Title

Variation among hemp (Cannabis sativus L.) analytical testing laboratories evinces regulatory and quality control issues for the industry

Authors

Azwell, Thomas
Ciotti, Chloe
Adams, Arthur
et al.

Publication Date

2022-12-01

DOI

10.1016/j.jarmap.2022.100434

Peer reviewed

Variation among hemp (*Cannabis sativus* L.) analytical testing laboratories evinces regulatory and quality control issues for the industry

Thomas Azwell[1], Chloe Ciotti[1], Arthur Adams[1], & Guido F. Pauli[2]

[1]Azwell Lab, College of Engineering, University of California, Berkeley, United States
[2]Pharmacognosy Institute and Department of Pharmaceutical Sciences, College of Pharmacy, University of Illinois at Chicago, United States

Keywords: cannabidiol, tetrahydrocannabinol, hemp regulation, USDA, THC

Abstract

Validated analytical testing of cannabinoid content for regulatory purposes is critical to farming high-cannabidiol (CBD) hemp (*Cannabis sativus* L.), as these methods are key to determine whether a crop is federally compliant by containing <0.3% THC or must be destroyed at the time of harvest. This report identifies the sources and extent of variation in reported cannabinoid content after flower selection and conducted a survey of ten accredited hemp regulatory testing laboratories, including one state-owned laboratory. The results indicate that total tetrahydrocannabinol (THC) content is reported inconsistently due to an insufficient standardization of sample preparation and testing methods, thus likely leading to erroneous data. This work represents the early phase of ongoing research to enhance the consistency of cannabinoid analyses of hemp flower samples as an essential tool in Cannabis crop development.

1. Introduction

Although hemp and marijuana plants belong to the same group of three species, *Cannabis sativa, C. indica,* and *C. ruderalis,* recent agricultural legislation in the U.S. has reclassified hemp, removing it from the Schedule 1 controlled substances list and distinguishing it from marijuana by

using the total THC level as a threshold. These levels are limited to "a Δ9-tetrahydrocannabinol concentration of not more than 0.3 percent on a dry weight basis" in any part of the hemp plant and any products derived from its cultivation (USDA Agricultural Marketing Service, 2019). This definition of hemp, tied to plant chemotype rather than plant genotype, stems from a 1976 taxonomical study distinguishing industrial hemp (used commonly at the time for fiber and food) from psychoactive, high-THC cultivars of (mostly) *C. sativa* marijuana (Small-Cronquist, 1976). However, hemp is known to be genetically distinguishable from marijuana because it lacks the allele coding for THCA-synthase (De Meijer et al., 2003; Sawler et al., 2015).

While industrial fibrous hemp remains an important crop, an increasing number of farmers are cultivating medicinal cultivars of hemp to meet consumer demand for CBD and other cannabinoid products such as cannabigerol (CBG) and cannabichromene (CBC). These products have been shown to have pharmacological activity providing purported therapeutic potential for antiepileptic, anxiolytic, antipsychotic, anti-inflammatory, and neuroprotective effects (Bridgeman-Abazia, 2017). Hemp flowers grown for medicinal use today contain up to 15% of CBD plus additional non-THC/non-CBD cannabinoids by dry weight. Depending on the growing conditions that favor cannabinoid production, these variants often approach or slightly exceed the 0.3% THC limit for classification as hemp at the time of harvest—despite having no psychotropic effects compared with marijuana, which is cultivated for the production of THC at dry-weight percentages commonly in excess of 20%. In general, a level of about 1% THC is considered the threshold for cannabis to have a psychotropic effect (Small-Marcus, 2002).

The industry faces several challenges as it attempts to scale cultivation to meet consumer demand for cannabinoid products. For example, categorizing cultivars becomes difficult due to variation in testing protocols and discrepancies in results between testing laboratories. However,

these results are critical when quantifying the amount of THC in hemp flower and hemp-derived extracts, and certificates of analysis (COA) are compulsory for legal hemp harvest. In most states, hemp farmers are required to submit a request to the state 30 days before the anticipated harvesting time, so an official sampler can collect flowers from multiple plants of the same strain on the farm for compliance testing (Cal. Code Reg., 2021).

The selection methods to control for the uniformity of flower sampling from each registered strain of hemp prepared for harvesting can vary slightly based on a local regulator's protocol. The samples are then sent to an analytical lab, either operated by the state or individually contracted, to determine the concentration of cannabinoids in each varietal (strain) of hemp sampled. Industry or other standard operating procedures for these labs are missing. In addition, none of the commercial labs we evaluated provided experimental details about the applied methods or validation data. This may be related to the lack of regulations and reflect a means of maintaining a competitive advantage in a competitive environment. The lack of validated convention methods means that generally accepted standards for instruments, calibration, or sampling replicates are missing. However, in any state, if the test results reveal a total THC concentration below 0.3%, the farmer is approved to harvest on their listed date. However, if the THC test result is above 0.3% it is considered "hot", and the farmer may have to destroy the entire crop.

The federal mandate distinguishes between hemp and marijuana based on plant chemotype, using a single marker compound (THC), but without establishing a convention method that is "fit for purpose" for this definition. This is despite the existence of fit-for-purpose chemotype identification convention methods in the literature (Sarma et al., 2020). The lack of a validated convention method representing the official standard is particularly problematic because measurement of the THC concentration involves quantitation at a relatively low level, which poses

analytical challenges, especially when performed outside highly controlled laboratory settings. With limited federal regulation over hemp analytical laboratory accreditation policies or standardized analysis protocols, the quantitation of cannabinoid composition reported by commercial hemp testing laboratories has the potential to vary greatly, including for identical samples. The sources of potential variation extend beyond analytic techniques and include sample storage upon receipt, homogenization and extraction methods, instrumentation, and calibration procedures. Collectively, these factors all influence the test results and, thereby, the ultimate decision about crop validation versus destruction. The development of different analytical methods to quantify cannabinoids in plant products allows individual laboratories to compete in a growing market for hemp analysis. However, this practice also interferes with transparency, exacerbates inter-laboratory variation in analysis results, diminishes the validity of crop regulation, and calls the overall value of COAs into question.

The attention focus on a single marker compound, THC, also raises the question about options and rationales for a more holistic analysis of the cannabinoid metabolome, which in fact is practically feasible. Potential approaches including methods for the quantitation of 13 individual cannabinoids have been compiled by the USP Cannabis Expert Panel (Sarma et al., 2020). However, with the legal definition of the Controlled Substance Act (CSA) in place, this does still not address the question of standardization of the analytical methodology used to determine cannabinoid content, including threshold CBD values. In this context, it should be pointed out that pharmacopeial approaches worldwide involve the establishment of convention methods, which are developed via a consensus building process and considered fit-for-purpose. While this approach provides a more definitive analytical framework, it does not necessarily address the levels of uncertainty involved in the analyses and are typically not embedded into metrological frameworks

by inclusion of primary reference standards (such as NIST benzoic acid PS1) and traceable certified reference materials (Nelson et al., 2018).

2. Experimental

2.1 Laboratory survey

A total of 45 laboratories were identified via a Google search of the term "analytical cannabis testing labs" and contacted about their participation in a research survey of their analytical methods, prioritizing labs that provided services to regulatory agencies above those used only by the industry. Of the 45 labs contacted, 11 agreed to participate. One of the participating laboratories was eliminated because it did not offer hemp testing. The 10 participating laboratories were asked to extract and analyze three hemp flower samples and to return their cannabinoid potency test results. Surveyed laboratories were also asked to provide information about their instrumentation, standards, extraction protocols, and methods of analysis. While all of the participating laboratories chose to withhold their extraction and analysis protocols, they disclosed the manufacturer and make of their analytical instrumentation and source of reference standards employed for calibration.

2.2 Preparation

Two strains of hemp, Sample A, a high-CBG chemotype, and Sample B, a high-CBD chemotype, were cultivated for this survey. Samples from each strain were harvested on the same day. The samples were prepared from buds removed from the stalks and trimmed of surrounding leaves. The samples from each strain were collected into a bag and shaken to randomly distribute flowers collected from individual plants for distribution to the testing facilities. Homogenized samples were ground using a Spex 2010 Geno/Grinder 115V. Flowers from each sample were

added to 50 mL centrifuge tubes along with three steel balls. Each sample was run at 1,000 strokes per minute for one minute, then manually stirred and visually inspected to ensure all flowers were ground. The grinding process was repeated at 1,000 strokes per minute for one minute. All ground flower from each sample was combined in a designated container and mixed again to ensure a completely homogenized distribution, minimizing variability due to flower selection from one individual plant or a portion thereof to another given the same strain.

2.3 Extraction control experiment

To determine variation in cannabinoid results due to extraction solvents, an internal analysis was performed using various extraction solvents. Homogenized flowers from Sample B were extracted nine times using each of the solvents reportedly used by the labs in the survey. About 200 mg of homogenized flowers was added to each of nine 50 mL polypropylene centrifuge tubes. The exact mass of Sample B was recorded for each tube, and each sample was extracted with 20 mL ethanol (3), 20 mL isopropanol (3), or 20 mL 1:1 isopropanol and acetonitrile (3). Each tube was mixed with a FisherBrand VWR Multi-tube vortexer at speed 9 for 5 minutes. Each of the tubes was centrifuged in a Q-Sep 3000 Centrifuge 110V for 5 minutes to separate residual solids from the supernatant. 1.5 mL of the suspended supernatant was removed and syringe-filtered through a Millex GP 0.22 µm filter to remove solid particulates. 50 µl of the supernatant was added to a 1.5 mL amber HPLC vial and diluted with 950 µl of the respective extraction solvent. The samples were analyzed using a Shimadzu LC2030-3D plus with a PDA detector equipped with a NexLeaf CBX C18 column and mobile phase 0.085% phosphoric acid in acetonitrile. The sample was compared to calibration standards “Certified Standard-11 Components (CRM) in Acetonitrile” sourced from Cayman Chemicals. The 250 µg/mL standard mixture was diluted to

0.5 ppm, 1 ppm, 5 ppm, 10 ppm, 50 ppm, and 500 ppm (R2 >0.999 for each of the 11 cannabinoids).

2.4 Sample B extract preparation

200.4 mg of homogenized Sample B was added to a 50 mL polypropylene centrifuge tube and extracted with 20 mL ethanol. The mixture was mixed with a FisherBrand VWR Multi-tube vortex at speed 9 for 5 minutes. The sample was centrifuged in a Q-Sep 3000 Centrifuge 110V for 5 minutes. The suspended supernatant was drawn into a syringe and filtered using a 0.22 µm Millex GP filter. We sent each surveyed lab 1 mL of the filtered extract in a vial sealed using parafilm and shipped in insulated, light-proof packaging. Participating laboratories were instructed to dilute the sample further using ethanol as necessary prior to the cannabinoid potency determination.

3. Results and discussion

For the survey used in this report, two strains of hemp, Sample A, a high-CBG chemotype, and Sample B, a high-CBD chemotype, were collected as whole flowers and distributed to laboratories that agreed to participate in a blind study of their analytical results and a survey of their testing protocols. The samples were also analyzed in our laboratory to further study the effects of homogenization and extraction protocols, and compare the results from the labs against a carefully controlled and transparent sample preparation, calibration, and analytic procedure. The reported total THC content of Sample A varied among ten surveyed laboratories (Figure 1) with a mean of 0.11% (dry weight basis), a standard deviation of 0.10%, and three labs returning a non-detect (ND) result, indicating that total THC content was below the limits of their detection methods (for the purposes of statistical analysis, ND results were treated as a result of 0%). Total

THC/CBD is calculated using the following formulas to take into account the loss of a carboxyl group during decarboxylation step:

Total THC = Δ9THC +(THCa (0.877)) and Total CBD = CBD + (CBDa (0.877))

Of the ten surveyed laboratories, one reported a result of 0.29% of THC, a value just below the legal threshold for crop destruction. The range of reported total THC (including THCA) content in this sample therefore spans 96.7% of the legally permissible total THC detection range for hemp, with a relative standard deviation of 95% of the mean. This suggests that the applied analytical methods for total THC quantitation across surveyed laboratories are not sufficiently precise at the levels required for the analysis of THC content in hemp. Our own analysis yielded a total THC content of 0.09%, close to the mean reported by the sample of surveyed laboratories.

From harvest of flower samples to sample preparation, homogenization, extraction, analysis, and quantitation, multiple factors can contribute to the variation in the reported total THC content, independently of the analytical method used. We probed these potential sources of variation by sending participating laboratories a homogenized (ground) Sample A, which was identified to participating laboratories as a third sample, Sample C. Although the plant material was identical to Sample A, the variation in reported results significantly decreased (standard deviation $\sigma = 0.10\%$ in whole flower and $\sigma = 0.05\%$ in homogenized flower), indicating that, in addition to potential flower-to-flower variation, different homogenization methods between laboratories also contributed to significant variation in the quantified total THC content.

Sample B was shipped to participating laboratories homogenized using a GenoGrinder. In this sample, the total THC content measured in our laboratory (0.51%) was higher than that of Sample A. Again, only the mean of the total THC content reported by the ten participating laboratories yielded a result similar to control (0.48%), whereas the reported values ranged from 0.27% (compliant) to 0.59% (non-compliant), with a relative standard deviation of 21% of the mean. Although the relative standard deviation was lower in this sample relative to Sample A, the standard deviation remained as high as 0.10% dry weight in terms of total THC content).

While all external laboratories reported using HPLC with UV-vis photodiode array detectors, the solvent used to extract flower samples varied: nine laboratories used methanol, one laboratory a 1:1 mixture of isopropanol and acetonitrile. In our internal reference analysis, ethanol was employed as the extraction solvent. In an extraction control experiment, nine samples of homogenized sample B were extracted with either ethanol (3), methanol (3), or isopropanol/acetonitrile (3). It was found that ethanol yielded the most exhaustive extraction, regardless of whether only total THC (Table 1) or total cannabinoid content were compared (15.7% ± 0.5, 14.3% ± 0.2, 15.1% ± 0.8 for ethanol, methanol, and isopropanol/acetonitrile, respectively).

Given the success in reducing analytical variation by pre-homogenizing flower samples, we anticipated that extracting the homogenized flower and shipping the extract to participating laboratories would significantly reduce variation. However, the variation in reported total THC content among participating laboratories increased when pre-extracted samples versus whole flower samples were shipped to and analyzed by participating laboratories (Sample B homogenized flower $\sigma = 0.10\%$ and extracted $\sigma = 0.12\%$).

A pre-extracted sample should show the highest degree of homogeneity (lowest variation) between aliquots sampled by participating laboratories, as it controls for all aspects of sample

harvesting and preparation up to analytical method and quantitation. One reason for the observed variations of results may be that participating laboratories employ sample preparation and analysis protocols for whole flower samples rather than liquid samples sent from customers. Here, variation may be due to lack of consistency of liquid sample handling, dilution procedures, and time to complete the analysis.

The outcomes of this survey highlight the potential issues resulting from variation among laboratories testing THC content. Farmers cultivating hemp strains for medicinal applications, however, will be interested not only in the THC content reported in COAs, but in the content of many other cannabinoids quantified in hemp samples. The variation in total cannabinoid content is shown in Figure 2, where it can be seen that for the whole flower Sample A, reported cannabinoid content spans a range from 9.8% to 19.4%—a huge variation. This once more highlights the importance of considering a wider spectrum of cannabinoids when evaluating cannabis materials (Sarma et al., 2020).

In addition to unavoidable inter-laboratory variation, the observed inconsistencies of the analytical outcomes can be inherent to the method by which cannabinoids are measured: HPLC with photodiode array detection is a comparative technique that relies on efficient separation, which in turn depends on many factors such as the specific column and chromatographic conditions (mobile phase solvents, solvent gradient, flow rate, temperature). To achieve quantitation, the HPLC chromatogram of a sample is compared to a chemically identical reference material (“standard”) that is analyzed under identical conditions to establish internal or external calibration. Provided that multiple calibrants are available, complex mixtures can be separated and individual components quantified. However, peaks overlap is one major source of error, especially when analyzing chemically complex mixtures such as Cannabis extracts. For example, similar or

fully overlapping retention times under the applied conditions—as is often the case, for intance with Δ8- vs. Δ9-THC or the *cis* and *trans* isomers — , affect the specificity of the quantitation. Moreover, the reference materials used for calibration can undergo chemical change (degradation), which may or may not be captured in both externally and internally calibrated methods. In fact, as terpenoids, cannabinoids are known to be relatively unstable compounds and commonly change after exposure to heat, light, or air (not counting the common decarboxylation reaction of the acid forms).

It should be noted that analytical methods are available that are independent of chemically identical reference materials (calibrants). For example, quantitative nuclear magnetic resonance (qNMR) has this capability as it represents a (relative) primary analytical technique. Its suitability for natural product and pharmaceutical analysis is widely acknowledged and the fitness of the application of quantum mechanics-based qNMR has recently been demonstrated for CBD (Pauli et al., 2014; Nelson et al., 2020).

The laboratories surveyed in this study used reference standards from Restek, Cerilliant, LGC, and Cayman. While most laboratories employed 11 standards, up to 18 standards were used to quantify various cannabinoids in hemp samples. Facilities that quantified a larger number of individual cannabinoids tended to report greater means in total cannabinoid content. However, no correlation was found between the number of calibration standards and the reported total percent content quantified across all facilities and all four samples. This kind of inconsistency highlights the analytical challenges hemp farmers face when evaluating the reliability of COA results and assessing the value of their crops.

4. Conclusion

This study represents a preliminary survey of the variation in reported total THC content among ten commercial hemp testing facilities. While early, the presented evidence strongly suggests that inconsistencies exist in reported total THC content, which leads to a regulatory liability. While the total THC content is currently the key legal metric defining whether a crop is harvestable or must be destroyed, it is not tied to, for example, a convention method that ensures validity and reproducibility of analytical outcomes, such as the methods found in pharmacopeias. The present data suggest that total THC content is frequently reported imprecisely, most likely due to analytical inconsistency of sample preparation and testing methods. Even under the carefully controlled conditions of the in-house laboratory, the relative standard deviation of total THC content as measured by HPLC using commercially available calibrants varied between 2% and 7% of the mean, depending on the extraction protocol employed. This variation is within the range of what can be expected for the quantitation of a minor (<1%) constituent embedded in a complex analytical matrix and indicates the intrinsic limitations in peak purity as the likely key factor in HPLC-UV-based quantitation.

These results suggest a significant likelihood of a hemp crop being marked for destruction due to inaccurate COA reporting – as well as a crop being labeled as compliant despite actually containing elevated THC levels. To support this burgeoning industry those exposed to the risk of growing this relatively new crop in the U.S., it would be a step forward to revise enforcement of COA results such that they account for the variation in reported total THC content, introduce methodology that establishes a reproducible linkage with metrological reference materials, and eliminate the inherent variability of results from different analytical methods. To do this, however, several factors would need to be carefully considered. The current regulatory specifications are

simple numbers that do not include measurement uncertainties. Determining the threshold values may have involved accounting for measurement uncertainty, but the values do not express that. In pharmacopeial monographs and standards, general rules exist about precision and accuracy, for example for the difference between 0.3% and 0.30%. Threshold values such as NMT 0.3%, however, are agnostic to statistics as they are expressed with only one significant number. It is always good to define the specifications of an analytical method to avoid this confusion, for example by clarifying that threshold values or ranges include three significant numbers, two true and one uncertain, as is common in analytical chemistry (Eurachem, 2012). This will also allow testing labs to demonstrate their fitness for analytical purposes. Considering our results, from the viewpoint of statistics alone, it may be adequate to introduce ranges of two standard deviations of the legal limit when defining the thresholds that determine the legal designation of a crop as hemp vs. marijuana.

Until these points are addressed, the definition of hemp based on an arbitrary number with one significant figure as upper boundary ("0.3%") of total THC content remains flawed, due to the demonstrated variability in the reported measurements of that value and the difficulty and cost of making the measurement for each crop. It should also be borne in mind that genotyping, rather than chemotyping, offers a valuable approach to the distinction of hemp from marijuana. Legal definitions that consider the existing analytical evidence will help protect hemp farming from the undue liability of misrepresented or inaccurate strain identification and flawed chemical analysis, and will foster the rational development of hemp crops with expanded utility by means of specific chemotype.

5. Captions of figures

Figure 1. Total THC content distributions as reported by ten participating commercial hemp analysis laboratories in the U.S.

Figure 2. Total cannabinoid concentration reported (top) and a comparison between total cannabinoids reported and number of standards used in quantitation (bottom).

References

Bridgeman, M.B., and Abazia, D. 2017. Medical Cannabis: History, Pharmacology, and Implications for Acute Care Settings. Pharmacy and Therapeutics, 42(3), 180-188. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5312634/

California Industrial Hemp Law and Regulations. CA Code of Regulations Title 3, Div. 4, Ch. 8, Art. 3. January 19, 2021. https://www.cdfa.ca.gov/plant/industrialhemp/docs/CaliforniaIndustrialHempLawandRegulations.pdf

De Meijer, E.P.M., Bagatta, M., Carboni, A., Crucciti, P., Moliterni, V.M.C., Ranalli, P., Mandolino, G. 2003. The inheritance of chemical phenotype in Cannabis sativa L. Genetics, 163, 335-346. https://doi.org/10.1093/genetics/163.1.335

Ellison, S.L. and Williams, A., 2012. Quantifying uncertainty in analytical measurement. Eurachem/CITAC.

Sawler, J., Stout, J.M., Gardner, K.M., Hudson, D., Vidmar, J., Butler, L., Page, J.E. 2015. PLOS ONE. 10, 1-9. https://doi.org/10.1371/journal.pone.0133292

Small, E., and Marcus, D. 2002. Hemp: A New Crop with New Uses for North America. Trends in New Crops and New Uses, ed. J. Janick and A. Whipkey, Alexandria, VA. American Society for Horticultural Science Press: 284-326. https://www.hort.purdue.edu/newcrop/ncnu02/v5-284.html

Nelson, M. A.; Waters, J. F.; Toman, B.; Lang, B. E.; Rück, A.; Breitruck, K.; Obkircher, M.; Windust, A.; Lippa, K. A. 2018. A New Realization of SI for Organic Chemical Measurement: NIST PS1 Primary Standard for Quantitative NMR (Benzoic Acid). Analytical Chemistry. 90, 10510-10517. https://doi.org/10.1021/acs.analchem.8b02575

Nelson, K.M., Bisson, J., Singh, G., Graham, J.G., Chen, S.N., Friesen, J.B., Dahlin, J.L., Niemitz, M., Walters, M.A. and Pauli, G.F., 2020. The essential medicinal chemistry of cannabidiol (CBD). Journal of medicinal chemistry, 63(21), 12137-12155. https://doi.org/10.1021/acs.jmedchem.0c00724

Pauli, G.F., Chen, S.N., Simmler, C., Lankin, D.C., Gödecke, T., Jaki, B.U., Friesen, J.B., McAlpine, J.B. and Napolitano, J.G., 2014. Importance of purity evaluation and the potential of quantitative 1H NMR as a purity assay: miniperspective. Journal of medicinal chemistry, 57(22), 9220-9231. https://doi.org/10.1021/jm500734a

Sarma, N. D., Waye, A., ElSohly, M. A., Brown, P. N., Elzinga, S., Johnson, H. E., H.E., Marles, R.J., Melanson, J.E., Russo, E., Deyton, L. and Hudalla, C., Vrdoljak, G.A., Wurzer, J.H., Khan, I.A., Kim, Nam-Cheol & Giancaspro, G. I. 2020. Cannabis inflorescence for medical purposes: USP considerations for quality attributes. Journal of natural products, 83(4), 1334-1351. https:/doi.org/10.1021/acs.jnatprod.9b01200

Small, E. and Cronquist, A. 1976. A practical and natural taxonomy for Cannabis. Taxon, 25: 405-435. https://doi.org/10.2307/1220524

USDA Agricultural Marketing Service. 2019. Establishment of a domestic hemp production program. Fed. Reg., 84(211), 58522-58564.

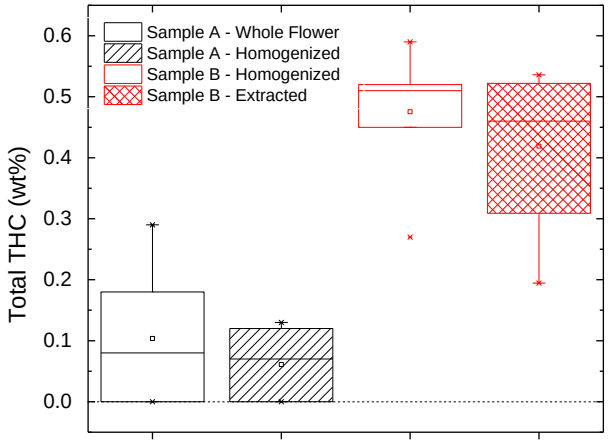

Figure 1. Total THC content distributions as reported by ten participating commercial hemp analysis laboratories in the U.S.

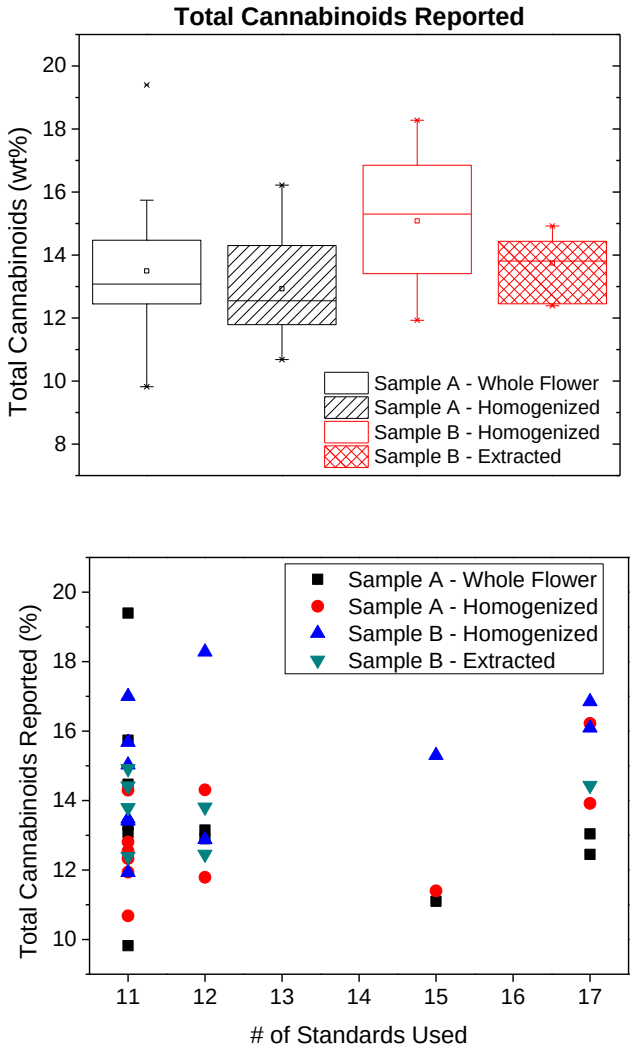

Figure 2. Total cannabinoid concentration reported (top) and a comparison between total cannabinoids reported and number of standards used in quantitation (bottom).

Table 1. Total THC content measured in identical samples extracted with various solvents

	Ethanol	Methanol	Isopropanol/acetonitrile (1:1)
Sample 1	0.520	0.445	0.45
Sample 2	0.507	0.463	0.51
Sample 3	0.502	0.459	0.47
Average	0.510 ± .007	0.456 ± .008	0.48 ± .03