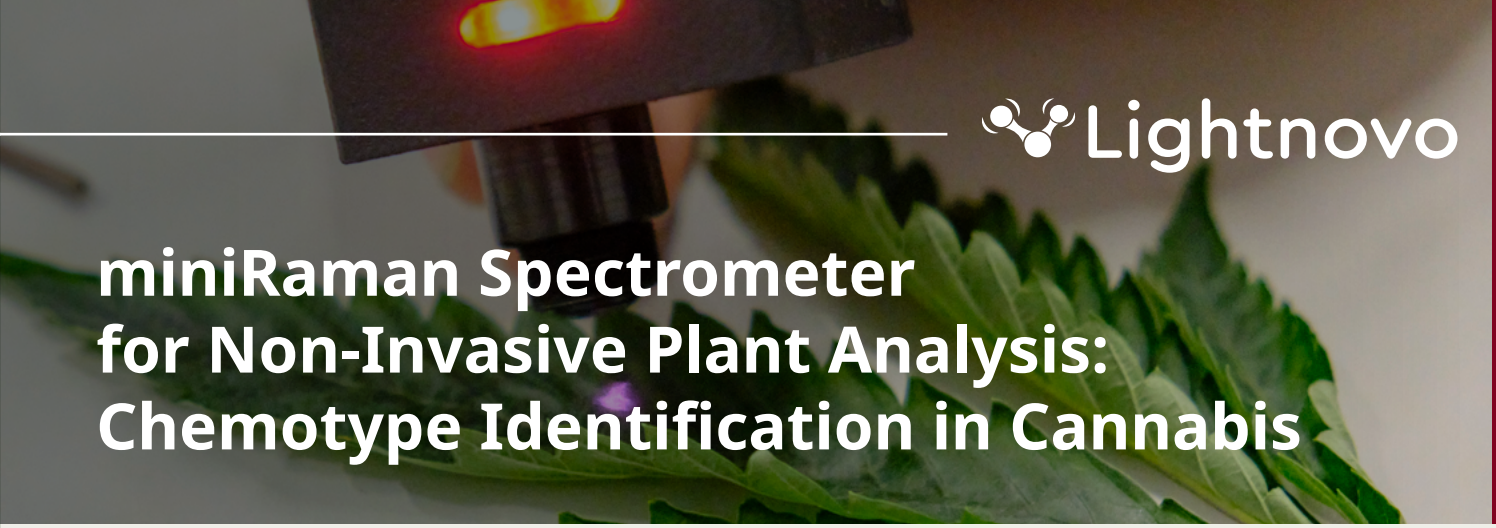

Lightnovo

APPLICATION NOTE

miniRaman Spectrometer for Non-Invasive Plant Analysis: Chemotype Identification in Cannabis

ApN-010

Key words

Cannabis; Raman spectroscopy; phytocannabinoids (pCBs); delta-9-tetrahydrocannabinol (THC), cannabidiol (CBD); chemotype identification; fluorescence suppression; chemometrics; principal component analysis (PCA); Partial Least Squares Discriminant Analysis (PLS-DA).

Preface

The measurements described in this Application Note were carried out in collaboration with Prof. Rime Bahij and Prof. Martin Aage Barsøe Hedegaard from the Department of Green Technology of the University of Southern Denmark (SDU), Odense, Denmark. SDU has a license to cultivate, manufacture, and distribute cannabis, issued by the Danish Medicines Agency in accordance with applicable legislation.

Introduction

Plants are a critical source of materials and biologically active compounds widely used in food production, agriculture, medicine, cosmetics, and bio-based industries. In modern field and greenhouse crop production, monitoring and early diagnosis of plant physiological state and identifying plant responses to biotic and abiotic stress are important for increasing productivity and crop stability. Early differentiation of plant biotypes and sexes is particularly important for breeding, regulation, and yield optimization. Identifying unsuitable plants at early growth stages helps prevent unwanted pollination, reduce low-productivity individuals, and optimize the plantation for the chosen cultivation purpose. These tasks require *in vivo* analytical methods and instruments that enable rapid, field-based measurements from plant to plant.

Raman spectroscopy is a powerful tool for plant analysis, enabling rapid characterization of plant tissues. It can be used to classify plant biotypes, assess their physiological state and nutrient content, monitor growth-related biochemical changes, and estimate crop quality [1–4]. Raman spectroscopy can also be used for plant health monitoring, enabling early detection of biotic and abiotic stress and capturing weak biochemical changes that appear long before any visible symptoms [5]. Such changes are reflected in Raman spectra through variations in the composition and content of key metabolites, including carotenoids, pectin, lignin, carbohydrates, amino acids, and nitrates.

The advantage of Raman spectroscopy is its ability to quickly obtain detailed molecular information about complex plant tissues in a non-destructive, non-contact manner without pre-treatment or chemical labeling of the studied material. Raman measurements can be performed directly on fresh or living plant tissues, without the use of chemical reagents or the time-consuming and expensive sample preparation, making the method particularly attractive for *in vivo* and *in situ* applications. These capabilities make Raman spectroscopy a promising analytical approach for precision modern agriculture and plant diagnostics.

In practice, applying Raman spectroscopy to plant tissues poses certain challenges due to strong autofluorescence. Additionally, many target compounds are present at significantly lower concentrations compared to the plant matrix. These factors complicate spectral interpretation and reduce sensitivity. However, the use of near-infrared excitation wavelengths, optimized optical designs, and advanced data analysis techniques such as Principal Component

Analysis (PCA) and Partial Least Squares Discriminant Analysis (PLS-DA) can effectively overcome these limitations [3,6].

This Application Note aims to demonstrate the potential of the hand-held Lightnovo miniRaman spectrometer [7], the most compact serial Raman instrument, for plant research and analysis. Using cannabis as a representative example, this study illustrates the feasibility of portable Raman for identifying spectral features associated with key bioactive components and non-invasive chemotype differentiation of living cannabis plants based on *in vivo* leaf spectra.

Background

Phytocannabinoids in cannabis

Cannabis plants are cultivated primarily for fibers, seeds, and phytocannabinoids (pCBs), including delta-9-tetrahydrocannabinol (THC), cannabidiol (CBD), cannabigerol (CBG), etc. Due to the psychoactive properties of THC, cannabis cultivation is strictly regulated. Based on THC content, the crop is classified into THC-rich "marijuana" and industrial hemp, which must not exceed a legal THC threshold, typically 0.3% (dry weight). Industrial hemp is cultivated primarily for the production of non-psychoactive pCBs, particularly CBD, and technical applications.

Improving efficiency and profitability in cannabis cultivation requires *in vivo* plant differentiation, particularly at the early vegetative growth stage. For instance, unfertilized female plants produce the highest levels of pCBs, whereas male plants yield superior fiber quality and greater biomass. The use of methods that can distinguish between male and female plants will help replace unsuitable plants, prevent unwanted pollination, and improve productivity, thereby optimizing the plantation for the chosen cultivation purpose.

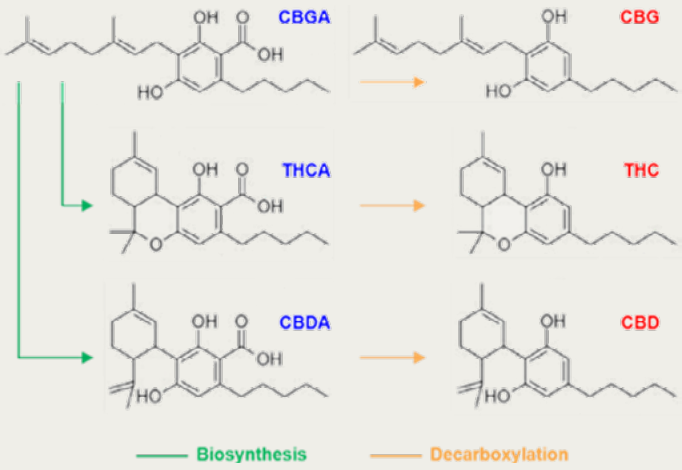

Figure 1. Scheme of main pCB biosynthesis and decarboxylation in cannabis. CBGA is a common biosynthetic precursor for THCA and CBDA. By a decarboxylation reaction, THCA and CBDA are converted into their neutral forms, THC and CBD, respectively.

Cannabis can produce at least 140 different pCBs, although most of these are usually present only in trace amounts [8]. Their biosynthesis proceeds via the common precursor cannabigerolic acid (CBGA), which is enzymatically converted into acidic cannabinoids such as THCA, CBDA, and CBCA (Fig. 1). All these reactions are biosynthetic, i.e., each is governed by the corresponding DNA-encoded enzyme(s). As a result, every plant receives a unique profile (composition and content) of pCBs depending on the features of its genome. In living plants, pCBs are present predominantly in acidic forms, while neutral cannabinoids (THC, CBD, CBG, CBC) are formed mainly via non-enzymatic decarboxylation reactions induced by harvesting or stress (heat, chemicals, etc.).

The pCB profile of cannabis is classified into a specific "chemotype," characterized by distinct compositions of major pCBs that consistently occur within the species. The most common chemotypes are: (I) THC-dominant; (II) THC/CBD-intermediate; (III) CBD-dominant; (IV) CBG-dominant; (V) pCBs-free. This classification reflects genetically determined differences in pCB biosynthesis and is stable within a given plant variety. Early chemotype differentiation can significantly enhance the efficiency of target pCB production. This is especially important for greenhouse cultivation, where production costs are comparatively high, and potential errors must be identified and eliminated as early as possible.

Note that although the pCBs composition is determined by chemotype, the content of each pCB changes during plant growth, and strongly depends on cultivation conditions. Analytical instruments capable of quantifying pCBs are useful for optimizing cultivation practices and determining the optimal harvest time. In particular, when growing hemp, they can help avoid penalties if there is a risk that the THC content in the plant will exceed the threshold value. Also, such analysis is valuable for quality control of cannabis products rich in various pCBs since they can readily undergo chemical transformation, forming undesirable compounds.

Main methods for analysis of phytocannabinoids in cannabis and its products

Currently, gas chromatography (GC) and high-performance liquid chromatography (HPLC) are primarily used for the chemical analysis of cannabis. These methods can quantify various pCBs with high accuracy, precision, sensitivity, selectivity, and linearity. However, they are time-consuming, invasive, destructive to plant samples, and rely on expensive, complex, and bulky devices. Sample preparation requires infrastructure and chemicals capable of degrading pCBs. Consequently, these techniques are unsuitable for rapid, *in vivo* analysis in field conditions.

However, not all spectroscopic methods are well-suited to *in vivo* plant analysis. Fluorescence spectroscopy is strongly affected by non-cannabinoid fluorophores and is mainly used for extracts, whereas near-infrared (NIR) spectroscopy is strongly affected by water and is therefore more suitable for dried material or hyperspectral imaging under controlled conditions.

Molecular spectroscopy presents a viable alternative or complementary approach to GC/HPLC. This is largely because pCBs are synthesized and stored externally on the surface of aerial plant structures within specialized hair-like outgrowths known as “glandular trichomes”, allowing optical spectroscopy directly from the surface without prior sample preparation. However, not all spectroscopic techniques are suitable for *in vivo* analysis of living plants. Fluorescence spectroscopy suffers from the presence of fluorophores other than pCBs, so it is mainly used for the analysis of cannabis extracts. In contrast, near-infrared (NIR) spectroscopy is limited by strong water absorption and is more effective for dried material. NIR hyperspectral imaging can partially mitigate this limitation but requires expensive instrumentation, controlled conditions, and analysis of large plant areas.

A good alternative to NIRS may be Raman spectroscopy, which also shows high efficiency in the analysis of cannabis [9,10]. Compared to NIRS, the key advantage of this method is that, in the Raman spectrum, the water-associated peaks are far from those of pCBs, making it well-suited for *in vivo* measurements of fresh plants.

For Raman measurement, one of the main issues is the fluorescence background. Optimizing the optical scheme enables mitigating its manifestations. However, complete removal usually requires post-processing of the spectrum. This is facilitated by the fact that the fluorescence peaks are much broader than the Raman peaks, thereby enabling the development of effective correction algorithms.

Raman spectroscopy has been shown to differentiate plants with varying THC/CBD content using spectra from their flowers and leaves, or even from individual trichomes. In most cases, the main spectral characteristics are in the wavenumber range of 600-1800 cm^{-1}. However, direct observation of most pCB fingerprints is difficult due to the presence of strong matrix peaks associated with common plant compounds (carotenoids, aliphatic compounds, etc.), e.g., at 1003 cm^{-1}, 1156 cm^{-1}, 1525 cm^{-1}, and others.

In addition, pCBs have similar molecular structures (Fig. 1), and their characteristic bands often overlap, particularly between acidic and neutral forms, and with strong contributions from the plant matrix. For rapid plant differentiation and assessment of their major pCBs content, it would be convenient to work with the corresponding unique, separate Raman peaks. However, this is quite difficult because many characteristic peaks in their spectra are quite close in position.

As a result, reliable chemotype differentiation of cannabis plants requires analysis of the full Raman spectrum rather than individual marker peaks, making the combination of Raman spectroscopy with multivariate data analysis methods, including PCA and PLS-DA, essential for practical applications.

Lightnovo miniRaman spectrometer

miniRaman is a compact dispersive spectrometer with a fixed transmission diffraction grating that operates in reflection mode (180 ° backscattering) [7], shown in Fig. 2. The instrument uses a diode laser to generate excitation light, which is delivered to the sample via a specialized optical probe (a Raman probe). The same optical components collect backscattered light. A CMOS image sensor is used to capture Raman spectra.

Figure 2. Two miniRaman spectrometers and a cannabis leaf.

A key peculiarity of the Lightnovo miniRaman spectrometer is the presence of a dedicated in-built reference channel, which allows for the "on the fly" correction of the measuring Raman spectrum [11]. At the same time, neither moving components, additional light sources, nor image sensors are used. In combination with other technical solutions, the reference channel eliminates the need for expensive high-end lasers and image sensors, as well as the forced cooling of these components, thereby reducing the background fluorescence signal. As a result, the instrument exhibits excellent measurement characteristics and exceptional compactness.

The following features make miniRaman an attractive tool for chemical analysis of plants, with particular relevance for cannabis:

- **Wide spectral range.** In the basic configuration, it is ν = (400 ... 2700) cm^{-1}, while the main Raman peaks of key plant metabolites and structural components are in the range ν = (500 ... 1800) cm^{-1} [3,9,10].
- **Record low size-mass characteristics.** With dimensions of 112×39×34 mm^3 and a mass of 400 g (200 g in aluminum housing), the instrument is ideal for *in vivo* analysis of plants "in the field", and in particular for rapid testing of cannabis products on production lines, for the activities of various regulatory authorities, customs, police, etc.
- **Flexibility of use.** The spectrometer is equipped with a battery and standard interfaces: USB, Bluetooth, and Wi-Fi. This allows it to be used as:
  (i) a portable instrument for *in vivo* analysis of plants at indoor and outdoor plantations;
  (ii) benchtop instrument in a laboratory;
  (iii) a remote instrument, e.g., as part of an automatic system.
- At the same time, a **reliable connection to a workstation** can be provided to collect and process large volumes of data.
- A large number of Raman probes enables analysis of plant parts at different distances, including direct contact. There is also a **specialized vial accessory** that makes working with plant extracts and other liquid samples more reliable and comfortable.
- **Special, precise focusing stages** are available that allow the spectrometer to be used as part of a confocal miniRaman microscope. This enables the analysis of the distributions of secondary metabolites and other chemicals in plants.

Materials, measurements and data processing

Samples

This Application Note presents data for two types of cannabis leaves with chemotypes I (THC-dominant) and III (CBD-dominant), provided by Schroll Medical ApS (Årslev, Denmark). The plants were grown in accordance with Good Agricultural and Collection Practices (GACP) and harvested at 11 weeks.

Chemotype assignment of the studied plants was independently confirmed by HPLC analysis, showing chemotype I plants with ~4.0% total THC and <0.1% total CBD, and chemotype III plants with ~7.3% total CBD and <0.7% total THC (dry weight). Other major pCBs (CBC/CBCA and CBG/CBGA) were present at low levels (<1%), and the combined content of additional measured pCBs remained below 0.3% [4].

As reference, THC-distillate obtained from Echo Pharmaceuticals (Leiden, Netherlands) and a CBD standard in crystalline form from Cannabis Pharm (Randers, Denmark), both 99% pure, were used.

Measurements

Raman measurements were performed using the Lightnovo miniRaman spectrometer [12], a compact Raman system designed for rapid on-site measurements. Two instrument models were used: 785 nm excitation at 46 mW and 830 nm excitation at 76 mW. The spectral resolution of both devices was 10 cm^{-1}. Spectra were recorded in the spectral range 550-2000 cm^{-1}. The exposure time was set to 600 ms for the 785 nm device and 2 s for the 830 nm device. In both cases, 5-repetition averaging and no image sensor gain were used. Leaf measurements were performed using contact probes, ensuring stable sample positioning without the need for complex alignment. The CBD reference standard was measured in direct contact mode, while the THC distillate was analyzed in a vial using a probe with a 1.5 cm focal distance.

To account for natural biological variability, Raman spectra for each cannabis leaf were collected from multiple locations: three from the upper (adaxial) surface and three from the lower (abaxial) surface of the central leaflet, with measurements taken at the apical, central, and basal regions (inset to Fig. 4a). In total, more than 60 spectra were recorded from 10 leaves of each chemotype.

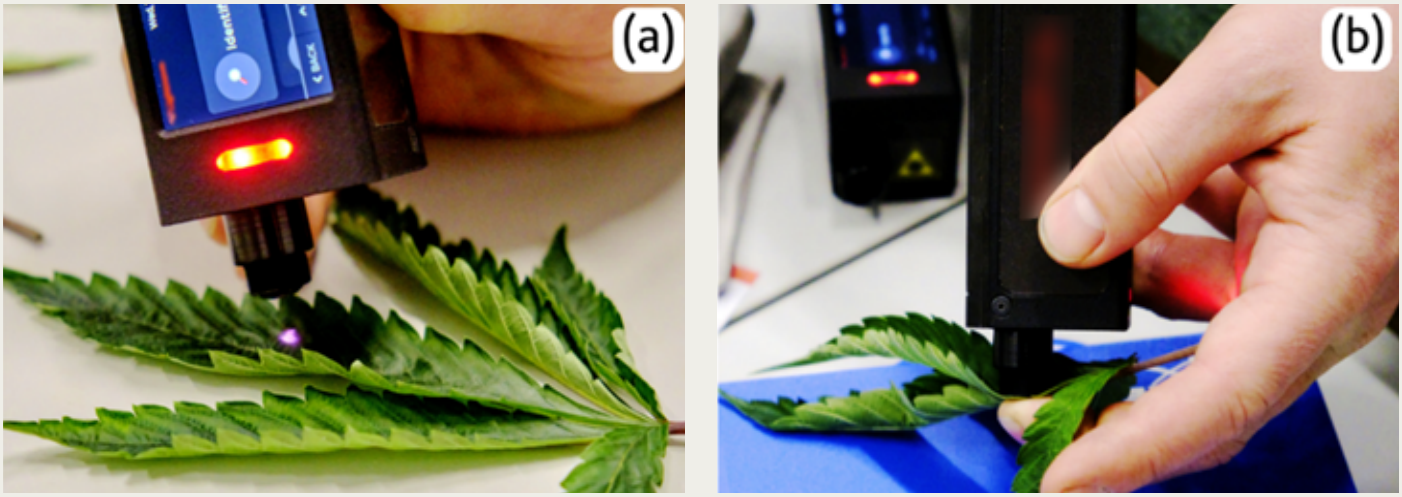

Figure 3. Measuring the spectrum from the adaxial (upper) surface of a cannabis leaf using the hand-held Raman spectrometer in (a) distant mode and (b) contact mode.

Raman spectra processing and analysis

The spectra were preprocessed using in-house MATLAB R2020b scripts (MathWorks, USA). First, Whittaker smoothing with $\lambda = 1$ was applied to reduce instrumental noise while preserving the narrow Raman peaks of cannabis. Second, Rubberband baseline correction was used to eliminate the contribution of fluorescence to the measured spectra. Finally, the Standard Normal Variate (SNV) scaling transformation was applied to each spectrum as a row-wise standardization, centering the data by subtracting the mean and scaling by the standard deviation to improve comparability across samples.

After preprocessing, PCA was used to decompose leaf spectra (whole dataset) for a preliminary data exploration. As an unsupervised dimensionality reduction method, PCA transforms complex spectral data into a set of orthogonal Principal Components (PCs). This allows visualization of natural groupings in the data and identification of the primary sources of spectral variance, specifically distinguishing between the dominant plant matrix (cellulose/lignin) and the weak pCB signatures.

Partial least squares discriminant analysis (PLS-DA)

To enable automated identification of cannabis chemotypes, PLS-DA was performed using Classification Toolbox 7.0 for MATLAB. In contrast to unsupervised methods such as PCA, PLS-DA is a supervised classification approach that uses prior knowledge of the sample classes (e.g., Chemotype I vs. Chemotype III) to maximize separation between them.

The PLS-DA algorithm reduces high-dimensional Raman spectral data to a smaller set of latent variables (LVs), which are linear combinations of the original wavenumbers, chosen to maximize the covariance between the spectral intensities (X-matrix) and the defined class labels (Y-matrix). This ensures the model focuses on spectral variances most relevant to the chemical differences between THC- and CBD-dominant plants, rather than general plant morphology.

A classification model was developed to distinguish spectral data from leaves of chemotypes I and III plants. Each dataset was randomly split into 70% for calibration (training and cross-validation) and 30% for the test set to evaluate the predictive performance of the developed model.

The optimal number of LVs was selected to minimize cross-validation prediction error [4]. Venetian blind cross-validation was used with five cross-validation groups. This step is critical to prevent "overfitting" and to ensure the model recognizes true molecular markers (such as the 1371 cm^{-1} THC peak or 1437 cm^{-1} CBD peak) rather than random noise.

Model performance was evaluated using standard classification metrics, including sensitivity (the correct classification of samples belonging to the class of interest) and specificity (the correct exclusion of samples not belonging to the class of interest). These metrics were evaluated for calibration, cross-validation, and test datasets.

Note that similar preprocessing and multivariate analysis methods, including spectral smoothing, baseline correction, normalization, PCA decomposition and PLS-based calibration are built-in features and can be performed directly in Miraspec software.

Results

Figures 4a and 4b show averaged Raman spectra of cannabis leaves registered at 785-nm and 830-nm excitations, respectively. The spectra were measured from the adaxial and abaxial surfaces of both chemotypes I and III. The mean processed spectra after smoothing, baseline correction, and SNV scaling are shown in Figs. 4c and 4d for chemotype III, and in Figs. 4e and 4f for chemotype I.

A strong fluorescence background, a characteristic of living plants, dominates the raw Raman spectra of cannabis leaves. As a matter of fact, plants exhibit strong fluorescence when excited in the visible spectral region. This fluorescence is mostly pronounced in the red range (650 - 750 nm) and decreases significantly toward the near-infrared region, which favors the use of long-wavelength Raman excitation for *in vivo* measurements. Therefore, using a longer excitation wavelength (830 nm) in our case reduces the fluorescence signal by 3-fold, making it more appropriate for the analysis of living plants.

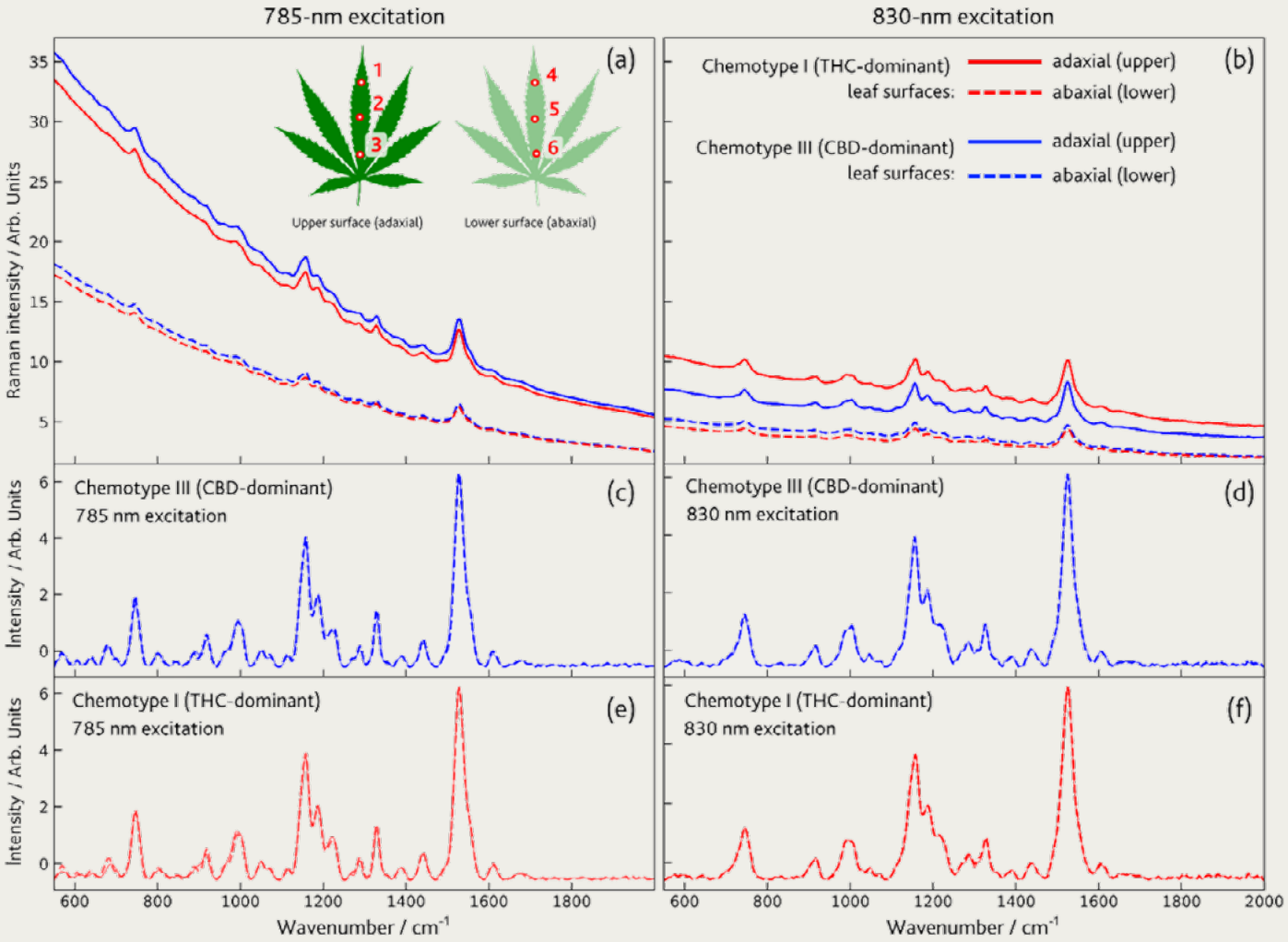

Figure 4. Raman spectra of cannabis leaves under 785 nm (left) and 830 nm (right) excitation.
(a)-(b): Averaged raw spectra from adaxial (solid line) and abaxial (dashed line) surfaces for chemotypes III (CBD-dominant) and I (THC-dominant).
(c)-(d): Smoothed, baseline-corrected, and SNV-scaled spectra for chemotype III.
(e)-(f): Processed spectra for chemotype I. Inset to (a) shows points of data collection from the leaf (adaxial, abaxial): apical (#1, #4); central (#2, #5); basal (#3, #6).

Nevertheless, both instruments allow Raman peaks to be clearly distinguished in the raw spectrum. It is worth noting that, for most industrial spectrometers operating at excitation wavelength ≤ 785 nm, raw spectra of cannabis have such a high background that the Raman spectrum's details become practically indistinguishable (e.g., see Figure 1 in [10]). In contrast, for miniRaman, many peaks are well-defined. This peculiarity is associated with the aforementioned reference channel, which allows partial fluorescence suppression immediately during the measurement.

Comparison of raw spectra suggests that the abaxial leaf surface produces a ~2-fold higher fluorescent signal, regardless of excitation wavelength or plant chemotype. This is most likely not related to pCB's properties, but rather due to morphological differences. In plants, fluorescence is thought to be due mainly to chlorophyll B and carotenoids (natural pigments). At the same time, the leaf surfaces differ in color, as can be seen in Fig. 3, so it can be assumed that they have various contents of carotenoids.

Spectral preprocessing with Rubberband baseline correction, following Whittaker smoothing, completely removes the fluorescent background for both spectrometer models and eliminates differences between the spectra of adaxial and abaxial surfaces, Figs. 4c-e.

Raman spectral features of cannabis leaves and products

The averaged Raman spectra of cannabis leaves of chemotypes I and III (Figs. 5c and 5d, respectively) exhibit a rich set of vibrational features typical of living plants. The spectra are dominated by strong plant matrix peaks at 1003 cm^{-1} (associated with carotenoids), 1156 cm^{-1} (cellulose), and 1524 cm^{-1} (lignin). These structural components mask weaker pCBs-associated signals, emphasizing the need for advanced spectral analysis methods to extract chemotype-specific features.

Reference Raman spectra of pure CBD powder (Fig. 5a) and THC distillate (Fig. 5b) provide useful information for identifying spectral features that may assist in differentiating THC- and CBD-dominant plants. Although the CBD powder spectrum represents a rich set of vibrational features consistent with literature data [9], only two of its peaks are most distinct in the spectra of leaves: 1339 cm^{-1} (associated with CH bending in CBD) and 1437 cm^{-1} (alkyl bending).

The spectrum of THC distillate contains typical THC-related peaks at 1371 cm^{-1} (CH bending) and 1625 cm^{-1} (aromatic C=C stretching), as well as a peak at 1664 cm^{-1} (carboxyl stretching), which is the THCA marker [9]. These features can also be traced in the leaves' spectra, but they are very weakly expressed against the background of other peaks. This prevents reliable chemotype discrimination based on individual peaks.

Raman spectra of THC distillate further illustrate the sensitivity of the method to chemical composition. In addition to Raman peaks reported for pure THC [9], the spectrum of THC distillate also shows pronounced peaks corresponding to CBD (1339 cm^{-1} and 1437 cm^{-1}) and, most probably, to CBG (988 cm^{-1} and 1309 cm^{-1}). This suggests that the distillate contains residual compounds, likely due to imperfect purification processes and/or partial degradation. Although this spectral complexity limits direct interpretation, it demonstrates the sensitivity and specificity of Raman spectroscopy for assessing purity and detecting residual compounds in cannabis products.

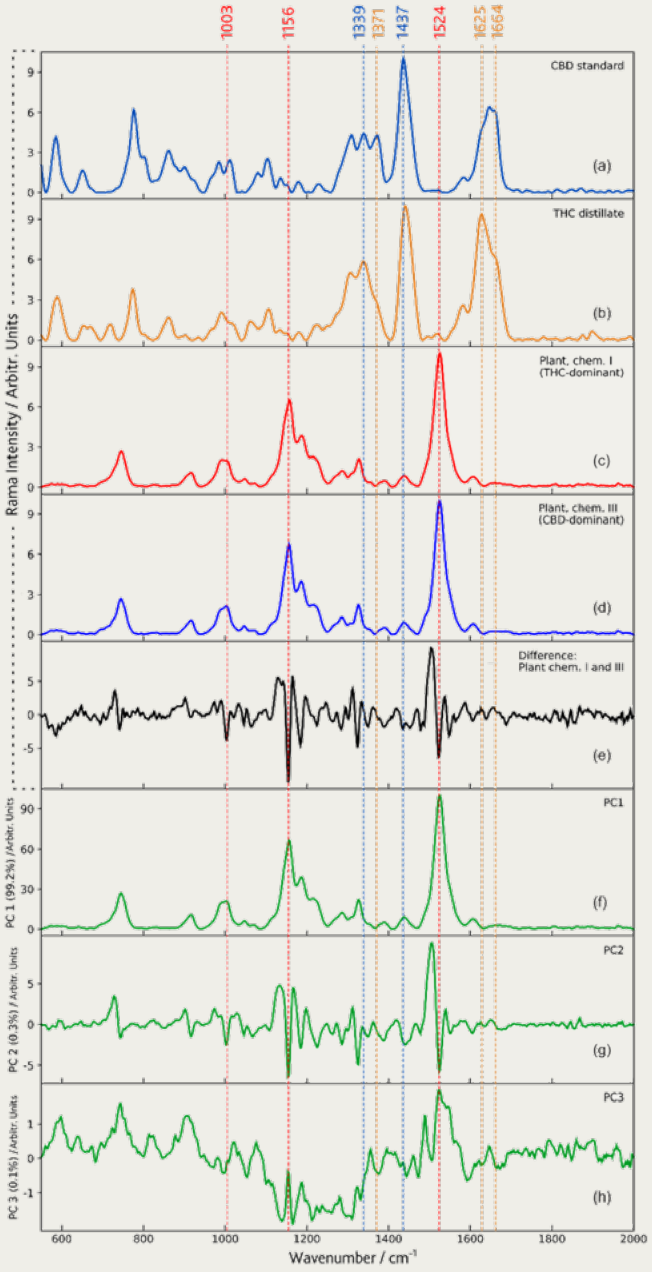

Figure 5. Raman spectra at 830 nm excitation for:
(a) CBD standard;
(b) THC-distillate;
(c) Chemotype I (THC-dominant) leaf;
(d) Chemotype III (CBD-dominant) leaf.
Panel (e) shows the difference spectrum between chemotypes I and III;
(f-h) PCA loadings spectra.
Relevant peaks distinguishing THC from CBD are highlighted: THC-associated (1371 cm^{-1}, 1625 cm^{-1} and 1664 cm^{-1}; brown), and CBD-associated (1339 cm^{-1} and 1437 cm^{-1}; blue). Key plant matrix peaks at 1003 cm^{-1} (carotenoids), 1156 cm^{-1} (cellulose), and 1524 cm^{-1} (lignin) are also indicated (red).

Table 1. Representative Raman bands relevant for cannabis leaf analysis and chemotype differentiation.

Raman shift (cm^{-1})	Major assignment	Observed in leaves	Analytical relevance
~775-790	THC/CBD/CBG ring modes	Weak-moderate	pCB fingerprint region
1003	Carotenoids, protein	Strong	General plant matrix; normalization
1156	Cellulose/carotenoids	Strong	Plant health; stress and pigment content
1273	CBD	Weak-moderate	CBD chemotype marker
1339	CBD	Weak–moderate	CBD chemotype marker
1371	THC	Weak–moderate	THC chemotype marker
1437	CBD (CH_2/CH_3 bending)	Moderate	Key CBD discriminator
1524	Lignin / carotenoids	Strong	Plant health / stress indicator
1625	THC/CBD (C=C stretching)	Weak	pCB-related feature contributing to chemotype differentiation
1646	CBD	Weak	CBD chemotype marker
1664	THC/THCA	Weak	THC chemotype marker

The most important Raman bands for cannabis leaf analysis and chemotype differentiation are summarized in Table 1, including plant matrix features and key THC- and CBD-associated bands. A complete set of 43 identified peaks for the spectra in Fig. 5, along with their attributions, can be found in ref. [4].

Additional contrast between THC- and CBD-dominant plants can be achieved by comparing the difference spectra of chemotypes I and III (Fig. 5e), where distinctive features of THC and CBD are amplified against high-amplitude peaks corresponding to the plant matrix.

Multivariate analysis further clarifies weak differences between chemotypes of cannabis plants. In particular, PCA applied to the 830-nm dataset results in PC1 explaining 99.2% of the variance, while PC2 and PC3 account for 0.3% and 0.1%, respectively. As shown in the PCA score plots (Fig. 6), PC1 primarily reflects plant matrix contributions with no significant separation between chemotypes. However, PC2 and PC3 captured key spectral differences, highlighting 1437 cm^{-1} (CBD) and 1371 cm^{-1} (THC) as dominant contributors (Fig. 5, f-h). The PC2 loadings spectrum closely resembles the difference spectrum, demonstrating that PCA successfully isolates chemotype-specific spectral variance. As a result, the two chemotypes show only partial separation in PCA space, indicating that chemotype-specific information is present but masked by the strong contribution of plant matrix components. Therefore, supervised methods such as PLS-DA are required for reliable chemotype classification.

In addition, PCA analysis (Fig. 6) confirms that spectra from different leaf sides (adaxial and abaxial) and regions (apical, central, and basal) did not cluster separately after preprocessing, indicating negligible variability due to these factors. Therefore, spectra obtained from all sides/regions for all leaves can be pooled and treated as equivalent.

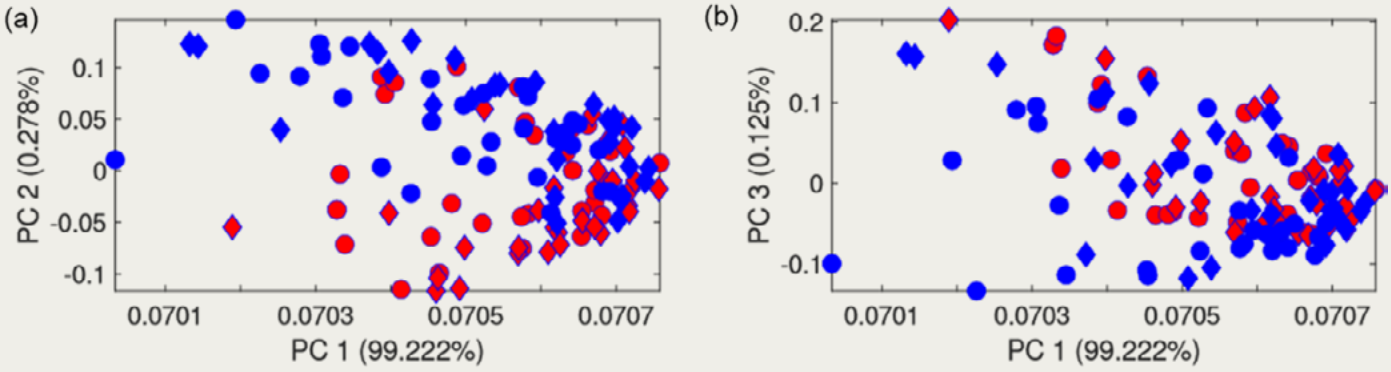

Figure 6. PCA score plots (830 nm excitation): (a) PC1 vs PC2 and (b) PC1 vs PC3. Data points correspond to chemotype III (CBD-dominant, red) and chemotype I (THC-dominant, blue) cannabis leaves; diamond symbols denote measurements from the adaxial (upper) leaf surface, and open circles denote the abaxial (lower) surface.

Chemotype differentiation of cannabis plants using PLS-DA classification

PLS-DA enables direct chemotype classification of cannabis plants based on their leaf Raman spectra. Classification models were developed to distinguish spectral data from leaves of plants of chemotypes I and III. The datasets included 130 spectra at 830 nm (60 for chemotype I and 70 for chemotype III) and 168 spectra at 785 nm (108 for chemotype I and 60 for chemotype III).

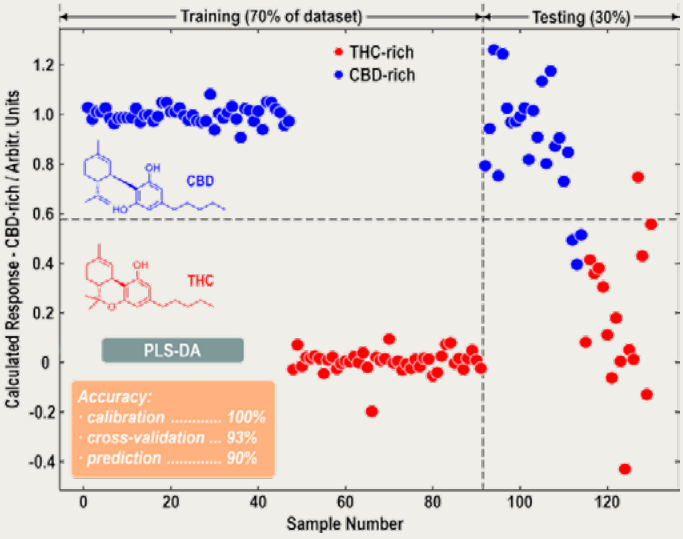

Figure 7. PLS-DA chemotype classification results: calculated response (class: chemotype III) for the 830-nm dataset, showing separation of chemotype I (THC-dominant, red) and chemotype III (CBD-dominant, blue): 70% - model training and cross-validation; 30% - testing. The prediction accuracy is 90% for the 830-nm dataset.

The PLS-DA model response for chemotype III obtained with the 830-nm dataset (Fig. 7) demonstrates clear separation between THC-dominant (chemotype I) and CBD-dominant (chemotype III) plants. The model with 830-nm excitation achieves approximately 90% prediction accuracy on the independent test set, while the 785-nm model reaches approximately 80%. A summary of model performance metrics for both datasets is shown in Table 2.

The optimized number of LVs was five for the 785-nm dataset and twelve for the 830-nm dataset, which reflects differences in signal-to-background ratio and fluorescence contribution between the two excitation wavelengths. Corresponding error rate plots are provided in the ref. [4].

Therefore, the PLS-DA model allows the extraction of weak but systematic spectral differences associated with pCB composition that are otherwise masked by dominant plant matrix signals, such as cellulose, lignin, and carotenoids. As a result, each Raman measurement can be reliably classified as “THC-dominant” or “CBD-dominant,” effectively converting complex Raman data into a simple and practical diagnostic tool.

Table 2. Performance metrics of Raman PLS-DA chemotype identification for Chemotype I (THC-dominant) and Chemotype III (CBD-dominant) cannabis leaves using 785 nm and 830 nm excitation.

Excitation λ	785 nm		830 nm	
Chemotype	Chemotype I (THC-dominant)	Chemotype III (CBD-dominant)	Chemotype I (THC-dominant)	Chemotype III (CBD-dominant)
Sensitivity (%)	90	65	94	87
Specificity (%)	65	90	87	94
Precision (%)	79	81	83	95
Accuracy (%)	80		90	

Conclusion

This Application Note demonstrates the capabilities of Lightnovo's miniRaman spectrometer for reliable *in vivo* analysis and chemotype differentiation of cannabis plants. Using two compact Raman instruments with 785- and 830-nm excitation enables chemotype classification accuracy of up to 90% based on Raman spectra collected from living plant leaves and analyzed using standard preprocessing and PLS-DA classification. In addition, the recorded spectra exhibit multiple pCB-associated features, suggesting strong potential for future pCB profiling and quantitative analysis with appropriate chemometric models. Moreover, intense Raman peaks conditioned by non-pCBs constituents by themselves could likely be used to differentiate plants by sex. The combination of non-invasive measurements, portability of the miniRaman spectrometer, and rapid spectral acquisition makes this approach particularly well-suited for field-based applications, including chemotype screening, pCB profiling, quality control, and forensic identification.

References

1. E. Jain, M. Rose, P.K. Jayapal, G.P. Singh, R.J. Ram, Harnessing Raman spectroscopy for the analysis of plant diversity, Sci. Rep. 14 (2024) 12692. https://doi.org/10.1038/s41598-024-62932-0.
2. M. Park, A. Somborn, D. Schlehuber, V. Keuter, G. Deerberg, Raman spectroscopy in crop quality assessment: focusing on sensing secondary metabolites: a review, Hortic. Res. 10 (2023). https://doi.org/10.1093/hr/uhad074.
3. W.Z. Payne, D. Kurouski, Raman spectroscopy enables phenotyping and assessment of nutrition values of plants: a review, Plant Methods. 17 (2021) 78. https://doi.org/10.1186/s13007-021-00781-y.
4. D. Komisar, A. Kutsyk, O. Vasyliev, Y. Aulin, S. Boisen, K. Stergiou, Y. Pilhun, L. Duelund, R. Bahij, M.A.B. Hedegaard, O. Ilchenko, Portable Raman Spectroscopy for Non-Invasive Chemotype Identification and pCB Profiling in Cannabis, J. Raman Spectrosc. 56 (2025) 1498–1507. https://doi.org/10.1002/jrs.70033.
5. A. Saletnik, B. Saletnik, G. Zaguła, C. Puchalski, Raman Spectroscopy for Plant Disease Detection in Next-Generation Agriculture, Sustainability. 16 (2024) 5474. https://doi.org/10.3390/su16135474.
6. H. Schulz, M. Baranska, Identification and quantification of valuable plant substances by IR and Raman spectroscopy, Vib. Spectrosc. 43 (2007) 13–25. https://doi.org/10.1016/j.vibspec.2006.06.001.
7. https://lightnovo.com/miniraman-spectrometer/, (n.d.).
8. R.M. Strongin, J. Meehan-Atrash, M. Vialpando, Recent Advances in the Science of Cannabis, CRC Press, Boca Raton, 2021. https://doi.org/10.1201/9780429274893.
9. L. Sanchez, D. Baltensperger, D. Kurouski, Raman-Based Differentiation of Hemp, Cannabidiol-Rich Hemp, and Cannabis, Anal. Chem. 92 (2020) 7733–7737. https://doi.org/10.1021/acs.analchem.0c00828.
10. S. Porcu, E. Tuveri, M. Palanca, C. Melis, I.M. La Franca, J. Satta, D. Chiriu, C.M. Carbonaro, P. Cortis, A. De Agostini, P.C. Ricci, Rapid In Situ Detection of THC and CBD in Cannabis sativa L. by 1064 nm Raman Spectroscopy, Anal. Chem. 94 (2022) 10435–10442. https://doi.org/10.1021/acs.analchem.2c01629.
11. R. Ilchenko, Oleksii; Rindzevicius, Tomas; Boisen, Anja; Slipets, An apparatus for carrying out raman spectroscopy, WO2019145005, 2019.
12. O. Ilchenko, Y. Pilhun, A. Kutsyk, D. Slobodianiuk, Y. Goksel, E. Dumont, L. Vaut, C. Mazzoni, L. Morelli, S. Boisen, K. Stergiou, Y. Aulin, T. Rindzevicius, T.E. Andersen, M. Lassen, H. Mundhada, C.B. Jendresen, P.A. Philipsen, M. Hædersdal, A. Boisen, Optics miniaturization strategy for demanding Raman spectroscopy applications, Nat. Commun. 15 (2024) 3049. https://doi.org/10.1038/s41467-024-47044-7.

Contact Details

Lightnovo ApS

Blokken 11, 1.
3460 Birkerød
Denmark (DK)

+45 61 46 37 10
info@lightnovo.com
CVR: 40979603

Scan to learn more about this and other products

Please consult your local sales representative for details