Cannabis: From Cultivar to Chemovar II—A Metabolomics Approach to Cannabis Classification

3 authors, including:

Arno Hazekamp

Bedrocan BV

50 PUBLICATIONS **975** CITATIONS

Some of the authors of this publication are also working on these related projects:

Project

The effects of inhaled delta-9-tetrahydrocannabinol (THC) and cannabidiol (CBD) on pain relief and subjective effects in fibromyalgia patients View project

Cannabis and
Cannabinoid Research

Mary Ann Liebert, Inc. publishers

ORIGINAL RESEARCH   **Open Access**

Cannabis:
From Cultivar to Chemovar II—A Metabolomics Approach to Cannabis Classification

Arno Hazekamp,[1,*] Katerina Tejkalová,[2] and Stelios Papadimitriou[2]

Abstract

Introduction: There is a large disparity between the "cultural" language used by patients using cannabis for self-medication and the "chemical" language applied by scientists to get a deeper understanding of cannabis effects in laboratory and clinical studies. The distinction between Sativa and Indica types of cannabis, and the different biological effects associated with them, is a major example of this. Despite the widespread use of cannabis by self-medicating patients, scientific studies are yet to identify the biochemical markers that can sufficiently explain differences between cannabis varieties.

Methods: A metabolomics approach, combining detailed chemical composition data with cultural information available for a wide range of cannabis samples, can help to bridge the existing gap between scientists and patients. Such an approach could be helpful for decision-making, for example, when identifying which varieties of cannabis should be made legally available under national medicinal cannabis programs. In our study, we analyzed 460 cannabis accessions obtained from multiple sources in The Netherlands, including hemp- and drug-type cannabis.

Results: Based on gas chromatography analysis of 44 major terpenes and cannabinoids present in these samples, followed by Multivariate Data Analysis of the resulting chromatographic data, we were able to identify the cannabis constituents that may act as markers for distinction between Indica and Sativa. This information was subsequently used to map the current chemical diversity of cannabis products available within the Dutch medicinal cannabis program, and to introduce a new variety missing from the existing product range.

Conclusion: This study represents the analysis of the widest range of cannabis constituents published to date. Our results indicate the usefulness of a metabolomics approach for chemotaxonomic mapping of cannabis varieties for medical use.

Keywords: cannabinoids; cannabis; gas chromatography; multivariate data analysis; terpenes; varieties

Introduction

Research into herbal cannabis poses serious challenges to modern medicine, which operates mainly according to the "single compound–single target" paradigm of pharmacology. Although it was once believed that THC (see Table 1 for abbreviations) was the unique active principle found in cannabis, it is becoming clear that a much wider range of cannabis constituents may be involved in its various therapeutic effects. Currently, many different subtypes of cannabis are known to exist, and the high number of (potential) active components significantly complicates a conventional reductionist approach using analytical chemistry, animal studies, and clinical trials, where typically a single active ingredient is identified before development of a final medicine is possible. With the recent growth in medicinal use of cannabis, the need to clearly distinguish between various cannabis plants and their

[1]Department of Research and Education, Bedrocan BV, Veendam, The Netherlands.
[2]Natural Products Laboratory, Institute of Biology, Leiden University, The Netherlands.

*Address correspondence to: Arno Hazekamp, PhD, Department of Research and Education, Bedrocan BV, Veendam 9640CA, The Netherlands, E-mail: a.hazekamp@bedrocan.com

Table 1. List of Analyzed Terpenes and Cannabinoids

	RT (min)	RRT	Compound
Internal standard	10.64	0.427	1-Octanol
Monoterpenes	6.37	0.256	Alpha-2-pinene
	7.59	0.304	Beta-2-pinene
	7.94	0.318	Myrcene
	8.38	0.336	Alpha-phellandrene
	8.60	0.345	Delta-3-carene
	9.18	0.368	Beta-phellandrene[1]
	9.23	0.370	R-limonene[1]
	9.33	0.374	Cineol
	9.88	0.396	Cis-ocimene
	10.29	0.413	Gamma-terpinene
	11.39	0.457	Terpinolene
	11.80	0.473	(–)-Linalool
	12.36	0.496	Beta-fenchol
	12.67	0.508	Cis-sabinene hydrate
	13.70	0.549	Camphor
	14.46	0.580	Borneol
	15.49	0.621	Alpha-terpineol
Sesquiterpenes	23.86	0.957	Cis-bergamotene
	24.93	1.000	Beta-caryophyllene
	25.57	1.026	Trans-bergamotene
	25.68	1.030	Alpha-guaiene
	25.70	1.031	Aromadendrene
	26.28	1.054	Alpha-humulene
	26.41	1.059	Trans-beta-farnesene
	28.44	1.141	Gamma-selinene
	28.56	1.146	Delta-guaiene
	29.45	1.181	Gamma-cadinene
	29.70	1.191	Eudesma-3,7(11)-diene
	30.27	1.214	Gamma-elemene
	30.69	1.231	Nerolidol[2]
	31.23	1.253	Trans-beta-caryophyllene
	31.29	1.255	Beta-caryophyllene oxide
	31.81	1.276	Guaiol
	32.62	1.308	Gamma-eudesmol
	33.67	1.351	Beta-eudesmol
	34.89	1.400	Alpha-bisabolol
Cannabinoids	54.46	2.185	Tetrahydrocannabivarin (THCV)
	57.29	2.298	Cannabidiol (CBD)
	57.58	2.310	Cannabichromene (CBC)
	58.83	2.360	Cannabigerol-monomethylether (CBGM)
	59.18	2.374	Delta-8-tetrahydrocannabinol (delta 8-THC)[3]
	59.89	2.402	Delta-9-tetrahydrocannabinol (THC)[3]
	61.37	2.461	Cannabigerol (CBG)
	61.57	2.469	Cannabinol (CBN)[3]

(1): beta-phellandrene and R-limonene were combined for analysis because of chromatographic peak overlap. (2): nerolidol may in fact be alpha-gurjunene. (3): delta-8-THC, CBN, and THC were combined for final analysis to one value for "THC total."

GC, gas chromatography; RT, GC retention time in minutes; RRT, relative retention time compared to beta-caryophyllene.

expected (therapeutic) effects has become more important than ever.

Medicinal cannabis may have therapeutic effects on various illnesses ranging from chronic pain and multiple sclerosis (MS), to epilepsy and anxiety.[1,2] An obvious question is how the chemical composition of cannabis reflects these various medicinal effects, and what types of cannabis should consequently be made available to pro-

vide patients with a full spectrum of therapeutic benefits. Already more than 700 different cultivated varieties (cultivars) of cannabis have been cataloged[3] and many more are thought to exist, each one with a potentially different effect on body and mind. Although pharmacopoeia monographs for the chemical analysis of cannabis have been developed in several countries,[4–6] these are intended for potency, quality, and adulteration issues only. They are not designed to map out the complex chemical variations found among different cannabis products. A better classification of cannabis varieties, and the chemical differences between them, would certainly promote further implementation of cannabis-based products into (pre)clinical research and modern medicine.

Science-based classification systems

The scientific debate about classification of the cannabis species has been going on for centuries. According to current scientific consensus, Cannabis is monotypic and consists only of a single species *Cannabis sativa L.*, as originally described by Leonard Fuchs in the 16th century.[3] Within this species, two important subdivisions are commonly made. One of them recognizes drug-type versus fiber-type cannabis based on the intended use of the plant, and is mostly relevant for legal purposes. The other subdivision is based on botanical principles and identifies Sativa versus Indica types of cannabis, both regarded as a subspecies of *Cannabis sativa L.*[7]

According to the botanical description of cannabis, Sativa types of cannabis were originally grown in the Western world on an industrial scale for fiber, oil, and animal feedstuff. They are characterized by tall growth with few, widely spaced, branches and long, thin leaves. In contrast, plants of the Indica type originated in South Asia and were known historically as Indian hemp. They are characterized by shorter bushy plants and broader leaves, typically maturing relatively fast. The two groups tend to have a different smell, which may reflect a different profile of fragrant terpenes. Most cannabis plants that are currently commercially available are in fact a hybrid (cross-breed) of Sativa and Indica ancestors. Cannabis-type *ruderalis* is sometimes also recognized as a separate subspecies. It is a smaller and "weedy" plant originally from Central Russia.[7]

Various scientific attempts have been made to classify cannabis plants based on their cannabinoid composition, as outlined in more detail in our previous article.[8] For forensic and legal purposes, the most important classification is that of the drug type (marijuana) versus the fiber type (hemp), with an emphasis on the total THC

content in the flowers of the plant. Until recently, cannabis products used for medicinal purposes in official programs all belonged to the drug type, because of their high content of the biologically active THC. However, it is becoming increasingly clear that multiple constituents may be involved in the overall effect of the drug.[9] This includes the cannabinoids CBD,[10,11] CBG,[12,13] and THCV,[14,15] as well as a variety of terpenes.[16]

Terpenes are volatile compounds responsible for the typical smell and taste of cannabis. More than 120 different types have been identified in cannabis and their relative composition can vary widely between varieties.[17,18] Terpenes present in cannabis have a wide range of known biological effects, and some may be involved in regulating and/or modifying the effects of THC and other cannabinoids.[16] Expanding our scientific understanding of the therapeutic benefits of terpenes is a budding new frontier of medical cannabis research. Nevertheless, no currently used cannabis classification system takes the terpenes into account.

Since the publication of the draft genome of cannabis,[19] various studies have looked into the genetic analysis of cannabis as a means of identifying distinct (sub)groups.[20,21] Such studies have been able to clearly differentiate hemp from marijuana-type cannabis,[21] but they find only a moderate correlation between the ancestry of cannabis strains, as reported by breeders, and the ancestry inferred from their DNA. As a result, the genetic identity of cannabis products could not be reliably inferred by its vernacular name or by its reported ancestry as supplied by cannabis growers.

Vernacular classification

Over the years, illicit growers of marijuana have created an enormous range of cultivated varieties, also popularly known as cultivars or strains. These are commonly distinguished, by plant breeders, recreational users, and medical cannabis patients alike, through the use of popularized vernacular names such as *White Widow*, *Northern Lights*, *AK-47*, or *Amnesia Haze*. The most common way to categorize such strains is based on plant morphology (phenotype), which considers features such as leaf shape, plant height, color, smell, and speed of growth. This has resulted in a vernacular system of distinguishing Sativa from Indica types of cannabis, which developed independent from scientific and taxonomic classification systems.[3] According to the vernacular system, Sativa strains are usually characterized as uplifting and energetic. The effects are mostly cerebral ("head-high"), also described

as spacey or hallucinogenic. This type gives a feeling of optimism and well-being. In contrast, Indica strains are primarily described as calming and grounding ("body-high"). This type is said to cause relaxation, stress relief, and an overall sense of serenity.[22]

Because of limited guidance and support from the medical community, self-medicating users of cannabis often adopt the vocabulary of cannabis subculture to navigate their search for cannabis medicine.[23] When buying cannabis for recreational as well as for medicinal reasons, consumers use the vernacular Sativa and Indica labels as a means to identify their preferred products. Unfortunately, it is largely unclear how these labels reflect an actual difference in chemical composition, and how they are related to the scientifically accepted botanical classification of cannabis cultivars, as described above. A major issue with vernacular classification is that there are no generally agreed standards or enforced regulations for variety labeling and naming. As a result, the Sativa/Indica distinction made by botanists and taxonomists may be very different from the one applied by consumers of cannabis.[22]

Several studies have challenged the usefulness of the current Sativa/Indica classification system.[3] For example, in a recent study, almost 500 cannabis samples available to California patients were analyzed for chemical composition.[24] The study concluded that popular product names as defined by the authoritative Leafly database (www.leafly.com) were no guarantee to obtain a product with a reproducible chemical composition, and that the provided product name was not a reliable indicator of THC potency or chemical profile. In fact, samples collected under the same name often did not even look similar in appearance. Alternative classifications have been proposed, such as Broad Leaflet versus Narrow Leaflet drug types.[20,25] Another recent article suggested a complete overhaul of the current botanical classification based on reinterpretation of historical taxonomic records.[26]

From cultivar to chemovar

For medicinal users of cannabis, the main challenge is to find a product with a chemical profile that matches their pharmacological need for treatment. The name of the cannabis product serves merely as a means to identify and purchase the desired product. What is needed is a practical manner to more directly visualize the chemical diversity present within the many cannabis products offered, and to make sure that the full range of diversity is accessible through legally available,

high-quality products under national programs safe-guarding quality and consistency.

A comprehensive overview of chemical diversity can help medicinal users and their physicians to successfully transition from a beneficial cannabis product obtained through illicit markets, to a comparable but high-quality variety that is available through official national programs. It may also help these national programs to narrow down the search for beneficial cannabis cultivars to be legally produced and introduced into healthcare. Such an approach fits exactly within the research field of metabolomics, which is recognized as a scientific methodology to better understand the complex effects of medicinal plants containing multiple active ingredients. A metabolomics approach would simultaneously identify and quantify all major cannabinoids and terpenes present in various cannabis types, and then use powerful statistical tools such as Multivariate Data Analysis to create a "map" of chemical diversity.

This methodology has already been successfully applied to cannabis for differentiation of cannabis products on a small scale as well as for quality control.[8,27] With this approach, it may be possible to move away from the current system of cannabis *cultivars*, with often vague and unsubstantiated characteristics dominated by Sativa/Indica labeling, toward a new classification based on chemical varieties (or "*chemovars*") with a complex, but well-defined and reproducible chemical profile.

Goal of our study

In this article, we address the question whether cannabis vernacular product names can be replaced with chemical profiles as a means of classifying cannabis products into distinct groups. A metabolomics approach was used to investigate the chemical variation present within 460 accessions of marijuana as well as hemp collected in The Netherlands. We were particularly interested in the differences in composition between the Indica and Sativa types. Samples were chemically profiled for 44 different major cannabinoids and terpenes using gas chromatography with flame ionization detection (GC/FID), followed by statistical evaluation of the chromatographic data. The study results indicate the usefulness of the *chemovar* approach (full chemical analysis for chemotaxonomic mapping of cannabis varieties) and may assist in a better identification of cannabis products with a potential for medicinal use.

To highlight the feasibility of our proposed *chemovar* mapping, we describe an example from the Dutch medicinal cannabis program where the experiences of cannabis-using patients were combined with our chem-ical profiling approach to introduce a new standardized cannabis variety for Dutch patients.

Materials and Methods

Solvents and chemicals

Authentic standards were available for the following terpenes: alpha-pinene, beta-pinene, beta-caryophyllene, ocimene, alpha-phellandrene, R-(+)-limonene, gamma-terpinene, (+)-borneol, (−) caryophyllene oxide, (−)-*trans*-caryophyllene, cineole, fenchol, alpha-bisabolol, camphor, myrcene, farnesene (mixture of isomers), S-(−)-limonene, gamma-terpineol, phytol, carvacrol, and (+)-aromadendrene were obtained from Sigma-Aldrich; alpha-humulene, terpineol, (−)-linalool, and *trans*-nerolidol were from Fluka; beta-eudesmol and geraniol were from Chromadex; delta-3-carene was from Roth.

Calibrated standards were available for the following cannabinoids: delta-9-tetrahydrocannabinol (THC), cannabidiol (CBD), cannabigerol (CBG), cannabichromene (CBC), *trans*-(−)-delta-9-tetrahydrocannabivarin (THCV), and cannabinol (CBN). They were purified and quantified as previously described.[8,28,29] A standard for cannabigerol-monomethylether (CBGM) was not available.

All organic solvents and 1-octanol (used as internal standard [IS]) were of analytical reagent grade or HPLC grade (Sigma-Aldrich).

Sample collection

A total of 460 accessions were collected from three different sources (coffeeshops, Bedrocan, HempFlax) as described below. Vernacular names and presumed botanical type (Sativa, Indica, hybrid, hemp) were recorded as provided by the source. All samples were kept in a freezer at −20°C until analyzed.

Cannabis cultivation and sales are formally illegal under Dutch law. However, through the famous outlets known as "coffeeshops" the sales of small quantities of cannabis are tolerated (condoned) under strict conditions. There are currently about 600 coffeeshops in The Netherlands, with the majority located in the bigger cities.[30] We collected a total of 56 accessions from various coffeeshops present in several Dutch municipalities. One gram of each product was purchased and stored in the package provided (typically a small plastic zip-lock bag). In general, we requested to purchase the most typical Sativa-dominant and Indica-dominant products available in the shop. No verifiable phenotypic information about the original plants is available for these samples.

Bedrocan BV is the official cultivator of pharmaceutical-grade cannabis within the Dutch medicinal cannabis program. A total of 37 different accessions were obtained from its breeding program. Included were six standardized varieties currently available to patients on prescription (official product names: Bedrocan, Bedrobinol, Bediol, Bedica, Bedrolite, and Bedropuur). Voucher specimens have been deposited at the Naturalis National Herbarium in Leiden, The Netherlands (L) under sample codes "Hazekamp 1–6." All accessions were grown from female clones under strictly standardized indoor conditions. Flowertops were harvested and air-dried for 1 week under controlled temperature and humidity. Final samples were delivered in airtight, triple-laminate aluminum foil bags. Detailed descriptions and pictures of the plants are available.

HempFlax BV produces semi-finished and finished products made from ecologically grown fiber hemp and flax. Under their current research license, the company is propagating a wide range of cannabis varieties, including hemp, Sativa, and Indica types. Seeds of these accessions are sold through the Dutch company Sensi Seed Bank. From this source, we received a total of 367 accessions in the form of dried flowers. Plants were grown under indoor conditions. Samples were delivered in plastic zip-lock bags. Detailed descriptions and pictures of these plants are available.

Sample extraction

Samples were homogenized by manually cutting the plant material in smaller pieces with scissors. Of each sample, 500 mg (accuracy ±5 mg) was weighed and transferred to a 50-mL plastic serum tube. To each tube, 40 mL of ethanol (absolute) was added, and the tube was agitated for 10 min with a mechanical shaker (Yellow Line Orbital Shaker OS 2 Basic; IKA GmbH) at 300 rpm. The tube was then centrifuged and clear supernatant was poured into a 100-mL volumetric flask. For exhaustive extraction, the procedure was repeated twice more with 25 mL of ethanol, and supernatants were combined. At this point, 200 μL of IS, consisting of a 1% solution of 1-octanol in ethanol, was added to each extract, resulting in a final IS concentration of 0.002% (v/v). Finally, volume was adjusted to 100 mL with ethanol and mixed well. The solution was centrifuged to remove solid particles and stored at −20°C until analyzed by GC. All accessions were analyzed in duplicate.

GC analysis

An Agilent GC 6890 series (Agilent Technologies, Inc.) equipped with a 7683 autosampler and an FID was used for simultaneous analysis of monoterpenes, sesquiterpenes, and cannabinoids, as previously described.[8] The instrument was equipped with a DB5 column (30 m length, 0.25 mm internal diameter, film thickness 0.25 μm; J&W Scientific, Inc.). The injector temperature was 230°C, with an injection volume of 1 μL, a split ratio of 1:20, and a carrier gas (N_2) flow rate of 1.2 mL/min. The temperature gradient started at 60°C and increased at a rate of 3°C/min until 240°C, which was held for 5 min making a total run time of 65 min/sample. The FID detector temperature was set to 250°C. The GC-FID was controlled by Agilent GC Chemstation software version (rev. B.04.01).

In each sample chromatogram, typically 30–40 chromatographic peaks were visible. All peaks were manually integrated and resulting peak area values were entered into an Excel datasheet. Raw integration data were then corrected for IS (1-octanol) peak area and used for principal component analysis (PCA) as described below. It should be noted that for each sample, the peak areas for CBN, delta-8-THC, and THC, were combined to one single value for "THC total." Because CBN and delta-8-THC are known degradation products of THC (resulting from long-term storage and aging) and the cannabis plant does not originally produce these degradation products as part of its metabolism, they are considered to be part of the actual (original) content of THC. In general, peak areas of CBN and delta-8-THC were present in trace amounts only.

Peak identification and quantification

To confirm peak identification, selected samples were further analyzed by GC-MS using a single quadrupole MS-detector in total ion count mode as previously described.[27] Compounds were compared based on their mass spectra and retention times with authentic standards as well as literature reports.[31–36] In addition, the NIST library was used to assist in compound identification (version 2.0f; Standard Reference Data Program of the National Institute of Standards and Technology, as distributed by Agilent Technologies).

For quantitative analysis, peak area values were quantified (in mg/g of plant material) with the use of calibration curves. Monoterpenes, sesquiterpenes, and cannabinoids present in the samples were quantified by using calibrated standards of beta-pinene, alpha-humulene, and CBD, respectively, comparable to our method previously described.[8] Each calibration curve consisted of four different concentration levels in the range of 0.005–0.1 mg/mL (beta-pinene,

alpha-humulene) or 0.25–1.0 mg/mL (CBD) dissolved in ethanol. Calibration curves were regularly prepared throughout the duration of the study. The resulting quantitative data were not corrected for residual moisture content of the samples.

Multivariate data analysis

After correcting for IS and averaging of duplicate analyses, peak integrals from the GC-FID chromatograms were used for Multivariate Data Analysis with the aid of SIMCAP+ version 13.0.3 software (Umetrics AB). All values were scaled with unit variance before modeling, which gives equal weight to all variables.

First, *unsupervised* analysis was performed by applying PCA to identify clusters of samples that have similar chemical composition. PCA is a statistical way of identifying patterns in complex data sets, by turning the data set into linearly uncorrelated values, which are the principal components (PCs).[37] The resulting plot highlights the major similarities and differences between the analyzed samples.[38] With this method, there is no *a-priori* expectation of the nature of such clusters, and we leave it up to the computer program to tell us if clusters of data points actually exist. For this analysis, we used GC data of all 460 accessions. The first two principal components (PC1 and PC2) were visualized in a scatter plot, in combination with its accompanying loading plot, to show the contribution of each chemical component to the diversity observed in the sample collection.

Next, a *supervised* analysis was performed by applying orthogonal projection to latent structure-discriminant analysis (OPLS-DA) on the GC data obtained from the 131 accessions in our data set that were labeled as pure Indica or Sativa. OPLS-DA is a form of partial least-squares regression analysis (a dimensionality reduction technique) combined with discriminant analysis (which is a form of classification). OPLS-DA is a further development from PCA and a method of choice when working with a data set consisting of known distinct classes (in our case, Sativa versus Indica).[39] This method shows which variables are responsible for class discrimination in an easy to interpret 2D visualization.[40]

In addition, for a better understanding of the quantitative differences between Sativa, Indica, and hemp types of cannabis, GC data were quantified with the use of calibration curves to determine the absolute content (in milligrams) of each component per gram of plant material. Statistical analysis was done by *T*-test to identify significant differences in cannabinoid or terpene composition between these three cannabis types. Finally, PCA map-ping of the entire cannabinoid and terpene profile was shown to be a useful way to identify missing cannabis types within the Dutch medicinal cannabis program.

Results and Discussion
Cannabis sampling and analysis

In the present study, we chemically analyzed 460 different cannabis accessions from various sources, representing one of the largest cannabis sample sets analyzed to date. Based on information obtained from suppliers, accessions were labeled as Hemp ($n=121$), Sativa ($n=68$), Indica ($n=63$), or Hybrid ($n=208$). With the use of GC analysis, 44 different compounds could be positively identified with a single chromatographic run, including 17 monoterpenes, 19 sesquiterpenes, and 8 cannabinoids (Table 1). It should be noted that during GC analysis, all acidic cannabinoids (e.g., THC-acid and CBD-acid) are fully converted to their neutral counterparts (THC, CBD, etc.) as a result of the high temperatures (up to 250°C) applied during injection and separation.

Since our previous study,[8] we were able to extend our list of positively identified terpenes from 20 to 36, by more closely studying the compounds previously identified by Ross and ElSohly[33] and Hillig,[34] as well as by the use of newly acquired standards and performing more detailed GC-MS analysis of unresolved peaks. Only one of the newly added compounds could not be unequivocally identified: the compound we putatively identified as nerolidol (in accordance with Ross and ElSohly[33]) may perhaps be alpha-gurjunene.[34] Unfortunately, we did not have the pure standard for alpha-gurjunene needed to distinguish between these two. Because of lack of resolution between the chromatographic peaks of beta-phellandrene and R-limonene, these two compounds were analyzed as a single peak.

Compared to our previous study on cannabis composition,[8] we made one modification to the extraction and analysis protocol, that is, the addition of an IS in the form of 1-octanol. In each sample, the IS peak area was used to compensate for small deviations in the injection volume of the GC, to obtain more accurate data. The IS did not overlap with other chromatographic peaks present in our samples, as indicated in Table 1.

Unsupervised data analysis: visualizing patterns in the data set

An unsupervised PCA approach was used to visualize any potential clustering in the data set, based on the analysis of a large number of terpenes and cannabinoids.

The data set consisted of 460 samples containing up to 44 identifiable compounds each, which equals over 20.000 variables (i.e., 460×44) for evaluation. For the statistical analysis, each analyzed compound represented a variable while its quantity (i.e., peak area compensated for IS) represented the observation. Figure 1a shows an evaluation of the first two principal components (PC1 and PC2) in the form of a score plot. It is clearly shown that the main distinction observed in our data set was between hemp-type samples on the left versus drug-type samples (Sativa + Indica + hybrids) on the right. PC1 (28.3%) and PC2 (8.5%) together explained 36.8% of the total variance found in the sample set; almost twice as high as the model proposed by Hillig[25] (PC1 + PC2 = 19.6%) and equal to the model by Elzinga et al.[24] (PC1 + PC2 = 34.6%). Analysis of PC3 added an additional 7.7% to the total variance explained by our model.

The accompanying loading plot (Fig. 1b) visualizes how samples differ from each other, and which compounds (markers) contribute the most to explaining the observed grouping. The two components showing the largest contribution to PC1 are CBD (far left of the plot) and THC (far right). It is striking that all components, except CBD, in the loading plot appear on the right side of the horizontal (PC1) axis, which means that all are more prominently present in drug cultivars. Although the ratio between THC and CBD is a known marker for the differentiation between hemp- and drug-type cannabis,[41] it may be a surprise that not a single terpene turns out to be a clear marker associated with hemp.

Among the drug-type accessions, a further separation of samples is seen along the vertical (PC2) axis, separating samples with relatively higher content of sesquiterpenes such as alpha-humulene, beta-caryophyllene, gamma-cadinene, and eudesma-3,7(11)-diene (upper half of the plot) versus those typified by relatively higher content of monoterpenes, such as gamma-terpinene, alpha-phellandrene, cis-ocimene, alpha-terpineol, and terpinolene (lower half). Interestingly, cannabinoids only play a limited role in this distinction, as they are mainly aligned along the PC1 axis, between the two groups. The only exception was the minor cannabinoid CBC, which was more strongly associated with the subgroup defined by higher sesquiterpene content.

Supervised data analysis: identifying markers in Sativa versus Indica types

When we applied a supervised (discriminatory) form of Multivariate Data Analysis, we were able to test the hypothesis that certain clusters must exist, based

on *a-priori* knowledge we possess about the sample set. More specifically, we wanted to verify whether clear differences exist between the chemical profiles of Sativa versus Indica samples, assuming these two classes actually exist. In other words, if indeed meaningful chemical differences exist between the vernacular Sativa and Indica types, OPLS-DA modeling must be able to identify the markers that are associated with the two classes. We therefore included only the 68 accessions labeled as pure Sativa and 63 accessions labeled as pure Indica, while hemp varieties and hybrids were excluded from this particular analysis.

Results of the modeling are visualized in the score plot shown in Figure 2a. Using the OPLS-DA method, the horizontal (X) axis visualizes maximum variation *between* classes (Sativa versus Indica), while the vertical (Y) axis shows maximum variation *within* these classes. Indeed, the score plots of the OPLS-DA model showed that the Sativa samples were distinguishable from the Indica samples with only very mild overlap. Good model quality was indicated by the values of R2X (0.387) and Q2 (0.582). The coefficient loading plot of the OPLS-DA model was used to identify the components responsible for sample differentiation on the score plot (Fig. 2b).

By subsequently constructing an S-plot (figure not shown) based on the OPLS-DA data, the most significant markers associated with our samples could be identified. We found that the presence of the terpenes *trans*-bergamotene, *trans*-beta-farnesene, delta-3-carene, and terpinolene was most strongly associated with the Sativa-type samples, while beta- and gamma-eudesmol, guaiol, myrcene, and gamma-elemene were the most prominent markers for the Indica-type samples. Interestingly, we discovered that all hydroxylated terpenes (marked in Fig. 2b) were most strongly associated with the Indica type of samples. This begs the question whether there is a functional correlation between the presence of the hydroxylated forms of terpenes and the plant morphology of the Indica type.

Quantitative analysis of Hemp, Sativa, and Indica

All samples labeled as Hemp ($n = 121$), Sativa ($n = 68$), or Indica ($n = 63$) were quantified with the use of GC calibration curves. The use of only three standards (beta-pinene for monoterpenes, alpha-humulene for sesquiterpenes, and CBD for cannabinoids) for quantification of all sample components greatly facilitated our analysis. This approach was based on the observation that the differences in FID-detector response within
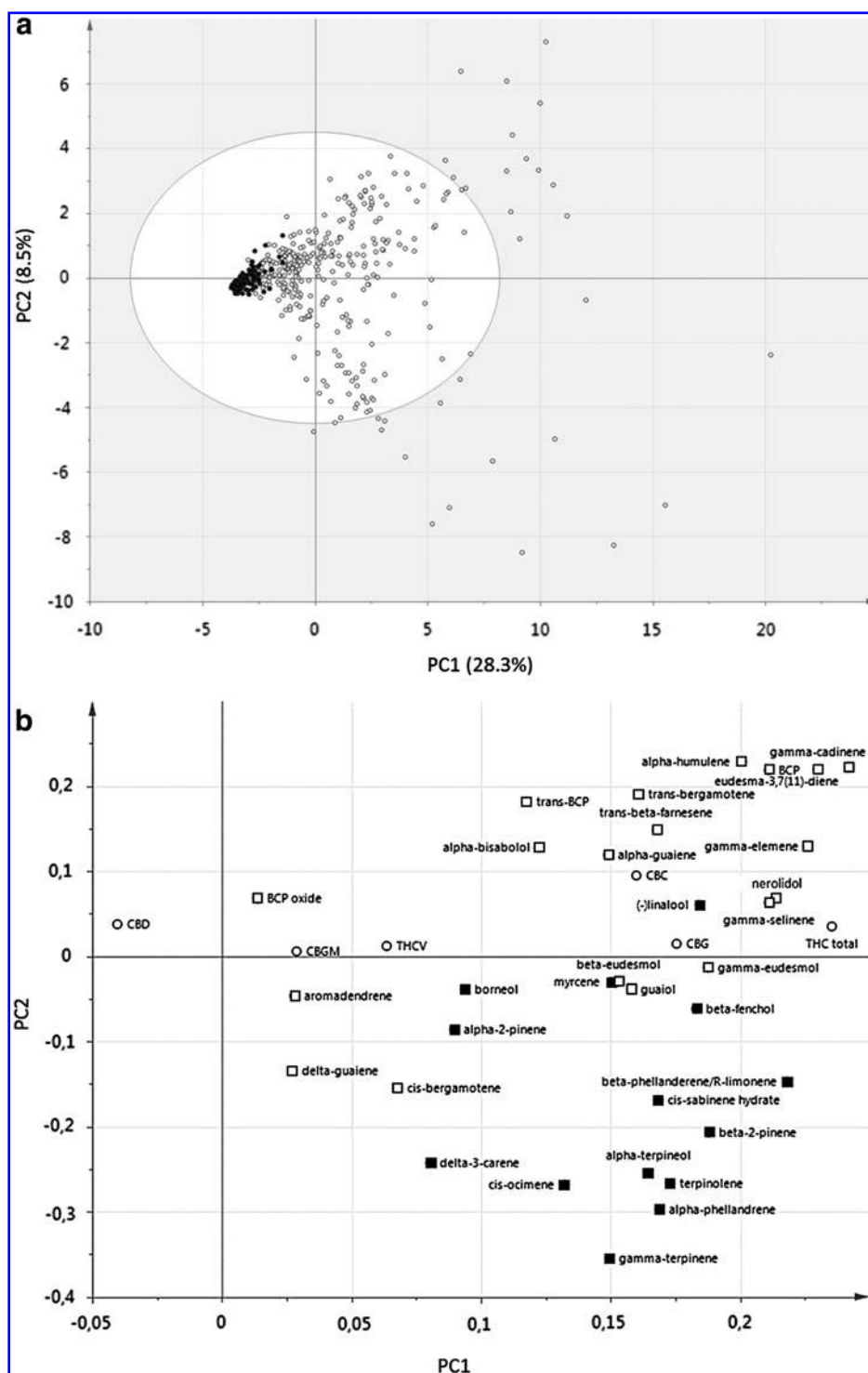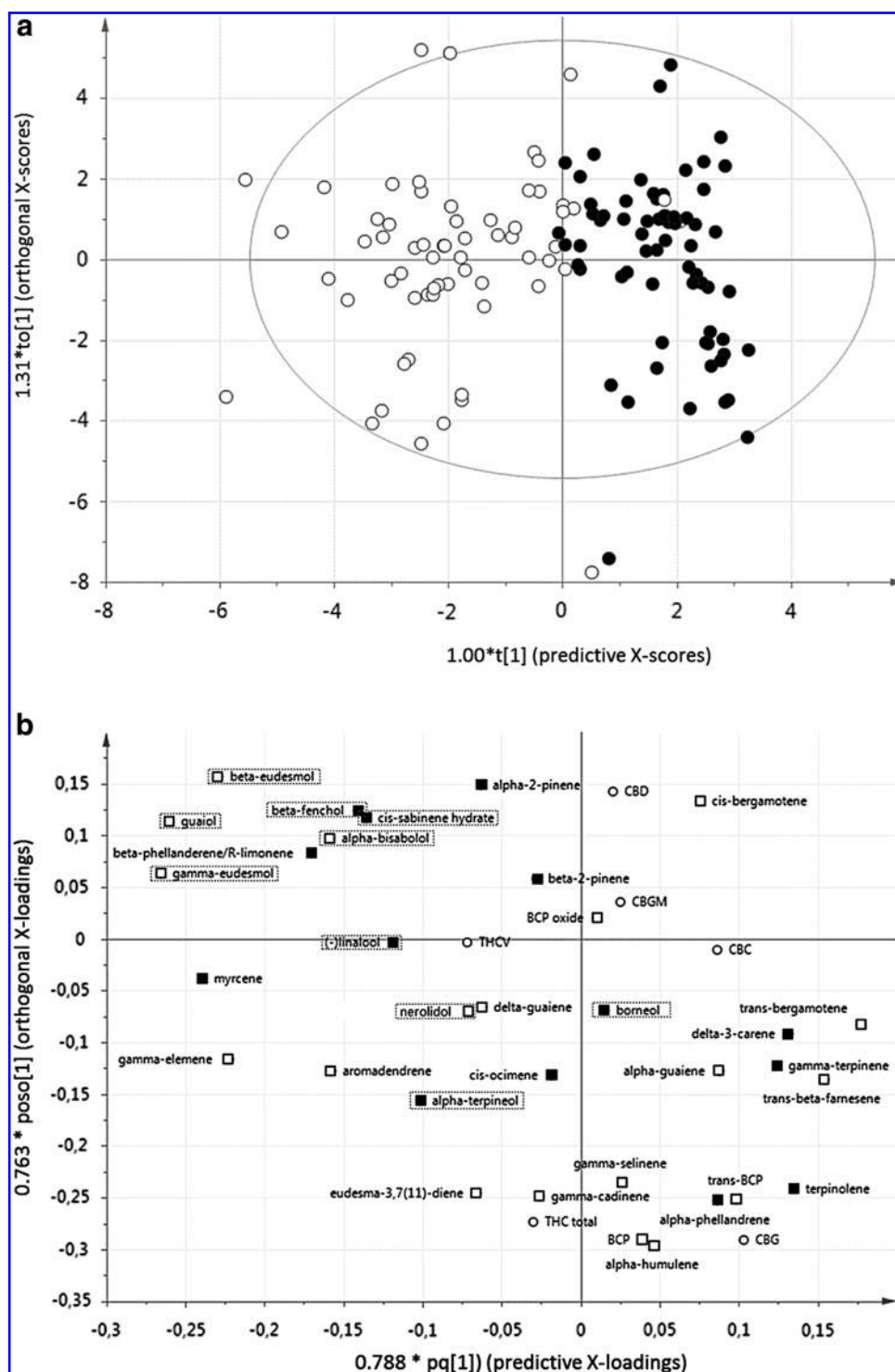
FIG. 1. **(a)** Unsupervised PCA score plot of all samples in our sample set ($n = 460$). Solid ellipse indicates 95% confidence interval based on Hotelling's T^2, which can be used to identify outliers. Hemp-type (●; $n = 121$) and drug-type (○; $n = 339$) samples are indicated. **(b)** PCA loading plot of all samples. Cannabinoids (○), monoterpenes (■), sesquiterpenes (□). PCA, principal component analysis.

FIG. 2. **(a)** OPLS-DA score plot discriminating Sativa-labeled (●; $n=68$) from Indica-labeled (○; $n=63$) samples. The ellipse in the score plot is a 95% Hotelling's T^2, which can be used to identify outliers. **(b)** OPLS-DA loading plot. Cannabinoids (○), monoterpenes (■), sesquiterpenes (□). Hydroxylated terpenes are marked with a dotted box. OPLS-DA, orthogonal projection to latent structure-discriminant analysis.

the chemical class of the monoterpenes, sesquiterpenes, or cannabinoids, respectively, are relatively small.[27] Because analyzing all 44 sample components by preparing their own standard curves would be an immense and costly task, one representative compound was chosen to quantify the others in the same chemical class. Of course, this approach only works if all observed components have a linear FID response within the analyzed concentration range. Indeed, this was shown to be true by analyzing selected extracts at different dilution levels and plotting GC peak area versus dilution for each observed peak.

The mean calculated content of cannabinoids and terpenes, reported in mg/g of sample weight, is shown in Table 2. When comparing cannabinoid profiles, all analyzed cannabinoids were found to be present in significantly lower quantities in hemp types compared to Sativa or Indica types ($p < 0.05$). The only exception was for CBD, with hemp accessions (mean: 1.1% CBD) showing a significantly higher mean content than Sativa or Indica (drug) accessions (mean: 0.3–0.4% CBD). When comparing Sativa- to Indica-type samples, we found only significant differences for the content of CBC and CBG, both of

Table 2. Quantitative Analysis of Cannabinoids, Monoterpenes, and Sesquiterpenes (in mg/g) in accessions labeled as Hemp, Indica, or Sativa

	Mean content (mg/g) $\pm$ SD		
	Hemp (n = 121)	Indica (n = 63)	Sativa (n = 68)
Alpha-2-pinene	0.52 ± 0.55a	1.21 ± 1.42b	0.82 ± 1.40b
Beta-2-pinene	0.19 ± 0.18a	0.73 ± 0.55b	0.64 ± 0.57b
Myrcene	1.06 ± 0.92a	8.67 ± 6.95b	2.03 ± 2.45c
Alpha-phellandrene	0.00 ± 0.01a	0.01 ± 0.04ab	0.03 ± 0.07b
Delta-3-carene	0.01 ± 0.05a	0.01 ± 0.06a	0.12 ± 0.23b
Beta-phellandrene/R-Limonene	0.09 ± 0.10a	1.38 ± 1.32b	0.51 ± 0.37c
Cineol*	0.00 ± 0.02a	0.00 ± 0.03a	0.00 ± 0.00a
Cis-ocimene	0.22 ± 0.24a	0.80 ± 0.93b	0.79 ± 0.91b
Gamma-terpinene	0.00 ± 0.01a	0.00 ± 0.00a	0.02 ± 0.06b
Terpinolene	0.23 ± 0.32a	0.27 ± 0.84a	0.94 ± 1.29b
(–)Linalool*	0.01 ± 0.02a	0.31 ± 0.36b	0.15 ± 0.12c
Beta-fenchol*	0.00 ± 0.00a	0.11 ± 0.17b	0.01 ± 0.04c
Cis-sabinene hydrate*	0.00 ± 0.00a	0.08 ± 0.14b	0.00 ± 0.03a
Camphor*	0.00 ± 0.00a	0.00 ± 0.00a	0.01 ± 0.06a
Borneol*	0.00 ± 0.00a	0.01 ± 0.05b	0.02 ± 0.14ab
Alpha-terpineol*	0.22 ± 0.05a	0.41 ± 0.40b	0.28 ± 0.14c
Cis-bergamotene	0.21 ± 0.05a	0.07 ± 0.11b	0.11 ± 0.11b
Beta-caryophyllene	0.98 ± 0.67a	1.94 ± 1.24b	2.16 ± 1.57b
Trans-bergamotene	0.04 ± 0.06a	0.12 ± 0.14b	0.28 ± 0.24c
Alpha-guaiene	0.00 ± 0.02a	0.17 ± 0.16b	0.29 ± 0.25c
Aromadendrene	0.00 ± 0.00a	0.13 ± 0.19b	0.02 ± 0.07c
Alpha-humulene	0.27 ± 0.19a	0.61 ± 0.40b	0.70 ± 0.61b
Trans-beta-farnesene	0.05 ± 0.07a	0.31 ± 0.26b	0.55 ± 0.37c
Gamma-selinene	0.03 ± 0.06a	0.26 ± 0.30b	0.31 ± 0.33b
Delta-guaiene	0.00 ± 0.02a	0.03 ± 0.15b	0.01 ± 0.03ab
Gamma-cadinene	0.02 ± 0.05a	0.41 ± 0.39b	0.36 ± 0.36b
Eudesma-3,7(11)-diene	0.05 ± 0.10a	0.56 ± 0.44b	0.43 ± 0.43b
Gamma-elemene	0.04 ± 0.09a	0.64 ± 0.52b	0.14 ± 0.26c
Nerolidol*	0.01 ± 0.03a	0.02 ± 0.06a	0.00 ± 0.03a
Trans-beta-caryophyllene	0.06 ± 0.06a	0.02 ± 0.05b	0.06 ± 0.11a
Beta-caryophyllene oxide*	0.00 ± 0.00a	0.01 ± 0.05a	0.01 ± 0.04a
Guaiol*	0.01 ± 0.03a	0.41 ± 0.38b	0.02 ± 0.08a
Gamma-eudesmol*	0.01 ± 0.04a	0.46 ± 0.36b	0.06 ± 0.12c
Beta-eudesmol*	0.00 ± 0.02a	0.25 ± 0.20b	0.05 ± 0.10c
Alpha-bisabolol*	0.03 ± 0.07a	0.36 ± 0.32b	0.15 ± 0.17c
THCV	0.14 ± 0.40a	2.17 ± 5.26b	0.97 ± 0.79b
CBD	10.93 ± 9.47a	2.97 ± 13.40b	3.77 ± 15.98b
CBC	0.64 ± 0.56a	2.06 ± 1.31b	2.53 ± 1.35c
CBGM	0.18 ± 0.17a	0.36 ± 0.50b	0.73 ± 4.47ab
THC	3.53 ± 5.40a	137.10 ± 53.64b	127.39 ± 81.21b
CBG	0.28 ± 0.38a	5.09 ± 4.22b	7.66 ± 7.19c

Compounds are listed in order of GC elution according to Table 1. Mean values are shown ± SD. In each row: values marked with different letters (a, b, c) are significantly different at $p < 0.05$. Hydroxylated terpenes are marked with an asterisk (*).
SD, standard deviation.

which tend to be somewhat more prominently present in Sativa types ($p < 0.05$). The mean THC content of both types was found to be comparable, with values of about 127–137 mg/g (or 12.7–13.7%). These values correspond very well with the average content reported for Dutch indoor-grown recreational cannabis offered through coffeeshops.[42]

Quantitative analysis of the terpenes confirmed the observation made by PCA that hydroxylated terpenes (marked with * in Table 2) are significantly more prevalent in Indica than in Sativa accessions. In particular, the sesquiterpenes guaiol, beta- and gamma-eudesmol, and alpha-bisabolol were strongly associated with the Indica type ($p < 0.001$). Indeed, the presence of such sesquiterpene alcohols has previously been reported to be important for the chemotaxonomic discrimination of Indica cannabis varieties originating from Afghanistan.[34]

Interestingly, in our study, we found myrcene, on average the most abundant terpene in our samples, to be strongly associated with the Indica type of cannabis ($p < 0.001$). In contrast, Hillig[34] measured relatively higher levels of myrcene in Sativa types, while Lynch et al.[20] found hemp samples to contain significantly more myrcene than drug-type cannabis. These conflicting results indicate that selecting single compounds as markers for specific cannabis types may have its limitations.

Mapping diversity: a practical application of the chemovar approach

The Dutch medicinal cannabis program currently provides six different standardized varieties to patients, with the original one (variety Bedrocan®) having been introduced as early as 2003. All six products were analyzed in the current study, together with another 31 experimental, nonstandardized accessions. Occasionally, Dutch patients have claimed that the current product range does not cover the full spectrum of varieties needed to treat all medical indications. Such opinions are commonly based on trial-and-error experiences by individual patients using cannabis obtained from a variety of sources, including coffeeshops and other illicit sources, or home-grown cannabis. Although new varieties have been introduced regularly over the years (most recently the CBD-rich variety Bedrolite® in 2014), the discussion about the need for new varieties regularly reappears. A clear and practical chemovar approach to this discussion may be useful to identify cannabis varieties that could be added to the current choice of products.

In this study, the analysis of a large number of samples available in The Netherlands has provided us with a sort of "chemical diversity map." In Figure 3, the PCA plot already shown in Figure 1a is repeated, but now with the six standardized pharmacy products (varieties Bedrocan, Bedrobinol, Bediol, Bedica, Bedrolite, and Bedropuur) highlighted to indicate their position on the map. This visualization makes it clear that two major sections of the map are not covered, marked as regions A and B. Based on the sample information obtained from the original source, we know that region A corresponds with a cluster of samples popularly labeled as *Amnesia*.

Parallel to this development, a Dutch patient activist group (Foundation PGMCG, Tilburg, The Netherlands) was campaigning for the introduction of a new cannabis type into the national cannabis program. Based on their personal experiences with the effects of various types of cannabis, they independently identified an Amnesia-type of cannabis to be missing from the spectrum of currently available varieties. Interestingly, this group was not aware of this study we were performing at the same time, which means that our own chemical testing and independent patient experiences both pointed toward Amnesia as a missing variety, simultaneously. This example indicates how real-life experience with medicinal cannabis can be fully compatible with the chemical analysis highlighted in our study, indicating the practical use of the chemovar approach. Together with the patient group, we were able to identify a cannabis variety that displayed the desired chemical profile. The variety is currently being prepared for standardized cultivation and is intended to be introduced as a new medicinal product on prescription later this year.

We are currently working on identifying a cannabis variety that has the proper chemical content of terpenes and cannabinoids to match with region B (Fig. 3). Such a variety would simultaneously need to display the desired botanical characteristics to be cultivated on a large scale (e.g., yield of flowers per plant, pest resistance, growth speed, and maximum size).

Conclusion

With modern analytical techniques, the rapid and comprehensive analysis of all cannabinoids and terpenes present in cannabis products has become a feasible option. This opens up the way to move from a confusing system of using vernacular *cultivar* names for the identification of cannabis products toward a more reliable and informative system of science-based *chemovar*
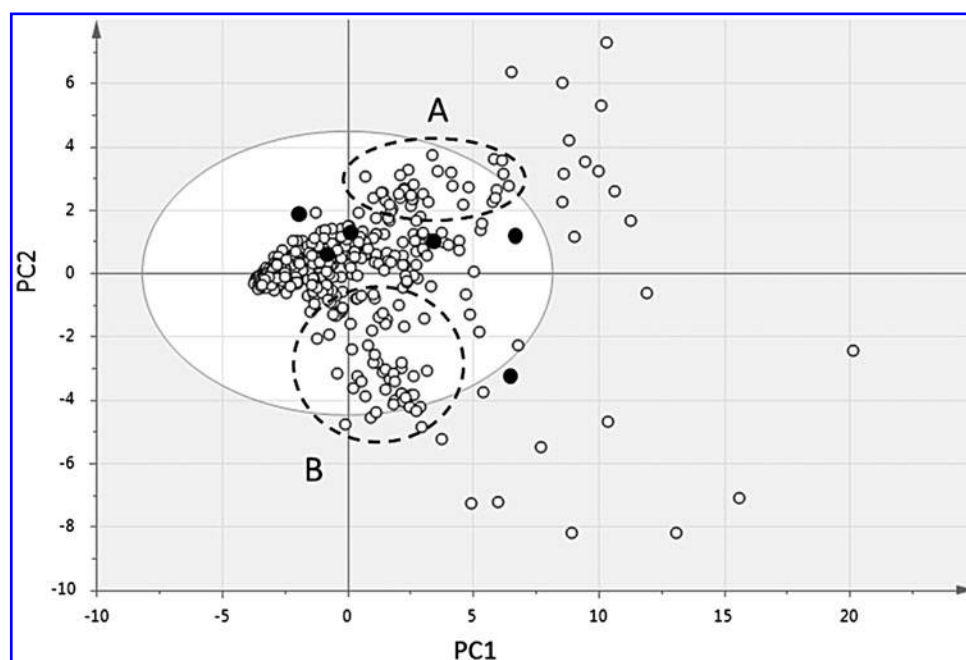
FIG. 3. Same plot as Figure 1, but indicating currently available Dutch pharmacy products (●). Two regions not covered by currently available products are indicated by dotted circles (**A** and **B**).

classification. In our study, we successfully applied this approach for cannabis differentiation. By analyzing 44 major cannabinoids and terpenes, hemp types of cannabis could be efficiently distinguished from drug-type cannabis based on PCA of the resulting GC data. Hemp accessions showed a relatively conserved chemical profile characterized by the presence of relatively more CBD. In contrast, drug accessions showed a much wider chemical diversity with all cannabinoids (except CBD) as well as terpenes present in much higher concentrations than in hemp. An additional distinction was found between drug accessions characterized by higher content of monoterpenes, versus a group that contained significantly higher content of sesquiterpenes. Potentially, this difference may be caused by evaporation of the relatively more volatile monoterpenes during drying and storage of some of the analyzed cannabis accessions.[33]

Using a supervised PCA methodology, vernacularly labeled Sativa and Indica accessions could be well separated into two distinct groups, which means that true differences seem to exist in chemical composition between these two types of cannabis. In contrast to popular belief, the Sativa and Indica types did not differ in their average content of the major cannabinoids THC or CBD. Instead, only CBC and CBG content was slightly but significantly higher in the Sativa group ($p < 0.05$). More prominent differences between the two groups were found in the terpene composition, specifically in the presence of hydroxylated terpenes associated with the Indica group. Interestingly, several other studies have associated the Indica type of cannabis with such terpenes, including guaiol and eudesmol,[34] fenchol, alpha-terpineol, linalool, geraniol, and camphor,[24] and linalool.[20] Based on these results, we believe that hydroxylated terpenes are good chemical markers for the identification of (vernacular) Indica cannabis types.

We speculate that the functional link between cannabis phenotype and hydroxylated forms of terpenes may be found in gibberellic acid (GA); a pentacyclic diterpene acid, and therefore, a hydroxylated terpene itself. It should be noted that many subtypes of GA exist. These hormones promote growth and elongation of cells, and their effect on cannabis development has been observed by scientists[43–45] as well as cannabis breeders.[46] Possibly the biosynthetic pathways for GAs and for hydroxylated terpenes share a common precursor in cannabis. Alternatively, the same catalytic enzymes may be involved in their biosynthesis, such as

a cytochrome P450 monoxygenase[47] or terpene hydroxylase.[48] Human selection over centuries of cannabis use may have led to a coevolution of plant morphology (plant shape and size) and terpene profile (reflected in smell, taste, and medicinal effects) with GAs as the connecting factor. Further research in live cannabis plants should be performed to further test this hypothesis.

As a major limitation, we should note that our study was based on the analysis of cannabis accessions collected specifically in The Netherlands. The obtained results may be different when samples from other regions are analyzed for chemical content. Although in total, more than 120 terpenes have been identified in various cannabis products so far,[16] our data suggest that most of these are only present in trace amounts. The 36 terpenes analyzed in our study (Table 1) seem to cover the whole range of terpenes found in significant quantities, at least in the samples obtained from our sources in The Netherlands. Nevertheless, it is possible that samples from other regions would contain other terpenes as prominent constituents. As a result, the mapping of cannabis diversity may be a constantly evolving effort. New accessions added will make the current map more complete but could also open up new chemical "territories" that are currently unknown.

Of course, the basis for any evaluation system based on chemical profiling is that all cannabis constituents of interest are unambiguously identified. Because of the large number of terpenes that exist (currently >20.000 are known) and the large overlap in their chemical structures and molecular weight, there is a risk that reported compounds may not be properly identified. For example, despite earlier reports of their presence in cannabis samples, the terpenes gamma-terpineol,[8] phytol,[16,20] carvacrol,[49] and geraniol[24,49] could not be detected in any of our samples. Furthermore, a recent article by Aizpurua-Olaizola et al.[50] used a GC method almost identical to ours, but reported multiple terpenes that were not found or identified in our study. To address such discrepancies, there is a need for a widely accepted and unified method for terpene analysis in cannabis, to make comparison between different studies more reliable.

To conclude, it is essential that future cannabis commerce allows complete and accurate cannabinoid and terpene profiles to be available.[51] When the emerging cannabis industry adopts a common chemical language across producers, processors, and retailers, the winners will be the patients. Based on our study, it may be con-

cluded that GC-FID-based metabolomics is a comprehensive and effective method to monitor chemical diversity in cannabis. Achieving a practical, accurate, and reliable classification system for cannabis, including a variety registration system, will require significant scientific investment and a legal framework that accepts both licit and illicit forms of this plant. Such a system is essential to realize the enormous potential of cannabis as a multiuse agricultural crop and medicinal plant. Chemovar mapping may be exactly the tool needed to achieve this goal.

Acknowledgments

Author Disclosure Statement
No competing financial interests exist.

References

1. Kowal MA, Hazekamp A, Grotenhermen F. Review on clinical studies with cannabis and cannabinoids 2010–2014. Cannabinoids. 2016;11(special issue):1–18.
2. Hazekamp A, Grotenhermen F. Review on clinical studies with cannabis and cannabinoids 2005–2009. Cannabinoids. 2010;5:1–21.
3. Erkelens JL, Hazekamp A. That which we call Indica, by any other name would smell as sweet. Cannabinoids. 2014;9:9–15.
4. Monograph Cannabis flos Version 7.1. Office of Medicinal Cannabis (OMC): The Hague: The Netherlands, 2014.
5. The Advisability and Feasibility of Developing USP Standards for Medical Cannabis. Stimuli to the revision process. USP-NF (US Parmacopeia–National Formulary): Rockville, MD, 2015.
6. Upton R, Craker L, ElSohly M, et al. Cannabis Inflorescence (cannabis spp.)–standards of identity, analysis, and quality control (revision 2014). American Herbal Pharmacopoeia (AHP): Soquel, CA, 2014.
7. Small E, Cronquist A. A practical and natural taxonomy for Cannabis. Taxon. 1976;25:405–435.
8. Hazekamp A, Fischdick JT. Cannabis—from cultivar to chemovar. Drug Test Anal. 2012;4:660–667.
9. Brodie JS, Di Marzo V, Guy GW. Polypharmacology shakes hands with complex aetiopathology. Trends Pharmacol Sci. 2015;36:802–821.
10. Szaflarski JP, Bebin EM. Cannabis, cannabidiol, and epilepsy—from receptors to clinical response. Epilepsy Behav. 2014;41:277–282.
11. Crippa JA, Derenusson GN, Ferrari TB, et al. Neural basis of anxiolytic effects of cannabidiol (CBD) in generalized social anxiety disorder: a preliminary report. J Psychopharmacol. 2011;25:121–130.
12. Pagano E, Montanaro V, Di Girolamo A, et al. Effect of non-psychotropic plant-derived cannabinoids on bladder contractility: focus on cannabigerol. Nat Prod Commun. 2015;10:1009–1012.
13. Borrelli F, Pagano E, Romano B, et al. Colon carcinogenesis is inhibited by the TRPM8 antagonist cannabigerol, a Cannabis-derived non-psychotropic cannabinoid. Carcinogenesis. 2014;35:2787–2797.
14. Englund A, Atakan Z, Kralj A, et al. The effect of five day dosing with THCV on THC-induced cognitive, psychological and physiological effects in healthy male human volunteers: a placebo-controlled, double-blind, crossover pilot trial. J Psychopharmacol. 2016;30:140–151.
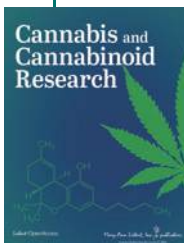
15. Tudge L, Williams C, Cowen PJ, et al. Neural effects of cannabinoid CB1 neutral antagonist tetrahydrocannabivarin on food reward and aversion in healthy volunteers. Int J Neuropsychopharmacol. 2014;18;pii: pyu094.
16. Russo EB. Taming THC: potential cannabis synergy and phytocannabinoid-terpenoid entourage effects. Br J Pharmacol. 2011;163:1344–1364.
17. Hazekamp A, Fischedick JT, Diez ML, Lubbe A, Ruhaak RL. Chemistry of Cannabis. In: Mander L, Lui H-W (eds.). Comprehensive Natural Products II Chemistry and Biology (vol. 3). Elsevier: Oxford, UK, 2010, pp. 1033–1084.
18. Brenneisen R, ElSohly MA. Chromatographic and spectroscopic profiles of cannabis of different origins: Part I. J Forens Sci. 1988;33:1385–1404.
19. Bakel H van, Stout JM, Cote AG, et al. The draft genome and transcriptome of Cannabis sativa. Genome Biol. 2011;12:R102.
20. Lynch RL, Vergara D, Tittes S, et al. Genomic and Chemical Diversity in Cannabis. BioRxiv. 2015.
21. Sawler J, Stout JM, Gardner KM, et al. The genetic structure of marijuana and hemp. PLoS One. 2015;10:e0133292.
22. Pearce DD, Mitsouras K, Irizarry KJ. Discriminating the effects of Cannabis sativa and Cannabis indica: a web survey of medical cannabis users. J Alt Comp Med. 2014;20:787–791.
23. Hazekamp A, Pappas G. Self-medication with Cannabis. In: Handbook of Cannabis. Oxford University Press: Oxford, UK, 2014.
24. Elzinga S, Fischedick J, Podkolinski R, et al. Cannabinoids and terpenes as chemotaxonomic markers in Cannabis. Nat Prod Chem Res. 2015;3:181.
25. Hillig KW. Genetic evidence for speciation in Cannabis (cannabaceae). Genet Resour Crop Evol. 2005;52:161–180.
26. McPartland J, Guy G. Correct(ed) vernacular nomenclature of Cannabis. O'Shaugnessy's online journal. Edition 4, January 2015.
27. Fischedick JT, Hazekamp A, Erkelens T, et al. Metabolic fingerprinting of Cannabis sativa L., cannabinoids and terpenoids for chemotaxonomic and drug standardization purposes. Phytochemistry. 2010;71:2058–2073.
28. Hazekamp A, Simons R, Peltenburg-Looman A, et al. Preparative isolation of cannabinoids from Cannabis sativa by centrifugal partition chromatography. J Liq Chrom Rel Technol. 2004;27:2421–2439.
29. Hazekamp A, Choi YH, Verpoorte R. Quantitative analysis of cannabinoids from Cannabis sativa using 1H-NMR. Chem Pharm Bull. 2004;52:718–721.
30. Coffeeshops in Nederland 2014 [Coffeeshops in the Netherlands 2014]. Bureau Intraval: Groningen-Rotterdam, The Netherlands, 2015.
31. Komori T, Nohara T, Hosokawa I, et al. Cannabigerol monomethyl ether, a new component of hemp. Chem Pharm Bull. 1968;16:1164–1165.
32. Adams RP. Identification of essential oils by ion trap mass spectrometry. Academic Press Inc.: New York, 1989.
33. Ross SA, ElSohly MA. The volatile oil composition of fresh and air-dried buds of Cannabis sativa. J Nat Prod. 1996;59:49–51.
34. Hillig KW. A chemotaxonomic analysis of terpenoid variation in cannabis. Biochem Syst Ecol. 2004;32:875–891.
35. Hillig KW, Mahlberg PG. A chemotaxonomic analysis of cannabinoid variation in cannabis (cannabaceae). Am J Bot. 2004;91:966–975.
36. Rothschild M, Bergström G, Wångberg S. Cannabis sativa: volatile compounds from pollen and entire male and female plants of two variants, Northern lights and Hawaian indica. Bot J Linnean Soc. 2005;147:387–397.
37. Jolliffe IT, Cadima J. Principal component analysis: a review and recent developments. Philos Trans A Math Phys Eng Sci. 2016;374:20150202.
38. Eriksson L, Johansson E, Kettaneh-Wold N, et al. Multi- and megavariate data analysis part 1: basic principles and applications, second edition. Umetrics Ltd.: Umeå, Sweden, 2006.
39. Trygg J, Wold S. Orthogonal projections to latent structures (O-PLS). J Chemometr. 2002;16:119–128.
40. Barshan E, Ghodsi A, Azimifar Z, et al. Supervised principal component analysis: visualization, classification and regression on subspaces and submanifolds. Pattern Recognit. 2011;44:1357–1371.
41. Brenneisen R, Kessler T. Psychotrope Drogen. V. Die Variabilitat der Cannabinoidführung von Cannabispflanzen aus Schweizer Kulturen in Abhängigkeit von genetischen und ökologischen Faktoren [Psychotropic drugs. V. Variability of cannabinoid liberation from Cannabis plants grown in Switzerland in relation to genetic and ecological factors]. Pharm Acta Helv. 1987;62:134–139.
42. National Drug Monitor, annual report 2015. Trimbos Institute: Utrecht, The Netherlands, 2015.
43. Mansouri H, Asrar Z, Amarowicz R. The response of terpenoids to exogenous gibberellic acid in Cannabis sativa L. at vegetative stage. Acta Physiologiae Plantarum. 2011;33:1085–1091.
44. Chailakhyan MK, Khryanin VN. The influence of growth regulators absorbed by the root on sex expression in hemp plants. Planta. 1978;138:181–184.
45. Heslop-Harrison J, Heslop-Harrison Y. Studies on flowering-plant growth and organogenesis: IV. Effects of gibberellic acid on flowering and the secondary sexual difference in stature in Cannabis sativa. Proc Roy Irish Acad., Sect B. 1960/1961;61:219–232.
46. Rosenthal E, Chong T. Marijuana grower's handbook: your complete guide for medical and personal marijuana cultivation. Quick American Archives: San Francisco, CA, 2010.
47. Janocha S, Schmitz D, Bernhardt R. Terpene hydroxylation with microbial cytochrome P450 monooxygenases. Adv Biochem Eng Biotechnol. 2015;148:215–250.
48. Takahashi S, Yeo Y, Greenhagen BT, et al. Metabolic engineering of sesquiterpene metabolism in yeast. Biotechnol Bioeng. 2007;97:170–181.
49. Turner CE, Elsohly MA, Boeren EG. Constituents of Cannabis sativa L. XVII. A review of the natural constituents. J Nat Prod. 1980;43:169–234.
50. Aizpurua-Olaizola O, Soydaner U, Öztürk E, et al. Evolution of the cannabinoid and terpene content during the growth of Cannabis sativa plants from different chemotypes. J Nat Prod. 2016;79:324–331.
51. Piomelli D, Russo EB. The Cannabis sativa versus Cannabis indica debate: an interview with Ethan Russo, MD. Cannabis and Cannabinoid Research. 2016;1:44–46.

Abbreviations Used

CBC = cannabichromene
CBD = cannabidiol
CBG = cannabigerol
CBN = cannabinol
CBGM = cannabigerol-monomethylether
FID = flame ionization detection
GA = gibberellic acid
GC = gas chromatography
IS = internal standard
MS = multiple sclerosis
OPLS-DA = orthogonal projection to latent structure-discriminant analysis
PCA = principal component analysis
PCs = principal components
THC = tetrahydrocannabinol
THCV = tetrahydrocannabivarin