF	308	10	0.032
PF	308	11	0.032
PF	308	13	0.177
PF	CS1	14	0.048
PF	CS1	21	0.113
PF	CS1	30	0.016
PF	302	22	0.018
PF	302	29	0.054
PF	301	24	0.065
PF	301	25	0.016
APR	305	7	0.100

N15 = *Cannabis* samples obtained in crops from Northeastern Brazil in 2015; N17 = *Cannabis* samples obtained in crops from Northeastern Brazil in 2017; PF = *Cannabis* samples cultivated from foreign seeds seized by Brazilian post services; APR = regular drug enforcement seizures from Rio Grande do Sul State, Brazil

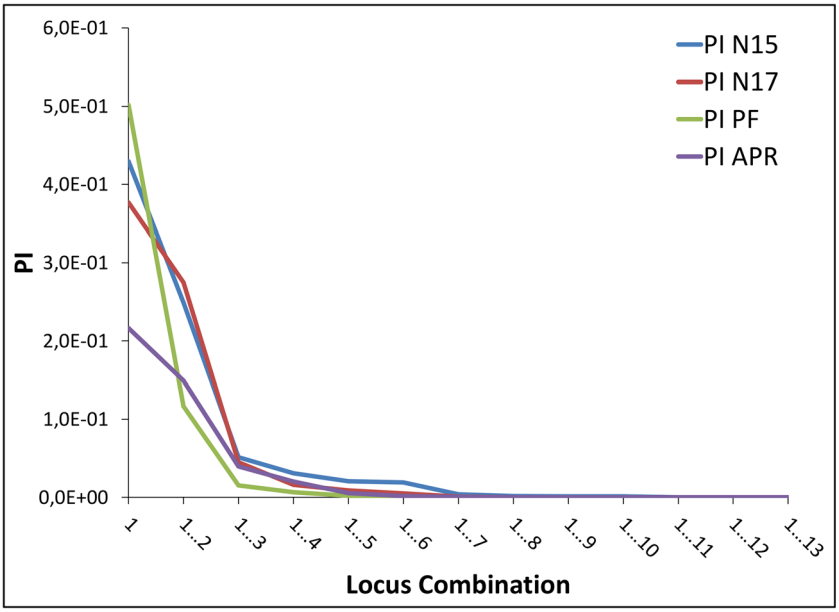

Fig. 2 Probability of genotypic identity of two individuals randomly sampled within each set, based on the combination of STR loci (locus 1; locus 1 + locus 2; locus 1 + locus 2 + l locus 3, and so on). N15 = *Cannabis* samples obtained in crops from Northeastern Brazil in 2015; N17 = *Cannabis* samples obtained in crops from Northeastern Brazil in 2017; PF = *Cannabis* samples cultivated from foreign seeds seized by Brazilian postal services; APR = regular drug enforcement seizures from Rio Grande do Sul State, Brazil

through Internet sites and probably originate from Europe, because all brands identified by their label could be found on European websites. However, some seeds were not labeled, so their precise origin could not be determined.

The five APR pressed samples apprehended in Southern Brazil presented an intermediary position between these two groups. Moreira (2005) [3] and Shibuya et al. (2007) [36] believe that Paraguay supplies the Southern Brazilian market, which led our group to consider that country as the source for these samples. However, these samples could also have been acquired from Uruguay, a country that shares a significant land border with Rio Grande do Sul State, and recently legalized recreational *Cannabis* use.

Table 6 Discriminant function used to identify the ability of each tracer to distinguish the drug source

Fingerprint selected (Marker)	Correctly classified samples (%)
D02	69.7
C11	73.7
H09	84.2
B01	68.4
E07	64.5
305	58.7
308	88.1
B05	76.0
H06	81.6
501	64.5
CS1	92.1
302	88.1
301	88.1

At the individual level, the PCoA analysis was able to discriminate each single sample, indicating an absence of plants with the same genotype. This supports the hypothesis of sexual rather than clonal propagation of the apprehended plants and seeds. Houston et al. (2015) [15], Houston et al. (2017) [37], and Shirley et al. (2013) [35] also found similar results for their samples.

As expected, the pairwise genetic distance (for both Nei's genetic distance and F_{ST}) was smaller between the N15 and N17 seizures, which were apprehended in plantations from the

Table 7 The multiple discriminant analysis results with the number of samples classified per group, the percentage of samples correctly classified, and the percentage of uncertainty for 13 analyzed *Cannabis* STR loci

Groups/origin	Number of samples classify per group			Correct classification (%)/mean
	N15 + N17	PF	APR	
N15 + N17	36	0	0	100
PF	0	31	0	100
APR	0	0	5	100
Total	36	31	5	100
Correct classification (%)	100	100	100	100
Uncertainty associated with samples classification (%)	0	0	0	0

N15 = *Cannabis* samples obtained in crops from Northeastern Brazil in 2015; N17 = *Cannabis* samples obtained in crops from Northeastern Brazil in 2017; PF = *Cannabis* samples cultivated from foreign seeds seized by Brazilian postal services; APR = regular drug enforcement seizures from Rio Grande do Sul State, Brazil

Table 8 The 11 steps of the multiple discriminant analysis, the marker, and its respective allele addition, the Wilks' Lambda, and percentage of samples correctly classified for each step

Step	Marker (allele)	Wilks' Lambda	*P* to remove	% of samples correctly classified (accumulated)
1	CS1	0.0254	< 1E−17	92.1
2	301	0.0068	< 1E−17	97.4
3	H06	0.0036	< 1E−17	100.0
4	H09	0.0023	< 1E−17	100.0
5	302	0.0009	< 1E−17	100.0
6	B01	0.0002	< 1E−17	100.0
7	501	0.0002	< 1E−17	100.0
8	E07	0.0001	< 1E−17	100.0
9	305	0.0001	< 1E−17	100.0
10	308	0.0000	< 1E−17	100.0
11	B05	0.0000	< 1E−17	100.0

same geographic region (Table 3). The largest distances were observed between PF and N15 followed by PF and N17 seizures. Such large distances between these groups are not unexpected due to the sample origins that compound each seizure.

The AMOVA revealed that 38% of the total differentiation at allelic frequency level is due to within-population differences and 62% due to differences among samples within groups (Table 4). Usually, plant populations present higher differentiation within populations than among groups.

This capability of revealing high differentiation among groups suggests that this set of 13 STR markers is, along with a robust individualization efficacy demonstrated by the proposed method, valuable for evidencing the geographic origin of seized plants of *Cannabis*, being a useful tool for forensic investigations.

Private alleles were observed for seven out of the 13 STR loci and for all seizure sets (Table 5). The largest number of private alleles (13) was observed for the PF samples, likely due to the European origin of the seeds. Allele 10 from locus B01 was observed with high frequency in sample sets N15 and N17, originating from seizures in different years, both in the same geographic area. This allele was not reported in previous studies using this STR locus for genotyping samples from other countries, suggesting it is a private allele for samples from Northeastern Brazil.

The PI analysis (Fig. 2) demonstrated that the combination of the 13 STR loci is quite informative for discriminating two samples randomly collected within each seizure set. For all different groups analyzed in the present study, a combination of seven loci was enough to differentiate sample genotypes.

The probability of a match between the genetic profiles of two unrelated samples is one in 2 million for N17, one in 23 million for N15, one in 5 billion for PF, and one in 2 million for APR. Considering the combined frequencies for all datasets (including all groups and samples evaluated in this study), the PI for a random sample seized by BFP is approximately one in 50 billion. Concerning the power of exclusion (PE) and the power of discrimination (PD), the combined set of STR loci generated values near the unity for both parameters ($PE_{comb} = 0.996$ and $PD_{comb} = 0.999$). At individual locus, these parameters are directly correlated with the estimations of heterozygosity (Table 2). Seeing these results, the high power of this set of STR markers for individualization and origin differentiation of *Cannabis sativa* L. samples seized in Brazil by the BFP can be considered kindly confident.

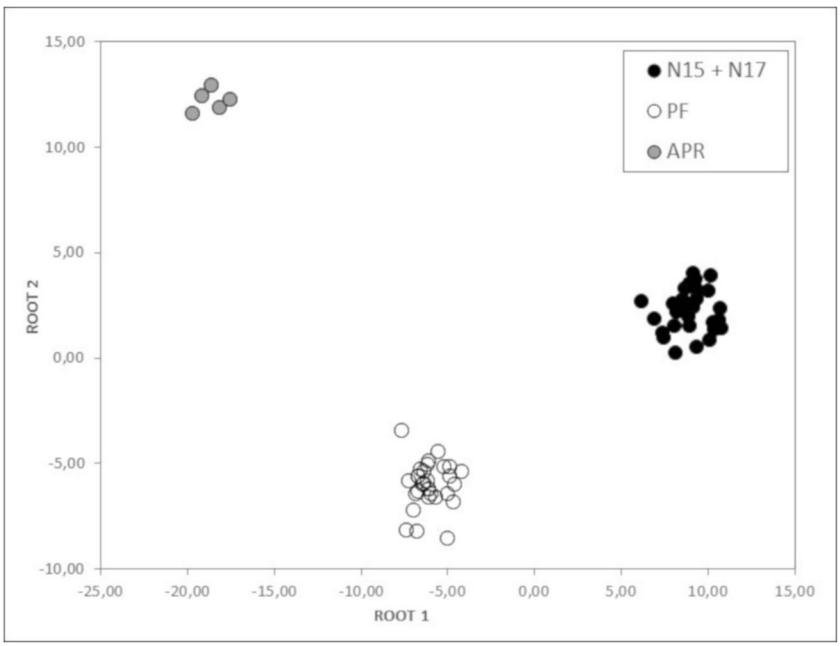

Fig. 3 Two-dimensional scatter plot of the first and second discriminant functions from stepwise discriminant function analysis according to geographic origin of samples. N15 = *Cannabis* samples obtained in crops from Northeastern Brazil in 2015; N17 = *Cannabis* samples obtained in crops from Northeastern Brazil in 2017; PF = *Cannabis* samples cultivated from foreign seeds seized by Brazilian postal services; APR = regular drug enforcement seizures from Rio Grande do Sul State, Brazil

Multiple discriminant analysis

Table 6 shows that all selected markers are able to correctly classify over 50% of the processed samples. Note that only the CS1 marker has the ability to correctly classify 92% of the samples analyzed in the proposed groups. However, for forensic samples, indexes close to 100% are desirable, and it is necessary to use a set of markers to reduce the error with the prediction of the sample origins.

Table 7 shows that all 72 *Cannabis* samples were correctly classified by Kruskal-Wallis *H* followed by MDA. Actually, the classification agrees with the predefined groups. Endorsing these results, the percentage of samples correctly classified was 100%, and the percentage of uncertainty was 0%.

The MDA performed combinations excluding the independent variables that did not contribute to classifying the samples in their predefined group. Accordingly, the markers that properly classified the samples were CS1, 301, H06, H09, 302, B01, 501, E07, 305, 308, and B05. Only C11 and D02 markers were not selected to be a part of the analysis, as they did not contribute to the samples' classification.

The MDA also revealed the contribution of each marker associated to certain alleles for the samples' classification. In Table 8, all 11 analysis steps are listed, including the marker and its respective allele for each step. It is distinctly noticed that Wilks' lambda decreased while percentage of samples correctly classified increased along the MDA.

Canonical variables and discriminant functions can be expressed as a linear combination of the markers included in the particular model [38, 39]. A group discrimination evaluation was performed based on a scatter plot between two first variable roots. Sample distribution in the plane of the two calculated discriminant functions is illustrated in Fig. 3.

The MDA confirmed the PCoA results (Fig. 3) and showed that the 13 STR markers clearly discriminated the seizure sets analyzed on the present work based on their geographic origin. The samples were 100% correctly classified, associated with a null percentage of uncertainty. The MDA results show the 13 STR system's efficiency in evidencing the geographic origin of seized plants of *Cannabis* and suggest that it can be a useful tool for forensic investigations.

Conclusion and perspectives

The 13-loci STR multiplex system successfully achieved the aim of sample individualization and origin differentiation for the *Cannabis* specimens examined in the present work. This system could be a useful tool to help police intelligence in tracing back-trade routes of particular drug syndicates or dealers, and in linking different *Cannabis* plants to a single crop or crime scene. However, a robust database is necessary to increase methodologic efficacy, especially in refining the resolution of the determination of any specific sample's geographic origin, as well as adding other *Cannabis* geographic sources and supply routes. Notably, Uruguayan samples should be of special interest in future studies, since illicit drug traffic has a potential to be markedly affected by recent legalization of *Cannabis* recreational use in that country, which could significantly impact drug consumption patterns in Southern Brazil. Finally, new markers proposed by Valverde et al. (2014) [40] and used by Houston et al. (2017) [37] will be tested, replacing some markers and combining tetranucleotide loci instead.

Acknowledgments We would like to thank Dr. David Gangitano and Dr. Rachel Houston for provide the panel, the bins, the allelic ladder, and two *Cannabis*-positive control samples. We also want to thank Dr. Claudia Paiva Nunes, Lucas Ribeiro, and Pietra Graebin for their technical assistance.

Funding information This work was supported by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) of Brazil (grant PRO-FORENSE 25/2014) and National Institute of Forensic Science and Technology (CNPq/INCT Forense grant 465450/2014-8).

Compliance with ethical standards

Disclaimer The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Conflict of interest The authors declare that they have no conflict of interest.

References

1. United Nations Office on Drugs and Crime (UNODC) (2016) World drug report 2016. UNODC, New York
2. United Nations Office on Drugs and Crime (UNODC) (2018) World drug report 2018. UNODC, New York
3. Moreira EM (2005) A despenalização do trabalhador rural preso no cultivo da Cannabis Sativa. Confluência - Rev Interdiscip Sociol e Direito 4(1):7–21
4. Fraga PCP (2014) Illicit crops in Brazil. The rural workers between income and the exploitation of drug trafficking. Humanit Soc Sci 2: 211–222. https://doi.org/10.11648/j.hss.20140206.20
5. United Nations Office on Drugs and Crime (UNODC) (2015) World drug report 2015. UNODC, New York
6. De Backer B, Maebe K, Verstraete AG, Charlier C (2012) Evolution of the content of THC and other major cannabinoids in drug-type Cannabis cuttings and seedlings during growth of plants. J Forensic Sci 57:918–922. https://doi.org/10.1111/j.1556-4029.2012.02068.x
7. Sawler J, Stout JM, Gardner KM, Hudson D, Vidmar J, Butler L, Page JE, Myles S (2015) The genetic structure of marijuana and hemp. PLoS One 10(8):e0133292. https://doi.org/10.1371/journal.pone.0133292
8. Laursen L (2015) Botany: the cultivation of weed. Nature 525:S4–S5. https://doi.org/10.1038/525S4a
9. Miller Coyle H, Palmbach T, Juliano N, Ladd C, Lee HC (2003) An overview of DNA methods for the identification and individualization of marijuana. Croat Med J 44:315–321

10. Ribeiro A, Dias V, Mello I, Silva R, Sabino B, Garrido R, Seldin L, Moura-Neto RS (2013) O gene rbcL como barcode para identificação forense de *Cannabis sativa*. Saúde, Ética & Justiça 18 (Ed. Especial): 67–71. https://doi.org/10.11606/issn.2317-2770.v18ispep67-71
11. Gilmore S, Peakall R, Robertson J (2003) Short tandem repeat (STR) DNA markers are hypervariable and informative in Cannabis sativa: implications for forensic investigations. Forensic Sci Int 131:65–74. https://doi.org/10.1016/S0379-0738(02)00397-3
12. Howard C, Gilmore S, Robertson J, Peakall R (2008) Developmental validation of a *Cannabis sativa* STR multiplex system for forensic analysis. J Forensic Sci 53:1061–1067. https://doi.org/10.1111/j.1556-4029.2008.00792.x
13. Howard C, Gilmore S, Robertson J, Peakall R (2009) A *Cannabis sativa* STR genotype database for Australian seizures: forensic applications and limitations. J Forensic Sci 54:556–563. https://doi.org/10.1111/j.1556-4029.2009.01014.x
14. Köhnemann S, Nedele J, Schwotzer D, Morzfeld J, Pfeiffer H (2012) The validation of a 15 STR multiplex PCR for Cannabis species. Int J Legal Med 126:601–606. https://doi.org/10.1007/s00414-012-0706-6
15. Houston R, Birck M, Hughes-Stamm S, Gangitano D (2015) Evaluation of a 13-loci STR multiplex system for Cannabis sativa genetic identification. Int J Legal Med 130:635–647. https://doi.org/10.1007/s00414-015-1296-x
16. DNeasy. (2013). Plant handbook. Valencia, CA: Qiagen
17. Dneasy. (2010). mericon food handbook. Valencia, CA: Qiagen.
18. Valverde L, Lischka C, Scheiper S, Nedele J, Challis R, de Pancorbo MM, Pfeiffer H, Köhnemann S (2014) Characterization of 15 STR *Cannabis* loci: nomenclature proposal and SNPSTR haplotypes. Forensic Sci Int Genet 9:61–65. https://doi.org/10.1016/j.fsigen.2013.11.001.
19. Gill P, Brinkmann B, D'Aloja E, Andersen J, Bar W, Carracedo A, Dupuy B, Eriksen B, Jangblad M, Johnsson V, Kloosterman AD, Lincoln P, Morling N, Rand S, Sabatier M, Scheithauer R, Schneider P, Vide MC (1997) Considerations from the European DNA profiling group (EDNAP) concerning STR nomenclature. Forensic Sci Int 87:185–192. https://doi.org/10.1016/S0379-0738(97)00111-4
20. Olaisen B, Bär W, Brinkmann B, Budowle B, Carracedo A, Gill P, Lincoln P, Mayr WR, Rand S (1998) DNA recommendations 1997 of the International Society for Forensic Genetics. Vox Sang 74:61–63
21. Scientific Working Group on DNA Analysis Methods. (2016) SWGDAM Validation Guidelines for DNA Analysis Methods—Approved 12/05/2016. Available at: https://www.swgdam.org/
22. Peakall R, Smouse PE (2012) GenALEx 6.5: genetic analysis in excel. Population genetic software for teaching and research-an update. Bioinformatics 28:2537–2539. https://doi.org/10.1093/bioinformatics/bts460
23. Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in excel. Population genetic software for teaching and research. Mol Ecol Notes 6:288–295. https://doi.org/10.1111/j.1471-8286.2005.01155.x
24. Nei M (1972) Genetic distance between populations. Amer Nat 106:283–292
25. Brenner C, Morris J (1990) Paternity index calculations in single locus hypervariable DNA probes: validation and other studies. Proc Int Symp Hum Identif 1989:21–53
26. Akram NA, Farooqi SR (2012) DNASF: a statistical package to analyze the distribution and polymorphism of CODIS STR loci in a heterogeneous population. J Forensic Res 3:1–6. https://doi.org/10.4172/2157-7145.1000170
27. Budowle B, Onorato AJ, Callaghan TF, Della MA, Gross AM, Guerrieri RA, Luttman JC, Mcclure DL (2009) Mixture interpretation: defining the relevant features for guidelines for the assessment of mixed dna profiles in forensic casework. J Forensic Sci 54:810–821. https://doi.org/10.1111/j.1556-4029.2009.01046.x
28. Pasquale AA De, Tumino G, Pasquale RC De, Pasquale RC De (1974) Micromorphology of the epidermic surfaces of female plants of *Cannabis* sativa L. Bull Narcotics 26: 27–40.
29. Metz SG, Sharma HC, Armstrong TA, Mascia PN (1988) Chromosome doubling and aneuploidy in anther-derived plants from two winter wheat lines. Genome 30:177–181
30. Sharma CB (1990) Chemically induced aneuploidy in higher plants. Mutagenesis 5:105–125. https://doi.org/10.1093/mutage/5.2.105
31. Adams KL, Wendel JF (2005) Polyploidy and genome evolution in plants. Curr Opin Plant Biol 8:135–141. https://doi.org/10.1016/j.pbi.2005.01.001
32. Bagheri M, Mansouri H (2015) Effect of induced polyploidy on some biochemical parameters in Cannabis sativa L. Appl Biochem Biotechnol 175:2366–2375. https://doi.org/10.1007/s12010-014-1435-8
33. Knight G, Hansen S, Connor M, Poulsen H, McGovern C, Stacey J (2010) The results of an experimental indoor hydroponic *Cannabis* growing study, using the 'Screen of Green' (ScrOG) method—yield, tetrahydrocannabinol (THC) and DNA analysis. Forensic Sci Int 202:36–44. https://doi.org/10.1016/j.forsciint.2010.04.022
34. Hsieh HM, Hou RJ, Tsai LC, Wei CS, Liu SW, Huang LH, Kuo YC, Linacre A, Lee JCI (2003) A highly polymorphic STR locus in *Cannabis sativa*. Forensic Sci Int 131:53–58. https://doi.org/10.1016/S0379-0738(02)00395-X
35. Shirley N, Allgeier L, Lanier T, Coyle HM (2013) Analysis of the NMI01 marker for a population database of *Cannabis* seeds. J Forensic Sci 58:176–182. https://doi.org/10.1111/1556-4029.12005
36. Shibuya EK, Sarkis JES, Negrini-Neto O, Martinelli LA (2007) Carbon and nitrogen stable isotopes as indicative of geographical origin of marijuana samples seized in the city of São Paulo (Brazil). Forensic Sci Int 167:8–15. https://doi.org/10.1016/j.forsciint.2006.06.002
37. Houston R, Birck M, Hughes-Stamm S, Gangitano D (2017) Developmental and internal validation of a novel 13 loci STR multiplex method for *Cannabis sativa* DNA profiling. Legal Med 26:33–40. https://doi.org/10.1016/j.legalmed.2017.03.001
38. Hair Jr JF, Black WC, Babin BJ, Anderson RE, Tatham RL (2009) *Análise multivariada de dados* (6th ed.). Porto Alegre: Book
39. Szczerbowska-Boruchowska M, Lankosz M, Adamek D (2011) First step toward the "fingerprinting" of brain tumors based on synchrotron radiation X-ray fluorescence and multiple discriminant analysis. J Biol Inorg Chem 16:1217–1226. https://doi.org/10.1007/s00775-011-0810-y
40. Valverde L, Lischka C, Erlemann S, De Meijer E, De Pancorbo MM, Pfeiffer H, Köhnemann S (2014) Nomenclature proposal and SNPSTR haplotypes for 7 new *Cannabis sativa* L. STR loci. Forensic Sci Int Genet 13:185–186. https://doi.org/10.1016/j.fsigen.2014.08.002