



ORIGINAL RESEARCH

Absorbance-Transmittance Excitation Emission Matrix Method for Quantification of Major Cannabinoids and Corresponding Acids: A Rapid Alternative to Chromatography for Rapid Chemotype Discrimination of *Cannabis sativa* Varieties

Adam M. Gilmore,¹ Mostafa A. Elhendawy,^{2,3} Mohamed M. Radwan,⁴ Linda H. Kidder,^{1,*} Amira S. Wanas,^{4,5,*} Murrell Godfrey,² Jana B. Hildreth,¹ A. Edward Robinson,¹ and Mahmoud A. ElSohly^{4,6}

Abstract

Background: Phytocannabinoids naturally occur in the cannabis plant (*Cannabis sativa*), and Δ^9 -tetrahydrocannabinol (THC) and cannabidiol (CBD) predominate. There is a need for rapid inexpensive methods to quantify total THC (for statutory definition) and THC–CBD ratio (for classification into three chemotypes). This study explores the capabilities of a spectroscopic technique that combines ultraviolet-visible and fluorescence, absorbance-transmittance excitation emission matrix (A-TEEM).

Methods: The A-TEEM technique classifies 49 dry flower extracts into three *C. sativa* chemotypes, and quantifies the total THC–CBD ratio, using validated gas chromatography (GC)-flame ionization (FID) and High-Performance Liquid Chromatography (HPLC) methods for reference. Multivariate methods used are principal components analysis for a chemotype classification, extreme gradient boost (XGB) discriminant analysis (DA) to classify unknown samples by chemotype, and XGB regression to quantify total THC and CBD content using GC-FID and HPLC data on the same samples.

Results: The A-TEEM technique provides robust classification of *C. sativa* samples, predicting chemotype classification, defined by THC–CBD content, of unknown samples with 100% accuracy. In addition, A-TEEM can quantify total THC and CBD levels relevant to statutory determination, with limit of quantifications (LOQs) of 0.061% (THC) and 0.059% (CBD), and high cross-validation (>0.99) and prediction (>0.99), using a GC-FID method for reference data; and LOQs of 0.026% (THC) and 0.080% (CBD) with high cross-validation (>0.98) and prediction (>0.98), using an HPLC method for reference data. A-TEEM is highly predictive in separately quantifying acid and neutral forms of THC and CBD with HPLC reference data.

Conclusions: The A-TEEM technique provides a sensitive method for the qualitative and quantitative characterization of the major cannabinoids in solution, with LOQs comparable with GC-FID and HPLC, and high values of cross-validation and prediction. As a spectroscopic technique, it is rapid, with data acquisition <45 sec per measurement; sample preparation is simple, requiring only solvent extraction. A-TEEM has the sensitivity to resolve

¹HORIBA Instruments, Inc., Piscataway, New Jersey, USA.

²Department of Chemistry and Biochemistry, University of Mississippi, University, Mississippi, USA.

³Department of Agriculture Chemistry, Faculty of Agriculture, Damietta University, Damietta, Egypt.

⁴National Center for Natural Products Research, University of Mississippi, University, Mississippi, USA.

⁵Department of Pharmacognosy, Faculty of Pharmacy, Minia University, Minia, Egypt.

⁶Department of Pharmaceutics and Drug Delivery, University of Mississippi, University, Mississippi, USA.

***Correction added** on August 26, 2022 after first online publication of April 29, 2022: The affiliation *Department of Pharmacognosy, Faculty of Pharmacy, Minia University, Minia, Egypt* has been added for author Dr. Amira S. Wanas.

*Address correspondence to: Linda H. Kidder, PhD, HORIBA Instruments, Inc., 20 Knightsbridge Road, Piscataway, NJ 08854, USA, E-mail: linda.kidder@horiba.com

and quantify cannabinoids in solution based on their unique spectral characteristics. Discrimination of legal and illegal chemotypes can be rapidly verified using XGB DA, and quantitation of statutory levels of total THC and total CBD comparable with GC-FID and HPLC can be obtained using XBD regression.

Keywords: A-TEEM; cannabinoids; GC; HPLC; fluorescence spectroscopy; fluorescence EEM

Introduction

Cannabis sativa is a dioecious plant that produces > 500 compounds of varying medical importance. The therapeutic properties of cannabis are mainly attributed to cannabinoids,¹ which are mostly concentrated in the female inflorescence (buds) of this species.² Cannabinoids are terpenophenolic compounds unique to cannabis.³ To date, 125 cannabinoids have been identified.⁴

Δ^9 -Tetrahydrocannabinol (THC) and cannabidiol (CBD) are the most well-known cannabinoids for their therapeutic properties.^{5,6} THC has been approved by the Food and Drug Administration (FDA) for the control of nausea and vomiting associated with chemotherapy and for appetite stimulation of AIDS patients suffering from the wasting syndrome. The drug, however, shows other biological activities that lend themselves to possible therapeutic applications, such as in the treatment of glaucoma,⁶ migraine headaches,^{7,8} spasticity,⁹ anxiety,¹⁰ and as an analgesic.⁹ The second major phytocannabinoid, CBD, has attracted much attention for development as a medicine for several conditions.

CBD has long been known for its anticonvulsant activity¹¹ and more recently recognized as a treatment for epilepsy.¹² CBD was also reported to show neuroprotective activity,^{13–17} and has just been shown to prevent entry of live SARS-CoV-2 into cells.¹⁸ Recently, the FDA has approved CBD as a treatment for two rare forms of childhood epilepsy namely Dravet syndrome and Lennox-Gastaut syndrome that are resistant to current drugs.^{19,20}

Although many of the minor cannabinoids are of interest and commonly analyzed,^{4,21} cannabis is classified into three major chemotypes based on the ratio of only THC and CBD, as most recently recommended by the USP expert panel.²² The three chemotypes are defined as (1) THC-dominant, (2) THC-CBD-intermediate, and (3) CBD-dominant.²² THC-dominant chemotype consists of samples in which the ratio of total THC to total CBD is not less than (NLT) 5:1. In the THC-CBD-intermediate chemotype, the ratio of the total THC content to total CBD content is NLT 0.2:1 and not more than (NMT) 5:1. The ratio of the total THC content to total CBD content is NMT 0.2:1 for the CBD-dominant chemotype.²²

A variety of analytical techniques have been developed for the detection and quantification of cannabinoids and other compounds such as the terpenes in the cannabis plant. Most analytical methods for cannabinoids rely on highly specialized and expensive equipment,^{23,24} mainly based on liquid chromatography with tandem mass spectrometry (LC-MS/MS)^{25–28} and gas chromatography (GC),^{29,30} (although other methods are also available, such as electrochemical sensors³¹), High-Performance Liquid Chromatography (HPLC) with photodiode array (PDA) detection,³² high performance thin layer chromatography,³³ ultra-performance liquid chromatography,³² and matrix-assisted laser desorption ionization mass spectrometry.³⁴

GC with flame ionization (GC/FID) or MS detection (GC/MS) is the most commonly used chromatographic methods in the analysis of cannabinoids.^{29,30,35–37} In GC/MS, a standardized electron ionization method is used to fragment the analytes permitting the use of compound libraries for compound identification.

Without derivatization, GC methods provide quantification of the total (the sum of the acidic and neutral) cannabinoids.³⁶ The acidic cannabinoids undergo decarboxylation directly after injection due to the high temperature at the injection port.^{36–38} Thus, acidic cannabinoids cannot be determined by GC unless they are derivatized before analysis.³⁵ Unlike GC, HPLC allows the differentiation between acidic and neutral cannabinoids³⁹ without derivatization before analysis.^{40,41}

Therefore, HPLC provides a more comprehensive chemical report of cannabis samples compared with GC. Recently, a validated LC-MS/MS method for the quantification of 17 cannabinoids in cannabis and hemp has been reported with high sensitivity, specificity, accuracy, precision, and stability during analysis of cannabis and hemp samples and can be adapted to concentration ranges and additional matrices as required.²⁵

Cost is another point of comparison in evaluating methods. Set up, maintenance, and running costs are often considered in selecting analytical techniques, especially in industry settings. The running cost of GC, HPLC, and LC/MS can be quite high, given the expenses associated with the mobile phase, stationary phase, standards,

and waste disposal fees. Moreover, these methods require highly trained personnel to perform accurate and reliable analyses and require high-level laboratory settings.

A variety of vibrational spectroscopic approaches, such as Raman, near infrared (NIR), and Fourier transform infrared (FT-IR), have been proposed to overcome the mentioned limitations of speed and expense of chromatographic measurements, but they present their own limitations, specifically specificity and sensitivity.²⁵ Fluorescence excitation emission matrix (EEM) spectroscopy has sensitivities comparable with chromatography, but its primary limitation has been a lack of specificity, in large part due to the inner filter effect (IFE), which distorts the measured fluorescence spectrum in a concentration-dependent manner, due to reabsorption that occurs at higher concentrations.⁴²

The standard IFE correction protocol uses concentration values from a separate ultraviolet-visible (UV/Vis) measurement to correct the fluorescence signal. Absorbance data recorded on a separate instrument provide an imperfect correction, with the result that an EEM for a single component is often not reproducible.⁴³ This has limited the utility of traditional fluorescence EEM spectroscopy for robust component identification and quantification.

A-TEEM^{TM42} (absorbance, transmittance and fluorescence EEM), is a 2-in-1 technique that combines UV/Vis/NIR with fluorescence spectroscopy in the same instrument. The absorption spectrum used for IFE correction is collected at the same time on the same sample as the fluorescence measurement, resulting in a very reproducible correction. The A-TEEM technique can be used as a robust and extremely sensitive analytical technique, comprising the best attributes of both molecular spectroscopy (speed, low per measurement cost, laboratory-to-line placement) and chromatography (sensitivity and selectivity).

A-TEEM, therefore, has the potential to displace separations methods (HPLC, GC, LCMS, and GCMS) for certain analyses as a simpler and faster analytical tool, with significantly lower per-sample measurement costs. The specific applications where A-TEEM will provide the most benefit are yet to be determined, but its attributes point to many potential applications in food and natural products such as potency, authenticity, and detection of adulteration.

This study describes the development of an A-TEEM method for (1) rapid and robust classification of the three specific chemotypes of *C. sativa* already described; (2) robust quantification of the major cannabinoids relating to statutory levels⁴³ (<0.3% by dry

weight of THC) of total THC and total CBD. The total THC and total CBD levels in the evaluated samples are determined with validated (GC/FID)⁴⁴ and (HPLC)⁴⁵ methods run on the same samples.

A-TEEM data with GC-FID reference data are used to verify chemotype classification and to create and test a multivariate regression model for quantification of major total THC and CBD levels. A-TEEM data with HPLC reference data are used to create and test a multivariate regression model for quantification of the individual neutral and acid (THC and THCA) forms, and neutral and acid (CBD and CBDA) forms. Total THC and total CBD quantification with HPLC are performed through the summation of the acid and neutral components (THC + THCA) and (CBD + CBDA).

Materials and Methods

Dry flower samples

C. sativa plants of the three major chemotypes (THC-dominant, THC-CBD intermediate, and CBD-dominant) were grown at the Marijuana garden, the University of Mississippi. The harvested flower buds were dried and ground before analysis. Dry flower samples from 49 separate plants (44 for calibration, 5 for validation) representative of the three chemotypes were prepared as extracts, for measurement with three different techniques: GC-FID, HPLC, and A-TEEM.

Cannabinoid standards and calibration curves

Standard solutions of pure THC, THCA, CBD, and CBDA were isolated in our laboratory⁴⁶ from plants grown in the medicinal plants garden, with purity of >95%. They were used to calibrate the HPLC and GC methods as previously described.^{44,45}

GC-FID analysis

Dry flower materials (100 mg) were extracted by 3 mL of IS solution [1 mg/mL of 4-Androstene-3,17-dione in CH₃OH-CHCl₃ (9:1)] for 1 h followed by filtration (working solution for GC). Data were collected on 49 samples, all samples were analyzed using a Varian (Palo Alto, CA) 3380 GC equipped with a Varian CP-8400 automatic liquid sampler, dual capillary injectors, and dual flame ionization detectors (GC/FIDs).

The following instrument parameters were employed: air, 30 psi (300 mL/min); hydrogen, 30 psi (30 mL/min); column head pressure, 14 psi (1.0 mL/min); split flow rate, 100 mL/min; split ratio, 50:1; septum purge flow rate, 5 mL/min; makeup gas pressure, 20 psi (30 mL/min); injector temperature, 240°C; detector temperature,

270°C; oven program, 170°C (hold 1 min) to 250°C at 10°C/min (hold 3 min); run time, 12 min; injection volume, 1 μ L, the GC column was J&M Agilent (Santa Clara, CA) (15 $\times$ 0.25 mm DB-1, 0.25 μ m).⁴⁴

HPLC analysis

Dry flower materials (1 g) were extracted with 10 mL of IS solution [1 mg/mL of 4-Androstene-3,17-dione in CH₃OH–CHCl₃ (9:1)] for 1 h followed by filtration then dilution of the extract (100 mg/mL) (3:1) by methanol (working solution for HPLC). Data were collected on 49 samples and HPLC analysis was performed on a Waters (Milford, MA) 2695 separations module (Empower 3 software) connected to a Waters 2996 PDA detector (190–500 nm). The HPLC column was Phenomenex (Torrance, CA), Luna C18(2) (150 mm $\times$ 4.6 mm $\times$ 3 μ m), the mobile phase composed of ACN–H₂O–formic acid 77:23:0.1. The detector is PDA at wavelength 220 nm, injection volume 10 μ L, flow rate 1.2 mL/min, and the run time 20 min.⁴⁵

A-TEEM analysis

Dry flower materials (1 g) were prepared for A-TEEM analysis by extraction with 10 mL of CH₃OH–CHCl₃ (9:1) for 1 h followed by filtration then dilution of the extract (100 mg/mL) to 25 mg/mL with MeOH then further dilution to 0.083 mg/mL (working solution for A-TEEM). All samples were run using the A-TEEM technique on Aqualog-UV-800-C (HORIBA Scientific, Piscataway, NJ). Standards and extracts were measured in 3 mL of 1 cm quartz cuvettes with methanol as the blank at 20°C. The excitation/emission wavelength range was 220–800 nm/250–800 nm with 5 nm intervals and an integration time of 0.25 sec.

Correction for the solvent background, dark detector signals, Rayleigh and Raman scattering, normalization to water Raman scattering area, and inner-filter effects were automatically applied using HORIBA Aqualog v4.2 software. Extracts were prepared in quadruplicate and each sample measurement was repeated three times. For the calibration samples, there was one for which only one replicate per repeat was recorded, for a total of 517 files on 44 unique plant samples. For validation, five unique plant samples were prepared in quadruplicate, with each sample measurement repeated three times. The validation files omitted one sample repeat for a total of 59 files. Table 1 summarizes this information for the calibration and validation of A-TEEM data sets.

Table 1. Summary of Experimental Design for Absorbance-Transmittance Excitation Emission Matrix Measurements

	Calibration samples	Validation samples
Dry flower extract (no. of samples)	44	5
Replicates	4	4
Sample repeats	3	3
Measurements (no. of files)	517 ^a	59 ^b

^aCalibration set contained 1 sample with only 1 replicate per repeat (total 517 files).

^bValidation set omitted one sample repeat due to data anomaly (total 59 files).

Data processing—multivariate analysis and linear regression

All multivariate analyses were performed with Eigen-vector, Inc. (Wenatchee, WA) Solo v8.8. The A-TEEM analysis in Solo started with unfolding the EEM data into a two-way array for both principal components analysis (PCA) and extreme gradient boost (XGB) discrimination and regression; unfolded data were further preprocessed using automatic clutter removal, centering, and scaling. For regression analysis, the concentration data were mean centered. The validation (prediction) data included only unique samples and their A-TEEM replicates compared with the calibration data. Linear regression was subsequently performed on the combined calibration and validation data, using HORIBA Aqualog v4.2 software.

Classification is performed with two different techniques: (1) a PCA for a simple classification based on spectral variance and (2) XGB discriminant analysis to generate a method that can be used to classify unknown samples, with GC-FID results used as reference data. Two quantitative methods to determine total THC and total CBD using A-TEEM are developed, one with GC-FID reference data and the other with HPLC reference data.

Table 2. Cannabidiol (CBD) and Tetrahydrocannabinol (THC) Concentrations as Determined with Gas Chromatography-Flame Ionization Method as Reference (Dry weight %), and the THC–CBD Concentration Ratios Defining the Class Groupings for THC-Dominant Chemotype, THC–CBD Intermediate Chemotype, and CBD-Dominant Chemotype

	Results
max (CBD), wt %	6.49
min (CBD), wt%	0
max (THC), wt%	7.02
min (THC), wt %	0.05
THC-dominant chemotype	THC–CBD $\geq$ 29.45:1
THC–CBD-intermediate chemotype	0.225:1 $\geq$ THC–CBD $\leq$ 0.668:1
CBD-dominant chemotype	THC–CBD $\leq$ 0.085:1

CBD, cannabidiol; THC, tetrahydrocannabinol.

Both rely on XGB regression to establish the model. For the HPLC method, total THC and total CBD content are determined by adding the neutral and acid forms of each, respectively, as determined by HPLC.

Results and Discussion

GC-FID and HPLC results

Table 2 gives the cannabinoid concentration parameters measured by GC-FID for data, calibration, and validation combined, characterizing the three chemotypes: maximum and minimum CBD-THC ratios and the absolute maxima and minima for CBD and THC measured. The samples have typical ranges of THC-CBD ratios,²² which tend to be well separated from the defined limits, in other

words typical THC-dominant samples have a THC-CBD ratio of $\geq 20:1$, even though a ratio of $\geq 5:1$ is sufficient for classification.

Testing samples that are closer to the limit definitions is outside the scope of this article. Based on the previously reported GC/FID method,⁴⁴ the limit of detection (LOD) value for total THC/THCA and total CBD/CBDA is 0.01%, whereas their limit of quantification (LOQ) value is 0.025%. Based on the previously published HPLC method,⁴⁵ the LODs of the four tested cannabinoids are 0.037% (CBD), 0.010% (CBDA), 0.014% (THC), and 0.007% (THCA), whereas the LOQ values are 0.110% (CBD), 0.029% (CBDA), 0.041% (THC), and 0.020% (THCA).

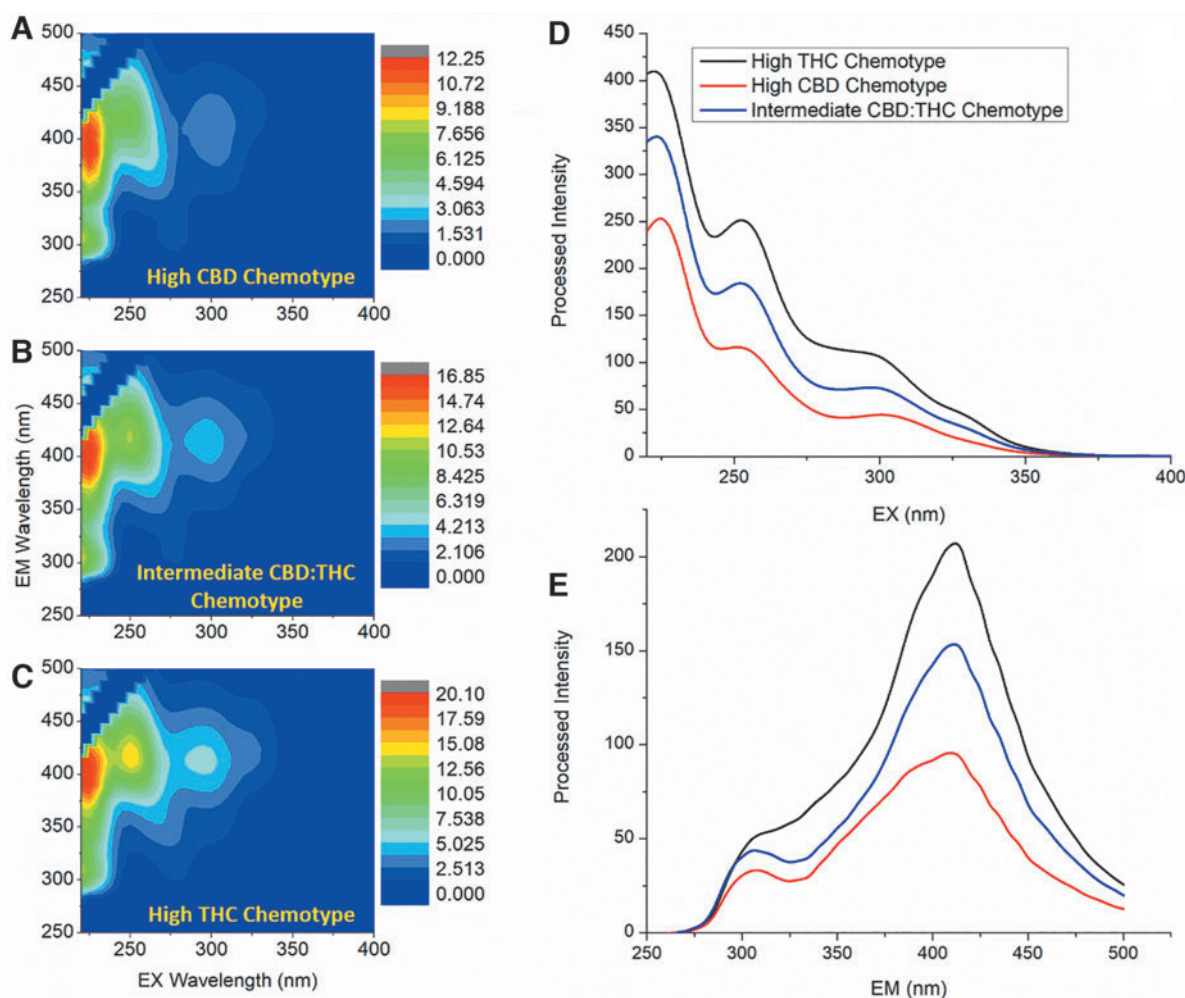


FIG. 1. Peak-normalized A-TEEM contours (A–C) and naturally scaled integral excitation- (D) and emission- (E) profiles for flower extracts of HTHC, hybrid and HCBD cannabis varieties. A-TEEM, absorbance-transmittance excitation emission matrix; HCBD, high CBD; HTHC, high THC.

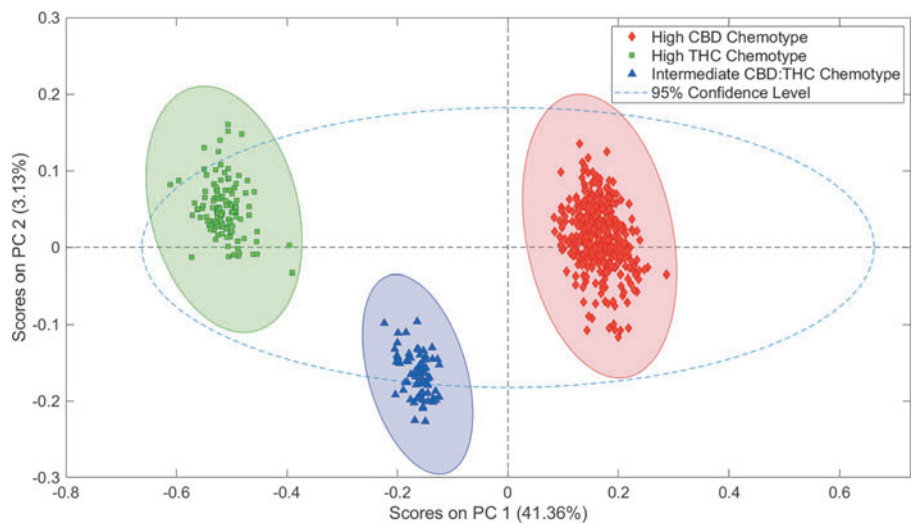


FIG. 2. PCA of THC-dominant (HTHC), THC–CBD-intermediate (HCBHYBRID), and CBD-dominant (HCB) cannabis varieties. Represented in each 99% confidence level ellipse are all calibration and validation samples for each class group. PCA, principal components analysis.

A-TEEM for chemotype identification—
qualitative discrimination

The first step to establish the capabilities of the A-TEEM method is to demonstrate that it can qualitatively distinguish the three *C. sativa* chemotypes based on relative THC–CBD content of dry flower extracts. Looking at representative data is a good starting point, and Figure 1 shows A-TEEM contours (A–C) and excitation (D) and emission (E) profile spectra of representative ex-

amples (same w:v extraction) for the three THC–CBD variety extracts designated as THC-dominant, THC–CBD-intermediate, and CBD-dominant from the GC-FID reference data, respectively.

The A-TEEM contours are each scaled to the peak excitation–emission intensity ($\lambda_{\text{ex}}/\lambda_{\text{em}} = 225/400 \text{ nm}$). THC levels increase from A to B to C, and there is a feature ($\lambda_{\text{ex}}/\lambda_{\text{em}} = 250/420 \text{ nm}$) in the A-TEEM contour whose intensity increases similarly. The integrated excitation

Table 3. (A–C) Confusion Matrices Highlighting 100% Predictive Accuracy of Absorbance-Transmittance Excitation Emission Matrix Method (Using Gas Chromatography-Flame Ionization Reference Data) to Classify Samples into Three Cannabis Varieties: Tetrahydrocannabinol (THC)-Dominant Chemotype, THC–Cannabidiol (CBD) Intermediate Chemotype, and CBD-Dominant Chemotype

Actual	Predicted			Total	Correct %
	CBD-dominant chemotype	THC–CBD-intermediate chemotype	THC-dominant chemotype		
(A). Calibration					
CBD-dominant chemotype	346	0	0	346	100
THC–CBD-intermediate chemotype	0	60	0	60	100
THC-dominant chemotype	0	0	109	109	100
(B). Cross-validation					
CBD-dominant chemotype	346	0	0	346	100
THC–CBD intermediate chemotype	0	60	0	60	100
THC-dominant chemotype	0	0	109	109	100
(C). Validation					
CBD-dominant chemotype	24	0	0	24	100
THC–CBD intermediate chemotype	0	12	0	12	100
THC-dominant chemotype	0	0	23	23	100

Results from XGBDA. Calibration (A), cross-validation (B), and prediction (C) all reflect 100% predictive accuracy. XGBDA, extreme gradient boost discrimination analysis.

(D) and emission (E) profiles are unscaled, and also show systematic differences with increasing THC. Both excitation and emission profiles increase in intensity corresponding to increasing THC content from CBD-dominant to THC–CBD-intermediate to THC-dominant.

To classify the hundreds of measurements collected, multivariate methods are required to automate classifications, determine how well the discrimination is working, and to create models against which unknowns may be classified. Figure 2 shows how effectively a simple PCA separates the 556 calibration and validation A-TEEM data points from the 49 unique plants into 3 separate clusters. As a PCA is totally unsupervised, these clusters are based on spectral variance alone. The GC-FID data were used to assign each data point to a specific chemotype, and this confirms that that unique clusters correspond to unique chemotype.

The CBD-dominant and THC-dominant samples are differentiated by PC1, THC-dominant samples have negative PC1 scores, whereas CBD-dominant samples have positive PC1 scores. The THC–CBD-intermediate chemotype has intermediate PC1 scores, as would be anticipated, but are more similar to the CBD-dominant samples than the THC-dominant samples. The GC-FID classification results given in Table 2 show the same trend, that the TCH–CBD ratio for intermediate samples is closer to the CBD-dominant than to the THC-dominant ratio. There is no overlap between the chemotypes in this scatter plot, with >99% confidence, indicating that A-TEEM data are sensitive to the spectral differences between these samples.

To further evaluate the ability of A-TEEM to discriminate these three varieties, the calibration and validation data were evaluated using the extreme gradient boost discrimination analysis algorithm. In previous studies with wine and grape classification, the approach has been shown to yield accurate classification of wine varie-

tals.^{47–49} The benefit of this approach is that a discriminant model can be used to classify unknown samples, in contrast to PCA that simply analyzes spectral variance.

Table 3A–C gives the confusion matrices for calibration, cross-validation, and prediction classifications using the “strict rule.” The cross-validation was performed using ($k=10$) Venetian blinds method described previously.^{47–49} The tables show that 100% of samples are classified correctly, there are no false positive or false negatives, all classifications are true positives and true negatives.

Table 4. Results of Absorbance-Transmittance Excitation Emission Matrix Model for Total Cannabidiol and Total Tetrahydrocannabinol Quantification with Gas Chromatography-Flame Ionization Reference Data

XGB fit	CBD	THC	<i>n</i>
Calibration	1.000000	0.99811	517
Cross-validation	0.997978	0.99904	517
Prediction	0.991433	0.996602	59

Coefficients of determination (R^2) for the XGB regression for total CBD and total THC measured by GC-FID for calibration, cross-validation, and prediction (validation) data. High values (>0.99) for all indicate the model is working well for the prediction of these components.

FID, flame ionization; XGB, extreme gradient boost.

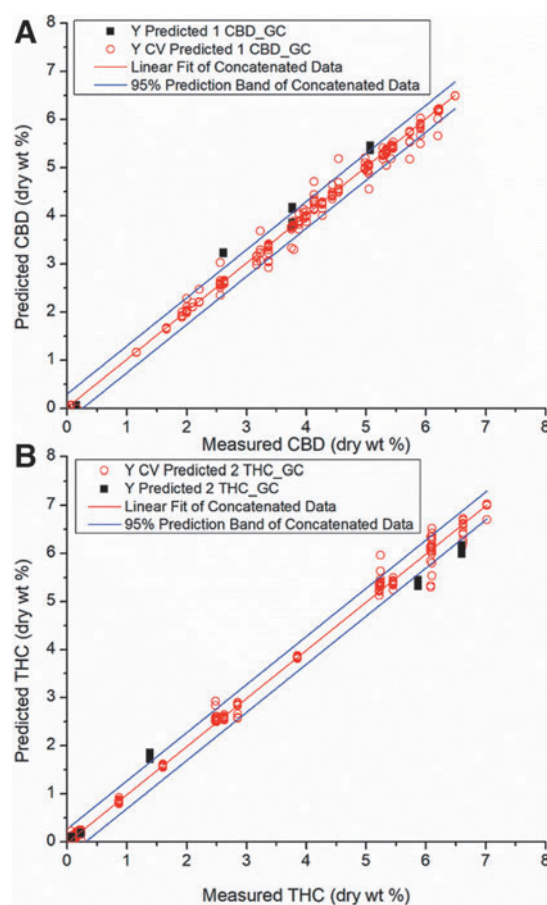


FIG. 3. Concatenated XGB regression fits of cross-validated (open red) and validation (black) A-TEEM sample data for total CBD (**A**) and total THC (**B**) measured by GC-FID and expressed as dry weight %. Also shown are 95% prediction interval band lines (blue) and the fit line (red). FID, flame ionization; GC, gas chromatography; XGB, extreme gradient boost.

Table 5. Linear Regression Statistics from Extreme Gradient Boost Model That Includes All 576 Prediction and Cross-Validation Data Points for Total Cannabidiol and Total Tetrahydrocannabinol Measured by Gas Chromatography-Flame Ionization

Technique	Parameter	Intercept	SE	Slope	LOD % (w/w)	LOQ % (w/w)	Adjusted R^2	<i>n</i>
A-TEEM	CBD + CBDA	0.019	0.006	1	0.019	0.059	0.9954	576
	THC + THCA	−0.016	0.006	1	0.020	0.061	0.9958	576
GC-FID	CBD + CBDA	-2.872×10^{-5}	NA	5.424×10^{-4}	0.01	0.025	1	NA
	THC + THCA	-2.877×10^{-6}	NA	5.546×10^{-4}	0.01	0.025	1	NA

The slope for A-TEEM data was constrained to unity. Defined are the y-intercept (intercept), SE of the y-intercept (SE), $\text{LOD} = 3.3 \times \text{SE}$, $\text{LOQ} = 3.3 \times \text{LOD}$, the adjusted R^2 , and number of files (*n*) that combines calibration and validation.

A-TEEM, absorbance-transmittance excitation emission matrix; CBDA, cannabidiolic acid; GC, gas chromatography; LOD, limit of detection; LOQ, limit of quantification; NA, not applicable; SE, standard error; THCA, tetrahydrocannabinolic acid.

Quantification of total CBD and THC by A-TEEM regression calibration with GC-FID

The next test of the capabilities of A-TEEM is to evaluate its powers of quantitation. Regression models for A-TEEM data were created and evaluated for total CBD (combined acid and neutral forms: CBDA and CBD) and total THC (combined acid and neutral forms: THCA and THC) using GC-FID reference data, noting that GC-FID is unable to distinguish between the acid and neutral forms of CBD or THC for the underivatized samples used in this study. The regressions were performed using the XGB regression algorithm. The multivariate fit coefficient of determination (R^2) for the actual calibration, cross-validation, and prediction data is given in Table 4. These values are all >0.99 , showing that the model is performing reliably.

Figure 3A and B shows the cross-validation data (open red circles) and validation data from five unique plants (solid black squares) with linear regression lines for total CBD and total THC, respectively. This includes all data in the calibration and prediction sets. As noted previously, the primary limitation of standard spectroscopic methods (FT-IR, Raman, and NIR) is that the LOD and LOQ are typically not sufficient to determine statutory levels of THC and CBD. An evaluation of the LOQ for A-TEEM can be calculated from the y-intercept and standard error of the y-intercept (SE) of the linear fit, if the slope of the regression is constrained to unity.

Table 5 presents the regression fits for A-TEEM and GC-FID methods for both THC and CBD, including all calibration and validation data, along with the values determined using the GC_FID reference method. Using the following formula, $\text{LOD} = 3.3 \times \text{SE}$, and $\text{LOQ} = 3.3 \times \text{LOD}$, the predicted LODs by component for the A-TEEM models are 0.019% (CBD) and 0.020% (THC). The predicted LOQs by component for the A-TEEM models are 0.059% (CBD) and 0.061% (THC), sufficiently low for statutory purposes.

Quantification of CBD, CBDA, THC, and THCA by A-TEEM regression calibration with HPLC

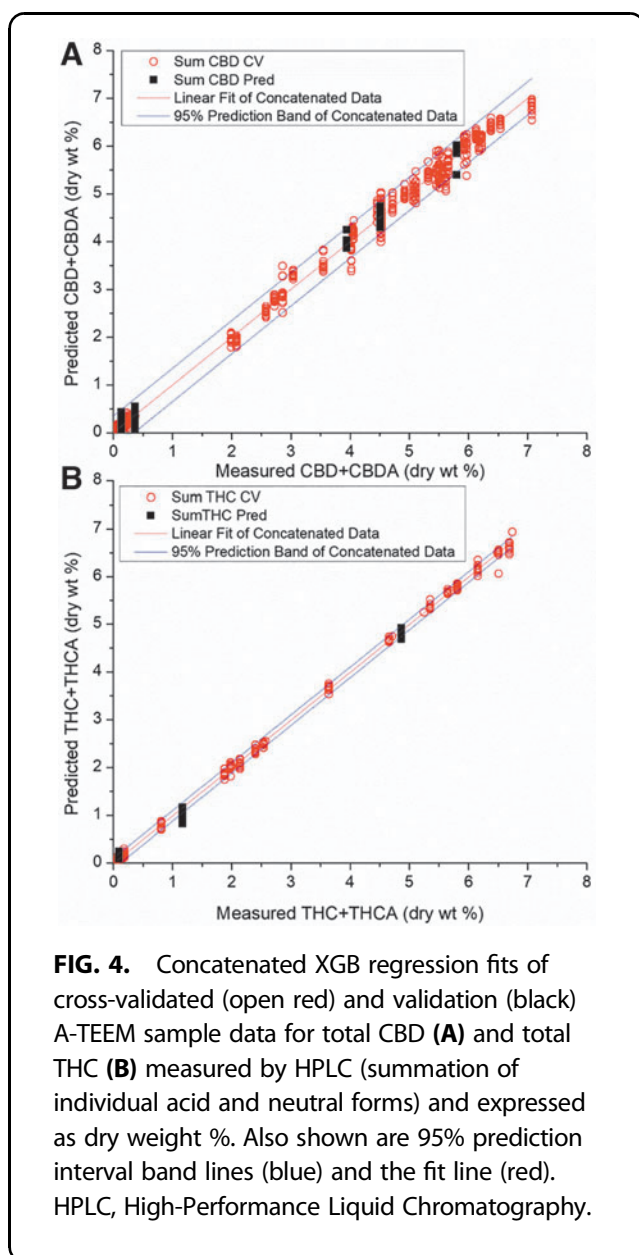
As HPLC methods are commonly used for quantification of THC and CBD levels, it is important to establish an A-TEEM method using HPLC reference data as well. In addition, as individual THC, THCA, CBD, and CBDA components are resolvable by HPLC, an A-TEEM model for the quantification of these components can also be developed. In previous studies with wine and grape classification and phenolic compound quantification, the XGB regression algorithm has documented superior fit statistics compared with the partial least squares algorithm.^{47–49} An XGB regression model for A-TEEM was created with HPLC reference data, to quantify and predict individual acid and base THC and CBD compounds, as well as combined total THC and CBD levels important for statutory purposes.

Table 6 gives the multivariate coefficients of determination (R^2) values for the calibration, cross-validation, and prediction data for the individual acid and neutral forms. Calibration values for all four components are >0.99 , with cross-validation values all ≥ 0.99 . For prediction, only one value was <0.99 , specifically the R^2 for CBD prediction was >0.98 , still very high. These

Table 6. Results of Absorbance-Transmittance Excitation Emission Matrix Model for Cannabidiol, Cannabidiol Acid, Tetrahydrocannabinol, and Tetrahydrocannabinol Acid Quantification with High-Performance Liquid Chromatography Reference Data

XGB fit	Calibration	Cross-validation	Prediction
CBD	0.992538	0.993634	0.981454
CBDA	0.990559	0.99237	0.996454
THC	0.998983	0.989893	0.993457
THCA	0.998296	0.999509	0.995188

Coefficients of determination (R^2) for the XGB regression for calibration and cross-validation ($n=517$) and prediction (validation) ($n=59$). High values (>0.98) for all indicate the model is working well for the prediction of these components.



numbers indicate that the A-TEEM method for quantification of these components using HPLC reference data is highly predictive.

Figure 4A and B is the HPLC-based analogue to the GC-FID model results presented in Figure 3A and B, and, respectively, show the prediction and cross-validation results and linear fit of predicted versus measured data for total THC and CBD. These values are determined by adding results for the acid and neutral forms (THCA + THC) and (CBDA + CBD), in contrast to the GC-FID data where the totals were directly measured. Open red circles are the summed (acid and neutral) cross-validated values, and solid black squares are the summed (acid and neutral) validation values from five unique plants.

Using the same formula already referenced in the GC-FID quantification section, the predicted LODs by component for the A-TEEM models are 0.024% (CBD/CBDA) and 0.008% (THC/THCA). The predicted LOQs by component for the A-TEEM models are 0.080% (CBD/CBDA) and 0.026% (THC/THCA) as given in Table 7, along with the values determined using the HPLC reference method.

A-TEEM for substitutory quantitative prediction on THC and CBD

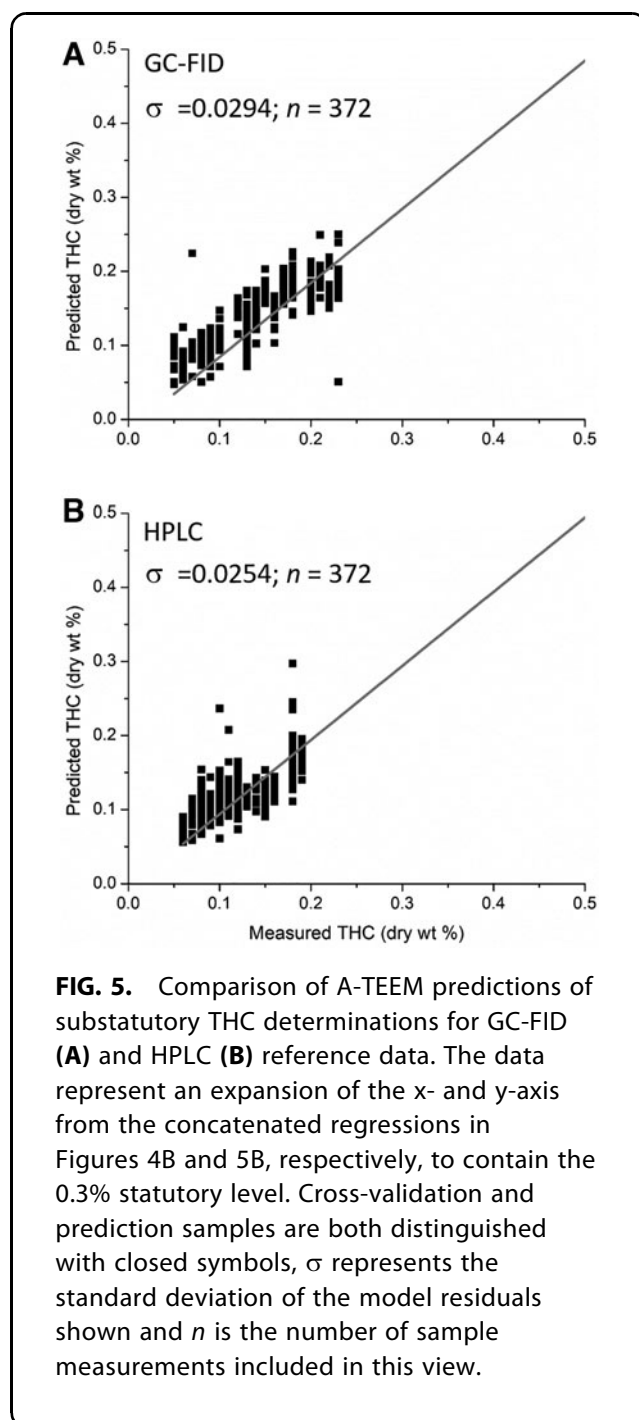
To assess the A-TEEM performance in the determination of low levels of THC, a measurement that is important for statutory determination, CBD-dominant samples are examined separately. Figure 5 compares A-TEEM predictions of the substitutory levels of THC against GC-FID reference data (A) and HPLC reference data (B), respectively. There were 372 individual data points classified as CBD-dominant (Table 3A, C), 348 from calibration and 24 from validation samples, corresponding to 31 individual plants. For these 372 A-TEEM measurements, there were no false positive

Table 7. Linear Regression Statistics from Extreme Gradient Boost Regression Model That Includes All 576 Prediction and Cross-Validation Data Points for Total Cannabidiol and Tetrahydrocannabinol Measured by HPLC

Technique	Parameter	Intercept	SE	Slope	LOD % (w/w)	LOQ % (w/w)	Adjusted R^2	n
A-TEEM	CBD + CBDA	0.001	0.007	1	0.024	0.080	0.9936	576
	THC + THCA	-0.006	0.002	1	0.008	0.026	0.9993	576
HPLC	CBD	-1.215×10^{-4}	NA	0.022	0.037	0.110	0.9998	NA
	CBDA	0.009	NA	0.025	0.010	0.029	0.9999	NA
	THC	-0.001	NA	0.022	0.014	0.041	0.9997	NA
	THCA	-0.006	NA	0.025	0.007	0.020	0.9996	NA

The A-TEEM slope was constrained to unity. Defined are the y-intercept (intercept), SE of the y-intercept (SE), $\text{LOD} = 3.3 \times \text{SE}$, $\text{LOQ} = 3.3 \times \text{LOD}$, the adjusted R^2 , and number of files (n) that combine calibration and validation.

HPLC, High-Performance Liquid Chromatography; NA, not applicable.



statutory assignments, and the standard deviation of the residual errors σ was 0.0294 (GC-FID reference data) and 0.0254 (HPLC reference data).

Conclusions

A-TEEM is a rapid (<45 sec per measurement) method, with simple sample preparation (extraction

and dilution), and low per-run cost compared with GC-FID and HPLC methods. Sensitivity and specificity are comparable with chromatography, and LODs and LOQs are low enough for statutory classification. A-TEEM with GC-FID reference data is able to classify cannabis varieties based on THC–CBD ratio and classify independent validation samples with 100% accuracy. Furthermore, A-TEEM regression models calibrated with both GC-FID and HPLC data yield LOQ values significantly <0.1%, indicating A-TEEM is suitable for evaluating statutory THC concentrations.

In practice, the models performed with 100% accuracy, yielding no statutory misclassifications in samples with THC levels significantly below the 0.3% w/w limit. Importantly, the A-TEEM model application can be automated for rapid analysis reporting in a laboratory environment. HORIBA products and solutions are intended to be used for quality control and safety testing in facilities where such use is permitted under applicable local or federal law.

Acknowledgment

We acknowledge the help of Sean M. Travers in reviewing the article.

Authors' Contributions

A.M.G., J.B.H., and A.E.R. conceived of the experiments; A.M.G., M.M.R., and M.A. Elshohly designed the experiments; M.A. Elhendawy, M.M.R., and A.S.W. performed the experiments; A.M.G. and M.A. Elhendawy analyzed the data; M.A. Elhendawy, M.M.R., A.M.G., and L.H.K. drafted the article with critical intellectual input to the article from J.B.H.; and A.M.G., J.B.H., M.G., and M.A. Elshohly supervised the project. All authors reviewed the article and approved its submission.

Author Disclosure Statement

No competing financial interests exist.

Funding Information

This study was supported in part by the National Institute on Drug Abuse (contract no. N01DA-15-7793).

References

1. Hazekamp A, Ware MA, Muller-Vahl KR, et al. The medicinal use of cannabis and cannabinoids—an international cross-sectional survey on administration forms. *J Psychoact Drugs*. 2013;45:199–210.
2. Fishedick JT, Hazekamp A, Erkelens T, et al. Metabolic fingerprinting of *Cannabis sativa* L., cannabinoids and terpenoids for chemotaxonomic and drug standardization purposes. *Phytochemistry*. 2010;71:2058–2073.

3. Hillig KW, Mahlberg PG. A chemotaxonomic analysis of cannabinoid variation in *Cannabis* (Cannabaceae). *Am J Bot*. 2004;91:966–975.
4. Radwan MM, Chandra S, Gul S, et al. Cannabinoids, phenolics, terpenes and alkaloids of cannabis. *Molecules*. 2021;26:2774.
5. Aizpurua-Olaizola O, Soydaner U, Öztürk E, et al. Evolution of the cannabinoid and terpene content during the growth of *Cannabis sativa* plants from different chemotypes. *J Nat Prod*. 2016;79:324–331.
6. Elsohly MA, Harland EC, Benigni DA, et al. Cannabinoids in glaucoma II: the effect of different cannabinoids on intraocular pressure of the rabbit. *Curr Eye Res*. 1984;3:841–850.
7. El-Mallakh RS. Marijuana and migraine. *J Headache Pain*. 1987;27:442–443.
8. Volfe Z, Dvilansky A, Nathan I. Cannabinoids block release of serotonin from platelets induced by plasma from migraine patients. *Int J Clin Pharmacol Res*. 1985;5:243–246.
9. Maurer M, Henn V, Dittrich A, et al. Delta-9-tetrahydrocannabinol shows antispastic and analgesic effects in a single case double-blind trial. *Eur Arch Psychiatry Clin Neurosci*. 1990;240:1–4.
10. McLendon DM, Harris RT, Maule WF. Suppression of the cardiac conditioned response by Δ -9-tetrahydrocannabinol: a comparison with other drugs. *J Psychopharmacol*. 1976;50:159–163.
11. Karler R, Cely W, Turkianis SA. The anticonvulsant activity of cannabidiol and cannabinol. *J Life Sci*. 1973;13:1527–1531.
12. Arzimanoglou A, Brandt U, Cross JH, et al. Epilepsy and cannabidiol: a guide to treatment. *J Epileptic Disord*. 2020;22:1–14.
13. Castillo A, Tolón M, Fernández-Ruiz J, et al. The neuroprotective effect of cannabidiol in an *in vitro* model of newborn hypoxic-ischemic brain damage in mice is mediated by CB2 and adenosine receptors. *Neurobiol Dis*. 2010;37:434–440.
14. Hayakawa K, Mishima K, Nozako M, et al. Repeated treatment with cannabidiol but not Δ -9-tetrahydrocannabinol has a neuroprotective effect without the development of tolerance. *Neuropharmacology*. 2007;52:1079–1087.
15. El-Remessy AB, Khalil IE, Matragoon S, et al. Neuroprotective effect of (–) Δ -9-tetrahydrocannabinol and cannabidiol in N-methyl-D-aspartate-induced retinal neurotoxicity: involvement of peroxynitrite. *Am J Pathol*. 2003;163:1997–2008.
16. Hampson AJ, Grimaldi M, Axelrod J, et al. Cannabidiol and (–)- Δ -9-tetrahydrocannabinol are neuroprotective antioxidants. *Proc Natl Acad Sci U S A*. 1998;95:8268–8273.
17. Iuvone T, Esposito G, Esposito R, et al. Neuroprotective effect of cannabidiol, a non-psychoactive component from *Cannabis sativa*, on β -amyloid-induced toxicity in PC12 cells. *J Neurochem*. 2004;89:134–141.
18. van Breemen RB, Muchiri RN, Bates TA, et al. Cannabinoids block cellular entry of SARS-CoV-2 and the emerging variants. *J Nat Prod*. 2022;85:176–184.
19. Devinsky O, Cross JH, Laux L, et al. Trial of cannabidiol for drug-resistant seizures in the Dravet syndrome. *NEJM*. 2017;376:2011–2020.
20. Devinsky O, Patel AD, Cross JH, et al. Effect of cannabidiol on drop seizures in the Lennox–Gastaut syndrome. *NEJM*. 2018;378:1888–1897.
21. Chandra S, Lata H, Elsohly MA. *Cannabis sativa* L.-botany and biotechnology. Cham, Switzerland: Springer, 2017.
22. Sarma ND, Waye A, Elsohly MA, et al. Cannabis inflorescence for medical purposes: USP considerations for quality attributes. *J Nat Prod*. 2020;83:1334–1351.
23. Citti C, Russo F, Sgrò S, et al. Pitfalls in the analysis of phytocannabinoids in cannabis inflorescence. *J Anal Bioanal Chem*. 2020;412:4009–4022.
24. Mercolini L, Mandrioli R, Protti M. New trends in the analysis of Cannabis-based products. *J Recent Adv Sci Cannabis*. 2021;111–130, DOI: 10.1201/9780429274893-5.
25. McRae G, Melanson JE. Quantitative determination and validation of 17 cannabinoids in cannabis and hemp using liquid chromatography-tandem mass spectrometry. *Anal Bioanal Chem*. 2020;412:7381–7393.
26. Nemeškalová A, Hájková K, Mikulů L, et al. Combination of UV and MS/MS detection for the LC analysis of cannabidiol-rich products. *Talanta*. 2020;219:121250.
27. Protti M, Brighenti V, Battaglia MR, et al. Cannabinoids from *Cannabis sativa* L.: a new tool based on HPLC–DAD–MS/MS for a rational use in medicinal chemistry. *ACS Med Chem Lett*. 2019;10:539–544.
28. Ternelli M, Brighenti V, Anceschi L, et al. Innovative methods for the preparation of medical Cannabis oils with a high content of both cannabinoids and terpenes. *J Pharm Biomed Anal*. 2020;186:113296.
29. Basas-Jaumandreu J, de Las Heras FXC. GC-MS metabolite profile and identification of unusual homologous cannabinoids in high potency *Cannabis sativa*. *J Planta Medica*. 2020;86:338–347.
30. Zekič J, Križman M. Development of gas-chromatographic method for simultaneous determination of cannabinoids and terpenes in hemp. *Molecules*. 2020;25:5872.
31. Cirrincione M, Zanfognini B, Pigani L, et al. Development of an electrochemical sensor based on carbon black for the detection of cannabidiol in vegetable extracts. *Analyst*. 2021;146:612–619.
32. Brighenti V, Pellati F, Steinbach M, et al. Development of a new extraction technique and HPLC method for the analysis of non-psychoactive cannabinoids in fibre-type *Cannabis sativa* L.(hemp). *J Pharm Biomed Anal*. 2017;143:228–236.
33. Corni G, Brighenti V, Pellati F, et al. Effect-directed analysis of bioactive compounds in *Cannabis sativa* L. by high-performance thin-layer chromatography. *J Chromatogr A*. 2020;1629:461511.
34. Beasley E, Francese S, Bassindale T. Detection and mapping of cannabinoids in single hair samples through rapid derivatization and matrix-assisted laser desorption ionization mass spectrometry. *J Anal Chem*. 2016;88:10328–10334.
35. Nahar L, Guo M, Sarker SD. Gas chromatographic analysis of naturally occurring cannabinoids: a review of literature published during the past decade. *Phytochem Anal*. 2020;31:135–146.
36. Citti C, Braghiroli D, Vandelli MA, et al. Pharmaceutical and biomedical analysis of cannabinoids: a critical review. *J Pharm Biomed Anal*. 2018;147:565–579.
37. Hazekamp A, Simons R, Peltenburg-Looman A, et al. Preparative isolation of cannabinoids from *Cannabis sativa* by centrifugal partition chromatography. *J Liq Chromatogr Relat Technol*. 2004;27:2421–2439.
38. De Backer B, Debrus B, Lebrun P, et al. Innovative development and validation of an HPLC/DAD method for the qualitative and quantitative determination of major cannabinoids in cannabis plant material. *J Chromatogr B*. 2009;877:4115–4124.
39. Romano LL, Hazekamp A. Cannabis oil: chemical evaluation of an upcoming cannabis-based medicine. *Cannabinoids*. 2013;1:1–11.
40. Aizpurua-Olaizola O, Omar J, Navarro P, et al. Identification and quantification of cannabinoids in *Cannabis sativa* L. plants by high performance liquid chromatography-mass spectrometry. *J Anal Bioanal Chem*. 2014;406:7549–7560.
41. Leghissa A, Hildenbrand ZL, Schug KA. A review of methods for the chemical characterization of cannabis natural products. *J Sep Sci*. 2018;41:398–415.
42. Gilmore AM, Tong XM. System and method for fluorescence and absorbance analysis. In United States Patent and Trademark Office, 2019.
43. Marinotti O, Sarill M. Differentiating full-spectrum hemp extracts from CBD isolates: implications for policy, safety and science. *J Diet Suppl*. 2020;17:517–526.
44. Elsohly MA, Mehmedic Z, Foster S, et al. Changes in cannabis potency over the last 2 decades (1995–2014): analysis of current data in the United States. *Biol Psychiatry*. 2016;79:613–619.
45. Gul W, Gul SW, Radwan MM, et al. Determination of 11 cannabinoids in biomass and extracts of different varieties of Cannabis using high-performance liquid chromatography. *J AOAC Int*. 2015;98:1523–1528.
46. Ahmed SA, Ross SA, Slade D, et al. Cannabinoid ester constituents from high-potency *Cannabis sativa*. *J Nat Prod*. 2008;71:536–542.
47. Ranaweera RK, Gilmore AM, Capone DL, et al. Authentication of the geographical origin of Australian Cabernet Sauvignon wines using spectrofluorometric and multi-element analyses with multivariate statistical modelling. *Food Chem*. 2021;335:127592.
48. Ranaweera RK, Gilmore AM, Capone DL, et al. Spectrofluorometric analysis combined with machine learning for geographical and varietal authentication, and prediction of phenolic compound concentrations in red wine. *Food Chem*. 2021;361:130149.
49. Gonzaga LS, Bastian SE, Capone DL, et al. Modelling Cabernet-Sauvignon wine sensory traits from spectrofluorometric data. *J OENO One*. 2021;55:19–33.

Cite this article as: Gilmore AM, Elhendawy MA, Radwan MM, Kidder LH, Wanas AS, Godfrey M, Hildreth JB, Robinson AE, ElSohly MA (2023) Absorbance-transmittance excitation emission matrix method for quantification of major cannabinoids and corresponding acids: a rapid alternative to chromatography for rapid chemotype discrimination of *Cannabis sativa* varieties, *Cannabis and Cannabinoid Research* 8:5, 911–922, DOI: 10.1089/can.2021.0165.

Abbreviations Used

A-TEEM = absorbance-transmittance excitation emission matrix
CBD = cannabidiol
CBDA = cannabidiolic acid
DA = discriminant analysis
EEM = excitation emission matrix
FDA = Food and Drug Administration
FID = flame ionization detector
FT-IR = Fourier transform infrared
GC = gas chromatography

HCBD = high CBD
HPLC = high-performance liquid chromatography
HTHC = high THC
IFE = inner filter effect
LC-MS/MS = liquid chromatography
with tandem mass spectrometry
LOD = limit of detection
LOQ = limit of quantification
NA = not applicable
NIR = near infrared
NLT = not less than
NMT = not more than
PCA = principal components analysis
PDA = photodiode array
PLS = partial least squares
THC = tetrahydrocannabinol
THCA = tetrahydrocannabinolic acid
UV/Vis = ultraviolet-visible
XGB = extreme gradient boost
XGBDA = extreme gradient boost discrimination analysis