The Terroir of Cannabis: Terpene Metabolomics as a Tool to Understand Cannabis sativa Selections

Authors
Elizabeth M. Mudge1,2, Paula N. Brown2,3, Susan J. Murch1

Affiliations
1 Chemistry, University of British Columbia, Kelowna, British Columbia, Canada
2 Natural Health & Food Products Research, British Columbia Institute of Technology, Burnaby, British Columbia, Canada
3 Biology, University of British Columbia, Kelowna, British Columbia, Canada

Key words
Cannabaceae, Cannabis sativa, metabolomics, terpenes, headspace GC-MS, entourage

ABSTRACT
The phytochemical diversity of Cannabis chemovars is not well understood, and many chemovars were created in informal breeding programs without records of parentage or the criteria for selection. Key criteria for selection sometimes included aroma notes and visual cues, which some breeders associated with pharmacological activity. We hypothesized that the process of selection for scents believed to be related to specific tetrahydrocannabinol levels has resulted in modified terpene biosynthesis in these chemovars. Thirty-two cannabinoids, 29 monoterpenes and 38 sesquiterpenes were measured in 33 chemovars from 5 licensed producers. A classification system based on cannabinoid content was used with targeted metabolomic tools to determine relationships in the phytochemistry. Three monoterpenes, limonene, β-myrcene, and α-pinene, and two sesquiterpenes, Caryophyllene and Humulene, were abundant in the majority of chemovars. Nine terpenes were present in tetrahydrocannabinol-dominant chemovars. Three monoterpenes and four sesquiterpenes were predominantly found in cannabinoid-containing chemovars. Low abundance terpenes may have been the aromatic cues identified by breeders. The medicinal activity of some of the terpenes is likely to contribute to the pharmacological effect of specific chemovars. Together, these data demonstrate the synergy of compounds in Cannabis chemovars and point to the need for additional research to understand the phytochemical complexity.

Introduction
Until recently, informal breeding has been the only source of Cannabis varieties (chemovars), where normal crop breeding protocols have not been strictly followed [1–3]. Key selection criteria used to select Cannabis chemovars included aroma, morphology, and other visual cues that some growers associated with tetrahydrocannabinol (THC) potency [1,2]. The aroma of chemovars has been shown to play a significant role in the selection, preference, and quality indication of chemovars [4,5]. The resultant marijuana chemovars may not be genetically distinct from one another but can have different chemistry, and some chemovars with the same name are not genetically similar [6,7]. Genetic variation within chemovars is highlighted by the expression of phytochemicals present, and terpenes are one of the major classes of compounds responsible for aroma, and, therefore, are impacted by these breeding practices.

As with many high value products such as wine and hops, the variation in aroma notes is the result of variation in the volatile constituents, including monoterpenes and sesquiterpenes [5,8,9]. Terpenes are particularly interesting in Cannabis because they are sequestered in glandular trichomes and co-accumulate with the cannabinoids. Both terpenes and cannabinoids are derived from the same precursor molecule, geranyl pyrophosphate, and more than 240 different cannabinoids and terpenes have been described in Cannabis [10–13]. Recent data has signified the pres-
ence of several terpene synthases in Cannabis, mainly producing the major monoterpenes and sesquiterpenes identified in Cannabis [14].

Plant metabolomics provides the ability to study small molecules within samples to understand the underlying impacts of genetics, environment, or stressors [15, 16]. Approaches can be used to combine information from targeted and untargeted metabolites to discover relationships, clusters, families, biochemical pathways, genetic expression, and post-translational modifications that would be missed when performing univariate analysis of single metabolites [16–19]. Several data reduction strategies and unsupervised classification techniques have been developed that reduce complex phytochemical diversity issues like those found in Cannabis chemovars and can be used to identify relationships between metabolites [17, 20, 21]. Multivariate statistics provide avenues to explore these relationships, which have recently been used to describe the impacts of domestication and breeding on cannabinoid biosynthesis but has not been used to evaluate terpene biosynthesis [22].

It has been suggested that Cannabis breeders selected for scent notes that they believe are indicative of high potency chemovars. We hypothesize that this process of selection for scents believed to be related to specific THC levels has resulted in modified terpene biosynthesis in these chemovars. To investigate this hypothesis, we assembled a collection of 33 Cannabis chemovars from 5 different producers and profiled the terpenes. Previous analysis had classified the chemovars as THC-dominant or cannabinoid (CBD)-THC hybrid chemovars [22]. Our data indicate that there are groups of terpenes with characteristic aromas that are associated with major cannabinoid content, which was the major focus of many clandestine breeding programs.

### Results

A total of 67 terpenes were detected and comprised 29 monoterpenes and 38 sesquiterpenes. Monoterpenes accounted for 87.1 to 99.5% of the terpene profiles, while sesquiterpenes accounted for the remaining 0.5 to 12.9%. Four chemovars had less than 1% sesquiterpenes, while the average content was 5.4%.

The classification system based on cannabinoid content, which has been used to highlight breeding based on cannabinoid potency described by Mudge et al. was used to identify relationships of different terpenes across the classes [22]. These classes are summarized in Table 1. To assess the relationships between THC, CBD, and the terpenes, each terpene was graphed according to THC content from lowest to highest and color coded to chemovar class [22]. Five monoterpenes and seven sesquiterpenes were ubiquitous across all chemovars (Fig. 1S and 2). Three monoterpenes, limonene, β-myrcene, and α-pinene, were abundant in the majority of chemovars, while the two most abundant sesquiterpenes, caryophyllene and humulene, ranged from 0.2 to 5.5% and 0.3 to 1.5% respectively.

Seventeen terpenes were found in chemovars from a range of cannabinoid groupings, but not in all chemovars (Fig. 1S, Supporting Information). β-Cubebeene was found in all chemovars except the very low THC, high CBD chemovar (Fig. 1Sd, Supporting Information). There were considerable correlations among the lower abundance sesquiterpenes with correlation coefficients above 0.8, as visually represented in Fig. 3. Correlations were observed between γ-murolone, copaene, β-cubebeene, elemol, germacrene A, guaia-3,9-diene, β-maaliene, γ-maaliene, selina-3,7(11)-diene, α-selinene, and δ-selinene (Fig. 3), for which many of these metabolites were observed in either all or almost all cannabinoid chemovars and clusters (Fig. 1S, Supporting Information).

Eight sesquiterpenes and one monoterpe were present in THC-dominant chemovars (Fig. 4S, Supporting Information). Four were found to be in chemovars identified as mid-range THC (Fig. 3S, Supporting Information). (Z,Z)-α-Farnesene was found only in the chemovar CAN36, and β-sesquiphellandrene was found only in the chemovar CAN27 (Fig. 4SFH). δ-Cadinene and an unidentified sesquiterpene were found only in one chemovar, CAN23 (Fig. 3S, Supporting Information). Santolina triene (tentative identification) was one of two monoterpenes observed to have correlations with sesquiterpenes sesquiterpene (unidentified) and δ-cadinene, all present in this grouping (Fig. 3 and 3S, Supporting Information).

There were 18 terpenes present in high abundance in chemovars identified as high THC and mid-level THC/CBD. Ten monoterpenes were identified as strongly correlated to one another, and are shown in Fig. 5. The remaining eight monoterpenes and sesquiterpene profiles observed in this group are summarized in Fig. 4S, Supporting Information. Terpinolene was the most dominant monoterpe, which was less than 0.3% in 27 of the 33 chemovars, but ranged from 13.4 to 41.2% in the six chemovars that have this distinctive monoterpe profile: CAN16, CAN17, CAN19, CAN21, CAN32, and CAN33. Terpinolene was correlated with other monoterpenes, such as α-tigliene, α-phellandrene, 3-carene, α-terpinene, p-cymene, β-phellandrene, α-terpinene, and terpinen-4-ol, with correlation coefficients ranging from 0.95 to 0.99.

<table>
<thead>
<tr>
<th>Group</th>
<th>Color code</th>
<th>CBD range (% w/w)</th>
<th>THC range (% w/w)</th>
<th># Chemovars</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Blue</td>
<td>&lt; MDL – 0.08</td>
<td>11.3–19.1</td>
<td>20</td>
</tr>
<tr>
<td>B</td>
<td>Purple</td>
<td>&lt; MDL – 0.02</td>
<td>8.0–9.9</td>
<td>3</td>
</tr>
<tr>
<td>C</td>
<td>Orange</td>
<td>7.1–9.7</td>
<td>5.0–6.7</td>
<td>6</td>
</tr>
<tr>
<td>D</td>
<td>Green</td>
<td>5.3–8.8</td>
<td>1.7–3.1</td>
<td>3</td>
</tr>
<tr>
<td>E</td>
<td>Red</td>
<td>16.1</td>
<td>0.7</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 1 Chemovars of Cannabis were clustered into five distinct groups that could be separated by their CBD and THC contents.
0.99 (Fig. 3). Two sesquiterpene alcohols were also classed in this group and were highly correlated to one another (Fig. 4S, Supporting Information).

The final three monoterpenes and four sesquiterpenes were predominantly found in CBD-containing chemovars (Fig. 5S, Supporting Information). Two sesquiterpene alcohols, guaiol and 10-epi-γ-eudesmol, were highly correlated to one another (Fig. 3).

The aromas that describe each of the terpenes detected and identified in this collection were compiled from published sources [23] and are grouped according to their presence within the ter-
Fig. 2 Sesquiterpene profiles identified as those present across the entire dataset. a β-Caryophyllene, b guaia-3,9-diene, c α-guaiene, d humulene, e β-maaliene, f selina-3,7(11)-diene, and g valencene.
pene groupings (▶Table 2). The aromas range from pine and woody to spicy, floral, and citrus. While Group 1 terpenes are present in all chemovars and invoke most of the major aromas, Groups 2 to 5 are considered undertones, contributing to unique aromas within each terpene cluster. Group 2 is a combination of woody, floral, and herbal undertones. Group 3, which is found only in THC-dominant chemovars, contained herbal, floral, woody, sweet, and spicy undertones. Group 4 appears to have a considerable amount of citrus, woody, musty, floral, and sweet undertones and Group 5, which has the CBD-dominant chemovars, has primarily citrus, tropical, and sweet undertones.

A principal component analysis (PCA) was performed in the autoscaled terpene profiles to evaluate the clustering and multivariate correlations between the metabolites. The PCA is shown in ▶Fig. 6A. The first two principal components (PCs) describe 47.53% of the variance within the data. There is no clear clustering of the chemovars according to their THC/CBD classifications as all five cluster groups overlap significantly. Based on the loading plots of the first two PCs (▶Fig. 6B), the majority of the sesquiterpenes cluster together in the top right quadrant of the plot, while the terpinolene-correlated monoterpenes appear to cluster separately from the Cannabis groups in the top left quadrant. PC2 appears to have some influence by different monoterpenes; α-pinene and β-myrcene are negatively correlated from the terpinolene-correlated terpenes on this PC.

It was previously noted than many of the monoterpenes and sesquiterpenes were identified across every cannabinoid class. Therefore, a data reduction strategy was undertaken to remove these metabolites and identify any unique clustering of the chemovars when removing these terpenes. In this case, the number of metabolites was reduced from 67 to 38 and then subjected to PCA (Fig. 6S, Supporting Information). The first two PCs of this reduced dataset describe 40.02% of the data. The loading plots indicate that the first PC is clearly influenced by the terpinolene-correlated monoterpenes, for which the chemovars all cluster together on the right side of the scores plot. PC2 appears to cluster a few chemovars on the top left and bottom left quadrant from the majority of the remaining chemovars. These are influenced by the contents of several sesquiterpenes. The chemovars in the top left quadrant are impacted by δ-selinene, germacrene B, α-cubenene, and y-elemene, all metabolites identified to be present only in THC-dominant chemovars.
Fig. 4  Sesquiterpene profiles present primarily in THC-dominant chemovars.  

a  α-amorphene,  

b  caryophyllene oxide,  
c  α-cubenene,  
d  β-elemene,  
e  γ-elemene,  
f  (Z,Z)-α-farnesene,  
g  germacrene B, and  
h  β-sesquiphellandrene.
Discussion

We hypothesized that the practice of selecting Cannabis chemovars by aromas thought to be indicative of THC content would result in a set of common scent tones characteristic of high-THC chemovars, and that the comprehensive and sensitive analysis of terpene profiles in Cannabis chemovars could then provide new insights into the impact of the domestication on Cannabis. Anecdotal evidence suggests that the informal breeding history of the crop predicted the potency of THC chemovars based on slight

▶ Fig. 5 Monoterpen profiles representing a unique group of terpenes that dominate both THC-dominant and CBD-THC hybrid chemovars found to be strongly correlated with terpinolene. a 3-Carene, b p-cymene, c p-cymenene, d α-phellandrene, e β-phellandrene, f α-terpinene, g γ-terpinene, h terpinen-4-ol, i terpinolene, and j α-thujene. continued next page
aromatic undertones and breeders selected for CBD-containing chemovars by choosing to clone individuals with specific aromas believed to predict these metabolites [1, 3]. Cannabis aromas play many roles in chemovar selection, euphoria, and product quality, and are strongly associated with clandestine breeding [1, 4]. Many of the terpenes have similar characteristic aromas, which can be impacted by concentration, synergy with other aromatic compounds, and subjective interpretation of the aroma [24]. Subjective interpretation of “desirable” Cannabis aromas during breeding could impact terpene profiles, and many chemovar names are indicative of their aroma. Several dominant aromas described in Cannabis names include lemon, sour, skunk, berry/fruit, diesel, or cheese. There has been considerable variation observed between the chemovar name and composition, suggesting that some chemovar names may not accurately describe aroma due to phytochemical variance [25].

Headspace GC-MS analysis was employed for the profiling of monoterpenes and sesquiterpenes in Cannabis because of its sensitivity in comparison to solvent extraction methods and the ability to highlight the aromatic expression (headspace) of the chemovars. This method detected 67 metabolites identified with reference standards and the NIST spectral database with considerable matching capabilities. In many previous characterizations of Cannabis, the number of terpenes ranged from 14 to 37, focusing only on high abundance terpenes and leaving a considerable number not evaluated [25–29]. Over 120 different terpenes have been previously detected in Cannabis, but many of those not detected in this study are typically present in trace levels [29]. The implementation of this more sensitive technique provides a deeper insight into the phytochemical variation within chemovars and the underlying variances that would otherwise be overlooked in traditional solvent extraction-based methods. Evaluating terpene profiles and potential aromatic characteristics provides a deeper insight into aroma selection by breeders and patients [4].

Positive correlations were observed of many low abundance terpenes with the cannabinoid classes. High-THC chemovars had a higher prevalence of herbal and floral undertones and a higher prevalence of several sesquiterpenes. Interestingly, caryophyllene oxide was correlated with high-THC chemovars and is a sesquiterpene identified by canine enforcement officers to detect drugs

Fig. 5 continued
<table>
<thead>
<tr>
<th>Group</th>
<th>Terpene</th>
<th>Scent descriptors</th>
<th>Aroma</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Group 1</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>α-pinene</td>
<td>woody/pine</td>
<td>woody, pine, citrus, spicy, floral</td>
</tr>
<tr>
<td></td>
<td>β-pinene</td>
<td>woody/pine</td>
<td></td>
</tr>
<tr>
<td></td>
<td>trans-2-pinanol</td>
<td>pine</td>
<td></td>
</tr>
<tr>
<td></td>
<td>camphene</td>
<td>woody/camphor</td>
<td></td>
</tr>
<tr>
<td></td>
<td>α-gurjunene derivative</td>
<td>woody/balsamic</td>
<td></td>
</tr>
<tr>
<td></td>
<td>β-maaliene</td>
<td>woody</td>
<td></td>
</tr>
<tr>
<td></td>
<td>selina-3,7(11)-diene</td>
<td>possible woody</td>
<td></td>
</tr>
<tr>
<td></td>
<td>camphene hydrate</td>
<td>woody/camphor</td>
<td></td>
</tr>
<tr>
<td></td>
<td>α-bergamotene</td>
<td>woody</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4,11-selinadiene</td>
<td>woody</td>
<td></td>
</tr>
<tr>
<td></td>
<td>endo-borneol</td>
<td>camphor</td>
<td></td>
</tr>
<tr>
<td></td>
<td>fenchone</td>
<td>camphor</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Z-sabinene hydrate</td>
<td>balsam</td>
<td></td>
</tr>
<tr>
<td></td>
<td>γ-gurjunene</td>
<td>musty</td>
<td></td>
</tr>
<tr>
<td></td>
<td>β-myrcene</td>
<td>spicy/balsamic/peppery</td>
<td></td>
</tr>
<tr>
<td></td>
<td>caryophyllene</td>
<td>spicy/cloves/robes</td>
<td></td>
</tr>
<tr>
<td></td>
<td>copaene</td>
<td>spicy/honey</td>
<td></td>
</tr>
<tr>
<td></td>
<td>γ-muurolene</td>
<td>spicy</td>
<td></td>
</tr>
<tr>
<td></td>
<td>D-limonene</td>
<td>citrus</td>
<td></td>
</tr>
<tr>
<td></td>
<td>α-terpineol</td>
<td>citrus</td>
<td></td>
</tr>
<tr>
<td></td>
<td>β-cubebene</td>
<td>citrus</td>
<td></td>
</tr>
<tr>
<td></td>
<td>valencene</td>
<td>citrus</td>
<td></td>
</tr>
<tr>
<td></td>
<td>guaia-3,9-diene</td>
<td>floral, rose, geranium</td>
<td></td>
</tr>
<tr>
<td></td>
<td>germacrene A</td>
<td>floral</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ylangene</td>
<td>ylang ylang</td>
<td></td>
</tr>
<tr>
<td></td>
<td>humulene</td>
<td>hoppy</td>
<td></td>
</tr>
<tr>
<td></td>
<td>α-selinene</td>
<td>celery</td>
<td></td>
</tr>
<tr>
<td></td>
<td>exo-fenchol</td>
<td>basil</td>
<td></td>
</tr>
<tr>
<td></td>
<td>β-selinene</td>
<td>celery</td>
<td></td>
</tr>
<tr>
<td><strong>Group 2</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>α-gurjunene</td>
<td>woody/balsamic</td>
<td>floral, woody, herbal</td>
</tr>
<tr>
<td></td>
<td>santolina triene</td>
<td>floral</td>
<td></td>
</tr>
<tr>
<td></td>
<td>sesquiterp-1</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td></td>
<td>δ-cadinene</td>
<td>herbal/thyme</td>
<td></td>
</tr>
<tr>
<td><strong>Group 3</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>α-amorphene</td>
<td>woody</td>
<td>herbal, woody, floral, citrus</td>
</tr>
<tr>
<td></td>
<td>caryophyllene oxide</td>
<td>spicy, woody/carrot</td>
<td></td>
</tr>
<tr>
<td></td>
<td>germacrene B</td>
<td>floral/robes</td>
<td></td>
</tr>
<tr>
<td></td>
<td>γ-elemene</td>
<td>floral</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2-carene</td>
<td>sweet</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(Z,Z)-α-farnesene</td>
<td>citrus</td>
<td></td>
</tr>
<tr>
<td></td>
<td>α-cubebene</td>
<td>herbal</td>
<td></td>
</tr>
<tr>
<td></td>
<td>β-elemene</td>
<td>herbal</td>
<td></td>
</tr>
<tr>
<td></td>
<td>β-sesquiphellandrene</td>
<td>herbal/oregano</td>
<td></td>
</tr>
</tbody>
</table>

*continued next page*
The CBD-containing chemovars were higher in citrus and tropical undertones, which were attributed to several monoterpenes and sesquiterpene alcohols. Aromas are determined based on volatility, threshold, concentration, and interactions with other aromatic compounds, therefore, the data described are only a preliminary estimation of aromatic characteristics from each compound [24].

Cannabis chemovars with similar THC/CBD contents exhibit varying pharmacological effects [3, 22, 29, 31], and previous authors have proposed an "entourage effect" theory, suggesting that cannabinoids and terpenes act synergistically to invoke varying pharmacological effects [5, 31, 32]. There are over 30,000 known terpenes in plants [23]. A summary of the pharmacological activities of terpenes identified in this collection that have been described in the literature through in vitro, in vivo, and clinical studies are presented in Table 3. Major monoterpenes such as α-pinene, β-myrcene, and limonene have been shown to have anti-inflammatory, analgesic, and sedative properties evaluated in animal models, respectively [33–36] (Table 3). Terpinolene, present in high abundance in only a select few chemovars, also showed anti-inflammatory and sedative properties in animal models [37, 38] (Table 3). Minor terpenes may also play a significant role. Linalool has been shown to have anti-inflammatory, sedative, anxiolytic, anticonvulsant, and antidepressant activities [31, 39]. Cymene has antinociceptive activity [40]. Terpinen-4-ol has been studied extensively for its anticonvulsant and anticancer activities [41, 42].

Data from other medicinal plants can aid in understanding the pharmacological effects of many of the terpenes in Cannabis. For example, Salvia sp. and Ocimum santum (holy basil) are used for their analgesic, antidepressant, anxiolytic, and anti-inflammatory activities [43, 44]. These plants have many similar terpenes including borneol, β-pinene, α-pinene, camphene α-thujene, β-caryophyllene, sabine, limonene, p-cymene, terpinolene, ocimen, α-cubebene, linalool, β-elemene, β-caryophyllene, α-guaiene, α-amorphene, α-humulene, isoborneol, borneol, α-selinene, β-selinene, and α-muurolene. Myrcia spp. have many similar terpenes and exhibit anti-inflammatory, antiproliferative, and antinociceptive activities [45]. Similarly, Ocium basilicum has reported antidepressant and anticonvulsant activities and similar terpene chemistry [46]. Further research is needed to understand the synergy of these bioactive compounds and the pharmacological significance for humans.

### Table 2 Continued

<table>
<thead>
<tr>
<th>Group</th>
<th>Terpene</th>
<th>Scent descriptors</th>
<th>Aroma</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 4</td>
<td>α-thujene</td>
<td>woody/frankincense</td>
<td>citrus, woody, sweet, spicy</td>
</tr>
<tr>
<td></td>
<td>α-terpinene</td>
<td>woody</td>
<td></td>
</tr>
<tr>
<td></td>
<td>cis-β-terpineol</td>
<td>woody</td>
<td></td>
</tr>
<tr>
<td></td>
<td>α-santalene</td>
<td>woody</td>
<td></td>
</tr>
<tr>
<td></td>
<td>α-bulnesene</td>
<td>Patchouli</td>
<td></td>
</tr>
<tr>
<td></td>
<td>α-fenchene</td>
<td>camphor</td>
<td></td>
</tr>
<tr>
<td></td>
<td>cis-β-farnesene</td>
<td>citrus/sweet</td>
<td></td>
</tr>
<tr>
<td></td>
<td>p-cymene</td>
<td>citrus/sweet</td>
<td></td>
</tr>
<tr>
<td></td>
<td>α-phellandrene</td>
<td>citrus/pepper</td>
<td></td>
</tr>
<tr>
<td></td>
<td>γ-terpinene</td>
<td>citrus</td>
<td></td>
</tr>
<tr>
<td></td>
<td>terpinolene</td>
<td>sweet/pine/citrus</td>
<td></td>
</tr>
<tr>
<td></td>
<td>linalool</td>
<td>citrus/floral/sweet</td>
<td></td>
</tr>
<tr>
<td></td>
<td>δ-selinene</td>
<td>floral</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3-carene</td>
<td>sweet</td>
<td></td>
</tr>
<tr>
<td></td>
<td>α-eudesmol</td>
<td>sweet</td>
<td></td>
</tr>
<tr>
<td></td>
<td>terpinen-4-ol</td>
<td>peppery/musty/sweet</td>
<td></td>
</tr>
<tr>
<td></td>
<td>p-cymenene</td>
<td>spicy/cloves</td>
<td></td>
</tr>
<tr>
<td></td>
<td>β-phellandrene</td>
<td>minty</td>
<td></td>
</tr>
<tr>
<td></td>
<td>bulnesol</td>
<td>spicy</td>
<td></td>
</tr>
<tr>
<td>Group 5</td>
<td>alloaromadendrene</td>
<td>woody</td>
<td>citrus, woody, sweet, tropical</td>
</tr>
<tr>
<td></td>
<td>guaiol</td>
<td>rose wood</td>
<td></td>
</tr>
<tr>
<td></td>
<td>10-epi-γ-eudesmol</td>
<td>sweet</td>
<td></td>
</tr>
<tr>
<td></td>
<td>cis-α-bisabolene</td>
<td>citrus/myrrh/balsamic</td>
<td></td>
</tr>
<tr>
<td></td>
<td>cis-β-ocimene</td>
<td>citrus/tropical</td>
<td></td>
</tr>
<tr>
<td></td>
<td>trans-β-ocimene</td>
<td>citrus/tropical</td>
<td></td>
</tr>
<tr>
<td></td>
<td>sabinene</td>
<td>citrus/pine/spicy</td>
<td></td>
</tr>
</tbody>
</table>
Many of the sesquiterpenes detected with the headspace method are not commonly evaluated in Cannabis. The highly abundant and commonly evaluated sesquiterpenes present in all Cannabis chemovars include β-caryophyllene and humulene. β-Caryophyllene and humulene have anti-inflammatory properties, while β-caryophyllene has been shown to be a CB2 receptor agonist, which contributes to its anxiolytic and antidepressant activities [5, 31,47, 48]. The data collected indicated that β-maaliene, guaia-3,9-diene, and selina-3,7(11)-diene were present in proportions as high as humulene in many of the chemovars and may have therapeutic potential. For example, β-maaliene was isolated from Nardostachys chinensis and provided to mice by inhalation. Results found locomotor activity was reduced due to its sedative effect [49]. Sedation is a common effect noted for many "indica" Cannabis chemovars. Guaia-3,9-diene is prevalent in many different plants including Atractylodes spp., Curcuma wenyujin, Blumea balsamitera, Eucalyptus spp., and Piper longum L., but has not been isolated to determine its pharmacological significance as a single entity. Extracts of these other plants have activities including immune boosting, digestive aid, anti-inflammation, rheumatism healing, and many more [50 –53]. Selina-3,7(11)-diene [also known as eudesma-3,7(11)-diene] was detected in considerable levels in Brazilian green propolis, commonly used in folk medicine to fight infections, but again it has not been evaluated as a single entity for pharmacological effects [54].

Some sesquiterpenes of note include β-elemene, which has strong antitumor activity, eudesmol, which is antiangiogenic, and bisabolene, which has anticonvulsant activity [55 –57]. Bisabolene was more prevalent in CBD-containing chemovars, and there is considerable evidence that CBD has seizure reduction activity [58]. This correlated terpene could signify a synergistic effect, suggesting the benefits of whole plant efficacy versus isolated cannabinoids. This demonstrates the need for in-depth metabolomic evaluations of chemovars for preclinical and clinical studies as well as to determine relationships and generate hypotheses to explain medicinal efficacy.

These data also contribute to our understanding of the domestication of the Cannabis crop. We have previously hypothesized that the variability in the chemotaxonomy arises from domestication syndrome in the Cannabis genome [22,59]. In many of the chemovars, there could have been a loss of phytochemical diversity, as breeders emphasized aromas and pharmacological activity. Breeding selection will impact the fitness of this genus, as these terpenes may be responsible for enhancing pest resistance, improving pollination, or other natural survival mechanisms [60]. Efforts to expand the genetic diversity in the cultivated crop may lead to new medicinal uses and pharmacological activities.

Several limitations of this study have been identified. Headspace analysis of the chemovars evaluated in this dataset are a subset of those available in the marketplace, which may have additional terpene profiles and/or aromas different from those provided herein. Given that aromas are observed from the interaction of many volatile constituents, there may be other metabolites responsible for the aromas of unique chemovars [24]. Aroma is a product of many factors including concentration, volatility, synergy, etc., and therefore aroma information for each terpene identified is provided, while the aroma of each chemovar cannot be deduced. Another consideration is that the samples were evaluated for terpene composition quickly after receipt in order to minimize the potential for losses due to volatility during storage, and the data collected is limited to dried samples. There is limited information available for drying, storage, and handling of these materials prior to receipt that may have impacted the terpene profiles. Additionally, information pertaining to the breeding, selection, and desirable quality attributes that growers were aiming for with each chemovar is unavailable, and therefore assumptions based on pertinent information in the industry around desirable attributes were used to classify strains and identify trends in terpene distributions.
The sensitive analytical method employed in this work allowed a significantly higher number of terpenes to be detected and identified in the chemovars. This expansion of chemical composition allowed for increased chemical characterization and identified several low abundance terpenes associated with cannabinoid potency. The data suggest that domestication syndrome, resulting from informal breeding and selection, has impacted phytochemical diversity, which may be associated with the pharmaco-logical variance observed across chemovars. Future research is needed to understand the activities of low abundance terpenes and synergistic effects in Cannabis chemovars and to determine the importance for medical efficacy and their roles in plant biosynthesis.

### Materials and Methods

#### Reagents

HPLC grade methanol and acetonitrile were purchased from VWR International. Water was deionized and purified to 18.2 MΩ using a Barnstead Smart2Pure nanopure system (Thermo Scientific). Ammonium formate, HPLC grade (> 99.0%), was purchased from Sigma-Aldrich and formic acid (98%) was HPLC grade and purchased from Fisher Scientific. Cannabinoid certified reference material standards purchased as 1 mL solutions in ampules were purchased from Cerilliant Corp. They included tetrahydrocannabinolic acid (THCA, 1.000 mg/mL), Δ9-tetrahydrocannabinol (THC, 1.001 mg/mL), cannabidiolic acid (CBDA, 1.000 mg/mL), cannabidiol (CBD, 1.000 mg/mL), cannabigerol (CBG, 1.000 mg/mL), cannabichromene (CBC, 1.000 mg/mL), tetrahydrocannabivarin (THCV, 1.00 mg/mL), and cannabinol (CBN, 1.000 mg/mL). Additional standards were purchased for peak identification from Cerilliant Corp., which included Δ8-THC (1.000 mg/mL), cannabi-
divaricin acid (CBDVA, 1.000 mg/mL), cannabidivaricin (CBDV, 1.000 mg/mL), cannabigerolic acid (CBGA, 1.000 mg/mL), and cannabicyclol (CB, 1.000 mg/mL). All cannabinoid standards were provided in either methanol or acetonitrile. Cannabis Terpene Mix A and Mix B containing 20 and 15 terpenes, respectively, at 20,000 mg/L in methanol were purchased from Sigma-Aldrich. Cannabis Terpene Mix A contained α-pinene, camphene, β-pinene, 3-carene, α-terpinene, limonene, γ-terpinene, fenchone, fenchol, camphor, isoborneol, menthol, citronellol, pulegone, geranyl acetate, α-cedrene, α-humulene,nerolidol, cedrol, and α-bisabolol. Cannabis Terpene Mix B contained β-pinene, 3-carene, p-cymene, limonene, terpinolene, linalool, camphor, bornene, α-terpineol, geraniol, β-caryophyllene, cis-nerolidol, β-eudesmol, and phytol.

**Plant materials**

Thirty-three chemovars of Cannabis sativa L. were purchased from five licensed producers in Canada under the Access to Cannabis for Medical Purposes Regulation (ACMPR), and laboratory analysis was performed under a Health Canada Research License. The test samples were provided as whole or milled flowers in 5-, 10-, and 15-g packages and stored at room temperature until use. Due to the legal restrictions pertaining to the storage of Cannabis chemovars, submission of voucher specimens to an herbarium was not possible, but given the regulatory framework associated with these plants, their identity has been confirmed as C. sativa L.

**Cannabinoid analysis**

The content of 32 cannabinoids was determined previously [22, 61]. In brief, ground Cannabis flowers (0.200 g) were extracted with 25 mL of 80% methanol for 15 min, followed by centrifugation at 4500 g for 5 min and filtration with a 0.22-µm PTFE filter. Extracts were diluted to within the calibration range using the extraction solvent and placed in the 4 °C sample holder for same-day analysis. Chromatographic separation was performed on an Agilent 1200 UHPLC with a Kinex C18 100 mm × 3.0 mm, 1.8 µm column (Phenomenex) using a gradient elution with 10 mM ammonium formate (pH 3.6) and acetonitrile with detection at 200 nm. Chemovars were classified into five clusters based on the range of CBD/THC values determined [22].

**Evaluation of volatile constituents**

Terpene profiles were determined using an in-house developed method, adapted from a previous terpene method [13]. Immediately after grinding, Cannabis flowers (0.100 g) were added to a 20-mL gas tight headspace vial. Samples were prepared in triplicate. Using a CTC Analytics Combi-PAL headspace autosampler, the vials were transferred to a heated incubator at 80°C for 15 min and agitated at 500 rpm. Next, 1000 µL of the vial headspace was injected using a syringe at 120°C. The injector temperature was 230°C with a split ratio of 10:1. GC analysis was undertaken on an Agilent 7890A GC coupled to a 5975B mass spectrometer (MS). Separation was achieved on a 20 m × 180 µm ID, 0.18 µm film thickness J&W DB-5MS column. Helium was used as the carrier gas at a flow rate of 1.3 mL/min. The column was held at 50°C for 3 min followed by a ramp to 170°C at 5°C/min for a total run time of 27 min. MS detection with electron impact ionization at 70 eV was used to collect mass spectra from m/z 50 to 500. The MS quad and source temperatures were 230 and 150°C, respectively.

**Chemometrics**

**Metabolite profiling**

Terpenes were identified based on comparison of mass spectra with the NIST spectral database (NIST 11). Additionally, retention indices were compared to published literature to confirm elution order and identity [62]. Several monoterpenes standards were also analyzed to confirm identity. Multivariate curve resolution using SOLO+MIA software (version 8.5; Eigenvector Research) was employed to separate coeluting monoterpenes, and peak areas were determined using the software program R, version 3.5.2 [63]. Peaks were manually aligned based on compound identity and retention time using Excel. Missing values (zeros) were replaced with half of the lowest value in the dataset.

**Identification of metabolite relationships**

Individual terpenes were plotted according to their cannabinoid profiles, previously described by Mudge et al. to identify trends within the datasets and classify them into unique groups [22]. Trends evaluated included those present across all chemovars, those found primarily in THC-dominant chemovars, those present primarily in CBD-THC hybrid chemovars, and other terpene correlations independent of cannabinoid content. Correlations between terpenes and cannabinoids were confirmed by evaluating Pearson correlation coefficients using the R program cor.

**Multivariate classification**

The data were autoscaled by mean centering and scaling to unit variance in order to give each metabolite equal weight prior to multivariate analyses. PCA was subsequently performed using Solo+MIA.

**Supporting information**

Terpene profiles identified across different cannabinoid classes, but not present in all chemovars (endo-borneol, camphene hydrate, copeane, β-cubebene, exo-fenchol, fenchone, germacrene A, α-gurjunene derivative, γ-gurjunene, γ-muurolene, trans-2-pinanol, z-sabinine hydrate, 4,11-selinadiene, α-selinene, β-selinene, α-terpineol, ylangene) are described in Fig. 15. The profile of the monoterpine 2-carene, detected primarily in THC-dominant chemovars is described in Fig. 25. The terpene profiles for those found primarily in mid-level THC-dominant chemovars (β-cadinene, α-gurjunene, santolina triene, sesquiterp-1) are summarized in Fig. 35. Several additional terpenes representing a unique group of terpenes that dominate both THC-dominant and THC-CBD hybrid chemovars (α-bulnesene, bulnesol, eudesmol, cis-β-farnesene, α-fenchene, linalool, α-santalene, δ-selinene) are summarized in Fig. 45. The profiles of terpenes found predominately in higher CBD chemovars (alloaromadendrene, cis-α-bisabolene, 10-epi-γ-eudesmol, guaol, cis-β-ocimene, trans-β-ocimene, sabine) are summarized in Fig. 55. A PCA of the monoterpene and sesquiterpene profiles after undertaking a data reduction strategy, and the associated loading plots are summarized in Fig. 65.
Acknowledgements

This research was undertaken, in part, thanks to funding from the Canada Research Chairs program.

Conflict of Interest

The authors declare no conflicts of interest.

References

[34] Do Vale TG, Furtado EC, Santos JS, Viana GSB. Central effects of citral, myrcene and limonene, constituents of essential oil chemotypes from Lippia alba ( mill.) N. E. Brown. Phyto medicine 2002; 9: 709–714


Bahi A, Al Mansouri E, Al Ameiri A, Nurulain SM, Ojha S. β-Caryophyllene, a CB2 receptor agonist produces multiple behavioral changes relevant to anxiety and depression in mice. Physiol Behav 2014; 135: 119–124


Peng W, Han T, Xin WB, Zhang XC, Zhang QY, Jia M, Qin LP. Comparative research of chemical constituents and bioactivities between petroleum ether extracts of the aerial part and the rhizome of Arctomyodes macrocephala. Med Chem Res 2011; 20: 146–151


Ali-Snafi PDAE. The pharmacological and therapeutic importance of Eucalyptus species grown in Iraq. IOSR J Pharm 2017; 7: 72–91

Hieu LD, Hoi TM, Thang TD, Erenasya OA. Fonenol, the main constituent of the essential oil of Piper longum L. Am J Essent Oil Nat Prod 2018; 6: 16–19


Ciftci O, Ozdemir I, Tanyildzi S, Yildiz S, Ogutkur H. Antioxidative effects of curcumin, β-mycene and 1,8-cineole against 2,3,7,8-tetra-
chlorodibenzo-p-dioxin-induced oxidative stress in rats liver. Toxicol Ind Health 2011; 27: 447–453


[85] Hsu HC, Yang WC, Tsai WJ, Chen CC, Huang HY, Tsai YC. α-Bulnesene, a novel PAF receptor antagonist isolated from Pogostemon cablin. Biochem Biophys Res Commun 2006; 345: 1033–1038


