RAMAN SPECTROSCOPY ENABLES HIGHLY ACCURATE DIFFERENTIATION BETWEEN YOUNG MALE AND FEMALE HEMP PLANTS

An Undergraduate Research Scholars Thesis

by

SAMANTHA HIGGINS

Submitted to the LAUNCH: Undergraduate Research office at
Texas A&M University
in partial fulfillment of requirements for the designation as an

UNDERGRADUATE RESEARCH SCHOLAR

Approved by
Faculty Research Advisor: Dr. Dmitry Kurouski

May 2022

Major: Genetics

RESEARCH COMPLIANCE CERTIFICATION

Research activities involving the use of human subjects, vertebrate animals, and/or biohazards must be reviewed and approved by the appropriate Texas A&M University regulatory research committee (i.e., IRB, IACUC, IBC) before the activity can commence. This requirement applies to activities conducted at Texas A&M and to activities conducted at non-Texas A&M facilities or institutions. In both cases, students are responsible for working with the relevant Texas A&M research compliance program to ensure and document that all Texas A&M compliance obligations are met before the study begins.

I, Samantha Higgins, certify that all research compliance requirements related to this Undergraduate Research Scholars thesis have been addressed with my Research Faculty Advisor prior to the collection of any data used in this final thesis submission.

This project did not require approval from the Texas A&M University Research Compliance & Biosafety office.

TABLE OF CONTENTS

ABSTRACT

Raman Spectroscopy Enables Highly Accurate Differentiation Between Young Male and Female Hemp Plants

Samantha Higgins
Department of Biochemistry and Biophysics
Texas A&M University

Research Faculty Advisor: Dr. Dmitry Kurouski
Department of Biochemistry and Biophysics
Texas A&M University

Determination of hemp (*Cannabis sativa*) sexuality is an integral part of the hemp industry. Current methods of analysis can be costly and invasive. Genetic testing is most often required which involves sending leaf samples to a laboratory for results. These wet laboratory methods require taking samples from immature hemp plants which can pose a risk due to the sensitivity of the young plants. Not only can collecting leaf samples damage the plant but the wait time to see results of genetic testing can take valuable time. This issue in the hemp industry can be solved by the emerging technology of Raman spectroscopy. Raman spectroscopy provides a way to accurately and non-invasively differentiate between young male and female hemp plants. A portable, hand-held Resolve Agilent Raman spectrometer was used as our instrument of analysis which does not disturb the plants in any way.

Highly accurate and nonintrusive hemp differentiation is exceedingly important to hemp growers due to the preference for female hemp. Female hemp plants have a higher concentration of cannabinoids than male plants. Current efforts to minimize cross-pollination are not as

effective as predetermining males or females. Hemp producers seeking higher cannabinoid concentrations desire fields that contain 100% female plants. If a male plant is within the field, it can unwantedly cross-pollinate female plants thus not allowing for the desired amount of cannabinoid production. Our results show that through Raman Spectroscopy, mature male and female hemp can be distinguished with an accuracy of 94% and even more importantly, that young hemp plants can be differentiated with 90% accuracy. In conclusion, our findings will allow hemp growers to save valuable time and expenses. This discovery broadens the many applications of Raman spectroscopy and can expand how plant sex determination is being conducted.

DEDICATION

To my family, friends, peers, and Dr. Kurouski who have never ceased to encourage me through my academic journey and this research process.

ACKNOWLEDGEMENTS

Contributors

I would like to extend my thanks to my faculty advisor, Dr. Kurouski for his encouragement, guidance, mentorship, and support during my time in his laboratory and throughout this research program.

I would also like to thank my friends, peers, department faculty, and staff for their support and for making my time at Texas A&M University a joyous experience that will forever be held dear in my heart. I cannot express enough the overwhelming support and inspiration I have found at this University.

Finally, thanks to my family for never ceasing to believe in me and for all their love.

The HPLC data in Figure 3.4 for RAMAN SPECTROSCOPY ENABLES HIGHLY ACCURATE DEFFERENTAION BETWEEN YOUNG MALE AND FEMALE PLANTS were provided by Dr. Stanislav Rizevsky. The HPLC analyses depicted in RAMAN SPECTROSCOPY ENABLES HIGHLY ACCURATE DEFFERENTAION BETWEEN YOUNG MALE AND FEMALE HEMP PLANTS were conducted in part by Dr. Stanislav Rizevsky and these data are unpublished.

Thank you to Dr. Jessup for growing and providing the hemp plants for analysis in College Station, Texas.

The Raman spectroscopy Figure 3.1 for RAMAN SPECTROSCOPY ENABLES HIGHLY ACCURATE DEFFERENTAION BETWEEN YOUNG MALE AND FEMALE HEMP PLANTS was provided by Dr. Kurouski, faculty advisor. The data for RAMAN

SPECTROSCOPY ENABLES HIGHLY ACCURATE DEFFERENTAION BETWEEN YOUNG MALE AND FEMALE HEMP PLANTS in Table 3.1 were provided by Dr. Kurouski.

All other work conducted for the thesis was completed by the student independently.

Funding Sources

Undergraduate research was supported by Texas A&M AgriLife Research at Texas A&M University.

This work was also made possible in part by Texas A&M University Governor's University Research Initiative (GURI) under Grant Number 12-2016/M1700437. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the GURI.

NOMENCLATURE

RS	Raman Spectroscopy
HPLC	High-Performance Liquid Chromatography
GURI	Governor's University Research Initiative
THC	Tetrahydrocannabinol
CBD	Cannabidiol
CBG	Cannabigerol
PCR	Polymerase Chain Reaction
LUT	Lutein
CHR	Chlorophyll
LYC	Lycopene
BCR	Beta-cryptoxanthin
ZEA	Zeaxanthin
BCA	Beta-carotene
PLS-DA	Partial Least Squares Discriminant Analysis
TPR	True Positive Rate

1. INTRODUCTION

1.1 Background

Throughout the last few years, the hemp (*Cannabis sativa*) industry has been growing exponentially with an expected revenue upwards of $1.9 billion by 2022[1]. As one can tell from this prediction, the hemp market is on the rise. Therefore, it is of great importance to have scientific technology that will save these hemp producers time and expenses. The hemp plant is deciduous and is also classified as an angiosperm, meaning that there are two genders of the hemp plant. Once the male and female plants become mature enough to flower, the female plants are pollinated by the male plants and so the next generation of hemp is produced.

Within the hemp industry, female plants are desired due to their higher concentration of cannabinoids. To keep the preferred cannabinoid concentration, hemp producers attempt to plant only females in a production field. Doing so ensures the highest amount of cannabinoid production that is possible for harvesting. *Cannabis sativa* contains Δ^9-tetrahydrocannabinol (Δ^9-THC), cannabidiol (CBD), and cannabigerol (CBG) [2,3]. When a hemp plant contains more than 0.3% of Δ^9-THC the plant is then referred to as marijuana, which is an illicit drug in some states, however, it is medicinally legal in many states due to its analgesic property[4].

There are many purposes of the hemp plant aside from the psychoactive Δ^9-THC and pain-killing effects such as the medicinal purposes of CBD and CBG. Studies show that CBG can reduce inflammation, has antimicrobial properties, and can assist in pain relief, whereas CBD is shown to aid in sleep, relieve pain, reduce nausea, and lessen anxiety and depression[3,5,6]. These are very attractive benefits as CBD and CBG have no psychotropic effects[7]. Therefore, the cultivation of hemp plants with less than 0.3% of THC is needed to keep growth legal and to

culture plants with as much CBD concentration as possible to use for manufacturing products with the medicinal benefits as discussed above. Since female plants are known to produce higher concentrations of CBD when unpollinated, known as sinsemilla, hemp producers want a field of female plants without the risk of crosspollination[8]. Cross-pollination can occur rapidly and have quite a large impact on the hemp crop as hemp is an anemophilous plant, meaning pollen can be spread by the wind, quickly covering a vast area of land[9]. There are methods to influence the gender of the plant such as the application of outside growth regulators and various chemicals that affect the plants' hormones[10]. Especially in the first few weeks of growth, hemp is extremely sensitive to exogenous growth factors and can be influenced by external sources.

In a recent study, spraying colloidal silver on female plants induced the growth of male flowers and pollen[11]. The resulting pollination of only female plants is called "feminization" and the next generation produced will be female as desired by the hemp famers[11]. While this is an attractive means to producing a field of 100% female hemp plants, it is a tedious process.

One can determine the gender of a hemp plant by genetic testing. However, hemp plant sex is a trait that is polygenetic meaning the phenotypic expression of gender is determined by more than one gene[12]. The sex of hemp plants is also greatly influenced by epigenetic traits making sex determination difficult through genetic testing when there is no control to the exogenous factors the plants might be influenced by[10]. This makes testing a difficult process due to the high variability of the gender of hemp plants arising from chromosomal components and epigenetic factors.

The problem that arises within the hemp industry is in identifying exactly what hemp plants are being planted in the field. Male and female hemp plants appear visually identical when they are still very young. Thus, planting a field of hemp without being sure of the gender of the

plant beforehand can result in unwanted cross-pollination that leads to less cannabinoid production. To prevent this from occurring, the male and female plants must first be differentiated before planting.

1.1.1 Current Methods of Differentiation Between Male and Female Hemp

Current methods for determining the sex of hemp plants are invasive and can be damaging and costly or they are extremely time-consuming and tedious work. Most methods for differentiating between male and female hemp require genetic testing to determine XX homogametic chromosomes (female) or XY heterogametic chromosomes (male)[11]. Genetic testing first requires samples to be collected from the hemp plant which can be damaging to its growth and even then, cannot be extremely accurate if there is the variability of the growth environment that could predetermine plant gender. Male hemp plants have specific genetic markers that can be identified through the use of PCR[13]. Though this method requires sending samples to a laboratory for analysis which can be expensive and will require much time for results depending on factors such as transportation, the current work being done by the laboratory, and the number of hemp samples to be tested.

Visualization of the gender-specific phenotypical characteristics of hemp plants is another option for differentiating between male and female plants. However, as one can imagine this would be a heavily laborious process as hemp experts would need to examine each hemp plant before planting them in a field. This method could take too much time to cover an entire crop of hemp and thus lead to the maturation of the hemp plant before even identifying all specimens and posing risk for cross-pollination. It is highly impractical to visually determine the gender of hemp plants and so it is necessary to use technology as the demand grows for an increased hemp supply.

1.2 Raman Spectroscopy

The issues that arise from the current methods of hemp sex differentiation can be solved by using Raman spectroscopy (RS). Raman spectroscopy is a spectroscopic technique that is non-damaging, label-free, and non-invasive[14]. This means of spectroscopic analysis by inelastic light scattering was first assumed in 1923 by Semkal using quantum theory[15]. Five years later this theory was confirmed experimentally by Raman and Krishnan[16]. Thus, was born Raman spectroscopy. The phenomenon of RS works by the inelastic scattering of photons as the light meets molecular components.

The contact of liquid or gas with laser light results in an altered molecular state composed of vibrations or rotational vibrations, and electronic excitement[17]. When a solid comes into contact with the laser light, there is an electronic excitation and optical phonons, which is a unit of vibrational energy that occurs when atoms within a crystalline structure oscillate[17,18]. These molecular excitations and vibrations can then be translated into a spectrum that shows the vibrational peaks which are unique to each analyzed specimen[19]. These spectra of each analyzed sample act as a fingerprint of sorts that is explicitly produced by the sample analyzed.

Many experiments before have utilized Raman spectroscopy with extraordinary results. It has been previously determined that RS can be used as a diagnostic tool for abiotic and biotic stresses in plants, determine the presence of diseases in plants, differentiate between cannabis and hemp, and many other applications[4]. This is extremely attractive to horticulturists for the many uses of RS. This technology would be especially attractive to hemp farmers for more efficient and faster sex differentiation in hemp plants.

The spectral fingerprint allows scientists to identify exact chemical components of analytes which can identify, differentiate, and provide insight into the sample of interest. The

sample of interest for this experiment was the hemp plant. The question of how hemp producers can know that there is 100% female hemp in the field in a time-efficient and non-destructive manner, without risk of cross-pollination can finally be answered. The solution to this overarching problem is RS. Raman spectroscopy is the emerging technology that will save many hemp farmers time, funding, and protect their crops from being damaged by invasive sexing procedures. We found that with high accuracy Raman spectroscopy was able to distinguish the male and female sex of the hemp plants with no damage at all to the fragile specimens.

1.3 HPLC Analysis

To corroborate the results of the Raman spectral analysis of the hemp plants, HPLC was used to determine the molecular components in the male and female hemp samples. HPLC, high-performance liquid chromatography, is a method of analysis that uses a column and highly pressurized sample dissolved in a mobile phase that runs through the column containing the stationary phase[20]. As the sample runs through the column it exits based on its size and polarity thus allowing the analyte to be determined.

From this method of analysis, we were able to determine which carotenoids are within the hemp plant leaves. After determination of the specific carotenoids within the hemp plant leaves, we were able to use statistical analysis to see which carotenoids are significantly different between the male and female hemp plants.

1.4 Key Findings

For this experiment, a hand-held Resolve Agilent Raman spectrometer was used to take spectra from young and mature hemp plants. This non-invasive process allowed analysis of the hemp plants before and after flowering. By scanning the young plants before they can be visually distinguished and then later identifying the sex of the plant upon flowering and taking more

spectra, we could distinguish male from female hemp plants in young and adult hemp. Young hemp gender can be differentiated with an accuracy of about 90%, while adult hemp gender can be differentiated with about 94% accuracy. This is an extremely significant discovery as now there is a portable, non-invasive, label-free technology that differentiates between male and female hemp plants with high accuracy.

Through HPLC analysis we discovered that the concentration of most carotenoids is nearly equivalent, and the concentrations of chlorophyll and beta-carotene are the same in both male and female hemp plants. However, lutein is significantly different between male and female plants. This and other factors can be attributed to the results of our Raman spectroscopic analysis.

2. METHODS

2.1 Data Collection of Hemp Plants

The hemp plants that Raman spectra were collected from, were grown in a greenhouse in College Station, Texas. The ones that were part of this experiment were all the same hemp variety and all planted at the same time. The plants were also grown in the same environment to ensure optimal growth. These measures were taken to limit any extraneous factors that could cause variability within the experiment. Five spectra were collected from the adaxial side of each hemp plant. After the hemp plants matured, the same number of Raman spectra were collected from the adult hemp plants. Also, two to three leaves, about 150 mg, from three female plants and three male plants were collected, placed in Ziploc bags, and stored in a -20°C freezer for future HPLC analysis.

2.2 Raman Spectroscopy

A hand-held Resolve Agilent Raman Spectrometer with an 831 nm laser was used to collect spectra from young and adult hemp plants. The power on the spectrometer was set to 495 mW power. The acquisition time was one second. The Raman spectrometer is preset to perform baseline subtraction from each spectrum. No damage occurred upon collection of these spectra.

2.3 Multivariate Statistical Analysis

To statistically analyze the Raman spectra collected from young and adult hemp plants, a PLS_Toolbox in MATLAB was used to quantify results. The young hemp spectra were labeled and sorted by the software once the correct gender of the plant was visually determined. The adult hemp spectra were also labeled and sorted within the software. A partial least squares discriminant analysis (PLS-DA) was used on the young hemp and adult hemp spectra.

To preprocess the spectra the second derivative of the intensities was taken using a second-order polynomial and filter length of 15, followed by mean and median centering. A cross-validation matrix was used. After, a true positive prediction rate (TPR) was generated for how accurately male and female hemp plants could be differentiated. This method of statistical analysis provided non-bias results of the spectral comparisons and evaluation within the PLS-DA Toolbox.

2.4 Carotenoid Extraction

To analyze the hemp plant leaves using HPLC, carotenoids were first extracted. Three male hemp plants and three female hemp plants were used for analysis. Two to three leaves were collected from each plant to be used (about 150 mg from each). To create a homogenous liquid solution of the hemp carotenoids, 0.15 g of leaves were ground using a mortar and pestle with 0.75 ml of 2:1 chloroform dichloromethane solution[21].

Upon complete homogenization, the solution was added to a vial and suspended in 1.5 ml of the 2:1 chloroform dichloromethane mixture. This mixture was then vortexed at 500 rpm for thirty minutes at 4°C temperature[21]. After, 0.5 ml of 1 M sodium chloride was added to the mixture and then centrifuged at 5000 g for ten minutes[21]. After centrifugation, the aqueous and organic phases were separated. The aqueous phase was then suspended in 0.75 ml of 2:1 chloroform dichloromethane mixture and centrifuged again at 5000 g for 10 minutes [21]. In the second round of separation, the aqueous phase was discarded and both organic phases were mixed. Centrifugal evaporation was used to dry the sample into a pellet. The dried pellet was then redissolved in a solution of 200 µL of methanol/tert-methyl butyl ether (MTBE) (60/40, v/v). The re-dissolved sample was used for HPLC injection.

2.5 HPLC

A Waters 1525 pump with both a 2707 autosampler and a 2489 photodiode array detector (PDA) was used for the HPLC analysis of the extracted leaf carotenoids. A reverse-phase C30, 3 μm column with dimensions 250 x 4.6 mm was utilized to separate the carotenoids (Thermofisher, part number, 075723). The two mobile phases were (A) methanol and water in a 95:5 ratio by volume and (B) tert-methyl butyl ether. A 97% (A) and 3% (B) elution gradient was used with a set linear increase of B to 100% at twenty minutes. Following this, the elution gradient was set to initial parameters after twenty-three minutes. A constant temperature of 20°C was maintained in the column. Diode-array detectors (PDA) monitored elution peaks. To quantify results software, namely Breeze, was used to analyze the peak area of sample peaks with those of standard reference peak areas.

3. RESULTS

3.1 Raman Spectroscopy

The hemp industry is expanding at a rapid rate. According to *Hemp Industry Daily*, the expected economic growth of the hemp market is projected to bring in revenue of around $16 billion by 2025[22]. Not only is the hemp industry forecasted to bring in a large income, but there are also many uses of the hemp plant. The hemp plant, which originated in Asia, can adapt to many environments, and has a wide variety of uses[23,24]. Among the many uses of hemp, it is known that cannabis can be beneficial in alleviating symptoms of disease, improving overall health, and acting as a calming and pain-reducing agent[25]. Though cannabis contains a higher concentration of THC than hemp, hemp contains very similar medicinal properties.

The hemp industry faces a major issue in the differentiation of male and female hemp plants. Current methods are costly and time-consuming. There is also a risk of damage to the hemp plants when collecting leaf samples to be sent off for genetic testing. To solve this problem, an experiment was conducted to see if Raman spectroscopy could differentiate between male and female hemp plants. The hemp plants used for analysis were grown in a greenhouse in College Station, Tx.

Raman spectra were collected twice on the same group of hemp plants. Spectra were collected from each hemp plant at two weeks old and again at four weeks old, after the maturation of the plant. There were five spectra collected from each young hemp plant and mature plant. To do this a portable hand-held Resolve Agilent Raman spectrometer was used to collect spectra on the hemp leaves. After the first round of spectral collection, the plants remained in the same growth conditions to reach the flowering stage. At four weeks, the hemp

plants were mature enough to collect the adult hemp spectra. After the spectra were collected, two to three leaves totaling ~150 mg were collected from three male plants and three female plants. These collected leaves were placed in separate Ziploc bags, one bag for each plant to be sampled and frozen at -20°C for HPLC analysis.

The resulting spectra of the Raman analysis will show the vibrational bands of the carotenoids. Certain bands also signify the presence of aromatic compounds and cellulose. Aliphatic vibrations cannot be assigned to any specific molecular components of the hemp leaves however the vibrational bands rage between 1280-1440cm^{-1}. The carotenoid vibrations can be assigned to the following Raman shifts 1000 cm^{-1}, 1115-1218 cm^{-1}, and 1525 cm^{-1}. The Raman shift assignments for cellulose is 747 cm^{-1}, 915 cm^{-1}, and 1047 cm^{-1}. The aromatic compounds have a Raman shift between 1601 cm^{-1} and 1630 cm^{-1}. The assignments to the vibrational bands can be found in Table 3.1.

Table 3.1: Vibrational band assignments of the components found within the hemp leaves. Table provided with permission by Dr. Kurouski.

Band	Vibrational mode	Assignment
747	γ(C–O-H) of COOH	Pectin[26]
915	ν(C-O-C) In plane, symmetric	Cellulose, lignin[27]
1000	-C=C- (in plane)	Carotenoids[28]
1047	ν(C-O)+ν(C-C)+δ(C-O-H)	Cellulose, lignin[27]
1115	-C=C- (in plane)	Carotenoids[27]
1155	-C=C- (in plane)	Carotenoids[28]
1185	ν(C-O-H) Next to aromatic ring+σ(CH)	Carotenoids[29]
1218	δ(C-C-H)	Carotenoids[29]
1288	δ(C-C-H)	Aliphatics[30]
1326	δCH_2 Bending	Aliphatics, cellulose, lignin[27]
1382	δCH_2 Bending	Aliphatics[30]
1440	δ(CH_2)+δ(CH_3)	Aliphatics[30]
1525	-C=C- (in plane)	Carotenoids[31,32]
1601-1630	ν(C-C) Aromatic ring+σ(CH)	Lignin[33,34]

The intensities of the carotenoids (1000 cm^{-1}, 1115-1218 cm^{-1}, and 1525 cm^{-1}), aromatic compounds (between 1601 cm^{-1} and 1630 cm^{-1}), and cellulose (747 cm^{-1}, 915 cm^{-1}, and 1047 cm^{-1}) are higher in female hemp plants than in male hemp plants as seen in Figure 3.1. These vibrational intensity differences suggest that the concentrations of carotenoids differ in hemp plants depending on the gender of the plant. These differences detected by Raman spectroscopy are what can be used to differentiate males from females.

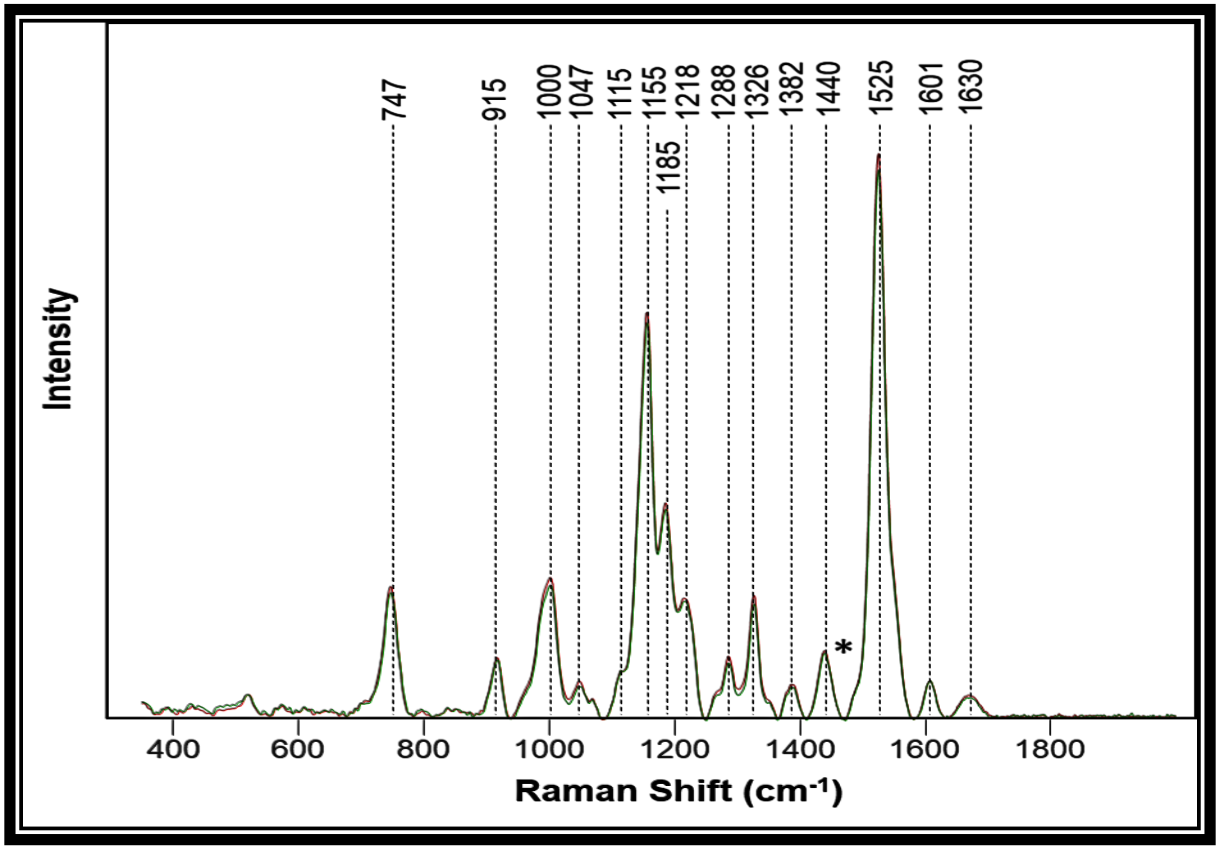

Figure 3.1: Raman spectra collected from leaves of male (green) and female (red) young hemp plants. Figure provided with permission by Dr. Kurouski.

To inspect the differences between male and female hemp plant spectra, a multivariate statistical analysis method using PLS-DA Toolbox with preprocessing described in the methods section was utilized to obtain statistical differences. We found that with Raman spectroscopy, young hemp plants can be distinguished between males and females with high accuracy of

around 90% as seen in Table 3.2, and adult hemp plant sex can be differentiated with about 94% accuracy as seen in Table 3.3.

Table 3.2: Vibrational band assignments of the components found within young hemp plant leaves.

	Number of spectra	TPR	Predicted as male	Predicted as female
Male	33	90.6%	31	2
Female	67	89.7%	1	66

Table 3.3: Vibrational band assignments of the components found within adult hemp plant leaves.

	Number of spectra	TPR	Predicted as male	Predicted as female
Male	37	94.3%	33	4
Female	68	94.3%	2	66

These results are promising and very positive as this technology will save hemp producers overly laborious processes involved in gender determination. Now fields can contain 100% female hemp plants without the risk of any cross-pollination to maximize the production of the essential cannabinoids used in the hemp industry.

3.2 HPLC

To understand exactly which carotenoids differ in concentration in male and female hemp plants, chromatographic analysis was used. When HPLC analysis was ready to be carried out on the hemp specimens, the leaves were removed from the freezer and thawed. Once thawed, the hemp leaves were homogenized separately following the same procedures. A total of six samples

were homogenized, three male and three female. Following homogenization, carotenoids were extracted in the manner described in the methods section of this manuscript. To ensure no external variability, the samples' carotenoids were extracted all in the same manner. Once all extractions were complete, HPLC was used to separate the carotenoids and provide means of comparison between male and female hemp plants.

The software, Breeze, was used for the quantification of the HPLC results. Within Breeze software, the peak areas and time of elution can be acquired. Standard carotenoids were injected before any hemp carotenoids to use as standards for comparison. The following standards were used: beta-carotene, alpha-carotene, chlorophyll, neoxanthin, viol, zeaxanthin, and lutein. Methanol was used to wash the column between the standards and the samples and between the male and female samples.

Our analysis shows that the carotenoids found within the hemp leaves are Lutein (LUT), Chlorophyll (CHR), Lycopene (LYC), Beta-cryptoxanthin (BCR), Zeaxanthin (ZEA), and Beta-carotene (BCA), Figure 3.2 and Figure 3.3. To find the differences, if any, of the carotenoids between male and female hemp plants, the Breeze software was used to measure the intensity and time of elution. The elution times of the peaks present on the chromatogram of the hemp samples were compared with the elution times of the standards to determine the identity of the carotenoids present in the hemp samples.

To determine the differences between the carotenoids within the male and female hemp species, the average peak intensities of male and female hemp plants were compared with one another to detect any significant differences. The results conclude that there are some differences in carotenoids in male and female hemp plants. We found from a study preceding this one that chlorophyll is highly fluorescent and thus cannot be attributed to any differences found in the

Raman spectra between male and female hemp plants. There were no differences in intensities of Lycopene (LYC), Beta-cryptoxanthin (BCR), Zeaxanthin (ZEA), or Beta-carotene (BCA) that could be attributed to a significant difference seen in the Raman spectra. However, as seen in Figure 3.4, there is a significant difference between the intensities of male and female lutein. This suggests that there is a difference between the concentrations of lutein within male and female hemp plants.

In the figure below, the following carotenoids are lutein eluting at 12.030, chlorophyll eluting at 13.000, lycopene eluting at 14.121, beta-cryptoxanthin eluting at 14.881, zeaxanthin eluting at 15.111, and lastly beta-carotene eluting at 17.543.

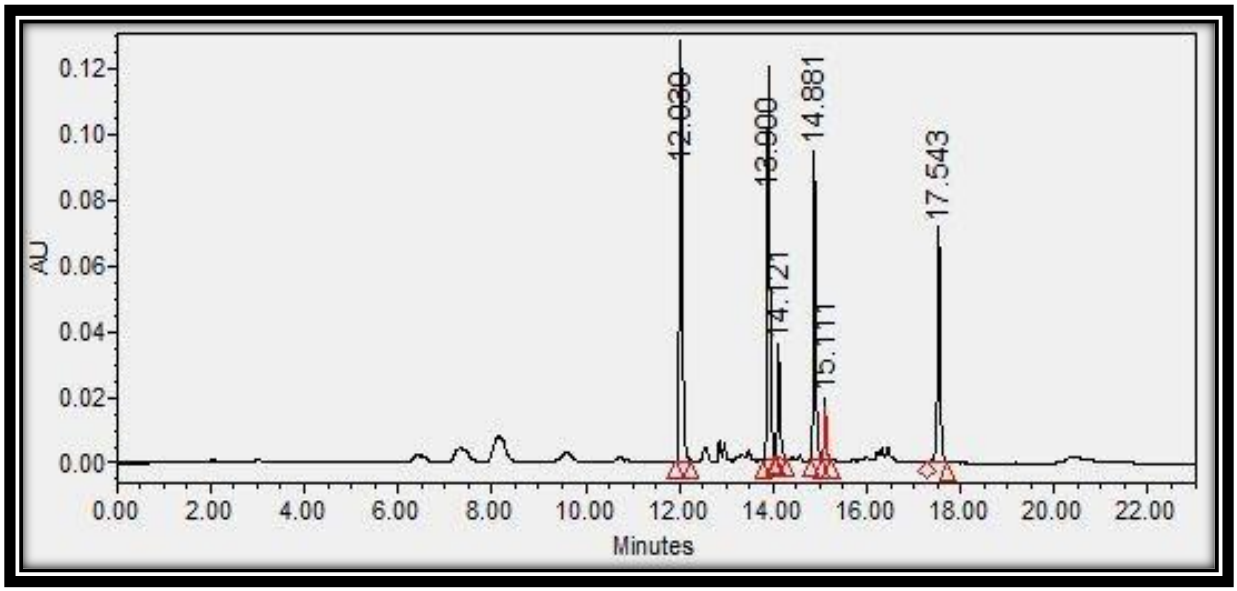

Figure 3.2: HPLC profiles of leaves collected from a mature female Hemp plant.

The following carotenoids in Figure 3.3 are lutein eluting at 11.936, chlorophyll eluting at 12.882, lycopene eluting at 14.106, beta-cryptoxanthin eluting at 14.867, zeaxanthin eluting at 15.098, and lastly beta-carotene eluting at 17.539. The other chromatograms for the two other female and two other male hemp samples are located in Appendix: A/HPLC Breeze Data.

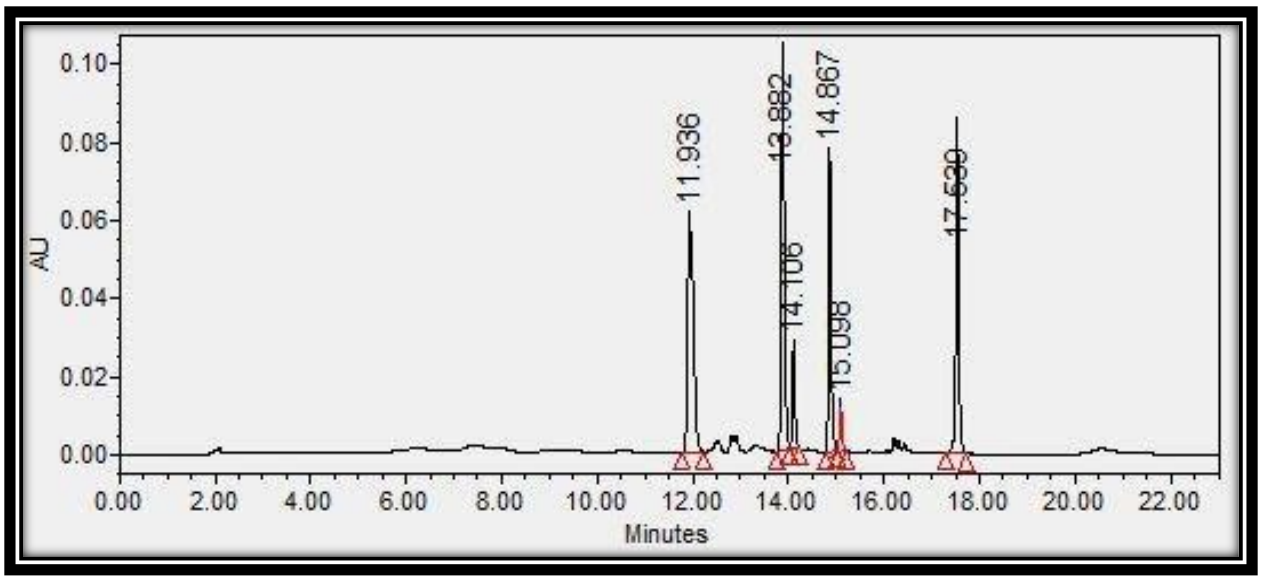

Figure 3.3: HPLC profiles of leaves collected from a mature male Hemp plant.

In Figure 3.4, the significant difference between the male (orange) and female (blue) carotenoids is seen in lutein (LUT). It is also worthy to note that all the carotenoids (except beta-cryptoxanthin) have a higher intensity in females than in males which is an interesting finding since the female peaks are higher than the male peaks in the Raman spectra. Though these differences are noteworthy they are not significant enough to contribute to any of the significant changes within the differences between male and female hemp plants using Raman spectroscopy.

Carotenoids have various functions within higher plants which are important for the overall health and growth of the plant. Within a plant, carotenoids are used in photosynthetic processes such as those involving light absorption for pigmentation and protective functions to ensure photosynthesis can occur without any harm to the plant[35]. The carotenoids found in the hemp leaves using HPLC analysis have specific roles in the photosynthetic mechanisms within the hemp plant. The carotenoid, zeaxanthin, protects the hemp plant during conditions in which the plant is under stress. Chlorophyll captures the solar energy that is used in photosynthesis and is also responsible for carrying electrons and separating charges[36]. Lycopene is a carotenoid with

antioxidant properties and is also believed to protect the plant from oxygen radicals or other harmful molecular substances in plants[37]. Studies show that beta-carotene has been successfully used to treat cancer and is a predecessor of vitamin A which has its benefits[38]. Within a plant, beta-carotene is largely present and partly responsible for pigmentation along with other carotenoids[39]. Beta-cryptoxanthin also plays a role in blocking lipid oxidization and is one of the radical scavengers within a plant[40]. The last carotenoid within the HPLC analysis is lutein which may be associated with the differences seen in male and female hemp spectra.

In a previous study, lutein was found to be the most prominent carotenoid in the hemp plant[41]. In this experiment, it can also be seen in Figure 3.4 that lutein, along with chlorophyll, is the most abundant carotenoid in the hemp sample extractions. The largely prominent xanthophyll in plants is lutein which is responsible for specific roles within photosynthesis[42]. Other experiments also found that in many hemp varieties, lutein was the most abundant carotenoid[43]. Lutein is derived from the carotenoid alpha-carotene. It has antioxidant properties and is also known to aid in vision health[44]. As lutein is an important carotenoid for plant health and biochemical pathway function, its prominence within the hemp plant is beneficial for using Raman spectroscopy in sex differentiation of male and female hemp plants as lutein is most likely responsible for the differences shown in male versus female spectra.

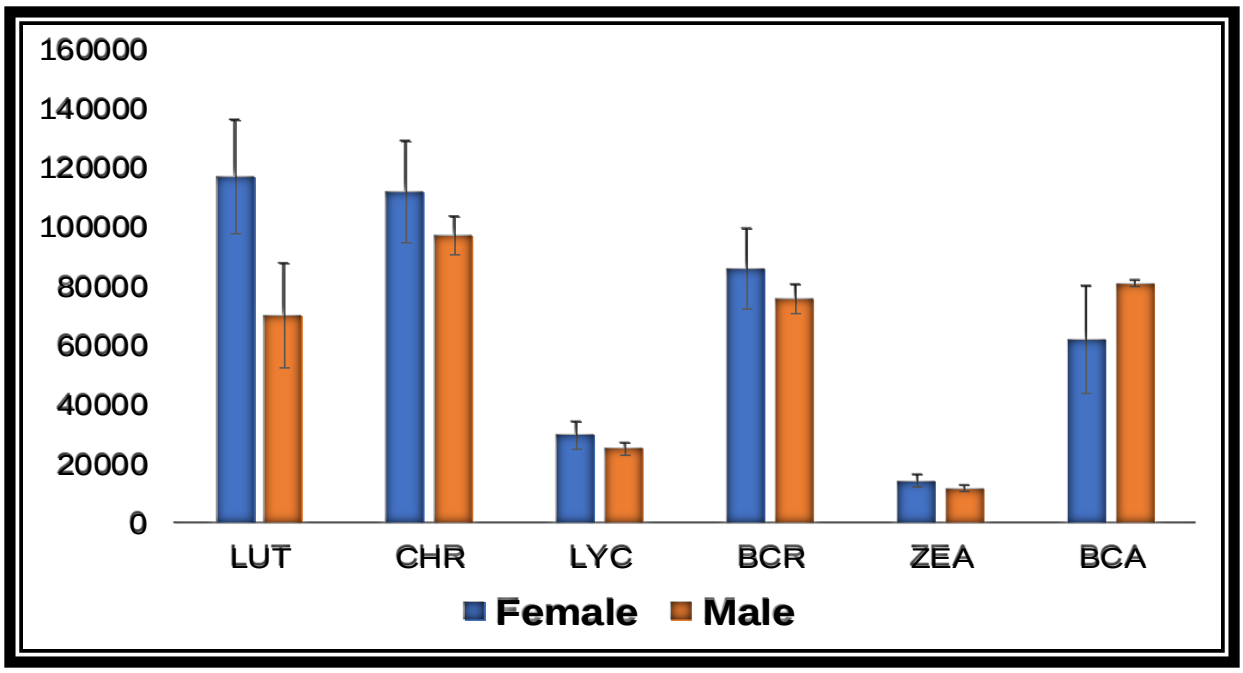

Figure 3.4: Histogram of the average areas of the peaks of Lutein (LUT), Chlorophyll (CHR), Lycopene (LYC), Beta-cryptoxanthin (BCR), Zeaxanthin (ZEA), Beta-carotene (BCA). Figure provided with permission by Dr. Rizevsky.

4. CONCLUSION

As the Hemp industry continues its expansion and is projected to bring in a copious amount of revenue in the near future, it becomes exceedingly important to have a method of differentiation between male and female hemp plants that is time-efficient and effective. As it is now known, there are many medicinal properties of the hemp plant that are highly beneficial. Countless other uses of the hemp plant such as the fiber industry, cosmetics, clothing, biofuel and so much more are also known.

Due to the biochemical variability between male and female hemp plants, the issue of sex differentiation continues to become a problem. The plants may have similar uses however they are not identical and thus their sexual identity must be predetermined before the crop is cross-pollinated. Hemp farmers seek to grow a field of strictly female hemp plants due to their higher concentrations of cannabinoids. Cross-pollination will cause the female crop to no longer be as useful within the aspect of harvesting their cannabinoids for manufacturing products or use for consumption. It becomes a burden to the farmers to have to send samples off to a laboratory for genetic testing, or visually inspect the plants without there being the certainty of classification.

To devise a method for a less laborious, time-efficient, and highly accurate sex determination process we first asked the question of whether or not Raman spectroscopy can differentiate between male and female hemp plants. To do this, hemp plants were grown in a greenhouse in College Station, Tx under normal conditions for healthy growth. Once the plants reached two weeks of age, spectra were collected from each of the plants. At this point, there were no visible phenotypic characteristics to allude to the sex of the plants. After two more weeks, the plants reached one month of age. At this time, they were flowered and could be

differentiated visually. The same number of spectra were again collected from the same group of plants.

Once all spectra were collected from both young and mature hemp plants. The spectral data were imported into MATLAB software where a PLS-DA Toolbox was used with set preprocessing methods to determine statistically significant differences in male and female hemp plant spectra. As seen in Table 3.2 and Table 3.3, the young hemp plants can be differentiated with high accuracy of around 90%, and mature hemp plant sex can be differentiated with around 94%. This remarkable discovery opens the door to a wider opportunity in the use of Raman spectroscopy technology. For sex differentiation, it would save time and cut costs. To reinforce our discovery, HPLC was used to confirm biochemical differences.

After spectra were collected from the mature hemp plants, leaves were collected and frozen for subsequent HPLC analysis. Before HPLC analysis, carotenoids were extracted from the hemp plants as described in the methods section of this manuscript. Six total samples were analyzed, three from the male hemp plants and three from the female hemp plants. Using HPLC and the Breeze software together allowed us to see which carotenoids were prominent within the hemp samples as well as which carotenoids could be responsible for the differences we see in the Raman spectra, Figure 3.1. We found that the carotenoid, lutein, is most prominent in the hemp plants and has the greatest difference between male and female hemp plants, Figure 3.4. Thus, we concluded that lutein was the probable cause for differences seen in the Raman spectra.

This technology can revolutionize the way that sex determination occurs within the hemp industry as well as in other horticulture industries. The predetermination of hemp plant sex allows for the hemp field to contain only female plants as desired by hemp growers. Many opportunities can arise from this discovery in way of the multitude of uses of Raman

spectroscopy. Further research should be conducted on a larger scale to ensure the benefits of Raman spectroscopy in hemp sex differentiation. Though from this experiment we know now that Raman spectroscopy can be used to differentiate between male and female hemp plants with high accuracy.

REFERENCES

1 Mark, T. B. & Snell, W. Economic issues and perspectives for industrial hemp. *Industrial hemp as a modern commodity crop*, 107-118 (2019).

2 Pellati, F. *et al.* Cannabis sativa L. and Nonpsychoactive Cannabinoids: Their Chemistry and Role against Oxidative Stress, Inflammation, and Cancer. *BioMed Research International* **2018**, 1691428, doi:10.1155/2018/1691428 (2018).

3 Borrelli, F. *et al.* Beneficial effect of the non-psychotropic plant cannabinoid cannabigerol on experimental inflammatory bowel disease. *Biochem Pharmacol* **85**, 1306-1316, doi:10.1016/j.bcp.2013.01.017 (2013).

4 Sanchez, L., Baltensperger, D. & Kurouski, D. Raman-Based Differentiation of Hemp, Cannabidiol-Rich Hemp, and Cannabis. *Analytical Chemistry* **92**, 7733-7737, doi:10.1021/acs.analchem.0c00828 (2020).

5 Russo, E. B., Guy, G. W. & Robson, P. J. Cannabis, pain, and sleep: lessons from therapeutic clinical trials of Sativex®, a cannabis-based medicine. *Chemistry & biodiversity* **4**, 1729-1743 (2007).

6 Izzo, A. A., Borrelli, F., Capasso, R., Di Marzo, V. & Mechoulam, R. Non-psychotropic plant cannabinoids: new therapeutic opportunities from an ancient herb. *Trends in Pharmacological Sciences* **30**, 515-527, doi:https://doi.org/10.1016/j.tips.2009.07.006 (2009).

7 Martínez, V. *et al.* Cannabidiol and Other Non-Psychoactive Cannabinoids for Prevention and Treatment of Gastrointestinal Disorders: Useful Nutraceuticals? *Int J Mol Sci* **21**, 3067, doi:10.3390/ijms21093067 (2020).

8 Toth, J. A. *et al.* Development and validation of genetic markers for sex and cannabinoid chemotype in Cannabis sativa L. *Gcb Bioenergy* **12**, 213-222 (2020).

9 Schluttenhofer, C. & Yuan, L. Challenges towards Revitalizing Hemp: A Multifaceted Crop. *Trends in Plant Science* **22**, 917-929, doi:https://doi.org/10.1016/j.tplants.2017.08.004 (2017).

10 Truta, E., Olteanu, Z., Surdu, S., Zamfirache, M.-M. & Oprica, L. Some aspects of sex determinism in hemp. *Analele Stiintifice ale Universitatii" Alexandru Ioan Cuza" din Iasi Sec. II a. Genetica si Biologie Moleculara* **8** (2007).

11 Flajšman, M., Slapnik, M. & Murovec, J. Production of Feminized Seeds of High CBD Cannabis sativa L. by Manipulation of Sex Expression and Its Application to Breeding. *Frontiers in plant science* **12**, 718092-718092, doi:10.3389/fpls.2021.718092 (2021).

12 Petit, J., Salentijn, E. M. J., Paulo, M.-J., Denneboom, C. & Trindade, L. M. Genetic Architecture of Flowering Time and Sex Determination in Hemp (Cannabis sativa L.): A Genome-Wide Association Study. *Frontiers in Plant Science* **11**, doi:10.3389/fpls.2020.569958 (2020).

13 Mandolino, G. & Carboni, A. Potential of marker-assisted selection in hemp genetic improvement. *Euphytica* **140**, 107-120, doi:10.1007/s10681-004-4759-6 (2004).

14 Cardona, M. & Merlin, R. in *Light Scattering in Solid IX* (eds Manuel Cardona & Roberto Merlin) 1-14 (Springer Berlin Heidelberg, 2007).

15 Smekal, A. Zur quantentheorie der dispersion. *Naturwissenschaften* **11**, 873-875 (1923).

16 Raman, C. V. & Krishnan, K. S. A New Type of Secondary Radiation. *Nature* **121**, 501-502, doi:10.1038/121501c0 (1928).

17 Jones, R. R., Hooper, D. C., Zhang, L., Wolverson, D. & Valev, V. K. Raman Techniques: Fundamentals and Frontiers. *Nanoscale Res Lett* **14**, 231-231, doi:10.1186/s11671-019-3039-2 (2019).

18 Perkowitz, S. in *Encyclopedia Britannica* (2018).

19 Das, R. S. & Agrawal, Y. K. Raman spectroscopy: Recent advancements, techniques and applications. *Vibrational Spectroscopy* **57**, 163-176, doi:https://doi.org/10.1016/j.vibspec.2011.08.003 (2011).

20 Petrova, O. E. & Sauer, K. High-Performance Liquid Chromatography (HPLC)-Based Detection and Quantitation of Cellular c-di-GMP. *Methods Mol Biol* **1657**, 33-43, doi:10.1007/978-1-4939-7240-1_4 (2017).

21 Dou, T. *et al.* Biochemical Origin of Raman-Based Diagnostics of Huanglongbing in Grapefruit Trees. *Frontiers in Plant Science* **12**, doi:10.3389/fpls.2021.680991 (2021).

22 Staff, H. I. D. in *Hemp Industry Daily* (2021).

23 Merlin, M. D. Archaeological Evidence for the Tradition of Psychoactive Plant Use in the Old World. *Economic Botany* **57**, 295-323 (2003).

24 Poniatowska, J., Wielgus, K., Szalata, M., Ozarowski, M. & Panasiewicz, K. Contribution of Polish agrotechnical studies on Cannabis sativa L. to the global industrial hemp cultivation and processing economy. *Herba Polonica* **65** (2019).

25 Aguilar, S., Gutiérrez, V., Sánchez, L. & Nougier, M. Medicinal cannabis policies and practices around the world. *Int Drug Policy Consort. Briefing paper*, 1-32 (2018).

26 Synytsya, A., Čopíková, J., Matějka, P. & Machovič, V. Fourier transform Raman and infrared spectroscopy of pectins. *Carbohydrate Polymers* **54**, 97-106 (2003).

27 Edwards, H., Farwell, D. & Webster, D. FT Raman microscopy of untreated natural plant fibres. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* **53**, 2383-2392 (1997).

28 Schulz, H., Baranska, M. & Baranski, R. Potential of NIR-FT-Raman spectroscopy in natural carotenoid analysis. *Biopolymers: Original Research on Biomolecules* **77**, 212-221 (2005).

29 Dou, T. *et al.* Biochemical Origin of Raman-Based Diagnostics of Huanglongbing in Grapefruit Trees. *Frontiers in Plant Science* **12** (2021).

30 Yu, M. M. S., H. G.; Jetter, R.; Blades, M. W.; Turner, R. F. Raman microspectroscopic analysis of triterpenoids found in plant cuticles. *Applied Spectroscopy* **61**, 32-37 (2007).

31 Devitt, G., Howard, K., Mudher, A. & Mahajan, S. Raman spectroscopy: an emerging tool in neurodegenerative disease research and diagnosis. *ACS chemical neuroscience* **9**, 404-420 (2018).

32 Adar, F. Carotenoids-their resonance raman spectra and how they can be helpful in characterizing a number of biological systems. *Spectroscopy* **32**, 12-20 (2017).

33 Kang, L., Wang, K., Li, X. & Zou, B. High pressure structural investigation of benzoic acid: raman spectroscopy and x-ray diffraction. *The Journal of Physical Chemistry C* **120**, 14758-14766 (2016).

34 Agarwal, U. P. Raman imaging to investigate ultrastructure and composition of plant cell walls: distribution of lignin and cellulose in black spruce wood (Picea mariana). *Planta* **224**, 1141-1153 (2006).

35 Young, A. J. The photoprotective role of carotenoids in higher plants. *Physiologia Plantarum* **83**, 702-708, doi:https://doi.org/10.1111/j.1399-3054.1991.tb02490.x (1991).

36 Tanaka, A. & Tanaka, R. Chlorophyll metabolism. *Current Opinion in Plant Biology* **9**, 248-255, doi:https://doi.org/10.1016/j.pbi.2006.03.011 (2006).

37 Collins, J., Perkins-Veazie, P. & Roberts, W. Lycopene: from plants to humans. *HortScience* **41**, 1135-1144 (2006).

38 do Monte, D. S. *et al.* Chemical and biological studies of β-carotene after exposure to Cannabis sativa smoke. *Toxicol Rep* **3**, 516-522, doi:10.1016/j.toxrep.2016.06.001 (2016).

39 Wang, L., Liu, Z., Jiang, H. & Mao, X. Biotechnology advances in β-carotene production by microorganisms. *Trends in Food Science & Technology* **111**, 322-332, doi:https://doi.org/10.1016/j.tifs.2021.02.077 (2021).

40 Ghosh, A., Hazra, U. & Dutta, D. Role of β-Cryptoxanthin as an Antioxidant and Its Ability to Bind with Transferrin. *International Journal of Bioscience, Biochemistry and Bioinformatics* **9** (2019).

41 Irakli, M. *et al.* Effect of Genotype and Growing Year on the Nutritional, Phytochemical, and Antioxidant Properties of Industrial Hemp (Cannabis sativa L.) Seeds. *Antioxidants* **8**, 491 (2019).

42 Dall'Osto, L. *et al.* Lutein is needed for efficient chlorophyll triplet quenching in the major LHCII antenna complex of higher plants and effective photoprotection in vivo under strong light. *BMC Plant Biol* **6**, 32-32, doi:10.1186/1471-2229-6-32 (2006).

43 Cattaneo, C. *et al.* Biochemical aspects of seeds from Cannabis sativa L. plants grown in a mountain environment. *Scientific Reports* **11**, 3927, doi:10.1038/s41598-021-83290-1 (2021).

44 Cheng, X., Zhao, X., Huang, C., Zhang, X. & Lyu, Y. Lutein content in petals and leaves of marigold and analysis of lutein synthesis gene expression. *Acta Physiologiae Plantarum* **41**, 128, doi:10.1007/s11738-019-2913-y (2019).

APPENDIX: A/HPLC BREEZE DATA

HPLC Standards:

blk data:

	Name	Retention Time	Area	% Area	Height	Peak Codes	Int Type	Amount	Units	Peak Type
1		2.305	2231	2.13	1048		tt			Unknown
2		2.838	28201	26.90	690		Bv			Unknown
3		17.744	2122	2.02	177		BV			Unknown
4		20.646	72280	68.95	1219		vB			Unknown

alpha:

	Name	Retention Time	Area	% Area	Height	Peak Codes	Int Type	Amount	Units	Peak Type
1		17.475	1815266	100.00	480044		BB			Unknown

beta:

	Name	Retention Time	Area	% Area	Height	Peak Codes	Int Type	Amount	Units	Peak Type
1		17.528	3132519	100.00	810221		VB			Unknown

lutein:

	Name	Retention Time	Area	% Area	Height	Peak Codes	Int Type	Amount	Units	Peak Type
1		12.070	2687919	100.00	553167		VV			Unknown

neo:

	Name	Retention Time	Area	% Area	Height	Peak Codes	Int Type	Amount	Units	Peak Type
1		7.988	3310784	100.00	159906		Vv			Unknown

chlor:

	Name	Retention Time	Area	% Area	Height	Peak Codes	Int Type	Amount	Units	Peak Type
1		12.904	26054	2.39	6545		bb			Unknown
2		13.083	52108	4.78	11003		bb			Unknown
3		13.273	27576	2.53	8061		bb			Unknown
4		13.397	44015	4.03	12888		bb			Unknown
5		13.896	895947	82.11	218584		bv			Unknown
6		14.115	45430	4.16	11835		vb			Unknown

viol:

	Name	Retention Time	Area	% Area	Height	Peak Codes	Int Type	Amount	Units	Peak Type
1		8.797	3055068	100.00	159367		VV			Unknown

HPLC Hemp Samples:

Female U1:

	Name	Retention Time	Area	% Area	Height	Peak Codes	Int Type	Amount	Units	Peak Type
1	lutein	12.092	338219	30.46	75528		BB			Unknown
2		12.917	11831	1.07	3996		bb			Unknown
3		13.025	12628	1.14	4407		bb			Unknown
4	chlor	13.933	303064	27.30	75974		VV			Unknown
5	lycopene	14.151	76118	6.86	20017		bb			Unknown
6	Beta-cryptoxanthin	14.908	202638	18.25	56774		bb			Unknown
7	zeaxanthin	15.136	32206	2.90	9569		bb			Unknown
8		16.333	16058	1.45	2830		bb			Unknown
9		16.470	14162	1.28	2450		bb			Unknown
10	Beta-carotene	17.554	103404	9.31	23130		BB			Unknown

Female U12:

	Name	Retention Time	Area	% Area	Height	Peak Codes	Int Type	Amount	Units	Peak Type
1	lutein	12.120	649619	28.91	151725		bb			Unknown
2		12.911	20505	0.91	6722		bb			Unknown
3		13.013	16498	0.73	5540		bb			Unknown
4	chlor	13.931	560575	24.95	143984		bb			Unknown
5	lycopene	14.148	146705	6.53	38160		bb			Unknown
6	Beta-cryptoxanthin	14.896	405503	18.05	109856		bb			Unknown
7	zeaxanthin	15.123	63523	2.83	17445		bb			Unknown
8	Beta- carotene	17.546	384229	17.10	94332		VB			Unknown

Female U28:

	Name	Retention Time	Area	% Area	Height	Peak Codes	Int Type	Amount	Units	Peak Type
1	lutein	12.030	564150	30.97	124261		bb			Unknown
2	chlor	13.900	457626	25.12	116022		bb			Unknown
3	lycopene	14.121	121677	6.68	31169		bb			Unknown
4	Beta-cryptoxanthin	14.881	335985	18.44	91423		bb			Unknown
5	zeaxanthin	15.111	58381	3.20	16137		bb			Unknown
6	Beta- carotene	17.543	284001	15.59	68511		VB			Unknown

Male U2:

	Name	Retention Time	Area	% Area	Height	Peak Codes	Int Type	Amount	Units	Peak Type
1	lutein	12.118	491438	28.61	109991		bb			Unknown
2	chlor	13.931	420774	24.50	107682		bb			Unknown
3	lycopene	14.148	113033	6.58	29242		bb			Unknown
4	Beta-cryptoxanthin	14.892	315334	18.36	85785		bb			Unknown
5	zeaxanthin	15.119	49320	2.87	14238		bb			Unknown
6	Beta- carotene	17.542	327830	19.09	81788		bB			Unknown

Male U7:

	Name	Retention Time	Area	% Area	Height	Peak Codes	Int Type	Amount	Units	Peak Type
1	lutein	11.819	466335	29.98	41662		bb			Unknown
2	chlor	13.852	384063	24.69	83202		bb			Unknown
3	lycopene	14.079	92368	5.94	21151		bb			Unknown
4	Beta-cryptoxanthin	14.851	259032	16.65	66442		bb			Unknown
5	zeaxanthin	15.084	38106	2.45	10080		bb			Unknown
6	Beta- carotene	17.534	315800	20.30	78704		bB			Unknown

Male U8:

	Name	Retention Time	Area	% Area	Height	Peak Codes	Int Type	Amount	Units	Peak Type
1	lutein	11.936	512836	29.95	59141		bb			Unknown
2	chlor	13.882	437749	25.56	100940		bb			Unknown
3	lycopene	14.106	102594	5.99	25067		bb			Unknown
4	Beta-cryptoxanthin	14.867	285413	16.67	75339		bb			Unknown
5	zeaxanthin	15.098	40032	2.34	10984		Vb			Unknown
6	Beta- carotene	17.539	333771	19.49	82903		BB			Unknown