



The differentiation of fibre- and drug type Cannabis seedlings by gas chromatography/mass spectrometry and chemometric tools

Julian Broséus, Frédéric Anglada¹, Pierre Esseiva^{*}

Institut de Police Scientifique, School of Criminal Sciences, Batochime, University of Lausanne, 1015 Lausanne-Dorigny, Switzerland

ARTICLE INFO

Article history:

Received 11 December 2009
Received in revised form 1 March 2010
Accepted 20 March 2010
Available online 24 April 2010

Keywords:

Cannabis
Cannabis chemotype
Gas chromatography–mass spectrometry
Support vector machines

ABSTRACT

Cannabis cultivation in order to produce drugs is forbidden in Switzerland. Thus, law enforcement authorities regularly ask forensic laboratories to determinate cannabis plant's chemotype from seized material in order to ascertain that the plantation is legal or not. As required by the EU official analysis protocol the THC rate of cannabis is measured from the flowers at maturity. When laboratories are confronted to seedlings, they have to lead the plant to maturity, meaning a time consuming and costly procedure. This study investigated the discrimination of fibre type from drug type Cannabis seedlings by analysing the compounds found in their leaves and using chemometrics tools. 11 legal varieties allowed by the Swiss Federal Office for Agriculture and 13 illegal ones were greenhouse grown and analysed using a gas chromatograph interfaced with a mass spectrometer. Compounds that show high discrimination capabilities in the seedlings have been identified and a support vector machines (SVMs) analysis was used to classify the cannabis samples. The overall set of samples shows a classification rate above 99% with false positive rates less than 2%. This model allows then discrimination between fibre and drug type Cannabis at an early stage of growth. Therefore it is not necessary to wait plants' maturity to quantify their amount of THC in order to determine their chemotype. This procedure could be used for the control of legal (fibre type) and illegal (drug type) Cannabis production.

© 2010 Elsevier Ireland Ltd. All rights reserved.

1. Introduction

More than 400 chemical compounds have been isolated in *Cannabis sativa* L. of which more than 60 are cannabinoids. Cannabinoids are a characteristic class of substances unique to *Cannabis*. The most abundant are Δ^9 -tetrahydrocannabinol (Δ^9 -THC or THC, the main psychoactive cannabinoid), cannabidiol (CBD), cannabigerol (CBG) and cannabichromene (CBC). In fresh plant material, most of cannabinoids exist in the form of their acids precursors. For example, in fresh plant material, 95% of THC and CBD consist of their precursors THCA-A and CBDA, respectively [1]. They are converted (by a decarboxylation) to their corresponding neutral cannabinoids in course of time, upon heating or under alkaline conditions. In this study we perform this decarboxylation process directly in the injector (in the liner) of the GC–MS [2] used for the analysis of the samples. In addition to cannabinoids, many constituents commonly encountered in nature like terpenes, alkanes, flavonoids and nitrogenous compounds are also present [3,4].

Cannabis sativa subspecies are divided into several chemical phenotypes. The relative proportions of THC, CBN and CBD have been used by various authors for distinguishing three predominant chemotypes; chemotypes I, II and III. Small and Beckstead [5] use the proportion of THC and CBD for classifying cannabis samples. For example, the main cannabinoid of the fibre type is CBD and the one of the drug type is THC. The following table summarizes the criteria used to discriminate chemotypes (Table 1).

The United Nations Office on Drugs and Crime [6] proposes to use a combination of three cannabinoids in the following index:

$$X = \frac{[\text{THC}] + [\text{CBN}]}{[\text{CBD}]}$$

with [THC], [CBN] and [CBD] being the relative proportion (peak area) of respectively THC, CBN and CBD in the chromatogram. Cannabinol (CBN) is not a natural cannabinoid, it is a degradation product of THC. If this ratio is higher than 1, the plant is classified as drug type cannabis or as chemotype I, while if index is lower than 1, the plant is classified as fibre type cannabis or as chemotype III. A study performed in Greece [7] highlighted the difficulty for differentiating drug from fibre type chemotypes with an error rate of classification of about 20% depending on index used.

This chemotype classification is mainly carried out with mature cannabis and few researches have been focused on the possibility to attribute a chemotype using cannabis seedlings. Previous

^{*} Corresponding author. Tel.: +41 21 692 46 52; fax: +41 21 692 46 05.

E-mail address: pierre.esseiva@unil.ch (P. Esseiva).

¹ We dedicate this article to the memory of our friend Frédéric Anglada, who unfortunately passed away on the 1st January 2010.

Table 1
Predominant Cannabis chemotypes according to THC and CBD content [5].

Chemotype	Designation products	Predominant cannabinoids	THC to CBD ratio
I: Drug type	Illegal hemp, resinous Cannabis Marijuana, hashish	THC	High ratio THC content >0.30% ^a CBD content <0.50% ^a
II: Intermediate type		THC, CBD	Ratio near 1 (0.5–2)
III: Fibre type	Legal hemp, industrial (textile) cannabis, Edible seed oil, essential oils, cosmetics, lubricants, fuels, fibres	CBD	Low ratio THC < 0.30% ^a

^a Dry weight of the reproductive part of the plant at flowering.

studies showed that the chemotype does not change in plants at different ages or in different sexes throughout the entire life cycle until flowering (even if the cannabinoids quantities are not constant). In fact, THC in chemotype I and CBD in chemotype III plants become the predominant cannabinoids very early (since 1 month). According to cannabinoids amounts and their ratio, it's possible to analyse Cannabis leaves at a very early phase of growth and to classify their chemotype [8,9].

Cannabis cultivation in order to produce illicit drugs is forbidden in several countries, of which Switzerland. However, fibre type cannabis cultivation is legal in some countries, under restrictions. For example, European Union requires that seeds intended for sowing of fibre cannabis varieties show in the mature plant a THC amount lower than 0.2% (w/w) [10] while Switzerland requires a THC amount lower than 0.3% (w/w) [11]. Thus, law enforcement authorities regularly ask forensic laboratories to determinate cannabis plant's chemotype from seized material in order to ascertain if the plantation is legal or not.

We propose then to develop a GC–MS method combined with a chemometric approach to discriminate drug type from fibre type cannabis based on the analysis of the relative proportions of the major compounds found in cannabis seedlings' leaves. Nowadays, this question is of utmost importance for law enforcement authorities in Switzerland, given that persons cultivating cannabis always advocate that the plants are fibre type and then legal. In such situations, the prosecutors ask forensic laboratories to provide a statement about the chemotype of the plant. When the cannabis is mature this problem is solved by measuring the THC concentration from the flowers. However, when laboratories are dealing with seedlings they have to lead the plant to maturity, as required by the EU official cannabis analysis protocol [12], resulting in a time consuming and costly procedure.

2. Materials and methods

2.1. Sampling and growth conditions

Seeds from 11 Cannabis varieties approved by the Swiss Federal Office for Agriculture [13] were purchased from the Fédération Nationale des Producteurs de Chanvre (Le Mans, France) (fibre type, chemotype III). Seeds from 13 illegal ones were obtained in particular websites (drug type; chemotype I). Cannabis varieties for each chemotype are listed in Table 2.

In order to determine the chemical profile of each seedling, 15 plants of each variety were greenhouse grown under the following environmental conditions: the temperature varied between 20 °C and 22 °C during the growth of the plants, relative humidity was fixed to 50% automatically and the light/dark photoperiod was hold for 15 and 9 h, respectively, using artificial lighting when necessary.

Table 2
Cannabis varieties used for chemotype determination of drug type and fibre type Cannabis [13].

Chemotype	Varieties
I: 13 drug type cannabis varieties	Afghan (A), AK47 (AK47), bubble gum (BG), chronic (CH), fraise sativa (FS), granflora (G), ice (I), maple leaf (ML), monster bloom (MB), northern light (NL), red diesel (RD), skyflight (S) and top 44 (T44)
III: 11 fibre type cannabis varieties	Beniko (BEN), epsilon 68 (EPS), fasamo (FAS), fedora 17 (FED 17), felina 32 (FEL), férimon (FER), futura 75 (FUT), lovrin 110 (LOV), kompolti (KOM), uniko B (UB) and USO31 (USO31)

After 28 days of growth, leaves were picked up from each plant of each variety. Leaves were dried at room temperature and powdered. Three samples by plant (100 mg each) were individually extracted with 5 mL of hexane containing 35 mg squalane per 100 mL as internal standard (IS). The tube containing the sample and extraction solution was placed in an ultrasonic water bath for 15 min and then in a rotating beater for 60 min. After filtration, 1 mL of the solution is then transferred into a vial. Each sample is then individually analysed by GC–MS to separate and to identify the compounds.

2.2. Analytical methods

An Agilent 7890A gas chromatograph interfaced with an Agilent 5975C mass selective detector was used for the analysis. Separation was accomplished on a HP-5ms capillary column (30 m length, 0.25 mm i.d. and 0.25 µm film thickness). Injections were carried out in split mode using a general purpose split/splitless liner packed with glass wool (Agilent Technologies No. 5183-4711). The temperature program starts at 100 °C, increases to 260 °C (at 10 °C/min) and holds for 10 min for a total run of 26 min. 2 µL of each sample was injected with helium as carrier gas (constant flow mode, 1 mL/min) using a split ratio of 1:20. Temperatures applied are 280 °C for injector, 250 °C for transfer line, 230 °C for ion source and 150 °C for quadrupole. Electron multiplier voltage was fixed to 1976 V. Data were acquired in the full scan mode (30–450 *m/z* mass range) with a sampling rate of 3 (1.77 scans/s) and were analysed using MSD Enhanced ChemStation v. D.02.00.275 (Agilent Technologies).

Compounds found in samples were identified using both mass spectrum computerized databases, such as WILEY and NIST05, and data coming from literature [14,15].

2.3. Chemometrics tools

2.3.1. Software

Data processing was performed using Microsoft Excel 2003 and R version 2.9.0 in which Support Vector Machines (SVMs) classification was performed using the package e1071.

2.3.2. Data processing

Areas of targets ions in the chromatograms were integrated for each identified compound. Peaks areas were normalized to the internal standard and the square root was performed in order to reduce the influence of larger peaks and thus to have the variables on a comparable scale. Then, data are scaled to zero mean and unit variance [16].

2.3.3. Classification methods: LDA and SVMs

The objective of the methodology is to classify a cannabis sample into drug or fibre type Cannabis. As these two classes could be linearly or non-linearly separable, two classification methods respectively managing these questions have been tested. The potential of linear discriminant analysis (LDA) and SVMs to predict the belonging of a plant to chemotype I (drug type) or chemotype III (fibre type) has been then evaluated. These two methods try to find optimal boundaries between classes in order to discriminate them. LDA produces a linear decision boundary between two classes while SVMs models can be used for complex non-linear situations using an appropriate kernel function. The aim of the SVMs is to find the unique separating hyperplane having the maximum margin of separation between the two classes. A peculiar characteristic of SVMs is that datapoints which lie on the

margins – named support vectors – determine the solution of the classification problem [17]. The best kernel is selected after experimentation. This step is necessary as no theoretical tools exist to determine which kernel performs the best for a given dataset. In our case, the radial basis function (RBF) kernel was an excellent alternative due to its only one hyperparameter to optimize [18], γ , leading to a reasonable complexity of the SVMs model as pointed out in the RBF equation:

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2) \quad (1)$$

with $\gamma > 0$ the parameter which determines the width of the RBF and x_i, x_j the vectors of the i th and j th training samples. As explained in [17], the complexity of the SVMs model can be controlled with an additional parameter, C . This penalty parameter controls the trade-off between allowing training errors (low value of C leads to a simpler prediction function) and looking for a more accurate model (high value of C leads to a complex prediction function). C can be adjusted by the user or automatically optimized using some SVMs packages. For a given dataset, γ and C have to be optimized in order to obtain best performance results. Detailed explanation can be found in [17–20] for SVMs and [21,22] for LDA methods.

2.3.4. Optimisation

While there is no requirement to optimize LDA model, it is necessary for SVMs [23]. In order to find the most performing SVMs model we have determined the optimal γ and C values with the function `tune.svm` in `R` (library `e1071`). This function carries out a grid search over specified ranges of γ and C . The values ranges tested were 0.001–100 for γ and 0.1–1000 for C . The best performance is computed using a 10-fold cross validation on the whole dataset. The parameter combination with the best performance is chosen and then a suitable SVMs model can be trained using these optimal parameters [20].

2.3.5. Assessing the model quality

To assess the quality of the model, cannabis samples were divided into two groups: the training samples on which the model is built, and the test samples on which the predictive capability is tested. The class membership is known for the two groups. The classification rate for each model is then calculated by measuring the percentages of samples correctly classified for training and test samples. The percentage of samples incorrectly classified for each class was also evaluated. They have been named as false positive drug (FPD) and false positive fibre (FPF). False positive drug is defined as the percentage of samples classified as drug type cannabis whereas they are fibre type cannabis. False positive fibre is defined as the percentage of samples classified as fibre type cannabis whereas they are drug type cannabis. To ensure that the results are well representative of the model classification ability and do not depend on the samples used, the separation into training and test sets is repeated 1000 times. Samples are randomly chosen each time. Classification measures are recorded over all 1000 iterations and the reported values are mean values (cf. Fig. 1). Finally, after the creation of the model a validation test is performed with an independent validation set. The latter consists of 44 cannabis seedlings coming from police seizures where growing conditions are

DEVELOPMENT OF CLASSIFICATION MODELS

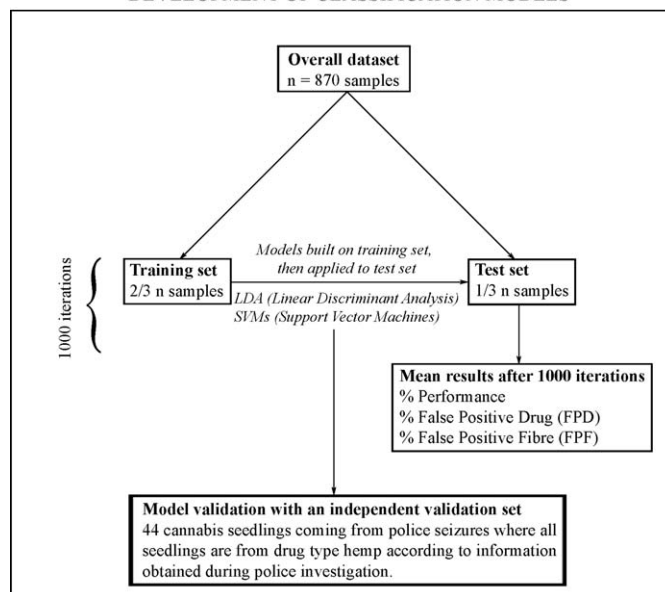


Fig. 1. Statistical methodology applied during this work.

not known but likely different from ours. All seedlings are from drug type cannabis according to information obtained during police investigation. Indeed, Cannabis seedlings were labelled by raisers using the name of their variety. Furthermore, defendants made a full confession and listed all Cannabis varieties that they cultivated.

3. Results and discussion

In this study, we have focused our attention on the overall set of compounds presents at detectable amounts in the leaves of cannabis seedlings. Fifteen compounds have been selected taking into account their presence in drug and fibre type seedlings cannabis. Fig. 2 shows a typical chromatogram obtained for a cannabis seedling analysis. Table 3 shows compounds retention times (relative to squalane) as well as their specific target ions and qualifiers.

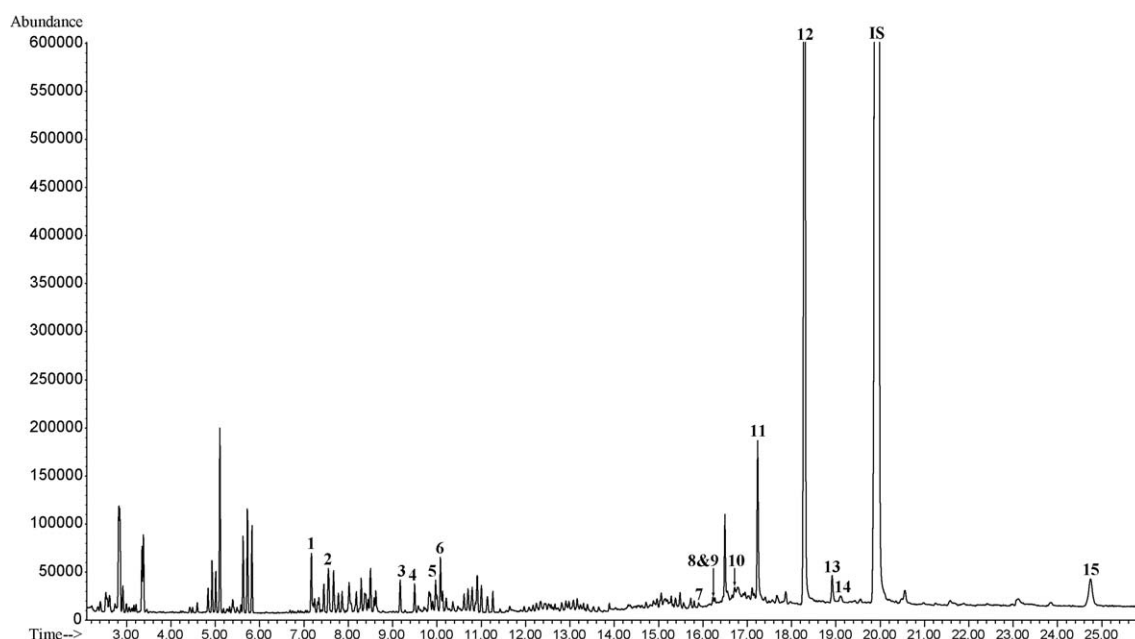


Fig. 2. Total ion current (TIC) chromatogram of a typical drug type cannabis sample. Number on the chromatogram refers to Table 3.

Table 3

Compounds found in leaves, their relative retention time (RRT), target ions and qualifiers.

Peak numbers	Compounds	RRT (min)	Target ion, qualifiers (<i>m/z</i>)
1	Beta caryophyllene	0.360	93 , 133, 41, 69
2	Alpha caryophyllene	0.379	93 , 80, 121, 147
3	Guaiol	0.460	161 , 107, 59, 93
4	γ -Eudesmol	0.476	189 , 161, 204, 133
5	Bulnesol	0.500	107 , 135, 93, 161
6	α -Bisabolol	0.506	109 , 119, 69, 43
7	Cannabinoid 1	0.797	231 , 314, 299, 271
8	Tetrahydrocannabivarin (THV)	0.813	271 , 286, 203, 243
9	Cannabinoid 2	0.823	231 , 314, 174, 243
10	Cannabicyclol (CBL)	0.839	231 , 232, 174, 314
11	Cannabidiol (CBD)	0.864	231 , 174, 314, 299
12	Δ^9 -THC (THC)	0.918	299 , 314, 231, 271
13	Cannabigerol (CBG)	0.949	193 , 231, 123, 316
14	Cannabinol (CBN)	0.959	295 , 238, 310, 223
15	Nonacosane	1.241	57 , 71, 85, 43

In this work discrimination between fibre type and drug type cannabis using $([THC] + [CBN])/[CBD]$ ratio leads to the misclassification of fibre type (0.6% of fibre type samples misclassified) as well as drug type cannabis (7.5% of drug type samples misclassified) as illustrated in Fig. 3. Drug type samples which are misclassified are coming in the majority from one variety.

Therefore, it is necessary to select compounds allowing a better differentiation between fibre type and drug type cannabis. Principal component analysis (PCA) and boxplots were then performed [21]. Eight compounds having high discrimination capabilities have been selected: they are guaiol, γ -eudesmol, bulnesol, α -bisabolol, THV, CBD, THC and CBN. With these compounds a PCA has been performed in order to visualize the representation of each sample using the scores plot against the first principal component (PC1) and second principal component (PC2). Fig. 4 illustrates the scores of cannabis samples against the first two principal components. Some fibre type and drug type cannabis samples are non-linearly separable leading to an overlapping area (circle in Fig. 4). Results for the classification performed with LDA and SVMs models are shown in Table 4. The high FPF rate observed for LDA model highlights the difficulty to separate cannabis samples with a linear

classifier. In order to decrease this value it has been decided to use a non-linear algorithm that is implemented with the SVMs model. SVMs performed much better than LDA with a classification performance above 99% over 1000 iterations. Furthermore, as expected, the FPF rate has dramatically decreased. SVMs used 122 support vectors to create the classification model, 95 in drug type samples and 27 in fibre type samples. The hyperparameters for RBF kernel were 1 for γ and 100 for C .

The main idea of the model was to maximise the classification of fibre type cannabis (minimise the false positive drug rate), in such a way that a person who legally grow Cannabis would not be judicially prosecuted by error. The low false positive rates observed are in agreement with this concept. As plants have been cultivated in the same environmental conditions, variations in cannabis samples' composition are mainly due to genetics characteristics (the seed). These variations have been well evaluated by the SVMs model given that discrimination between fibre type and drug type cannabis has been observed using selected compounds.

Furthermore, 100% of the validation set has been correctly classified and thus samples predicted as drug type cannabis by the SVMs model. As far as LDA, about 11% of the samples were erroneously classified as fibre type cannabis. SVMs model's ability to correctly classify the validation set demonstrates that the model can handle the case when new samples have grown under different environmental conditions. Therefore genetics characteristics are more important factors than environmental ones. Moreover, it confirms that the SVMs classification power is better than the linear algorithm one.

As presented in a recent study the chemotype does not change in plants at different ages or in different sexes throughout the entire life cycle until flowering [8,24]. Thus, the SVMs model that has been built can be used to discriminate between fibre and drug type cannabis at an early stage of growth. Consequently, it is not necessary to cultivate during several months seized seedlings to sample their flowering portions and to quantify the amount of THC. This actual procedure is time consuming (time of growth) and expensive (cost of the standard used for quantification, cost of a secured greenhouse and cost of materials required for culture). For chemotype identification, sampling leaves at an early stage of development and the determination of the relative amount of guaiol, γ -eudesmol, bulnesol, α -bisabolol, THV, CBD, THC and CBN

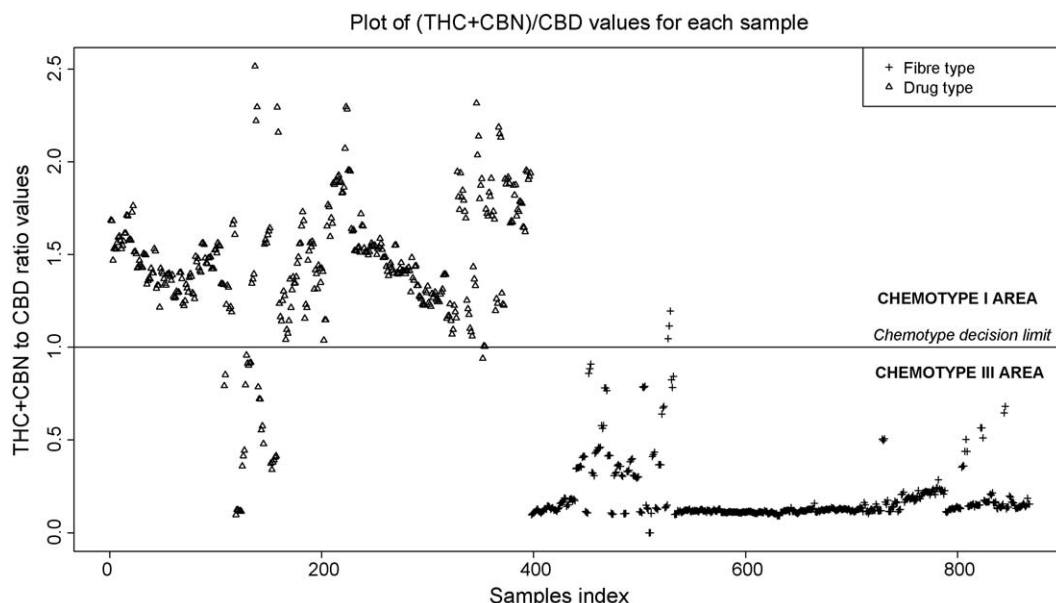


Fig. 3. Plot of Δ^9 -tetrahydrocannabinol (THC) + cannabinol (CBN) to cannabidiol (CBD) ratio values for each cannabis sample.

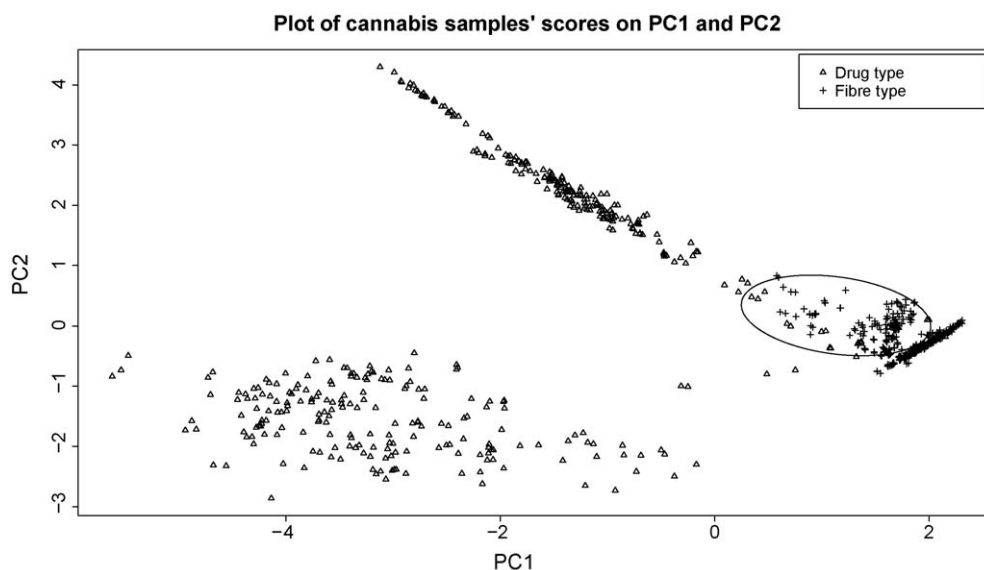


Fig. 4. Plot of cannabis samples scores against the first principal component (PC1) and second principal component (PC2).

Table 4

Results for LDA and SVMs showing mean values of the classification measures over 1000 iterations with the standard deviation in brackets.

	Training set			Test set		
	%Performance	%FPD	%FPF	%Performance	%FPD	%FPF
LDA	97.3 (± 0.4)	0.3 (± 0.2)	5.5 (± 0.9)	97.0 (± 1.0)	0.3 (± 0.5)	6.0 (± 2.3)
SVMs	99.7 (± 0.1)	0.0	0.7 (± 0.3)	99.2 (± 0.7)	0.3 (± 0.8)	1.3 (± 1.2)

would be a potential modification of the EU official cannabis analysis protocol. The present method could be applied to the production control of non-scheduled (fibre type) and scheduled (drug type) Cannabis. In addition, an option to confirm the results obtained by the SVMs model is the quantification of THC once the seedling has reached the stage of maturity and florescence.

4. Conclusions

In this work an original Cannabis classification model has been developed allowing the discrimination between fibre- and drug type Cannabis seedlings by GC/MS based on non-cannabinoid and cannabinoid leaf constituents. Using selected compounds the SVMs model correctly classifies samples with a classification performance above 99% and with low false positive rates (less than 2%). The GC–MS method is consequently valid, fast and cheap regarding the classification of a cannabis seedling in comparison to the EU official procedure, i.e. cultivation of seedlings to mature plants. Thus, it is not necessary to wait for the plant's florescence, to sample their flowering portions and quantify the amount of THC – the actual time consuming and expensive procedure – in order to determine their chemotype. The SVMs model that has been built can be used to discriminate between fibre and drug type cannabis at an early stage of growth (1 month of growth in this work).

Chemotype identification by sampling leaves at an early stage of development and determination of the relative amount of guaiol, γ -eudesmol, bulnesol, α -bisabolol, THV, CBD, THC and CBN is a valuable tool that could be included in EU official cannabis analysis protocol.

Acknowledgments

The authors would like to thank Daniel Alvarez, Cédric Liengme, Kevin Rudaz and Blaise Tissot-Dit-Sanfin from the département de

biologie moléculaire végétale of the University of Lausanne for the technical support.

References

- [1] C.E. Turner, M.A. Elsohly, E.G. Boeren, Constituents of Cannabis-Sativa L. 17. A review of the natural constituents, *Journal of Natural Products* 43 (2) (1980) 169–234.
- [2] F.E. Dussy, C. Hamberg, M. Luginbühl, T. Schwerzmann, T.A. Briellmann, Isolation of [Delta]9-THCA-A from hemp and analytical aspects concerning the determination of [Delta]9-THC in cannabis products, *Forensic Science International* 149 (1) (2005) 3–10.
- [3] R. Brenneisen, Chemistry and analysis of phytocannabinoids and other cannabis constituents, in: M.A. Elsohly (Ed.), *Marijuana and the Cannabinoids*, Humana Press, Totowa, NY, 2007, pp. 17–49.
- [4] M.A. Elsohly, D. Slade, Chemical constituents of marijuana: the complex mixture of natural cannabinoids, *Life Sciences* 78 (5) (2005) 539–548.
- [5] E. Small, H. Beckstead, Common cannabinoid phenotypes in 350 stocks of Cannabis, *Lloydia-The Journal of Natural Products* 36 (2) (1973) 144–165.
- [6] United Nations Office on Drugs and Crime (UNODC), Recommended Methods for the Identification and Analysis of Cannabis and Cannabis Products, (2009) <http://www.unodc.org/documents/scientific/ST-NAR-40-Ebook.pdf> (retrieved October 2009).
- [7] M. Stefanidou, A. Dona, S. Athanasis, I. Papoutsis, A. Koutselinis, The cannabinoid content of marihuana samples seized in Greece and its forensic application, *Forensic Science International* 95 (2) (1998) 153–162.
- [8] D. Pacifico, F. Miselli, A. Carboni, A. Moschella, G. Mandolino, Time course of cannabinoid accumulation and chemotype development during the growth of *Cannabis sativa* L., *Euphytica* 160 (2) (2008) 231–240.
- [9] A.F. Vogelmann, J.C. Turner, P.G. Mahlberg, Cannabinoid composition in seedlings compared to adult plants of Cannabis-Sativa, *Journal of Natural Products* 51 (6) (1988) 1075–1079.
- [10] Official Journal of the European Communities of 29 February 2000 (C 56 E/18), Commission Regulation (EC) modifying regulation EC n°1251/1999, <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:C:2000:056E:0017:0018:FR:PDF> (retrieved October 2009).
- [11] Ordonnance du Département Fédéral de l'Economie sur les semences et les plants des espèces de grandes cultures et de plantes fourragères, p. 59, <http://www.adm.ch/ch/fr/rs/9/916.151.1.fr.pdf> (retrieved October 2009).
- [12] Official Journal of the European Communities of 21 February 2002 (L 51/21), Commission Regulation (EC) No. 327/2002, <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2002:051:0014:0022:FR:PDF> (retrieved October 2009).

- [13] Ordonnance de l'Office Fédéral de l'Agriculture sur le catalogue des variétés de céréales, de pommes de terre, de plantes fourragères, de plantes oléagineuses et à fibres ainsi que de betteraves, RS 916.151.6, <http://www.admin.ch/ch/f/as/2009/2441.pdf> (retrieved October 2009).
- [14] A. Hazekamp, A. Peltenburg, R. Verpoorte, C. Giroud, Chromatographic and spectroscopic data of cannabinoids from *Cannabis sativa* L., *Journal of Liquid Chromatography & Related Technologies* 28 (15) (2005) 2361–2382.
- [15] R. Lewis, S. Ward, R. Johnson, D.T. Burns, Distribution of the principal cannabinoids within bars of compressed cannabis resin, *Analytica Chimica Acta* 538 (1/2) (2005) 399–405.
- [16] C. Reimann, P. Filzmoser, R. Garrett, R. Dutter, *Statistical Data Analysis Explained: Applied Environmental Statistics with R*, John Wiley & Sons, Chichester, England/Hoboken, NJ, 2008.
- [17] O. Ivanciuc, Applications of support vector machines in chemistry, in: K.B. Lipkowitz, T.R. Cundari (Eds.), *Reviews in Computational Chemistry*, vol. 23, Wiley-VCH, Weinheim, 2007, pp. 291–400.
- [18] C.W. Hsu, C.C. Chang, C.J. Lin, *A Practical Guide to Support Vector Classification*, Department of Computer Science and Information Engineering, National Taiwan University, 2003.
- [19] H.D. Li, Y.Z. Liang, Q.S. Xu, Support vector machines and its applications in chemistry, *Chemometrics and Intelligent Laboratory Systems* 95 (2) (2009) 188–198.
- [20] A. Karatzoglou, D. Meyer, K. Hornik, Support vector machines in R, *Journal of Statistical Software* 15 (9) (2006).
- [21] D.L. Massart, B.G.M. Vandeginste, L.M.C. Buydens, S. De Jong, P.J. Lewi, J. Smeyers-Verbeke, *Handbook of Chemometrics and Qualimetrics: Part A*, Elsevier, Amsterdam, 1997.
- [22] B.G.M. Vandeginste, D.L. Massart, L.M.C. Buydens, S. De Jong, P.J. Lewi, J. Smeyers-Verbeke, *Handbook of Chemometrics and Qualimetrics: Part B*, Elsevier, Amsterdam, 1998.
- [23] S.J. Dixon, R.G. Brereton, Comparison of performance of five common classifiers represented as boundary methods: Euclidean Distance to Centroids, Linear Discriminant Analysis, Quadratic Discriminant Analysis, Learning Vector Quantization and Support Vector Machines, as dependent on data structure, *Chemometrics and Intelligent Laboratory Systems* 95 (1) (2009) 1–17.
- [24] G. Fournier, J. Bausset, O. Beherec, M. Desvals, S. Bertucelli, The simplification of the industrial hemp crops control is possible, *Annales de Toxicologie Analytique* 19 (3) (2007) 201–209.